

The implied volatility smirk in SPY options

Wei Guo, Sebastian A. Gehricke, Xinfeng Ruan & Jin E. Zhang

To cite this article: Wei Guo, Sebastian A. Gehricke, Xinfeng Ruan & Jin E. Zhang (2021)
The implied volatility smirk in SPY options, *Applied Economics*, 53:23, 2671-2692, DOI:
[10.1080/00036846.2020.1866159](https://doi.org/10.1080/00036846.2020.1866159)

To link to this article: <https://doi.org/10.1080/00036846.2020.1866159>



Published online: 05 Jan 2021.



Submit your article to this journal 



Article views: 366



View related articles 



View Crossmark data 



Citing articles: 1 View citing articles 



The implied volatility smirk in SPY options

Wei Guo , Sebastian A. Gehrcke , Xinfeng Ruan and Jin E. Zhang

Department of Accountancy and Finance, Otago Business School, University of Otago, Dunedin, New Zealand

ABSTRACT

We provide a comprehensive study of the implied volatility (IV) smirk in the SPDR S&P 500 Exchange-Traded Fund (SPY ETF) option market. In general, the IV curves are downward sloping with little curvature, exhibiting an almost straight line. However, the shape of the IV curves becomes more curved during the global financial crisis (GFC) period, indicating that the commonly accepted IV smirk shape is driven by the GFC. In addition, based on in-sample, out-of-sample tests and asset allocation analysis, we show that the first difference of the slope factor can predict the next month's SPY excess returns.

KEYWORDS

Implied volatility (IV); IV smirks; SPY options; global financial crisis (GFC); prediction

JEL CLASSIFICATION

Code: G12 G13

1. Introduction

In this paper, we quantify the Implied Volatility (IV) curves for the SPDR S&P 500 Trust Exchange Traded Fund (ETF) options, which allows us to study the shape and dynamics of the IV curves. Further, we investigate whether the IV factors can predict future SPY ETF returns. Studying the IV curves' dynamics is important for risk management and asset allocation, since the IV curves contain risk information for predicting future market returns. We apply the method proposed in Zhang and Xiang (2008) to quantify the SPY IV curves, and provide a theory to explain the left-skewed shape during the global financial crisis (GFC) period. Then, we investigate IV curve factors' predictive ability for SPY returns.

ETF markets have become increasingly popular with investors in recent years. By the end of 2018, there are more than 1800 unique ETF products covering numerous conceivable market sectors and trading strategies, with a capital value of approximately 3.4 trillion dollars. The low cost and easy diversification of ETFs make them attractive tools to investors. Unlike stocks, ETFs provide access to diversified portfolios without needing to purchase individual stocks separately, therefore lowering the costs to hold highly diversified portfolios and investors can easily and cheaply enter and exit ETFs through trading on an exchange, unlike closed-end

funds. Among ETFs, the SPY ETF is the largest and most liquid ETF, as presented in **Table 1**, in terms of the capital value under management and trading volume. **Figure 1** presents the dynamics of the SPY ETF from January 2005 to December 2017. It shows an increasing trend with time, along with its increasing popularity. The movement of the SPY ETF is the same as that of the S&P 500 index, as it is a replication of the S&P 500 index.

Based on the prosperity of ETFs, the options written on ETFs have started to become popular in the early 21st century. Among all of the ETF option markets, SPY options are the most popular. The Chicago Board Options Exchange (CBOE) first issued options on the SPY ETF in January 2005. As presented in **Table 2**, the trading volume of SPY options is the largest among all ETF options in the year of 2017, with around 22 million contracts. SPY options normally have no more than 1 penny bid-ask spreads, which minimizes the transaction cost for investors to hedge or speculate on the S&P 500 index.

Figure 2 shows the trading volume and value of SPY options from January 2005 through December 2017. We can see that its trading volume and value show an upward trend since January 2005 through December 2017. The highest spike of the trading volume and value happens around June 2011 due to the European Sovereign Debt Crisis. When

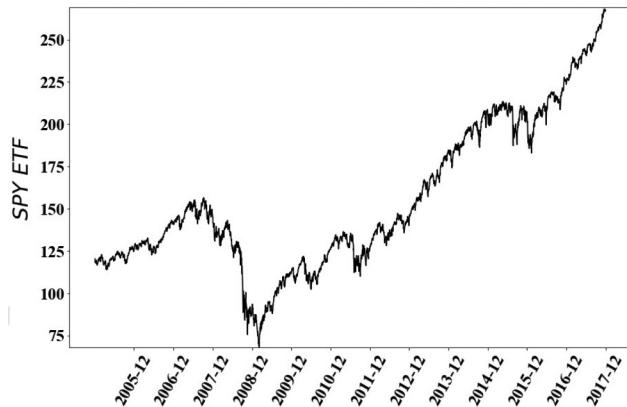


Figure 1. Daily SPY ETF. This figure shows the daily SPY ETF price adjusted for dividends from January 2005 to December 2017 from OptionMetrics.

