Can Equity Option Returns Be Explained by a Factor Model? IPCA Says Yes

Amit Goyal **Alessio Saretto***

February 2024

Abstract

We show that much of the profitability in equity option return strategies, that try to capture option mis-pricing by taking exposure to underlying volatility, can be explained by an IPCA model. The economic magnitude of the return-adjustment produced by IPCA is important—even before transaction costs, the average IPCA alpha of 46 long-short trading strategies constructed on previously discovered signals is close to zero and contrasts with average realized returns of over 80 basis points per month.

JEL classification: G11, G12, G13

*Amit Goyal is from Swiss Finance Institute at the University of Lausanne, email: Amit.Goyal@unil.ch and Alessio Saretto is from the Federal Reserve Bank of Dallas, email: alessio.saretto@dal.frb.org. We thank Andreas Fuster, Bryan Kelly, and participants at the Virtual Derivative Workshop and the Frontiers of Factor Investing Conference at Lancaster University for valuable comments and suggestions, and Seth Pruitt for providing his code. The views expressed in this paper do not necessarily reflect those of the Federal Reserve System, the Federal Reserve Bank of Dallas or its staff.

1 Introduction

There is an extensive literature on factor models for stocks (see, for example, Fama and French, 2015 and Hou, Mo, Xue, and Zhang, 2021) and a growing literature on factor models for corporate bonds (see, for example, Chung, Wang, and Wu, 2019; Lin, Wang, and Wu, 2011; and Kelly, Palhares, and Pruitt, 2023). However, despite the proliferation of studies that document portfolio trading strategies in options that seem to generate abnormal profits (we discuss these studies later in the section on related literature), literature on factor models for equity options is scant. In this paper, we propose an option factor model that can explain a substantial portion of realized returns for most option trading strategies. Expected returns from our model are very close to their realized counterparts. Accordingly, the most commonly studied option trading strategies display little abnormal profitability.

We study delta-hedged option returns. In this case, most (if not all) option pricing models show that option returns depend on volatility dynamics of the underlying stock returns, interest rate dynamics, and jump risks. A factor structure of stock returns implies a factor structure for stocks' total volatility. Thus, part of the change in a stock's volatility is due to change in aggregate volatility. Accordingly, the literature often uses the aggregate volatility factor of Coval and Shumway (2001) to explain option returns. At the same time, Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) provide evidence of a factor structure in idiosyncratic volatility. Therefore, it is likely that another part of the change in a stock's volatility is due to change in idiosyncratic volatility and, thus, option returns are related to commonality in idiosyncratic volatility. Option returns are also related to changes in interest rates. Muravyev, Pearson, and Pollet (2022) show that the interest rate used to price options may be driven in part by borrow fees, which differ across stocks. These borrow fees may also have systematic and idiosyncratic components.[1] Finally, to the extent that there is commonality in jump risk (possibly related to default risk as in Vasquez and Xiao, 2023), option returns may share a common component related to jump and/or higher order risks. It is, thus, possible that characteristics that explain the cross-section of option returns do so because they proxy for betas on latent common factors related to volatility, borrow rates, and/or jump risks.

Instrumented principal component analysis (IPCA) of Kelly, Pruitt, and Su (2019) is, thus, ideally suited for our search for a factor model for options. IPCA helps in reducing the dimension of the many characteristics that are related to option returns and allows a parsimonious, yet intuitive, way to model time varying betas. These betas are allowed

[1] While option Greeks capture sensitivity of option prices to some of these risks, these Greeks are not option 'betas' to option risk factors because they measure sensitivities to stock-specific risks only.

to be functions of stock, option, and contract characteristics. IPCA eschews the need to construct portfolios on an ad-hoc basis and, instead, performs a dimension reduction of many individual assets into a few latent factors. Therefore, IPCA can be used to extract factors from individual option returns potentially including multiple option contracts on the same underlying stock.

In our empirical analysis, we construct a library of predictor variables for option returns over the sample period from 1996 to 2022. Our characteristics come from previous studies such as An, Ang, Bali, and Cakici (2014), Boyer and Vorkink (2014), Bali and Murray (2013), Cao and Han (2013), Christoffersen, Goyenko, Jacobs, and Karoui (2018), Goyal and Saretto (2009), Hu and Jacobs (2020), Vasquez (2017), and Zhan, Han, Cao, and Tong (2022), and can be classified into four categories: (i) contract characteristics such as maturity, open interest, option Greeks, etc.; (ii) descriptors of risk-neutral return distribution such as model free implied volatility, skewness, and kurtosis; (iii) measures of physical return distribution such as realized volatility, skewness, kurtosis, and autocorrelation; and (iv) firm level characteristics such as size, price, and book-to-market. In total, we use 46 characteristics as predictors of option returns and as IPCA conditioning variables.

We study expiration to expiration returns defined as the percentage change between an option position from Monday after expiration till expiration Friday next month. Thus, the option contract at the initiation of the position has residual maturity of around 25 days. The option positions are delta-hedged daily up to the contract expiration. We show that many predictors generate statistically significant average returns to long-short delta-hedged call option portfolios obtained by sorting stocks based on the set of 46 characteristics. A few predictors produce extremely significant average returns with t-statistics larger than 10. For example, the average monthly return from sorts on the Goyal and Saretto (2009) realized minus implied volatility spread (RV−IV) is 2.87% (t-statistic = 24.51).

Because it is well-known that transaction costs in option markets are high (see, for example, Christoffersen, Goyenko, Jacobs, and Karoui, 2018 and Muravyev, 2016), we also calculate portfolio returns net of transaction costs, calculated following Muravyev and Pearson (2020). Note that our expiration to expiration portfolios incur option transaction costs only at the initiation leg of the portfolio. Even after controlling for option transaction costs, we find significant returns for 16 strategies. For example, the net average return to the RV−IV strategy is 1.73% (t-statistic = 14.94). our findings largely confirm results in earlier studies of option strategies. These studies often propose risk-adjustments using factor models from stocks and find large alphas from such factor models. Thus, the first pass could suggest that option markets exhibit huge inefficiencies.

Our contention is that this conclusion is premature and most likely a function of the factor models. While the joint hypothesis problem afflicts all empirical studies, it is possible that the extant factor models exacerbate this issue because option returns are highly nonlinear and should have dynamic factor loadings. Our IPCA factor model is designed to ameliorate these problems by extracting latent factors with dynamic betas from individual option returns.

In estimating the IPCA model, we vary the number of factors from one to five but find that three factors are sufficient to explain delta-hedged call option returns. For example, our baseline 3-factor IPCA model produces time-series R^2 (cross-sectional average of time-series R^2 for each delta-hedged call position) of 15.6% and cross-sectional R^2 (time-series average of the cross-sectional R^2 for each month) of 11.3%. These R^2 numbers are non-trivial and compare favorably with similar numbers reported for stocks (Kelly, Pruitt, and Su, 2019), corporate bonds (Kelly, Palhares, and Pruitt, 2023), and cryptocurrencies (Babiak and Bianchi, 2023). When we consider the pricing performance of managed portfolios (i.e., zero-cost portfolios that weight individual positions based on their ranks, transformed into a z-score) constructed on the same 46 characteristics that we use in our IPCA, these R^2s are around 75%.

One concern with IPCA models is that the factors and model parameters are estimated using the full-sample. While unconditional betas (or parameters of conditional betas) are estimated using full sample in most empirical studies, the extraction of latent factors could raise doubts about data over-fitting. To allay these concerns, we also perform an out-of-sample analysis. In this analysis, we estimate latent factors and model parameters on an expanding window basis and use these estimates to make real-time forecasts for the subsequent month. We find only a modest decrease in out-of-sample R^2 relative to their in-sample counterparts for managed portfolios. For example, for the case of 3-factor model, the cross-section R^2 is 75.70% in-sample and 69.62% out-of-sample. As in Kelly, Palhares, and Pruitt (2023), the reason for this modest deterioration in performance is the relatively few parameters that are estimated in IPCA.

Applying the estimated IPCA model to the list of trading strategies, we find that, accounting for multiple hypothesis testing (based on the 5% FDR of Benjamini and Hochberg, 1995), only 2 strategies have statistically significant IPCA alphas (5 at conventional levels). The magnitude of the alpha, even for considering only those that are significant, is much lower compared to the magnitude of the long-short raw returns. For example, the alpha of the Goyal and Saretto (2009) RV−IV long-short strategy is only 0.34% (t-statistic = 4.58) compared to the return of 2.87% (t-statistic = 24.51). Notably, no trading strategy has a significant alpha after accounting for transaction costs.

We conclude that the extant evidence on option trading strategies presents little significant challenges to market efficiency. It is important to bear in mind, however, that IPCA factors are statistically constructed to capture systematic variation in the cross-section of delta-hedged returns. The notion of anomalies and/or market efficiency should, therefore, be viewed merely from the lens of investors who must bear the risk to earn returns to these factors, regardless of their economic origins.

In analyzing the mapping from characteristics to covariances/dynamic betas, we find an important role of RV−IV. This characteristic is not only important for explaining the returns to the Goyal and Saretto (2009) strategy based on RV−IV (which may be somewhat mechanical), but it also captures commonality in many other predictors of option returns.

Although our baseline IPCA model uses the full set of 46 characteristics, we also analyze restricted versions of IPCA that use subsets of these characteristics based on their economic and/or statistical significance. We find that a subset of characteristics contains the bulk of the information that matters to explain the cross-section of all delta-hedged returns.

In the last part of our analysis, we investigate properties of the statistical IPCA factors. The average return of the three factors are 2.06%, 0.26%, and 0.11% per month. The pairwise correlations between factors are all negative, with the correlation between the first two factors being the most negative at −27%. The annualized Sharpe ratios of the mean-variance tangency portfolio constructed from the three factors are 5.8 and 4.4 for in-sample and out-of-sample respectively. After accounting for transaction costs, these Sharpe ratios decrease to 4.1 in-sample and 2.7 out-of-sample. Thus, the IPCA factors span a wide region of the option-return space although the caveat of notion of economic risk remains, as noted before.

We also analyze how the factors' realized returns behave in periods of particular stress in the market. In particular, we analyze (a) negative aggregate shocks to market return and VIX proposed by Büchner and Kelly (2022) and Karakaya (2013), (b) proxies that measure changes in the tails of the aggregate distribution such as tail risk measure of Kelly and Jiang (2014) and conditional risk premia of the skewness and kurtosis of S&P500 return distribution following Carr and Wu (2009) and Kozhan, Neuberger, and Schneider (2013), and (c) states of world related to the role of intermediaries (Baltussen, Da, Lammers, and Martens (2021) and He, Kelly, and Manela (2017)). We find that the three factors span different states of nature characterized by different dimensions of risk, but caution against over-interpreting these results as showing that IPCA factors returns represent risk-premia related to adverse conditions.

The rest of the paper proceeds as follows. Section 2 discusses related literature. Section 3 presents the data and descriptive statistics on the basic option portfolio strategies. We provide a brief overview of the IPCA method in Section 4. The main results of the paper are discussed in Section 5. Section 6 provides an interpretation of the IPCA betas and factors. Section 7 concludes. The internet appendix provides many other robustness checks.

2 Related literature

One can classify option return predictors as broadly related to signals coming from the underlying stocks (or firms) or those coming from options themselves. Regarding stock predictors of option returns, Cao and Han (2013) show that delta-hedged option returns decrease with the stock idiosyncratic volatility. Galai and Masulis (1976) and Hu and Jacobs (2020) show, both theoretically and empirically, that the expected returns of calls (puts) decrease (increase) with underlying stock volatility. Aretz, Lin, and Poon (2023) show that it is important to distinguish the effect of idiosyncratic and systematic volatility in studying these relations. An, Ang, Bali, and Cakici (2014) show that implied volatility increases for options on stocks with high past returns. Ramachandran and Tayal (2021) report a monotonic relation between various measures of short-sales constraints and delta-hedged put returns on overpriced stocks. Zhan, Han, Cao, and Tong (2022) show that many stock characteristics (such as profit margin, firm profitability, cash holding, and shares issuance) have a strong predictive power for delta-hedged option returns.

There are also many studies that use option characteristics to predict option returns. Boyer and Vorkink (2014) find that ex-ante option implied skewness predicts option returns negatively. Relatedly, Bali and Murray (2013) find a negative relation between risk-neutral stock skewness and the returns of skewness assets constructed from options. Byun and Kim (2016) report that call options written on the most lottery-like stocks under-perform otherwise similar call options written on the least lottery-like stocks. Christoffersen, Goyenko, Jacobs, and Karoui (2018) find a positive risk-adjusted return spread for illiquid over liquid equity options. Goyal and Saretto (2009) find that the difference between historical volatility and implied volatility predicts delta-hedged and straddle returns. Heston, Jones, Khorram, Li, and Mo (2023) find that momentum works in option markets too. Muravyev (2016) documents that option market order-flow imbalance significantly predicts daily option returns. Ruan (2020) finds that volatility-of-volatility has predictive power for option returns. Vasquez (2017) shows that difference in long-term and short-term implied volatility has predictive power for delta-hedged option returns. Bali, Beckmeyer, Moerke, and Weigert (2023) and Shafaati, Chance, and Brooks (2021) apply machine learning techniques to uncover which characteristics are important for explaining the cross-section of option returns (see also Goyenko and Zhang, 2022 for using machine learning to predict the joint cross-section of stock and option returns).

The literature on factor models for option returns is small. Coval and Shumway (2001) propose a factor constructed from zero-beta straddles to explain S&P500 index option returns. Büchner and Kelly (2022) apply IPCA techniques to construct factors for index options. Karakaya (2013) identifies three factors based on level, slope, and value in individual naked option returns. Broadie, Chernov, and Johannes (2009), however, suggest that straddles or delta-hedged option returns are more immune than naked option returns to sampling problems. More recently, Horenstein, Vasquez, and Xiao (2022) construct option factors by using RP-PCA of Lettau and Pelger (2020). Their factor model presents an improvement over option factors constructed from index options of Coval and Shumway (2001) and Büchner and Kelly (2022). While we view their model as a complementary approach to modeling option returns, IPCA has two advantages over RP-IPCA: first, IPCA can handle information at the firm (and potentially at the option contract) level rather than just at the portfolio level; second, IPCA allows for dynamic betas.

3 Data

3.1 Sample and variable construction

Options data are from OptionMetrics and cover the period from January 1996 through December 2022. We impose two sets of filters. The first type of filter applies to all observations and is aimed at eliminating contracts that are non-standard or prices that violate basic no-arbitrage bounds or market minimum specifications. We eliminate contracts (i) for which the underlying is not a share code 10 or 11 stock, or (ii) that have non-standard settlements (i.e., not the closing of the expiration day), or (iii) that have non-standard expiration (i.e., weeklies), or (iv) that settle for a number of underlying shares different from 100, or (v) for which one of the bid or ask quotes is missing, or (v) for which the ask is lower than the bid, or the ask is $5 above the bid, or (vi) the bid-ask spread is lower than the minimum tick size (i.e., 5 cents for option prices below $3 and 10 cents for option prices above $3), or (vii) for which the mid-point price is below or $100 above the current exercise payoff, or (viii) option delta is less than equal to zero for calls, or greater than equal to zero for puts.

The second set of filters only applies to prices that are used to construct the first leg of the return. We study expiration to expiration returns defined as the percentage change between an option position from Monday after expiration till expiration Friday next month, where the option contract at the initiation of the position has residual maturity of around 25 days.[2] For example, for calculation of return for the "month" of February 2022, we initiate the position on Monday January 24, 2022 and terminate the position on Friday February 18, 2022. We apply the following filters only to the prices on January 24.[3] We consider options (i) that are closest to ATM with moneyness (defined as the ratio of strike to underlying price) between 0.8 and 1.2, and (ii) that have positive volume or positive open interest, (iii) for which the mid-point price is higher than 25 cents, and (iv) the percentage bid-ask spread is less than 50%. In addition, to avoid issues related to early exercise of American options, we remove all options on underlying stocks that pay a dividend during the holding period (i.e., between January 24 and February 18 in the example above).

