Shrinking the Cross-Section of Index Option Returns

Christian Fieberg * Gerrit Liedtke † Paige Michael-Shetley ‡ Thorsten Poddig §
Thomas Walker ¶

Abstract

This paper shows that sparse factor models fail to capture the predictive patterns in S&P500 index options; rather a high-dimensional stochastic discount factor (SDF) is necessary. Studying a set of 54 option-based factors, our empirical findings suggest that all of them contribute to the SDF. A non-sparse SDF-implied mean-variance efficient portfolio spans the highest mean-variance efficient frontier among all benchmark models, yielding fewer pricing errors for option return anomalies. Factors such as the maturity slope and factor momentum, along with characteristic-based factors exploiting the spreads in embedded leverage, vega, theta, time-to-maturity, and options prices, contribute the most to the SDF.

Keywords: option return predictability, stochastic discount factor, cross-section, asset pricing, machine learning, big data

JEL classification: G10, G11, G12, G13

*HSB Hochschule Bremen - City University of Applied Sciences, 28199 Bremen, Werderstr. 73, Germany, christian.fieberg@hs-bremen.de.

†Corresponding author. University of Bremen, 28359 Bremen, Max-von-Laue-Str. 1, Germany, gliedtke@uni-bremen.de

‡University of Bremen, 28359 Bremen, Max-von-Laue-Str. 1, Germany, paige@uni-bremen.de

§University of Bremen, 28359 Bremen, Max-von-Laue-Str. 1, Germany, poddig@uni-bremen.de

¶Concordia University, John Molson Building, 1450 Guy, Montreal, Canada, thomas.walker@concordia.ca

1 Introduction

Research on option pricing frequently centers around the question of what drives the variation in option returns. Contradicting the view that options are redundant assets (Black and Scholes, 1973), the literature finds that option returns are driven by factors other than simply their exposure to variations in the underlying. Specifically, a growing strand of literature documents patterns in option returns that can be attributed to *characteristics* of the underlying. For example, Goyal and Saretto (2009) find that the difference between the historical volatility of the underlying and its implied volatility predicts delta-hedged stock option returns. Hu and Jacobs (2020) find a significant relationship between the return of a stock option and its volatility, while Cao and Han (2013) document that the idiosyncratic volatility of the underlying is a cross-sectional predictor for option returns. Zhan et al. (2022) test the predictability of stock option returns using ten well-known stock return anomalies and find that these have significant predictive power. Bali et al. (2023) test the 94 stock characteristics used in Green et al. (2017) and Gu et al. (2020) and find similar results. Beyond the underlying's characteristics, the literature finds that option characteristics themselves have predictive power. Boyer and Vorkink (2014) show a negative relationship between the option implied skewness and option returns. Relatedly, Christoffersen et al. (2018) and Kanne et al. (2023) document a significant relation between liquidity and option returns. Vasquez (2017) finds that the difference in long-term and short-term implied volatility predicts delta-hedged option returns and Ruan (2020) documents that the volatility of the implied volatility is a significant predictor. Recently, Heston et al. (2022) find a momentum effect in option returns and Käfer et al. (2023) show that option momentum originates from factor momentum in option-based factors.

The expanding body of literature on option return predictability suggests that, like in the equity or bond markets, there are numerous signals that are potentially informative about future returns. As it is common in other asset classes (Fama and French, 1993, 2015; Hou et al., 2020), researchers aim to capture the cross-sectional patterns in option returns by assuming low-dimensional factor models: Horenstein et al. (2022) use the risk-premium principal component analysis (RP-PCA) proposed by Lettau and Pelger (2020) to uncover the factor structure in stock option returns and find that a three-factor model including the equally weighted market portfolio, a historical-minus-implied volatility, and a volatility-of-volatility-factor best approximates the principal components. Bali et al. (2022) find that a six-factor model including the market factor, a stock price, option price, implied-minus-realized volatility, implied-minus-realized skewness, and implied-minus-realized kurtosis factor best captures the variation in option returns. Karakaya (2014) proposes a level, maturity slope, and moneyness skewness factor and Frazzini and Pedersen (2022) show that a betting-against-beta factor is priced in the cross-section. Büchner and Kelly (2022) estimate a conditional latent factor model using instrumented principal component analysis (IPCA) for S&P500 index option returns and find that a three-factor model best captures the return variation in delta-hedged options. They show that the identified latent factors are highly correlated with the level, moneyness skewness, and maturity slope factors proposed in Karakaya (2014); however, they capture information beyond the three observable factors.

In light of the large number of predictable patterns in the cross-section of option returns, it is essential to identify the signals that truly provide incremental information for future option returns. Independently of the respective asset class studied, the no-arbitrage condition that all alphas equal zero implies the existence of a stochastic discount factor (SDF) M_t that satisfies the condition $E_{t-1}\left[M_t \boldsymbol{f}_t\right] = 0$, with $\boldsymbol{f}_t$ denoting an $L \times 1$ vector of factor returns at time t. The Euler equation and the absence of arbitrage are the motivation for researchers to propose low-dimensional factor models, assuming that option returns are sparse in only a few factors (e.g., Karakaya, 2014; Büchner and Kelly, 2022; Horenstein et al., 2022; Bali et al., 2022); however, given the many option sensitivities (i.e., “Greeks”) and other cross-sectional effects documented in the literature, does a sparse representation of the SDF actually suffice to capture the entire variation in option returns?

Sparsity in the SDF implies high redundancy among the candidate factors; however, our empirical findings do not support the sparsity argument. Specifically, we study delta-hedged returns of S&P500 index options over a period of almost three decades. We replicate 54 factors studied in the previous literature and, using the agnostic approach of Kozak et al. (2020), we examine—in a purely data-driven fashion—whether factors earn a premium because they proxy for a source of systematic risk or because they are correlated with the true SDF but do not provide incremental information. The approach proposed by Kozak et al. (2020) allows the SDF to be estimated while regularizing the factor weights in terms of shrinkage (L^2) to reduce overfitting and sparsity (L^1) to zero out factors in the SDF. Our empirical findings suggest that that many factors earn a significant premium that is not captured by either the capital asset pricing model (CAPM) or other models previously proposed in the literature. Figure 1 shows the alphas obtained from a time-series regression of the 54 option factors on the ex ante mean-variance efficient portfolios (MVPs) implied by the CAPM, Karakaya (2014) three-factor (K3), Horenstein et al. (2022) three-factor (HVX3), Bali et al. (2022) five-factor (BCCSZ5),[1] RP-PCA three-factor (RPPCA3), and IPCA three-factor (IPCA3) models.[2] The filled dots represent the alphas against the MVP implied by a non-sparse SDF, estimated using L^2 regularization. The adjacent model allows for sparsity in the SDF by regularizing the SDF coefficients with both L^1 and L^2 regularization. The dashed red lines mark the t-statistic thresholds at which the alpha is statistically significant at the 5% level, after adjusting the p-values to control for the multiple hypothesis testing problem using the Benjamini and Hochberg (1995) method. The numbers above the circles refer to the number of factors that earn a significant alpha compared to the respective model. Considering the CAPM, sixteen option-based factors earn a significant abnormal premium, supporting the finding of cross-sectional predictability of index option returns documented in the literature. The results for the non-sparse SDF, which employs L^2 regularization, suggest that the index option return puzzle is highly multidimensional so that a few characteristic-based factors cannot capture the entire return variation. Specifically, the non-sparse SDF fails to explain the returns of two factors, while the benchmark models leave between nine and twenty factors unexplained. This result emphasizes the need for a non-sparse

[1] Bali et al. (2022) propose a six-factor model for stock options; however, one factor is based on stock prices. Because we analyze index options for which all options have the same underlying, we cannot create a stock price factor. Thus, we analyze the cross-sectional variation explained by the five option-based factors.

[2] Note that we use the terms MVP and tangency portfolio interchangeably.

model for option returns.

[Insert Figure 1 about here]

Our extensive asset pricing tests highlight the superiority of a non-sparse SDF. The MVP implied by the non-sparse SDF achieves an annualized CAPM alpha of 1.40%. The alphas against other empirical benchmark models range from 0.66% to 1.53%. Not even the statistical benchmark models, namely, RP-PCA and IPCA, can explain the MVP returns of the non-sparse SDF, suggesting that it exhibits mean-variance efficiency. In contrast, we find that the other MVPs are not mean-variance efficient. Throughout our empirical study, we further show that the extension from a set of 54 factors that are linear in option characteristics to an extensive set of 702 factors, including nonlinear transformations and interactions, does not further improve the span of the efficient frontier. Our findings are robust across several market states, against alternative estimation windows, and against the exclusion of illiquid options. Finally, MVP returns cannot be explained by asset pricing factors from the stock market, supporting the view that options are not redundant assets but have their own dynamics. Our empirical findings that the SDF in index option returns is non-sparse is in line with the findings documented in Kozak et al. (2020) for equities. Furthermore, the results presented in Büchner and Kelly (2022) likewise suggest that many option characteristic-based contain incremental information about future option returns; however, our findings reveal that IPCA is not able to capture all information with a handful of factors.

Although our empirical findings indicate that many factors contribute to the SDF, some factors dominate. The maturity slope factor of Karakaya (2014) that captures variations in option returns associated with shifts in the implied volatility term structure is by far the most important factor (after controlling for the market portfolio). Next, factor momentum and a theta-based factor along with characteristic-based factors constructed from put options, namely embedded leverage, option vega, change in implied volatility, time-to-maturity, and option price, are of great importance. While other factors weigh less in the SDF, they are not of negligible relevance; rather, the linear combination of many factors contributes to the pricing performance of the SDF. In fact, sparse SDFs have higher pricing errors, suggesting that the option pricing puzzle is highly multidimensional.

Our findings align with the recent evidence in the literature for other asset classes that the SDF is dense in many characteristics, hence, it cannot be spanned by simple linear models covering only a few factors. Examples of such studies include Bryzgalova et al. (2023) for equities, Dickerson et al. (2023) for corporate bonds, and Käfer et al. (2024) for single-name equity options. The recent evidence that a "complex" SDF best prices assets is in line with the theoretical evidence provided by Didisheim et al. (2024). They refer to this phenomenon as "virtue of complexity", which is also shown to exist in time-series analyses (Kelly et al., 2022, 2024).

This paper contributes to the existing literature in several ways. First, we extend the literature on the cross-sectional predictability of option characteristics to index option returns. While most cross-sectional studies relate to stock options with different underlyings (Goyal and Saretto, 2009; Cao and Han, 2013; Ruan, 2020; Hu and Jacobs, 2020; Zhan et al., 2022; Bali et al., 2022, 2023), studies on the cross-section of index options are limited (Cao and Huang, 2007; Hu and Liu, 2022; Büchner and Kelly, 2022), despite their high relevance for risk management. While Zhan et al. (2022) show that the underlying's characteristics predict future option returns, Bali et al. (2023) find that contract-level characteristics are the most important signals for future returns, although they show that the underlying's characteristics play a minor yet not unimportant role. By studying index option returns, the underlying is the same for all contracts, thus allowing to study the risk factors that drive option returns in isolation, i.e., without having to take differences in the respective underlying into account. This also eliminates "idiosyncratic noise" due to stock-specific events or incomplete information in the underlying's characteristics.

Second, our paper is related to the strand of literature that applies machine learning techniques to the cross-section of option returns. Cao and Huang (2007) use principal component analysis to uncover latent risk factors in S&P500 index option returns. Horenstein et al. (2022) use RP-PCA to identify the empirical factors that best approximate the most informative principal components in stock option returns. Büchner and Kelly (2022) and Goyal and Saretto (2022) use IPCA to uncover a conditional linear latent factor structure in index and stock option returns, respectively. Closely related to the approach used in this paper is the work of Shafaati et al. (2022) who use a penalized linear regression to uncover which characteristics provide incremental information in stock option returns. Within this high-dimensional setting, they show that a penalized estimator yields more accurate predictions. Bali et al. (2023) additionally allow for nonlinearities and demonstrate that predictability is further enhanced. Similarly, Goyenko and Zhang (2022) train machine learning techniques using option and stock characteristics as inputs to predict option and stock returns.

The remainder of this paper is organized as follows: Section 2 briefly reviews the estimation of the robust SDF as proposed in Kozak et al. (2020) and Section 3 describes the data used in our empirical study. Section 4 presents the main results while Section 5 reviews the findings for a series of robustness checks. Section 6 concludes the article.

2 Method

This section briefly reviews the estimation of the robust SDF estimation approach proposed in Kozak et al. (2020). To introduce notation, let $\boldsymbol{r}_t$ be an $N \times 1$ vector of option excess returns, with $i = 1, \ldots, N$ denoting the number of option contracts. $\boldsymbol{Z}_{t-1}$ is an $N \times L$ matrix of option characteristics, with $l = 1, \ldots, L$ being the number of characteristics. Note that $\boldsymbol{Z}_{t-1}$ may include any linear or nonlinear transformation of the characteristics (Kozak et al., 2020).

Assuming that the law of one price holds, one can find an SDF linear in excess returns

$$M_t = 1 - \boldsymbol{b}_{t-1}^{'} \left(\boldsymbol{r}_t - E_{t-1} \left[\boldsymbol{r}_t \right] \right) \tag{1}$$

such that all individual options are priced according to:

$$E_{t-1} \left[M_t \boldsymbol{r}_t \right] = 0 \tag{2}$$

with $\boldsymbol{b}_{t-1}$ denoting an $N \times 1$ vector of SDF loadings. The SDF loadings can be obtained by

$$\boldsymbol{b}_{t-1} = \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \tag{3}$$

with $\boldsymbol{\Sigma}^{-1}$ denoting the true $N \times N$ variance-covariance matrix of asset returns and $\boldsymbol{\mu}$ denoting an $N \times 1$ vector of true asset risk premia $E_{t-1}[\boldsymbol{r}_t]$. However, the true population moments $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are not known and must be estimated. Estimating $\boldsymbol{\Sigma}$ from empirical data poses a significant challenge for two main reasons: First, individual return data is unbalanced, introducing complexities in the estimation process. Second, as the sample size N increases, the estimation of equation (3) may become infeasible.

To address these issues, researchers frequently try to summarize the variation in asset returns with few characteristic-based factors (Fama and French, 1993, 2015; Horenstein et al., 2022; Bali et al., 2022; Büchner and Kelly, 2022). That is, the characteristic-based SDF can be expressed as:

$$M_t = 1 - \left(\boldsymbol{Z}_{t-1} \boldsymbol{b}\right)' \left(\boldsymbol{r}_t - E\left[\boldsymbol{r}_t\right]\right) \tag{4}$$

with $\boldsymbol{b}$ denoting an $L \times 1$ vector of time-invariant coefficients (Kozak et al., 2020). Mapping the asset excess returns into the L-dimensional factor space, i.e.,

$$\boldsymbol{f}_t = \boldsymbol{Z}_{t-1}' \boldsymbol{r}_t \tag{5}$$

with $\boldsymbol{f}_t$ denoting an $L \times 1$ vector of factor returns, the SDF in equation (4) can be rewritten as

$$M_t = 1 - \boldsymbol{b}' \left(\boldsymbol{f}_t - E\left[\boldsymbol{f}_t\right]\right). \tag{6}$$

Likewise, equation (2) can be reformulated to

$$E_{t-1}\left[M_t \boldsymbol{f}_t\right] = 0. \tag{7}$$

Analogous to equation (3), one can solve equation (6) and (7) by estimating

$$\boldsymbol{b} = \boldsymbol{\Sigma}_f^{-1}\boldsymbol{\mu}_f \tag{8}$$

with $\boldsymbol{\Sigma}_f$ denoting the true $L \times L$ variance-covariance matrix of factor returns and $\boldsymbol{\mu}_f$ denoting an $L \times 1$ vector of true factor premia. By first mapping individual returns into the factor space, one can always create balanced factor return data, which makes the estimation of $\boldsymbol{\Sigma}_f$ possible. Likewise, by reducing the dimensionality from N to L assets, the estimation of $\boldsymbol{\Sigma}_f$ becomes feasible. However, as outlined in Kozak et al. (2020), the estimation of the population moments may become infeasible if L is large. Similarly, estimation may become inefficient in the presence of highly correlated factors.

Instead of estimating $\boldsymbol{b}$ as in equation (8), Kozak et al. (2020) propose a penalized estimator for $\boldsymbol{b}$ to efficiently operate in high-dimensional datasets. Specifically, their dual-penalty estimator allows for L^1 and L^2 regularization. L^2 regularization adds the squared magnitude $\boldsymbol{b}'\boldsymbol{b}$ to the loss function in order to reduce the magnitude of less important factors. While L^2 regularization shrinks the coefficients towards but not exactly to zero, the L^1 regularization adds a penalty for the absolute value of magnitude $\sum_{l=1}^{L} |b_l|$ to the loss function, and thus allows shrinking the coefficients to exactly zero. Accordingly, L^2 and L^1 regularization are also referred to as shrinkage and sparsity, respectively. While L^1 regularization allows the coefficients to be shrunk to exactly zero, it performs poorly if the factors are highly correlated (Tibshirani, 1996; Zho and Hastie, 2005); therefore, the L^1 penalty is often blended with the L^2 penalty, which shrinks the coefficients of correlated factors toward each other, thereby allowing them to borrow strength from each other (Hastie et al., 2011). The combined L^1 and L^2 is often referred to as elastic net, and is frequently applied in empirical finance (Gu et al., 2020; Dong et al., 2022).

