Option Mispricing and Alpha Portfolios

Andras Fulop* Junye Li† Mo Wang‡

This version: September 2023

Abstract

Relying on a latent factor model with time-varying dependence of systematic risk and mispricing on firm and option characteristics, we reveal economically substantial mispricing in the options market. The portfolio based on individual options alphas related to characteristics earns an out-of-sample annualized Sharpe ratio of 1.62 in call option returns and of 1.86 in put option returns. Commonly used risk factors in the stock and options markets are incapable of explaining abnormal returns of option alpha portfolios. We show that characteristics related to risk-neutral moments and liquidity and their interactions largely contribute to option mispricing and that many characteristics that contribute to systematic risk also contribute to mispricing at times.

Keywords: Option Return Predictability, Mispricing, Factor Model, Alpha Portfolio, Projected PCA

JEL Classification: *G11, G12, G13*

*ESSEC Business School, fulop@essec.edu
†Fudan University, li_junye@fudan.edu.cn
‡ESSEC Business School, mo.wang@essec.edu

1. Introduction

Options are not redundant assets. Since Black and Scholes (1973), a majority of the literature in option pricing has focused on an approach that first specifies the full dynamics of the underlying price and its instantaneous variance and then employs risk-neutral valuation. However, such a parametric approach may suffer from model misspecification and is incapable to explain empirically observed variation in option returns (Israelov and Kelly, 2017). Investors may not hold options up to maturity and may be more interested in risk factors that drive variation of option returns and in quantifying any mispricing if there exists in the option holding period. However, this question is far less understood in the options market than in the corresponding underlying equity market.

In this paper, we follow the factor pricing approach and answer an important question: whether options are mispriced under very general assumptions on systematic risk, and if so, how we can quantify such mispricing. Different from equities, the factor pricing approach is not straightforward to apply in the options market for at least three important reasons. First, there do not exist well-established factors that can explain time series comovements of options returns and average return spreads across options on individual equities. Second, the unique option contractual features and relative underdevelopment of the market make option return mispricing and exposures to systematic risk time-varying and dependent on both underlying and option characteristics. Third, the dependence structure of systematic risk and mispricing on characteristics may well change over time as well, due to, for example, structural changes of the options market and different types of arbitrage opportunities appearing and disappearing at different times.

Therefore, in this paper, we adopt the latent factor pricing approach and further make alpha and beta depending on firm and option characteristics. To accommodate the time-varying dependence structure, we resort to the projected principal component analysis (PPCA) originally proposed by Fan, Liao, and Wang (2016) and then extended by Kim, Korajczyk, and Neuhierl (2021). Unlike the standard statistical factor model, which requires a large time series sample size for consistent estimation, the PPCA estimator of the factor loading function converges to the true one as the dimension of the cross section increases, even for short time series. This property allows us for estimating the model using a short rolling window over time and hence introducing time-varying dependence structure. We assume that characteristics are relatively stable over a short time interval and estimate alphas of individual option returns that relate to characteristics. Based on those alpha estimates, we construct option alpha portfolios that are completely out-of-sample. This approach gives characteristics maximal explanatory power for systematic risk before estimating alphas and therefore, the alpha portfolios constructed in this way can be regarded as realizations of mispricing in the options market. Two recent papers by Buchner and Kelly (2022) and Goyal and Saretto (2022) follow similar latent factor pricing approach and apply the instrumental principal component analysis (IPCA) of Kelly, Pruitt, and Su (2019) to examining the factor structure in S&P 500 index option returns and individual equity option returns, respectively. However, a key difference is that both assume that the dependence structures of systematic risk and mispricing on characteristics are constant over time.

We implement a conservative investigation of option mispricing by focusing on short-term at-the-money options contingent on individual stocks. Those options are the most liquid and hence are less mispriced. If there exists any mispricing in these options, it is likely to be even more serious in less liquid segments of the options market. Using both volatility surface data and single equity options data provided by *OptionMetrics*, we construct a large cross-section of short-maturity at-the-money call and put options for the period ranging from March 1996 to December 2021. Option returns in our sample are delta-hedged, removing the part that can be explained by local variation of the underlying stock price. We further construct a large panel of firm and option characteristics (in total, 113 characteristics), on which option betas and alphas may potentially depend.

There is a tradeoff when we choose the size of the rolling window in the extended PPCA. A larger window size could give us more observations to estimate the functions of factor loadings and mispricing, whereas a smaller window size makes our assumption of stable characteristics in the window more acceptable. Hence, in the paper, we try different window sizes ranging from 12 to 24 months. Within each window, we use all option returns observations and the initial observations of characteristics to obtain functional structure of mispricing and then estimate alphas of individual option returns using the end-of-window characteristics. Based on those alpha estimates, we construct an alpha portfolio whose weights are proportional to individual option alphas that purely relate to characteristics and hold this portfolio for one month and compute its realized returns. We repeat this procedure over time to the end of sample. For a window size of 12 months, we find that for the number of the latent factors ranging from 1 to 10, the annualized Sharpe ratio of the call alpha portfolio ranges from 1.17 to 1.62, and the Sharpe ratio of the put alpha portfolio ranges from 1.28 to 1.86. Note that the exact number of the latent factors is unknown and needs to be estimated. Using the projected-PC eigenvalue ratio test of Fan, Liao, and Wang (2016), we find that the number of the latent factors varies from 1 to 3 over time. The smaller number of factors may be insufficient to capture all systematic risks; however, Kim, Korajczyk, and Neuhierl (2021) show that a slight increase of the number of extracted factors does not harm the model performance materially. Therefore, we select the number of the latent factors equal to 3 in most of our analyses. In fact, the Sharpe ratios of both call and put alpha portfolios reach the highest values of 1.62 and 1.86, respectively, when the number of the latent factors is set to 3.

We then ask whether the commonly used risk factors in the stock and options markets can explain the abnormal returns of option alpha portfolios. We consider multiple stock market factors, namely, the Fama-French five factors (Fama and French, 1992, 1996, 2015), the momentum factor (Carhart, 1997), the q five factors (Hou et al., 2021), and the mispricing four factors (Stambaugh and Yuan, 2017), and various options market factors, namely, the Karakaya three factors (Karakaya, 2013), the liquidity factor (Christoffersen, Goyenko, Jacobs, and Karoui, 2018), and the idiosyncratic volatility factor (Cao and Han, 2013); we also construct several latent factors from option returns using principal component analysis (PCA) and risk-premium PCA (RP-PCA) of Lettau and Pelger (2020). We find that none of those factor models can explain the abnormal returns of option alpha portfolios. The smallest risk-adjusted return is from the mispricing four-factor models, which is 1.89% ($t = 6.93$) for the call alpha portfolio and 2.34% ($t = 7.30$) for the put alpha portfolios. The risk-adjusted return from a factor model that includes all five options market factors is about 2.92% ($t = 3.95$) for the call alpha portfolio and it is about 3.03% ($t = 4.04$) for the put alpha portfolio.

The contributions of characteristics to systematic risk and mispricing vary substantially over time, in support of our argument that temporal dependence structures of systematic risk and mispricing on characteristics are time-varying. Furthermore, most characteristics are only related to mispricing for a very short while, reflecting that such mispricing may disappear very quickly after its discovery. We find that characteristics related to risk-neutral moments of the underlying stock and liquidity play more important roles in capturing both mispricing and systematic risk. Characteristics that significantly contribute to systematic risk also contribute to mispricing at times. According to the broad classification of Chen and Zimmermann (2022) and Bali, Beckmeyer, Moerke, and Weigert (2022), we find that most of the top 20 characteristics contributing to systematic risk and mispricing are in the groups of liquidity and risk. Bali, Beckmeyer, Moerke, and Weigert (2022) also find importance of liquidity- and risk-lated characteristics in predicting the cross-section of option returns.

Introducing nonlinearity captured by interaction terms of characteristics slightly improves the performance of alpha portfolios. We find that in comparison with the linear case, with the 12-month window size and three latent factors, the annualized Sharpe ratio of the call alpha portfolio increases from 1.62 to 1.69, and the Sharpe ratio of the put alpha portfolio increases from 1.86 to 1.88. We further find that when the interaction terms are introduced, the contributions of many characteristics become negligible and most of the top 20 characteristics contributing to systematic risk and mispricing are those related to interactions of characteristics. However, consistently with the previous finding, the liquidity- and risk-related characteristics still play the most important roles.

We implement various robustness checks. Our results remain when using different estimation window sizes and using the post-2004 sample. Furthermore, while our analysis in the main text is based on the volatility surface option data, we find that our results are robust to using the single equity options data.

Our paper relates and contributes to several strands of literature. First, it closely relates to the literature that searches for the factor structure in options returns. As mentioned above, Buchner and Kelly (2022) and Goyal and Saretto (2022) apply the IPCA method to searching for the factor structures of the S&P index option returns and individual equity option returns, respectively. Both find that the mispricing in option returns is relatively small. However, our paper differs in important aspects. First, we allow for time-varying dependence structures of both systematic risk and mispricing on characteristics, whereas in the IPCA procedure, the relations between characteristics and systematic risk/mispricing are constant over time. Time-varying dependence structure can better help us identify mispricing, given that a particular mispricing is usually short-lived. Second, we consider a large panel of 113 characteristics that have been shown to predict stock returns (Neuhierl, Tang, Varneskov, and Zhou, 2021) and/or to predict option returns (Bali, Beckmeyer, Moerke, and Weigert, 2022), whereas Buchner and Kelly (2022) employ only 7 option-related variables and Goyal and Saretto (2022) use 44 characteristics. Indeed, using a large set of characteristics is important in spanning expected returns and covariance and revealing the underlying risk-return relationship. Given such differences, our paper reveals economically substantial mispricing in the options market.

Early studies on the factor structure in option returns include Coval and Shumway (2001) and Jones (2006). Karakaya (2013) focuses on individual equity options and proposes three factors of level, maturity, and value. Christoffersen, Fournier, and Jacobs (2018) find a strong factor structure in individual equity options using the principal component analysis approach and show the first principal components of the equity volatility levels, skews, and term structures explain a substantial fraction of the cross-sectional variation. Working with option portfolios constructed from a certain number of firm characteristics, Horenstein, Vasquez, and Xiao (2020) find that a four-factor model can explain the cross-section and time-series of equity option returns.

Our paper also contributes to the literature on cross-sectional option return predictability. A number of papers investigate whether those firm characteristics that forecast stock returns can also predict option returns (see, e.g., Bakshi and Kappadia, 2003; Goyal and Saretto, 2009; Cao and Han, 2013; Hu and Jacobs, 2020; Vasquez, 2017; Christoffersen, Goyenko, Jacobs, and Karoui, 2018; Zhan, Han, Cao, and Tong, 2022). Using a large panel of firm and option characteristics, two recent studies of Brooks, Chance, and Shafaati (2018) and Bali, Beckmeyer, Moerke, and Weigert (2022) apply machine learning methods to investigate option return predictability in the cross-section of individual equity options. However, these papers do not clearly answer the question on the sources of option return predictability: whether predictability is from characteristics predicting alpha or predicting beta. Our paper quantifies how much of expected excess option return is from mispricing.

Another contribution of our paper is to clearly show importance of risk- and liquidity-related characteristics in predicting option returns. Those studies that use only a limited number of characteristics may suffer from the issue of omitted variables. Further, the assumption of constant temporal dependence of alpha and beta can not capture the structural changes of the data generating process, and therefore cannot capture time-series variation of contributions of characteristics. Instead, we build up the latent factor model by using information contained in a wide range of firm and option characteristics and estimate both cross-sectional and the temporal relations between characteristics and systematic risk and mispricing. We find that characteristics related to risk-neutral moments and stock and option liquidity and their interactions are more closely associated with options mispricing and factor loadings. Bali, Beckmeyer, Moerke, and Weigert (2022) apply machine learning approaches to option return predictability and find that the top three relevant groups of characteristics are those of option contractual features, liquidity, and risk.

The paper proceeds as follows. Section 2 presents the model and estimation method based on the PPCA method. Section 3 introduces the data on options and firm characteristics. Section 4 presents our main empirical results based on the volatility surface options data. Section 5 provides several robustness checks. Section 6 concludes the paper.

Extra results are presented in the Internet Appendix.

2. Model and Estimation

2.1. Delta-Hedged Option Returns

We focus on individual stock option returns that are delta-hedged. Given that the underlying stock price variation is a key factor that affects option returns, delta-hedging results in the fraction of option returns that is not explained by local-linear exposure to fluctuations of the underlying stock price. Following the literature (e.g., Bakshi and Kappadia, 2003; Cao and Han, 2013), we first calculate the delta-hedged option gain that is a daily-rebalanced zero-investment portfolio consisting of a long option (call or put) position, hedged by a short position in the underlying stock.

Let $t = t_0 < \cdots < t_N = t + \tau$ represent all trading days between time t and $t + \tau$, and in this paper, we choose τ equal to one month. The daily delta-hedged option gain over the one-month investment period is given by

$$\begin{aligned} \Pi(t, t+\tau) &= (V_{t+\tau} - V_t) - \sum_{n=0}^{N-1} \Delta_{t_n} \left(S_{t_{n+1}} - S_{t_n} \right) \\ &\quad - \sum_{n=0}^{N-1} \frac{a_n r_n}{365} \left(V_{t_n} - \Delta_{t_n} S_{t_n} \right), \end{aligned} \tag{1}$$

where V_t represents the option price at time t, S_t is the stock price at time t, Δ_t is the Black-Scholes delta of the option at time t, a_n is the number of calendar days between two successive trading dates, and r_t is the annualized risk-free interest rate.

As we see from Equation (1), the delta-hedged option gain can be separated into three components: the first term, $V_{t+\tau} - V_t$, is the difference between the end-of-period and initial option prices; the second term, $\sum_{n=0}^{N-1} \Delta_{t_n} \left(S_{t_{n+1}} - S_{t_n} \right)$, is the amount invested in the underlying stock for hedging the first-order stock price movement; and the last term, $\sum_{n=0}^{N-1} \frac{a_n r_n}{365} \left(V_{t_n} - \Delta_{t_n} S_{t_n} \right)$, is the difference between the option and stock positions that investors can borrow from or lend to the market. To make the delta-hedged option gain comparable across options, similar to Cao and Han (2013), we scale the dollar delta-hedged option gain by the absolute position in the options and underlying stock:[1]

$$r_{t,t+\tau} = \frac{\Pi(t, t+\tau)}{|\Delta_t S_t - V_t|}. \tag{2}$$

2.2. The Model

The literature on option pricing primarily employs parametric no-arbitrage models to investigate risk factors that affect option prices. However, such an approach needs a full specification of dynamics of the underlying stock price. Israelov and Kelly (2017) show that such parametric no-arbitrage models are unable to explain a large part of empirically observed variation in option returns.

In this paper, we rely on the factor pricing approach to investigate the risk-return relationship in the options market. We assume that the stochastic discount factor (SDF), m_{t+1}, is a linear function of some return-based risk factors, f_{t+1},

$$m_{t+1} = 1 - b_t'\big(f_{t+1} - \mu_{f,t}\big), \tag{3}$$

where b_t is a vector of (time-varying) SDF loadings and $\mu_{f,t}$ is expected values of factors. The condition of no-arbitrage ensures existence of a positive SDF such that

$$E_t[m_{t+1} r_{i,t+1}] = 0 \Rightarrow r_{i,t+1} = \alpha_{i,t} + \beta_{i,t}' f_{t+1} + \epsilon_{i,t+1}, \tag{4}$$

where the factor loadings, $\beta_{i,t} \equiv cov_t(r_{i,t+1}, f_{t+1}) var_t(f_{t+1})^{-1}$, $\alpha_{i,t}$ captures any mispricing in option returns, and the factor risk premium is given by $\lambda_t = E_t[F_{t+1}]$. Correct specification of the model results in zero mispricing and the expected option returns should be determined by $E_t[r_{i,t+1}] = \beta_{i,t}\lambda_t$.

It may be desirable to use some pre-specified observable factors. However, this is difficult for the options market as there do not exist well-established observable factors. Therefore, we rely on a factor model that treats the pricing factors as latent. The arbitrage

[1] In a recent study, Buchner and Kelly (2022) scale the dollar delta-hedged option gain by the underlying stock price.

pricing theory of Ross (1976) (also see, Huberman, 1982; Chamberlain and Rothschild, 1983; Ingersoll, 1984) guarantees the existence of a factor model and bounded pricing errors. Under some mild conditions, the model can be written in matrix form as

$$\mathbf{R} = \boldsymbol{\alpha}\mathbf{1}_T' + \mathbf{B}\mathbf{F}' + \mathbf{E}, \tag{5}$$

where $\mathbf{R}$ is a $N \times T$ matrix containing N option returns for a time period $t = 1, \ldots, T$, $\mathbf{F}$ is a $T \times K$ matrix of latent factors summarizing the systematic risk in the option market, $\boldsymbol{\alpha}$ is a $N \times 1$ vector capturing the pricing errors, $\mathbf{B}$ is a $N \times K$ factor loading matrix, and $\mathbf{E}$ is a $N \times T$ matrix containing the error term for each return at each time.

The option contractual structure with finite maturity and different moneyness suggests that both alpha and beta in Equation (5) should depend on option characteristics, and more importantly, the dependence structure of alpha and beta on characteristics should also be changing over time due to, for example, structural changes of the market, and different types of alpha appearing and disappearing quickly at different times. Therefore, we assume that $\boldsymbol{\alpha}$ and $\mathbf{B}$ are functions of firm and option characteristics, and for a short time period, those characteristics are relatively stable. Let $\mathbf{X}$ be a $N \times L$ matrix of characteristics and assume that both $\boldsymbol{\alpha}$ and $\mathbf{B}$ are affine functions of $\mathbf{X}$ as follows,

$$\boldsymbol{\alpha} = \mathbf{X}\boldsymbol{\theta} + \Gamma_\alpha, \qquad \mathbf{B} = \mathbf{X}\boldsymbol{\beta} + \Gamma_\beta, \tag{6}$$

where θ is a $L \times 1$ vector, $\boldsymbol{\beta}$ is a $L \times K$ matrix of constant, and the $N \times 1$ vector Γ_α and the $N \times K$ matrix Γ_β are orthogonal to $\mathbf{X}$, absorbing the parts of $\boldsymbol{\alpha}$ and $\mathbf{B}$ that are not related to characteristics. It becomes clear from the above setup that the option return predictability is either due to characteristics predicting alpha (mispricing) or due to characteristics predicting beta (systematic risk). By further exploring $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$, we can learn contributions of characteristics to $\boldsymbol{\alpha}$ and $\mathbf{B}$ and deepen our understanding of exact sources of option return predictability. Combining Equations (5) and (6), we rewrite the factor model as

$$\mathbf{R} = (\mathbf{X}\boldsymbol{\theta} + \Gamma_\alpha)\mathbf{1}_T' + (\mathbf{X}\boldsymbol{\beta} + \Gamma_\beta)\mathbf{F}' + \mathbf{E}. \tag{7}$$

2.3. Estimation

Our aim is not only to introduce time-varying alpha and beta, but also to make the dependence of alpha and beta on characteristics changing over time. These two features are important for investigating the risk-return relationship in the option market. For this purpose, we rely on the projected principal component (PPCA) approach originally proposed by Fan, Liao, and Wang (2016) to estimate the model. The PPCA estimator of the factor loading function converges to the true one as the dimension (N) of the cross section increases, even for small time-series samples (T). This property allows us for estimating the model using a short rolling window over time and hence introducing time-varying dependence structure. While the original approach of Fan, Liao, and Wang (2016) explicitly imposes zero restriction on alpha, a recent study by Kim, Korajczyk, and Neuhierl (2021) extend this approach to estimate not only the latent factors and their loadings, but also the mispricing function that relates alpha to characteristics.

We assume that characteristics are relatively stable over a short-time interval and estimate the model following the rolling procedure of Kim, Korajczyk, and Neuhierl (2021), which actually treats the model of Equation (7) as a locally unconditional approximation to a conditional model. Within each estimation window, the dependence of $\boldsymbol{\alpha}$ and $\mathbf{B}$ on characteristics are constant; however, the rolling estimation of the model enables us to study both the cross-sectional and the temporal relations of characteristics with systematic risk and mispricing.

Under some standard regularities of factor models (see, Fan, Liao, and Wang, 2016; Kim, Korajczyk, and Neuhierl, 2021), within each time window, the estimation with the extended PPCA is implemented in the following four steps. First, we demean option returns to remove alpha and then project the demeaned time-series option returns on characteristics to obtain the demeaned projected option returns that only depend on factor loadings ($\mathbf{X}\boldsymbol{\beta}$) and demeaned latent factors. Second, we apply the standard PCA to the demeaned projected option returns to obtain estimates of factor loadings and factors. Up to here, the procedure gives characteristics maximal explanatory power to systematic risk. Third, we estimate the mispricing function ($\mathbf{X}\boldsymbol{\theta}$) by regressing average returns of individual options on characteristics orthogonal to the estimated factor loading ($\mathbf{X}\hat{\boldsymbol{\beta}}$), which is equivalent to project residuals of average option returns to risk premia ($\mathbf{X}\hat{\boldsymbol{\beta}}\bar{\mathbf{F}}$) onto characteristics. Finally, we use the estimated mispricing function to construct an alpha portfolio,

$$\hat{\boldsymbol{w}} = \frac{1}{N}\mathbf{X}\hat{\boldsymbol{\theta}}. \tag{8}$$

To be specific, for the time window, $t = 0, 1, \ldots, T$ (e.g., $T = 12$ months), we use the option returns from $t = 1, \ldots, T$ and characteristics at time 0, $\mathbf{X}_0$, to obtain $\hat{\boldsymbol{\theta}}$. The alpha portfolio weights are then computed using the updated characteristics at time T, $\mathbf{X}_T$, as $\hat{\boldsymbol{w}}_{\boldsymbol{T}} = \frac{1}{N_T}\mathbf{X}_T\hat{\boldsymbol{\theta}}$, which are used to compute the alpha portfolio return at $T+1$, $R^a_{p,T+1} = \hat{\boldsymbol{w}}'_{\boldsymbol{T}}\mathbf{R}_{T+1}$. The characteristics $\mathbf{X}$ are standardized to have zero means, suggesting that the portfolio weights $\hat{\boldsymbol{w}}_{\boldsymbol{T}}$ sum to zero and the portfolio is a zero-cost investment. This procedure is repeated recursively at each month up to the end of sample. The portfolio constructed in this way is an alpha portfolio without exposure to systematic risk based on pure alpha that is related to characteristics,[2] is completely out-of-sample, and hence can be regarded as a realization of mispricing in the option market.