As the literature mostly focuses on daily delta-hedged call returns as the norm for investigating cross-sectional patterns in stock option returns (see, for example, Cao and Han, 2013 and Zhan, Han, Cao, and Tong, 2022), we also consider daily rebalanced delta-hedged calls.[4] We define returns to buying a delta-hedged call as:

$$\text{DHCall}_{t+1} = \frac{(C_{t+1} - C_t) - \sum_{n=0}^{N-1} \Delta_{d_n}(S_{d_{n+1}} - S_{d_n})}{\Delta_t S_t - C_t} - R_{ft+1}, \tag{1}$$

where we rebalance the hedge at each of the days $d_n, n = 0, ..., N-1$, $d_0 = t$, $d_N = t+1$, $C_{t+1} = \max(S_{t+1} - K, 0)$, and R_f is the riskfree return. Following standard convention, we scale the delta-hedged gains by $(\Delta_t S_t - C_t)$ to guarantee that the denominator is always positive. Therefore, our return is the negative of the return to writing a delta-hedged call.

Stock level information is obtained from CRSP/Compustat data, which are matched and aligned in time with OptionMetrics information so that no look ahead biased is introduced. All accounting information is assumed to be available six months after the fiscal year end. For each stock-month return position we construct monthly characteristics that can generally be divided into four categories:

1. Contract characteristics such as moneyness, bid-ask spread, open interest, volume, option price, option delta, option vega, option gamma, and implied volatility.

2. Measures of risk-neutral distribution. Model-free implied volatility, model-free implied skewness, and model-free implied kurtosis are constructed using options with the same

[2]Initiating the position on Monday allows us to consider option chains that are introduced on Monday and have had some volume/open interest to satisfy our sample construction filters.

[3]Duarte, Jones, Khorram, and Mo (2023) highlight issues when researchers inadvertently apply filters to the data that are infeasible from a trading strategy perspective.

[4]We present an analysis of delta-hedged puts and straddles in the Internet Appendix Section IA3.

moneyness and approximate maturity. We also construct variables using the implied volatility surface: the ATM level (of implied volatility) for approximate maturity of 30 days; the slope of the 30 days curve, as the difference between implied volatility for moneyness of 0.8 and the corresponding value for moneyness of 1.0; the ATM term, as the difference between the implied volatility value for 30 days and the corresponding value for 360 days maturity; the volatility of implied volatility.

3. Measures of the physical distribution of returns such as the stock returns: volatility, autocorrelation, skewness, kurtosis, cumulative return of the previous eleven months, the largest absolute return in the past 10 days (Max10), turnover, and idiosyncratic volatility.

4. Stock returns predictors that have been found to predict also option returns such as the RSI, BM, asset, debt, leverage, profit margin, profitability, equity issuance, external financing cash, cash flow variance, analyst dispersion, lever of institutional ownership, and Z score.

We refer to Appendix A for detailed descriptions about variables constructions and their source references.

3.2 Empirical regularities

We construct equal-weighted portfolios of daily rebalanced delta-hedged call positions by sorting stocks into deciles based on the characteristics detailed in the previous section. Following the bulk of the literature (see, for example, Aretz, Lin, and Poon, 2023; Bali and Murray, 2013; Cao and Han, 2013; Goyal and Saretto, 2009; and Hu and Jacobs, 2020), we consider expiration to expiration returns and sort stocks on Monday following expiration Friday (i.e., we do not skip a day between portfolio formation and trading initiation).[5]

For each sorting variable, we calculate the return of long-short portfolios that invest in the extreme deciles. These returns are calculated using mid-point prices (average of bid and ask). To ease interpretation, we construct the trading strategy so that the full-sample average return is positive. In other words, for some variables, we present results of a 1−10 portfolio, while for others we present results of a 10−1 portfolio. The second column of Table 1 shows the construction methodology used for each variable. We refer to these returns as those on long-short strategies. The third column reports average return and

[5]Some authors analyze option returns constructed from month-end to month-end to align with the extensive literature on stock return predictability (see, for example, Zhan, Han, Cao, and Tong, 2022). We consider this alternate return construction in the Internet Appendix Section IA4.

Table 1: Long-short delta-hedged call portfolios

The table reports average returns with and without accounting for transaction costs of 46 long-short portfolios. We construct the trading strategies so that average returns are positive by either buying options positions in decile 10 and writing option positions in decile 1, or by buying options in decile 1 and writing options in decile 10. We tabulate the portfolio construction in the second column. Transaction costs are calculated using a ratio of effective to quoted spread equal to 30%. t-statistics larger than the Benjamini and Hochberg (1995) 5% FDR adjusted threshold (i.e., 2.25 for mean return and 2.44 for transaction cost adjusted returns) are reported in bold. The sample period is 1996 to 2022.

	Construction	Return		Return w/ TC	
Moneyness	1−10	**0.38**	**(3.60)**	−0.86	(−8.13)
Bid-ask	1−10	**0.71**	**(8.07)**	−0.13	(−1.50)
Open interest	1−10	0.04	(0.66)	−0.59	(−9.84)
Delta	10−1	**0.22**	**(2.58)**	−0.68	(−7.52)
Vega	10−1	**1.44**	**(14.83)**	**0.47**	**(5.02)**
Gamma	1−10	**0.67**	**(7.09)**	−0.25	(−2.69)
Volume	10−1	**0.23**	**(3.11)**	−0.40	(−5.28)
Option price	10−1	**0.65**	**(5.72)**	−0.12	(−1.09)
IV ATM	1−10	**2.34**	**(17.94)**	**1.40**	**(10.95)**
IV slope	10−1	**2.15**	**(25.53)**	**0.91**	**(10.46)**
IV term	10−1	**1.42**	**(15.94)**	**0.43**	**(4.84)**
IV vol	1−10	**1.45**	**(15.65)**	**0.61**	**(6.81)**
MFvol	1−10	**1.16**	**(5.16)**	**0.69**	**(3.13)**
MFskew	1−10	**0.83**	**(4.98)**	**0.42**	**(2.54)**
MFkurt	10−1	**0.74**	**(4.41)**	0.33	(1.97)
Stock price	10−1	**1.73**	**(16.09)**	**0.74**	**(7.03)**
Stock return	10−1	0.04	(0.46)	−0.93	(−9.30)
Stock return11	10−1	**0.29**	**(2.91)**	−0.71	(−7.09)
RV	1−10	**1.06**	**(9.23)**	0.22	(1.94)
Rskew	1−10	**0.14**	**(2.23)**	−0.74	(−12.62)
Rkurt	1−10	**0.91**	**(14.48)**	**0.15**	**(2.45)**
Turnover	1−10	0.13	(1.32)	−0.66	(−6.68)
IdiosynVol	1−10	**1.15**	**(11.18)**	**0.33**	**(3.26)**
Max10	1−10	**0.94**	**(8.35)**	0.11	(0.99)
MarketCap	10−1	**1.59**	**(16.74)**	**0.67**	**(7.22)**
Autocorrelation	10−1	0.07	(1.24)	−0.72	(−12.25)
RV−IV	10−1	**2.87**	**(24.51)**	**1.73**	**(14.94)**
RV−MFvol	10−1	**1.64**	**(7.49)**	**1.11**	**(5.16)**
Rskew−MFskew	1−10	0.11	(0.73)	−0.31	(−2.13)
Rkurt−MFkurt	1−10	**1.00**	**(8.03)**	**0.65**	**(5.22)**
BM	10−1	0.13	(1.39)	−0.75	(−8.33)
Profitability	10−1	**0.92**	**(11.71)**	0.03	(0.41)
InstOwn	10−1	**0.93**	**(12.10)**	0.03	(0.38)
RSI	10−1	0.10	(1.56)	−0.63	(−9.77)
Assets	10−1	**1.46**	**(14.59)**	**0.64**	**(6.71)**
Debt	10−1	**0.68**	**(8.82)**	0.04	(0.50)
Leverage	10−1	0.00	(0.06)	−0.84	(−15.11)
CashFlowVar	1−10	**0.72**	**(9.11)**	−0.01	(−0.16)
Cash to asset	1−10	**0.87**	**(9.42)**	−0.04	(−0.43)
AnalystDisp	10−1	**0.52**	**(8.18)**	−0.49	(−7.82)
1yr NewIss	1−10	**0.43**	**(5.58)**	−0.34	(−4.45)
5yr NewIss	1−10	**0.65**	**(8.33)**	−0.09	(−1.25)
Profit margin	10−1	**1.15**	**(13.37)**	**0.25**	**(3.05)**
ROE	10−1	**0.82**	**(11.09)**	−0.07	(−0.99)
ExternalFin	1−10	**0.52**	**(6.10)**	−0.31	(−3.70)
Z score	10−1	**0.68**	**(7.99)**	−0.29	(−3.62)

related t-statistics. To caution against the interpretation of single test statistical significance we provide multiple hypothesis testing (MHT) adjusted significance accounting for a 5% FDR following Benjamini and Hochberg (1995) (see Harvey, Liu, and Saretto, 2020 for a review of multiple hypothesis methods).

Table 1 shows that many characteristics are strong predictors of delta-hedged call returns. Of the 46 predictors, 39 have statistically significant average returns to long-short portfolios (MHT adjustment versus conventional level does not change this number). A few predictors produce extremely significant average returns with t-statistics larger than 10. For example, the average monthly return resulting from sorts on RV−IV, IV ATM, and IV slope is 2.87%, 2.34%, and 2.15%, respectively, with t-statistics of around 20. Relative to the Black-Scholes idealized framework of expected excess return on delta-hedged call return of zero, the magnitude of many of these returns is large.[6]

It is well-known that transaction costs in option markets are high (see, for example, Christoffersen, Goyenko, Jacobs, and Karoui, 2018; Muravyev, 2016; and Muravyev and Pearson, 2020). If investors cannot profit from option strategies after accounting for transaction costs, then one may not learn much from factor models that price options in the absence of transaction costs. Therefore, we pay special attention to transaction costs in our analysis. Our empirical choice of studying expiration to expiration returns is already a first step in this direction as our portfolios incur option transaction costs only at the initiation leg of the portfolio.

We adjust the option initial prices to account for transaction costs following the evidence presented by Muravyev and Pearson (2020). We consider a ratio of effective to quoted spread, ESPR/QSPR, equal to 30%. We obtain effective bid and ask prices at which an investor could either buy or sell a contract. For example, if the bid price of an option is \$3 and the ask price is \$4, then a ratio of ESPR/QSPR = 30% implies that we buy the option at \$3.65 for a long position or sell the option at \$3.35 for a short position (instead of the mid-point price of \$3.50). Net returns are reported in the fifth column of Table 1. Since, by construction, the gross returns to all strategies are positive, a negative net return imply that the strategy is not profitable after transaction costs. We find that even net of transaction cost 16 trading strategies still have significant returns. For example, the average monthly net return of the RV−IV strategy is still 1.73% (t-statistic = 14.94).

[6]For some characteristics, the magnitude of returns in Table 1 is different from that in the original studies (for instance, Christoffersen, Goyenko, Jacobs, and Karoui, 2018 and Zhan, Han, Cao, and Tong, 2022). There are two reasons for this difference. One, we study expiration to expiration returns while some studies use month-end to month-end returns. Second, Duarte, Jones, Khorram, and Mo (2023) raise sample selection issues with some studies. As discussed in Section 3.1, our sample does not have any look-ahead bias.

Can these returns be explained by a factor model? We turn to answering this question next.

4 IPCA

In this section, we present a detailed description of the IPCA method introduced by Kelly, Pruitt, and Su (2019). Returns on N assets over time period $t = 1, \ldots, T$ are assumed to be generated by the model:

$$R_{i,t+1} = \alpha_{i,t} + \beta'_{i,t} F_{t+1} + E_{i,t+1} = \left(Z'_{i,t}\Gamma_\alpha\right) + \left(Z'_{i,t}\Gamma_\beta\right) F_{t+1} + E_{i,t+1}, \tag{2}$$

where F's are K latent factors and E's are the residuals. $\alpha_{i,t}$ and $\beta_{i,t}$ are asset specific time-varying intercept and factor loadings that are both assumed to be linearly related to L observable characteristics summarized in $Z_{i,t}$ (including a constant). The $L \times 1$ vector Γ_α and the $L \times K$ matrix Γ_β define the mapping from a potentially large number of characteristics to the mispricing and a small number of risk factor exposures. The parameters Γ are global in the sense that are assumed to be constant across all stocks and all time periods.

Kelly, Pruitt, and Su (2019) emphasize that the estimation of Γ_β amounts to finding a few linear combinations of candidate characteristics that best describe the latent factor loading structure. Thus, IPCA eschews the need to construct portfolios on an ad-hoc basis and instead performs a dimension reduction of many individual assets into a few latent factors.

Denote R_{t+1} as $N_{t+1} \times 1$ vector, Z_t as a $N_{t+1} \times L$ matrix, α_t as a $N_{t+1} \times 1$ vector, β_t as a $N_{t+1} \times K$ matrix. Here N_{t+1} is the number of stocks with a valid return at time $t+1$ and all observable characteristics at time t. Then, we can rewrite the model in matrix form as:

$$R_{t+1} = \alpha_t + \beta_t F_{t+1} + E_{t+1} = (Z_t\Gamma_\alpha) + (Z_t\Gamma_\beta) F_{t+1} + E_{t+1}. \tag{3}$$

It will be useful to define a $L \times L$ matrix $W_t = Z'_t Z_t / N_{t+1}$ and a $L \times 1$ vector $X_{t+1} = Z'_t R_{t+1} / N_{t+1}$. We can interpret X_{t+1} as the $t+1$ returns on a set of L managed portfolios. The return on the lth portfolio is a weighted average of asset returns with weights determined by the value of lth characteristic for each asset at time t (normalized by the number of non-missing observations each month, N_{t+1}).

The first order conditions for the estimation of the system are given by:

$$\begin{aligned} \hat{F}_{t+1} &= \left(\hat{\Gamma}'_\beta W_t \hat{\Gamma}_\beta\right)^{-1} \hat{\Gamma}'_\beta (X_{t+1} - W_t \hat{\Gamma}_\alpha) \\ \text{vec}\left(\hat{\Gamma}'\right) &= \left(\sum_{t=1}^{T-1} W_t \otimes \hat{\tilde{F}}_{t+1} \hat{\tilde{F}}'_{t+1}\right)^{-1} \sum_{t=1}^{T-1} X_{t+1} \otimes \hat{\tilde{F}}_{t+1}, \end{aligned} \tag{4}$$

where $\tilde{F}_{t+1} = [F_{t+1} : 1]'$, and $\Gamma = [\Gamma_\beta : \Gamma_\alpha]$. As Kelly, Pruitt, and Su (2019) show, the first order conditions involve OLS regressions only and, therefore, the estimation proceeds very quickly. As is evident from equation (4), the estimation requires only the L-dimensional objects W_t and X_{t+1} rather than original N-dimensional asset returns. A further advantage of this method, thus, is the ease with which missing data are handled.

Identifying a unique set of parameters is important in latent factor models, as they are identified only up to a rotation. Models $\Gamma_\beta F$ and $\Gamma_\beta P P^{-1} F$ are identical for any rotation matrix P. Following Kelly, Pruitt, and Su (2019), we impose the normalization that $\Gamma'_\beta \Gamma_\beta$ is the identity matrix, that the unconditional second moment matrix of F is diagonal with descending diagonal entries, and that the time-series average of F is positive. In models with unrestricted alpha, we impose an additional identification assumption of $\Gamma'_\alpha \Gamma_\beta = 0$. The identification assumptions do not restrict the model's ability to explain returns, but merely serve as a way to pin down unique parameters of the model. We initialize the algorithm by setting Γ_β equal to the first K eigenvectors of the sample second moment of the managed portfolio returns and factors F to be the first K principal components of the managed portfolio panel. We iterate the two first-order conditions in equation (4) until convergence. Since $Z'_t Z_t$ is not too volatile, the initial conditions are a good starting guess and convergence is achieved rapidly.[7]

We also estimate a restricted IPCA model where $\Gamma_\alpha = 0$ to test whether risk compensation in returns solely arises from exposure to systematic factors, or whether returns partially line up with characteristics directly, hence constituting compensation without risk.