Thus, the dual-penalty estimator in Kozak et al. (2020) minimizes the Hansen and Jagannathan

(1991) distance subject to an L^2 and L^1 penalty:

$$\hat{\boldsymbol{b}} = \arg\min_{\hat{\boldsymbol{b}}} \left(\bar{\boldsymbol{\mu}} - \bar{\boldsymbol{\Sigma}}\hat{\boldsymbol{b}}\right)' \bar{\boldsymbol{\Sigma}}^{-1} \left(\bar{\boldsymbol{\mu}} - \bar{\boldsymbol{\Sigma}}\hat{\boldsymbol{b}}\right) + \gamma_2 \hat{\boldsymbol{b}}'\hat{\boldsymbol{b}} + \gamma_1 \sum_{l=1}^{L} |\hat{b}_l| \tag{9}$$

with γ_2 and γ_1 controlling the degree of L^2 and L^1 regularization, respectively, and $\bar{\boldsymbol{\mu}}_f$ and $\bar{\boldsymbol{\Sigma}}$ denoting the sample estimates of the true vector of factor premia and the variance-covariance matrix of factor returns.

3 Data

This section discusses the sample data and our preprocessing approach. Specifically, Section 3.1 introduces our sample and discusses the filters used to prepare the data, Section 3.2 illustrates how we calculate delta-hedged returns, and Section 3.3 describes the factor construction.

3.1 Data Source and Preprocessing

We obtain daily option data from OptionMetrics during the period from January 1996 to December 2022, including option-specific characteristics, underlying index values, and option sensitivity measures such as the Black-Merton-Scholes (BMS) delta, gamma, vega, and theta. Data for the VIX index are obtained through the Chicago Board Options Exchange (CBOE).

To ensure the cleanliness of our data, we apply a number of filters to the OptionMetrics data that were previously proposed in the literature. Specifically, we exclude observations in which the bid price is negative, the bid price is greater than the ask price, or no-arbitrage conditions are violated (Karakaya, 2014; Lemmon and Ni, 2014; Büchner and Kelly, 2022; Frazzini and Pedersen, 2022).

In our study, we rely on daily data to increase the estimation efficiency of the SDF (see Kozak et al., 2020, for a discussion on the relevance of a large number of time-series observations); however, we do not rebalance portfolios daily to reduce turnover (Kozak et al., 2020).[3] Instead, on each third Friday of a month, which is the standard expiration day for the options studied here, we assign options to factor portfolios (see Section 3.3) and hold these factor portfolios until the third Friday of the next month (Büchner and Kelly, 2022). To further enhance the quality of our data, we then apply the following filters at the beginning of each holding period: First, we exclude observations for which the implied volatility measure is not available. Second, we control for outliers by excluding observations with embedded leverage below (above) the 1st (99th) percentile of the cross-sectional embedded leverage distribution, with embedded leverage $\Omega_{i,t}$ of option i at time t defined as:

$$\Omega_{i,t} = |\Delta_{i,t} \frac{S_t}{F_{i,t}}| \tag{10}$$

with $\Delta_{i,t}$ denoting the BMS delta of option i at time t, S_t being the spot price at time t (the same for all options), and $F_{i,t}$ being the mid price of the option i at time t. We apply this filter to put and call options separately (Karakaya, 2014; Büchner and Kelly, 2022). Third, we increase the liquidity of our sample by limiting it to option contracts with an absolute forward delta between 0.01 and 0.5 and, fourth, we require a time-to-maturity of one to twelve months (Israelov and Kelly, 2017; Büchner and Kelly, 2022). As shown in Büchner and Kelly (2022), options outside of this range account for only a marginal fraction of the option market and are of minor economic importance; however, they introduce noise into the return data. Finally, because we use daily data, we require for an option to be included in the portfolio that its characteristic is available at the beginning of the holding period and that all return observations within that holding period are non-missing.

Figure 2 shows the number of options and the total open interest over time. At the beginning of our sample, there are only a few hundred options, but from 2008 onwards, the number of contracts and the trading volume increases rapidly, so that the cross-section at the end of the sample consists of approximately 4,000 unique options per holding period.

[3]In unreported results, we also analyze monthly data but find that the results are qualitatively unchanged. To be consistent with the empirical analysis in Kozak et al. (2020), we therefore stick to daily data.

[Insert Figure 2 about here]

3.2 Delta-Hedged Option Returns

We follow the common literature (e.g., Bakshi and Kapadia, 2003; Cao and Han, 2013; Büchner and Kelly, 2022) and focus on delta-hedged option returns, so that the options are immune to linear price changes of the underlying. That is, we delta-hedge the returns on a daily basis to isolate the fraction of option returns that is not due to price changes in the underlying (Büchner and Kelly, 2022; Bali et al., 2023). Specifically, the discrete delta-hedged profit-and-loss (P&L) from day t to $t+1$ with delta-hedging at the end of each trading day is given by:

$$\Pi_{[t,t+1]} = (F_{t+1} - F_t) - \Delta_t (S_{t+1} - S_t) - \frac{a_{t,t+1} r_t^f}{365} (F_t - \Delta_t S_t) \tag{11}$$

with F_t denoting the option's mid-price at time t, Δ_t the option's delta at time t, S_t the closing price of the underlying at time t, r_t^f the risk-free rate at time t, and $a_{t,t+1}$ the number of days between trading days t and $t+1$. The first term is thus the raw P&L of the option from day t to $t+1$, the second term adjusts the raw P&L by delta-hedging the position, and the third term adjusts for the cost of funding the delta-hedged portfolio at the risk-free rate. Finally, as in Büchner and Kelly (2022), the delta-hedged option return is obtained by dividing the delta-hedged option P&L $\Pi_{[t,t+1]}$ by the closing price of the underlying:[4]

$$r^{\Delta}_{i,t+1} = \frac{\Pi_{[t,t+1]}}{S_t} \tag{12}$$

Table 1 provides summary statistics for our sample, including the mean, median, and standard deviation of option characteristics and returns for call (Panel A) and put (Panel B) options. The last row in each panel reports the number of observations for the call and put options. In total, our sample covers 106,990 unique option contracts (36,527 call and 70,463 put options) over a

[4] In Internet Appendix A, we show that the Euler equation shown in equation (2) also holds for funded delta-hedged option returns as defined in equations (11) and (12), justifying the use of the method described above.

sample period of almost three decades, resulting in 6,722 daily observations. The mean and median delta-hedged return is negative, which is in line with the previous literature (Cao and Han, 2013; Büchner and Kelly, 2022). For call options, the annualized average delta-hedged return is -4.67% while put options earn an average return of -7.05%.

[Insert Table 1 about here]

3.3 Candidate Option-Based Factors

Our set of candidate factors includes 48 characteristic-based factors and six non-characteristic-based factors that were previously proposed in the literature. The characteristic-based factors are zero-investment long-short portfolios with individual option weights depending on the cross-sectional rank of the options. Specifically, we follow Kozak et al. (2020) and first transform all characteristics into normalized cross-sectional ranks as described in equation (13):

$$\tilde{z}^l_{i,t-1} = \frac{\text{rank}\left(z^l_{i,t-1}\right)}{n_{t-1}+1} \tag{13}$$

with $z^l_{i,t-1}$ denoting the l-th characteristic of the i-th option, observed at time $t-1$, n_{t-1} denotes the number of available options at time $t-1$, and $\tilde{z}^l_{i,t-1}$ denotes the i-th option's normalized rank at time $t-1$. Then, we center the characteristic ranks cross-sectionally and divide them by the sum of absolute deviations from the mean of all options to obtain the portfolio weights:

$$w^l_{i,t-1} = \frac{\tilde{z}^l_{i,t-1} - \bar{\tilde{z}}^l_{t-1}}{\sum_{i=1}^{n_{t-1}} |\tilde{z}^l_{i,t-1} - \bar{\tilde{z}}^l_{t-1}|} \tag{14}$$

where $w^l_{i,t-1}$ represents the weight of option i at time $t-1$ in the factor portfolio l. The term $\bar{\tilde{z}}^l_{t-1}$ denotes the cross-sectional mean of characteristic ranks, i.e., $\bar{\tilde{z}}^l_{t-1} = \frac{1}{n_{t-1}} \sum_{i=1}^{n_{t-1}} \tilde{z}^l_{i,t-1}$. Finally, we stack the option weights into a matrix $\boldsymbol{W}_{t-1}$ of dimension $N \times L$ and obtain the factor portfolio returns $\boldsymbol{f}_t$ at time t as:

$$\boldsymbol{f}_t = \boldsymbol{W}_{t-1}^{'} \boldsymbol{r}_t^{\Delta} \tag{15}$$

with $\boldsymbol{r}_t^{\Delta}$ denoting the $N \times 1$ vector of delta-hedged option returns at time t, as defined in equation (12). As emphasized in Kozak et al. (2020), these factor portfolios are insensitive to outliers and keep the amount of long and short positions invested in the characteristic-based factor fixed. Thus, a change in the number of options at any time t has no effect on the strategy's gross exposure.

The list of characteristic-based factors includes the option sensitivities *delta*, *gamma*, *vega*, and *theta* as well as the higher-order sensitivities *speed*, *vanna*, and *volga*. The set is augmented by embedded leverage (*emb_lev*), moneyness (*mness*), time-to-maturity (*ttm*), implied volatility (*implvol*), change in implied volatility (*implvol_ch*), volatility of implied volatility (*volvol*), maximum implied volatility over the previous holding period (*maxivol*), implied skewness (*iskew*), implied kurtosis (*ikurt*), volume (*volume*), open interest (*openint*), turnover (*turnover*), mid price (*midprice*), market capitalization (*mcap*), bid-ask spread (*bidask*), the last holding period's option return (*ret1*), and the maximum option return over the previous holding period (*max*). In line with Büchner and Kelly (2022), we interact all characteristics with an indicator variable equal to one if the option contract is a put option and zero otherwise. As a result, we have a set of characteristic-based factors that invests in both put and call contracts and a set investing in put options only. Thus, the final set of option characteristics includes 48 characteristics. A detailed description of all option characteristics is provided in Internet Appendix B.

Beyond the characteristic-based factors defined above, we include six option-based factors that were previously proposed in explaining option returns. These include a level (*level*), moneyness skewness (*skewness*), and maturity slope (*maturity slope*) factor (Karakaya, 2014; Büchner and Kelly, 2022), the betting-against-beta (*bab*) factor (Frazzini and Pedersen, 2022), and two factor momentum factors (Käfer et al., 2023). Specifically, the level factor is an equally weighted portfolio that shorts ATM options with an absolute delta between 0.4 and 0.5. The moneyness skewness factor goes long in call options with a delta between 0.1 and 0.2 and short in put options with deltas between -0.2 and -0.1, and the maturity slope factor buys options with maturities between six and twelve months and sells options with one month to maturity. The betting-against-beta (BAB) factor is constructed following Frazzini and Pedersen (2022) and Büchner and Kelly (2022), that is, options are sorted into high and low embedded leverage portfolios based on the cross-sectional median. Options are weighted by their market capitalization (open interest times price), and the portfolios are rescaled by the value-weighted embedded leverage of the resulting long and short portfolios, respectively. Lastly, the *bab* factor is the equally weighted average of the *bab* factors constructed for call and put options separately (Frazzini and Pedersen, 2022; Büchner and Kelly, 2022). Finally, we include two factor momentum factors. Recently, Heston et al. (2022) document a momentum effect in option returns, however, Käfer et al. (2023) find that option momentum is not driven by momentum in individual options but by momentum in option factors, i.e., factor momentum (Ehsani and Linnainmaa, 2022). Thus, to account for momentum in option returns, we augment our factor set by two time-series factor momentum strategies that buy (sell) all factors described above with a positive (negative) return over the previous one (*fmom1*) and six (*fmom6*) months, respectively. We choose these formation periods because they were found in Käfer et al. (2023) to be the most important frequencies.[5]

In sum, we study a comprehensive set of 54 option-based factors and their contribution to the SDF. Figure 3 plots the annualized average CAPM-adjusted returns against the average raw factor returns. The market portfolio is proxied by the equally weighted average delta-hedged return of all available options (Horenstein et al., 2022). Table 1 in Internet Appendix C shows the average raw returns and alphas corresponding to each factor. Factors that have a significant premium after controlling for the CAPM at the 5% level are marked with filled circles. The p-values are corrected to account for the multiple testing problem (Harvey et al., 2016; Chordia et al., 2020) by controlling for the false discovery rate using the Benjamini and Hochberg (1995) method. The circles are aligned around the 45-degree line, suggesting that the alphas increase linearly with the

[5]In unreported results we find that the actual choice of the formation period or whether to use time-series or cross-sectional factor momentum does not qualitatively affect the results.

respective average returns. Thus, the CAPM fails to capture the variation associated with many factors. In fact, out of twenty-four factors that earn a significant premium, the CAPM can only explain two, thus leaving twenty-two factors unexplained. As such, the single-factor CAPM is not sufficient for pricing index options. In the following analyses, we examine how many factors are necessary and which factors capture most of the variation.

[Insert Figure 3 about here]

4 Empirical Results

4.1 A Stochastic Discount Factor for Index Option Returns

We begin our empirical analysis by studying the structure of the optimal SDF by comparing the pricing ability of SDFs with varying levels of L^1 and L^2 regularization. Because we aim to identify the drivers of cross-sectional return variation, we first orthogonalize all factors with respect to the equally weighted market portfolio (Kozak et al., 2020; Avramov et al., 2023). We then perform a 5-fold cross-validation to determine the optimal level of sparsity (L^1) and shrinkage (L^2). Note that we employ a time-series cross-validation; that is, we preserve the temporal order of the data by estimating the SDF coefficients using only information available at time t. This approach, also known as "walk-forward cross-validation" (Kaastra and Boyd, 1996; Kohzadi et al., 1996), ensures that the performance of time-series models is effectively tested without suffering from a forward looking bias. Specifically, we divide the sample into $(K+1)$ chunks of equal size. Then, we estimate the SDF coefficients using the first chunk k_1 and apply the weights to the next sample k_2. Next, we roll forward the estimation window and estimate the parameters using the data of chunk k_2 and evaluate the model on the test sample k_3. We continue this procedure until the sample ends. For each test sample, we calculate the following objective:

$$R^2_{cv} = 1 - \frac{\left(\bar{\mu}_{cv} - \bar{\Sigma}_{cv}\hat{b}\right)' \left(\bar{\mu}_{cv} - \bar{\Sigma}_{cv}\hat{b}\right)}{\bar{\mu}'_{cv}\bar{\mu}_{cv}} \tag{16}$$

with $\bar{\mu}_{cv}$ and $\bar{\Sigma}_{cv}$ denoting the mean and variance-covariance matrix of the withheld sample data. The estimated vector of SDF loadings is represented by $\hat{b}$. Because the first chunk was never used for validation, we average over $K = 5$ R^2_{cv}s and choose the regularization strengths that maximize the average R^2_{cv} (Kozak et al., 2020).

Figure 4 shows the average R^2_{cv}s for different levels of sparsity and shrinkage. Larger values on the y-axis indicate low sparsity (i.e., low L^1 regularization), while larger values on the x-axis represent a higher degree of L^2 shrinkage. As pointed out in Kozak et al. (2020), the L^2 penalty term in equation (9), γ_2, can be expressed as the root expected squared Sharpe ratio κ. Therefore, the degree of shrinkage is labeled κ. Warmer colors indicate higher R^2_{cv}s. Our results resemble those reported in Kozak et al. (2020) for stocks; that is, models without shrinkage (bottom left corner) have poor out-of-sample performance. Penalizing the estimator remarkably improves the R^2_{cv}s. We also find that the vector of SDF coefficients $\boldsymbol{b}$ is far from being sparse. Non-sparse representations having between thirty and fifty nonzero weights in the SDF produce the highest R^2_{cv}s. This finding indicates that there is no redundancy in option factor returns, rather many factors matter, thus rejecting the idea of a characteristic-sparse SDF. In our subsequent analyses, we therefore analyze the properties and pricing performance of the SDF estimator employing only L^2 regularization and consider the dual-penalty estimator as a benchmark.

[Insert Figure 4 about here]

Figure 5 shows the estimated SDF coefficients using the optimal L^2-only estimator. By far the most important factor is the maturity slope factor of Karakaya (2014), capturing the cross-sectional variation in option prices associated with changes in the implied volatility term structure (Büchner and Kelly, 2022). The other factors proposed in Karakaya (2014)—level and moneyness skewness—do not have large coefficients, indicating that they capture a smaller fraction of cross-sectional variation in option returns. As discussed in Karakaya (2014), the level factor represents a compensation for market-wide volatility and jump shock; therefore, it is closely related to the market factor (the Pearson correlation coefficient is -0.90). Because our analyses focus on the cross-sectional predictability of option returns and we consider factors that are orthogonal to the market, the level factor captures only a smaller fraction of information about cross-sectional differences in option returns. Likewise, the moneyness skewness factor of Karakaya (2014) also has a small weight in the SDF, suggesting that it also captures only limited incremental information.

The next most important factor is the one-month factor momentum factor, which achieves—according to Table 1 in Internet Appendix C—an annualized CAPM-adjusted return of 0.27% (t-stat = 4.30). Its large coefficient in the SDF indicates that factor momentum represents an incremental phenomenon that cannot be explained by other factors in the data set. The observation that buying past winner and selling past loser factors yields significant abnormal returns was first documented by Ehsani and Linnainmaa (2022). Numerous subsequent studies confirm that factor momentum is a pervasive effect, existing in industry-adjusted portfolios (Arnott et al., 2023), cryptocurrencies (Fieberg et al., 2023), and more related to this study, in stock options (Käfer et al., 2023). Thus, our finding that factor momentum exists in index options aligns with this recent strand of literature.