3. Data and Summary Statistics

3.1. Option Data

Our option data are obtained from OptionMetrics IvyDB, a primary database for research on options. This dataset includes options on individual equities, indexes, and exchange-traded funds. In the paper, we focus on options on individual common equities listed on NYSE, NYSE American (formerly AMEX), and NASDAQ for the period of March 1996 to December 2021. The information provided by OptionMetrics includes option maturities, strikes, implied volatilities, and sensitivity measures (Greeks), as well as information on underlying stocks. We use both volatility surface data and single equity options data and mainly focus on short-maturity at-the-money options given that those options are the

[2] There exists the other source of mispricing that is unrelated to characteristics due to the constant term Γ_α.

most liquid and hence should suffer less from mispricing.

Volatility Surface Data. We download the volatility surface file, which contains the interpolated implied volatility on standardized options with respect to fixed deltas and maturities for each security on each day. The implied volatility is computed using binomial trees that take into account the early exercise of individual stock options and the dividends expected to be paid over the life of the options. Volatility surface options can be regarded as portfolios composed of single equity options used in interpolation, and we believe that those data should contain the same but clean information as single equity options. The interpolation leaves us fewer missing values, and we can build our empirical analyses on a larger sample. We select at-the-money options (delta equal to 0.5 for calls and -0.5 for puts) with time-to-maturity of 30 days, and we then construct one-month delta-hedged option returns at the end of each month.

Single Equity Options Data. We also extract single equity options from the option price file. Following the literature (see, e.g., Goyal and Saretto, 2009; Cao and Han, 2013; Zhan, Han, Cao, and Tong, 2022; Bali, Beckmeyer, Moerke, and Weigert, 2022), we apply several filters to the data. First, we remove those options whose implied volatility and Greeks are unavailable. Second, we eliminate those options whose underlying stocks pay dividends during the one-month investment period. Third, we drop options with zero trading volume over the last seven calendar days to ensure the options are liquid enough. Fourth, to avoid the microstructure noise, we delete the options whose best bid price is zero, the best offer price is not strictly larger than the best bid price, the option price is smaller than \$0.125, or the relative bid-ask spread is larger than 50%. Fifth, we drop those options that violate the American option bounds and the convexity condition (Bollerslev, Todorov, and Xu, 2015). Finally, from those options that survive, we select one call and one put whose moneyness are closest to one and that have the shortest maturity but longer than one month at the end of each month for each underlying stock.

Summary Statistics. To show the distribution of option returns, we pool all observations of the option-date pairs and present the summary statistics in Table 1. Panel A presents the summary statistics of option returns constructed from the volatility surface data, and Panel B and C reports summary statistics of call and put option returns, respectively, constructed from the single equity options. In general, the sample sizes of option returns from the volatility surface data are much larger than those from the single equity options data. Based on the volatility surface data, there are 293,459 observations of call returns and 297,275 observations of put returns, whereas based on single equity options data, there are only 90,356 observations of call returns and 68,045 observations of put returns.

When we use the volatility surface data, we see from Panel A that the monthly mean return of delta-hedged call options is negative, -1.07%, with a relatively high standard deviation of 16.68%, and the monthly mean return of delta-hedged put options is positive, 0.61%, also with a large standard deviation of 15.50%. However, from Panels B and C, we see that when we use single equity options data, the monthly mean returns of both call and put option returns are negative, -0.19% and -0.30%, respectively, and their standard deviations become smaller, 5.02% and 4.24%, respectively.

While for the volatility surface data, the maturity and delta of call and put options are fixed at 30 days and 0.5 (absolute value), respectively, the mean maturity of both call and put options in the single equity options data is slightly longer, about 49 days, with a standard deviation of about 2, and their mean moneyness is about 1, with a very small standard deviation of 0.03. We also present liquidity and size of the underlying stocks in the last two rows of each panel, we see that the mean Amihud illiquidity ratio in the volatility surface data (Panel A) is much larger than those in the single equity options data (Panels B and C), suggesting that the underlying stocks in the single equity options data are on average more liquid, and that the average size of the underlying stocks in the single equity options data (Panels B and C) is larger than that in the volatility surface data.

3.2. Characteristics

The characteristics used in the paper contain both firm- and option-specific variables. Our firm characteristics are mainly from Chen and Zimmermann (2022) who construct over

200 variables documented in the literature that have predictive power in the cross-section of stock returns.[3] Several studies have shown that a number of firm characteristics that can predict stock returns can also forecast option returns (see, e.g., Cao and Han, 2013; Hu and Jacobs, 2020; Karakaya, 2013; Vasquez, 2017; Zhan, Han, Cao, and Tong, 2022). We only choose those variables that have continuous values with missing values less than 15%, and we also drop several characteristics that are very highly correlated with other variables. As a result, we choose 79 firm characteristics from Chen and Zimmermann (2022) database.

Furthermore, the literature has found that a number of option-related variables have predictive power for stock returns and/or option returns (see recent studies by Neuhierl, Tang, Varneskov, and Zhou (2021) for stock returns and Bali, Beckmeyer, Moerke, and Weigert (2022) for option returns). We therefore construct a total of 31 option-related characteristics following this strand of literature. Given their relevance to option returns, size, stock price and stock return are also included in our study. The complete list of firm and options characteristics and the corresponding references are presented in Table A1 in the Internet Appendix. In total, there are 113 characteristics to be considered in our empirical analysis.

We cross-sectionally rank and standardize all characteristics at each month such that they are in the range of $[-0.5, 0.5]$. The missing values are then imputed to be 0. One advantage of using the cross-sectional ranks of characteristics is that the impact of potential data errors and outliers in individual characteristics can be largely alleviated (see, e.g., Kelly, Pruitt, and Su, 2019; Freyberger, Neuhierl, and Weber, 2020; Kozak, Nagel, and Santosh, 2020).

4. Empirical Results

In this section, we present our main empirical results based on the volatility surface data. To save space, we move results based on single equity options data to the Internet Appendix. Subsection 4.1 presents the overall performance of option alpha portfolios;

[3] Those data are publicly available at Andrew Y. Chen's website, www.openassetpricing.com.

Subsection 4.2 discusses whether the commonly used factors in the stock market and in the options market can explain abnormal returns of option alpha portfolios; and Subsection 4.3 examines contributions of characteristics to systematic risk and mispricing in the options market.

4.1. Performance of Option Alpha Portfolios

Following the extended PPCA procedure presented in Section 2, we construct option alpha portfolios with different numbers of latent factors and examine how the number of latent factors (K) affect performance of option alpha portfolios. We implement the extended PPCA using a rolling window approach with the window size of 12 months to construct alpha portfolios, for which individual option's weights are proportional to the corresponding estimated alphas as in Equation (8). To make the results comparable with common equity and option factors, we normalize the in-sample annualized standard deviation of alpha portfolios to 20%. Panels A and B of Table 2 summarizes the performance results of call and put alpha portfolios, respectively.

From Panel A, we see that the annualized Sharpe ratios of call alpha portfolios are relatively high, ranging from 1.17 to 1.62. The Sharpe ratio is about 1.52 when we choose only one latent factor; it then slightly increases to 1.62 when the number of factors is chosen to be 3; thereafter, the Sharpe ratio is steeply declining to 1.17 when the number of factors reaches 10. We also notice that the increase of Sharpe ratio from $K = 1$ to $K = 3$ is due to an increase of the alpha portfolio mean, and that the out-of-sample standard deviation is smaller than the in-sample standard deviation, which is normalized to 20%, when the mean or Sharpe ratio of the alpha portfolio reaches the highest level. The call alpha portfolio returns are in general right-skewed and leptokurtic.

Panel B on put alpha portfolios reveals quite similar implications. The annualized Sharpe ratios are also high, ranging from 1.28 when $K = 10$ to 1.86 when $K = 3$. Similar to call alpha portfolios, when the number of latent factors is larger than 3, the Sharpe ratio decreases. The put alpha portfolio returns with the maximal Sharpe ratio are also right-skewed and leptokurtic

The basic message delivered from Table 2 is that the Sharpe ratios of both call and put option alpha portfolios are relatively high no matter which number of latent factors we choose. What do high Sharpe ratios of call and put alpha portfolios imply? They suggest that mispricing may be prevailing in the options market. As we show, we build alpha portfolios based on estimated individual option alphas related to characteristics by firstly giving them maximal explanatory power for systematic risk premia. The exposures of alpha portfolios to systematic risks should approach zero when the number of individual options increase. Therefore, we regard high Sharpe ratios of option alpha portfolios as a realization of mispricing in the options market. We see that in general, Sharpe ratios of put alpha portfolios are higher than those of call alpha portfolios, suggesting that mispricing in put options may be more severe than in call options.

The number of true factors is unknown and needs to be estimated. As shown in Kim, Korajczyk, and Neuhierl (2021), if the number of extracted factors is smaller than the number of true factors, the extended PPCA can not guarantee orthogonality between alpha portfolio weights and factor loadings, and furthermore, too many factors may lead to inaccurate model estimates and harm the performance of alpha portfolios. Therefore, for precisely determining the number of latent factors, we follow the projected-PC eigenvalue-ratio method of Fan, Liao, and Wang (2016) to estimate the number of latent factors by maximizing the ratio of adjacent eigenvalues of the demeaned projected option returns (also see, Ahn and Horenstein, 2013; Lam and Yao, 2012). The basic idea of this method is that the K largest eigenvalues of sample covariance grow with respect to sample size, whereas the rest remains bounded or grows very slowly. We apply this projected-PC eigenvalue-ratio test in each rolling window in both call and put option samples and find that the optimal number of latent factors in both samples ranges from 1 to 3. The smaller number of factors may be insufficient to capture all systematic risks and is unable to ensure estimated options alphas not exposed to systematic risks. Jones (2006) argues that for explaining S&P 500 index option returns, the best-performing model should have 2-3 factors, and that more complex models are very unstable out of sample. Moreover, Kim, Korajczyk, and Neuhierl (2021) show that a slight increase of the number of extracted factors does not harm the model performance materially. We therefore select the number of factors equal to 3 in most of our analyses. We have already seen that at $K = 3$, the Sharpe ratios of both call and put alpha portfolios reach the highest values, 1.62 and 1.86, respectively.

The last two columns of Table 2 present the minimum and maximum returns of alpha portfolios, moderate values of which reveal that high Sharpe ratios of alpha portfolios are not driven by extreme values. We further check if the performance of alpha portfolios could be attributed to any small subsamples of options. We presents the distributions of individual option weights in call and put alpha portfolios for $K = 3$ at each month in the upper and lower panels of Figure 1, respectively. We see that for both call and put alpha portfolios, the ranges of weights of individual options change over time and they are always within narrow bounds, suggesting that any individual options can not have dominant impact on the performance of both call and put alpha portfolios. Therefore, we can conclude that high Sharpe ratios of both call and put alpha portfolios originate most likely from economically substantial mispricing in the options market.

We further explore how the performance of alpha portfolios evolves over time. Figure 2 shows the natural logarithm of cumulative returns of call and put alpha portfolios with $K = 1$ and 3, their corresponding market portfolios, and the stock market portfolio in the upper and lower panels, respectively. The market portfolio of call (put) options is simply the equal-weighted average of all call (put) option returns using all data in the volatility surface file. To ensure the returns of different portfolios are comparable, we normalize the alpha portfolios to have the same standard deviations as the stock market portfolios. We see that the cumulative returns of both call and put alpha portfolios increase dramatically over time, whereas the cumulative returns of the call market portfolio are decreasing over time and the cumulative returns of the put market portfolio hardly display any increasing or decreasing tendency. The cumulative returns of the stock market portfolio slightly increase over time. The price paths of alpha portfolios with $K = 1$ and 3 look very similar. We further notice that the increase of cumulative returns of both call and put alpha portfolios is much faster in the beginning period of the sample, suggesting that mispricing is even more serious in this period. As a robustness check in Section 5, we reconstruct alpha portfolios using the post-2004 sample and still find considerable Sharpe ratios for both call and put alpha portfolios.

4.2. Option Alpha Portfolios and Risk Factors

Our alpha portfolios are constructed based on estimates of alphas that are related to characteristics and are orthogonal to systematic risk in the options market. In this part, we further examine whether the commonly used risk factors in the stock and options markets can explain abnormal returns of both call and put alpha portfolios.

Option Market Factors. Different from the stock market, there do not exist well-established observable risk factors in the options market, and studies examining the factor structure on option returns are relatively sparse. Following the literature, we rely on the three factors of Karakaya (2013), the liquidity factor (see, e.g., Christoffersen, Goyenko, Jacobs, and Karoui, 2018), and the idiosyncratic risk factor (see, e.g., Cao and Han, 2013). Karakaya (2013) proposes three factors, namely, level, maturity, and value, to explain variations in the cross-section of individual option returns. The options used in Karakaya (2013) have various moneyness and maturities, whereas in our paper, we mainly focus on short-maturity at-the-money options that are the most liquid and are less mispriced. To give those three factors maximal explanatory power, we modify them slightly by using at-the-money options only. Similar to Karakaya (2013), we define the level factor as the equal-weighted average excess returns of at-the-money options; we define the maturity factor as the equal-weighted average of return difference between 6-month and 1-month at-the-money options; and the value factor is defined as the high-minus-low returns of decile portfolios sorted on the value, defined as the difference between average implied volatility of at-the-money call and put options and realized volatility of the underlying stock (Goyal and Saretto, 2009).

Cao and Han (2013) find that delta-hedged option returns decreases monotonically with an increase in the underlying idiosyncratic volatility. Another study by Christoffersen, Goyenko, Jacobs, and Karoui (2018) find that there exist significant illiquidity premia in the equity options market. A recent paper by Zhan, Han, Cao, and Tong (2022) show that the option idiosyncratic volatility and illiquidity factors can explain option return spreads constructed based on well-studied 10 firm characteristics. Therefore, using all at-the-money call and put options in our sample, we construct these two factors as high-minus-low spread returns of decile option portfolios based on the Amihud illiquidity ratio and idiosyncratic volatility of the underlying stock.

Table 3 presents the regression results of the call and put alpha portfolio returns ($K = 3$) on those option market factors in Panels A and B, respectively. The first column of each panel shows that the alpha portfolios earn statistically significant average monthly excess returns, 2.31% ($t = 7.42$) for the call alpha portfolio and 2.67% ($t = 7.90$) for the put alpha portfolio. When we use the level factor alone, akin to the CAPM in the options market, we find that the risk-adjusted return is about 2.25% for the call alpha portfolio and is about 2.62% for the put alpha portfolio, both of which are highly statistically significant ($t = 6.90/7.12$, respectively). However, when we use the Karakaya three factors, the risk-adjusted return increases to 3.17% for the call alpha portfolio and to 3.56% for the put alpha portfolio, both of which are still highly statistically significant. We then combine the level factor and the illiquidity and idiosyncratic volatility factors together. We still find highly statistically significant risk-adjusted returns for both call and put alpha portfolios (2.06% and 2.41%, respectively). Finally, in the last column, we combine all five factors together and find that the risk-adjusted return is about 2.92% ($t = 3.95$) for the call alpha portfolio and is about 3.03% ($t = 4.04$) for the put alpha portfolio. We also find that none of the factors is statistically significant in explaining the call and put alpha portfolio returns.

We also check how the risk-adjusted returns change with an increase in the number of latent factors in Figure 3. In general, when the model is more likely misspecified, i.e., the number of latent factors is too small (e.g., only one factor) or too large (e.g., 10 factors), the risk-adjusted returns become relatively small for both call and put alpha portfolios. We also find that the risk-adjusted returns from the factor models that contain Karakaya factors are larger than those that exclude them.

Stock Market Factors. Given its high relevance to the options market, we also try the commonly used factors in the stock market, namely, the Fama-French five factors (Fama and French, 1992, 1996, 2015), the Carhart momentum factor (Carhart, 1997), the q factors (Hou, Xue, and Zhang, 2015; Hou, Mo, Xue, and Zhang, 2021), and the mispricing factors (Stambaugh and Yuan, 2017). We run the time-series regressions of returns of alpha portfolios based on three latent factors in PPCA on those common equity factors and check if the superior performance of option alpha portfolios can be attributed to risks related to the stock market. The regression results are presented in Table 4. Panel A is for the call alpha portfolio and panel B for the put alpha portfolio.

In columns 1-6, we introduce the market factor, the Fama-French five factors, the FF5 augmented by momentum factors, the q four factors, the q five factors, and the mispricing four factors, respectively. We see that the alpha estimates from those factor models are very close to the average monthly excess returns of alpha portfolios and are all highly statistically significant. Columns 2 and 3 reveal that except the stock market factor, the four Fama-French factors and the momentum factor are statistically significant in the regression of the call alpha portfolio returns, and except the stock market and momentum factors, the four Fama-French factors are statistically significant in the regression of the put alpha portfolio returns. Both call and put alpha portfolio returns are negatively related to the value and profitability factors and are positively related to the size and investment factors. Columns 4 and 5 show that the size factor is significant in explaining both call and put alpha portfolios in q factors. In the last column, we find that when using mispricing factors, the alpha estimates are the smallest for both call and put alpha portfolios, but the magnitudes are still economically substantial, 1.89% and 2.34% in call and put alpha portfolios, respectively. The explanatory power of mispricing factors for the call alpha portfolio returns mainly comes from the size and performance factors, while that for the put alpha portfolio only steps from the size factor. The above results suggest that the commonly used risk factors in the stock market have trivial explanatory power for abnormal returns of both call and put alpha portfolios. Figure A1 in the Internet Appendix presents how the alpha estimates (risk-adjusted returns) resulting from the different stock factor models vary with respect to the number of the latent factors. We see that the risk-adjusted returns of alpha portfolios are relatively stable with respect to K.

Option Latent Factors. We also try the latent option factors to explain returns on option alpha portfolios. For this purpose, we rely on the Principal Component Analysis (PCA) and the Risk-Premium Principal Component Analysis (RP-PCA, Lettau and Pelger, 2020) to extract the latent factors of call (put) option returns. Note that even though our option alpha portfolios are constructed in real-time recursively using a small window, we use the whole sample to extract latent factors to give them more information. Following Lettau and Pelger (2020), we set the hyperparameter of λ in RP-PCA to 10. Increasing λ to larger values has negligible impact on the results. Table A2 and Figure A2 in the Internet Appendix present the main results, which show that those latent factors are also incapable of explaining abnormal returns of both call and put alpha portfolios.

4.3. Contributions of Characteristics to Mispricing and Systematic Risk

In the extended PPCA, within each estimation period, factor loadings and mispricing are functions of characteristics observed at the beginning of the period. A natural question is then how and which of those characteristics affect factor loadings and mispricing over time. For this purpose, we regress the estimated factor loadings and mispricing on characteristics in each period, and the estimated $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ (see Equation (6)) capture the relationships between characteristics and systematic risk and mispricing, respectively. The estimated coefficients of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ are a $L \times K$ matrix and a $L \times 1$ vector, respectively. We evaluate the contribution of characteristic l to systematic risk as $\sum_{k=1}^{K} |\hat{\boldsymbol{\beta}}_{k,l}|$ and to mispricing as $|\hat{\boldsymbol{\theta}}_l|$. To make contributions of characteristics comparable over time, we normalize the contributions cross-sectionally at each period such that the largest contribution of a characteristic is always equal to one.

Figure 4 presents the heatmap of characteristics to systematic risk for call option returns. The heatmap for put option returns is similar and presented in Figure A3 in the Internet Appendix. We clearly see that the contributions of characteristics vary both in the cross section and in the time series. In the cross-section, we see that there are a number of characteristics whose contributions are sizable, whereas many others have negligible contributions. For example, those variables related to underlying equity and option liquidity (e.g., illiquidity, VolMkt, vol/dvol) are closely related to systematic risk; and the risk-neutral moments (e.g., 9m/12m skewness and kurtosis, upside and downside semivariances of Huang and Li (2019)) also play important roles in factor loadings. We also find that the characteristics with significant contributions in the heatmap for put option returns are very similar to those for call option returns. We further notice that the contributions of those key characteristics change over time. The time-varying contributions of characteristics suggest that the temporal dependence between characteristics and factor loadings is not constant. The constant temporal dependence is a standard assumption usually adopted in conditional factor pricing models (including the IPCA approach of Kelly, Pruitt, and Su (2019)), and our approach based on the extended PPCA easily enable us to capture such a time-varying relation between characteristics and risk and mispricing.

Figure 5 presents the heatmap of characteristics to mispricing for call option returns (see Figure A4 in the Internet Appendix for put option returns). Compared with the heatmaps of systematic risk, we see that there are fewer characteristics that are consistently related to mispricing. Many characteristics are only related to mispricing for a few periods, reflecting that such mispricing disappears quickly. It seems that the characteristics related to option liquidity play very important role in determining mispricing: the variables such as vol, dvol, so, and dso periodically have sizable contributions. The mispricing heatmaps further highlight the importance of time-varying dependence between characteristics and mispricing.

The heatmaps are convenient for visually checking the cross-sectional contributions of characteristics and the time-series variations of dependence of systematic risk and mispricing on characteristics. We further quantify the aggregate contributions of characteristics by taking the time-series average of contributions of each characteristic. Table 5 summarize the time-series average contributions of top 20 characteristics to mispricing and factor ladings. Panel A is for call option returns and Panel B for put option returns. From both panels, we see that most characteristics contributing to both systematic risk and mispricing can be roughly categorized into Liquidity and Risk, according to the classification of Bali, Beckmeyer, Moerke, and Weigert (2022) and Chen and Zimmermann (2022). The liquidity group contains measures of liquidity of options and underlying stocks, e.g., option trading volume (vol), trading volume of a stock scaled by its market equity (VolMkt), the Amihud illiquidity ratio of the underlying (Illiquidity), and the zero trading days (zerotrade) of the underlying. The underlying liquidity may also relate to options liquidity. The risk group includes characteristics that measure riskiness of the underlying stocks, e.g., long-term risk-neutral skewness and kurtosis (rns9m, rns12m, rnk9m, rnk12m), the difference between implied volatility and historical volatility (ivrv), and risk-neutral semivariance (mfvu, mfvd). Other characteristics such size and so (the stock-option trading volume ratio) are also closely related to option/stock liquidity. We refer readers to Table A1 in the Internet Appendix for the complete classification of all 113 characteristics. Our results in Table 5 suggest that liquidity- and risk-related characteristics contribute substantially to the factor loadings and are crucial in capturing mispricing, suggesting that option return predictability stems from the predicting power of characteristics to both mispricing and systematic risk. It seems that there exists systematic mispricing in the options market and those options with certain risk and liquidity features may expose higher mispricing than others.