4.1 Statistical inference

4.1.1 Γ_α and number of factors

We first estimate the unrestricted model to obtain parameters $\hat{\Gamma}_\alpha$ and $\hat{\Gamma}_\beta$ and the time-series of the latent factors $\hat{F}_t$. Note that equation (3) and the definitions of W_t and X_{t+1} imply

[7] We thank Seth Pruitt for providing code for IPCA estimation on his website.

that:

$$X_{t+1} = W_t\Gamma_\alpha + W_t\Gamma_\beta F_{t+1} + Z_t' E_{t+1}/N_{t+1}. \tag{5}$$

Define the residuals of managed portfolio returns as $U_{t+1} \equiv Z_t' E_{t+1}/N_{t+1}$ and their fitted values as $\hat{U}_{t+1} = X_{t+1} - W_t\hat{\Gamma}_\alpha - W_t\hat{\Gamma}_\beta \hat{F}_{t+1}$. The returns on managed portfolio for the bth bootstrap are generated by:

$$X_{t+1}^b = W_t\hat{\Gamma}_\beta \hat{F}_{t+1} + q_{t+1}^b \hat{U}_{t+1}^b, \tag{6}$$

by imposing the null of zero alpha ($\Gamma_\alpha = 0$). In equation (6), $\hat{U}_{t+1}^b$ is the estimated residual from a random time period drawn uniformly from the set of all possible dates and q_{t+1}^b is a Student-t random variable with unit variance and five degrees of freedom, as in Kelly, Pruitt, and Su (2019). Using the bootstrapped sample, we reestimate the unrestricted IPCA model and record the estimated parameter $\hat{\Gamma}_\alpha^b$. Recall that since the estimation of IPCA model requires only W and X matrices, we do not need to bootstrap the asset returns themselves. Finally, we calculate Wald statistics $W_\alpha = \hat{\Gamma}_\alpha' \hat{\Gamma}_\alpha$ in the data and $W_\alpha^b = \hat{\Gamma}_\alpha^{b\prime} \hat{\Gamma}_\alpha^b$ in the bth bootstrap. Inferences are drawn by calculating a p-value as the fraction of bootstrapped W_α^b statistics that exceed the value of W_α from the actual data. We use $B = 1{,}000$ repetitions in the bootstrap.

4.1.2 Characteristic importance

We modify the procedure described above to test significance of individual characteristic. Since $\hat{\Gamma}_\beta = (\hat{\gamma}_{\beta,1}, \ldots, \hat{\gamma}_{\beta,l-1}, \hat{\gamma}_{\beta,l}, \hat{\gamma}_{\beta,l+1}, \ldots \hat{\gamma}_{\beta,L})'$, we can test for the significance of the lth characteristic by setting $\hat{\gamma}_{\beta,l} = 0$, defining $\hat{\Gamma}_\beta^l = (\hat{\gamma}_{\beta,1}, \ldots, \hat{\gamma}_{\beta,l-1}, 0_{K\times 1}, \hat{\gamma}_{\beta,l+1}, \ldots \hat{\gamma}_{\beta,L})'$, and then bootstrapping managed portfolio returns as

$$X_{t+1}^b = W_t\hat{\Gamma}_\beta^l \hat{F}_{t+1} + q_{t+1}^b \hat{U}_{t+1}^b. \tag{7}$$

We calculate Wald statistics $W_{\beta,l} = \hat{\gamma}_{\beta,l}' \hat{\gamma}_{\beta,l}$ in the data and $W_{\beta,l}^b = \hat{\gamma}_{\beta,l}^{b\prime} \hat{\gamma}_{\beta,l}^b$ in the bth bootstrap. Inferences are drawn by calculating a p-value as the fraction of bootstrapped $W_{\beta,l}^b$ statistics that exceed the value of $W_{\beta,l}$ from the actual data.

4.2 Pricing metrics

We calculate the pricing performance of the model by calculating a variety of R^2's and pricing error, following Kelly, Palhares, and Pruitt (2023) as follows:

$$\begin{aligned}
\text{Total } R^2 &= 1 - \frac{\sum_{i,t}\left(R_{i,t+1} - Z'_{i,t}\hat{\Gamma}_\beta \hat{F}_{t+1}\right)^2}{\sum_{i,t} R^2_{i,t+1}} \\
\text{Time Series } R^2 &= \frac{1}{\sum_i T_i}\sum\nolimits_i T_i R_i^2;\ \text{where } R_i^2 = 1 - \frac{\sum_t\left(R_{i,t+1} - Z'_{i,t}\hat{\Gamma}_\beta \hat{F}_{t+1}\right)^2}{\sum_t R^2_{i,t+1}} \\
\text{Cross Section } R^2 &= \frac{1}{T}\sum\nolimits_t R_t^2;\ \text{where } R_t^2 = 1 - \frac{\sum_i\left(R_{i,t+1} - Z'_{i,t}\hat{\Gamma}_\beta \hat{F}_{t+1}\right)^2}{\sum_i R^2_{i,t+1}} \\
\text{Relative Pricing Error} &= \frac{\sum_i \alpha_i^2}{\sum_i \bar{R}_i^2};\ \text{where } \alpha_i = \frac{1}{T_i}\sum\nolimits_t\left(R_{i,t+1} - Z'_{i,t}\hat{\Gamma}_\beta \hat{F}_{t+1}\right).
\end{aligned} \tag{8}$$

Note that the denominator in the Total R^2 is the sum of squared returns without demeaning following Gu, Kelly, and Xiu (2020). Total R^2 measures how well the model describes realized returns in a panel context. The Time Series R^2 summarizes pricing performance as the cross-sectional average of the performance of each test asset i (we weight each asset's performance by the number of time-series observations for that asset, T_i). The Cross Section R^2 quantifies the cross-sectional predictive performance as in a Fama-MacBeth context. Relative Pricing Error analyzes IPCA's ability to price average return. In contrast to R^2 measures, lower value of pricing error indicates better performance.

Following Kelly, Pruitt, and Su (2019), we also calculate R^2 and relative pricing error in an out-of-sample context. The exact procedure is as follows. For each month t, we first estimate IPCA on expanding window that includes data up to month t. Denote the resulting parameter estimates as $\hat{\Gamma}_{\beta,t}$. Then, we calculate the out-of-sample factor return at time $t+1$ as $\hat{F}_{t+1,t} = \left(\hat{\Gamma}'_{\beta,t} Z'_t Z_t \hat{\Gamma}_{\beta,t}\right)^{-1} \hat{\Gamma}'_{\beta,t} X_{t+1} = \hat{\theta}'_t X_{t+1}$. In this way, IPCA factor return at time $t+1$ uses portfolio weights, $\hat{\theta}_t$ ($L \times K$ matrix), observable at time t only. The out-of-sample measures in equation (8) are then calculated by replacing $\hat{\Gamma}_\beta$ with $\hat{\Gamma}_{\beta,t}$ and $\hat{F}_{t+1}$ with $\hat{F}_{t+1,t}$.

5 IPCA performance

We include all the stock and option level characteristics described in Section 3.1 in our baseline results. As is standard in IPCA analysis, we convert all characteristics to cross-sectional ranks from 0 to 1, and then take a normal inverse cumulative distribution function to convert these to z-scores (roughly between -3 and 3).[8] The advantage of normal inverse transformation is that, relative to ranks, it increases the weights of observations in the tails. This is important in our setting because decile returns to the strategies studied in Section 3.2 are monotonic for the most part but deliver extreme returns in the tails (see Internet Appendix Table IA2 for details). Using the z-score weights makes the resulting managed portfolio return (X using the notation of Section 4) closer to that of long-short portfolio.

It is unclear how to estimate an IPCA model on returns that are net of transaction cost. IPCA factors are portfolio combinations of managed portfolios. Ex-ante, it is not obvious whether a particular managed portfolio will have a positive or a negative position in a particular IPCA factor in a particular month; the sign of the position could be different for different IPCA factors and/or for different months. As such, it is not possible to use net option returns as inputs to the IPCA model. In line with all the studies using the IPCA method (for instance those of Babiak and Bianchi, 2023; Büchner and Kelly, 2022; Kelly, Palhares, and Pruitt, 2023; and Kelly, Pruitt, and Su, 2019), we sidestep the issue of incorporating transaction costs in the *estimation* of IPCA. We do, however, analyze factor returns and alphas from net of transaction cost returns.

5.1 In-sample

We start our analysis with a constrained IPCA where we force the alpha to be zero ($\Gamma_\alpha = 0$). Since theory provides little guidance on the number of factors, we perform the IPCA analysis for up to five latent factors. Table 2 tabulates total, time-series and cross-sectional R^2s as well as relative pricing errors from equation (8) for this specification. We report the statistics for individual equity option position as well as for the managed portfolios.

As is typical of other IPCA applications (see, for example, Kelly, Palhares, and Pruitt, 2023), we find that increasing the number of factors monotonically improves performance measures in-sample. The Total R^2 for individual option positions increases from 8.77% to 13.22% when going from 1 to 5 latent factors. The Time Series R^2 and Cross Section R^2 are of roughly the same magnitude as the Total R^2. As is usually the case, we also find that the performance measures are higher for managed portfolios than those for individual options. This result obtains partly because the managed portfolios have much lower cross-

[8] There are some variables (for example, volume) for which we observe the same values of Z across stocks. These observations are assigned the same rank and the transformed weight. For purely programming convenience, we assign a value of zero to missing Z. This has, obviously, no impact on any of the results as the option return, even if available, for a firm with missing Z is not incorporated in the managed portfolio return.

sectional and time-series dispersion in returns than individual delta-hedged option returns. For example, the Total R^2 for managed portfolios is 63.93% for 1-factor model and increases to 91.23% for 5-factor model.

Table 2: IPCA performance

The table presents performance measures of IPCA models for delta-hedged call returns. We report Total, Time Series, and Cross Section R^2, as well as the Relative Pricing Error from equation (8) for individual stock option positions and managed portfolios for the constrained ($\Gamma_\alpha = 0$) model. We also report p-values on the bootstrap Wald test of $\Gamma_\alpha = 0$. The last row of the top panel reports number of unpriced factors calculated as the number of estimated IPCA factors whose returns are statistically indistinguishable from zero. In the bottom panel, we report out-of-sample performance measures, obtained by splitting the sample in half and using the first half to estimate the model. We then use estimated factors and conditional betas to produce a one month out-of-sample forecast. We then roll the estimation procedure forward one month at a time, until the end of the sample. Results are reported for IPCA models with $K = 1$ to 5. Conditional betas are estimated for each of the characteristics used in Table 1. The sample period is 1996 to 2022.

	Number of factors				
	1	2	3	4	5
	In-sample				
Stock Level Option Positions:					
Total R^2	8.77	11.54	12.22	12.75	13.22
Time Series R^2	12.11	15.04	15.68	15.87	16.16
Cross Section R^2	8.69	10.84	11.38	11.86	12.27
Relative Pricing Error	78.46	73.68	72.74	71.80	71.20
Managed Portfolios:					
Total R^2	63.93	82.45	87.60	89.41	91.23
Time Series R^2	40.65	65.15	73.43	76.48	80.33
Cross Section R^2	53.45	70.50	75.70	78.38	80.58
Relative Pricing Error	24.46	0.38	0.23	0.18	0.15
Bootstrap test:					
W_α p-value	3.10	92.80	86.00	80.20	99.90
Economic test:					
# Unpriced Factors	0	0	0	1	3
	Out-of-sample				
Stock Level Option Positions:					
Total R^2	6.70	8.74	9.32	9.77	10.26
Time Series R^2	0.08	−6.76	−6.15	−9.12	−19.25
Cross Section R^2	6.65	8.32	8.87	9.28	9.77
Relative Pricing Error	81.75	77.73	76.99	76.39	75.32
Managed Portfolios:					
Total R^2	54.39	78.23	83.27	85.01	87.65
Time Series R^2	28.22	57.52	66.42	68.95	73.94
Cross Section R^2	47.19	63.92	69.62	72.14	75.00
Relative Pricing Error	48.16	9.47	5.26	6.06	5.16

The Relative Pricing Error is around 70% for individual option positions and remains relatively high even when we increase the number of factors. The Relative Pricing Error is high at 24.46% for managed portfolios for the case of 1 latent factor but declines to around 0.20% when the number of latent factors is bigger than 2. Overall, we conclude that, in-sample, IPCA is able to explain about 8%–10% of the variability in stock level option returns and about 80%–90% of the variation in managed portfolios returns.

The IPCA option return performance measures compare well with similar estimates for stocks (Kelly, Pruitt, and Su, 2019), corporate bonds (Kelly, Palhares, and Pruitt, 2023), and index equity options (Büchner and Kelly, 2022). Our measures are a bit lower (lower R^2s and higher relative pricing error) than those in these studies but are very similar to the respective measures extracted by applying IPCA to cryptocurrencies (Babiak and Bianchi, 2023).

Although we tabulate results only for the constrained version, we also estimate an unconstrained IPCA model where we place no restriction on alpha. Following the procedure described in Section 4.1.1, we calculate a Wald test and its bootstrapped p-value for the null hypothesis of zero alpha, and report this in the last row of Table 2. We find that the p-value is 3.10 for one latent factor, 92.80 for two factors, and remains high when more factors are added. Thus, we reject the hypothesis of zero alpha for the one factor specification, but we fail to reject it for any other specification. We conclude that at least two factors are needed to describe the variation in returns with only the time-varying loadings on factors without an intercept.

We also notice that IPCA models with higher K do not always result in factors that carry a significant average return. In other words, some of the estimated factors are not priced (have an average return that is statistically insignificantly different from zero). For each IPCA specification, we report the numbers of unpriced factors in the last row of Table 2. We refer to this as the economic test. For specifications with $K \geq 4$, we always find at least one factor with an average return that is not statistically different from zero. Combining the bootstrap test with the economic test, we conclude that a specification with 3 factors is preferable.

5.1.1 IPCA vs. PCA

Given the fact that IPCA extracts latent factors which are ultimately a combination of managed portfolios, one might wonder how the performance of IPCA factors compares to that of factors derived from a traditional PCA. We extract principal components from the set of 47 managed portfolios (i.e., one for each characteristics plus a constant). Conveniently, a visual inspection of the eigenvalues supports the idea that the first three principal components absorbs a large portion of the variability in the data (i.e., the first three eigenvalues capture approximately 85% of the total variance).

We compare the first 3 principal components (henceforth PCA factors), with the 3 IPCA factors of the benchmark model. To get a sense of how the two sets of factors are related to each other, we compute the canonical correlation coefficient (see Boivin and Ng, 2006) and find it to be 0.74. In other words, while the two sets of factors are positively related to each other, they seem to span a sufficiently different space.

Second, we run a conditional beta model similar to IPCA where the factors are pre-specfied as the first 3 PCA factors.[9] The conditional beta PCA model performs relatively well, thus highlighting the importance of conditional betas. However, the performance of the PCA factors is worse than that of the latent IPCA factors, thus highlighting the importance of joint estimation of Γ_β and factors in the IPCA estimation. For example, the Total R^2 for managed portfolios is 85.70% for the PCA factors, versus 87.60% for IPCA. The difference in relative pricing errors is more dramatic: it is 4.53% versus 0.23%, for PCA and IPCA factors respectively. Increasing the number of factors from 3 to 5 does not improve PCA's performance. For example, even with 5 PCA factors, the relative pricing error is still one order of magnitude larger (at 2.30%) than that of the 3-factor IPCA model.

These results are somewhat in line with those of Kelly, Pruitt, and Su (2019) who find that IPCA does better than PCA with static loadings for stock returns. Our analysis shows that the outperformance of IPCA for delta-hedged option returns carries over even when the PCA factors are allowed to have dynamic loadings.

5.2 Out-of-sample

The results in the previous sections are based on in-sample estimation where we estimate both the factors and the model parameters using the entire sample. This is the usual practice in asset pricing studies where unconditional betas (or parameters of conditional betas in conditional models) are estimated using the full sample. Nevertheless, given that our latent factors are also extracted using full-sample estimation, this could raise doubts about data over-fitting. Accordingly, we also consider how the IPCA model performs out-of-sample.

We calculate the statistics from equation (8) in an out-of-sample fashion as described

[9] It is easy to adapt the estimation machinery of Section 4 for pre-specified factors. See Kelly, Pruitt, and Su (2019) for details. Recall that the IPCA procedure initializes the factors using PCA factors and then computes Γ_β using these factors in the first iteration. Therefore, one can view the conditional beta PCA model as an IPCA model where only Γ_β is updated from one step of the iteration to the other.

in Section 4.2. The out-of-sample period is the second half of our sample, from 2009 to 2022 (we use expanding windows in the out-of-sample analysis). The results are tabulated in the bottom panel of Table 2. We find a modest increase in out-of-sample R^2s to their in-sample counterparts. For example, for $K = 3$, the Cross Section R^2 for managed portfolios is 75.70% in-sample and 69.62% out-of-sample. Time Series R^2 show similar decline in the out-of-sample relative to in-sample. Pricing errors for individual option positions also change only by modest amount in the out-of-sample test. Although the relative increase is more substantial for managed portfolios, out-of-sample Relative Pricing Errors for portfolios remain low. For example, with 3 latent factors, Relative Pricing Errors are 0.23% and 5.26%, in-sample and out-of-sample, respectively.