Next, the theta-based factor, which earns a large and statistically significant CAPM-adjusted premium of 0.34% (t-stat = 2.89), has a large coefficient in the SDF, supporting the findings in Bali et al. (2023) that theta plays a major role in the prediction of option returns. The next most important factors are characteristic-based factors formed only from put options, beginning with embedded leverage. Frazzini and Pedersen (2022) argue that investors demand lower returns for high embedded leverage options because buying options that embed leverage allows them to increase their market exposure without violating leverage constraints. The negative SDF coefficient and the negative CAPM-adjusted return of -0.54% (t-stat = -5.20) as reported in Table 1 in Internet Appendix C support this hypothesis. In line with the findings in Shafaati et al. (2022) and Büchner and Kelly (2022) that vega is a relevant driver of option returns, its large nonzero coefficient indicates that its CAPM-adjusted premium of 0.52% (t-stat = 4.65) cannot be captured by other factors. Furthermore, other factors, including the change in implied volatility factor, the put-based time-to-maturity factor, and the put-based midprice factor have large nonzero coefficients—all earning significant premia (see Table 1 in Internet Appendix C).

Interestingly, factors constructed solely from put options tend to contribute the most to the SDF, while factors constructed from both call and put options are of lower importance. Because put options act as insurance products for the underlying, they are, on the one hand, more liquid and over a wider range of strikes. Therefore, they tend to have more interpretable information priced in, whereas call options tend to have more noise. On the other hand, as insurance products, they earn returns in left-tail events with respect to the general market—especially given that the underlying is a market index. Thus, they will tend to contain a premium that is associated with jump/tail risks (and even volatility risks via the leverage effect) in the distribution of the underlying perceived by the market. Information regarding tail and jump risks is relevant to expectations of the volatility of the underlying—which is the key driver of all options—as well as shorter maturity options, for which jumps cause a significant reaction in prices and also risk sensitivities that drive subsequent returns after a jump.

In summary, our findings suggest that the option pricing puzzle is multidimensional and that many factors capture unique information about future option returns. The observation that they have nonzero coefficients in the SDF does reject the idea of a sparse SDF, whose covariances with option factors cause them to earn significant premia. The subsequent sections focus on analyzing the pricing ability of this non-sparse SDF compared to sparse models.

[Insert Figure 5 about here]

4.2 Asset Pricing Tests

Having determined that a non-sparse SDF achieves the best pricing performance in terms of statistical measures, we inquire about the economic asset pricing performance of the SDF-implied MVP. Because the SDF coefficients are equivalent to the weights in the MVP (Back, 2010), our asset pricing tests focus on the efficient frontier spanned by the SDF. Comparing asset pricing models in terms of MVP performance instead of focusing on alphas of some test assets is also favored in, for example, Barillas and Shanken (2018), Soebhag et al. (2022), and Kozak et al. (2020). However, in Section 4.3, we also test the pricing ability of the candidate SDFs for the set of factors studied in this paper.

We compare the performance of the MVP implied by the non-sparse SDF employing L^2 regularization with various benchmark MVPs, beginning with an MVP obtained from the dual-penalty estimator. The other benchmark models include characteristic-sparse representations implied by empirical factor models, including the Karakaya (2014) three-factor (K3), Horenstein et al. (2022) three-factor (HVX3), and Bali et al. (2022) five-factor (BCCSZ5) models, as well as statistical asset pricing models, namely, RP-PCA (RPPCA3), and IPCA (IPCA3), both including three factors.[6] For these benchmark models, the SDF-implied portfolio is obtained by first orthogonalizing the factors (or their projections) with respect to the market. Then, MVP weights are estimated according to equation (8), i.e., $\hat{\boldsymbol{b}} = \bar{\boldsymbol{\Sigma}}_f^{-1}\bar{\boldsymbol{\mu}}_f$ (Kozak et al., 2020).

To avoid the potential of overfitting, we compare out-of-sample performance metrics of the candidate MVPs. Specifically, our analysis focuses on recursive model estimation using only data known at time t (Lewellen, 2015; Kelly et al., 2019; Büchner and Kelly, 2022). That is, we estimate the SDF coefficients using in-sample data and apply these to out-of-sample observations. Our out-of-sample period begins ten years after our sample period starts, i.e., in January 2006, and ends in December 2022. Instead of re-estimating the MVP weights daily, we rebalance the portfolios each third Friday of a month to reduce turnover and computational costs.[7] That is, we use all data available from January 1996 to the 20th of January 2006 (the beginning of our first out-of-sample holding period) to estimate the MVP weights. The optimal level of regularization is selected using a 5-fold time-series cross-validation as described above. We obtain the out-of-sample MVP returns by multiplying the weights with the factor returns over the holding period. Then, we re-estimate

[6] Note that we estimate RP-PCA and IPCA with three factors because Büchner and Kelly (2022) find that a three-factor IPCA best prices S&P500 index options and Horenstein et al. (2022) identify a three-factor RP-PCA model for stock options. In unreported results, we also estimate models with $K = 1, \ldots 6$ factors, however, we can confirm their results that the three-factor models perform best. Both RPPCA3 and IPCA3 are estimated based on the same set of CMPs used for estimating the penalized SDF.

[7] This rebalancing period aligns with our factor construction described in Section 3.3.

all parameters, including the most recent observations used for estimation.[8] Note that, as in our previous analyses, we first orthogonalize the factors with respect to the market return using only in-sample information. Furthermore, we follow Kozak et al. (2020) and Avramov et al. (2023) and re-scale the MVP weights to target the in-sample volatility of the market portfolio.[9]

Table 2 shows the annualized alphas of the non-sparse MVP against the benchmark models. Note that because the factors are orthogonal to the market portfolio, the average return of the MVP equals the CAPM alpha, which amounts to 1.40% p.a. and is statistically significant at the 1% level (t-stat = 4.80). The alphas against the empirical benchmark models are also large and statistically significant: The abnormal return compared to the HVX3 and BCCSZ5 models is 1.53% (t-stat = 6.65) and 1.40% (t-stat = 4.81), respectively, whereas the alpha against the K3 model is only about half as large, i.e., 0.66%. However, the alpha against the K3 model is statistically significant at the 1% level (t-stat = 4.07). The next two columns show the abnormal returns compared to the two statistical asset pricing models that first rotate the factors into a low-dimensional latent factor space to achieve a dimensionality reduction from 54 to three factors, namely, RPPCA3 and IPCA3. The alpha against the IPCA-based MVP is 1.15% (t-stat = 4.77), while the alpha against the RPPCA3 is only 0.30% but statistically significant (t-stat = 2.35). The results indicate that RPPCA3 is a strong competitor, however, it fails to explain the returns of the MVP implied by the non-sparse SDF.

The last column shows the performance compared to an MVP implied by an SDF allowing for both sparsity and shrinkage. The non-sparse MVP employing only L^2 regularization outperforms the dual-penalty estimator with an alpha of 0.34%, which is significant at the 1% level (t-stat = 2.68). The finding that the MVP obtained from the L^2 estimator outperforms the MVP estimated using the dual-penalty thus supports the findings from Figure 4; rejecting the idea of a sparse SDF that prices option returns.

[8] In Table 2 in Internet Appendix C, we show results using a rolling fixed-length estimation window.

[9] Because the loadings in the SDF are proportional to the inverse of $\hat{\boldsymbol{\Sigma}}_f$ times $\hat{\boldsymbol{\mu}}_f$, the coefficients can become extremely large, producing an MVP with unrealistic mean and volatility.

[Insert Table 2 about here]

Figure 6 visualizes the abnormal performance of the non-sparse SDF-implied MVP over time. The outperformance against the CAPM, HVX3, BCCSZ5, and IPCA3 models is clearly visible. Likewise, although weaker, the non-sparse SDF consistently generates abnormal returns against the dual-penalty estimator and the K3 and RPPCA3 models. Interestingly, the abnormal returns relative to alternative MVPs do not move in the same way. Taking into account the drop in abnormal returns at the beginning of 2020, the abnormal returns compared to all models, except the dual-penalty and K3 model, decrease, while the abnormal returns against the MVP implied by the dual-penalty estimator and K3 model do not decline. This indicates that these MVPs perform poorly in times of crises. We further address the time-series variation in MVP profitability in Section 5.3.

[Insert Figure 6 about here]

The statistically significant alphas against benchmark models indicate that the non-sparse SDF-implied MVP cannot be spanned by other models; however, is it really mean-variance efficient? To address this question, we perform a test of mean-variance efficient frontier expansion in the spirit of Novy-Marx and Velikov (2016) and Soebhag et al. (2022), based on the out-of-sample performance improvement obtained by combining two MVPs.

Consider two MVPs, labeled as MVP^A and MVP^B, respectively. If MVP^A exhibits mean-variance efficiency, it should achieve a higher Sharpe ratio compared to all other available investment opportunities (Barillas and Shanken, 2018). In other words, a portfolio that combines investments from both MVP^A and MVP^B (referred to as $MVP^{A,B}$) should not surpass the performance of MVP^A. To assess the mean-variance efficiency of an MVP, we obtain the generalized alpha from the following time-series regression (Novy-Marx and Velikov, 2016; Soebhag et al., 2022):

$$MVP_t^{A,B} = \alpha + \beta MVP_t^A + \epsilon_t \tag{17}$$

with α and β denoting the regression coefficients and ϵ the residuals at time t. The generalized alpha (α) can be interpreted as the abnormal return due to the addition of MVP^B to the investment opportunity set. Therefore, a positive and statistically significant α indicates that the performance of MVP^A can be improved by augmenting the investment opportunity set with MVP^B, i.e., MVP^A is not mean-variance efficient (Novy-Marx and Velikov, 2016; Barillas and Shanken, 2018; Soebhag et al., 2022). Denote $\mathbf{\Sigma}^{A,B}$ as the covariance matrix and $\boldsymbol{\mu}^{A,B}$ as the vector of the average returns of the portfolios MVP^A and MVP^B. The optimal weights of both assets in $MVP_t^{A,B}$ are obtained according to equation (3), i.e., $\hat{\boldsymbol{b}}^{A,B} = \left(\mathbf{\Sigma}^{A,B}\right)^{-1} \boldsymbol{\mu}^{A,B}$. Again, the weights are rescaled to target the in-sample volatility of the market portfolio.

Table 3 reports the results. The first row shows the non-sparse SDF employing L^2 shrinkage. When adding other MVPs to the non-sparse MVP, the alphas are small and insignificant, ranging between -0.17% and 0.25% p.a., indicating that no other MVP improves the span of the efficient frontier. Hence, the non-sparse SDF-implied MVP is mean-variance efficient. In contrast, when added to any benchmark model—except to the MVP obtained from the dual-penalty estimator—the non-sparse SDF improves the span of the efficient frontier, suggesting that benchmark models are not mean-variance efficient. However, the MVP obtained from the dual-penalty estimator has a statistically significant generalized alpha of 0.48% (t-stat = 2.60) against the RPPCA3 model, rejecting the mean-variance efficiency of this portfolio.

[Insert Table 3 about here]

4.3 Pricing Performance for Anomalies

The previous analyses show that the MVP implied by the non-sparse SDF exhibits mean-variance efficiency, while benchmark models fail to span the same or an even higher efficient frontier. According to Barillas and Shanken (2018), this should imply that a non-sparse SDF best prices assets. To test this, we estimate the abnormal returns of the option factors against the MVPs for the non-sparse SDF, dual-penalty SDF, K3, BCCSZ5, RPPCA3, and IPCA3 models, i.e., we regress each of the option factors—which are orthogonal to the equally weighted market portfolio—on the time series of ex ante MVP returns and obtain the alpha.

Figure 7 plots the factor alphas (y-axis) against the CAPM alphas (x-axis) over the out-of-sample period. The alphas of the empirical benchmark models—K3 and BCCSZ5—are large and aligned around the 45-degree line, indicating that the alphas increase linearly with the CAPM alphas. The average absolute alphas are 0.18% and 0.21%, respectively, and, assuming a 5% significance level, the models leave fourteen and seventeen factors unexplained. Similarly, the RPPCA3 (IPCA3) model fails to explain eighteen (sixteen) factors and produces average absolute alphas of 0.19% (0.20%). Thus, the rejection rates are far above the 5% false discovery rate.

The non-sparse model in Panel A does a much better job. It has an average absolute alpha of 0.13% and leaves only two factors unexplained, namely, the implied kurtosis (p-value = 0.48%) and the one-month factor momentum (p-value = 1.17%). The relatively week performance of the SDF obtained from the dual-penalty estimator is also reflected in its explanation for option factors: The SDF employing both L^1 and L^2 regularization has an average absolute alpha of 0.16%, while leaving nine factors unexplained. Thus, the results confirm our earlier finding that a non-sparse SDF is best suited to price delta-hedged index option returns.

[Insert Figure 7 about here]

4.4 Out-of-Sample Importance

In Section 4.1, we studied the structure of the optimal SDF based on parameter estimation using the entire sample period from 1996 to 2022. We identified that several factors have nonzero coefficients in the SDF, including the *maturity slope*, *embedlev:put*, *vega:put*, *theta*, *ttm:put*, *midprice:put*, and *fmom1* factors. To analyze whether these factors explain the superior performance of the SDF-implied MVP found in the previous analyses, we perform an out-of-sample importance analysis as a robustness check. Specifically, for each out-of-sample holding period, we obtain the MVP weights following the same recursive model estimation scheme described above and set the weight of one factor to zero, while fixing the weight of the other factors (Kelly et al., 2019; Gu et al., 2020, 2021; Büchner and Kelly, 2022). We then obtain the SDF-implied MVP returns and measure the contribution of a factor to the SDF-implied MVP by the reduction of the annualized Sharpe ratio as a result of setting the weight to zero.

Figure 8 reports the results, which confirm our overall results. The factors that mostly contribute to the performance of the MVP are the maturity slope factor of Karakaya (2014), the put-based embedded leverage, vega, and midprice factors, the change in the implied volatility factor, and the one-month factor momentum factor.

[Insert Figure 8 about here]

5 Robustness Checks

This section tests the robustness of our findings and provides further insights on the pricing performance of the identified SDF. Specifically, in Section 5.1, we expand the set of candidate factors to include interactions and nonlinear transformations of the option characteristics. Section 5.2 analyzes the pricing ability of the SDF for a subset of liquid options. The performance of the MVP over time is analyzed in more detail in Section 5.3 and in Section 5.4, we compare the performance of the MVP against stock market factors.

5.1 Large Set of Nonlinear Characteristics

Estimating the SDF as described in Section 2 allows it SDF to depend on a large set of candidate factors, including hundreds or even thousands of factors. In the previous analyses, we tested a moderate set of 54 factors and found that many of them contribute to the SDF, thus spanning the highest mean-variance efficient frontier among all benchmark factor pricing models. In constructing the characteristic-based factors, we relied on the linear relationship of future delta-hedged option returns and the options' characteristics; however, Bali et al. (2023) show that nonlinearities matter when modeling option returns. As such, the SDF-implied MVP may not span the highest achievable efficient frontier if we neglect nonlinear information. Therefore, we next test whether adding nonlinearities and interactions of characteristics to the factor set enhances the performance of the tangency portfolio.

Our construction of nonlinear characteristics and interactions follows Kozak et al. (2020), that is, we include the second and third power of all characteristics, as well as their first-order interactions. Note that our initial factor set is based on 24 option characteristics, doubled by creating interactions with a put indicator variable. For all $L = 24$ characteristics and for the 24 characteristics interacted with the put indicator variable, we include the first-order interactions, resulting in $L\,(L-1) = 552$ additional factors.

Denote $z^k_{i,t-1}$ and $z^m_{i,t-1}$ as two rank-transformed characteristics k and m of contract i at time $t-1$. The first-order interaction characteristic $z^{k,m}_{i,t-1}$ is the product of these characteristics. We additionally include nonlinear transformations of the characteristics by taking the second and third powers of them, thus leading to additional $4L = 96$ factors. The interactions and nonlinear transformations are normalized according to equation (14); however, we do not re-rank the options. As Kozak et al. (2020) point out, this approach is closely related to bivariate portfolio sorts and further not re-ranking the transformed characteristics allows, for example, the cubic transformation to have larger exposure to options with extreme realizations of the base characteristic but with the same leverage. In summary, our extended factor set consists of 48 characteristic-based factors used in the previous analyses, 552 interaction-based factors, 96 nonlinear characteristic-based factors, and six non-characteristic-based factors, amounting to 702 factors.

Figure 1 in Internet Appendix C shows the OOS R^2s of the dual-penalty estimator depending on the degree of L^1 and L^2 regularization, when applied to the extended set of 702 factors, and Figure 2 shows the fifteen largest coefficients in the dual-penalty estimator. Qualitatively, the results support the findings from Section 4.1: A sparse SDF is not sufficient for pricing index option returns. We find that the same factors as identified before as well as their nonlinear transformations contribute the most to the SDF. Specifically, the maturity slope factor weighs the most, followed by the third power of the *embedlev:put* and *vega:put* factors. However, their linear counterparts also have a large coefficient in the SDF. Likewise, the linear *midprice:put*, *ttm:put*, and *implvol_ch* factors and their cubic transformations also appear in the SDF. However, does the nonlinear extension of the factor set improve the span of the efficient frontier? To answer this question, we perform several mean-variance efficient frontier expansion tests as described in equation (17). Specifically, we test whether the span of the efficient frontier can be improved by extending the linear factor set (z_i) by a) quadratic (z_i^2), b) cubic (z_i^3), or c) quadratic and cubic (z_i^2, z_i^3) characteristic transformations as well as by d) first-order interactions ($z_i z_j$), and e) quadratic and cubic extensions and interactions ($z_i^2, z_i^3, z_i z_j$). Thus, we successively analyze whether nonlinear effects in characteristics or their interactions improve the performance of the MVP, instead of testing all 702 factors at once, which could lead to an inefficient estimation of the covariance matrix or its inverse, respectively.

Table 4 reports the generalized alphas as defined in equation (17). When extending the linear factor set to include nonlinear characteristic-based factors, the absolute alphas are small, ranging between 0.04% and 0.44% p.a. These alphas are not significant at any conventional significance level, suggesting that considering nonlinear factors in addition to linear ones does not improve the span of the efficient frontier.