Goyal and Saretto (2009) document option return predictability by the implied and historical variance spread, and they interpret it as a measure of volatility mispricing; however, Karakaya (2013) regards the implied and historical variance spread as a measure of option value. We find that the variables of ivrv and ivrv_ratio, two measures of the implied-historical volatility spread, contribute to both mispricing and systematic risk. Christoffersen, Goyenko, Jacobs, and Karoui (2018) find existence of significant illiquidity premia in the options market, and Zhan, Han, Cao, and Tong (2022) employ the Amihud illiquidity ratio of the underlying stock as a measure of the option liquidity. Our results suggest that the illiquidity premium is related to both systematic risk and mispricing. In a recent study, using the machine learning approach, Bali, Beckmeyer, Moerke, and Weigert (2022) also find that those characteristics related to liquidity and risk play important roles in predicting future option returns. However, all the above studies are silent and do not quantify how much option return predictability is from predicting mispricing or systematic risk or both.

To better understand the relationships between alpha portfolios and characteristics, we compare ranks of characteristics in the long and short legs. The long (short) leg is the group of option returns that receive the positive (negative) weights in the alpha portfolios. We calculate the average value of the rank-transformed characteristics in each leg at each time. We present time-series ranks of size and characteristics related to Illiquidity and Risk in long and short legs for call alpha portfolios in Figure A5 in the Internet Appendix. The ranks of characteristics in the long and short legs change over time, further supporting that the temporal dependence of mispricing on characteristics are time-varying. We find that the call alpha portfolio generally long (short) options contingent on large-cap (small-cap) underlying stocks; however, for some periods, e.g., during the 2008 financial crisis, the alpha portfolio takes long positions on call options of small-cap firms. The ranks of the Amihud Illiquidity ratio are similar to those of size except for the reverse of the long and short legs, which suggests that the contribution of size may result from its correlation with the illiquidity. The spreads of ranks in 9-month risk-neutral skewness, upside and downside semivariances are relatively large, consistent with the findings in Table 5 that the risk-neutral moments are closely related to mispricing. For put options, the magnitudes of average ranks of characteristics in long and short legs are akin to those in call options (see Figure A6 in the Internet Appendix), but the signs are reversed. Given that the values of call and put options depend on the stock price movements in the opposite direction, the same characteristics should have the opposite impact on the call and put alpha portfolios.

To further investigate the source of performances of option alpha portfolios, we construct alpha portfolios using the subsample of options based on a certain number of characteristics. We choose 14 important risk- and liquidity-related characteristics in predicting option alphas from Table 5. We estimate the options alphas each month using the previous 12-month data with three latent factors, then divide all options into two groups based on the median of each characteristic. The high (low) group contains options with the characteristic above (below) the median. Within each group, we build alpha portfolios with option weights proportional to estimated option alphas. The alpha portfolios on high (low) groups are noted as high (low) alpha portfolios. The results are reported in Table 6. H-L is the Sharpe ratio difference between the high and low alpha portfolios. The annualized Sharpe ratios of alpha portfolios on option subsamples are in general smaller than those on full samples, which may result from the decrease in the sample size. The Sharpe ratio differences between high and low alpha portfolios are substantial. The maximum difference for call (put) options is 0.83 (0.82).

5. Robustness Checks

In this section, we implement several robustness checks. Subsection 5.1 expands the space of characteristics by adding interaction terms and examine how nonlinearity affects the performance of option alpha portfolios. Subsection 5.2 checks how the performance of option alpha portfolios changes with respect to different rolling window size used in the extended PPCA. Subsection 5.3 investigates how option alpha portfolios perform in the post-2004 sample. Subsection 5.4 explores whether our findings remain when we use single equity options data.

5.1. Nonlinearity

In the above analysis, we assume that alpha and factor loadings in option returns are linear functions of firm and option characteristics. Several recent studies have shown that nonlinearity, in particular, interactions of characteristics, play important roles in explaining equity excess returns (see, e.g., Freyberger, Neuhierl, and Weber, 2020; Kozak, Nagel, and Santosh, 2020). It then raises a question whether our results remain when we introduce interaction terms in functions of mispricing and factor loadings. For this purpose, we expand the space of characteristics by adding interaction terms. Given that the number of characteristics is large, interacting all characteristics leads to voluminous combinations. As we have already seen from the above (see Table 5) that the risk- and liquidity-related characteristics play much more important roles than others, we decide to interact those risk-neutral moments, namely, rns9m, rnk9m, ivrv, and mfvu, with those liquidity-related variables, namely, vol, illiquidity. We also add size and so to liquidity-related variables since they are highly correlated with underlying stock liquidity. Based on these expanded characteristics, we then reestimate the mispricing function using the same extended PPCA and construct the corresponding alpha portfolios.

The performance of the alpha portfolios is presented in Table 7. We see that when interaction terms are introduced, Sharpe ratios of alpha portfolios are still economically substantial. In fact, when $K = 3$, the Sharpe ratio is about 1.69 for the call alpha portfolio, and it is about 1.88 for the put alpha portfolio, both of which are larger than those values in the linear case observed in Table 2. We also notice that for the call and put alpha portfolio, the Sharpe ratio reaches the highest level when $K = 3$. The increase of Sharpe ratios mainly comes from the rise of portfolio returns. Those results may suggest that introducing interaction terms can help estimate mispricing and systematic risk more effectively.

The cross-sectional and temporal contributions of characteristics and their interactions to factor loadings and mispricing for call option returns are presented in Figure 6 and Figure 7, respectively. For brevity, we present corresponding figures for put option returns in the Internet Appendix (see Figure A7 and Figure A8). We see that these two figures are quite different from those in the linear case. Now the contributions concentrate only on a very small number of characteristics, in particular, on interaction terms. The feature of "on-and-off" of characteristics in the alpha heatmap becomes even more obvious. It seems that when interaction terms are introduced, contributions of many individual characteristics become negligible, highlighting importance of interaction terms in predicting option returns. Table 8 presents the top 20 characteristics that contribute to factor loadings and mispricing for call and put returns in Panels A and B, respectively. With comparison to Table 5, we find that many interaction terms jump in and the most important variables are still those of liquidity- and risk-related characteristics and their combinations. We notice that most of those characteristics belonging to the group of Others disappear, and only so and size remain since they are closely related to liquidity.

5.2. Size of Estimation Windows

In constructing alpha portfolios, we face a tradeoff: the longer estimation window gives us more observations to estimate systematic factors and mispricing, whereas the shorter one makes the assumption of stable characteristics within the window more acceptable. We check how the estimation window size affects the performance of alpha portfolios. We consider the window size ranging from 12 months to 24 months, and the number of latent factors is equal to 3. The performance of alpha portfolios is presented in Table 9. We see that for both call and put alpha portfolios, though the annualized standard deviations do not change that much, the annualized mean returns decline with respect to an increase of the window size. For call alpha portfolios, the Sharpe ratio decreases from 1.62 for the window size of 12 months to 1.24 for the window size of 24 months, and for put alpha portfolios, the Sharpe ratio decreases from 1.86 for the window size of 12 months to 1.32 for the window size of 24 months. The performance deterioration of alpha portfolios for the longer window size may result from our assumption that characteristics in the estimation window are constant. The longer the estimation window is, the worse this assumption becomes. Fortunately, the PPCA approach enables us to accurately estimate the latent factors even for a small sample size (T). We find that the characteristics significantly contributing to mispricing and systematic risks are very similar across different estimation windows.

5.3. Subsamples

We have seen that the increase of cumulative returns of both call and put alpha portfolios is much faster in the initial period of our sample. This may raise a concern that the high Sharpe ratios we have found are driven by option observations in this period. In this subsection, we explore the performance of alpha portfolios using the sample after 2004.

The results are summarized in Table 10, in which we choose the number of the latent factors equal to three. The means of post-2004 alpha portfolio returns are smaller than those of the entire sample; however, the standard deviations of post-2004 alpha portfolios also become smaller. Hence, the Sharpe ratios are comparable. The Sharpe ratios for call alpha portfolios range from 1.08 for the window size of 24 months to 1.70 for the window size of 18 months, and the Sharpe ratios for put alpha portfolios range from 1.20 for the window size of 24 months to 1.69 for the window size of 21 months.

5.4. Using Single Equity Options Data

All the above results are based on the volatility surface options data provided by OptionMatrics. The volatility surface options in our sample have the homogeneous moneyness and days-to-maturity; however, in practice such options do not exist and investors have only access to actual equity options. Therefore, we ask whether our findings remain when we use the single equity options data. Given that we require that each firm in our sample has successive non-missing observations in each estimation period, the option observations used to build alpha portfolios are fewer than those from the volatility surface options data (see Table 1).

We implement the same empirical exercises as above and report all the results in the Internet Appendix (Tables A3–A9). In short, we still find economically substantial mispricing in single equity options. For the model with three latent factors and with a window size of 12 months, the Sharpe ratio of the call alpha portfolio is about 1.13, and the Sharpe ratio of the put alpha portfolio is about 1.35. The abnormal returns of both call and put alpha portfolios can not be explained by commonly used observable risk factors. When constructing alpha portfolios using subsamples, the Sharpe ratio differences between risk-related characteristics subsamples are larger on average than those between liquidity-related subsamples. The decreases in the importance of liquidity-related characteristics are as expected since the single equity options we used are more liquid.

We further construct the alpha portfolios only using liquid options with positive trading volume, positive open interest, and option bid-ask spread smaller than its 75% cross-sectional quantile (which is 0.15 for call and 0.13 for put). The results are presented in Table A8 in the Internet Appendix. We again find economically substantial Sharpe ratios for both call and put alpha portfolios.

6. Conclusion

Options are not redundant assets. In the options market, investors may not hold options up to maturity and may be more interested in risk factors that drive variation of option returns and in quantifying any mispricing if exists in the holding period. However, this question is far less understood in the options market than in the corresponding underlying equity market.

In this paper, relying on the latent factor pricing approach, we examine whether options are mispriced under very general assumptions on systematic risk, and if so, how we can quantify such mispricing. Based on the projected principal component analysis approach of Fan, Liao, and Wang (2016) and Kim, Korajczyk, and Neuhierl (2021), our latent factor model can take into account time-varying dependence of systematic risk and mispricing on a large panel of firm and option characteristics. We construct option alpha portfolios based on the estimated individual option mispricing. When we choose the number of the latent factors equal to 3 and an estimation window with size of 12 months, the out-of-sample annualized Sharpe ratios of call and put alpha portfolios are 1.62 and 1.86, respectively, revealing economically substantial mispricing in the options market.

Commonly used risk factors in the stock and options markets cannot explain abnormal returns of option alpha portfolios. Our results are robust with different estimation window sizes and the post-2004 sample; the results remain when we use single equity options instead of the volatility surface option data. Characteristics related to risk-neutral moments and liquidity and their interactions play very important roles in capturing option mispricing; most characteristics that contribute to mispricing also contribute to systematic risk.

References

Ahn, S. and A. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica 81*, 1203–1227.

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

Bali, T. G., H. Beckmeyer, M. Moerke, and F. Weigert (2022). Option return predictability with machine learning and big data. Technical report, Georgetown University.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–659.

Bollerslev, T., V. Todorov, and L. Xu (2015). Tail risk premia and return predictability. *Journal of Financial Economics 118*, 113–134.

Brooks, R., D. Chance, and M. Shafaati (2018). The cross-section of individual equity option returns. Technical report, Lousiana State University.

Buchner, M. and B. Kelly (2022). A factor model for option returns. *Journal of Financial Economics 143*, 1140–1161.

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

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance 52*, 57–82.

Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica 51*, 1281–1304.

Chen, A. Y. and T. Zimmermann (2022). Open source cross-sectional asset pricing. *Critical Finance Review 11*, 207–264.

Christoffersen, P., M. Fournier, and K. Jacobs (2018). The factor structure in equity options. *Review of Financial Studies 31*, 595–637.

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

Coval, J. and T. Shumway (2001). Expected option returns. *Journal of Finance 56*, 983–1009.

Fama, E. and K. French (1992). The cross-section of expected stock returns. *Journal of Finance 47*, 427–465.

Fama, E. and K. French (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance 51*, 55–84.

Fama, E. and K. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*, 1–22.

Fan, J., Y. Liao, and W. Wang (2016). Projected principal component analysis in factor models. *Annals of Statistics 44*, 219–254.

Freyberger, J., A. Neuhierl, and M. Weber (2020). Dissecting characteristics nonparametrically. *Review of Financial Studies 33*, 2326–2377.

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

Goyal, A. and A. Saretto (2022). Are equity option returns abnormal? ipca says no. Technical report.

Horenstein, A., A. Vasquez, and X. Xiao (2020). Common factors in equity option returns. Technical report, University of Miami.

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

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

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

Huang, T. and J. Li (2019). Option-implied variance asymmetry and the cross-section of stock returns. *Journal of Banking and Finance 101*, 21–36.

Huberman, G. (1982). A simple approach to arbitrage pricing theory. *Journal of Economic Theory 28*, 183–191.

Ingersoll, J. E. (1984). Some results on the theory of arbitrage pricing. *Journal of Finance 39*, 1021–1039.

Israelov, R. and B. Kelly (2017). Forecasting the distribution of option returns. Technical report, Yale University.

Jones, C. (2006). A nonlinear factor analysis of S&P 500 index option returns. *Journal of Finance 61*, 2325–2363.

Karakaya, M. (2013). Characteristics and expected returns in individual equity options. Technical report, University of Chicago.

Kelly, B., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics 134*, 501–534.

Kim, S., R. Korajczyk, and A. Neuhierl (2021). Arbitrage portfolios. *Review of Financial Studies 34*, 2813–2856.

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

Lam, C. and Q. Yao (2012). Factor modeling for high-dimensional time series: Inference for the number of factors. *Annals of Statistics 40*, 694–726.

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

Neuhierl, A., X. Tang, R. Varneskov, and G. Zhou (2021). Option characteristics as cross-sectional predictors. Technical report, Washington University in St. Louis.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*, 341–360.

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

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., B. Han, J. Cao, and Q. Tong (2022). Option return predictability. *Review of Financial Studies 35*, 1394–1442.

Table 1: Summary Statistics

This table reports the distribution of delta-hedged option returns and the liquidity measures of the underlying. We collect options with underlying stocks listed in NYSE, NYSE American (formerly AMEX), or NASDAQ from OptionMetrics. The sample period is from March 1996 to December 2021. The delta-hedged options returns are calculated as a portfolio with one option hedged by Δ shares of underling stocks, following Cao and Han (2013). Δ is the options delta from the Black-Scholes-Merton model. We hold the portfolio for one month, and delta hedges are rebalanced daily. Amihud Illiquidity is the Amihud illiquidity ratio in Amihud (2002) of the underlying stocks, and size is the natural logarithm of the firms' market capitalization. Panel A reports the summary statistics of the volatility surface options. The summary statistics of Amihud illiquidity and size are for underlying of call options. Panel B and C displays the summary statistics of single equity call and put options. Moneyness is defined as the ratio of strike price to spot stock price. Days-to-maturity is the days between the date of portfolio construction and the expiration date of options. Vega is the options vega from the Black-Scholes-Merton model scaled by the stock price. Bid-ask spread is the difference between the options bid price and ask price scaled by the mean of bid and ask prices.

Panel A: Volatility Surface Options									
Variables	n		mean	SD	p90	p75	median	p25	p10
Delta-hedged call return	293459	(%)	-1.07	16.68	17.46	7.94	-0.56	-9.20	-19.76
Delta-hedged put return	297275	(%)	0.61	15.50	17.61	8.11	0.17	-7.72	-16.53
Amihud Illiquidity	293459		1.48	4.05	3.40	1.06	0.29	0.08	0.03
Size	289764		6.91	1.41	8.76	7.79	6.81	5.91	5.16
Panel B: Single Equity Call Options									
	n		mean	SD	p90	p75	median	p25	p10
Delta-hedged call return	90356	(%)	-0.19	5.02	4.32	1.23	-0.77	-2.46	-4.37
Moneyness = K/S	90356		1.00	0.03	1.05	1.02	1.00	0.98	0.96
Days-to-maturity	90356		49.32	2.04	52	51	49	48	47
Vega	90356		0.14	0.01	0.15	0.15	0.14	0.14	0.13
Bid-ask spread	90356		0.12	0.08	0.04	0.06	0.10	0.15	0.22
Amihud Illiquidity	89399		0.13	0.26	0.31	0.13	0.05	0.02	0.01
Size	90356		8.08	1.27	9.74	8.85	7.99	7.19	6.51
Panel C: Single Equity Put Options									
	n		mean	SD	p90	p75	median	p25	p10
Delta-hedged Put return	68045	(%)	-0.30	4.24	3.56	0.95	-0.81	-2.31	-3.94
Moneyness = K/S	68045		1.00	0.03	1.04	1.02	1.00	0.98	0.95
Days-to-maturity	68045		49.30	2.03	52	51	49	48	47
Vega	68045		0.14	0.01	0.15	0.15	0.14	0.14	0.13
Bid-ask spread	68045		0.11	0.08	0.03	0.06	0.09	0.13	0.21
Amihud Illiquidity	67292		0.08	0.13	0.19	0.09	0.04	0.02	0.01
Size	68045		8.30	1.28	9.97	9.09	8.20	7.40	6.72

Table 2: Option Alpha Portfolios

This table shows the summary statistics of the option alpha portfolios. At each month, we estimate the mapping from options alpha to characteristics, $\boldsymbol{\theta}$, by Projected Principal Component Analysis (PPCA) introduced in section 3, and compute the estimated alphas as the product of $\widehat{\boldsymbol{\theta}}$ and one-month ahead characteristics. The estimation window is 12 months. The weights of individual options in the alpha portfolios are proportional to the estimated options alpha. K ranges from one to 10, and it is the number of latent factors used in PPCA. Mean, SD, and SR shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period.

Panel A: Alpha Portfolio of Call Options							
K	Mean	SD	SR	Skew	Kurt	Min	Max
1	0.22	0.14	1.52	1.35	6.19	-0.12	0.27
2	0.23	0.18	1.31	0.35	2.31	-0.15	0.24
3	0.28	0.17	1.62	1.35	6.92	-0.14	0.34
4	0.28	0.18	1.60	0.84	5.17	-0.2	0.31
5	0.28	0.18	1.52	0.71	4.70	-0.17	0.31
6	0.27	0.19	1.41	0.77	5.40	-0.19	0.35
7	0.25	0.19	1.31	0.68	3.92	-0.19	0.31
8	0.26	0.21	1.27	0.64	3.40	-0.19	0.31
9	0.26	0.21	1.24	0.64	4.20	-0.18	0.32
10	0.24	0.20	1.17	0.51	4.34	-0.19	0.33
Panel B: Alpha Portfolio of Put Options							
K	Mean	SD	SR	Skew	Kurt	Min	Max
1	0.27	0.15	1.80	2.10	11.14	-0.11	0.33
2	0.28	0.18	1.57	0.86	3.72	-0.17	0.25
3	0.32	0.17	1.86	1.11	4.73	-0.17	0.26
4	0.31	0.18	1.72	1.01	3.84	-0.14	0.25
5	0.29	0.17	1.68	0.63	2.55	-0.12	0.24
6	0.30	0.18	1.68	0.70	2.70	-0.12	0.26
7	0.29	0.18	1.58	0.83	2.89	-0.14	0.25
8	0.30	0.19	1.62	0.89	2.65	-0.13	0.25
9	0.29	0.19	1.59	0.86	2.68	-0.12	0.25
10	0.26	0.20	1.28	0.77	3.27	-0.14	0.29

Table 3: Explanatory Power of Options Risk Factors

This table reveals the explanatory power of option risk factors on the option alpha portfolios returns. We regress the monthly returns of the option alpha portfolios on the common risk factors in the options market and report the alpha (in percentage) and factor loadings. Level is the equal-weighted average excess returns of at-the-money options. Karakaya contains level, maturity and value factors. Maturity factor is the equal-weighted average return difference between six-month and one-month at-the-money options. Value factor is the high-minus-low returns of decile portfolios sorted on the value, the difference between average implied volatility of at-the-money call and put options and realized volatility of the underlying stock. Newey-West (1987) adjusted t-statistics are reported in the parathesis. ***/**/* indicate the significant at 1%, 5% and 10% confidence level, respectively.

Panel A: Call Option Alpha Portfolio

	Excess ret.	Level	Karakaya	Level + Illiq. + IdioRisk	Karakaya+ Illiq. + IdioRisk
Alpha	2.31***	2.25***	3.17***	2.05***	2.92***
	(7.42)	(6.90)	(4.44)	(6.67)	(3.95)
Level		-0.23	-0.09	0.13	0.19
		(-0.85)	(-0.35)	(0.42)	(0.68)
Maturity			-0.26		-0.23
			(-0.58)		(-0.51)
Value			-0.50*		-0.46
			(-1.67)		(-1.48)
Illiquidity				0.22	0.04
				(0.50)	(0.09)
IdioRisk				0.81*	0.70
				(1.65)	(1.39)
Adj. R^2	0.00	-0.00	0.01	0.01	0.01
Num. obs.	298	298	298	298	298

Panel B: Put Option Alpha Portfolio

	Excess ret.	Level	Karakaya	Level + Illiq. + IdioRisk	Karakaya+ Illiq. + IdioRisk
Alpha	2.67***	2.62***	3.56***	2.41***	3.03***
	(7.90)	(7.12)	(4.67)	(6.93)	(4.04)
Level		-0.17	-0.18	0.31	0.25
		(-0.47)	(-0.59)	(0.80)	(0.75)
Maturity			0.15		0.09
			(0.35)		(0.21)
Value			-0.51*		-0.33
			(-1.65)		(-1.12)
Illiquidity				0.69	0.53
				(1.58)	(1.17)
IdioRisk				0.95*	0.87*
				(1.94)	(1.73)
Adj. R^2	0.00	-0.00	0.01	0.02	0.02
Num. obs.	298	298	298	298	298

Table 4: Explanatory Power of Equity Risk Factors

This table reveals the explanatory power of equity risk factors on the option alpha portfolios returns. We regress the monthly returns of the option alpha portfolios on the common risk factors in the stock market and report the alpha (in percentage) and factor loadings. Market is the stock market factor. FF3 is the market, size, and value factors in Fama and French (1993). UMD is the momentum factor in Carhart (1997). FF5 is the market, size, value, investment, and profitability factors in Fama and French (2015). Q4 contains the market, size, investment and profitability factors proposed by Hou, Xue and Zhang (2015). Q5 is the Q4 plus the expected investment growth factors in Hou, Mo, Xue and Zhang (2021). MF4 consists of market, size, management and performance factors in Stambsugh and Yuan (2017). Newey-West (1987) adjusted t-statistics are reported in the parathesis. ***/**/* indicate the significant at 1%, 5% and 10% confidence level, respectively.