As noted before, some of the factors for $K \geq 4$ are statistically weak. It is interesting to see that, out-of-sample, relative pricing error for managed portfolios increases from 3 to 4 factors (but then declines again for 5 factors). This suggests that introducing weak factors has some consequences out-of-sample and provides some more comfort in our choice of $K = 3$ for the number of factors.

It may seem surprising that we do not see much deterioration in out-of-sample performance relative to in-sample performance. As Kelly, Palhares, and Pruitt (2023) argue, this is explained by the relatively few parameters that are estimated in IPCA. With 47 characteristics (including a constant), our three factor model requires 141 conditional beta parameters while a standard static beta model would require more than 100,000 parameter estimates for thousands of stock options.

5.3 Portfolio alphas

We now return to the long-short strategies presented in Table 1. The variables underlying these strategies are the conditioning variables in IPCA. In this section, we discuss how expected option returns extracted from the three factor IPCA model are useful at adjusting the performance of these strategies.

We calculate the average returns and the difference between the realized return and the IPCA expected return (i.e., the IPCA alpha). The expected return of the portfolio is constructed by aggregating the expected return of the individual option positions, at the stock level. For example, if the long leg is composed of 30 stocks in a particular month, then the expected return for that month is the average of the IPCA expected returns of those 30 stocks for the same month. Define the portfolio alpha as

$$\hat{\alpha}_p = \frac{1}{T} \sum_{t=1}^{T} \left(R_{p,t+1} - Z'_{p,t} \hat{\Gamma}_\beta \hat{F}_{t+1} \right), \tag{9}$$

where $Z_{p,t}$ is the $L \times 1$ vector of the portfolio p's characteristics at time t (calculated as weighted average of individual stock characteristics). Equation (9) shows that the alpha depends on estimated parameters $\hat{\Gamma}_\beta$ and estimated factors $\hat{F}$. Therefore, one needs to account for the estimation error in these model parameters to properly calculate the standard error of alpha estimates. We provide details on the procedure in Appendix B.

Table 3: IPCA alphas

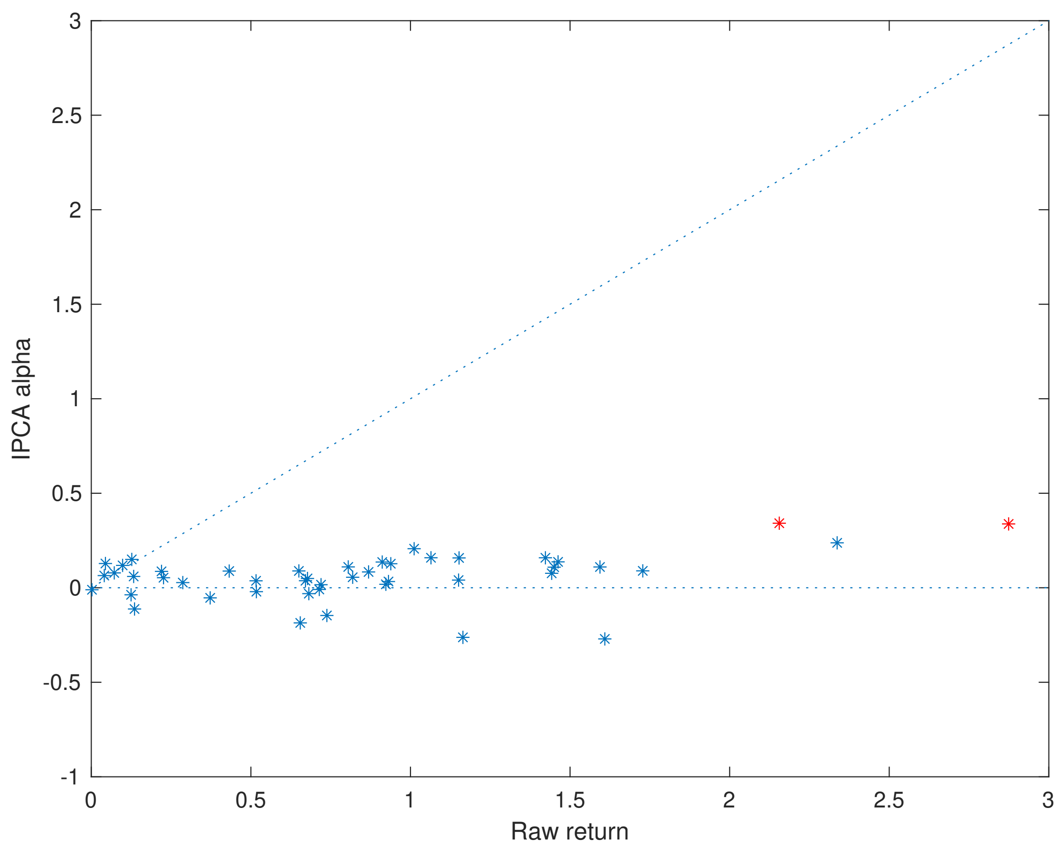
The table presents IPCA alphas and t-statistics in parentheses for long-short portfolios of delta-hedged call returns as described in Table 1. The IPCA model has 3 factors. Portfolios with t-statistics larger than the Benjamini and Hochberg (1995) 5% FDR adjusted threshold (i.e., 2.60 and 3.18 for mean return and alphas of long-short portfolios, respectively) are reported in boldface. The sample period is 1996 to 2022.

Moneyness	−0.05	(−0.52)	Max10	0.13	(1.24)
Bid-ask	−0.01	(−0.08)	MarketCap	0.11	(1.00)
Open interest	0.06	(0.86)	Autocorrelation	0.08	(1.34)
Delta	0.09	(0.90)	**RV−IV**	**0.34**	**(4.68)**
Vega	0.08	(0.58)	RV−MFvol	−0.27	(−1.24)
Gamma	0.04	(0.28)	Rskew−MFskew	−0.04	(−0.24)
Volume	0.05	(0.65)	Rkurt−MFkurt	0.21	(1.51)
Option price	−0.19	(−1.14)	BM	0.15	(2.10)
IV ATM	0.24	(2.11)	Profitability	0.02	(0.25)
IV slope	**0.34**	**(4.58)**	InstOwn	0.03	(0.46)
IV term	0.16	(2.11)	RSI	0.12	(1.71)
IV vol	0.11	(1.21)	Assets	0.14	(1.41)
MFvol	−0.26	(−1.14)	Debt	−0.03	(−0.36)
MFskew	0.11	(0.63)	Leverage	−0.01	(−0.15)
MFkurt	−0.15	(−0.82)	CashFlowVar	0.02	(0.18)
Stock price	0.09	(0.65)	Cash to asset	0.08	(1.14)
Stock return	0.13	(1.43)	AnalystDisp	-0.02	(-0.34)
Stock return11	0.03	(0.31)	1yr NewIss	0.09	(1.37)
RV	0.16	(1.41)	5yr NewIss	0.09	(1.34)
Rskew	−0.11	(−1.88)	Profit margin	0.04	(0.53)
Rkurt	0.14	(1.95)	ROE	0.06	(0.79)
Turnover	0.06	(0.70)	ExternalFin	0.04	(0.51)
IdiosynVol	0.16	(1.65)	Z score	0.05	(0.57)

Table 3 reports alphas and t-statistics. The average alpha across the 46 trading strategies is about 6 basis points, or equivalently about 7% of the average raw return. After controlling for MHT, 2 (5 at conventional levels) long-short portfolios still have a statistically significant IPCA alpha: IV slope and RV−IV. These strategies have a raw return of 2.15% and 2.87% respectively. Their IPCA expected returns are 1.81% and 2.53%, thus resulting in alphas of 34 basis points. In other words, the model explains more than 84% of the average realized return. From both a statistical and economic point of view, the performance of IPCA is quite remarkable.

Figure 1: IPCA alphas

The figure plots portfolio alphas from Table 3 against portfolio returns from Table 1 for long-short portfolios of delta-hedged call returns. The alpha is computed from the IPCA model with 3 factors. The sample period is 1996 to 2022.



We provide a visual exposition by plotting the portfolio alphas against their average realized returns in Figure 1. Alphas that are statistically significant at MHT levels are plotted in red star while alphas that are statistically indistinguishable from zero are plotted in blue stars. The figure shows that the IPCA model manages to eliminate most of the alphas, with most of the portfolios located on the horizontal line corresponding to zero alphas (as is already evident from the results in Table 3).

We also examine the alphas of strategies after accounting for transaction costs. Recall that Table 1 shows that 16 strategies still have statistically significant net of transaction cost returns. Internet Appendix Figure IA1 shows that net of transaction cost alphas from the

IPCA model are negative for all strategies.

We conduct several robustness checks. These include selecting particular cross-section of options based on liquidity and moneyness, pricing delta-hedged puts and straddles, and pricing month-end to month-end returns. We provide details of all these tests in the Internet Appendix Sections IA1-IA4.

There is an important caveat to our analysis, indeed to the IPCA model. As emphasized by Kelly, Pruitt, and Su (2019), the notion of risk in the statistical IPCA model is that of correlated fluctuations in the cross-section of asset returns arising from common IPCA factors. These factors may not necessarily correspond to those studied in macro-finance or q type models, and may arise from investor sentiment as in, for example, Kozak, Nagel, and Santosh (2018) and Stambaugh and Yuan (2016). Thus, IPCA factors capture non-diversifiable variation in returns regardless of their economic origin. As such, following standard terminology, while we use α as synonymous with anomalous returns, "there can be phenomena considered anomalous from the perspective of a canonical asset pricing theory that would appear as a risk factor in our analysis, so long as those phenomena produce covariation among large swathes of asset returns" Kelly, Pruitt, and Su (2019, pg. 506).

6 Interpreting IPCA

While our analysis so far shows that IPCA factors with dynamic conditional betas do well in explaining option returns, one limitation is that these factors are latent. In this section, we seek to understand what IPCA betas and factors might represent.

6.1 IPCA betas

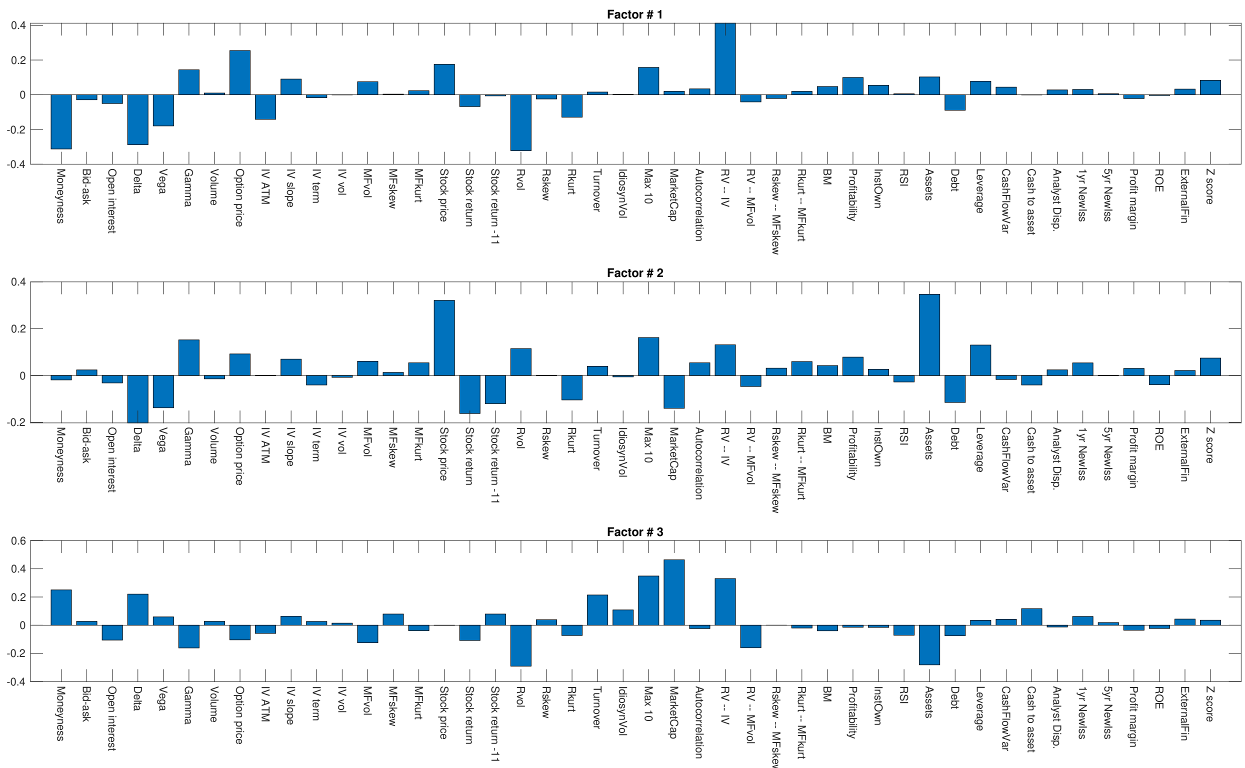
One advantage of IPCA is that it allows conditional betas to be a linear function of observable characteristics. This is summarized by the matrix of coefficients Γ_β. Define $\gamma_{\beta,l,k}$ to reflect the influence of the lth characteristic in determining the beta of an option position with respect to the kth factor. Figure 2 plots $\gamma_{\beta,l,k}$ for each factor in separate panels.

A quick look by singling out the absolute largest coefficient in each panel suggests a large role for RV−IV, total assets of the firm, and market capitalization, respectively for factors 1, 2, and 3. These characteristics are related to expected returns because they proxy for covariances/betas with respect to IPCA factors.

For example, in the context of the IPCA model, options with high RV−IV demand a higher expected return than those with low RV−IV because the former have high loadings on the first factor, everything else equal (the return to the long-short strategy on RV−IV in Table 1 is 2.87%). Stated differently, options with low RV−IV have low expected returns and high prices. Investors are willing to pay higher prices for options with low RV−IV because these securities are hedges (relative to options with high RV−IV) against bad realizations of the first factor.

Figure 2: IPCA gammas

The figure shows IPCA gammas for each of the 3 factors in the IPCA model from Table 2. The sample period is 1996 to 2022.



To shed some light on how IPCA functions, Table 4 shows several quantities of interest for all the RV−IV decile portfolios. Both realized and IPCA expected returns monotonically increase across the ten decile portfolios. The IPCA model prices very closely most of the 10 portfolios; only decile 1 has an alpha that fails to line up with the average return. Most of the variation in realized return across the ten portfolios is addressed by the first factor, where the spread in the betas is strongly related to the spread in RV−IV characteristic itself.

As another example, options on stocks with high total assets have high betas on the second IPCA risk factor, thereby, (at least partly) explaining the returns to the long-short strategy on Assets in Table 1 of 1.46%. However, the mapping from characteristics to betas is not always so straight-forward. For example, IV term does not have a high $\gamma_{\beta,l,k}$ on any of the three factors, yet has a return on 1.42% in Table 1. In unreported analysis, we find that deciles 1 and 10 of the IV term strategy have low and high betas on the first factor,

Table 4: RV−IV expected IPCA returns decomposition

The table presents a decomposition of the returns of decile portfolios constructed by sorting stocks based on the difference between realized and implied volatility (RV−IV). Expected return and IPCA alpha are derived from the IPCA model estimated with 3 factors from Table 3. The sample period is 1996 to 2022.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ret	−2.89	−1.14	−0.81	−0.59	−0.52	−0.37	−0.25	−0.20	−0.06	−0.02
E(Ret)	−2.41	−1.46	−1.04	−0.75	−0.59	−0.36	−0.22	−0.11	0.02	0.13
α_{IPCA}	−0.48	0.32	0.24	0.16	0.07	−0.01	−0.03	−0.09	−0.07	−0.14
β_{F1} E(F1)	−2.63	−1.70	−1.27	−0.99	−0.84	−0.61	−0.49	−0.40	−0.31	−0.27
β_{F2} E(F2)	0.20	0.20	0.19	0.19	0.19	0.19	0.20	0.20	0.22	0.25
β_{F3} E(F3)	0.02	0.04	0.04	0.04	0.06	0.06	0.07	0.09	0.11	0.15
RV−IV	−1.75	−1.04	−0.68	−0.39	−0.11	0.13	0.39	0.68	1.04	1.75

respectively, because the extreme deciles have large (in absolute terms) values of RV−IV. A somewhat similar example is the strategy based on Stock price that earns 1.73% in Table 1. Although stock price is a relatively important characteristic for the second factor in Figure 2, the variation in expected returns produced by the betas on the second factor is not enough to explain the realized returns. Rather, once again, we find that deciles 1 and 10 of the strategy on Stock price a spread in betas on the first factor because the extreme deciles have a spread in their RV−IV.