Table 1. Summary for popular ETFs. This table summarizes the basic information for the top ETFs in terms of asset under management (AUM), which is calculated by the product of the outstanding shares and market price per share. All the information is summarized as of December 2017.

Symbol	Name	AUM	Underlying	Inception	Issuer
SPY	SPDR S&P 500 ETF	320B	S&P 500 Index	01/22/93	State Street
IVV	iShares Core S&P 500 ETF	208B	S&P 500 Index	05/15/00	Blackrock
VTI	Vanguard Total Stock Market ETF	144B	CRSP U.S. Total Market	05/24/01	Vanguard
VOO	Vanguard S&P 500 ETF	135B	S&P 500 Index	9/07/10	Vanguard
QQQ	Invesco QQQ	93B	Nasdaq 100 Index.	3/10/99	Invesco
VEA	Vanguard FTSE Developed Markets ETF	81B	stocks markets outside the US	07/20/07	Vanguard
IEFA	iShares Core MSCI EAFE ETF	75B	index of developed-market stocks	10/18/12	Blackrock

Table 2. Summary of SPX, SPY and other options. This table summarizes the properties of the main options traded in the global markets. The mean volume and open interest for each options are calculated over 2017.

	SPX	SPY	IVV	VOO	XEO	OEX	OEF	NDX
Underlying	S&P 500	SPY ETF	IVV ETF	VOO ETF	S&P 100	S&P 100	OEF ETF	Nasdaq 100
Underlying benchmark	S&P 500	S&P 500	S&P 500	S&P 500	S&P 100	S&P 100	S&P 100	Nasdaq 100
Exercise-Style	European Cash	American Shares	American Shares	American Shares	European Cash	American Cash	American Cash	European Shares
Settlement	No	Yes	Yes	Yes	No	No	Yes	No
Dividend								
Mean Volume (000's)(2017)	1,160.30	2,478.49	0.16	0.17	0.13	1.90	0.17	15.52
Mean Open Interest(000,s) (2017)	13,823.26	22,336.29	4.28	4.48	14.13	5.48	5.48	118.93
Issuer	CBOE	State Street's SPDR	BlackRock iShares	Vanguard	CBOE	CBOE	BlackRock iShares	NASDAQ
Underlying Issuer	S&P Global	S&P Global	S&P Global	S&P Global	S&P Global	S&P Global	S&P Global	Nasdaq

examining the dynamics of the trading volume and value for the SPY call and put options separately, they show similar dynamics as those of the total markets, but the trading volume and value for the put options are larger than those for the call options, indicating that the put options are more popular in the market.

There are a lot of studies on modelling EFT volatility. Chen and Huang (2010) use the Generalized Autoregressive Conditional Heteroscedasticity-Autoregressive Moving Average (GARCH-ARMA)

and the exponentially Generalized Autoregressive Conditional Heteroscedasticity-Autoregressive Moving Average (EGARCH-ARMA) to study the leverage effects on the volatility of index ETFs. Huang, Tong, and Wang (2020) further investigate the performance of the popular discrete-time volatility models in pricing SSE 50 ETF options. However, they focus on the validity of the models. In our paper, we study IV smirk of SPY options, in particular its shape and dynamics. We also explore how