[Insert Table 4 about here]

Having shown that nonlinear factor extensions do not improve the span of the efficient frontier compared to the linear factor set, we also ask how this translates into pricing errors for the set of 702 factors. Figure 9 shows the factor alphas against the SDF-implied MVP including 54 linear factors (Panel A) and the MVP including all 702 factors (Panel B). The results support the findings from Table 4 that the nonlinear extension does not improve the pricing performance of the SDF. The SDF having nonzero coefficients for only linear factors fails to capture the alphas of ten factors, leaving an average absolute alpha of 0.20%. In contrast, the SDF including all 702 factors fails to explain 63 factors with an average absolute alpha of 0.24%.

[Insert Figure 9 about here]

To conclude, this robustness check shows that nonlinear extensions of the initial factor set do not improve the span of the efficient frontier and that the SDF estimated from only 54 factors proves to have low pricing errors for an extremely large set of test factors, thereby corroborating our main results.

5.2 Excluding Illiquid Options

Avramov et al. (2023) show for stock market data that the SDF takes extreme weights in stocks that are difficult to arbitrage, i.e., small and distressed stocks that are less likely to be of interest to investors and are difficult to trade. This robustness analysis examines the extent to which our results are driven by difficult-to-arbitrage options. Note that our filters described in Section 3.1 already ensure that we consider relatively liquid options; however, we now tighten our filter and only include those 50% of option contracts each holding period that have the lowest bid-ask spreads, as these options face lower trading costs and are more liquid.

Figure 10 plots the factor alphas of the 54 factors constructed from a subset of liquid options against the SDF-implied MVP from our main analyses. We find that the SDF fails to explain only one factor, i.e., the one-month factor momentum (p-value = 2.94%), which is in line with the false discovery rate of 5%. The average absolute alpha is only 0.09%, suggesting that the pricing performance of the SDF does not stem from a small subset of illiquid options, but that the SDF generalizes well to a set of liquid contracts.

[Insert Figure 10 about here]

5.3 Dependence on Market States

In Figure 6, we show that the non-sparse MVP consistently outperforms the benchmark MVPs over time. However, noting some declines in abnormal profitability in specific phases, we take a more formal look at the alphas by dividing the out-of-sample period into subperiods. Studies on stock return predictability find that the performance of factors depends on certain market states captured by investor sentiment (Stambaugh et al., 2012; Avramov et al., 2019), volatility (Nagel, 2012), and illiquidity (Chordia et al., 2014). In addition, Avramov et al. (2023) show that sophisticated machine learning methods derive profitability during periods when limits to arbitrage are high. To disentangle whether the performance of the SDF-implied MVP also depends on different market phases, we use a battery of proxies to split the time series into subsamples. First, we divide the sample into price (volatility) jump and non-jump periods, which are defined as holding periods in which the return of the S&P500 is smaller than -4% (the change in the VIX is larger than 4%) (Karakaya, 2014). Second, Kirchler (2009) and Asem and Tian (2010) link past market returns with investors' confidence. Therefore, we proxy for investor sentiment by defining bear and bull markets as periods in which the 12-month trailing market return is below (above) the sample median. Third, we define the low, medium, and high VIX regimes based on the tercile distribution of the historical VIX over the sample period. Fourth, we analyze the performance of the MVP in "recession" and "non-recession" periods defined according to the NBER recession dates.[10] Finally, we proxy for uncertainty in economic conditions by dividing the sample into low and high uncertainty periods based on the median of the Bekaert et al. (2022) uncertainty index.

Table 5 reports the annualized alphas of the MVP implied by the non-sparse SDF against the benchmark models. The non-sparse MVP achieves a significant market-adjusted abnormal return in all subperiods, except during recession periods. In market states that are characterized by price jumps, the alpha is significant at least at the 10% level (t-stat = 1.90), which may be due to the reduced number of time-series observations. The average returns are particularly high in bull markets (1.82%) and non-recession periods (1.67%); however, also in periods with bad economic conditions, including a high VIX (1.86%) and high uncertainty (1.99%). These findings are in line with Bali et al. (2023) who show that machine learning methods are more profitable in high

[10] https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions

volatility regimes.

The results in Table 5 highlight the superiority of the MVP against the benchmark portfolios even during these subperiods. Again, the MVP achieves remarkable alphas against benchmark models in any subperiod except in the price jump and recession periods. The K3 model has significantly lower alphas compared to the CAPM, resulting in alphas that are only half as large or even smaller; however, the non-sparse MVP attains significant alphas in eight out of thirteen market states. The RPPCA3 leads to even smaller alphas. Only in the price jump, non-volatility jump, and non-recession periods, the MVP is unexplained by RPPCA3. In summary, the non-sparse SDF-implied MVP attains remarkable alphas in many subperiods but fails to outperform the K3 and RPPCA3 in many market states. However, it is never outperformed by any competing MVP.

[Insert Table 5 about here]

5.4 Other Factor Sets

Due to the link of options to their underlying, studies on option return predictability often use stock market factors or asset pricing factors from other asset classes to capture the variation in option returns. Table 6 reports the annualized alphas of the SDF-implied MVP against the stock market CAPM, the Fama and French (1993) three-factor (FF3), Fama and French (2015) five-factor (FF5), and Fama and French (2018) six-factor (FF6) model,[11] as well as the factors of the augmented q-factor model of Hou et al. (2015, 2021) (HVX),[12] the mispricing factor model of Stambaugh and Yuan (2017) (SY),[13] the three factors of Daniel et al. (2020a) (DHS),[14] and, lastly, the five factors of Daniel et al. (2020b) (DMRS).[15] The results show that none of the stock market asset pricing models capture the returns of the MVP, suggesting that the option-based factors capture unique

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

[12] https://global-q.org/factors.html

[13] The Stambaugh and Yuan (2017) factor data is only available up to 2016, therefore, we obtain MGMT and PERF factors from Jensen et al. (2023) via https://jkpfactors.com/

[14] https://sites.google.com/view/linsunhome

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

information. The results are in line with the previous literature showing that stock factors perform poorly in explaining option returns (Büchner and Kelly, 2022; Bali et al., 2023).

[Insert Table 6 about here]

6 Conclusion

In contrast to prior attempts in the literature to capture the cross-sectional variation in delta-hedged option returns using low-dimensional models that include only a handful of factors, our empirical findings suggest that dozens of factors contribute to the SDF and, in turn, span the mean-variance efficient frontier. Using the dual-penalty estimator proposed in Kozak et al. (2020), our findings highlight that shrinking MVP weights helps to improve mean-variance efficiency of the implied tangency portfolio; however, the idea of a sparse SDF is rejected. A non-sparse SDF-implied MVP outperforms sparse representations and exhibits lower pricing errors, even in subperiods or after excluding illiquid options.

Of the extensive factor set studied, some factors dominate, including a maturity slope and factor momentum factor, along with put-based factors that exploit the characteristic spreads in embedded leverage, vega, time-to-maturity, and option price; however, theta- and implied volatility change-based factors constructed from both call and put options also contribute largely to the SDF. While the empirical findings reveal that many factors that are linear in option characteristics appear in the SDF, nonlinear transformations or interactions of characteristics do not add further information.

By studying S&P500 index options, we observe that the option pricing puzzle is highly multidimensional. Moreover, by studying options that have the same underlying, we minimized any effect related to the microstructure of different underlyings. The recent literature on, for example, stock options shows that stock characteristics also drive the variation in delta-hedged option returns (Zhan et al., 2022), although option-based characteristics appear to be the most important drivers (Bali et al., 2023). While our study identifies several relevant option-based factors as drivers of cross-sectional variation in index option returns, some questions remain unanswered. Do the results obtained from this study also hold for stock options? And, more interestingly, do option factors based on stock characteristics matter after accounting for option-based characteristics? These questions will be interesting to address in future studies.

References

Arnott, R.D., Kalesnik, V., Linnainmaa, J.T., 2023. Factor momentum. Review of Financial Studies 36, 3034–3070.

Asem, E., Tian, G.Y., 2010. Market dynamics and momentum profits. Journal of Financial and Quantitative Analysis 45, 1549–1562.

Avramov, D., Cheng, S., Metzker, L., 2023. Machine learning vs. economic restrictions: Evidence from stock return predictability. Management Science 69, 2587–2619.

Avramov, D., Chordia, T., Jostova, G., Philipov, A., 2019. Bonds, stocks, and sources of mispricing. Working Paper.

Back, K.E., 2010. Asset Pricing and Portfolio Choice. Oxford University Press.

Bakshi, G., Kapadia, N., 2003. Delta-hedged gains and the negative market volatility risk premium. Review of Financial Studies 16, 527–566.

Bali, T.G., Beckmeyer, H., Mörke, M., Weigert, F., 2023. Option return predictability with machine learning and big data. Review of Financial Studies 36, 3548–3602.

Bali, T.G., Cao, J., Chabi-Yo, F., Song, L., Zhan, X., 2022. A factor model for stock options. Working Paper.

Barillas, F., Shanken, J., 2018. Comparing asset pricing models. Journal of Finance 73, 715–754.

Bekaert, G., Engstrom, E.C., Xu, N.R., 2022. The time variation in risk appetite and uncertainty. Management Science 68, 3975–4004.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological) 57, 289–300.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Boyer, B.H., Vorkink, K., 2014. Stock options as lotteries. Journal of Finance 69, 1485–1527.

Bryzgalova, S., Huang, J., Julliard, C., 2023. Bayesian solutions for the factor zoo: We just ran two quadrillion models. Journal of Finance 78, 487–557.

Büchner, M., Kelly, B.T., 2022. A factor model for option returns. Journal of Financial Economics 143, 1140–1161.

Cao, C., Huang, J.Z., 2007. Determinants of S&P 500 index option returns. Review of Derivatives Research 10, 1–38.

Cao, J., Han, B., 2013. Cross section of option returns and idiosyncratic stock volatility. Journal of Financial Economics 108, 231–249.

Chordia, T., Goyal, A., Saretto, A., 2020. Anomalies and false rejections. Review of Financial Studies 33, 2134–2179.

Chordia, T., Subrahmanyam, A., Tong, Q., 2014. Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? Journal of Accounting and Economics 58, 41–58.

Christoffersen, P., Goyenko, R., Jacobs, K., Karoui, M., 2018. Illiquidity premia in the equity options market. Review of Financial Studies 31, 811–851.

Daniel, K., Hirshleifer, D., Sun, L., 2020a. Short- and long-horizon behavioral factors. Review of Financial Studies 33, 1673–1736.

Daniel, K., Mota, L., Rottke, S., Santos, T., 2020b. The cross-section of risk and returns. Review of Financial Studies 33, 1927–1979.

Dickerson, A., Julliard, C., Mueller, P., 2023. The corporate bond factor zoo. Working Paper.

Didisheim, A., Ke, S., Kelly, B.T., Malamud, 2024. Complexity in factor pricing models. Working Paper.

Dong, X., Li, Y., Rapach, D.E., Zhou, G., 2022. Anomalies and the expected market return. Journal of Finance 77, 639–681.

Ehsani, S., Linnainmaa, J.T., 2022. Factor momentum and the momentum factor. Journal of Finance 77, 1877–1919.

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

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E.F., French, K.R., 2018. Choosing factors. Journal of Financial Economics 128, 234–252.

Fieberg, C., Liedtke, G., Metko, D., Zaremba, A., 2023. Cryptocurrency factor momentum. Quantitative Finance 23, 1853–1869.

Frazzini, A., Pedersen, L.H., 2022. Embedded leverage. Review of Asset Pricing Studies 12, 1–52.

Goyal, A., Saretto, A., 2009. Cross-section of option returns and volatility. Journal of Financial Economics 94, 310–326.

Goyal, A., Saretto, A., 2022. Are equity option returns abnormal? IPCA says no. Working Paper.

Goyenko, R., Zhang, C., 2022. The joint cross section of option and stock return predictability with big data and machine learning. Working Paper.

Green, J., Hand, J.R.M., Zhang, X.F., 2017. The characteristics that provide independent information about average U.S. monthly stock returns. Review of Financial Studies 30, 4389–4436.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Review of Financial Studies 33, 2223–2273.

Gu, S., Kelly, B., Xiu, D., 2021. Autoencoder asset pricing models. Journal of Econometrics 222, 429–450.

Hansen, L.P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. Journal of Political Economy 99, 225–262.

Harvey, C.R., Liu, Y., Zhu, H., 2016. . . . and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

Hastie, T., Tibshirani, R., Friedman, J., 2011. The elements of statistical learning. Springer, New York.

Heston, S.L., Jones, C.S., Khorram, M., Li, S., Mo, H., 2022. Option momentum. Journal of Finance, Forthcoming.

Horenstein, A.R., Vasquez, A., Xiao, X., 2022. Common factors in equity option returns. Working Paper.

Hou, K., Mo, H., Xue, C., Zhang, L., 2021. An augmented q-factor model with expected growth. Review of Finance 25, 1–41.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28, 650–705.

Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. Review of Financial Studies 33, 2019–2133.

Hu, G., Jacobs, K., 2020. Volatility and expected option returns. Journal of Financial and Quantitative Analysis 55, 1025–1060.

Hu, G., Liu, Y., 2022. The pricing of volatility and jump risks in the cross-section of index option returns. Journal of Financial and Quantitative Analysis 57, 2385–2411.

Israelov, R., Kelly, B.T., 2017. Forecasting the distribution of option returns. Working Paper.

Jensen, T.I., Kelly, B.T., Pedersen, L.H., 2023. Is there a replication crisis in finance? Journal of Finance 78, 2465–2518.

Kaastra, I., Boyd, M., 1996. Designing a neural network for forecasting financial and economic time series. Neurocomputing 10, 215–236.

Käfer, N., Moerke, M., Wiest, T., 2023. Option factor momentum. Working Paper.

Kanne, S., Korn, O., Uhrig-Homburg, M., 2023. Stock illiquidity and option returns. Journal of Financial Markets 63, 100765.

Karakaya, M.M., 2014. Characteristics and expected returns in individual equity options. Working Paper.

Kelly, B.T., Malamud, S., Zhou, K., 2022. The virtue of complexity everywhere. Working Paper.

Kelly, B.T., Malamud, S., Zhou, K., 2024. The virtue of complexity in return prediction. Journal of Finance 79, 459–503.

Kelly, B.T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134, 501–524.

Kirchler, M., 2009. Underreaction to fundamental information and asymmetry in mispricing between bullish and bearish markets. An experimental study. Journal of Economic Dynamics and Control 33, 491–506.

Kohzadi, N., Boyd, M.S., Kermanshahi, B., Kaastra, I., 1996. A comparison of artificial neural network and time series models for forecasting commodity prices. Neurocomputing 10, 169–181.

Kozak, S., Nagel, S., Santosh, S., 2020. Shrinking the cross-section. Journal of Financial Economics 135, 271–292.

Käfer, N., Moerke, M., Weigert, F., Wiest, T., 2024. A bayesian stochastic discount factor for the cross-section of individual equity options. Working Paper.

Lemmon, M., Ni, S.X., 2014. Differences in trading and pricing between stock and index options. Management Science 60, 1985–2001.

Lettau, M., Pelger, M., 2020. Factors that fit the time series and cross-section of stock returns. Review of Financial Studies 33, 2274–2325.

Lewellen, J., 2015. The cross-section of expected stock returns. Critical Finance Review 4, 1–44.

Nagel, S., 2012. Evaporating liquidity. Review of Financial Studies 25, 2005–2039.

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

Novy-Marx, R., Velikov, M., 2016. A taxonomy of anomalies and their trading costs. Review of Financial Studies 29, 104–147.

Ruan, X., 2020. Volatility-of-volatility and the cross-section of option returns. Journal of Financial Markets 48, 100492.

Shafaati, M., Chance, D.M., Brooks, R., 2022. The cross-section of individual equity option returns. Working Paper.

Soebhag, A., van Vliet, B., Verwijmeren, P., 2022. Non-standard errors in asset pricing: Mind your sorts. Working Paper.

Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288–302.

Stambaugh, R.F., Yuan, Y., 2017. Mispricing factors. Review of Financial Studies 30, 1270–1315.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58, 267–288.

Vasquez, A., 2017. Equity volatility term structures and the cross section of option returns. Journal of Financial and Quantitative Analysis 52, 2727–2754.

Zhan, X., Han, B., Cao, J., Tong, Q., 2022. Option return predictability. Review of Financial Studies 35, 1394–1442.

Zho, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society. Series B (Methodological) 67, 301–320.

Figure 1: t-Statistics of Out-of-Sample Factor Alphas

The figure shows the t-statistics for alphas derived from a time-series regression of option factors on ex ante MVPs implied by various SDFs. These include the SDF proposed in this paper (highlighted by filled circles) that uses L^2 regularization, as well as an SDF estimated with both L^1 and L^2 regularization, and the SDF representations of the CAPM, K3, HVX3, BCCSZ5, RPPCA3, and IPCA3 models. The t-statistics are adjusted using the Newey and West (1987) method. Additionally, the multiple hypotheses testing problem is addressed by adjusting the p-values through the Benjamini and Hochberg (1995) method. The number above the circles denote the number of alphas that are significant at the 5% level. All results are based on out-of-sample estimates spanning from January 2006 to December 2022.