Although RV−IV is important for explaining many strategy returns, it is not the only story. An example is provided by the sorts on Profit margin that earn 1.15% in Table 1. Here it is the spread in RV leading to spread in betas on the first factor and to a slightly lesser extent, spread in Asset leading to spread in betas on the second factor that explains the variation in realized returns.

Since the $\gamma_{\beta,l,k}$ coefficient associated with each characteristic changes for each factor (and could even be of different signs across two factors), it is not possible to gauge the overall importance of each characteristic from the previous analysis. Nevertheless, following Kelly, Pruitt, and Su (2019), we compute a statistic that measures the impact of each characteristic by summing up the squares of its gammas (one for each factor) and then taking the square-root, $\sqrt{\sum_k \gamma^2_{\beta,l,k}}$. We evaluate the statistical significance of each statistic (there is one for each characteristic) using the bootstrap procedure described in Section 4.1.2. We report the impact statistics and their p-value in Table 5, and present results in descending order of the impact metric (i.e., larger statistic at the top). In general, characteristics with the largest impact metric are also more statistically significant. The characteristics with the larger impact metric that are also statistically significant are RV−IV, Assets, and Market cap. These are the same characteristics that are also visible in Figure 2. Therefore, options with high RV−IV, options on stocks with large assets, and options on stocks with high market capitalization, in general, command higher expected returns because these options have higher betas (everything else equal) on IPCA risk factors.

Table 5: Characteristics importance

The table presents measures of importance for each characteristic used in the IPCA estimation of Table 2 with 3 factors. Variables are sorted by the square root of the sum of squared Γ. We also report the p-value of the bootstrap test for each $\Gamma_\beta = 0$. The sample period is 1996 to 2022.

RV−IV	0.54 (0.00)	Open interest	0.12 (0.02)
MarketCap	0.49 (0.00)	Z score	0.12 (0.04)
Assets	0.46 (0.00)	IdiosynVol	0.11 (0.07)
RV	0.45 (0.05)	1yr NewIss	0.09 (0.02)
Max 10	0.42 (0.00)	MFskew	0.08 (0.55)
Delta	0.42 (0.03)	RSI	0.08 (0.15)
Moneyness	0.40 (0.02)	BM	0.08 (0.17)
Stock price	0.37 (0.08)	MFkurt	0.07 (0.60)
Option price	0.29 (0.69)	Autocorrelation	0.07 (0.02)
Gamma	0.26 (0.25)	Rkurt−MFkurt	0.07 (0.45)
Vega	0.23 (0.39)	CashFlowVar	0.06 (0.08)
Turnover	0.22 (0.01)	InstOwn	0.06 (0.14)
Stock return	0.21 (0.04)	ExternalFin	0.06 (0.11)
Rkurt	0.18 (0.00)	Profit margin	0.05 (0.20)
RV−MFvol	0.17 (0.36)	IV term	0.05 (0.22)
Debt	0.16 (0.06)	Bid-ask	0.05 (0.52)
MFvol	0.16 (0.48)	Rskew	0.05 (0.16)
Leverage	0.16 (0.07)	ROE	0.05 (0.20)
IV ATM	0.15 (0.45)	AnalystDisp	0.04 (0.17)
Stock return11	0.14 (0.15)	Rskew−MFskew	0.04 (0.69)
IV slope	0.13 (0.03)	Volume	0.03 (0.06)
Profitability	0.13 (0.04)	5yr NewIss	0.02 (0.62)
Cash to asset	0.12 (0.01)	IV vol	0.02 (0.79)

6.2 Which characteristics matter?

So far we have included the same characteristics in the IPCA estimation process as those that are used to construct the long-short strategies that we use to evaluate the model. This raises the question of whether a characteristic needs to be included in the IPCA model for the model to be able to price the associated long-short strategy. We investigate this next.

6.2.1 Removing insignificant characteristics

While motivated by existing literature, our choice of conditioning characteristic is still arbitrary. On the one hand, more information could be used, but, on the other, some of the information that we use appears to be redundant according to Table 5. Accordingly, we remove all characteristics that are statistically insignificant at the 5% level based on the metric in Table 5 and rerun the IPCA procedure using only the 17 significant characteristics as conditioning variables (plus a constant). We refer to the resulting model as the *restricted* IPCA model.

We evaluate the restricted IPCA model on two grounds: how the model prices single option level positions, and how the model prices long-short portfolios constructed using the characteristics that are left out of the estimation exercise. On the first count, the Total R^2 goes from 12.22% (third column of Table 2) in the original IPCA model to 11.90% in the restricted IPCA model with 17 characteristics. Similarly, the Relative Pricing Error increases from 72.74% to 74.60%. On the second count, we visualize the results in Panel A of Figure 3 by organizing the plot in a similar fashion to Figure 1. Of the 29 $(46 - 17)$ long-short strategies constructed on characteristics that are left out from the estimation, only 5 remain significant after controlling for MHT. The economic magnitude of the IPCA alpha is small compared to the raw returns: the average IPCA alpha is 4.5 basis points, while the average raw return is 80 basis. Thus, the deterioration (from full IPCA to restricted IPCA) in the ability to price non-included strategies is small. This exercise, thus, shows that the characteristics that are significant in Table 5 contain the bulk of the information that matters to explain the cross-section of delta-hedged returns.

6.2.2 Removing unimportant characteristics

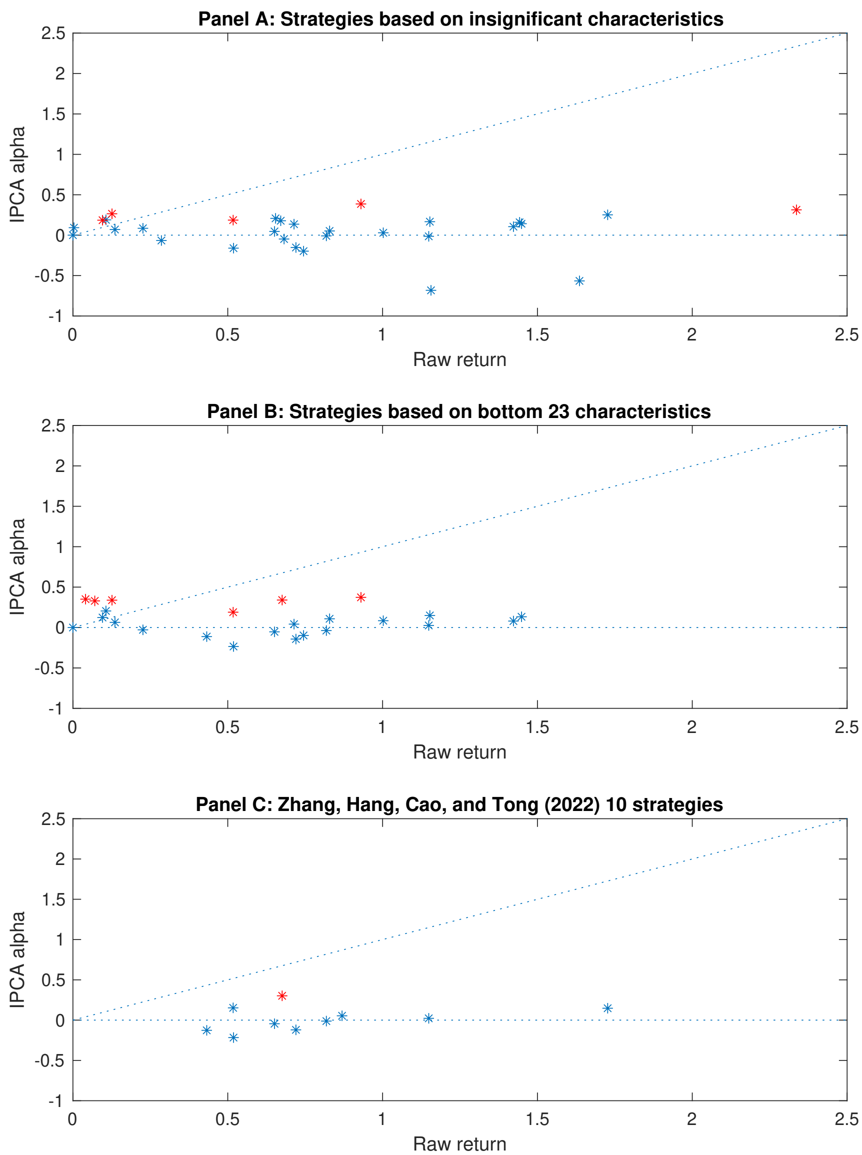
We next evaluate whether the magnitude of the impact estimates reported in Table 5 matters. This contrasts with the previous exercise where we only considered statistical significance.

We include the 23 most important characteristics (the ones on the right side of Table 5) in the estimation model and evaluate the strategies constructed based on the remaining 23 characteristics (the ones on the left-hand-side of Table 5). We find that the overall performance of the resulting restricted IPCA model remains strong: the relative pricing error of the restricted IPCA model is 73.90%. We again visualize the performance of the 23 strategies that are left out of the estimation in Panel B of Figure 3. We find that these 23 long-short strategies have an average IPCA alphas of about 10 basis points, relative to an average raw return of 60 basis points.

Thus, comparing the reduction in alpha relative to raw return across Panel A and Panel B of Figure 3, we find that importance of the characteristics is better gauged by statistical significance than by the magnitude of the statistics reported in Table 5.

Figure 3: Characteristics subsamples

The figure shows IPCA alphas versus raw returns for three different set of characteristics. In each case, we retain some characteristics and estimate a restricted IPCA model which we use to price the long-short portfolios constructed with the characteristics that are not included in the estimation (i.e., target long-short portfolios). In Panel A we use the 17 statistically significant characteristics from Table 5 to estimate the model and the remaining 29 statistically insignificant characteristics to construct the target long-short portfolios. In Panel B we use the 23 characteristics with the highest impact score from Table 5 to estimate the model and the remaining 23 characteristics to construct the target long-short portfolios. In Panel C, we use the 10 characteristics from Zhan, Han, Cao, and Tong (2022) to construct the target portfolios and the remaining 36 to estimate the IPCA model. The sample period is 1996 to 2022.



6.2.3 Removing ex-post characteristics

Since we use characteristics that other papers have advanced, one concern is that the success of the IPCA model is due to the fact that we already know that those characteristics work. The analysis in the previous section aids partly at addressing this issue. A related concern is that McLean and Pontiff (2016) show deterioration in strategy performance after publication. Thus, one might wonder whether the IPCA model is successful at explaining option returns simply because option return predictability has become weaker.

We address these issues by focusing on the 10 stock characteristics in Zhan, Han, Cao, and Tong (2022) (numbered 5.9 to 5.18 in Appendix A). We focus on these characteristics because (i) they are all grouped in one paper, (ii) the paper was the last one published in a top finance journal, and (iii) the characteristics contain stock information that, at first glance, may not be related to the main drivers of delta-hedged option returns (all of them do generate significant long-short returns as detailed in Table 1). We remove these 10 characteristics from the estimation exercise and re-estimate a restricted IPCA model with the remaining 36 characteristics. We visualize the results in Panel C of Figure 3. Only one long-short portfolio has a significant alpha (after controlling for MHT). Thus, similar to the results in previous two subsections, the IPCA model is capable of explaining the returns of trading strategies that are based on characteristics that are not included in the estimation.

This particular set up also allows us to also partially address the impact of publication bias. The sample period in Zhan, Han, Cao, and Tong (2022) is January 1996 to April 2016. We refer to this period as the in-sample period and the period between May 2016 and December 2012 as the out-of-sample period. We report returns and restricted IPCA alphas in Table 6 for the full sample and each of the two sub-samples.

The top panel of Table 6 shows raw returns across the three sample periods. We see that the returns in the out-of-sample period (average 1.43% across 10 characteristics) are larger than those in the in-sample period (average 0.61% across 10 characteristics). In fact, for each of the 10 strategies, we find that the returns are higher in-sample than those out-of-sample.

The bottom panel of Table 6 shows that the pattern carries over to alphas: alphas in the out-of-sample period are higher than those in the in-sample period. However, none of the alphas is significant at MHT levels in either of the two sub-samples (the strategy based on Z-score has a significant alpha at conventional levels in the out-of-sample period). From an economic point of view, the portion of return left unexplained by the IPCA alpha in the out-of-sample period is quite small at 20% (0.28/1.43).

Note, that our test is quite stringent. The restricted IPCA model does not use the 10 characteristics at all (neither in the publication in-sample period nor in the out-of-sample period). Therefore, the performance of the IPCA model in explaining the average realized returns, at least for the 10 characteristics of Zhan, Han, Cao, and Tong (2022), cannot be attributed to post-publication bias.

Table 6: Pricing the 10 portfolios in Zhan, Han, Cao, and Tong (2022)

We exclude the 10 characteristics in Zhan, Han, Cao, and Tong (2022) from the set of covariates and re-estimate the IPCA model on the remaining 36 (plus a constant). We refer to this estimated model as the "restricted IPCA." We then compute average raw returns and alphas in three sample periods: the full sample (January 1996 to December 2022), in-sample (January 1996 to April 2016) corresponding to the sample studied in Zhan, Han, Cao, and Tong (2022), and out-of-sample (from May 2016 to December 2020). t-statistics larger than the Benjamini and Hochberg (1995) 5% FDR adjusted threshold are reported in bold. MHT thresholds for raw returns are 2.22, 2.06, and 3.11, for each of the three periods respectively. MHT thresholds for restricted IPCA alphas are 3.12, 3.21, and 3.29, for each of the three periods respectively.

	Full sample 1996-2022		In-sample 1996-2016/5		Out-of-sample 2016/6-2022	
	Raw returns					
Stock price	**1.73**	**(16.06)**	**1.55**	**(14.14)**	**2.27**	**(8.30)**
CashFlowVar	**0.72**	**(9.10)**	**0.50**	**(6.19)**	**1.41**	**(7.42)**
Cash to asset	**0.87**	**(9.41)**	**0.60**	**(6.13)**	**1.69**	**(8.47)**
AnalystDisp	**0.52**	**(8.17)**	**0.50**	**(6.68)**	**0.59**	**(4.80)**
1yr NewIss	**0.43**	**(5.57)**	**0.17**	**(2.14)**	**1.25**	**(7.10)**
5yr NewIss	**0.65**	**(8.31)**	**0.36**	**(5.10)**	**1.55**	**(7.54)**
Profit margin	**1.15**	**(13.35)**	**0.95**	**(11.17)**	**1.77**	**(7.94)**
ROE	**0.82**	**(11.07)**	**0.69**	**(9.04)**	**1.20**	**(6.58)**
ExternalFin	**0.52**	**(6.09)**	**0.26**	**(2.78)**	**1.31**	**(8.07)**
Z score	**0.68**	**(7.98)**	**0.49**	**(5.72)**	**1.24**	**(5.86)**
	Restricted IPCA alphas					
Stock price	0.15	(1.11)	0.05	(0.39)	0.43	(1.33)
CashFlowVar	−0.12	(−1.40)	−0.26	(−3.23)	0.31	(1.32)
Cash to asset	0.05	(0.68)	−0.03	(−0.33)	0.31	(1.80)
AnalystDisp	0.15	(2.61)	0.16	(2.36)	0.13	(1.12)
1yr NewIss	−0.13	(−1.97)	−0.22	(−3.43)	0.17	(1.06)
5yr NewIss	−0.04	(−0.69)	−0.16	(−2.53)	0.32	(1.95)
Profit margin	0.02	(0.30)	−0.04	(−0.46)	0.21	(1.26)
ROE	−0.01	(−0.18)	−0.08	(−1.08)	0.20	(1.27)
ExternalFin	−0.22	(−2.97)	−0.30	(−3.75)	0.05	(0.35)
Z score	**0.30**	**(3.13)**	0.18	(1.77)	0.69	(2.96)

We acknowledge that our approach does not completely rule out concerns that published papers are data mined. However, the robustness of results of the restricted IPCA model does indicate that the mechanical advantage of IPCA in explaining a characteristic-based trading strategy by using the same characteristic as a conditioning variable is not that big.