Panel A: Call Option Alpha Portfolio						
	Market	FF5	FF5+umd	Q4	Q5	MF4
alpha	2.31***	2.28***	2.33***	2.41***	2.24***	1.89***
	(7.59)	(8.13)	(7.98)	(7.57)	(8.01)	(6.93)
mktrf	0.04	-0.06	0.00	-0.14	-0.10	0.11
	(0.53)	(-0.83)	(0.02)	(-1.56)	(-1.15)	(1.32)
smb		0.36***	0.33***			
		(3.09)	(3.00)			
hml		-0.48***	-0.37***			
		(-4.03)	(-3.23)			
rmw		-0.37**	-0.40***			
		(-2.39)	(-2.72)			
cma		0.52***	0.48***			
		(2.65)	(2.59)			
umd			0.18			
			(1.46)			
me				0.45**	0.48**	
				(2.24)	(2.37)	
ia				-0.20	-0.18	
				(-0.94)	(-0.83)	
roe				-0.22*	-0.33**	
				(-1.75)	(-2.19)	
eg					0.30	
					(1.46)	
size						0.48**
						(2.47)
mgmt						0.21
						(0.97)
perf						0.25**
						(2.01)
Adj. R^2	-0.00	0.22	0.25	0.12	0.12	0.11
Num. obs.	298	298	298	298	298	298

Panel B: Put Option Alpha Portfolio						
	Market	FF5	FF5+umd	Q4	Q5	MF4
alpha	2.67***	2.76***	2.71***	2.89***	2.70***	2.34***
	(7.99)	(8.36)	(7.99)	(7.57)	(7.86)	(7.30)
mktrf	0.00	-0.11	-0.07	-0.20	-0.16	0.06
	(0.00)	(-1.55)	(-0.85)	(-2.13)	(-1.75)	(0.61)
smb		0.32***	0.30**			
		(2.78)	(2.56)			
hml		-0.45***	-0.38***			
		(-3.45)	(-3.15)			
rmw		-0.36***	-0.38***			
		(-2.66)	(-2.90)			
cma		0.38**	0.36**			
		(2.05)	(2.03)			
umd			0.11			
			(0.83)			
me				0.33**	0.36**	
				(1.97)	(2.14)	
ia				-0.22	-0.20	
				(-1.12)	(-0.99)	
roe				-0.35**	-0.46***	
				(-2.21)	(-2.60)	
eg					0.32	
					(1.44)	
size						0.42**
						(2.49)
mgmt						0.17
						(0.74)
perf						0.20
						(1.45)
Adj. R^2	-0.00	0.19	0.20	0.10	0.11	0.07
Num. obs.	298	298	298	298	298	298

Table 5: Contribution of Characteristics to Options Alphas and Factor Loadings

This table reports the top 20 characteristics contributing most to options alpha and factor loadings. The contributions of characteristics on options alpha are measured by $\widehat{\boldsymbol{\theta}}$, the mapping from the options alpha to characteristics. The contributions of characteristics on options factor loadings are measured by $\widehat{\boldsymbol{\beta}}$, the mapping from the options beta to characteristics. PPCA estimate $\mathbf{X}\boldsymbol{\beta}$ as a whole, and we regress it on characteristics $\mathbf{X}$ to recover $\widehat{\boldsymbol{\beta}}$. The absolute values of betas of characteristics are summed to measure the total contribution. The contributions of characteristics to options alpha and factor loadings are scaled by the cross-section maximum, so the contributions in each period range from zero to one. According to Chen and Zimmermann (2020) and Bali et al. (2021), most characteristics are divided into risk and liquidity groups. The rest is divided into the Others group. The definitions of the characteristics can be found in the Internet Appendix.

	Panel A: Characteristics Contribution (Call)						Panel B: Characteristics Contribution (Put)					
	Factor Loadings			Alpha			Factor Loadings			Alpha		
	Char.	Cat.	Contr.	Char.	Cat.	Contr.	Char.	Cat.	Contr.	Char.	Cat.	Contr.
1	vol	Liquidity	0.6834	vol	Liquidity	0.584	dvol	Liquidity	0.6583	dvol	Liquidity	0.5701
2	VolMkt	Liquidity	0.6453	dvol	Liquidity	0.5689	vol	Liquidity	0.648	vol	Liquidity	0.54
3	dvol	Liquidity	0.6377	AM	Others	0.456	VolMkt	Liquidity	0.6447	AM	Others	0.4378
4	rns9m	Risk	0.5478	so	Others	0.4153	rns9m	Risk	0.5696	rns9m	Risk	0.4112
5	rnk9m	Risk	0.5336	rns9m	Risk	0.397	rnk9m	Risk	0.5221	rnk9m	RIsk	0.406
6	rnk12m	Risk	0.5031	rnk9m	RIsk	0.3797	Illiquidity	Liquidity	0.5011	so	Others	0.3813
7	mfvu	Risk	0.4947	rns12m	Risk	0.3781	rns12m	RIsk	0.4989	rns12m	RIsk	0.3685
8	Illiquidity	Liquidity	0.4883	dso	Others	0.3552	ivrv	Risk	0.4914	size	Others	0.362
9	zerotradeAlt12	Liquidity	0.4829	size	Others	0.3525	ivrv_ratio	RIsk	0.491	stock_price	Others	0.3522
10	ivrv_ratio	RIsk	0.4743	mfvu	Risk	0.327	zerotradeAlt12	Liquidity	0.471	rnk12m	RIsk	0.3487
11	rns12m	Risk	0.4735	mfvd	Risk	0.323	so	Others	0.4664	dso	Others	0.3364
12	AM	Others	0.4677	ivrv_ratio	Risk	0.3207	mfvu	Risk	0.464	mfvu	RIsk	0.3357
13	ivrv	Risk	0.4624	rnk12m	Risk	0.3204	AM	Others	0.4617	ivrv_ratio	RIsk	0.3347
14	so	Others	0.4504	ivrv	RIsk	0.3187	rnk12m	RIsk	0.4601	ivrv	Risk	0.3303
15	size	Others	0.424	Illiquidity	Liquidity	0.3102	size	Others	0.4436	Illiquidity	Liquidity	0.329
16	DolVol	Liquidity	0.42	stock_price	Others	0.299	Mom6m	Others	0.4342	mfvd	Risk	0.3224
17	Mom6m	Others	0.4181	zerotrade	Liquidity	0.2847	DolVol	Liquidity	0.4188	zerotrade	Liquidity	0.2951
18	mfvd	Risk	0.4161	DolVol	Liquidity	0.281	dso	Others	0.4155	DolVol	Liquidity	0.282
19	zerotrade	Liquidity	0.4095	SP	Others	0.2763	zerotrade	Liquidity	0.401	VolMkt	Liquidity	0.2662
20	dso	Others	0.39	VolMkt	Liquidity	0.2552	mfvd	Risk	0.4001	SP	Others	0.2496

Table 6: Alpha Portfolios on Subsamples of Options

This table shows the Annualized Sharpe ratios of alpha portfolios on options above and below characteristics median. Options alphas are estimated using all call or put options, while the high (low) alpha portfolios are constructed using options with the characteristic above (below) its median. Estimation window is 12 months and the latent factors are set to three. $abs(H - L)$ shows the absolute value of Sharpe rstio differences between high alpha portfolios and low alpha portfolios.

		Panel A. Call Options			Panel B: Put Options		
Char.	Cat.	High	Low	abs(H-L)	High	Low	abs(H-L)
rns9m	Risk	1.33	0.62	0.71	1.53	1.01	0.52
rnk9m	Risk	1.03	1.16	0.14	1.41	1.32	0.08
rns12m	Risk	1.33	0.61	0.72	1.55	0.96	0.59
rnk12m	Risk	1.05	1.14	0.09	1.33	1.32	0.01
mfvu	Risk	1.27	0.56	0.72	1.51	1.01	0.51
mfvd	Risk	1.26	0.69	0.57	1.42	1.15	0.27
ivrv	Risk	0.92	1.49	0.57	1.30	1.76	0.47
ivrv_ratio	Risk	1.44	1.00	0.45	1.71	1.37	0.33
vol	Liquidity	0.70	1.34	0.64	1.00	1.42	0.42
dvol	Liquidity	0.68	1.35	0.67	0.93	1.42	0.49
VolMkt	Liquidity	1.26	0.97	0.29	1.45	1.25	0.20
Illiquidity	Liquidity	1.42	0.70	0.72	1.69	0.87	0.82
zerotrade	Liquidity	1.27	0.84	0.43	1.47	1.09	0.39
DolVol	Liquidity	1.52	0.69	0.83	1.68	0.91	0.77

Table 7: Alpha Portfolios with Interactions of Characteristics

This table shows the summary statistics of the alpha portfolios when adding the interaction of characteristics. We interact rns9m, rnk9m, ivrv, and mfvu with vol, illiquidity, size, and so to allow the nonlinearity. At each month, we estimate the mapping from options alpha to characteristics, $\boldsymbol{\theta}$, by Projected Principal Component Analysis (PPCA) introduced in section 3, and compute the estimated alphas as the product of $\widehat{\boldsymbol{\theta}}$ and one-month ahead characteristics. The estimation window is 12 months. The weights of individual options in the alpha portfolios are proportional to the estimated options alpha. K ranges from one to 10, and it is the number of latent factors of PPCA. Mean, SD, and Sharpe shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period.

Panel A: Call Option Alpha Portfolio

K	Mean	SD	Sharpe	Skew	Kurt	Min	Max
1	0.23	0.14	1.60	1.06	5.54	-0.12	0.27
2	0.25	0.18	1.40	0.25	1.91	-0.16	0.23
3	0.30	0.18	1.69	1.02	4.99	-0.15	0.33
4	0.31	0.18	1.68	0.70	3.73	-0.20	0.29
5	0.30	0.18	1.63	0.76	3.47	-0.16	0.31
6	0.28	0.19	1.50	0.83	4.83	-0.16	0.35
7	0.28	0.19	1.42	0.60	3.87	-0.20	0.32
8	0.28	0.21	1.36	0.58	3.50	-0.20	0.31
9	0.28	0.21	1.38	0.43	3.17	-0.20	0.31
10	0.28	0.20	1.37	0.57	3.57	-0.20	0.31

Panel B: Put Option Alpha Portfolio

K	Mean	SD	Sharpe	Skew	Kurt	Min	Max
1	0.27	0.15	1.82	1.58	7.56	-0.11	0.29
2	0.29	0.18	1.62	0.80	3.26	-0.16	0.26
3	0.33	0.17	1.88	0.96	3.84	-0.16	0.26
4	0.32	0.18	1.76	1.13	4.05	-0.12	0.30
5	0.30	0.18	1.69	0.62	2.05	-0.12	0.22
6	0.30	0.18	1.68	0.72	2.47	-0.12	0.23
7	0.29	0.19	1.58	1.07	3.47	-0.13	0.26
8	0.30	0.19	1.57	1.10	3.68	-0.11	0.29
9	0.30	0.20	1.54	1.23	4.36	-0.11	0.32
10	0.27	0.21	1.31	1.32	5.29	-0.12	0.36

Table 8: Contributions of Characteristics with Interaction Terms

This table reports the top 20 characteristics contributing most to options alpha and factor loadings. The characteristics contain 16 interaction terms constructed in table 7. The contributions of characteristics on options alpha are measured by $\widehat{\boldsymbol{\theta}}$, the mapping from the options alpha to characteristics. The contributions of characteristics on options factor loadings are measured by $\widehat{\boldsymbol{\beta}}$, the mapping from the options beta to characteristics. PPCA estimate $\mathbf{X}\boldsymbol{\beta}$ as a whole, and we regress it on characteristics $\mathbf{X}$ to recover $\widehat{\boldsymbol{\beta}}$. The absolute values of betas of characteristics for different latent factors are summed to measure the total contribution. The contributions of characteristics to options alpha and factor loadings are scaled by the cross-section maximum, so the contributions in each period range from zero to one. According to Chen and Zimmermann (2020) and Bali et al. (2021), most characteristics are divided into risk and liquidity groups. Interaction terms are divided into the Interaction group. The rest is divided into the Others group. The definitions of the characteristics can be found in the Internet Appendix.

	Panel A: Characteristics Contributions (Call)						Panel B: Characteristics Contributions (Put)					
	Factor Loadings			Alpha			Factor Loadings			Alpha		
	Char.	Cat.	Contr.	Char.	Cat.	Contr.	Char.	Cat.	Contr.	Char.	Cat.	Contr.
1	vol	Liquidity	0.8542	vol	Liquidity	0.6969	vol	Liquidity	0.8252	vol	Liquidity	0.6687
2	vol_rnk9m	Interaction	0.7123	vol_rnk9m	Interaction	0.6044	vol_rnk9m	Interaction	0.7359	vol_rnk9m	Interaction	0.6041
3	Illiquidity	Liquidity	0.6624	Illiquidity	Liquidity	0.4802	Illiquidity	Liquidity	0.6502	Illiquidity	Liquidity	0.4988
4	Illiq_rnk9m	Interaction	0.5339	so	Others	0.4424	so	Others	0.5433	so	Others	0.4563
5	so	Others	0.5157	Illiq_rnk9m	Interaction	0.4375	Illiq_rnk9m	Interaction	0.5352	Illiq_rnk9m	Interaction	0.4416
6	so_rnk9m	Interaction	0.4743	so_rnk9m	Interaction	0.3946	so_rnk9m	Interaction	0.4796	so_rnk9m	Interaction	0.3726
7	vol_mfvu	Interaction	0.4283	vol_mfvu	Interaction	0.3261	Illiq_mfvu	Interaction	0.3519	rns9m	Risk	0.2642
8	Illiq_mfvu	Interaction	0.3513	Illiq_mfvu	Interaction	0.2677	vol_mfvu	Interaction	0.3299	ivrv	Risk	0.2631
9	rns9m	Risk	0.3068	ivrv	Risk	0.252	rns9m	Risk	0.3272	size_ivrv	Interaction	0.2615
10	ivrv	Risk	0.2879	size_ivrv	Interaction	0.2515	ivrv	Risk	0.3172	dvol	Liquidity	0.2553
11	size_ivrv	Interaction	0.2829	dvol	Liquidity	0.2371	size_ivrv	Interaction	0.2912	Illiq_mfvu	Interaction	0.2545
12	size_rns9m	Interaction	0.2705	rns9m	Risk	0.2367	size_rns9m	Interaction	0.289	vol_mfvu	Interaction	0.2515
13	mfvu	Risk	0.2471	mfvu	Risk	0.2317	mfvu	Risk	0.2572	mfvu	Risk	0.2402
14	dvol	Liquidity	0.2125	size_rns9m	Interaction	0.2239	dvol	Liquidity	0.2464	size_rns9m	Interaction	0.2234
15	VolMkt	Liquidity	0.2113	size	Others	0.1849	so_mfvu	Interaction	0.2218	rnk9m	Risk	0.1879
16	so_mfvu	Interaction	0.2075	so_mfvu	Interaction	0.1745	VolMkt	Liquidity	0.2175	size	Others	0.181
17	rnk9m	Risk	0.197	size_rnk9m	Interaction	0.1706	size_mfvu	Interaction	0.2084	size_mfvu	Interaction	0.1798
18	size_mfvu	Interaction	0.1851	size_mfvu	Interaction	0.1695	size	Others	0.2064	so_mfvu	Interaction	0.179
19	size	Others	0.1848	rnk9m	Risk	0.1637	rnk9m	Risk	0.2021	rns12m	Risk	0.1624
20	ivrv_ratio	Risk	0.1765	rns12m	Risk	0.1634	ivrv_ratio	Risk	0.1854	AM	Others	0.1578

Table 9: Alpha Portfolios with Different Estimation Windows

This table shows the summary statistics of the alpha portfolios with different estimation windows. The estimation windows are 12, 15, 18, 21, and 24 months. At each month, we estimate the mapping from options alpha to characteristics, $\boldsymbol{\theta}$, by Projected Principal Component Analysis (PPCA) introduced in section 3, and compute the estimated alphas as the product of $\widehat{\boldsymbol{\theta}}$ and one-month ahead characteristics. The weights of individual options in the alpha portfolios are proportional to the estimated options alpha. We set the number of latent factors K equal to three. We report the summary statistics of alpha portfolios in the entire and post-2004 samples. Mean, SD, and Sharpe shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period.

Panel A: Call Alpha Portfolios								
K	Window	Mean	SD	Sharpe	Skew	Kurt	Min	Max
3	12	0.28	0.17	1.62	1.35	6.92	-0.14	0.34
3	15	0.25	0.18	1.40	1.97	16.39	-0.18	0.45
3	18	0.26	0.16	1.61	1.22	6.31	-0.15	0.31
3	21	0.19	0.16	1.21	1.03	5.59	-0.13	0.32
3	24	0.22	0.18	1.24	0.60	2.40	-0.12	0.28
Panel B: Put Alpha Portfolios								
K	Window	Mean	SD	Sharpe	Skew	Kurt	Min	Max
3	12	0.32	0.17	1.86	1.11	4.73	-0.17	0.26
3	15	0.29	0.18	1.57	1.48	8.92	-0.13	0.40
3	18	0.27	0.17	1.60	1.21	5.18	-0.13	0.30
3	21	0.26	0.17	1.60	0.70	2.84	-0.12	0.28
3	24	0.24	0.18	1.32	0.61	1.67	-0.14	0.28

Table 10: Subsamples and Alpha Portfolios

This table shows the summary statistics of the alpha portfolios with different estimation windows. The estimation windows are 12, 15, 18, 21, and 24 months. At each month, we estimate the mapping from options alpha to characteristics, $\boldsymbol{\theta}$, by Projected Principal Component Analysis (PPCA) introduced in section 3, and compute the estimated alphas as the product of $\widehat{\boldsymbol{\theta}}$ and one-month ahead characteristics. The weights of individual options in the alpha portfolios are proportional to the estimated options alpha. We set the number of latent factors K equal to three. We report the summary statistics of alpha portfolios in the entire and post-2004 samples. Mean, SD, and Sharpe shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period.

Panel A: Call Alpha Portfolios (2004-2019)								
K	Window	Mean	SD	Sharpe	Skew	Kurt	Min	Max
3	12	0.20	0.13	1.49	0.36	1.14	-0.07	0.18
3	15	0.20	0.14	1.40	-0.27	2.66	-0.18	0.16
3	18	0.23	0.13	1.70	0.30	0.63	-0.08	0.16
3	21	0.16	0.14	1.16	0.16	0.13	-0.08	0.14
3	24	0.16	0.15	1.08	0.08	0.49	-0.11	0.16
Panel B: Put Alpha Portfolios (2004-2019)								
K	Window	Mean	SD	Sharpe	Skew	Kurt	Min	Max
3	12	0.23	0.14	1.62	0.00	2.28	-0.17	0.16
3	15	0.23	0.15	1.51	0.14	0.99	-0.12	0.15
3	18	0.23	0.14	1.63	0.25	0.65	-0.10	0.14
3	21	0.25	0.15	1.69	0.24	0.20	-0.12	0.14
3	24	0.20	0.16	1.20	0.32	0.20	-0.10	0.16

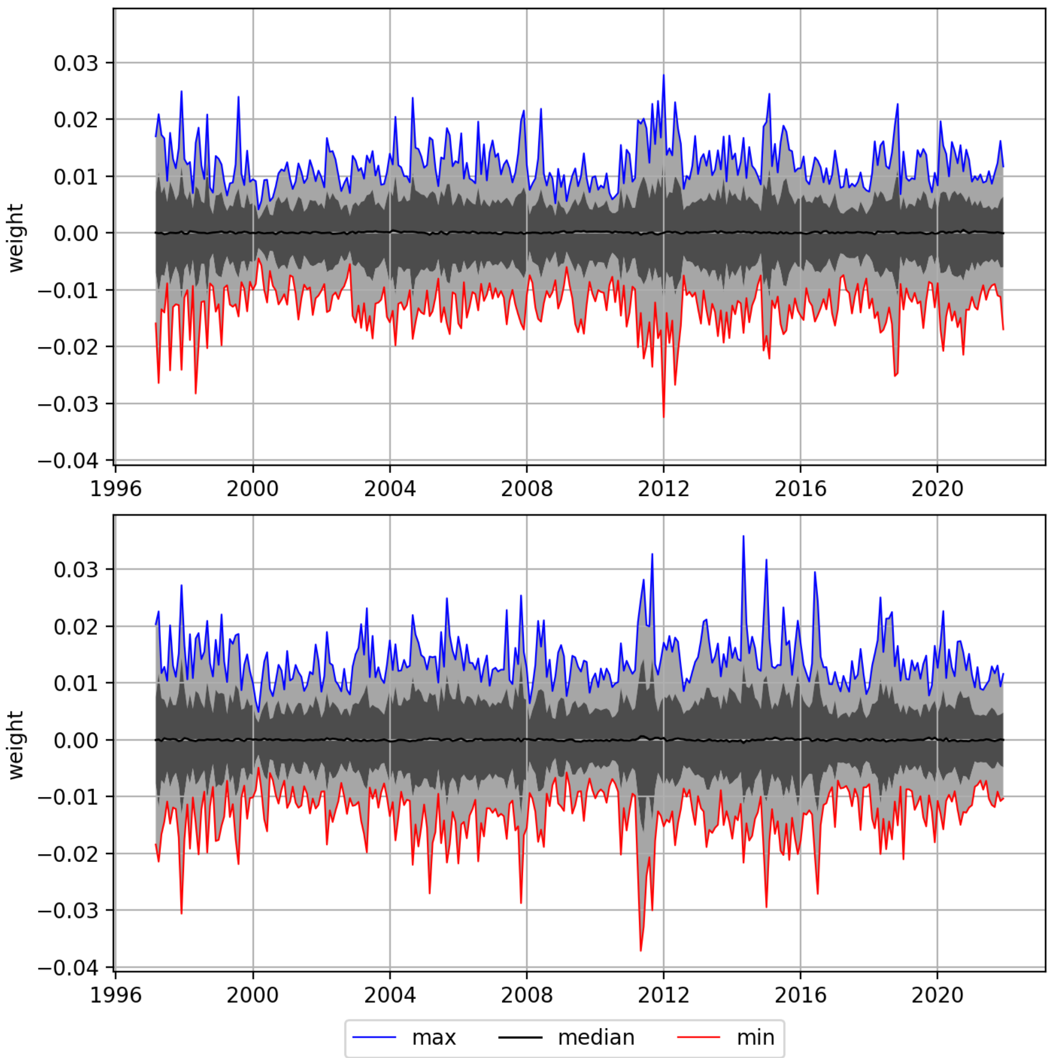


Figure 1: Distributions of Alpha Portfolio Weights

These two figures show the distribution of weights of individual options in the alpha portfolios. The upper (lower) panel shows weights for call (put) options.The alpha portfolios are constructed as introduced in table 2. The number of latent factors is three, and the estimation window is twelve months. The blue and red lines represent the maximum and minimum weights. The grey-shadowed areas display the weights ranked between maximum and 95%, and the weights ranked between minimum and 5%, respectively. The dark-shadowed area represents weights ranked between 5% and 95%. The black line in the middle of the figure shows the median weights of individual options.