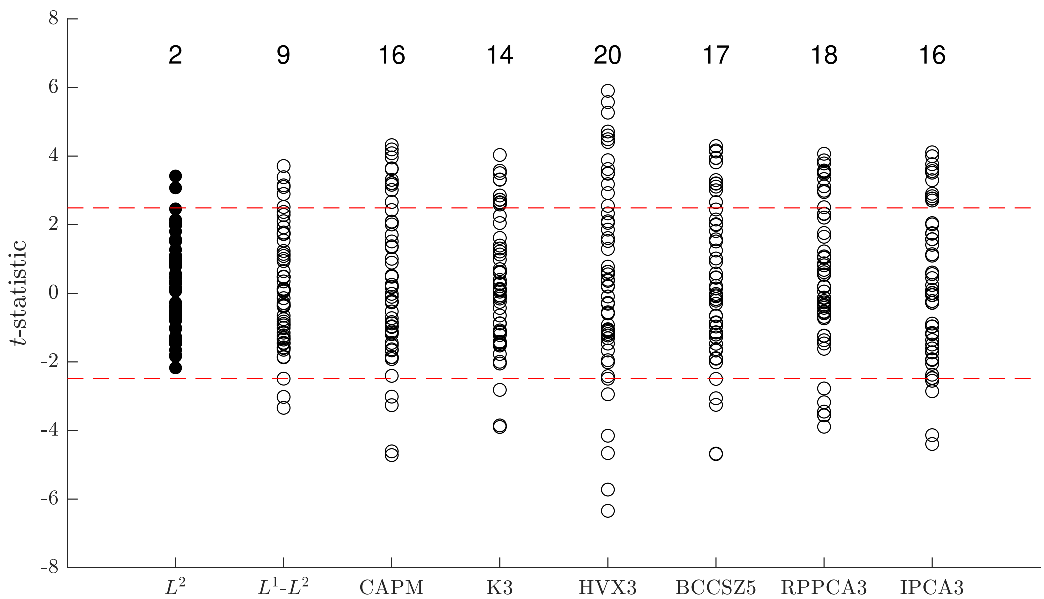
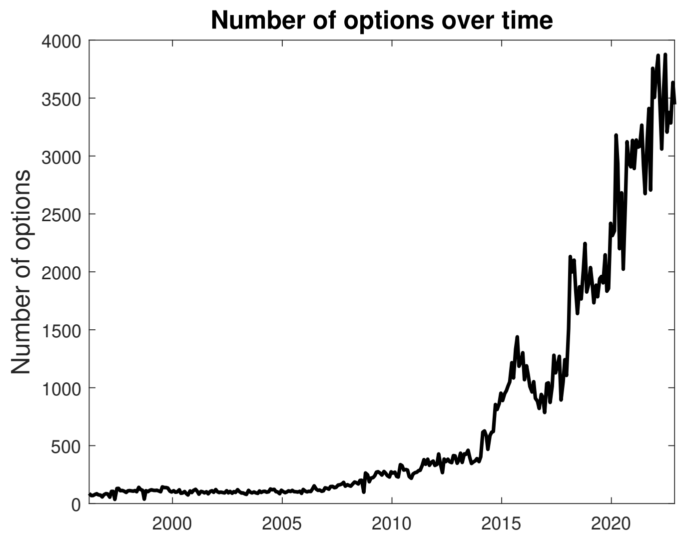


Figure 2: Sample Properties Over Time

The figure shows (a) the daily number and (b) the aggregate open interest of the options covered in this study. The study period runs from January 1996 to December 2022.

(a) Number of option contracts



(b) Open interest

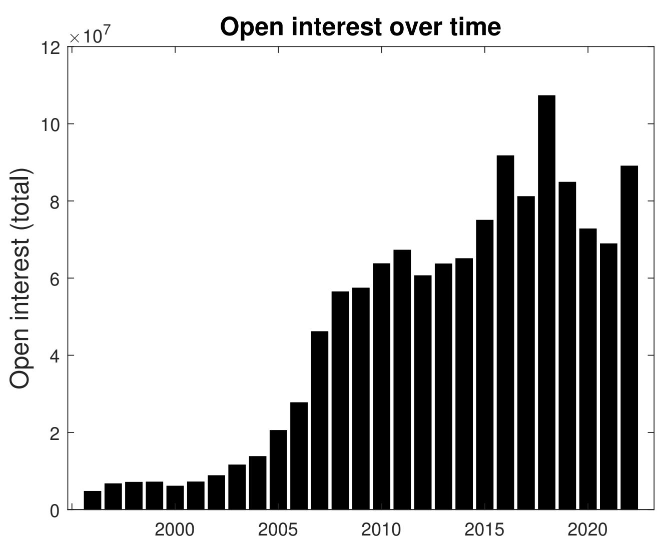


Figure 3: CAPM Alphas of Option-based Factors

The figure shows the annualized average factor returns (x-axis) and the annualized average CAPM-adjusted returns (y-axis) for all 54 option-based factors considered. Factors that have a significant CAPM alpha at the 5% level are marked by filled circles. The t-statistics are adjusted using the Newey and West (1987) method. To control for multiple hypothesis testing, the p-values are adjusted using the Benjamini and Hochberg (1995) method. All returns are expressed in percentage terms. The analysis covers the period from January 1996 to December 2022.

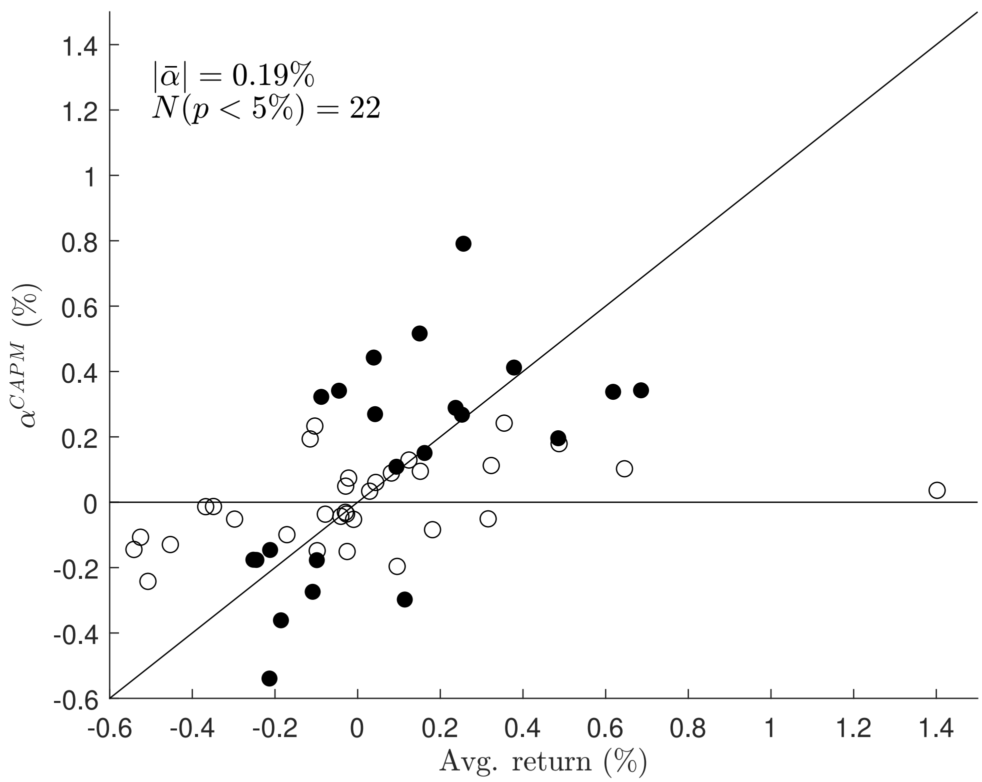


Figure 4: R^2_{cv}s from Dual-Penalty Specification

The figure presents cross-validation R^2_{cv} values for models employing varying degrees of L^1 and L^2 regularization, applied to 54 option-based factors. The x-axis represents the degree of L^2 regularization, while the y-axis portrays the degree of L^1 regularization. The level of shrinkage (L^2) is expressed as the root expected squared Sharpe ratio κ and the degree of sparsity (L^1) is represented by the number of factors that have a nonzero coefficient in the SDF. Elevated R^2_{cv} values are denoted by warmer colors. The analysis covers the study period from January 1996 to December 2022.

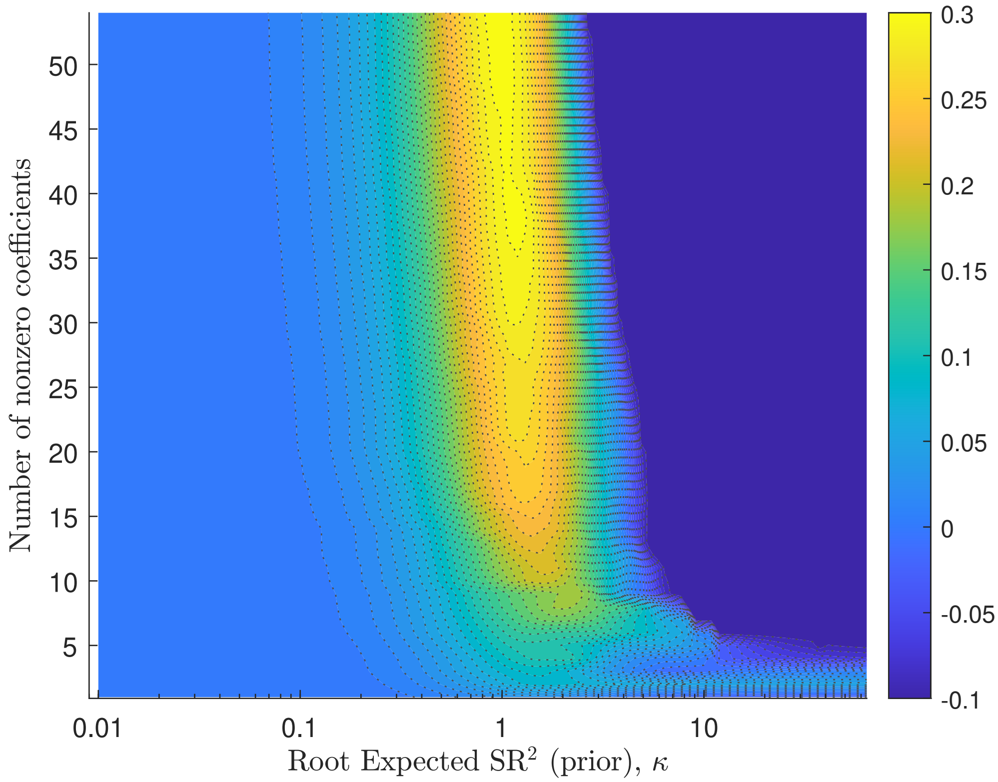


Figure 5: SDF Coefficients

The figure shows the contribution of all factors in the SDF estimated using L^2 regularization. Coefficients are sorted descending on their absolute SDF coefficient. The optimal level of L^2 regularization is selected using a 5-fold time-series cross-validation. The sample period is from January 1996 to December 2022.

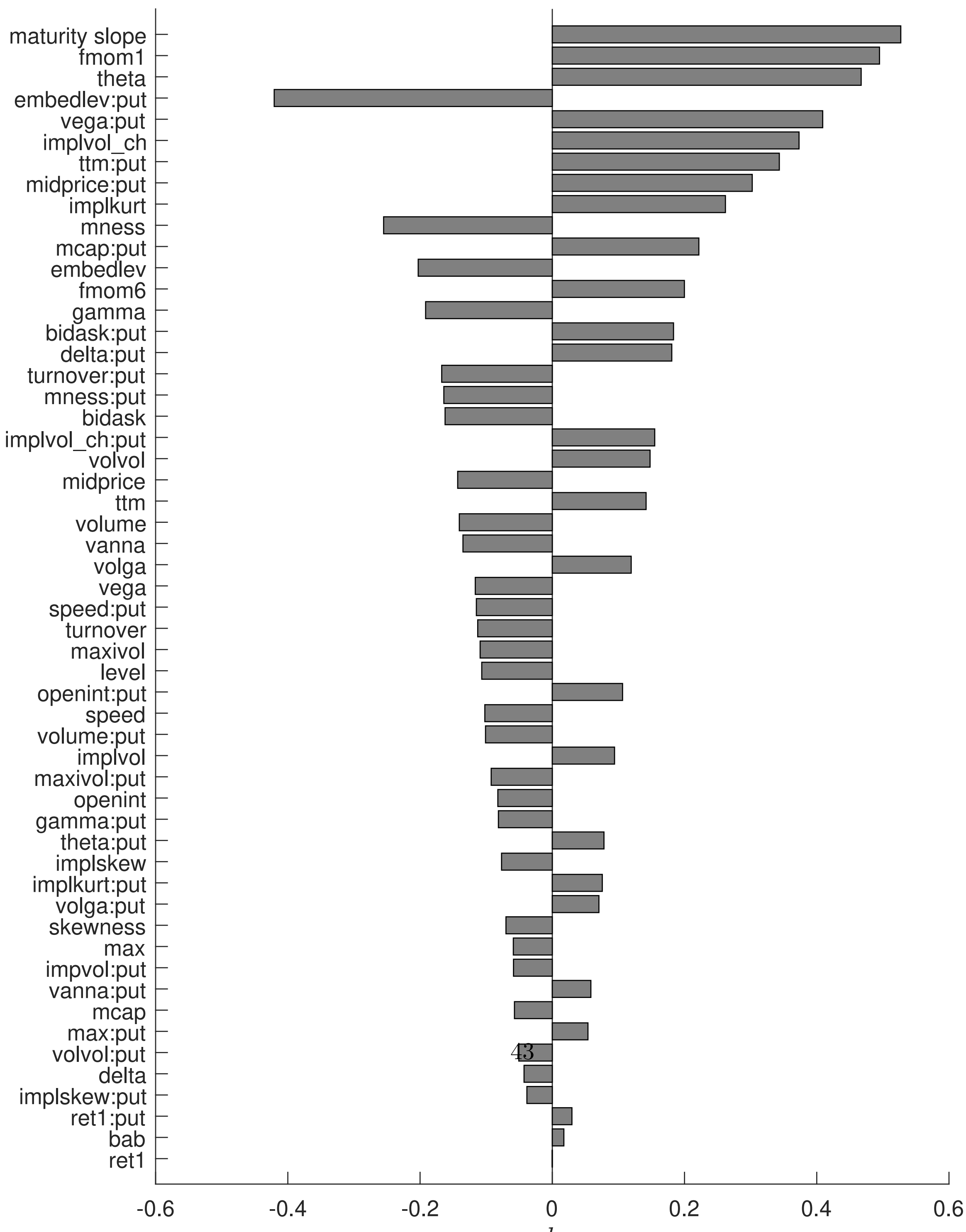


Figure 6: Cumulative Abnormal Return

The figure shows cumulative abnormal returns of the non-sparse SDF-MVP over the out-of-sample period from January 2006 to December 2022. The benchmark models are the CAPM, K3, HVX3, BCCSZ5, RPPCA3, and IPCA3 model, as well as the SDF-implied MVP obtained from the dual-penalty estimator ($L^1 - L^2$).

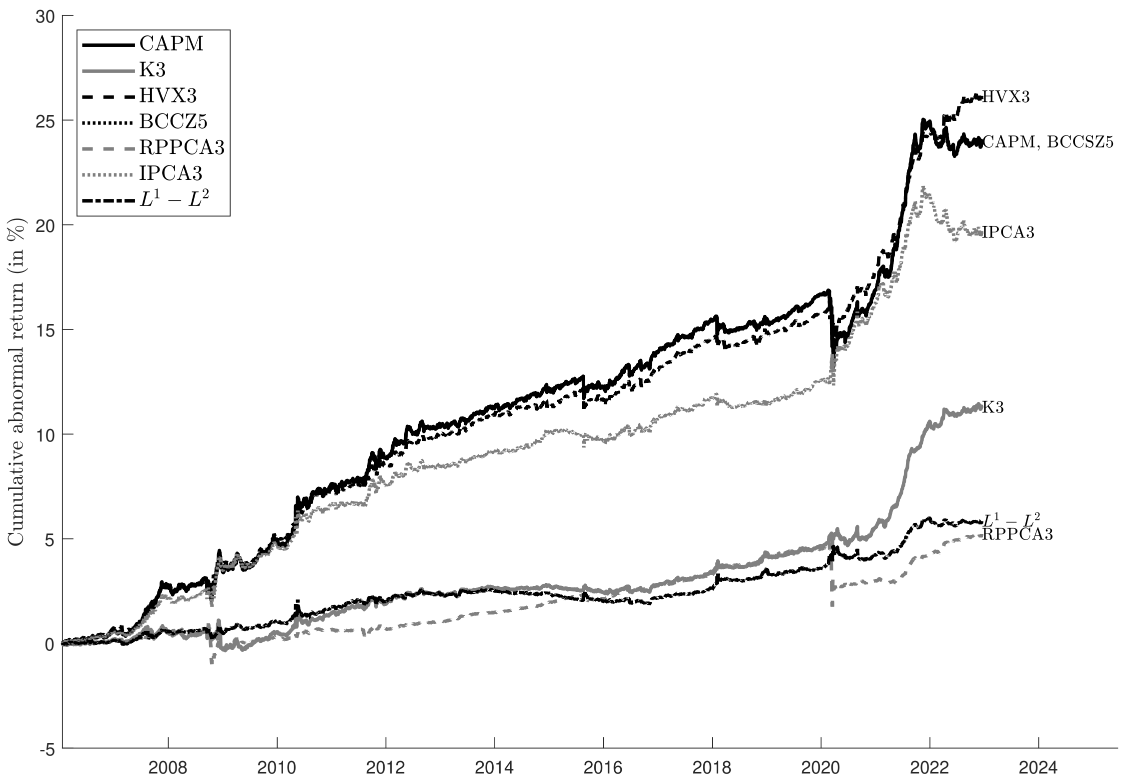
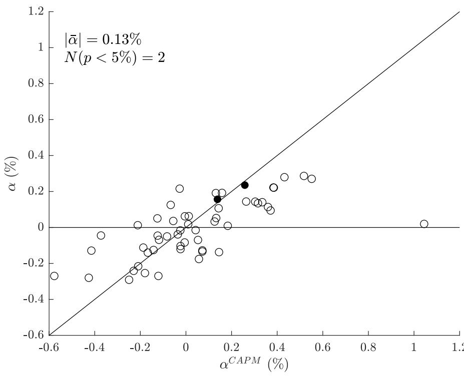


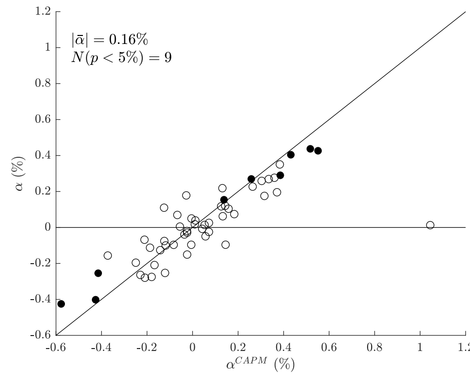
Figure 7: Out-of-Sample Factor Alphas

The figure shows the out-of-sample alphas (in %) of the 54 factors against a) the non-sparse b) the dual-penalty, c) K3, d) BCCSZ5, e) RPPCA3, and f) IPCA3 SDF-implied MVP. The MVP weights are estimated using only information available at time t, thus, the alphas are out-of-sample. Alphas that are significant at the 5% level are marked by filled circles. The t-statistics are adjusted using the Newey and West (1987) method. To control for multiple hypothesis testing, the p-values are adjusted using the Benjamini and Hochberg (1995) method. The analysis covers the study period from January 2006 to December 2022.