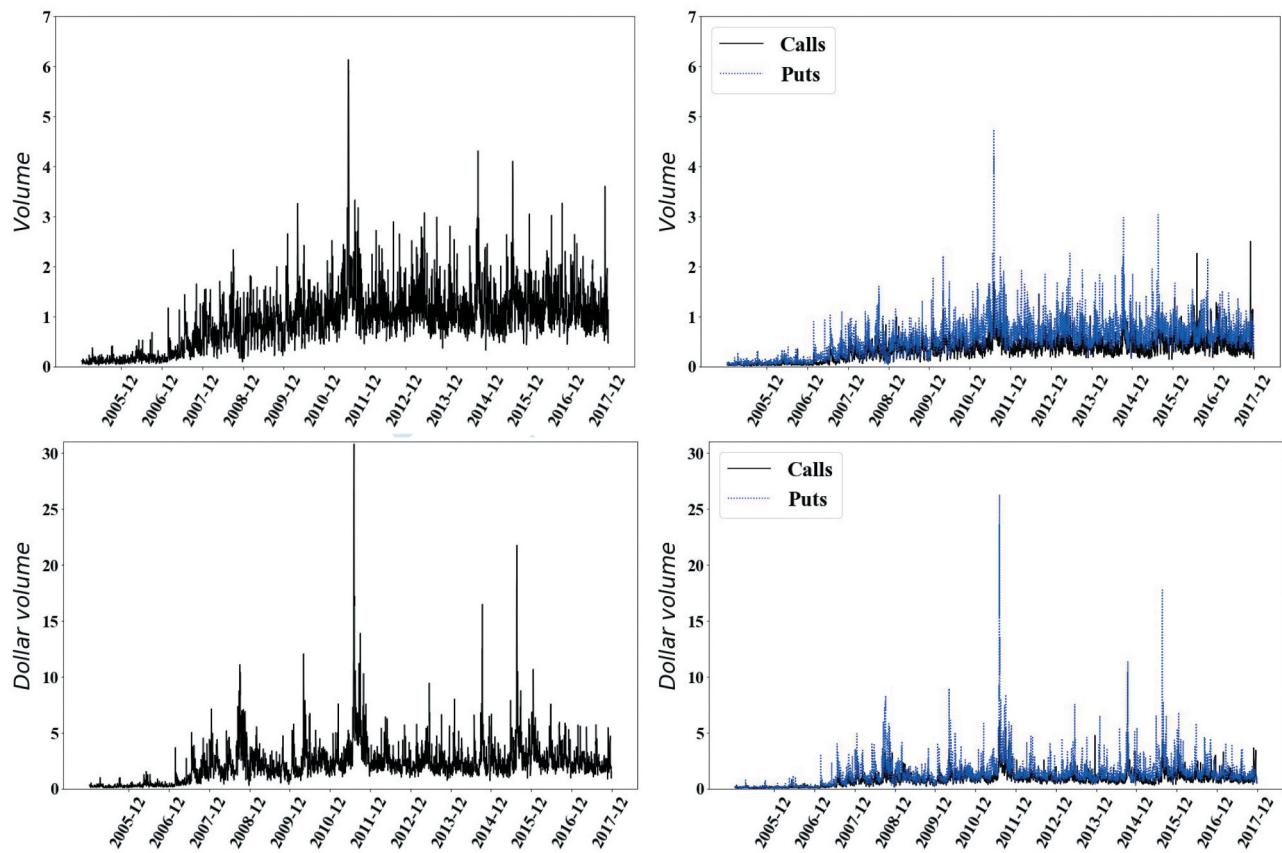


Figure 2. Daily trading activity of SPY options. This figure shows the total trading volume (in millions) and dollar trading value (in millions) for SPY options, the trading volume and dollar trading volume for call and put options for respectively.

SPY returns are related to their IV curve factors. There are exhaustive studies on the predictive power of IV curves on index returns and equity returns. For instance, Xing, Zhang, and Zhao (2010) study the cross-sectional predictive ability of a IV smirk measure on future equity returns. Ruan and Zhang (2018) study the time-series predictive power of risk-neutral moments in the crude oil market. In our paper, we focus on the predictive power of IV curve factors for SPY returns.

In this paper, we quantify the IV curves using a quadratic polynomial following the method developed by Zhang and Xiang (2008) and study the dynamics of IV curve factors with constant maturities. To have a clear look at the dynamics of the IV smirk during the GFC period, we also explore a subsample analysis. Then, we employ the theory of Zhang, Zhao, and Chang (2012) and Zhang, Chang, and Zhao (2020) to explain the significantly skewed shape of the IV curves during the GFC period. After that, we investigate the predictive power of the SPY IV curve factors on the monthly SPY excess returns. Beyond in

cluding the level, slope and curvature factors, we also create the third and fourth cumulants following Ruan and Zhang (2018), as predictors. In line with the procedure in Ruan and Zhang (2018), we test the predictive ability of each of the predictors for the excess SPY returns based on in-sample and out-of-sample results. We measure the economic value of each IV curve factors by investing a portfolio based on the factor's predictability following Rapach, Ringgenberg, and Zhou (2016). We further provide robustness checks along with control variables.