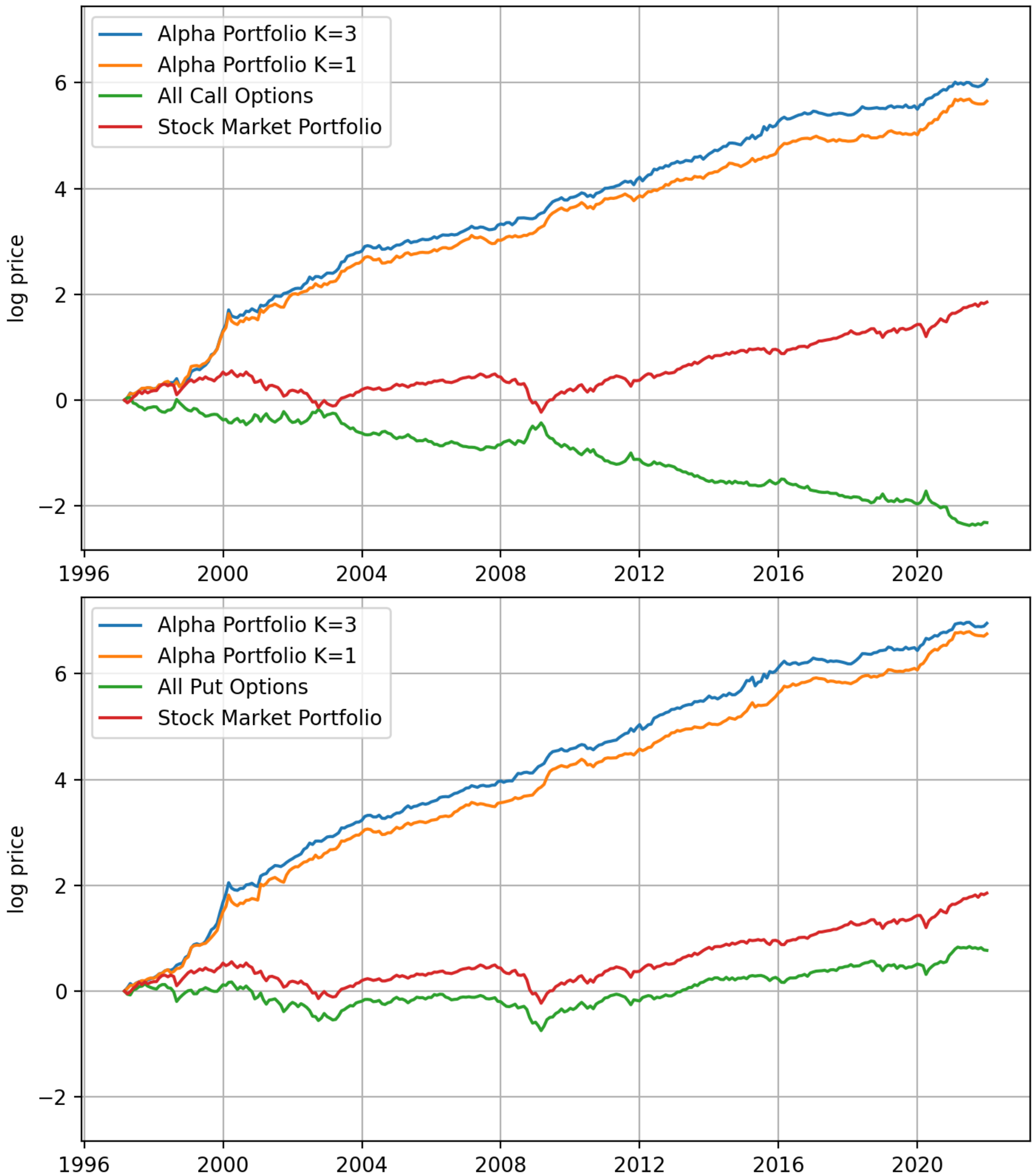


Figure 2: Cumulative Returns of Option Alpha and Market Portfolios

These two figures exhibit the natural logarithm of the cumulative returns (in excess of one) of two alpha and two market portfolios. The upper (lower) panel shows price path of portfolios of call (put) options. The blue and orange lines represent the log prices of the alpha portfolios with one and three latent factors, respectively. The estimation window is twelve months. The green and red lines are log prices of call (put) options and stock market portfolios, respectively. The call (put) options market portfolio is the equal-weighted average return of all one-month at-the-money call (put) options. The four portfolios are adjusted to have the same standard deviation as the stock market portfolio.

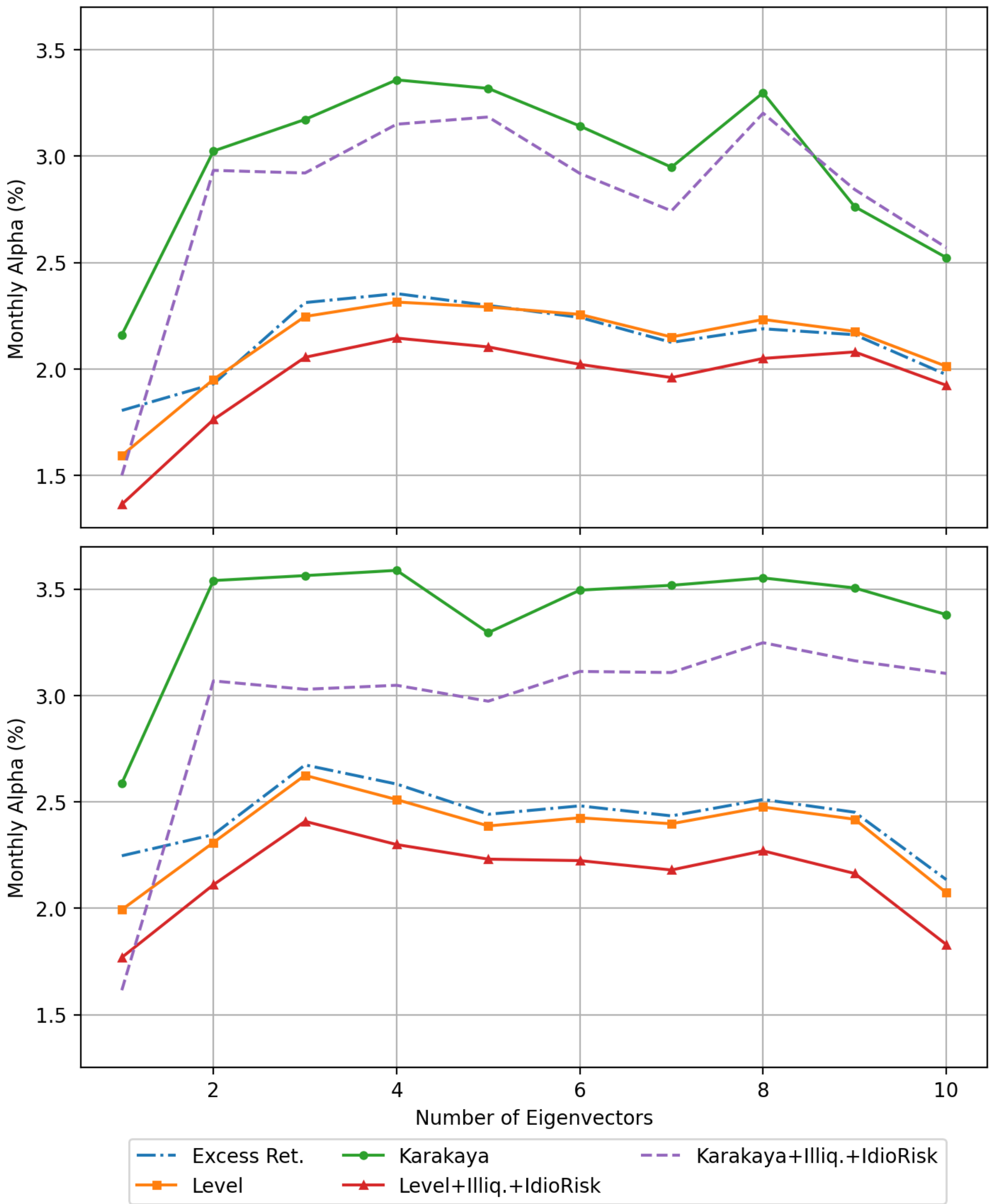


Figure 3: Alphas of Options Market Risk Factors with Varying K

This figure reveals the explanatory power of options market risk factors on the returns of alpha portfolios with the varying number of latent factors K. The upper (lower) panel is for call (put) options. The estimation window of the alpha portfolios is twelve months. We regress the returns of the alpha portfolios on options market risk factors and report alphas. Excess return is the monthly return of the alpha portfolios minus the risk-free rate. Level is the market portfolio that contains all call and put options. Karakaya represents the level, maturity, and value factors constructed following Karakaya (2013). Illiquidity and Idiosyncratic Risk are two factors constructed on Amihud's illiquidity ratio and idiosyncratic volatility of the underlying, respectively. Factors are constructed as the difference of extreme deciles of one-month at-the-money options.

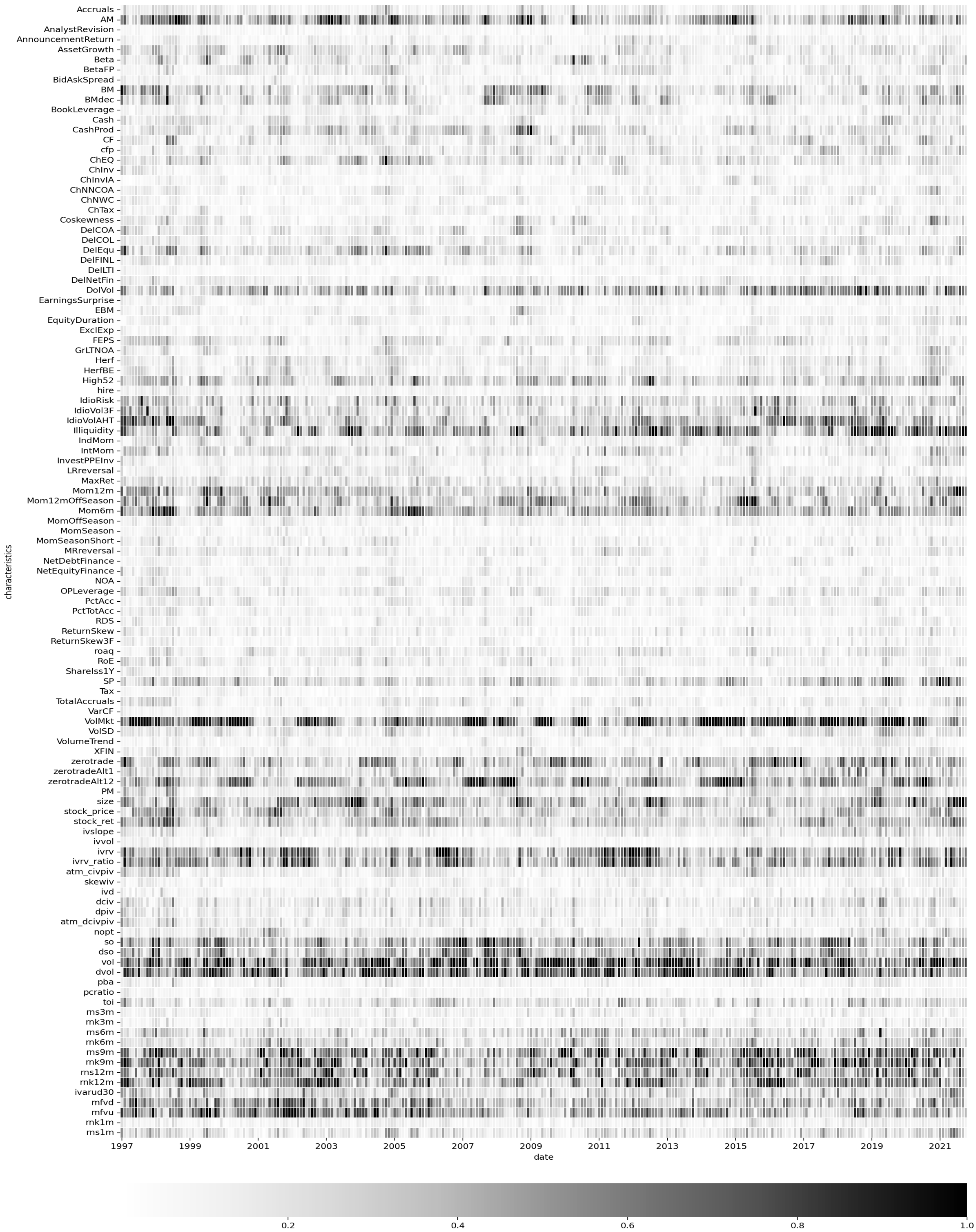


Figure 4: Contributions of Characteristics to Call Options Factor Loadings

The heatmap displays the contributions of the firm characteristics on options factor loadings. The number of latent factors is three, and the estimation window is twelve months. We estimate the options factor loadings by PPCA and regress the estimated factor loadings on characteristics to uncover the mapping from factor loadings to characteristics. The contribution of the firm characteristic is the summation of the magnitude of the mapping. The contributions are scaled by maximum contribution each month, so the range of contributions is zero to one. The heatmap shows the contributions of characteristics in each month. Darker (lighter) the mark, the closer the contribution to one (zero).

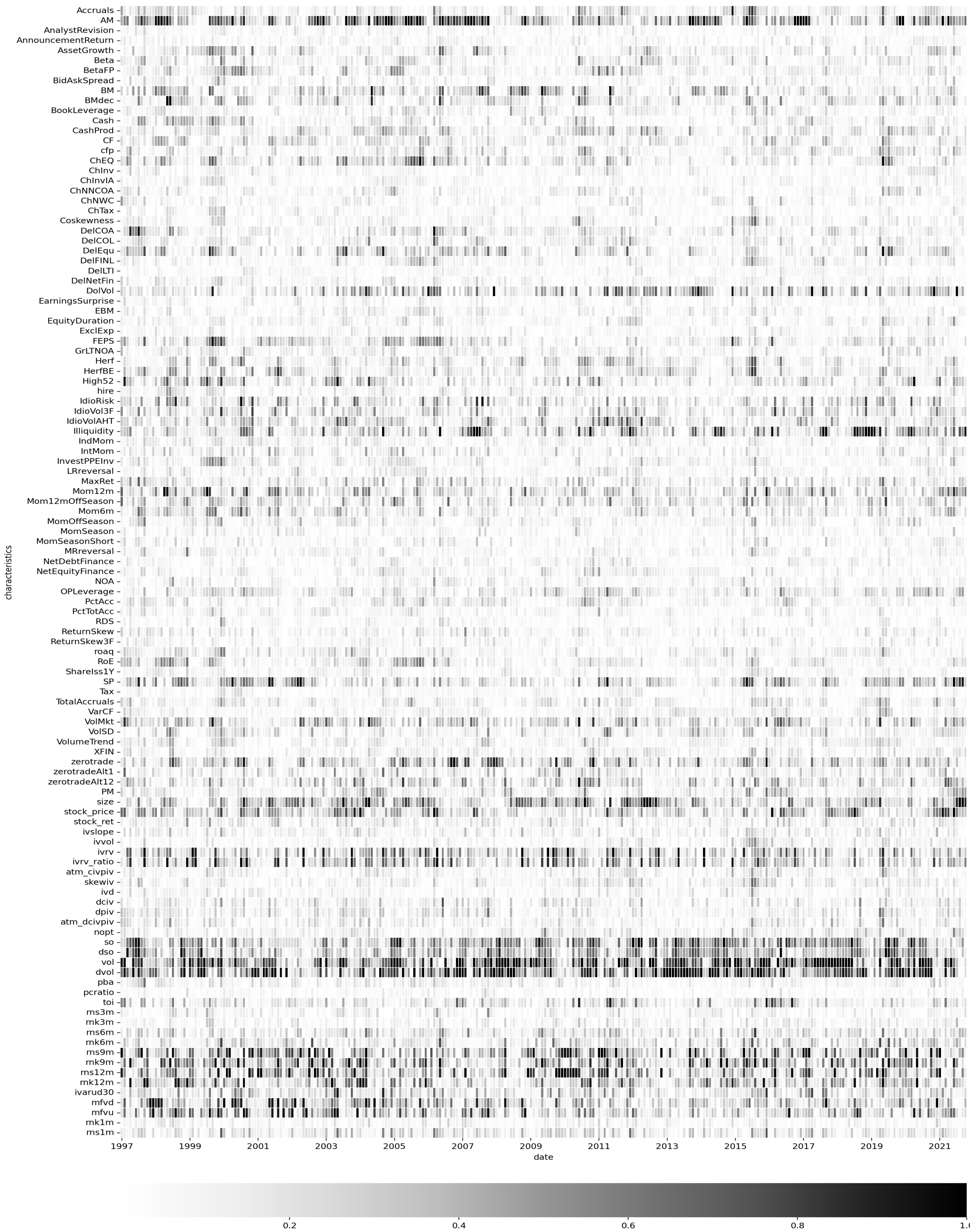


Figure 5: Contributions of Characteristics to Call Options Alphas

The heatmap displays the contributions of the firm characteristics on options alphas. The number of latent factors is three, and the estimation window is twelve months. The contributions are measured by $\boldsymbol{\theta}$, the mapping from options alphas to characteristics. The contributions are scaled by maximum contribution each month, so the range of contributions is zero to one. The heatmap shows the contributions of characteristics in each month. Darker (lighter) the mark, the closer the contribution to one (zero).

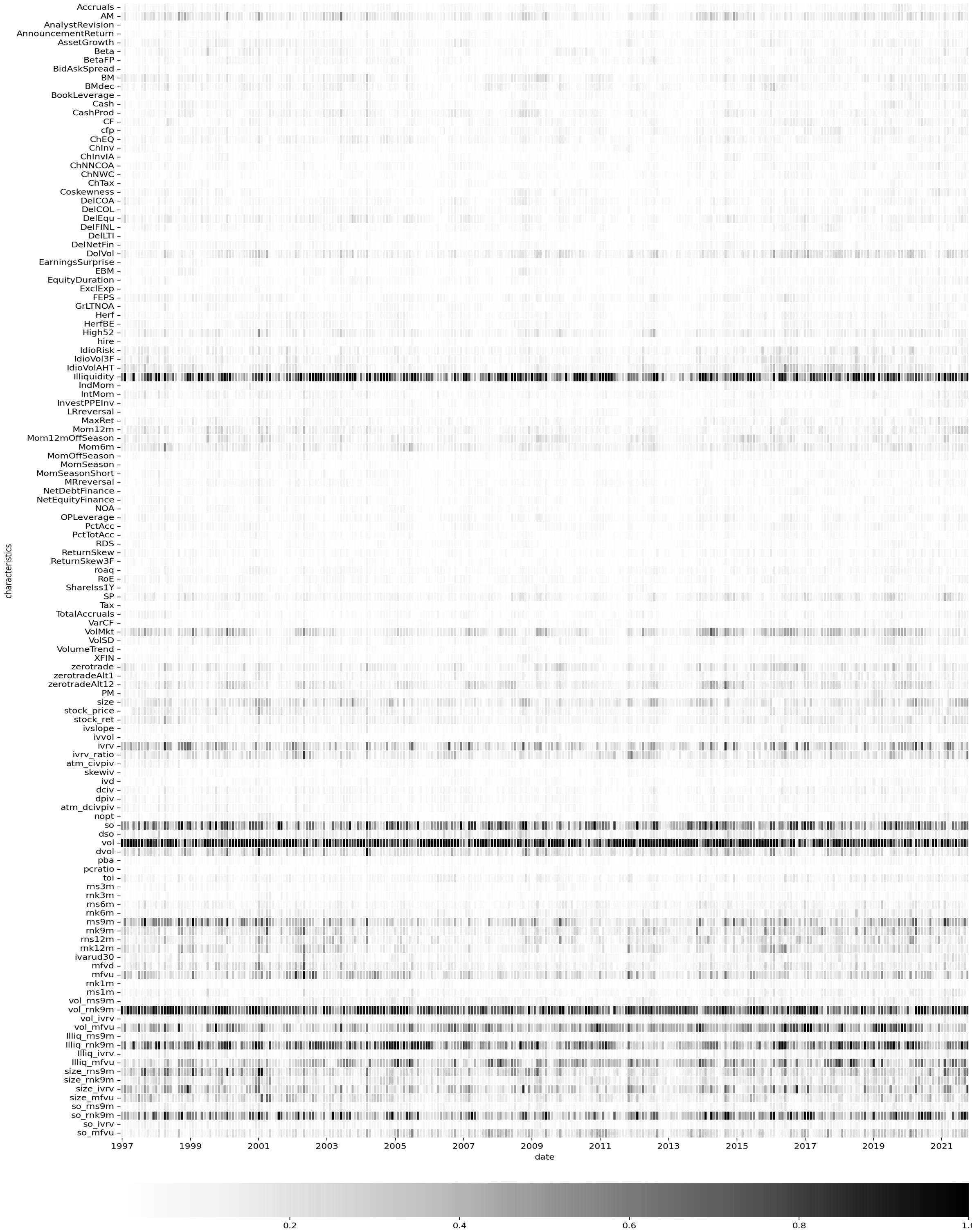


Figure 6: Contributions of Characteristics to Call Options Factor Loadings - Nonlinear Case

The heatmap displays the contributions of the firm characteristics on options factor loadings. The number of latent factors is three, and the estimation window is twelve months. We interact vol, illiquidity, size, and so with rns9m, rnk9m, ivrv, and mfvu. Contributions are calculated as in figure 4. Darker (lighter) the mark, the closer the contribution to one (zero).