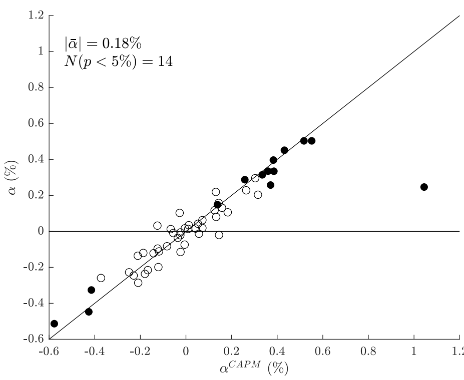
(a) L^2



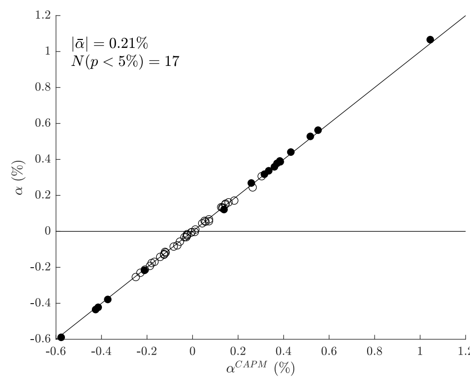
(b) $L^1 - L^2$



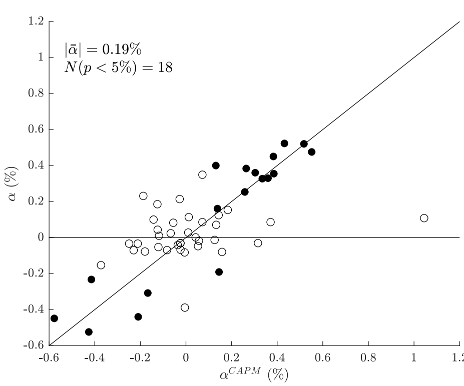
(c) K3



(d) BCCSZ5



(e) RPPCA3



(f) IPCA3

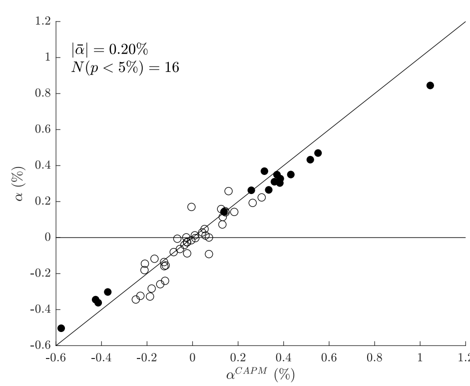


Figure 8: Out-of-Sample Importance Measure

The table reports the decrease in the annualized Sharpe ratio of the non-sparse SDF-implied MVP as a result of setting the weight pertaining to a factor to zero. All results are based on recursive model estimation over the sample period from January 2006 to December 2022.

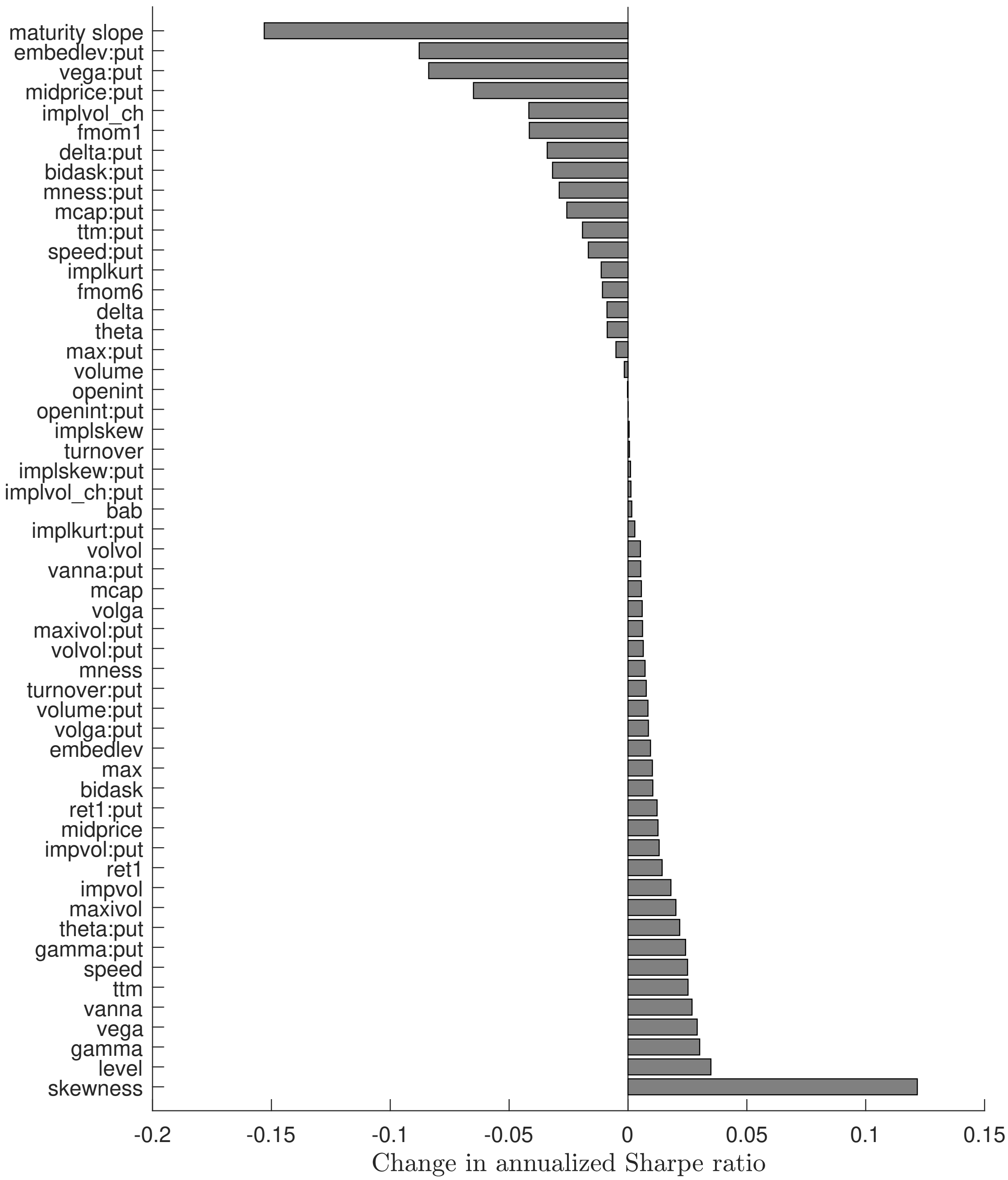


Figure 9: Out-of-Sample Alphas of Extended Factor Set

The figure shows the alphas (in %) of the 702 factors against a) the non-sparse SDF including only 54 factors and b) the non-sparse SDF including all 702 factors. The MVP weights are estimated using only information available at time t, thus the alphas are out-of-sample. Alphas that are significant at the 5% level are marked by filled circles. The t-statistics are adjusted using the Newey and West (1987) method. To control for multiple hypothesis testing, the p-values are adjusted using the Benjamini and Hochberg (1995) method. The analysis covers the study period from January 2006 to December 2022.

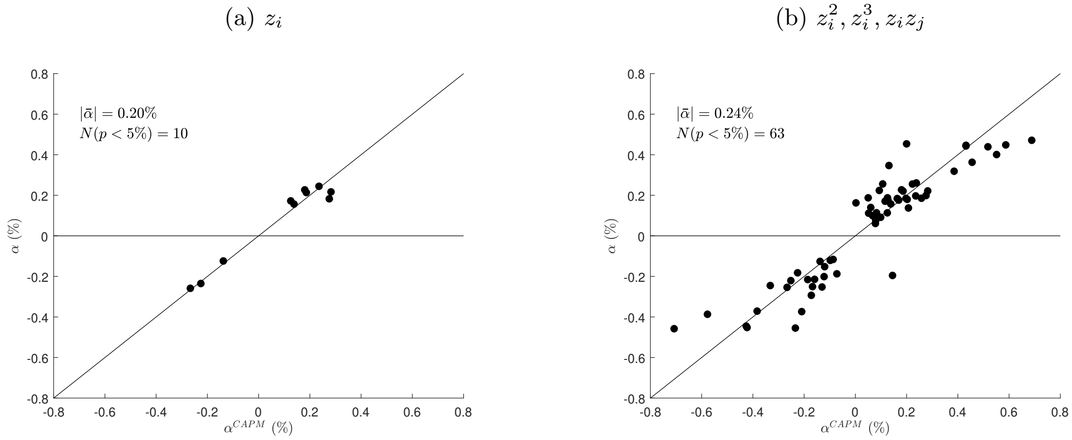


Figure 10: Out-of-Sample Factor Alphas for a Subset of Liquid Options

The figure shows the alphas (in %) of the 54 factors constructed from a subset of the 50% most liquid options at the beginning of each holding period. The MVP weights are estimated using only information available at time t, thus the alphas are out-of-sample. Alphas that are significant at the 5% level are marked by filled circles. The t-statistics are adjusted using the Newey and West (1987) method. To control for multiple hypothesis testing, the p-values are adjusted using the Benjamini and Hochberg (1995) method. The analysis covers the study period from January 2006 to December 2022.

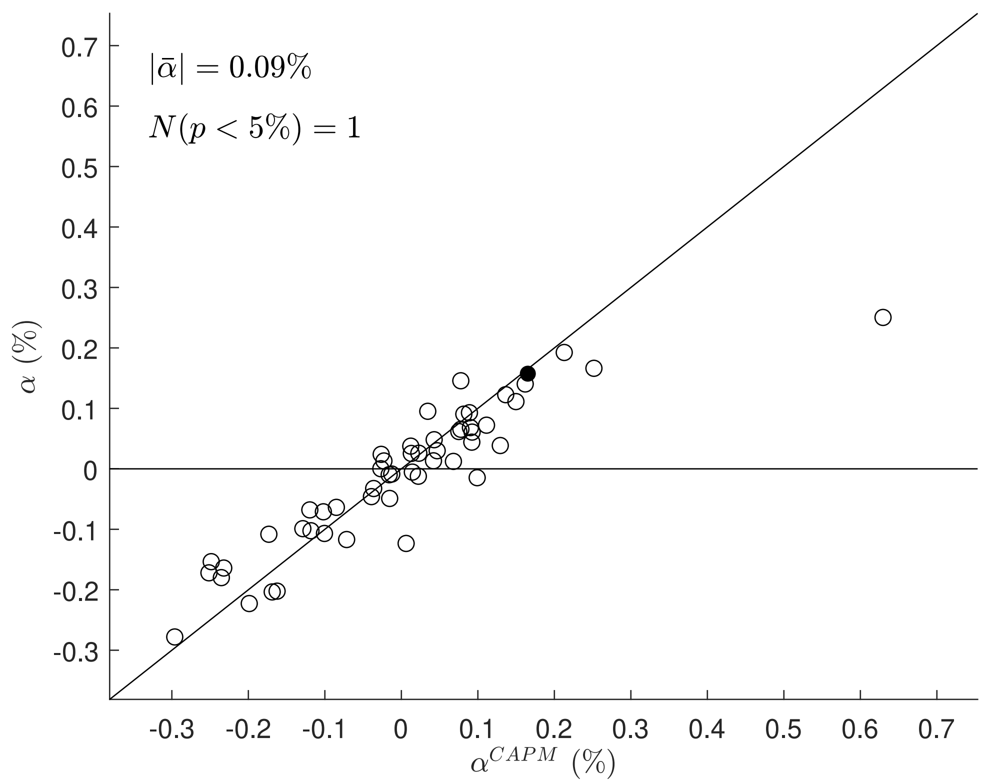


Table 1: Summary Statistics of Option Characteristics and Returns

The table reports summary statistics of option characteristics for the call (Panel A) and put (Panel B) option contracts analyzed in this study. The sample covers the period from January 1996 to December 2022. The table shows the time-to-maturity (ttm), moneyness (mness), embedded leverage (embedlev), BMS implied volatility (implvol), BMS delta, gamma, vega, and theta, as well as the delta-hedged option return as defined in equation (12).

Panel A: Call option contracts

	ttm	mness	embedlev	implvol	delta	gamma	vega	theta	$r^{\Delta}_{i,T}$
Mean	122.52	1.03	29.76	0.16	0.21	0.001	389.86	-117.88	-4.67
Median	91.00	0.94	24.35	0.15	0.18	0.001	312.40	-82.59	-2.56
Std	87.09	0.66	18.55	0.06	0.15	0.001	310.31	109.66	2.34
No. obs	71,253	71,253	71,253	71,253	71,253	71,253	71,253	71,253	71,253

Panel B: Put option contracts

	ttm	mness	embedlev	implvol	delta	gamma	vega	theta	$r^{\Delta}_{i,T}$
Mean	122.84	-1.08	15.91	0.27	-0.16	0.001	346.47	-156.44	-7.05
Median	91.00	-1.08	13.96	0.26	-0.12	0.000	263.99	-120.10	-4.13
Std	86.90	0.62	8.24	0.10	0.14	0.001	293.38	123.65	2.82
No. obs	148,868	148,868	148,868	148,868	148,868	148,868	148,868	148,868	148,868

Table 2: Performance of the SDF-implied MVP

The table reports annualized abnormal returns (in %) and t-statistics in parentheses of the non-sparse SDF-implied MVP, estimated using L^2 regularization, against the benchmark models. The benchmark models include the CAPM (α^{CAPM}), an SDF estimated employing dual-penalty ($\alpha^{L^1-L^2}$), K3 (α^{K3}), HVX3 (α^{HVX3}), BCCSZ5 (α^{BCCSZ5}), RPPCA3 (α^{RPPCA3}), and IPCA3 (α^{IPCA3}). The t-statistics, reported in parentheses, are adjusted using the Newey and West (1987) method. The out-of-sample period is from January 2006 to December 2022.

	α^{CAPM}	α^{K3}	α^{HVX3}	α^{BCCZ5}	α^{RPPCA3}	α^{IPCA3}	$\alpha^{L^1-L^2}$
L^2	**1.40**	**0.66**	**1.53**	**1.40**	**0.30**	**1.15**	**0.34**
	(4.80)	(4.07)	(6.65)	(4.81)	(2.35)	(4.77)	(2.68)

Table 3: Frontier Expansion Test

The table presents out-of-sample annualized generalized alphas (in %) derived from a time-series regression of MVP^A on $MVP^{A,B}$, with MVP^A representing the "base" portfolio and $MVP^{A,B}$ constituting a mean-variance efficient portfolio that invests in both MVP^A and MVP^B; that is, the optimal portfolio weights are obtained by estimating $\hat{\boldsymbol{b}}^{A,B} = \left(\boldsymbol{\Sigma}^{A,B}\right)^{-1} \boldsymbol{\mu}^{A,B}$. The rows show the MVP^A and the columns show the MVP^B, i.e., the row MVP is not mean-variance efficient if the alpha is statistically significant. The t-statistics, reported in parentheses, are adjusted using the Newey and West (1987) method. Any alpha statistically significant at the 5% level or higher is highlighted in bold. The out-of-sample period is from January 2006 to December 2022.

		MVP^B						
		L^2	L^1-L^2	K3	HVX3	BCCSZ5	RPPCA3	IPCA3
MVP^A	L^2		-0.10	-0.09	0.16	-0.17	0.25	0.05
			(-0.73)	(-0.53)	(1.95)	(-0.69)	(1.58)	(0.14)
	L^1-L^2	0.12		0.13	0.02	-0.16	**0.48**	0.16
		(1.26)		(1.15)	(0.18)	(-0.62)	(2.60)	(0.38)
	K3	**0.51**	**0.58**		-0.08	0.06	**0.57**	**1.35**
		(3.23)	(3.50)		(-1.36)	(0.23)	(3.13)	(2.36)
	HVX3	**1.71**	**1.91**	**1.50**		0.26	**1.62**	**3.45**
		(6.65)	(5.47)	(4.03)		(0.61)	(4.79)	(4.02)
	BCCSZ5	**1.17**	**1.37**	**1.24**	-0.22		**1.21**	**2.69**
		(3.58)	(3.37)	(3.01)	(-0.68)		(4.15)	(2.82)
	RPPCA3	**0.33**	**0.37**	0.09	0.07	-0.03		**0.70**
		(2.55)	(2.54)	(0.63)	(0.79)	(-0.13)		(2.19)
	IPCA3	**0.65**	0.68	**1.23**	0.75	0.66	**1.12**	
		(1.98)	(1.62)	(2.73)	(1.58)	(1.54)	(3.91)	

Table 4: Frontier Expansion Test using Extended Set of Factors

The table presents out-of-sample annualized generalized alphas (in %) derived from a time-series regression of MVP^A on $MVP^{A,B}$, with MVP^A representing the "base" portfolio and $MVP^{A,B}$ constituting a mean-variance efficient portfolio that invests in both MVP^A and MVP^B. The portfolios MVP^A and MVP^B are SDF-implied MVPs estimated using L^2 regularization based on a linear factor set (z_i) as well as a linear factor set, extended by quadratic (z_i^2) and cubic (z_i^3) characteristic transformations and first-order interactions $z_i z_j$ of the characteristics. The rows show the MVP^A and the columns show the MVP^B. The MVP based on the factor set shown in the rows is not mean-variance efficient if the alpha is statistically significant. The t-statistics, reported in parentheses, are adjusted using the Newey and West (1987) method. The out-of-sample period is from January 2006 to December 2022.