6.3 IPCA factors

We next turn our attention to IPCA factors. Although these factors are latent, we attempt to characterize their properties. All factors (labeled F1, F2, and F3) have positive average return by construction. The average return of F1 is 2.06% per month, while the remaining two factors have average returns of 0.26% and 0.11%, respectively. The volatility of the factors goes from 1.38% for F1 to 0.87% for F3. The annualized Sharpe ratio of the first factor is the highest at 5.1. The pairwise correlations between factors are all negative, with the correlation between F1 and F2 being the most negative at -27%.

Factors that can price all assets with zero alpha are also the factors that can be combined to obtain the maximum Sharpe ratio (Gibbons, Ross, and Shanken, 1989). We calculate Sharpe ratios of the tangency portfolio obtained by combining the three IPCA factors both in-sample and out-of-sample. The in-sample tangency portfolio calculation uses the sample average and the sample covariance matrix for portfolio weights. The out-of-sample calculation follows the same procedure as described in Section 4.2. In every month t during the out-of-sample period, we calculate the portfolio weights, $\hat{\theta}_t$, using sample moments calculated with data through t. The tangency portfolio return at time $t+1$ is then given by weights $\hat{\theta}_t$ and out-of-sample IPCA factor return, $\hat{F}_{t+1,t}$. We find that the IPCA tangency portfolio Sharpe ratio is 5.8 in-sample and only slightly lower at 4.4 out-of-sample. For comparison, the annualized Sharpe ratio of the equal-weighted delta-hedged option market portfolio is 1.5 in our sample period and about 0.8 when positions are weighted by dollar open interest. Recall that our estimation method uses gross returns to estimate IPCA factors. Armed with the portfolio weights in managed portfolios, we also calculate a tangency portfolio adjusting returns for transaction costs. We find that the Sharpe ratio for the tangency portfolio using net returns is 4.1 in-sample and 2.7 out-of-sample, and thus still economically large. Thus, the IPCA factors span a wide region of the option-return space. At the same time, we remind readers that IPCA factors are designed to capture systematic variation in cross-section of delta-hedged option returns and, thus, are not necessarily related to factors in economic models.

We next provide a statistical interpretation of the factors. Specifically, we link the realization of the IPCA factors to a few proxies for adverse conditions, and compare them to the rest of the sample (i.e., we tabulate the t-statistic for the difference in mean in the factor return in the months defined by the event relative to the rest of the sample). We report results in Table 7.

We first consider how the IPCA factors relate to the negative aggregate shocks proposed by Karakaya (2013) and Büchner and Kelly (2022). We consider months in which the market portfolio return is lower than the 33rd percentile of the sample distribution (Low MktRet), or the change in the VIX index from the previous month is larger than the 66th percentile of the sample distribution (High ΔVIX). Table 7 shows that F1 has lower realizations in states where the market return is negative or VIX increases; the average factor return in those states is lower than the rest of the sample, although the difference is not statistical significant in the case of market return. F2 and F3 behave as a hedge against these negative shocks: The average return in these two negative states is approximately 50 and 25 bps higher than that of the full sample, for F2 and F3 respectively, with the differences being highly statistically significant.

Table 7: IPCA factors

The table presents sample statistics for the 3 IPCA factors under different scenarios. We first consider months in which the market portfolio return is lower than the 33rd percentile of the sample distribution (Low Mkt Ret) or the change in the VIX index from the previous month is larger than the 66th percentile of the sample distribution (High ΔVIX). We consider states of the world using higher order moments of S&P500 returns: (a) level of the tail risk measure of Kelly and Jiang (2014) is higher than the 66th percentile (High Tail Risk); (b) difference of realized and implied volatility is lower than the 33th percentile (High Vol Risk); (c) difference of realized and implied skewness is higher than the 66th percentile (High Skew Risk); (d) and difference of realized and implied kurtosis is lower than the 33th percentile (High Kurt Risk). We also consider instances when the intermediary capital risk factor of He, Kelly, and Manela (2017) is lower than the 33rd percentile (Low ICRac), or the cross-sectional equal-weighted average of options gamma and vega are higher than the 66th percentile (High Avg Gamma and High Avg Vega). The last two states of the world are defined using beginning period values (the other states of the world are defined using realizations during the period). We report the t-statistic for the difference in mean in the factor return in the months defined by the event relative to the rest of the sample. The sample period is 1996 to 2022.

	Obs	Mean	t-stat (Diff)	Mean	t-stat (Diff)	Mean	t-stat (Diff)
		F1		F2		F3	
Full sample	323	2.06		0.26		0.11	
Low Mkt Ret	107	1.88	(−1.43)	0.74	(3.52)	0.37	(3.98)
High ΔVIX	107	1.79	(−2.26)	0.75	(3.72)	0.35	(4.31)
High Tail Risk	107	1.72	(−3.12)	0.44	(1.31)	−0.14	(−3.42)
High Var Risk	107	2.83	(7.18)	−0.10	(−3.43)	−0.15	(−3.33)
High Skew Risk	107	1.63	(−3.81)	0.30	(0.29)	−0.08	(−2.56)
High Kurt Risk	107	1.61	(−4.21)	0.16	(−1.00)	−0.08	(−2.60)
Low ICRFac	107	2.13	(0.63)	0.59	(2.57)	0.03	(−0.91)
High Avg Gamma	107	1.57	(−4.95)	−0.09	(−3.40)	0.11	(0.02)
High Avg Vega	107	1.83	(−1.77)	0.47	(1.62)	−0.29	(−5.50)

Second, because delta-hedged options have payoffs related to large movements in the underlying, we relate the factor returns to "extreme event risk," as in Kelly and Jiang (2014). We study how our IPCA factors are linked to the tail risk measure, and to differences between risk-neutral (model free) variance, skewness, and kurtosis extracted from S&P500

options with approximately one month to maturity (similar to the maturity of options in the dela-hedged option returns), and their corresponding realized future counterparts (e.g., the realized variance between the date on which option prices are sampled to compute implied moments and the expiration date of the same option contracts). The differences between realized and implied moments, thus, capture the return of volatility, skewness, and kurtosis swaps, respectively, à la Carr and Wu (2009) and Kozhan, Neuberger, and Schneider (2013).

As before we consider terciles: tail risk measure higher than the 66th percentile (High Tail Risk), difference of realized and implied volatility is lower than the 33th percentile (High Vol Risk), difference of realized and implied skewness is higher than the 66th percentile (High Skew Risk), and difference of realized and implied kurtosis is lower than the 33th percentile (High Kurt Risk). We find that F1 and F3 have lower realizations during periods when the tail risk measure is largest. F1 is also negatively related to high skewness and kurtosis risk, but is positively related to high volatility risk (which is not surprising as F1 is negatively related to ΔVIX). F2 is negatively related to high volatility risk. F3 is instead negatively related to all three dimensions of higher moments risk.

Another possible source of variation in factor returns is the availability of trading counterparts. We examine states of the world when the intermediary capital ratio factor of He, Kelly, and Manela (2017), ICRFac, is lower than its 33rd percentile. With the exception of F2 which is positively related to ICRFac, we do not find much variation of the factor returns related to intermediary capital constraints. Relatedly, a newer strand of literature discusses the temporary price impact of intermediaries' trading due to their hedging demand. For example, Baltussen, Da, Lammers, and Martens (2021) and Barbon and Buraschi (2022) document how negative gamma exposure creates intra-day momentum. Inspired by these papers, we consider instances of high Gamma or Vega, when the proper execution of delta-hedging might be challenging. We construct aggregate measures of Gamma and Vega by averaging option Greeks across all options in our sample. As before, we consider instances in which Gamma and Vega are higher than the 66th percentile of their respective distributions. Note that because Gamma is inversely related to volatility, High Avg Gamma spans different states of nature than ΔVIX. We find that F1 and F2 are negatively related to instances when Avg Gamma is high, while F3 is negatively related to instances when Avg Vega is high.

Thus, we find that the three factors span different states of nature characterized by different dimensions of risk. As mentioned earlier, we urge caution in over-interpreting these results as indicating that factors returns represent risk-premia related to adverse conditions.

7 Conclusion

Explaining the returns from option positions is a daunting task. Option returns are extremely volatile and carry sizable liquidity premia. We unify the knowledge accumulated in the extant literature about cross-sectional characteristics that predicts option returns and recent advances in asset pricing methods, and show how much of the profitability of delta-hedged option returns can be explained away by an IPCA factor model with conditional betas.

While both the latent factors and dynamic betas are necessary to explain option returns, one theme that emerges from our analysis is the importance of the characteristic RV−IV. We find that not only a strategy based on RV−IV produces high returns of 2.87% per month, but also that the characteristic RV−IV is related to covariances with respect to risk factors. While the importance of a characteristic related to volatility such as RV−IV is not entirely surprising in explaining returns to volatility-sensitive delta-hedged call positions, what our analysis highlights is that RV−IV captures commonality in many other predictors of option returns. Setting the covariance coefficients, Γ_β, related to RV−IV to zero greatly decreases the ability of the IPCA model to price any portfolio that is not related to RV−IV: for example, 32 strategies see an increase in alpha and that increase is 1% on average.

Our results are important because they reaffirm the idea that market efficiency is still a valid framework to think about how prices form in markets. The fact that this is true even in a market that is highly segmented and riddled with informational advantages, as is the stock option market, is quite remarkable.

Appendices

A Variable construction

1. We construct measures related to contract characteristics as follows:

   1.1. Moneyness: Moneyness defined as the ratio of strike to underlying price.

   1.2. Bid-ask spread: Bid-ask spread of the contract on the initiation day of the strategy.

   1.3. Open interest: Dollar open interest of the contract (i.e., open interest defined as number of contracts times the most recent option price).

   1.4. Delta: Option delta.

   1.5. Vega: Option vega.

   1.6. Gamma: Option gamma.

   1.7. Volume: Dollar option volume on the initiation day of the strategy (i.e., volume defined as number of contracts times the most recent option price).

   1.8. Option price: Option price calculated as the midpoint of the bid and ask quotes.

2. We construct measures related to risk neutral distribution of returns as follows:

   2.1. IV ATM: The ATM implied volatility extracted from the 30 days volatility surface.

   2.2. IV slope: The difference between the OTM implied volatility (i.e., ratio of strike to underlying equal to 0.8) extracted from the 30 days volatility surface and the corresponding ATM implied volatility.

   2.3. IV term: The difference between the ATM implied volatility extracted from the 360 days volatility surface and the corresponding implied volatility extracted from the 30 days surface as in Vasquez (2017).

   2.4. IV vol: Volatility of implied volatility from the volatility surface of OptionMetrics. We use implied volatility of calls with 30 days to maturity with a delta of 0.5 and calculate the standard deviation using daily data over the last month with a minimum of 15 days.

   2.5. MFvol: model-free implied volatility is constructed from 30 days OTM call and OTM put option prices as in Bakshi, Kapadia, and Madan (2003). In particular we follow the procedure described in Hansis, Schlag, and Vilkov (2010), whose code is available on Grigory Vilkov's page.

   2.6. MFskew: model-free implied skewness is constructed from 30 days OTM call and OTM put option prices as in Bakshi, Kapadia, and Madan (2003). In particular we follow the procedure described in Hansis, Schlag, and Vilkov (2010), whose code is available on Grigory Vilkov's page.

2.7. MFkurt: model-free implied kurtosis is constructed from 30 days OTM call and OTM put option prices as in Bakshi, Kapadia, and Madan (2003). In particular we follow the procedure described in Hansis, Schlag, and Vilkov (2010), whose code is available on Grigory Vilkov's page.

3. We construct measures related to physical distribution of returns as follows:

   3.1. Stock return: Monthly stock return.

   3.2. Stock return11: Stock return over the last 11 months skipping the most recent month.

   3.3. RV: Volatility of log returns calculated using daily data over the last 12 months with a minimum of 150 observations.

   3.4. Rskew: Skewness of log returns calculated using daily data over the last 12 months with a minimum of 150 observations.

   3.5. Rkurt: Kurtosis of log returns calculated using daily data over the last 12 months with a minimum of 150 observations.

   3.6. Turnover: Ratio of number of shares traded over the last month to total shares outstanding.

   3.7. IdiosynVol: Standard deviation of residuals from Fama and French (1993) three-factor model. We use daily data over the last month with a minimum of 10 days.

   3.8. Max10: Average of the 10 highest daily returns over the last 3 months, following Bali, Cakici, and Whitelaw (2011).

   3.9. Autocorrelation: Autocorrelation of returns calculated using daily data over the last 6 months with a minimum of 100 observations, following Jeon, Kan, and Li (2021).

4. We construct measures related to differences between physical and risk-neutral distribution of returns as follows:

   4.1. RV−IV: difference between realized and option implied volatility.

   4.2. RV−MFvol: difference between realized and model free volatility.

   4.3. Rskew−MFskew: difference between realized and model free skewness.

   4.4. Rkurt−MFkurt: difference between realized and model free kurtosis.

5. We construct stock level variables as follows:

   5.1. BM: Book-to-market ratio is calculated as the ratio of book value of equity, calculated as in Fama and French (1992) to the current market value of equity.

   5.2. Profitability: Profitability, as in Novy-Marx (2013), is calculated as the ratio of gross profits (`GP`) to total assets (`AT`).

   5.3. InstOwn: Institutional ownership percentage (`instown_perc`) from Thomson Reuters 13f holdings.

5.4. MarketCap: Market capitalization.

5.5. RSI: The ratio of shares that are sold short (`shortintadj` in Compustat short interest file) to the total shares outstanding, as in Ramachandran and Tayal (2021).

5.6. Assets: Total assets (`AT`).

5.7. Debt: Total debt defined as the sum of long-term debt (`DLTT`) and debt in current liabilities (`DLC`).

5.8. Leverage: Ratio of financial debt to assets.

5.9. CashFlowVar: Cash flow variance, as in Haugen and Baker (1996), is computed as the variance of the monthly ratio of cash flow to market value of equity over the last 60 months. Cash flow is net income (`IB`) plus depreciation and amortization (`DP`).

5.10. Cash to asset: The cash-to-assets ratio, as in Palazzo (2012), is the value of corporate cash holdings (`CHE`) over the value of the firm's total assets (`AT`).

5.11. AnalystDisp: Analyst earnings forecast dispersion, as in Diether, Malloy, and Scherbina (2002), computed as the standard deviation (`stdev`) of annual earnings-per-share forecasts (`EPS1`) scaled by the absolute value of the average outstanding forecasts (`meanest`).

5.12. 1yr NewIss: 1-year new issues, as in Pontiff and Woodgate (2008), measured as the change in log of shares outstanding (`shrout`) from 11 months ago. Shares outsanding are adjusted for splits using the cumulative adjustment factor to adjust shares (`cfacshr`).

5.13. 5yr NewIss: 5-year new issues, as in Daniel and Titman (2006), measured as five-year real change in log of number of shares outstanding.

5.14. Profit margin: Profit margin, as in Soliman (2008), calculated as earnings before interest and tax (`OIADP`) scaled by revenues (`SALE`).

5.15. Stock price: The log of stock price at the end of last month, as in Blume and Husic (1972).

5.16. ROE: Profitability, as in Fama and French (2006), calculated as earnings divided by book equity, in which earnings is defined as income before extraordinary items (`IB`), and book equity is calculated as in Fama and French (1992).

5.17. ExternalFin: Total external financing, as in Bradshaw, Richardson, and Sloan (2006), calculated as net share issuance plus net debt issuance minus cash dividends, scaled by total assets (`AT`). Net share issuance is computed as net cash received from the sale (and/or purchase) of common and preferred stock (`SSTK` less `PRSTKC`) less cash dividends paid (`DV`). Net debt issuance represents net cash received from the issuance (and/or reduction) of debt (`DLTIS` less `DLTR` plus `DLCCH`). `DLCCH` is set to zero if missing.

5.18. Z score: Z-Score, as in Dichev (1998) constructed as [1.2×(Working Capital/Assets) + 1.4×(Retained Earnings/Assets) + 3.3×(EBIT/Assets) + 0.6×(Market Value of Equity/Book Value of Total Liabilities) + (Revenues/Assets)]. Working capital is the difference between ACT and LCT, Retained earnings are RE, EBIT is OIADP, Revenues are SALE, Total Liabilities are LT, Assets are AT, and the market value of equity is the market capitalization at fiscal year-end.