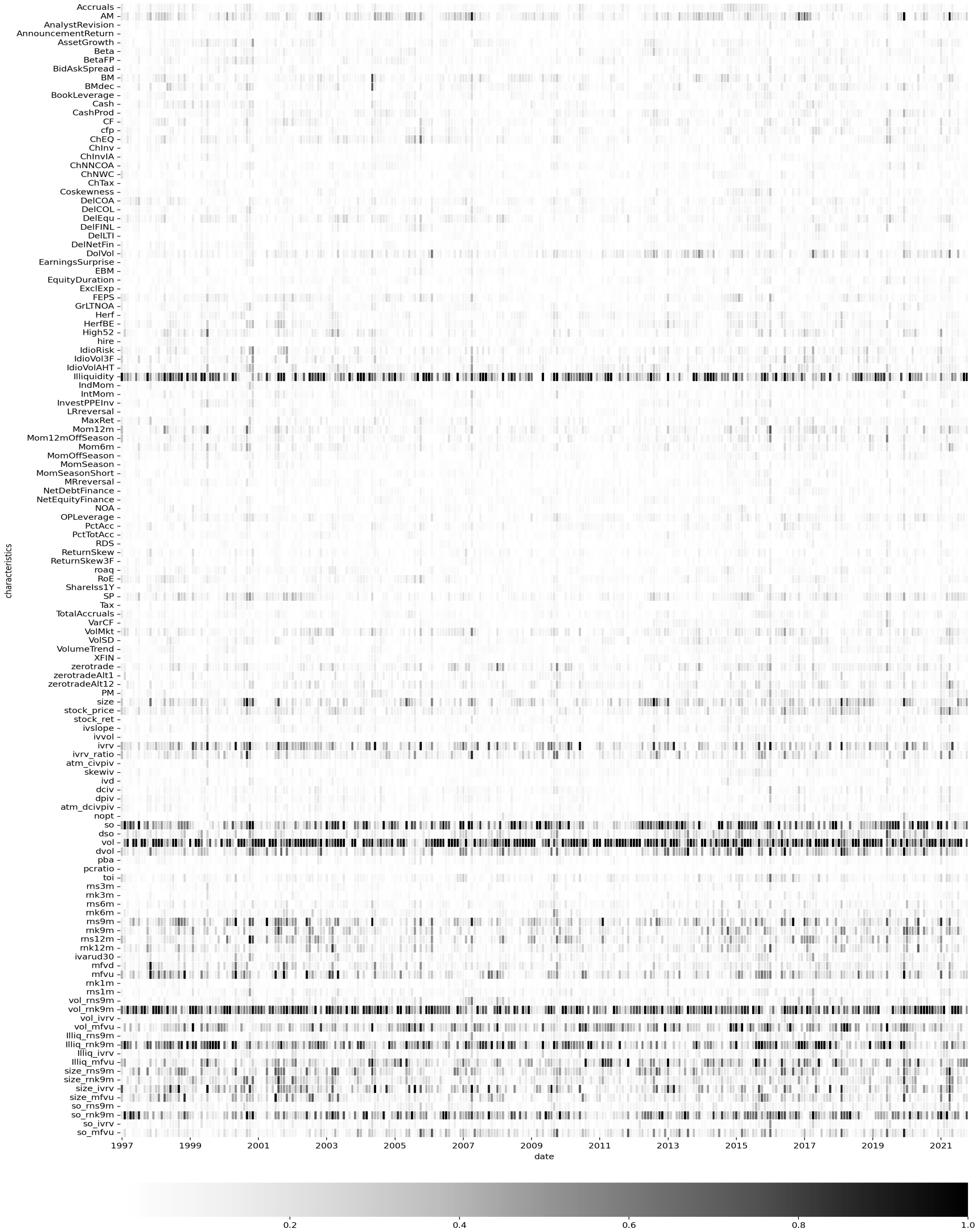


Figure 7: Contributions of Characteristics to Call Options Alphas - Nonlinear Case

The heatmap displays the contributions of the firm characteristics on options alphas. The number of latent factors is three, and the estimation window is twelve months. We interact vol, illiquidity, size, and so with rns9m, rnk9m, ivrv, and mfvu. Contributions are calculated as in figure 5. Darker (lighter) the mark, the closer the contribution to one (zero).

Internet Appendix
(not for publication)

Option Mispricing and Alpha Portfolios

Andras Fulop, Junye Li, and Mo Wang

I. Characteristics

Table A1 summarizes the firm and option characteristics used in the paper. The original authors of the characteristics and years of the paper published are displayed. The details of the construction of the characteristics can be found in Chen and Zimmermann (2022) and Bali, Beckmeyer, Moerke, and Weigert (2022). The characteristics are divided into 10 categories according to Bali, Beckmeyer, Moerke, and Weigert (2022). If the characteristic is not in Bali, Beckmeyer, Moerke, and Weigert (2022), we decide its category based on Chen and Zimmermann (2022) or the category of similar characteristics.

II. Additional Figures and Tables

Figure A1 presents how the alpha estimates (risk-adjusted returns) resulting from the different stock factor models vary with respect to the number of the latent factors in PPCA. Figure A2 presents how the alpha estimates resulting from the latent option factor models vary with respect to the number of the latent factors in PPCA.

Figures A3 and A4 present the cross-sectional and temporal contributions of characteristics to beta and alpha, respectively. The number of latent factors is three, and the estimation window is twelve months. The contribution of the firm characteristic is the summation of the magnitude of the mapping. The contributions are scaled by the maximum contribution each month, so the range of contributions is zero to one. The heatmap shows the contributions of characteristics in each month.

Figures A5 and A6 present the average ranks of characteristics in the long and short legs of the call and put alpha portfolios, respectively. The number of latent factors is three, and the estimation window is twelve months. The ranks of characteristics range from -0.5 to 0.5. The long and short legs contain options with positive and negative weights in the put portfolios, respectively. The light blue line represents the short leg, while the orange line represents the long leg.

Figures A7 and A8 present the cross-sectional and temporal contributions of characteristics to beta and alpha, respectively, when we take into account interaction terms of characteristics.

Table A2 reveals the explanatory power of option latent factors on the alpha portfolios' returns. We extract latent factors from options by PCA or RP-PCA (Risk Premium PCA). We regress the monthly returns of the alpha portfolios on the latent factors and report the alpha (in percentage) and factor loadings. PCA (RP-PCA) factors are constructed by applying PCA (RP-PCA) to at-the-money options. The λ in RP-PCA is set to 10.

III. Results from Single Equity Options Data

III.1. Overall Performance

Table A3 shows the performances of alpha portfolios built on single equity options. The estimation window is six months. We choose six months instead of twelve months because the underlying that satisfies the liquidity filter becomes insufficient when extending the estimation window. Although we normalize the in-sample annualized standard deviation of alpha portfolios to 20%, the standard deviations of the alpha portfolios are notably high, which may be due to the small number of options used. The annualized Sharpe ratios of the alpha portfolios for call options range from 1.66 to 2.23, and those for put options range from 1.26 to 1.79. Sharpe ratios are lower for put options than call options, which may step from the smaller sample size of put options. These Sharpe ratios are higher than those of volatility surface options. Annualized mean returns of alpha portfolios with three latent factors are 69% for call options and 62% for put options, which are approximately twice as large as those for volatility surface options. When $K = 3$, the kurtosis is 1.88 for call options and 4.26 for put options, indicating that the performances of the alpha portfolios are not driven by outliers. The ranges of portfolio returns also support this argument.

We investigate the impact of the length of estimation windows on the performance of alpha portfolios in the entire sample and post-2004 sample. Table A4 displays the results. Like the empirical pattern of the volatility surface options, the annualized Sharpe ratios decrease when prolonging the estimation window in the whole sample. The decrease is significant for call options: extending the estimation window from 6 months to 12 months leads to a decrease in the Sharpe ratio from 2.22 to 1.13. Sharpe ratios of put options are relatively low in general but still have a decreasing pattern when the length of the estimation window increases. Two issues may be responsible for the decreasing Sharpe ratios. The first one is the assumption of the constant characteristics discussed previously. When extending the length, assuming that the characteristics are unchanged within the estimation window becomes implausible. The second issue is the decreasing number of options used each month, as shown in the last columns of Table A4. PPCA requires sufficient options to get a consistent estimator. Given that the sample size of single equity options is limited, the loss of options caused by extending the estimation window hurt the Sharpe ratio more severely than that in volatility surface options. No matter the length of estimation windows, the put options available for PPCA are less than the call options, which is an important reason why the Sharpe ratios of put options are lower than that of call options.

At the beginning of the single equity options sample, the number of options available for PPCA is minimal, preventing the PPCA from consistently estimating the parameters. After dropping the alpha portfolios before 2004, the annualized Sharpe ratios of alpha portfolios with three latent factors and six-month estimation windows increase to 2.62 for single equity call options and 2.25 for single equity put options. The Sharpe ratios still decrease with the length of the estimation window, while the magnitudes of Sharpe ratios are broadly larger than those in the entire sample for call and put options.

III.2. Factor Analysis

We explore if the returns of alpha portfolios can be attributed to option factors documented in the literature. Option factors are constructed similarly as before, except we use single equity options instead of volatility surface options in construction. Table A5 displays the results. For the single equity call options, the coefficient of the value factor is highly significant. The alpha of the alpha portfolios decreases from 5.73% (t=9.58) to 3.82% (t=4.77) when introducing all five factors. No factor is significant in explaining the alpha portfolio returns of single euqity put options. We aggregate all the factors in the last columns, and the risk-adjusted return of the alpha portfolio for single equity call options is 3.82% (t=4.77) and that for single equity put options is 4.91% (t=4.30). Overall, option factor models cannot explain the performances of alpha portfolios on single equity options.

III.3. Mispricing, Systematic Risks, and Characteristics

We explore the association of characteristics with alpha and with factor loadings. We compute the contributions of the characteristics as in section 4.3. Table A6 shows the top 20 essential characteristics in their contribution to alpha and factor loadings. Characteristics that contribute to systematic risks also contribute to mispricing. This phenomenon is common in both single equity call and put options and is consistent with the findings in volatility surface options. We divide the most characteristics into a group related to options liquidity and a group linked to higher moments risks. The characteristics of the two groups are similar to those of volatility surface options. The ranks of the liquidity-related characteristics decline compared with their ranks in the results of volatility surface options, indicating that risk-related characteristics are more important in measuring systematic risks and mispricing among single equity options. The possible reason may be that the single equity options we used are highly liquid and have less variation in liquidity, so the liquidity characteristics are not as important as in the volatility surface options.

We also construct alpha portfolios using subsamples of single equity options to explore the impact of the characteristics. Table A7 in the appendix shows the Sharpe ratios of the high and low alpha portfolios on single equity options. The high and low alpha portfolios are constructed the same as those in Table 6. The results are similar to those for call and put options, except that the Sharpe ratio differences between risk-related characteristics subsamples are larger on average than those between liquidity-related subsamples. The decreases in the importance of liquidity-related characteristics are as expected since the single equity options we used are very liquid.

III.4. Liquidity and Transaction Cost

To explore the impact of option liquidity on alpha portfolio performance, we estimate the portfolio weights using all options, and construct portfolios using only liquid options. We define liquid options as options with positive trading volume, positive open interest and the bid-ask-spread below its 75% quantile (0.15 for call and 0.13 for put). The alpha portfolios are summarized in table A8. With only liquid options, the performances of alpha portfolios are even higher than those in table A3. When the number of latent factor equals 3, the annualized Sharpe ratio of single equity call (put) alpha portfolio increase from 2.22 (1.68) to 2.44 (2.45). Table A8 suggests that mispricing we document in the options market is not restricted to illiquid options.

To shed light on the transaction cost, we introduce the effective bid-ask spread. We consider effective bid-ask spread of 0, 10%, 20% and 40% of quoted bid-ask spread. The effective bid (ask) is the average of quoted bid and ask price minus (plus) one-half of the effective bid-ask spread. Given the single equity options in our sample are short-lived, to be more realistic, we hold the options to maturity when we construct the alpha portfolios. The portfolio weights are estimated by one-month single equity option returns as before. Table A9 displays the results. As we expected, the annualized Sharpe ratios of single equity options decrease with transaction cost. We find that when effective spread is 40%-60% of quoited bid-ask spread, the annualized Sharpe ratios of single equity option alpha portfolios are close to zero.

Table A1: Firm and Option Characteristics

Char	Ref	Description	Cat
Accruals	Sloan (1996)	Accruals	Accruals
AM	Fama and French (1992)	Total assets to market	Value
AnalystRevision	Hawkins, Chamberlin and Daniel (1984)	EPS forecast revision	Profitability
AnnouncementReturn	Chan, Jegadeesh and Lakonishok (1996)	Earnings announcement return	Profitability
AssetGrowth	Cooper, Gulen and Schill (2008)	Asset growth	Investment
Beta	Fama and MacBeth (1973)	CAPM beta	Risk
BetaFP	Frazzini and Pedersen (2014)	Frazzini-Pedersen Beta	Risk
BidAskSpread	Amihud and Mendelsohn (1986)	Bid-ask spread	Liquidity
BM	Rosenberg, Reid, and Lanstein (1985)	Book to market using most recent ME	Value
BMdec	Fama and French (1992)	Book to market using December ME	Value
BookLeverage	Fama and French (1992)	Book leverage (annual)	Quality
Cash	Palazzo (2012)	Cash to assets	Quality
CashProd	Chandrashekar and Rao (2009)	Cash Productivity	Profitability
CF	Lakonishok, Shleifer, Vishny (1994)	Cash flow to market	Quality
cfp	Desai, Rajgopal, Venkatachalam (2004)	Operating Cash flows to price	Quality
ChEQ	Lockwood and Prombutr (2010)	Growth in book equity	Investment
ChInv	Thomas and Zhang (2002)	Inventory Growth	Investment
ChInvIA	Abarbanell and Bushee (1998)	Change in capital inv (ind adj)	Investment
ChNNCOA	Soliman (2008)	Change in Net Noncurrent Op Assets	Investment
ChNWC	Soliman (2008)	Change in Net Working Capital	Investment
ChTax	Thomas and Zhang (2011)	Change in Taxes	Quality
Coskewness	Harvey and Siddique (2000)	Coskewness	Risk
DelCOA	Richardson et al. (2005)	Change in current operating assets	Investment
DelCOL	Richardson et al. (2005)	Change in current operating liabilities	Quality
DelEqu	Richardson et al. (2005)	Change in equity to assets	Investment
DelFINL	Richardson et al. (2005)	Change in financial liabilities	QUality
DelLTI	Richardson et al. (2005)	Change in long-term investment	Investment
DelNetFin	Richardson et al. (2005)	Change in net financial assets	Investment
DolVol	Brennan, Chordia, Subra (1998)	Past trading volume	Liquidity
EarningsSurprise	Foster, Olsen and Shevlin (1984)	Earnings Surprise	Profitability
EBM	Penman, Richardson and Tuna (2007)	Enterprise component of BM	Value
EquityDuration	Dechow, Sloan and Soliman (2004)	Equity Duration	Value
ExclExp	Doyle, Lundholm and Soliman (2003)	Excluded Expenses	Quality
FEPS	Cen, Wei, and Zhang (2006)	Analyst earnings per share	Profitability
GrLTNOA	Fairfield, Whisenant and Yohn (2003)	Growth in long term operating assets	Investment

Char	Ref	Description	Cat
Herf	Hou and Robinson (2006)	Industry concentration (sales)	Quality
HerfBE	Hou and Robinson (2006)	Industry concentration (equity)	Quality
High52	George and Hwang (2004)	52 week high	Past Prices
hire	Bazdresch, Belo and Lin (2014)	Employment growth	Profitability
IdioRisk	Ang et al. (2006)	Idiosyncratic risk	Risk
IdioVol3F	Ang et al. (2006)	Idiosyncratic risk (3 factor)	RIsk
IdioVolAHT	Ali, Hwang, and Trombley (2003)	Idiosyncratic risk (AHT)	RIsk
Illiquidity	Amihud (2002)	Amihud's illiquidity	Liquidity
IndMom	Grinblatt and Moskowitz (1999)	Industry Momentum	Past Prices
IntMom	Novy-Marx (2012)	Intermediate Momentum	Past Prices
InvestPPEInv	Lyandres, Sun and Zhang (2008)	change in ppe and inv/assets	Investment
LRreversal	De Bondt and Thaler (1985)	Long-run reversal	Past Prices
MaxRet	Bali, Cakici, and Whitelaw (2010)	Maximum return over month	Past Prices
Mom12m	Jegadeesh and Titman (1993)	Momentum (12 month)	Past Prices
Mom12mOffSeason	Heston and Sadka (2008)	Momentum without the seasonal part	Past Prices
Mom6m	Jegadeesh and Titman (1993)	Momentum (6 month)	Past Prices
MomOffSeason	Heston and Sadka (2008)	Off season long-term reversal	Past Prices
MomSeason	Heston and Sadka (2008)	Return seasonality years 2 to 5	Past Prices
MomSeasonShort	Heston and Sadka (2008)	Return seasonality last year	Past Prices
MRreversal	De Bondt and Thaler (1985)	Medium-run reversal	Past Prices
NetDebtFinance	Bradshaw, Richardson, Sloan (2006)	Net debt financing	Quality
NetEquityFinance	Bradshaw, Richardson, Sloan (2006)	Net equity financing	Quality
NOA	Hirshleifer et al. (2004)	Net Operating Assets	Quality
OPLeverage	Novy-Marx (2010)	Operating leverage	Quality
PctAcc	Hafzalla, Lundholm, Van Winkle (2011)	Percent Operating Accruals	Accruals
PctTotAcc	Hafzalla, Lundholm, Van Winkle (2011)	Percent Total Accruals	Accruals
PM	Soliman (2008)	Profit Margin	Profitability
RDS	Landsman et al. (2011)	Real dirty surplus	Quality
ReturnSkew	Bali, Engle and Murray (2015)	Return skewness	Risk
ReturnSkew3F	Bali, Engle and Murray (2015)	Idiosyncratic skewness (3F model)	Risk
roaq	Balakrishnan, Bartov and Faurel (2010)	Return on assets (qtrly)	Profitability
RoE	Haugen and Baker (1996)	net income / book equity	Profitability
ShareIss1Y	Pontiff and Woodgate (2008)	Share issuance (1 year)	Investment
SP	Barbee, Mukherji and Raines (1996)	Sales-to-price	Value
Tax	Lev and Nissim (2004)	Taxable income to income	Quality
TotalAccruals	Richardson et al. (2005)	Total accruals	Accruals
VarCF	Haugen and Baker (1996)	Cash-flow to price variance	Quality
VolMkt	Haugen and Baker (1996)	Volume to market equity	Liquidity
VolSD	Chordia, Subra, Anshuman (2001)	Volume Variance	Liquidity

Char	Ref	Description	Cat
VolumeTrend	Haugen and Baker (1996)	Volume Trend	Liquidity
XFIN	Bradshaw, Richardson, Sloan (2006)	Net external financing	Quality
zerotrade	Liu (2006)	Days with zero trades	Liquidity
zerotradeAlt1	Liu (2006)	Days with zero trades	Liquidity
zerotradeAlt12	Liu (2006)	Days with zero trades	Liquidity
atm_civpiv	Bali et al. (2021)	Implied volatility of atm call minus that of atm put	Informed Trading
atm_dcivpiv	An et al. (2014)	The change of atm call and put implied volatility spread	Informed Trading
dciv	An et al. (2014)	The change of the implied volatility of atm call	Informed Trading
dpiv	An et al. (2014)	The change of the implied volatility of atm put	Informed Trading
dso	Roll et al. (2010)	Stock-option dollar volume ratio	Informed Trading
dvol	Cao and Wei (2010)	Dollar trading volume	Liquidity
ivarud30	Huang and Li (2019)	Option-implied variance asymmetry	Risk
ivd	Schlag, Thimme, and Weber (2020)	Implied volatility duration	Risk
ivrv	Goyal and Saretto (2009)	Implied and historical volatility spread	Risk
ivrv_ratio	Bali et al. (2021)	Implied volatility divided by historical volatility	Risk
ivslope	Vasquez (2017)	Implied Volatility Slope	Risk
ivvol	Baltussen, van Bekkum, and der Grient (2018)	Volatility of atm implied volatility	Risk
mfvd	Huang and Li (2019)	Option-implied downside semivariance	Risk
mfvu	Huang and Li (2019)	Option-implied upside semivariance	Risk
nopt	Bali et al. (2021)	Number of traded options	Liquidity
pba	Cao and Wei (2010)	Proportional bid-ask spread	Liquidity
pcratio	Blau, Nguyen, and Whitby (2014)	Put-call ratio	Informed Trading
rnk12m	Borochin, Chang, and Wu (2020)	12-month risk-neutral kurtosis	Risk
rnk1m	Borochin, Chang, and Wu (2020)	1-month risk-neutral kurtosis	Risk
rnk3m	Borochin, Chang, and Wu (2020)	3-month risk-neutral kurtosis	Risk
rnk6m	Borochin, Chang, and Wu (2020)	6-month risk-neutral kurtosis	Risk
rnk9m	Borochin, Chang, and Wu (2020)	9-month risk-neutral kurtosis	Risk
rns12m	Borochin, Chang, and Wu (2020)	12-month risk-neutral skewness	Risk
rns1m	Borochin, Chang, and Wu (2020)	1-month risk-neutral skewness	Risk
rns3m	Borochin, Chang, and Wu (2020)	3-month risk-neutral skewness	Risk
rns6m	Borochin, Chang, and Wu (2020)	6-month risk-neutral skewness	Risk
rns9m	Borochin, Chang, and Wu (2020)	9-month risk-neutral skewness	Risk
size	Banz (1981)	Natural logrhythm of the market capitalization	Quality
skewiv	Xing, Zhang, and Zhao (2010)	Implied volatility of otm puts minus the that of atm calls	Informed Trading
so	Roll et al. (2010)	Stock-option volume ratio	Informed Trading
stock_price	Cao et al. (2021)	Natural logrhythm of the stock price	Informed Trading
stock_ret		Return of underlying stock	Frictions
toi	Bali et al. (2021)	Total open interest	Liquidity
vol	Cao and Wei (2010)	Trading Volume	Liquidity

Table A2: Explanatory Power of Options Latent Factors

This table reveals the explanatory power of option latent factors on the returns of alpha portfolios. We regress the monthly returns of the alpha portfolios on the option latent factors and report the alpha (in percentage) and factor loadings. PCA (RP-PCA) factors are constructed by applying PCA (RP-PCA) to at-the-money options. The λ in RP-PCA is set to 10. Newey-West (1987) adjusted t-statistics are reported in the parathesis. ***/**/* indicate the significant at 1%, 5% and 10% confidence level, respectively.