		MVP^B					
		z_i	z_i^2	z_i^3	z_i^2, z_i^3	$z_i z_j$	$z_i^2, z_i^3, z_i z_j$
MVP^A	z_i		0.11	-0.10	-0.22	-0.04	0.44
			(0.50)	(-0.70)	(-1.28)	(-1.37)	(1.28)
	z_i^2	-0.08		-0.02	0.03	-0.08	0.27
		(-0.98)		(-0.21)	(0.29)	(-0.98)	(1.40)
	z_i^3	0.02	0.25		0.27	0.02	0.34
		(0.09)	(1.06)		(0.92)	(0.09)	(1.04)
	z_i^2, z_i^3	-0.17	0.19	0.15		-0.17	0.20
		(-0.71)	(0.95)	(0.64)		(-0.71)	(0.75)
	$z_i z_j$	-0.04	0.11	-0.10	-0.22		0.44
		(-1.37)	(0.50)	(-0.70)	(-1.28)		(1.28)
	$z_i^2, z_i^3, z_i z_j$	-0.00	0.13	-0.15	-0.15	-0.00	
		(-0.01)	(1.19)	(-1.45)	(-1.55)	(-0.01)	

Table 5: SDF-implied MVP Performance Depending on Market States

The table reports annualized alphas (in %) of the SDF-implied MVP against the benchmark portfolios in different market states. The sample is divided into subperiods marked by (a) price jumps and non price jumps, (b) volatility and non volatility jumps, (c) bear and bull markets, (d) a low, medium, and high VIX, (e) low and high uncertainty, and (f) recessions and non recessions. Price (volatility) jump periods are defined as holding periods in which the return of the S&P500 (VIX) is below -4% (above 4%). Markets are classifies as either "bear" or "bull" markets if the 12-month trailing return is below or above the historical median 12-month trailing return, respectively. The classification into VIX regimes is based on the historical tercile distribution. The market is characterized by low (high) uncertainty if the Bekaert et al. (2022) uncertainty index at the beginning of the holding period is below (above) the historical median. Lastly, "recession" is defined using the NBER recession dates. The t-statistics, reported in parentheses, are adjusted using the Newey and West (1987) method. All results are based on out-of-sample estimates over the period from January 2006 to December 2022.

	# obs	α^{CAPM}	α^{K3}	α^{HVX3}	α^{BCCSZ5}	α^{RPPCA3}	α^{IPCA3}
Price jump	581	1.57	0.37	1.24	1.56	**0.73**	1.72
		(1.90)	(0.54)	(1.75)	(1.63)	(2.24)	(1.90)
Non price jump	3654	**1.39**	**0.71**	**1.56**	**1.38**	0.27	**1.08**
		(4.17)	(4.68)	(6.52)	(4.61)	(1.71)	(4.17)
Vol jump	1502	**1.29**	**0.55**	**1.55**	**1.29**	0.16	**1.37**
		(2.96)	(2.21)	(3.42)	(2.20)	(0.47)	(2.96)
Non vol jump	2733	**1.48**	**0.77**	**1.53**	**1.49**	**0.33**	**0.75**
		(3.00)	(3.43)	(6.00)	(4.76)	(3.37)	(3.00)
Bear market	2126	**1.01**	0.31	**1.24**	**1.03**	0.22	**0.76**
		(2.35)	(1.32)	(4.38)	(2.90)	(1.17)	(2.35)
Bull market	2109	**1.82**	**1.04**	**1.88**	**1.87**	0.43	**1.55**
		(4.50)	(4.78)	(5.64)	(4.46)	(1.80)	(4.50)
Low VIX	1280	**1.50**	0.41	**1.40**	**1.62**	0.06	**1.13**
		(2.69)	(0.99)	(4.06)	(3.29)	(0.23)	(2.69)
Medium VIX	1688	**1.02**	0.13	**0.94**	**1.06**	0.17	**0.63**
		(2.19)	(0.84)	(3.01)	(3.36)	(1.68)	(2.19)
High VIX	1267	**1.86**	**1.62**	**2.28**	**2.20**	0.68	**1.64**
		(3.09)	(5.13)	(4.36)	(3.55)	(1.79)	(3.09)
Low uncertainty	2135	**0.87**	**0.57**	**0.93**	**0.85**	0.29	**0.89**
		(3.78)	(4.67)	(3.51)	(2.50)	(1.27)	(3.78)
High uncertainty	2076	**1.99**	**0.77**	**2.23**	**1.99**	0.29	**1.19**
		(2.76)	(2.48)	(6.44)	(4.26)	(1.52)	(2.76)
Recession	418	-1.03	-0.38	-0.24	-0.85	-1.40	0.44
		(0.36)	(-0.47)	(-0.20)	(-0.49)	(-1.08)	(0.36)
Non recession	3836	**1.67**	**0.72**	**1.71**	**1.70**	**0.36**	**0.78**
		(3.41)	(4.73)	(8.02)	(6.60)	(5.11)	(3.41)

Table 6: Stock Market Factor Alphas

The table reports alphas (in %) and t-statistics of the non-sparse SDF-implied MVP against the stock market CAPM, the Fama and French (1993) three-factor (FF3), Fama and French (2015) five-factor (FF5), Fama and French (2018) six-factor (FFC6), the augmented q-factor (HXZ) model of Hou et al. (2015, 2021), as well as the four-factor Stambaugh and Yuan (2017) (SY), the three-factor Daniel et al. (2020a) (DHS), and the five-factor Daniel et al. (2020b) (DMRS) models. The t-statistics, reported in parentheses, are adjusted using the Newey and West (1987) method. The sample period is from January 2006 to December 2022.

	CAPM	FF3	FF5	FF6	HXZ	SY	DHS	DMRS
Full period	**1.34**	**1.34**	**1.37**	**1.37**	**1.36**	**1.37**	**1.34**	**1.40**
	(4.56)	(4.54)	(4.63)	(4.62)	(4.71)	(4.59)	(4.54)	(4.73)
Price jump	1.46	1.55	1.67	1.68	1.51	1.66	1.43	1.53
	(1.48)	(1.65)	(1.79)	(1.82)	(1.61)	(1.74)	(1.45)	(1.56)
Non price jump	**1.34**	**1.34**	**1.36**	**1.36**	**1.38**	**1.36**	**1.34**	**1.38**
	(4.43)	(4.41)	(4.46)	(4.46)	(4.67)	(4.43)	(4.43)	(4.49)
Vol jump	**1.22**	**1.20**	**1.22**	**1.24**	**1.23**	**1.24**	**1.22**	**1.27**
	(2.08)	(2.04)	(2.06)	(2.08)	(2.18)	(2.09)	(2.11)	(2.22)
Non vol jump	**1.44**	**1.44**	**1.48**	**1.48**	**1.46**	**1.48**	**1.44**	**1.51**
	(4.57)	(4.54)	(4.68)	(4.67)	(4.64)	(4.66)	(4.55)	(4.77)
Bear market	**0.93**	**0.93**	**0.99**	**0.98**	**0.92**	**0.99**	**0.91**	**1.05**
	(2.56)	(2.57)	(2.70)	(2.68)	(2.56)	(2.69)	(2.45)	(2.90)
Bull market	**1.79**	**1.78**	**1.79**	**1.80**	**1.81**	**1.80**	**1.80**	**1.79**
	(4.06)	(4.04)	(4.07)	(4.13)	(4.24)	(4.15)	(4.15)	(4.04)
Low VIX	**1.43**	**1.46**	**1.51**	**1.48**	**1.38**	**1.50**	**1.36**	**1.46**
	(2.80)	(2.88)	(2.95)	(2.90)	(2.74)	(2.92)	(2.70)	(2.95)
Medium VIX	**0.99**	**0.96**	**0.95**	**0.94**	**0.92**	**0.90**	**0.99**	**0.86**
	(2.90)	(2.79)	(2.77)	(2.79)	(2.59)	(2.59)	(2.96)	(2.41)
High VIX	**1.81**	**1.81**	**1.88**	**1.89**	**1.96**	**1.87**	**1.85**	**2.03**
	(2.74)	(2.75)	(2.86)	(2.85)	(3.10)	(2.84)	(2.82)	(3.13)
Low uncertainty	**0.83**	**0.80**	**0.79**	**0.80**	**0.85**	**0.83**	**0.84**	**0.75**
	(2.38)	(2.27)	(2.25)	(2.32)	(2.61)	(2.46)	(2.47)	(2.04)
High uncertainty	**1.91**	**1.92**	**2.02**	**2.00**	**1.91**	**2.00**	**1.91**	**2.06**
	(4.09)	(4.12)	(4.32)	(4.27)	(4.10)	(4.25)	(4.05)	(4.44)
Recession	-0.66	-0.66	-0.67	-0.65	-0.45	-0.71	-0.64	-0.27
	(-0.45)	(-0.45)	(-0.45)	(-0.45)	(-0.34)	(-0.48)	(-0.43)	(-0.19)
Non recession	**1.64**	**1.64**	**1.66**	**1.66**	**1.65**	**1.65**	**1.63**	**1.64**
	(6.19)	(6.21)	(6.31)	(6.33)	(6.29)	(6.24)	(6.17)	(6.22)

Internet Appendices for

"Shrinking the Cross Section of Index Option Returns"

Content

Appendix A shows that the Euler equation is valid for delta-hedged option returns. Appendix B provides a short description of the option-based characteristics analyzed in the study. Appendix C shows additional results from the study, including summary statistics in Appendix C.1 and estimation results using the extended factor set in Appendix C.2. In Appendix C.3, we show mean-variance frontier spanning test results using a rolling estimation window.

A Delta-Hedged Option Returns & the Euler Equation

This appendix demonstrates that the Euler equation shown in equation (2) applies directly to delta-hedged option returns as specified in equations (11) and (12), respectively. The reasoning follows directly from the foundational probabilistic concepts discussed in well-known literature on option pricing such as Musiela and Rutkowski (2004) and Shreve (2004). Given the option price F_{t_i} at time t_i, the risk neutral probability measure $\mathbb{E}^{\mathbb{Q}}$, the empirical probability measure $\mathbb{E}^{\mathbb{P}}$ and the time change $\delta t_i = t_i - t_{i-1}$, the expression for the option price is

$$\begin{aligned} F_{t_i} &= \exp\left(-r^f_{t_i}\delta t_{i+1}\right)\mathbb{E}^{\mathbb{Q}}\left[F_{t_{i+1}}\right] \\ &= \exp\left(-r^f_{t_i}\delta t_{i+1}\right)\mathbb{E}^{\mathbb{P}}\left[M_{t_{i+1}}F_{t_{i+1}}\right]. \end{aligned} \tag{IA1}$$

The first line is a consequence of the Feynman-Kac formula and a deterministic risk-free interest rate $r^f_{t_i}$, and the second line is a consequence of the Girsonav theorem via the fact that the stochastic discount factor $M_{t_{i+1}}$ is a Radon-Nikodym derivative. It follows that

$$\mathbb{E}^{\mathbb{P}}\left[M_{t_{i+1}}F_{t_{i+1}}\right] = \exp\left(r^f_{t_i}\delta t_{i+1}\right)F_{t_i}. \tag{IA2}$$

Similarly, the forward price of the underlying asset is[16]

$$\mathbb{E}^{\mathbb{Q}}\left[S_{t_{i+1}}\right] = \mathbb{E}^{\mathbb{P}}\left[M_{t_{i+1}}S_{t_{i+1}}\right] = \exp\left(r^f_{t_i}\delta t_{i+1}\right)S_{t_i}. \tag{IA3}$$

[16] Assuming no dividends, without loss of generality. However, it may be shown that this proof extends straightforwardly to the case of dividends.

Considering the bank account $B_{t_i} = \exp\left(\int_0^{t_i} r_s^f ds\right)$ with the deterministic risk-free interest rate $r_{t_i}^f$ that is fixed in time interval (t_i, t_{i+1}), we have

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[B_{t_{i+1}}\right] &= \mathbb{E}^{\mathbb{P}}\left[M_{t_{i+1}} B_{t_{i+1}}\right] \\
&= \exp\left(\int_0^{t_{i+1}} r_s^f ds\right) \\
&= \exp\left(\int_0^{t_i} r_s^f ds\right) \exp\left(\int_{t_i}^{t_{i+1}} r_s^f ds\right) \\
&= \exp\left(r_{t_i}^f \delta t_{i+1}\right) B_{t_i}.
\end{aligned} \tag{IA4}
$$

Under the assumption of an option PnL $\Pi_{t_i,t_{i+1}}$ in equation (11), the conditional expectation $\mathbb{E}_{t_i}$ given the filtration $\mathscr{F}_{t_i}$ up to time t_i, and that $\delta t_i = \frac{a_{t_i,t_{i+1}}}{365}$, it follows that that the the conditional expected delta-hedged PnL between time points t_i and t_{i+1} is

$$
\begin{aligned}
\mathbb{E}_{t_i}^{\mathbb{P}}\left[M_{t_{i+1}} \Pi_{t_i,t_{i+1}}\right] &= \mathbb{E}_{t_i}^{\mathbb{P}}\left[M_{t_{i+1}}\left(F_{t_{i+1}} - F_{t_i}\right)\right] - \mathbb{E}_{t_i}^{\mathbb{P}}\left[M_{t_{i+1}} \Delta_{t_i}\left(S_{t_{i+1}} - S_{t_i}\right)\right] \\
&- \mathbb{E}_{t_i}^{\mathbb{P}}\left[M_{t_{i+1}} r_{t_i}^f \delta_{t_{i+1}}\left(F_{t_i} - \Delta_{t_i} S_{t_i}\right)\right] \\
&= \mathbb{E}_{t_i}^{\mathbb{Q}}\left[\left(F_{t_{i+1}} - F_{t_i}\right)\right] - \mathbb{E}_{t_i}^{\mathbb{Q}}\left[\Delta_{t_i}\left(S_{t_{i+1}} - S_{t_i}\right)\right] \\
&- \mathbb{E}_{t_i}^{\mathbb{Q}}\left[r_{t_i}^f \delta_{t_{i+1}}\left(F_{t_i} - \Delta_{t_i} S_{t_i}\right)\right] \\
&= \left(\exp\left(r_{t_i}^f \delta_{t_{i+1}}\right) F_{t_i} - F_{t_i}\right) - \left(\exp\left(r_{t_i}^f \delta_{t_{i+1}}\right) S_{t_i} - S_{t_i}\right) \\
&- r_{t_i}^f \delta_{t_{i+1}}\left(F_{t_i} - \Delta_{t_i} S_{t_i}\right) \\
&= \left(\exp\left(r_{t_i}^f \delta_{t_{i+1}}\right) - 1 - \delta_{t_{i+1}} r_{t_i}^f\right)\left(F_{t_i} - \Delta_{t_i} S_{t_i}\right).
\end{aligned} \tag{IA5}
$$

After applying the Taylor series expansion of $\exp\left(r_{t_i}^f \delta_{t_i}\right) = 1 + r_{t_i}\delta_{t_i} + \mathcal{O}^2\left(r_{t_i}^f \delta_{t_i}\right) \approx 1 + r_{t_i}\delta_{t_i}$, it follows that

$$
\begin{aligned}
\mathbb{E}_{t_i}^{\mathbb{P}}\left[M_{t_{i+1}} \Pi_{t_i,t_{i+1}}\right] &= \left(\exp\left(r_{t_i}^f \delta_{t_{i+1}}\right) - 1 - \delta_{t_{i+1}} r_{t_i}^f\right)\left(F_{t_i} - \Delta_{t_i} S_{t_i}\right) \\
&= \left(1 + \delta_{t_{i+1}} r_{t_i}^f - 1 - \delta_{t_{i+1}} r_{t_i}^f\right)\left(F_{t_i} - \Delta_{t_i} S_{t_i}\right) = 0.
\end{aligned} \tag{IA6}
$$

The Euler equation thus holds for the delta-hedged option PnL between time periods t_i and t_{i+1}, and by extension for the option return. From the law of iterated expectation, the argument extends to the option return for the entire holding period.

B Variable Definitions

Table 1: Variable Description

The table provides definitions of the characteristics used to create characteristic-based option factors studied in this paper. "Symbol" in the first column refers to the abbreviations used in the paper.

Symbol	Characteristic	Description
bidask	Bid-ask spread	Following Bali et al. (2023), the bid-ask spread is defined as $\frac{2\times(bid-ask)}{bid+ask}$.
delta	Option delta	Following Black and Scholes (1973), the sensitivity of the option with respect to a linear change in the underlying asset.
embedlev	Embedded leverage	Following Karakaya (2014) and Frazzini and Pedersen (2022), the embedded leverage of an option is defined as the ratio of the underlying's spot price to the option's mid price, multiplied by the absolute option delta.
gamma	Option gamma	Following Black and Scholes (1973), the sensitivity of the option with respect to a quadratic change in the underlying. It is also the sensitivity of *delta* with respect to linear changes in the underlying.
implkurt	Implied kurtosis	Fourth moment of the probability distribution of the returns of the underlying asset.
implskew	Implied skewness	Third moment of the probability distribution of the returns of the underlying asset.
implvol	Implied volatility	Black and Scholes (1973) implied volatility.
implvol_ch	Change in *implvol*	The percentage change in Black and Scholes (1973) implied volatility.