B IPCA Alpha Standard Error Calculation

Consider a portfolio with returns R_{t+1} and characteristics Z_t. The alpha of this portfolio is defined in equation (9) as $\hat{\alpha} = \sum_t \left(R_{t+1} - Z_t' \hat{\Gamma}_\beta \hat{F}_{t+1} \right) / T$. Define $\hat{\alpha}_{t+1} = R_{t+1} - Z_t' \hat{\Gamma}_\beta \hat{F}_{t+1}$. Then we have:

$$\begin{aligned} \hat{\alpha}_{t+1} &= R_{t+1} - Z_t' \hat{\Gamma}_\beta \hat{F}_{t+1} \\ &= \left(R_{t+1} - Z_t' \overline{\Gamma}_\beta \hat{F}_{t+1} \right) - Z_t' E_\beta \hat{F}_{t+1}, \end{aligned} \tag{B1}$$

where we have accounted for the fact that the $L \times K$ matrix $\hat{\Gamma}_\beta$ has estimation error specified by E_β. The standard error of $\hat{\alpha}$ needs to correct for the estimation error in $\hat{\Gamma}_\beta$ and $\hat{F}$. Obtaining an analytical expression for the the variance of $\hat{\alpha}$ is a difficult, if not impossible, task. We make some simplifying conservative assumptions below to make some headway in this issue.

First, ignoring the estimation error in the latent factor estimates, we have:

$$\begin{aligned} \text{var}(\hat{\alpha}_{t+1}) &= \text{var}\left(R_{t+1} - Z_t' \overline{\Gamma}_\beta F_{t+1} \right) + \text{var}\left(Z_t' E_\beta F_{t+1} \right) \\ &\quad -2\text{cov}\left(R_{t+1} - Z_t' \overline{\Gamma}_\beta F_{t+1}, Z_t' E_\beta F_{t+1} \right) \\ &= \text{var}\left(R_{t+1} - Z_t' \overline{\Gamma}_\beta F_{t+1} \right) + \text{var}\left(Z_t' E_\beta F_{t+1} \right). \end{aligned} \tag{B2}$$

where, in the second line, we have made use of the second simplifying assumption that E_β is uncorrelated with R, Z, and F.

Note that $Z_t' E_\beta F_{t+1} = \sum_{l=1}^{L} Z_{t,l} E_{\beta,l}' F_{t+1}$, where $Z_{t,l}$ is the lth characteristic at time t and $E_{\beta,l}$ is $K \times 1$ vector of estimation errors in Γ_β estimates for the lth characteristics. We now make the third simplifying assumption that the estimation error in betas across different characteristics are uncorrelated. We then have:

$$\text{var}\left(Z_t' E_\beta F_{t+1} \right) = \sum_{l=1}^{L} \text{var}\left(Z_{t,l} E_{\beta,l}' F_{t+1} \right). \tag{B3}$$

Since $Z_{t,l} E_{\beta,l}' F_{t+1} = \sum_{k=1}^{K} Z_{t,l} E_{\beta,l,k} F_{t+1,k}$, we have:

$$\begin{aligned} \text{var}\left(Z_{t,l} E_{\beta,l,k} F_{t+1,k} \right) &= \text{var}(Z_{t,l} F_{t+1,k}) \times \left[\text{var}(E_{\beta,l,k}) + \overline{Z_{t,l} F_{t+1,k}}^2 \right] \\ \text{cov}\left(Z_{t,l} E_{\beta,l,k_1} F_{t+1,k_1}, Z_{t,l} E_{\beta,l,k_2} F_{t+1,k_2} \right) &= \text{cov}(E_{\beta,l,k_1}, E_{\beta,l,k_2}) \times \overline{Z_{t,l}^2 F_{t+1,k_1} F_{t+1,k_2}}. \end{aligned} \tag{B4}$$

Combining equations (B2) to (B4), we get:

$$\begin{aligned}\text{var}(\hat{\alpha}_{t+1}) &= \text{var}\left(R_{t+1} - Z_t'\overline{\Gamma}_\beta F_{t+1}\right) \\ &+ \sum_{l=1}^{L}\left\{\sum_{k=1}^{K}\text{var}(Z_{t,l}F_{t+1,k}) \times \left[\text{var}(E_{\beta,l,k}) + \overline{Z_{t,l}F_{t+1,k}}^2\right]\right. \\ &\left.+ \sum_{k_1=1}^{K}\sum_{k_2=1,k_2\neq k_1}^{K}\text{cov}(E_{\beta,l,k_1}, E_{\beta,l,k_2}) \times \overline{Z_{t,l}^2 F_{t+1,k_1}F_{t+1,k_2}}\right\}. \end{aligned} \quad \text{(B5)}$$

Finally, the standard error of $\hat{\alpha}$ is given as usual by $\sqrt{\text{var}(\hat{\alpha}_{t+1})/T}$. Equation (B5) requires variances of covariances of E_β. Fortunately, these are already available via our bootstrap procedure from Section 4.1. The other quantities in equation (B5) are replaced by their sample counterparts.

References

An, Byeong-Je, Andrew Ang, Turan G. Bali, and Nusret Cakici, 2014, The Joint Cross Section of Stocks and Options, *Journal of Finance* 69, 2279–2337.

Aretz, Kevin, Ming-Tsung Lin, and Ser-Huang Poon, 2023, Moneyness, Underlying Asset Volatility, and the Cross-Section of Option Returns, *Review of Finance* 27, 289–323.

Babiak, Mykola, and Daniele Bianchi, 2023, A Risk-based Explanation of Cryptocurrency Returns, Working paper, SSRN https://ssrn.com/abstract=3935934.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options, *Review of Financial Studies* 16, 637–659.

Bali, Turan G., Heiner Beckmeyer, Mathis Moerke, and Florian Weigert, 2023, Option Return Predictability with Machine Learning and Big Data, *Review of Financial Studies* 36, 3548–3602.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing Out: Stocks as Lotteries and the Cross-section of Expected Returns, *Journal of Financial Economics* 99, 427–446.

Bali, Turan G., and Scott Murray, 2013, Does Risk-Neutral Skewness Predict the Cross Section of Equity Option Portfolio Returns? *Journal of Financial and Quantitative Analysis* 48, 1145–1171.

Baltussen, Guido, Zhi Da, Sten Lammers, and Martin Martens, 2021, Hedging Demand and Market Intraday Momentum, *Journal of Financial Economics* 142, 377–403.

Barbon, Andrea, and Andrea Buraschi, 2022, Gamma Fragility, Working paper, SSRN https://ssrn.com/abstract=3725454.

Benjamini, Yoav, and Yosef Hochberg, 1995, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289–300.

Blume, Marshall E., and Frank Husic, 1972, Price, Beta, and Exchange Listing, *Journal of Finance* 28, 283–299.

Boivin, Jean, and Serena Ng, 2006, Are More Data Always Better for Factor Analysis? *Journal of Econometrics* 132, 169–194.

Boyer, Brian H., and Keith Vorkink, 2014, Stock Options as Lotteries, *Journal of Finance* 69, 1485–1527.

Bradshaw, Mark T., Scott A. Richardson, and Richard G. Sloan, 2006, The Relation Between Corporate Financing Activities, Analysts' Forecasts and Stock Returns, *Journal of Accounting and Economics* 42, 53–85.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2009, Understanding Index Option Returns, *Review of Financial Studies* 22, 4493–4529.

Büchner, Matthias, and Bryan T. Kelly, 2022, A Factor Model for Option Returns, *Journal of Financial Economics* 143, 1140–1161.

Byun, Suk-Joon, and Da-Hea Kim, 2016, Gambling Preference and Individual Equity Option Returns, *Journal of Financial Economics* 122, 155–174.

Cao, Jie, and Bing Han, 2013, Cross-Section of Option Returns and Idisoyncratic Stock Volatility, *Journal of Financial Economics* 108, 231–249.

Carr, Peter, and Liuren Wu, 2009, Variance Risk Premium, *Review of Financial Studies* 22, 1311–1341.

Christoffersen, Peter, Ruslan Goyenko, Kris Jacobs, and Mehdi Karoui, 2018, Illiquidity Premia in the Equity Options Market, *Review of Financial Studies* 31, 811–851.

Chung, Kee H., Junbo Wang, and Chunchi Wu, 2019, Volatility and the Cross-section of Corporate Bond Returns, *Journal of Financial Economics* 133, 397–417.

Coval, Joshua D., and Tyler Shumway, 2001, Expected Option Returns, *Journal of Finance* 56, 983–1009.

Daniel, Kent, and Sheridan Titman, 2006, Market Reactions to Tangible and Intangible Information, *Journal of Finance* 61, 1605–1643.

Dichev, Ilya D., 1998, Is the Risk of Bankruptcy a Systematic Risk? *Journal of Finance* 53, 1131–1147.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of Opinion and the Cross Section of Stock Returns, *Journal of Finance* 57, 2113–2141.

Duarte, Jefferson, Christopher S. Jones, Mehdi Khorram, and Haitao Mo, 2023, Too Good to Be True: Look-ahead Bias in Empirical Option Research, Working paper, SSRN https://ssrn.com/abstract=4590083.

Fama, Eugene F., and Kenneth R. French, 1992, The Cross-section of Expected Stock Returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2006, Profitability, Investment and Average Returns, *Journal of Financial Economics* 82, 491–518.

Fama, Eugene F., and Kenneth R. French, 2015, A Five-factor Asset Pricing Model, *Journal of Financial Economics* 116, 1–22.

Galai, Dan, and Ronald Masulis, 1976, The Option Pricing Model and the Risk Factor of Stock, *Journal of Financial Economics* 3, 53–81.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A Test of the Efficiency of a Given Portfolio, *Econometrica* 57, 1121–1152.

Goyal, Amit, and Alessio Saretto, 2009, Cross-section of Option Returns and Volatility, *Journal of Financial Economics* 94, 310–326.

Goyenko, Ruslan, and Chengyu Zhang, 2022, The Joint Cross Section of Option and Stock Returns Predictability with Big Data and Machine Learning, Working paper, SSRN https://ssrn.com/abstract=3747238.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical Asset Pricing via Machine Learning, *Review of Financial Studies* 33, 2223–2273.

Hansis, Alexandra, Christian Schlag, and Grigory Vilkov, 2010, The Dynamics of Risk-Neutral Implied Moments: Evidence from Individual Options, Working paper, SSRN https://ssrn.com/abstract=1470674.

Harvey, Campbell, Yan Liu, and Alessio Saretto, 2020, An Evaluation of Alternative Multiple Testing Methods for Finance Application, *Review of Asset Pricing Studies* 10, 199–248.

Haugen, Robert A., and Nardin L. Baker, 1996, Commonality in the Determinants of Expected Stock Returns, *Journal of Financial Economics* 41, 401–439.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary Asset Pricing: New Evidence from Many Asset Classes, *Journal of Financial Economics* 126, 1–35.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh, 2016, The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications, *Journal of Financial Economics* 119, 249–283.

Heston, Steven L., Christopher S. Jones, Mehdi Khorram, Shuaiqi Li, and Haitao Mo, 2023, Option Momentum, *Journal of Finance* 78, 3141–3192.

Horenstein, Alex, Aurelio Vasquez, and Xiao Xiao, 2022, Common Factors in Equity Option Returns, Working paper, SSRN https://ssrn.com/abstract=3290363.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2021, An Augmented q-Factor Model with Expected Growth, *Review of Finance* 25, 1–41.

Hu, Guanglian, and Kris Jacobs, 2020, Volatility and Expected Option Returns, *Journal of Financial and Quantitative Analysis* 55, 1025–1060.

Jeon, Yoontae, Raymond Kan, and Gang Li, 2021, Stock Return Autocorrelations and Expected Option Returns, Working paper, SSRN https://ssrn.com/abstract=3363331.

Karakaya, Mete, 2013, Characteristics and Expected Returns in Individual Equity Options, University of Chicago.

Kelly, Bryan T., and Hao Jiang, 2014, Tail Risk and Asset Prices, *Review of Financial Studies* 27, 2841–2871.

Kelly, Bryan T., Diogo Palhares, and Seth Pruitt, 2023, Modeling Corporate Bond Returns, *Journal of Finance* 78, 1967–2008.

Kelly, Bryan T., Seth Pruitt, and Yinan Su, 2019, Characteristics are Covariances: A Unified Model of Risk and Return, *Journal of Financial Economics* 134, 501–524.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting Factor Models, *Journal of Finance* 73, 1183–1223.

Kozhan, Roman, Anthony Neuberger, and Paul Schneider, 2013, Skew Risk Premia in the Equity Index Market, *Review of Financial Studies* 26, 2174–2203.

Lettau, Martin, and Markus Pelger, 2020, Factors That Fit the Time Series and Cross-Section of Stock Returns, *Review of Financial Studies* 33, 2274–2325.

Lin, Hai, Junbo Wang, and Chunchi Wu, 2011, Liquidity Risk and Expected Corporate Bond Returns, *Journal of Financial Economics* 99, 628–650.

McLean, R. David, and Jeffrey Pontiff, 2016, Does Academic Research Destroy Stock Return Predictability? *Journal of Finance* 71, 5–32.

Muravyev, Dmitriy, 2016, Order Flow and Expected Option Returns, *Journal of Finance* 71, 673–708.

Muravyev, Dmitriy, and Neil D. Pearson, 2020, Option Trading Costs Are Lower Than You Think, *Review of Financial Studies* 33, 4973–5014.

Muravyev, Dmitriy, Neil D. Pearson, and Joshua M. Pollet, 2022, Is There a Risk Premium in the Stock Lending Market? Evidence from Equity Options, *Journal of Finance* 77, 1787–1828.

Novy-Marx, Robert, 2013, The Other Side of Value: The Gross Profitability Premium, *Journal of Financial Economics* 108, 1–28.

Palazzo, Berardino, 2012, Cash Holdings, Risk, and Expected Returns, *Journal of Financial Economics* 104, 162–185.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share Issuance and Cross-sectional Returns, *Journal of Finance* 63, 921–945.

Ramachandran, Lakshmi Shankar, and Jitendra Tayal, 2021, Mispricing, Short-sale Constraints, and the Cross-section of Option Returns, *Journal of Financial Economics* 141, 297–321.

Ruan, Xinfeng, 2020, Volatility-of-volatility and the Cross-section of Option Returns, *Journal of Financial Markets* 48, 1–26.

Shafaati, Mobina, Don M. Chance, and Robert Brooks, 2021, The Cross-Section of Individual Equity Option Returns, Working paper, SSRN https://ssrn.com/abstract=3927410.

Soliman, Mark T., 2008, The Use of DuPont Analysis by Market Participants, *Accounting Review* 83, 823–853.

Stambaugh, Robert F., and Yu Yuan, 2016, Mispricing Factors, *Review of Financial Studies* 30, 1270–1315.

Vasquez, Aurelio, 2017, Equity Volatility Term Structures and the Cross Section of Option Returns, *Journal of Financial and Quantitative Analysis* 52, 2727–2754.

Vasquez, Aurelio, and Xiao Xiao, 2023, Default Risk and Option Returns, *Management Science* forthcoming.

Zhan, Xintong (Eunice), Bing Han, Jie Cao, and Qing Tong, 2022, Option Return Predictability, *Review of Financial Studies* 35, 1394–1442.

Internet Appendix

IA1 Impact of liquidity

Christoffersen, Goyenko, Jacobs, and Karoui (2018), Muravyev (2016), and Muravyev and Pearson (2020) show that liquidity plays an important role in determining option prices and option realized returns. Therefore, we perform sensitivity analysis of the base case results in three different ways. First, we consider only options which have a positive trading volume on the position initiation day (recall that the base case considers options where either the volume or the open interest is positive). Second, we split the sample by the option dollar open interest. Third, we split the sample by the size of the option bid-ask spread.

We perform IPCA with three latent factors for each of the sub-sample and report the results of the analysis in Panel A of Table IA1. To not overwhelm the reader with too many numbers, we report only select performance metrics: Relative Pricing Error, average absolute alpha and number of significant alphas of long-short portfolios. Regarding IPCA fit, Panel A shows that fit (measured by R^2 or the relative pricing error) improves with liquidity. This is to be expected given the prior literature's findings of higher volatility of option returns for less liquid options.

We also find that the average alpha of long-short produced by risk-adjustment through IPCA is relatively stable across open interest positions and volume restrictions, and varies a little across different bid-ask spread groups. Equally noteworthy is the fact that the IPCA model leaves at most 3 significant alphas (in the low open interest sub-sample). Overall, we conclude that partitioning the data along measures that proxy for liquidity does not impact the ability of IPCA to control for risk, and leaves very few trading strategies with significant alphas.