Panel A: Alpha Portfolio of Call Options				
	PCA 1F	PCA 3F	RP-PCA 1F	RP-PCA 3F
Alpha	2.20***	2.19***	2.20***	2.24***
	(7.80)	(8.23)	(7.80)	(7.71)
PCA1	0.00	0.00		
	(1.46)	(1.56)		
PCA2		0.02**		
		(2.37)		
PCA3		0.00		
		(-0.13)		
RP-PCA1			0.00	0.00
			(1.45)	(1.27)
RP-PCA2				0.02**
				(2.37)
RP-PCA3				-0.01
				(-0.64)
Adj. R^2	0.01	0.12	0.01	0.12
Num. obs.	298	298	298	298

Panel B: Alpha Portfolio of Put Options				
	PCA 1F	PCA 3F	RP-PCA 1F	RP-PCA 3F
Alpha	2.00***	1.99***	1.99***	2.09***
	(7.53)	(7.92)	(7.48)	(6.74)
PCA1	0.01***	0.01***		
	(4.05)	(4.86)		
PCA2		0.02***		
		(2.86)		
PCA3		-0.01		
		(-0.97)		
RP-PCA1			0.01***	0.01***
			(4.05)	(4.16)
RP-PCA2				0.02***
				(3.02)
RP-PCA3				-0.01
				(-1.04)
Adj. R^2	0.11	0.23	0.11	0.22
Num. obs.	298	298	298	298

Table A3: Alpha Portfolios of Single Equity Options

This table shows the summary statistics of the alpha portfolios built on single equity options. The alpha portfolios are constructed in the way introduced in table 2. The number of latent factors K ranges from one to ten. Mean, SD, and SR shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period.

Panel A: Alpha Portfolio of Single Equity Call Options

K	Mean	SD	SR	Skew	Kurt	Min	Max
1	0.66	0.30	2.21	1.00	2.73	-0.19	0.45
2	0.65	0.29	2.23	0.56	1.35	-0.18	0.37
3	0.69	0.31	2.22	0.98	1.88	-0.17	0.39
4	0.60	0.31	1.93	0.86	2.98	-0.25	0.46
5	0.48	0.29	1.66	0.57	3.72	-0.37	0.38
6	0.59	0.32	1.86	0.28	1.74	-0.36	0.38
7	0.58	0.32	1.79	0.48	1.34	-0.27	0.39
8	0.60	0.31	1.89	0.44	2.28	-0.34	0.43
9	0.59	0.32	1.85	0.42	1.41	-0.31	0.38
10	0.56	0.33	1.72	-0.07	3.38	-0.48	0.36

Panel B: Alpha Portfolio of Single Equity Put Options

K	Mean	SD	SR	Skew	Kurt	Min	Max
1	0.54	0.40	1.34	-2.86	34.71	-1.11	0.64
2	0.57	0.34	1.67	0.23	3.56	-0.36	0.47
3	0.62	0.37	1.68	0.51	4.26	-0.37	0.51
4	0.56	0.35	1.62	0.16	4.25	-0.43	0.45
5	0.51	0.38	1.33	-0.98	12.52	-0.77	0.54
6	0.56	0.34	1.63	-0.42	3.65	-0.44	0.42
7	0.57	0.32	1.79	-0.24	2.62	-0.41	0.36
8	0.46	0.37	1.26	-0.52	2.25	-0.40	0.35
9	0.52	0.34	1.52	-0.09	1.48	-0.29	0.36
10	0.53	0.32	1.64	-0.40	1.64	-0.32	0.29

Table A4: Alpha Portfolios of Single Equity Options with Different Estimation Windows

This table shows the summary statistics of the alpha portfolios of single equity options with different estimation windows. The estimation windows are 6, 9, 12, 15, and 18 months. We set the number of latent factors K equal to three. We report the summary statistics of alpha portfolios in the entire and post-2004 samples. Mean, SD, and SR shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period. Num. Obs. reports the monthly average number of options used to construct the alpha portfolios.

Panel A: Alpha Portfolio of Single Equity Call Options (Full Sample)

K	Window	Mean	SD	SR	Skew	Kurt	Min	Max	Num.Obs.
3	6	0.69	0.31	2.22	0.98	1.88	-0.17	0.39	309.44
3	9	0.41	0.24	1.70	-0.22	0.38	-0.16	0.24	255.78
3	12	0.32	0.28	1.13	0.15	0.99	-0.23	0.33	219.28
3	15	0.34	0.26	1.31	0.33	0.59	-0.16	0.32	192.76
3	18	0.37	0.25	1.46	0.41	1.84	-0.26	0.30	171.67

Panel B: Alpha Portfolio of Single Equity Call Options (2004-2021)

K	Window	Mean	SD	SR	Skew	Kurt	Min	Max	Num.Obs.
3	6	0.77	0.29	2.62	0.92	1.86	-0.11	0.38	365.14
3	9	0.43	0.24	1.74	-0.27	0.34	-0.16	0.24	307.36
3	12	0.36	0.29	1.23	0.15	0.99	-0.23	0.33	266.61
3	15	0.34	0.27	1.28	0.36	0.58	-0.16	0.32	236.14
3	18	0.32	0.26	1.25	0.41	2.10	-0.26	0.30	211.34

Panel C: Alpha Portfolio of Single Equity Put Options (Full Sample)

K	Window	Mean	SD	SR	Skew	Kurt	Min	Max	Num.Obs.
3	6	0.62	0.37	1.68	0.51	4.26	-0.37	0.51	233.03
3	9	0.47	0.30	1.56	-0.51	4.19	-0.48	0.30	192.98
3	12	0.39	0.29	1.35	0.31	1.18	-0.25	0.34	165.95
3	15	0.28	0.25	1.13	0.60	1.62	-0.17	0.35	146.33
3	18	0.31	0.25	1.21	0.19	1.45	-0.22	0.31	130.83

Panel D: Alpha Portfolio of Single Equity Put Options (2004-2021)

K	Window	Mean	SD	SR	Skew	Kurt	Min	Max	Num.Obs.
3	6	0.82	0.36	2.25	0.77	4.26	-0.32	0.51	292.55
3	9	0.57	0.31	1.83	-0.79	5.57	-0.48	0.30	244.12
3	12	0.47	0.30	1.59	0.25	1.03	-0.25	0.34	210.53
3	15	0.31	0.25	1.22	0.64	1.58	-0.14	0.35	185.62
3	18	0.32	0.26	1.26	0.04	1.56	-0.22	0.31	165.76

Table A5: Explanation Power of Options Market Risk Factors on Single Equity Options Alpha Portfolios

This table reveals the explanatory power of risk factors in the options market on the returns of the alpha portfolios of single equity options. We regress the monthly returns of the alpha portfolios on the common risk factors in the options market and report the alpha (in percentage) and factor loadings. Options market risk factors are introduced in table 3, but are constructed by single equity options. Newey-West (1987) adjusted t-statistics are reported in the parathesis. ***/**/* indicate the significant at 1%, 5% and 10% confidence level, respectively.

Panel A: Alpha Portfolio of Single Equity Call Options

	Excess Ret.	Level	Karakaya	Level + Illiq. + IdioRisk	Karakaya + Illiq. + IdioRisk
Alpha	5.73***	5.96***	3.99***	5.53***	3.82***
	(9.58)	(9.40)	(5.89)	(6.59)	(4.77)
Level		0.78*	0.49	0.90*	0.60
		(1.87)	(0.98)	(1.91)	(1.01)
Maturity			0.62		0.56
			(0.69)		(0.60)
Value			1.51***		1.51***
			(3.11)		(3.00)
Illiquidity				-0.73	-0.01
				(-1.39)	(-0.03)
IdioRisk				-0.16	0.18
				(-0.27)	0.34
Adj. R^2	0.00	0.02	0.07	0.02	0.06
Num. obs.	292	292	292	292	292

Panel B: Alpha Portfolio of Single Equity Put Options

	Excess Ret.	Level	Karakaya	Level + Illiq. + IdioRisk	Karakaya + Illiq. + IdioRisk
Alpha	5.20***	5.42***	5.07***	5.23***	4.91***
	(7.17)	(6.79)	(4.86)	(5.77)	(4.30)
Level		0.76	0.24	0.87	0.34
		(1.16)	(0.43)	(1.13)	(0.47)
Maturity			1.38		1.33
			(1.12)		(1.06)
Value			0.28		0.34
			(0.55)		(0.65)
Illiquidity				0.03	0.17
				(0.05)	(0.32)
IdioRisk				0.24	0.24
				(0.46)	(0.46)
Adj. R^2	0.00	0.01	0.01	0.00	0.00
Num. obs.	292	292	292	292	292

Table A6: Contribution of Characteristics in Single Equity Options

This table reports the top 20 characteristics contributing most to single equity options alpha and factor loadings. The contributions of characteristics on options alpha are measured by $\widehat{\boldsymbol{\theta}}$, the mapping from the options alpha to characteristics. The contributions of characteristics on options factor loadings are measured by $\widehat{\boldsymbol{\beta}}$, the mapping from the options beta to characteristics. PPCA estimate $\mathbf{X}\boldsymbol{\beta}$ as a whole, and we regress it on characteristics $\mathbf{X}$ to recover $\widehat{\boldsymbol{\beta}}$. The absolute values of betas of characteristics are summed to measure the total contribution. The contributions of characteristics to single equity options alpha and factor loadings are scaled by the cross-section maximum, so the contributions in each period range from zero to one. According to Chen and Zimmermann (2020) and Bali et al. (2021), most characteristics are divided into risk and liquidity groups. The rest is divided into the Others group. The definitions of the characteristics can be found in the Internet Appendix.

	Panel A: Characteristics Contribution (Single Equity Call)						Panel B: Characteristics Contribution (Single Equity Put)					
	Factor Loadings			Alpha			Factor Loadings			Alpha		
	Char.	Cat.	Contr.	Char.	Cat.	Contr.	Char.	Cat.	Contr.	Char.	Cat.	Contr.
1	mfvu	Risk	0.652	mfvd	Risk	0.4765	mfvd	Risk	0.6274	mfvu	Risk	0.4626
2	mfvd	Risk	0.6503	mfvu	Risk	0.4747	mfvu	Risk	0.6047	rnk9m	Risk	0.4545
3	rnk9m	Risk	0.5803	rnk9m	Risk	0.4357	rnk9m	Risk	0.5743	mfvd	Risk	0.4504
4	dvol	Liquidity	0.5649	vol	Liquidity	0.4147	rns9m	Risk	0.5515	rns9m	Risk	0.4344
5	vol	Liquidity	0.5508	rns9m	Risk	0.3999	rnk12m	Risk	0.4897	rnk12m	Risk	0.364
6	rns9m	Risk	0.5479	dvol	Liquidity	0.3962	vol	Liquidity	0.4772	rns12m	Risk	0.3569
7	rnk12m	Risk	0.5062	rnk12m	Risk	0.3628	dvol	Liquidity	0.4769	ivrv	Risk	0.3376
8	ivrv	Risk	0.5054	ivrv	Risk	0.3559	ivrv	Risk	0.4684	vol	Liquidity	0.3342
9	ivrv_ratio	Risk	0.5005	rns12m	Risk	0.3545	rns12m	Risk	0.4662	dvol	Liquidity	0.3324
10	rns12m	Risk	0.4755	ivrv_ratio	Risk	0.3507	ivrv_ratio	Risk	0.4402	ivrv_ratio	Risk	0.3322
11	VolMkt	Liquidity	0.4098	VolMkt	Liquidity	0.3235	rnk6m	Risk	0.3899	VolMkt	Liquidity	0.3224
12	zerotrade	Liquidity	0.3846	SP	Others	0.3093	VolMkt	Liquidity	0.3853	zerotrade	Liquidity	0.3029
13	AM	Others	0.3843	AM	Others	0.2984	AM	Others	0.3851	rnk6m	Risk	0.2954
14	dso	Others	0.3838	zerotrade	Liquidity	0.2937	rns6m	Risk	0.384	rns6m	Risk	0.2872
15	so	Others	0.3623	dso	Others	0.2895	zerotrade	Liquidity	0.3811	AM	Others	0.284
16	Illiquidity	Liquidity	0.3541	size	Others	0.2845	dso	Others	0.3611	Illiquidity	Liquidity	0.2757
17	IdioVol3F	Risk	0.3538	Illiquidity	Liquidity	0.2797	zerotradeAlt12	Liquidity	0.3602	zerotradeAlt12	Liquidity	0.2715
18	size	Others	0.3383	ivarud30	Risk	0.2774	Illiquidity	Liquidity	0.3574	SP	Others	0.2707
19	zerotradeAlt12	Liquidity	0.3327	rns6m	Risk	0.2759	IdioVol3F	Risk	0.3546	size	Others	0.2572
20	rnk6m	Risk	0.3318	zerotradeAlt12	Liquidity	0.2694	so	Others	0.3516	dso	Others	0.2565

Table A7: Alpha Portfolios of Single Equity Options Subsamples

This table shows the alpha portfolios constructed using the subsamples of single euqity options. The options alphas are estimated using all single equity options, while the high (low) alpha portfolios are constructed using option with the characteristic above (below) the median. We choose 3 latent factors and a 6-month estimation window. $abs(H-L)$ represents the absolute value of Sharpe ratio differences between high and low alpha portfolios.

		Single Equity Calls			Single Equity Puts		
Char.	Cat.	High	Low	abs(H-L)	High	Low	abs(H-L)
rns9m	Risk	1.94	0.77	1.17	1.81	0.23	1.59
rnk9m	Risk	0.54	2.11	1.56	-0.12	1.77	1.89
rns12m	Risk	2.05	0.69	1.35	1.69	0.32	1.36
rnk12m	Risk	0.55	2.09	1.54	-0.03	1.69	1.73
mfvu	Risk	2.03	0.35	1.68	1.73	-0.01	1.73
mfvd	Risk	2.15	0.24	1.90	1.85	-0.28	2.13
ivrv	Risk	1.10	1.87	0.77	0.87	1.38	0.51
ivrv_ratio	Risk	1.81	1.20	0.61	1.34	1.04	0.29
vol	Liquidity	2.04	1.05	0.98	1.78	0.47	1.31
dvol	Liquidity	2.00	1.08	0.92	1.83	0.46	1.37
VolMkt	Liquidity	1.19	2.01	0.83	0.53	1.55	1.02
Illiquidity	Liquidity	2.01	1.09	0.92	1.69	0.64	1.05
zerotrade	Liquidity	0.95	2.00	1.04	0.63	1.58	0.95
DolVol	Liquidity	1.85	1.28	0.57	1.52	0.86	0.67

Table A8: Alpha Portfolios of Liquid Single Equity Options

This table reveals the performance of alpha portfolios constructed by liquid single equity options. We relax the restriction that options should have positive trading volume, and estimate the options alphas using all options. We then construct alpha portfolios only using liquid options. Liquid call (put) options are defined as options with positive trading volume, positive open interest and option bid-ask spread less than 0.15 (0.13). The number of latent factors K ranges from one to ten. Mean, SD, and SR shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period.

Panel A: Alpha Portdolios of Single Equity Call Options							
K	Mean	SD	SR	Skew	Kurt	Min	Max
1	0.60	0.22	2.79	1.33	6.05	-0.13	0.45
2	0.60	0.23	2.66	1.30	4.51	-0.13	0.38
3	0.60	0.25	2.44	2.29	12.95	-0.12	0.60
4	0.54	0.23	2.40	1.83	7.52	-0.10	0.45
5	0.44	0.20	2.20	1.26	3.22	-0.12	0.31
6	0.51	0.21	2.40	1.82	9.32	-0.13	0.47
7	0.51	0.21	2.45	1.55	6.66	-0.12	0.41
8	0.50	0.21	2.38	1.69	8.09	-0.13	0.43
9	0.51	0.21	2.45	1.27	4.88	-0.14	0.36
10	0.51	0.22	2.35	1.93	10.10	-0.13	0.48
Panel B: Alpha Portdolios of Single Equity Put Options							
K	Mean	SD	SR	Skew	Kurt	Min	Max
1	0.50	0.20	2.55	0.92	1.86	-0.11	0.28
2	0.50	0.20	2.42	0.85	2.77	-0.20	0.31
3	0.48	0.20	2.45	1.21	2.39	-0.08	0.27
4	0.46	0.20	2.31	1.64	6.01	-0.08	0.40
5	0.32	0.17	1.91	1.36	5.12	-0.12	0.30
6	0.40	0.19	2.13	0.53	2.38	-0.18	0.21
7	0.40	0.19	2.15	0.54	2.47	-0.18	0.22
8	0.40	0.18	2.17	0.57	2.07	-0.16	0.21
9	0.39	0.18	2.15	0.57	2.23	-0.15	0.23
10	0.40	0.18	2.27	0.52	1.99	-0.16	0.23

Table A9: Alpha Portfolios of Liquid Single Equity Options

This table shows the performance of alpha portfolios when introducing the transaction cost. We introduce transaction cost by considering effective bid-ask spread of 0, 10%,20% and 40% of quoted bid-ask spread. The option returns used to estimate the alpha portfolio weights are one-month return, and ones used to contruct the alpha portfolio are hold-to-maturity returns (about 50 days on average). The number of latent factors in estimating portfolio weights is 3 and the estimatino window is 6 months. Mean, SD, and SR shows the annualized mean return, annualized standard deviation, and the annualized Sharpe ratio of the alpha portfolios, respectively. Skew and Kurt are skewness and kurtosis. Min and Max are minimum and maximum returns of the alpha portfolios during the sample period.

Panel A: Single Equity Call Alpha Portfolio with Transaction Cost									
K	window	Eff. Spread	Mean	SD	Sharpe	Skew	Kurt	Min	Max
3	6	0	1.04	0.40	2.58	0.96	2.23	-0.32	0.60
3	6	10	0.86	0.39	2.20	0.86	2.16	-0.34	0.56
3	6	20	0.68	0.38	1.79	0.76	2.11	-0.37	0.53
3	6	40	0.32	0.36	0.87	0.54	2.00	-0.41	0.46
Pane B: Single Equity Put Alpha Portfolio with Transaction Cost									
K	window	Eff. Spread	Mean	SD	Sharpe	Skew	Kurt	Min	Max
3	6	0	0.91	0.47	1.94	0.81	3.19	-0.38	0.67
3	6	10	0.76	0.46	1.64	0.66	2.95	-0.40	0.64
3	6	20	0.60	0.45	1.32	0.50	2.73	-0.41	0.60
3	6	40	0.28	0.44	0.64	0.17	2.38	-0.45	0.53

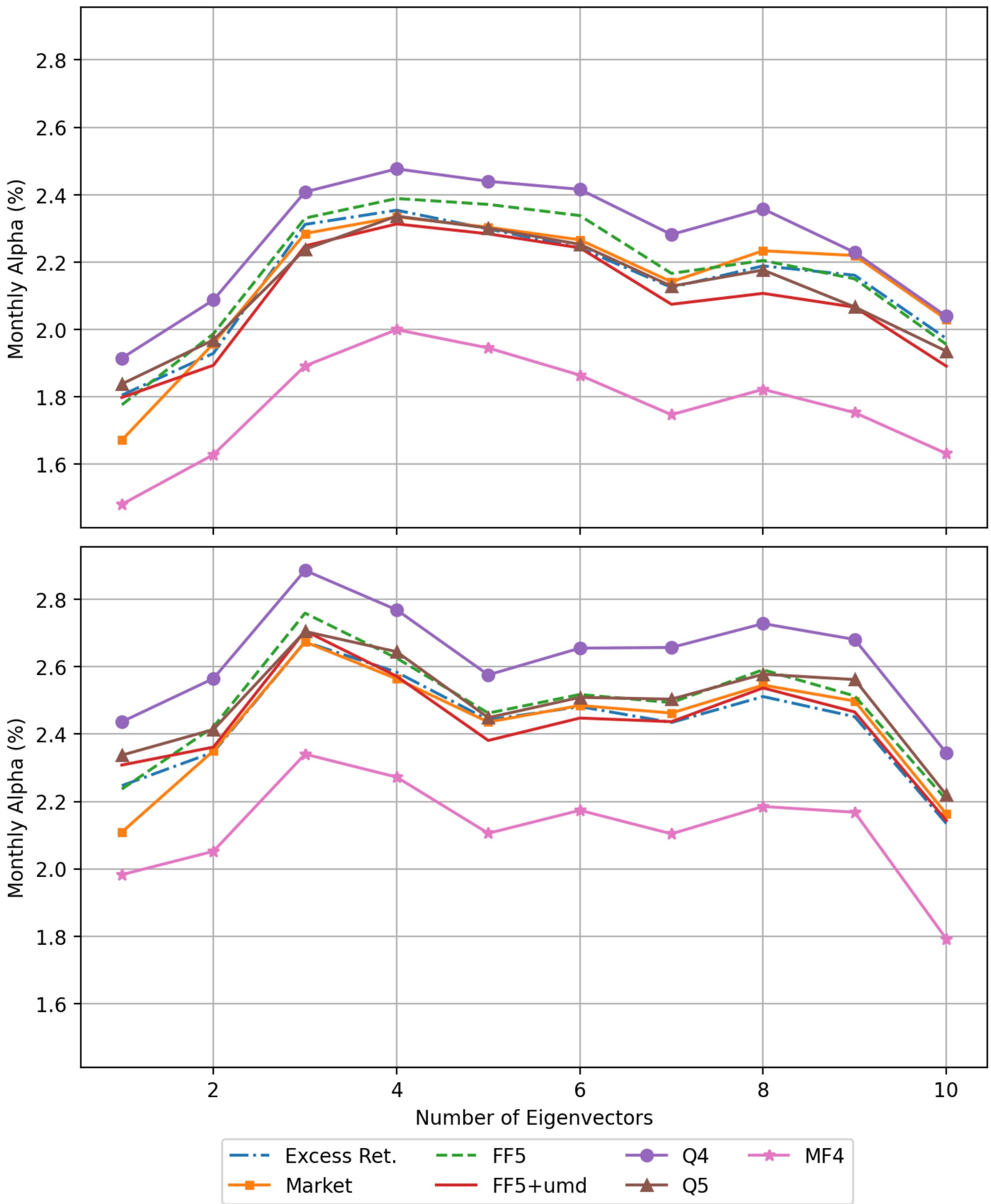


Figure A1: Alphas of Stock Market Risk Factors with Varying K

The two figures reveal the explanatory power of stock market risk factors on the returns of alpha portfolios with the varying number of latent factors K. The upper (lower) panel is for call (put) alpha portfolios. The estimation window of the alpha portfolios is twelve months. We regress the returns of the alpha portfolios on stock market risk factors and report alphas. Excess return is the monthly return of the alpha portfolios minus the risk-free rate. Market is the market portfolio of stocks in the CAPM model. FF5 is the market, size, value, investment, and profitability in Fama and French (2015). UMD is the momentum factor in Carhart (1997). Q4 is the market, size, investment and profitability in Hou, Xue and Zhang (2015). Q5 is Q4 augmented by expected investment growth in Hou, Mo, Xue and Zhang (2021). MF4 is market, size, management and performance factors in Stambaugh and Yuan (2017).