Table 1: Variable Definitions (continued)

Symbol	Characteristic	Description
max	Maximum option return	Following Liu et al. (2022), we measure the option lottery-like characteristics by using the highest daily option return in the most recent month.
maxivol	Maximum implied volatility	Following Liu et al. (2022), we measure the option lottery-like characteristics by using the highest daily implied volatility in the most recent month.
mcap	Market capitalization	Following Büchner and Kelly (2022) and Frazzini and Pedersen (2022), the market capitalization is the product of open interest and the option's mid price.
midprice	Option price	Following Bali et al. (2022), the mid price of the option.
mness	Moneyness	Following Bali et al. (2023), moneyness is the ratio of the strike price to the spot price of the underlying.
ret1	Past-month return	Following Jegadeesh (1990) and Heston et al. (2022), we include the return of the previous holding period.
openint	Open interest	Following Bali et al. (2023), the open interest of the option.
speed	Option speed	Following Black and Scholes (1973), the sensitivity of the option with respect to a cubic change in the underlying. It is also the sensitivity of *gamma* with respect to a linear change in the underlying and the sensitivity of delta with respect to a quadratic change in the underlying.
theta	Option theta	Following Black and Scholes (1973), the time-decay of the option value.

Table 1: Variable Definitions (continued)

Symbol	Characteristic	Description
ttm	Time-to-maturity	Following Büchner and Kelly (2022), time-to-maturity is the number of days to maturity.
turnover	Option turnover	Following Bali et al. (2023), turnover is the ratio of the option's volume to its open interest.
vanna	Option vanna	Following Black and Scholes (1973), the sensitivity of *delta* with respect to a change in the implied volatility.
vega	Option vega	Following Black and Scholes (1973), the sensitivity of the option with respect to a linear change in the implied volatility.
volga	Option volga	Following Black and Scholes (1973), the sensitivity of the option with respect to a quadratic change in the implied volatility. It is also the sensitivity of *vega* with respect to a linear change in the implied volatility.
volume	Option volume	Following Bali et al. (2023), the daily trading volume of an option.
volvol	Volatility of *implvol*	Following Ruan (2020) and Horenstein et al. (2022), the standard deviation of an option's implied volatility in the past month.

C Additional Results

C.1 Factor Summary Statistics

Table 1: Average Factor Returns

The table reports the average annualized raw and CAPM-adjusted factor returns (in %), as well as t-statistics and p-values (in %). The t-statistics are adjusted using the Newey and West (1987) method. To control for multiple hypothesis testing, p-values are adjusted using the Benjamini and Hochberg (1995) method. Average returns and CAPM alphas are reported in bold if they are statistically significant at the 5% level, after adjusting the p-values to control for multiple hypothesis testing. The study period is from January 1996 to December 2022.

	Raw			CAPM-adjusted		
	Avg (%)	t-stat	p-value	Avg (%)	t-stat	p-value
bab	-0.03	-1.09	39.05	0.05	2.25	5.82
bidask	**-0.54**	-4.50	0.00	-0.14	-1.09	41.25
bidask:put	-0.05	-0.35	78.53	**0.34**	3.09	0.90
delta	**0.62**	5.08	0.00	**0.34**	3.27	0.52
delta:put	-0.09	-0.65	58.93	**0.32**	2.72	1.95
embedlev	0.10	0.85	47.31	-0.20	-2.06	8.32
embedlev:put	-0.21	-1.72	14.75	**-0.54**	-5.20	0.00
fmom1	**0.25**	4.82	0.00	**0.27**	4.30	0.02
fmom6	**0.16**	3.16	0.43	**0.15**	2.64	2.22
gamma	**-0.51**	-3.91	0.04	-0.24	-1.14	40.37
gamma:put	**-0.35**	-3.08	0.54	-0.01	-0.11	93.22
implkurt	**0.09**	3.18	0.43	**0.11**	3.50	0.26
implkurt:put	0.03	0.80	49.56	0.03	0.87	47.98
implskew	-0.04	-2.02	8.30	-0.04	-1.96	9.36
implskew:put	-0.03	-1.16	37.77	-0.03	-1.07	41.56

Table 1: Average Factor Returns (continued)

	Raw			CAPM-adjusted		
	Avg (%)	t-stat	p-value	Avg (%)	t-stat	p-value
implvol	-0.02	-0.19	86.84	0.07	0.45	71.89
implvol_ch	**0.35**	3.70	0.07	0.24	1.99	8.93
implvol_ch:put	0.12	1.41	26.03	0.13	1.15	40.37
impvol:put	-0.11	-1.04	40.39	0.19	2.00	8.93
level	**1.40**	3.83	0.04	0.04	0.14	92.55
maturity slope	0.26	1.06	40.04	**0.79**	3.51	0.26
max	**-0.45**	-3.85	0.04	-0.13	-0.85	48.46
max:put	-0.10	-0.89	47.13	0.23	2.05	8.32
maxIVOL	-0.10	-1.22	35.59	-0.15	-1.30	33.50
maxIVOL:put	-0.10	-1.44	25.47	**-0.18**	-2.48	3.27
mcap	**-0.30**	-3.95	0.04	-0.05	-0.61	63.33
mcap:put	0.04	0.50	69.56	**0.27**	3.55	0.26
midprice	**-0.53**	-4.15	0.02	-0.11	-0.75	54.33
midprice:put	0.04	0.29	82.12	**0.44**	3.81	0.11
mness	-0.03	-0.21	86.23	-0.15	-0.93	47.55
mness:put	0.11	0.85	47.31	**-0.30**	-2.51	3.11
openint	-0.08	-2.08	7.75	-0.04	-0.90	47.62
openint:put	0.04	0.97	42.88	0.06	1.20	38.96
ret1	0.15	2.01	8.30	0.09	1.12	40.66
ret1:put	0.08	1.00	41.81	0.09	0.96	46.87
skewness	0.65	2.01	8.30	0.10	0.31	79.91
speed	**0.49**	3.83	0.04	0.18	1.75	14.28

Table 1: Average Factor Returns (continued)

	Raw			CAPM-adjusted		
	Avg (%)	t-stat	p-value	Avg (%)	t-stat	p-value
speed:put	-0.19	-2.09	7.75	**-0.36**	-4.88	0.00
theta	**0.69**	6.09	0.00	**0.34**	2.89	1.39
theta:put	**0.32**	2.66	1.81	-0.05	-0.43	71.89
ttm	**0.24**	3.29	0.30	**0.29**	3.92	0.08
ttm:put	**0.38**	5.03	0.00	**0.41**	5.40	0.00
turnover	**-0.24**	-6.05	0.00	**-0.18**	-2.86	1.44
turnover:put	**-0.21**	-5.06	0.00	**-0.15**	-3.00	1.12
vanna	**0.32**	2.65	1.81	0.11	0.90	47.62
vanna:put	-0.03	-0.43	73.72	-0.04	-0.48	70.78
vega	**-0.37**	-2.99	0.70	-0.01	-0.09	93.22
vega:put	0.15	1.14	38.11	**0.52**	4.65	0.00
volga	**0.49**	5.72	0.00	**0.20**	2.65	2.22
volga:put	0.18	1.94	9.43	-0.08	-1.00	45.07
volume	**-0.25**	-6.37	0.00	**-0.18**	-2.97	1.17
volume:put	**-0.17**	-4.17	0.02	-0.10	-2.13	7.56
volvol	-0.01	-0.13	90.05	-0.05	-0.57	65.62
volvol:put	-0.11	-1.12	38.11	**-0.27**	-2.80	1.64

C.2 Extended Factor Set

Figure 1: R^2_{cv}s from Dual-Penalty Specification for Extended Factor Set

The figure presents cross-validation R^2_{cv} values for models employing varying degrees of L^1 and L^2 regularization, applied to an extended factor set of 702 factors. The x-axis represents the degree of L^2 regularization, while the y-axis portrays the degree of L^1 regularization. The level of shrinkage (L^2) is expressed as the root expected squared Sharpe ratio κ and the degree of sparsity (L^1) is represented by the number of factors that have a non-zero coefficient in the SDF. Elevated R^2_{cv} values are denoted by warmer colors. The study covers the period from January 1996 to December 2022.

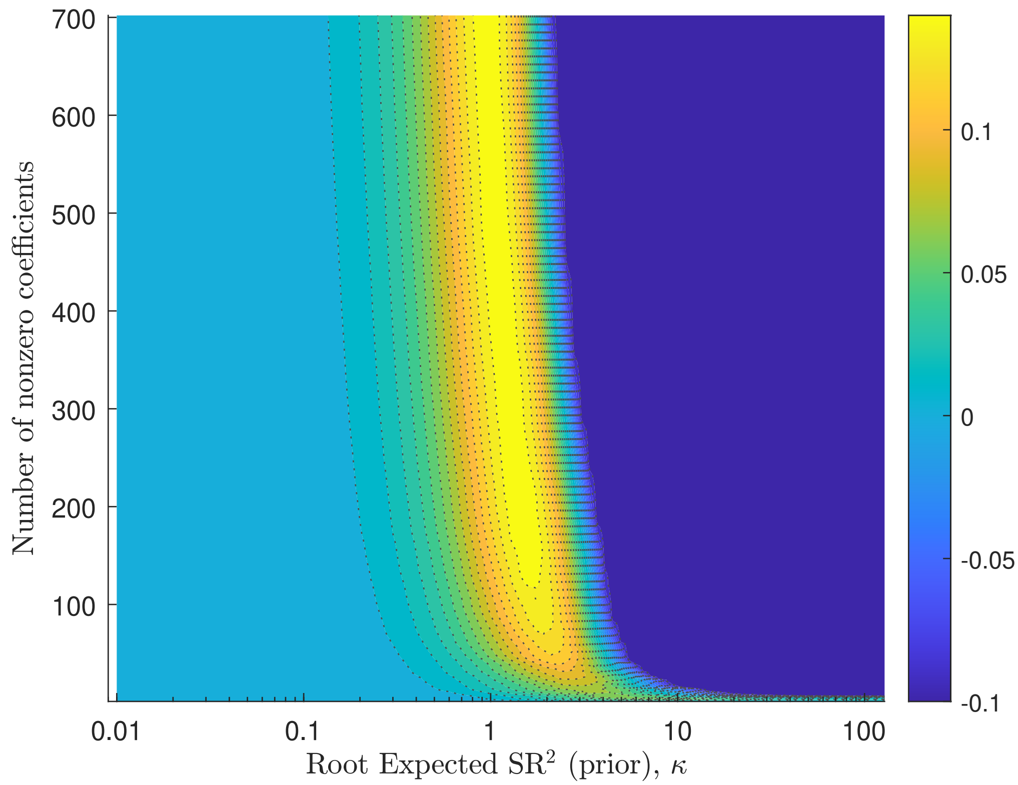
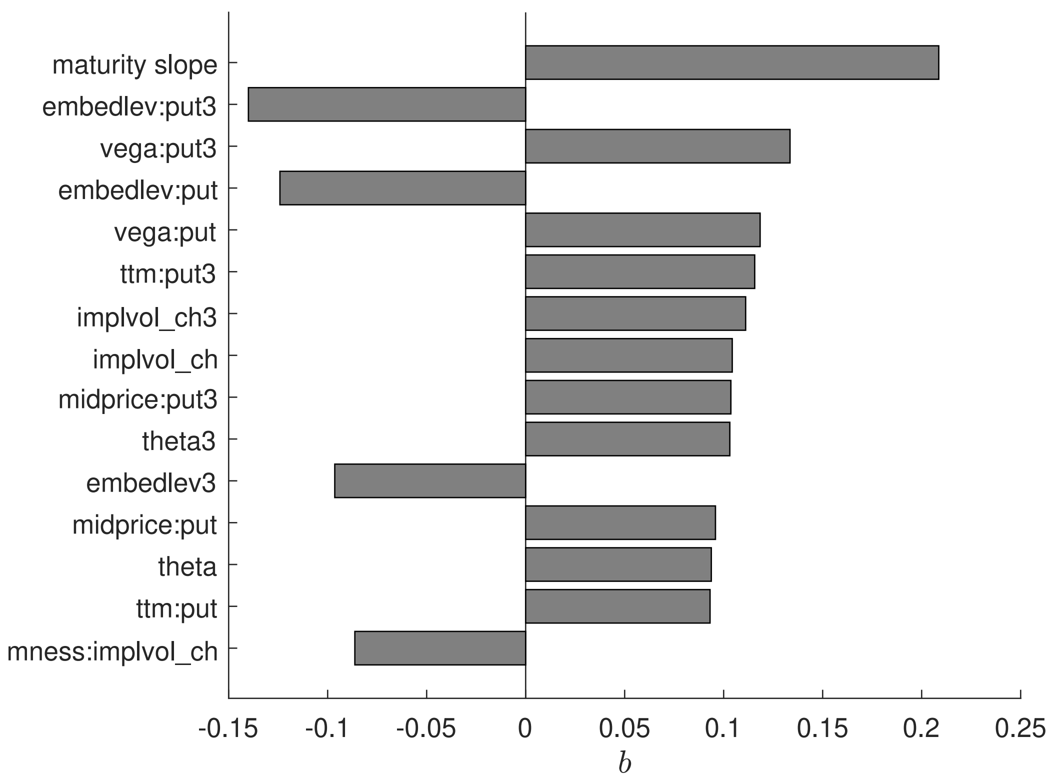


Figure 2: SDF Coefficients in Extended Factor Set

The figure shows the coefficients of the fifteen factors that have the largest (in absolute terms) weight in the SDF using the L^2-only penalty estimator, applied to the extended set of 702 factors. Second (third) order polynomials of the characteristics are denoted by *charname*2 (*charname*3). The coefficients are sorted in descending order based on their absolute SDF coefficient. The sample period is from January 1996 to December 2022.



C.3 Variation in Training Window

In the main analyses, we estimated the SDF coefficients using an expanding time window. However, how do the results change if we assume a fixed rolling window? Table 2 reports the generalized alphas of this exercise. The findings are qualitatively similar to those reported in Table 3; that is, a combination of the MVP implied by the non-sparse SDF and any other benchmark MVP does not statistically significantly outperform the MVP implied by the non-sparse SDF, indicating mean-variance efficiency. In contrast, the other MVPs do not exhibit mean-variance efficiency, which supports the findings from the main analyses.

Table 2: Frontier Expansion Test Using Rolling Estimation Window

The table presents annualized generalized alphas (in %) derived from a time-series regression of MVP^A on $MVP^{A,B}$, with MVP^A representing the "base" portfolio and $MVP^{A,B}$ constituting a mean-variance efficient portfolio that invests in both MVP^A and MVP^B. The weights in the MVP are estimated using a rolling estimation window with a fixed length of ten years. The t-statistics, reported in parentheses, are adjusted using the Newey and West (1987) method. Any alpha statistically significant at the 5% level or higher is highlighted in bold. The out-of-sample period is from January 2006 to December 2022.

MVP^A	MVP^B: L^2	L^1-L^2	K3	HVX3	BCCSZ5	RPPCA3	IPCA3
L^2		-0.06	-0.00	-0.20	-0.92	-0.94	-0.42
		(-0.13)	(-0.01)	(-0.59)	(-1.80)	(-1.47)	(-0.55)
L^1-L^2	-0.13		-0.34	-0.51	-0.67	-0.86	-0.21
	(-0.29)		(-1.44)	(-1.55)	(-1.50)	(-1.39)	(-0.36)
K3	**2.14**	**2.03**		-0.22	0.49	**1.49**	**2.18**
	(5.76)	(5.56)		(-0.49)	(0.93)	(4.62)	(3.08)
HVX3	**2.60**	**2.69**	0.67		0.32	**1.84**	**1.96**
	(7.50)	(7.08)	(1.52)		(0.50)	(6.27)	(2.04)
BCCSZ5	**1.54**	**1.96**	0.61	-0.44		**1.35**	1.33
	(2.91)	(3.73)	(1.03)	(-1.24)		(4.67)	(1.54)
RPPCA3	**2.14**	**2.22**	0.14	0.21	0.25		1.20
	(4.34)	(4.55)	(0.62)	(1.61)	(0.56)		(1.58)
IPCA3	0.73	1.08	0.74	-0.92	-0.70	0.70	
	(0.87)	(1.71)	(1.50)	(-1.70)	(-0.99)	(1.82)	

References

Bali, T.G., Beckmeyer, H., Mörke, M., Weigert, F., 2023. Option return predictability with machine learning and big data. Review of Financial Studies 36, 3548–3602.

Bali, T.G., Cao, J., Chabi-Yo, F., Song, L., Zhan, X., 2022. A factor model for stock options. Working Paper.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological) 57, 289–300.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Büchner, M., Kelly, B.T., 2022. A factor model for option returns. Journal of Financial Economics 143, 1140–1161.

Frazzini, A., Pedersen, L.H., 2022. Embedded leverage. Review of Asset Pricing Studies 12, 1–52.

Heston, S.L., Jones, C.S., Khorram, M., Li, S., Mo, H., 2022. Option momentum. Journal of Finance, Forthcoming.

Horenstein, A.R., Vasquez, A., Xiao, X., 2022. Common factors in equity option returns. Working Paper.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881.

Karakaya, M.M., 2014. Characteristics and expected returns in individual equity options. Working Paper.

Liu, C., Wang, T., Wang, Y., Xiang, H., 2022. Variation in put-call parity deviation and equity option returns. Working Paper .

Musiela, M., Rutkowski, M., 2004. Martingale Methods in Financial Modelling (Second Edition). Springer.

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

Ruan, X., 2020. Volatility-of-volatility and the cross-section of option returns. Journal of Financial Markets 48, 100492.

Shreve, S., 2004. Stochastic Calculus for Finance II: Continuous-Time Models. Springer.