IA2 Moneyness intervals

Our baseline analysis uses the options that are closest to ATM. Although the average moneyness (i.e., ratio of option strike to stock price) is in fact close to 1.0, the sample contains options which moneyness varies between 0.8 and 1.2. While our choice of this relatively large moneyness interval is in line with Zhan, Han, Cao, and Tong (2022), it is not the only approach in the literature. For example, Goyal and Saretto (2009) use a much tighter interval of between 0.975 and 1.025. Moreover, equity and index option returns show systematic variation along the moneyness dimension (Karakaya, 2013 and Büchner and Kelly, 2022). Therefore, it is worth investigating whether IPCA can account for such differences.

We check the robustness of our baseline choice by splitting the sample in different moneyness categories: between 0.8 and 0.9, between 0.9 and 0.975, between 0.975 and 1.025, between 1.025 and 1.1, and between 1.1 and 1.2. Due to the fact that strike prices are rather coarse, especially for stocks with low prices, some stocks have only one contract in the entire interval of 0.8 and 1.2. Other stocks have multiple positions, and thus potentially contribute one option position to each of the five sub-sample splits. Panel B of Table IA1 presents the results for 3-factor IPCA model.

As Table 1 suggests, delta-hedged option returns are decreasing with moneyness in our baseline sample. This is also true across the five moneyness groups: the average return (not reported in the table) in the first group (i.e., moneyness between 0.8 and 0.9) is −0.5% while the average return in the last group (i.e., moneyness between 1.1 and 1.2) is −1.0%. The standard deviation of option returns moves in the opposite direction (from 4% to 13% per month). The increase in volatility of option returns explains why the IPCA fit declines in general as we move from left to right columns (from low to high moneyness). For instance, the relative pricing error for managed portfolios increases from 0.37% to 1.06% across the five moneyness subsamples.

Turning to the analysis of alphas, we find that, as with average returns, the average absolute alpha increases as well from about 2bps to 9bps as we move along the moneyness subsamples. There is however an inverted U-shape pattern in the number of strategies that remain significant, with the peak being 4 for the strict ATM subsample (i.e., 0.975−1.025).

IA3 Alternative option positions

In the main analysis we consider delta-hedged call option (DHCall) returns as a way for investor to profit from volatility exposure of options. Delta-hedged put returns (DHPut) theoretically speaking should give the same return as DHCall, but in practice can differ because of demand induced price pressure (Ramachandran and Tayal, 2021). We repeat the IPCA analysis by considering DHPut and tabulate results in Panel C of Table IA1. Panel C shows that IPCA explains DHPut returns equally as well as DHCall returns in terms of relative pricing error. While the average absolute alpha is also similar for DHCall and DHPut, the number of statistically significant alphas at 5 is higher for DHPut.

We also consider a joint estimation model where we pool DHCall and DHPut. The IPCA model for the joint case has a unique set of factors, but the conditional betas are position specific (i.e., there are two sets of Γ_β, each associated with one of the possible option position). Thus, we have two sets of long-short portfolios for the joint sample. The third column of Panel C of Table IA1 shows that the joint estimation delivers intermediate results. For example the number of strategies that have significant alphas is 0 for the joint estimation.[10]

While delta-hedged positions are neutral to movements in the underlying, they are not the only way to obtain volatility exposure: ATM straddles are also commonly considered as, for example, in Vasquez (2017). Since straddle returns are much more volatile than DHCall returns, our preliminary checks show that straddles require five IPCA factors (in contrast to three factors as in the rest of the paper). The fourth column of Panel C of Table IA1 shows that Straddle returns are harder to explain for IPCA; they have larger pricing errors. Interestingly, however, there are no statistically significant alphas of Straddle returns.

[10]Since the IPCA factors in the joint estimation are different from those in individual estimation, there is no mechanical reason to expect that the number of statistically significant alphas for the joint estimation will be the same as the sum of statistically significant alphas in the individual estimations.

IA4 Alternative return construction

So far we have analyzed expiration to expiration option returns. However, some authors also consider returns that are constructed from month-end to month-end (see, for example, Zhan, Han, Cao, and Tong, 2022). Accordingly, we consider monthy option returns. The last column of Panel C of Table IA1 shows the results for these DHCall option positions. We find that managed portfolios show similar pricing performance. The number of statistically significant alphas for the IPCA model remains at 2 for these positions.

We conclude by noting that IPCA is a relatively flexible instrument, and can account for substantial differences in the data being explained. Nevertheless, we do urge a note of caution in interpreting the results in this appendix. The IPCA procedure is a data driven procedure: both the identification of the factors and the importance of the characteristics used to determine the structure of conditional betas depends on the inputs fed into the procedure. Therefore, if a researcher is interested in studying the cross-section of some other challenging positions such as strangles, calendar spreads, bull or bear spreads, etc., then one would need to run the IPCA on the specific type of position one is studying.

Table IA1: Robustness checks

The table presents results from IPCA estimation of subsamples of the data. Panel A imposes filtering constraints related to liquidity of the option contracts. The first column contains only the delta-hedged call returns for which the contract volume on the return origination date is greater than zero. The columns after this separate observations for which dollar open interest and bid-ask spread are above or below their respective cross-sectional medians (at the trade initiation date). Each estimation extracts four IPCA factors. Panel B restricts the moneyness of option contracts to narrower bands, than those used for the main analysis (i.e., closest moneyness to ATM in the 0.8 and 1.2 interval). We consider subgroups where we select the option contract that is closest to the middle of five intervals: between 0.8 and 0.9, between 0.9 and 0.975, between 0.975 and 1.025, between 1.025 and 1.1, and between 1.1 and 1.2. Panel C reports different option positions. In addition to ATM delta-hedged calls (DHCall), we consider ATM delta-hedged puts (DHPut), a joint estimation where we pool DHCall and DHPut (Joint), and ATM straddles (Straddle). Additionally we consider returns of DHCall constructed from month-end to month-end returns. IPCA on Straddle return uses five factors. All other models use three IPCA factors. In the joint estimation, there is a unique set of three factors, but conditional betas are position specific (i.e., there are two sets of Γ_βs, each associated with DHCall and DHPut). We report Relative Pricing Error for managed portfolios and alphas of long-short portfolios. The sample period is 1996 to 2022.

Panel A: Liquidity stratification

	Volume>0	Open Interest		Bid-Ask	
		Low	High	Low	High
Relative Pricing Error	0.24	0.32	0.16	0.13	0.68
Average Alpha	0.07	0.07	0.10	0.12	0.04
# Significant Alphas	2	2	2	1	2

Panel B: Moneyness stratification

	0.8–0.9	0.9–0.975	0.975–1.025	1.025–1.1	1.1–1.2
Relative Pricing Error	0.69	1.01	0.80	0.71	0.96
Average Alpha	0.01	0.04	0.06	0.12	0.15
# Significant Alphas	1	2	2	2	1

Panel C: Alternate option positions

	Expiration-to-expiration				MtoM
	DHCall	DHPut	Joint	Straddle	DHCall
Relative Pricing Error	0.25	0.17	0.27	3.42	0.25
Average Alpha	0.07	0.06	0.06	−0.09	0.09
# Significant Alphas	2	6	3	0	2

IA5 Additional tables and figures

Figure IA1: IPCA alphas accounting for transaction costs
The figure plots portfolio IPCA alphas after transaction costs against portfolio returns from Table 1 for long-short portfolios of delta-hedged call returns. The alpha is computed from the IPCA model with 3 factors. The sample period is 1996 to 2022.

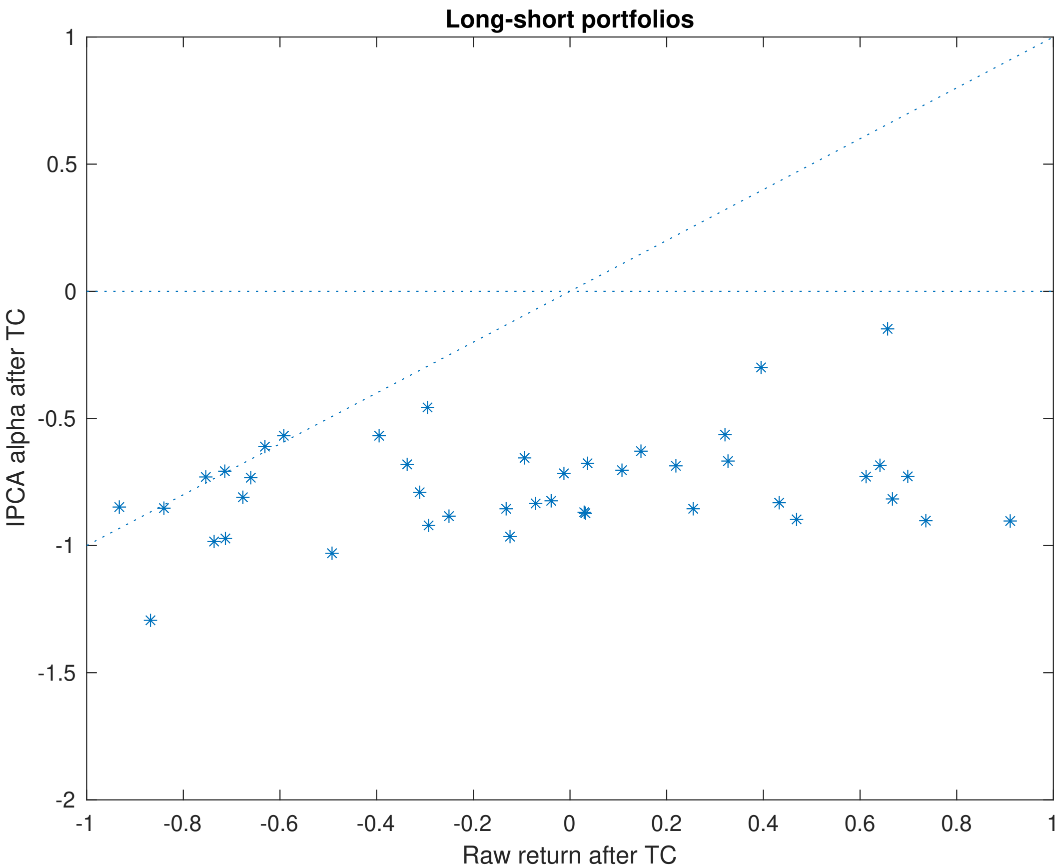


Table IA2: Deciles returns for delta-hedged portfolios

The table presents average returns for all deciles of the 46 trading strategies reported in Table 1. The sample period is 1996 to 2022.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Moneyness	−1.43	−0.69	−0.51	−0.45	−0.42	−0.46	−0.53	−0.61	−0.74	−1.05
Bid-ask	−1.06	−0.92	−0.82	−0.67	−0.62	−0.57	−0.46	−0.48	−0.42	−0.34
Open interest	−0.63	−0.71	−0.65	−0.70	−0.75	−0.74	−0.71	−0.74	−0.69	−0.60
Delta	−0.69	−0.75	−0.67	−0.60	−0.56	−0.66	−0.74	−0.86	−0.83	−0.47
Vega	−1.74	−1.06	−0.85	−0.70	−0.47	−0.51	−0.42	−0.41	−0.33	−0.30
Gamma	−1.08	−0.81	−0.80	−0.72	−0.72	−0.63	−0.59	−0.53	−0.53	−0.40
Volume	−0.73	−0.70	−0.69	−0.73	−0.70	−0.60	−0.63	−0.63	−0.58	−0.50
Option price	−0.98	−0.76	−0.70	−0.65	−0.53	−0.50	−0.44	−0.39	−0.34	−0.34
IV ATM	−2.57	−0.95	−0.62	−0.49	−0.45	−0.42	−0.39	−0.36	−0.32	−0.23
IV slope	−2.48	−1.19	−0.81	−0.77	−0.52	−0.56	−0.51	−0.42	−0.38	−0.32
IV term	−1.85	−0.89	−0.71	−0.60	−0.70	−0.49	−0.40	−0.36	−0.35	−0.43
IV vol	−1.78	−1.12	−0.81	−0.69	−0.54	−0.44	−0.41	−0.32	−0.36	−0.33
MFvol	−1.50	−0.74	−0.48	−0.32	−0.46	−0.58	−0.25	−0.35	−0.27	−0.33
MFskew	−1.17	−0.69	−0.52	−0.40	−0.43	−0.64	−0.24	−0.42	−0.27	−0.34
MFkurt	−1.13	−0.63	−0.44	−0.34	−0.77	−0.32	−0.34	−0.43	−0.26	−0.40
Stock price	−1.97	−0.98	−0.73	−0.52	−0.53	−0.46	−0.40	−0.36	−0.32	−0.24
Stock return	−0.99	−0.65	−0.63	−0.54	−0.57	−0.53	−0.62	−0.63	−0.71	−0.94
Stock return11	−0.92	−0.70	−0.75	−0.68	−0.69	−0.64	−0.65	−0.59	−0.57	−0.63
RV	−1.54	−0.86	−0.68	−0.62	−0.55	−0.55	−0.54	−0.53	−0.48	−0.48
Rskew	−1.17	−0.83	−0.59	−0.56	−0.58	−0.51	−0.46	−0.47	−0.64	−1.03
Rkurt	−1.31	−0.87	−0.85	−0.70	−0.66	−0.57	−0.54	−0.48	−0.44	−0.40
Turnover	−0.89	−0.68	−0.69	−0.61	−0.65	−0.62	−0.66	−0.62	−0.63	−0.76
IdiosynVol	−1.64	−0.96	−0.73	−0.59	−0.52	−0.50	−0.47	−0.45	−0.46	−0.48
Max10	−1.42	−0.77	−0.68	−0.65	−0.61	−0.59	−0.56	−0.52	−0.51	−0.49
MarketCap	−1.88	−1.01	−0.76	−0.64	−0.53	−0.47	−0.52	−0.40	−0.32	−0.29
Autocorrelation	−0.83	−0.71	−0.67	−0.67	−0.63	−0.66	−0.59	−0.62	−0.69	−0.76
RV−IV	−2.89	−1.14	−0.81	−0.59	−0.52	−0.37	−0.25	−0.20	−0.06	−0.02
RV−MFvol	−1.79	−0.76	−0.42	−0.41	−0.74	−0.34	−0.21	−0.15	−0.06	−0.20
Rskew−MFskew	−0.95	−0.56	−0.37	−0.31	−0.30	−0.75	−0.27	−0.34	−0.40	−0.82
Rkurt−MFkurt	−1.14	−0.79	−0.55	−0.47	−0.46	−0.71	−0.32	−0.28	−0.24	−0.14
BM	−0.94	−0.57	−0.55	−0.58	−0.62	−0.62	−0.67	−0.67	−0.77	−0.81
Profitability	−1.48	−0.61	−0.56	−0.65	−0.65	−0.58	−0.51	−0.54	−0.57	−0.55
InstOwn	−1.45	−0.89	−0.66	−0.58	−0.58	−0.50	−0.47	−0.48	−0.58	−0.52
RSI	−0.78	−0.65	−0.60	−0.62	−0.68	−0.64	−0.73	−0.72	−0.81	−0.68
Assets	−1.79	−0.80	−0.67	−0.62	−0.57	−0.53	−0.54	−0.51	−0.48	−0.33
Debt	−1.04	−0.83	−0.70	−0.69	−0.59	−0.57	−0.54	−0.56	−0.48	−0.36
Leverage	−0.85	−0.65	−0.67	−0.61	−0.57	−0.55	−0.58	−0.60	−0.71	−0.84
CashFlowVar	−1.14	−0.90	−0.69	−0.76	−0.70	−0.52	−0.51	−0.45	−0.46	−0.42
Cash to asset	−1.47	−0.78	−0.60	−0.52	−0.61	−0.61	−0.57	−0.50	−0.55	−0.60
AnalystDisp	−1.20	−1.14	−0.65	−0.48	−0.79	−0.44	−0.44	−0.49	−0.57	−0.68
1yr NewIss	−1.08	−0.79	−0.65	−0.60	−0.64	−0.61	−0.62	−0.58	−0.59	−0.65
5yr NewIss	−1.23	−0.87	−0.66	−0.59	−0.65	−0.57	−0.51	−0.54	−0.57	−0.58
Profit margin	−1.66	−0.63	−0.59	−0.72	−0.67	−0.50	−0.53	−0.46	−0.44	−0.51
ROE	−1.54	−0.86	−0.66	−0.61	−0.56	−0.52	−0.43	−0.43	−0.50	−0.72
ExternalFin	−1.16	−0.62	−0.68	−0.65	−0.72	−0.61	−0.61	−0.56	−0.57	−0.65
Z score	−1.45	−0.70	−0.63	−0.61	−0.59	−0.60	−0.51	−0.54	−0.57	−0.78