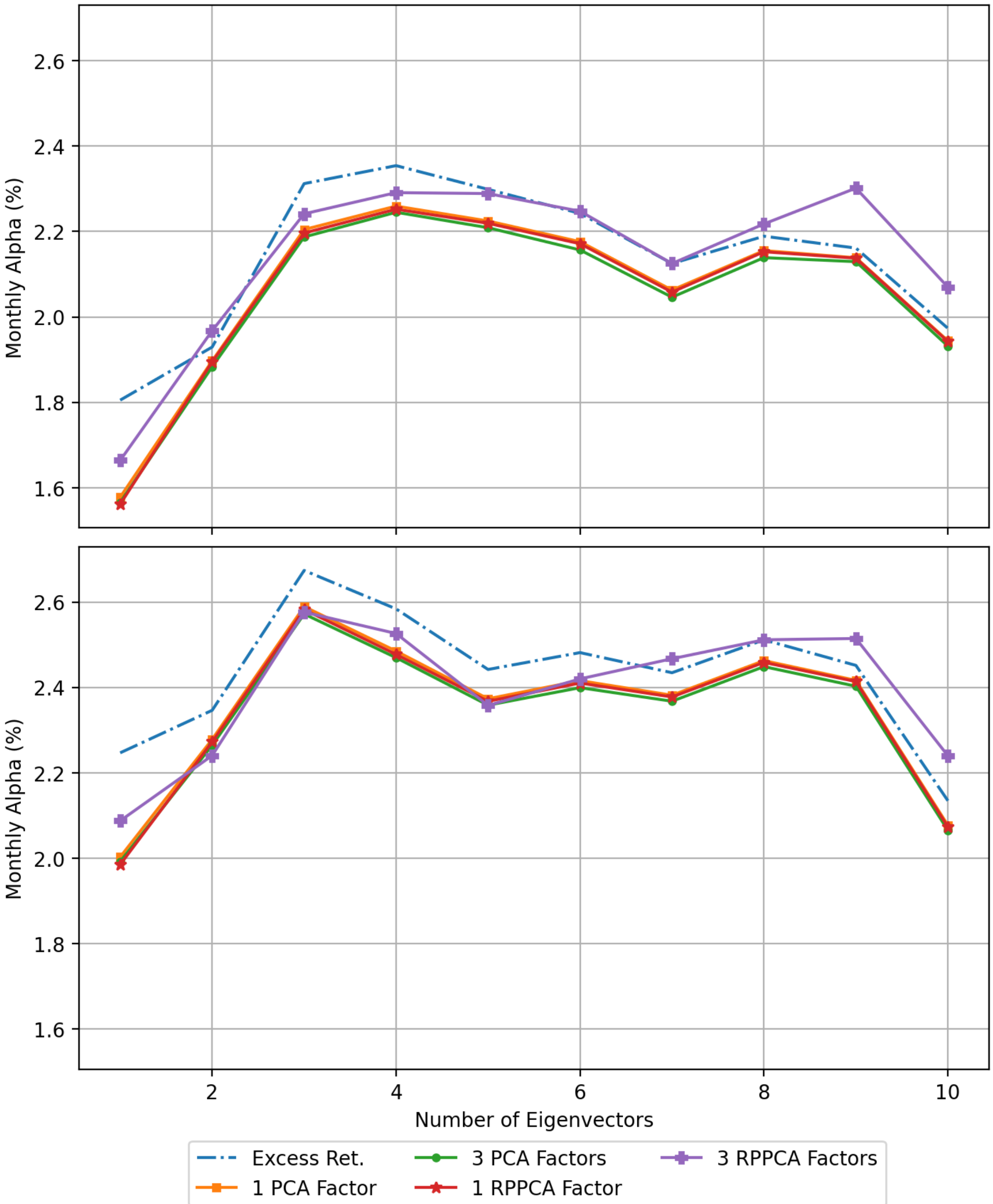


Figure A2: Alphas of Options Latent Factors with Varying K

This figure reveals the explanatory power of options latent factors on the returns of alpha portfolios with the varying number of latent factors K. The upper (lower) panel is for call (put) options. The estimation window of the alpha portfolios is twelve months. We regress the returns of the alpha portfolios on options latent factors and report alphas. Excess return is the monthly return of the alpha portfolios minus the risk-free rate. PCA (RPPCA) factors are constructed by applying PCA (RP-PCA) to option returns. The λ in RP-PCA is chosen to be 10.

Figure A3: Contributions of Characteristics to Put Options Factor Loadings

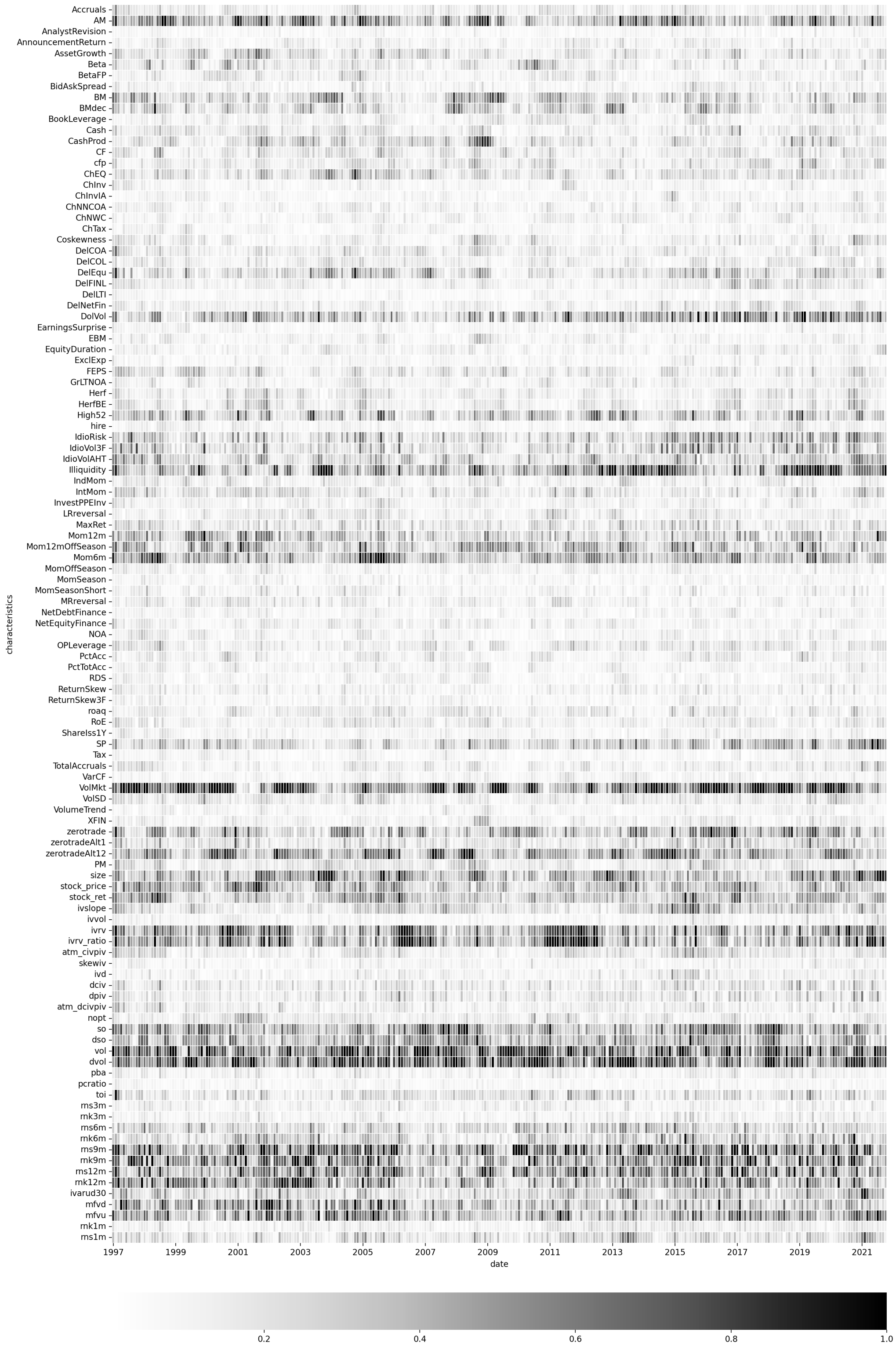
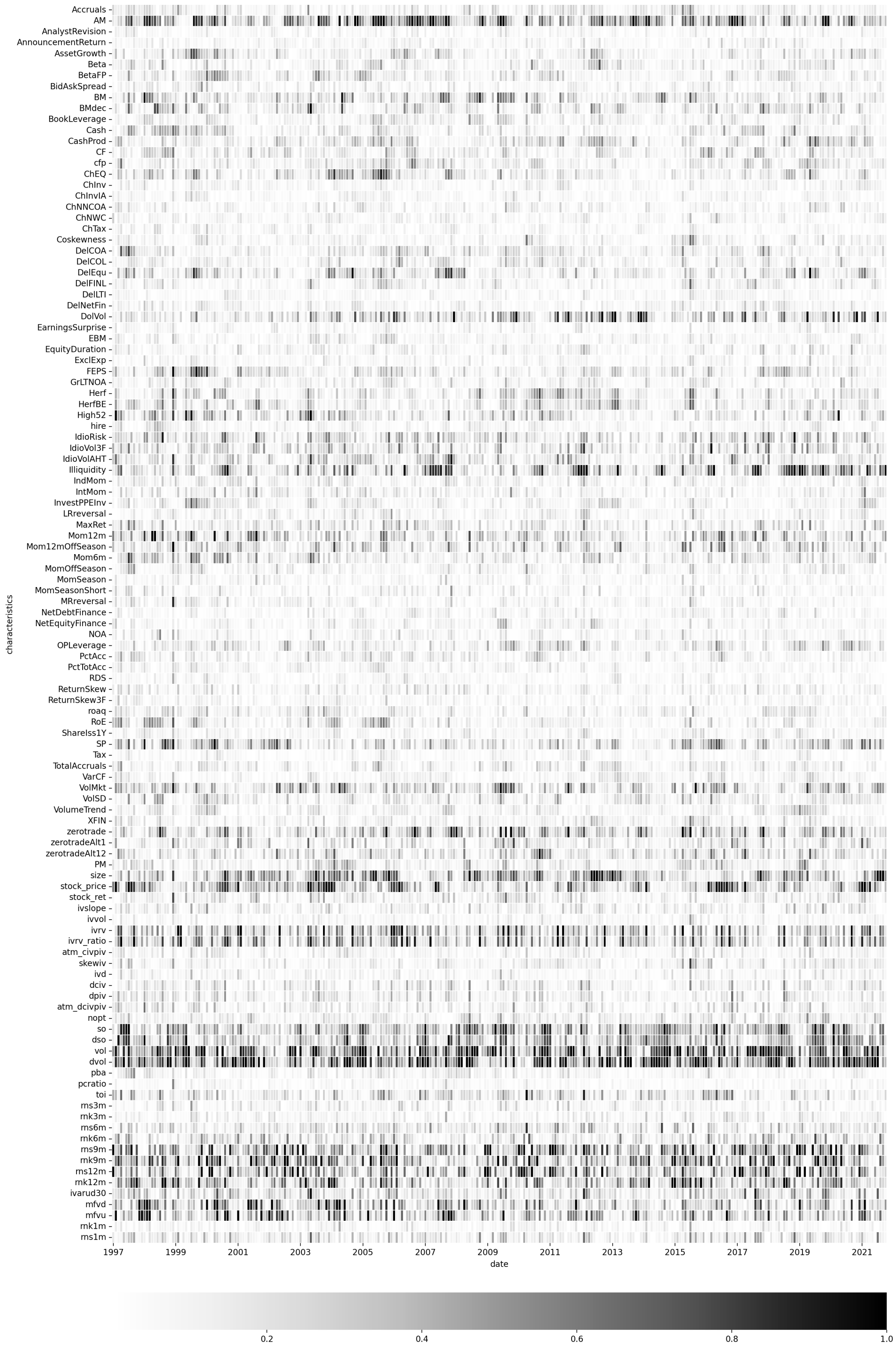


Figure A4: Contributions of Characteristics to Put Options Alphas



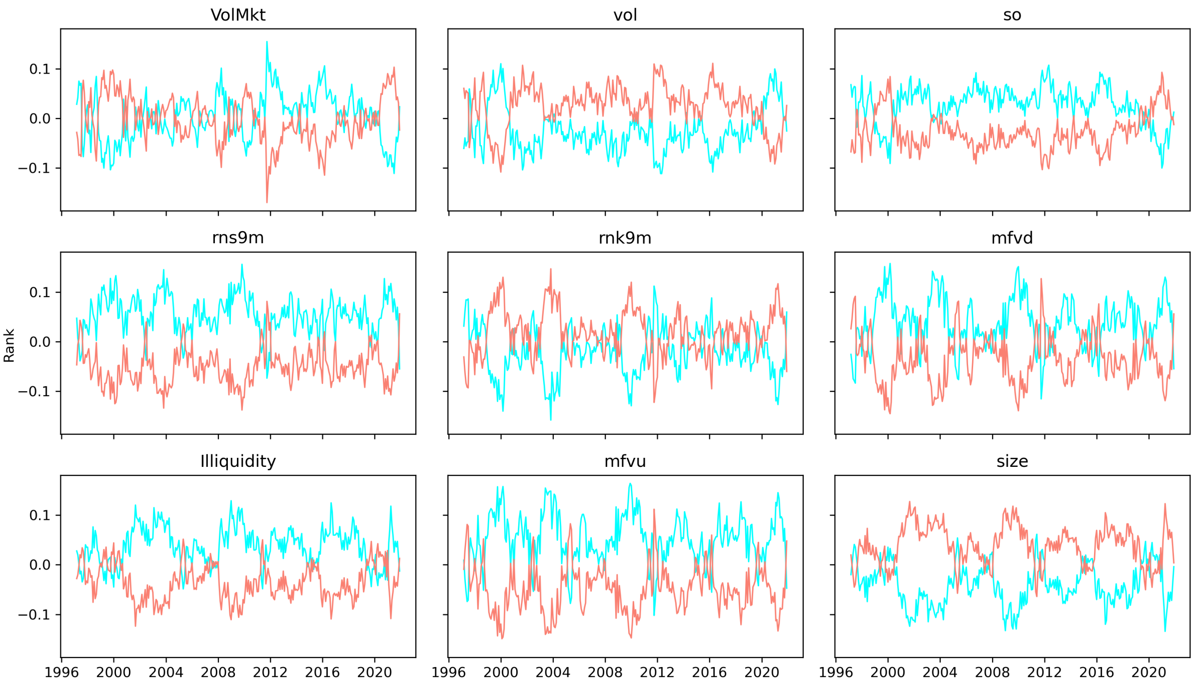


Figure A5: Ranks of Characteristics in the Long and Short Legs

The figures show the average ranks of characteristics in the long and short legs of the call options alpha portfolios. The number of latent factors is three, and the estimation window is twelve months. The ranks of characteristics range from -0.5 to 0.5. The long and short legs contain call options with positive and negative weights in the alpha portfolios, respectively. The light blue line represents the short leg, while the orange line represents the long leg.

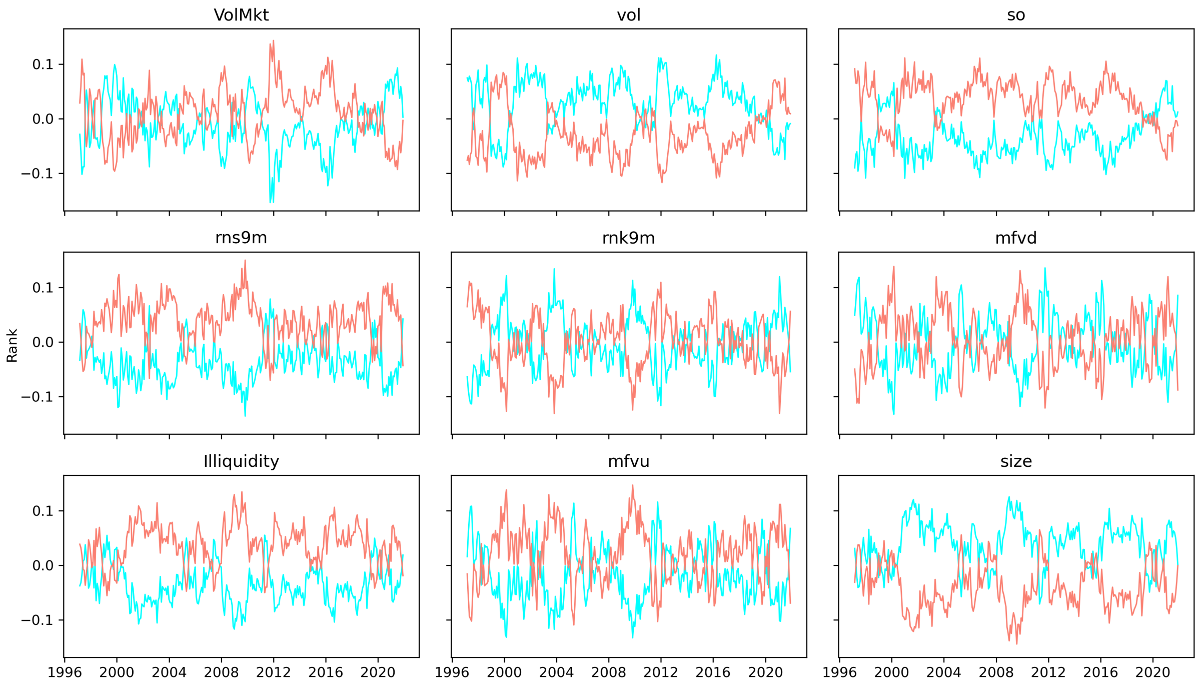


Figure A6: Ranks of Characteristics in the Long and Short Legs

The figures show the average ranks of characteristics in the long and short legs of the put options alpha portfolios. The number of latent factors is three, and the estimation window is twelve months. The ranks of characteristics range from -0.5 to 0.5. The long and short legs contain put options with positive and negative weights in the alpha portfolios, respectively. The light blue line represents the short leg, while the orange line represents the long leg.

Figure A7: Contributions of Characteristics to Put Options Factor Loadings - Nonlinear Case

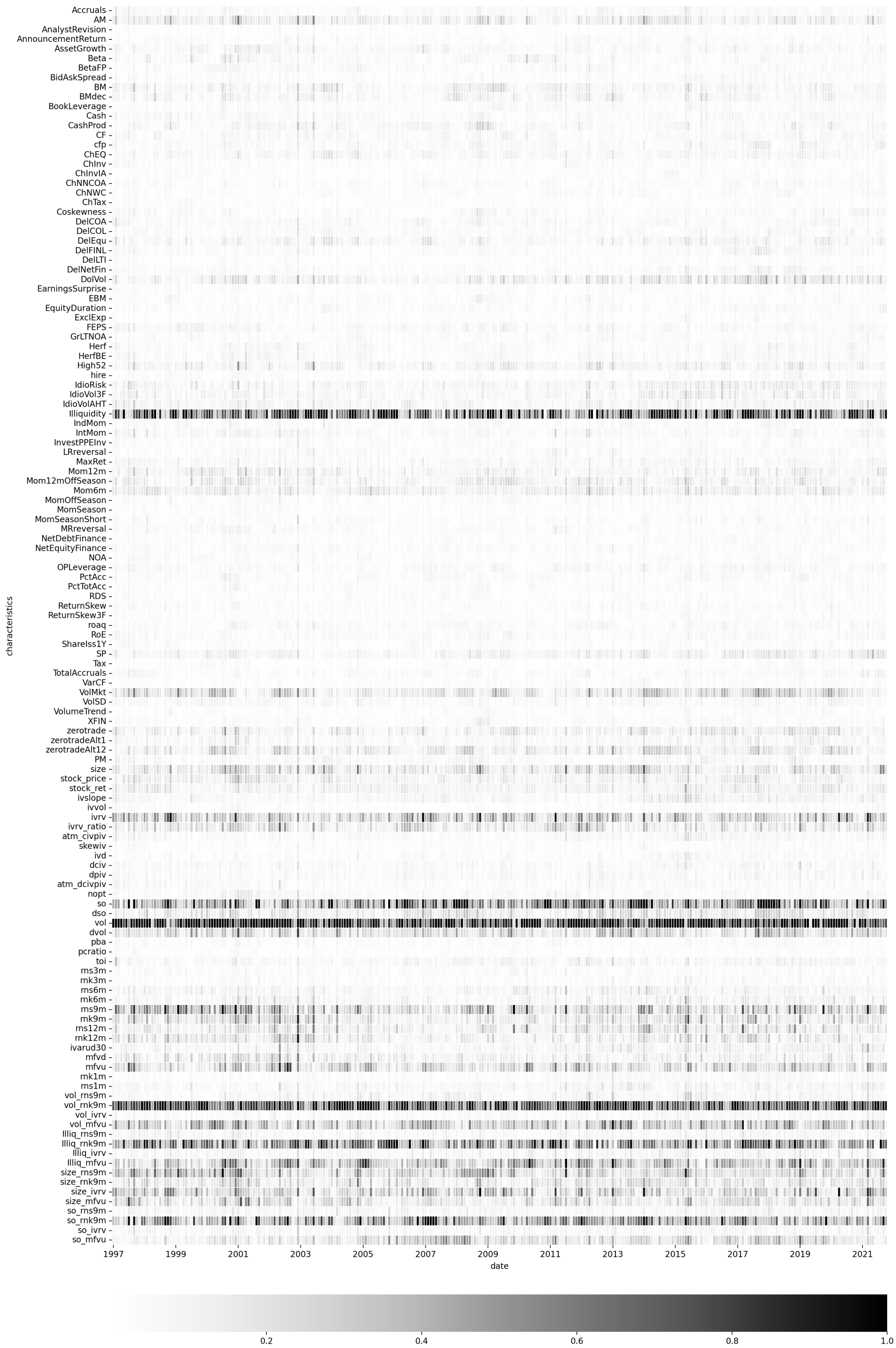


Figure A8: Contributions of Characteristics to Put Options Alphas - Nonlinear Case