The main results are summarized as follows: 1) On average, the IV curves of the SPY ETF options exhibit a downward sloping line with some minor curvature. 2) The well accepted smirk shape of IV curves is only pronounced during the GFC period, but not during normal times and previous findings of the smirk shape may be driven by crisis periods. 3) The first difference of the slope factor, obtained from the IV curves, has a statistically and economically significant ability to predict the next-month excess SPY ETF returns. 4) Our results are robust to popular predictors.

Our paper contributes to several strands of literature. First, we contribute to the study of the IV smirk. A voluminous literature documents the IV smirks in index option markets. Pan (2002) documents the IV smirk in the S&P 500 market and proposes that the jump risk can explain the IV smirk. Carr and Wu (2003) document the S&P 500 IV smirk shape and find that the IV smirk exists in options with maturity of up to 2 years. Benzoni, Collin-Dufresne, and Goldstein (2011) document the IV smirk shape for equity index options, and they propose that the arrival of jumps triggers agents' belief about future jump, thus generating the IV smirk. In this paper, we document the overall average IV smirk shape for the SPY market, which is the most liquid ETF market; however, we find that this shape is largely driven by the GFC in our sample period. We employ the theory of Zhang, Zhao, and Chang (2012) and Zhang, Chang, and Zhao (2020) to show that the downward sloping IV curve shape could be due to the jump effect.

Second, we employ the method proposed by Zhang and Xiang (2008) for quantifying the IV curves in SPY market. This method has been used in other markets. Li, Gehricke, and Zhang (2019) apply the method to quantify the IV smirks in the FXI (iShares China Large-Cap ETF) market. Gehricke and Zhang (2019) also apply this method to quantify the IV smirks in the VXX (S&P 500 VIX Short-Term Futures ETN) market. Aschakulporn and Zhang (2020) further employ the method to quantify the New Zealand whole milk powder options.

Finally, our paper contributes to studies on predicting equity returns. There are a huge volume of studies on predicting equity returns (Ang and Bekaert 2007; Welch and Goyal 2008; Campbell and Thompson 2008; Bollerslev, Tauchen, and Zhou 2009; Cochrane 2011; Neely et al. 2014; McLean and Pontiff 2016; Lin 2018; and many more). Welch and Goyal (2008) study 14 popular predictors' predictability on equity returns, exploring in-sample and out-of-sample tests, and find that all of these predictors perform poorly. Other authors also re-examine the predictive ability of other macroeconomic factors from different perspectives. Ang and Bekaert (2007) specifically examine the predictive power

of dividend yield at different horizons, while Gourio (2012) looks at the disaster risk and business cycle effect on equity risk premia. In our paper, we focus on using the IV curve factors to predict SPY returns. Several papers explore using IV factors to predict asset returns (for example, Ruan and Zhang 2018; Li, Gehricke, and Zhang 2019; Aschakulporn and Zhang 2020), however none have followed this approach for SPY returns.

The remainder of the paper is organized as follows: Section 2 describes the data. Section 3 elaborates on the methodology. Section 4 shows the results, and Section 5 concludes.

2. Data

We obtain the SPY ETF and options data from the OptionMetrics, from 5 January 2005 through 31 December 2017. We calculate the monthly SPY returns using the continuously compounded return over the month inclusive of dividends, which are also obtained from OptionMetrics. The risk-free rate, proxied by the one-month U.S. Treasury rate, is obtained from U.S. Treasury website. All the monthly data is sampled at the last day of the month.

The price of options we use is the middle of the bid and ask quote prices. Before we start our analysis, we clean the options data following the existing literature, as below:

- We filter out the options with zero bid price and those with mid quotes less than 0.2 as in Chang, Christoffersen, and Jacobs (2013).
- We discard those options without information in the OptionMetrics database as in Neumann and Skiadopoulos (2013).
- We exclude the options with maturity of less than 7 days following the VIX white paper (<https://www.cboe.com/micro/vix/vixwhite.pdf>).
- We discard the option chains which have less than 3 strike prices, because at least three observations are required to employ our quantification method.

Table 3 reports a summary of the daily trading activities of SPY options after cleaning. The statistics include mean strike numbers, mean trading volume, mean open interest and mean maturity days. In general, the mean number of



Table 3. Summary of trading activity for SPY options. This table summarizes the mean number of strike prices, volume, open interest and number of observations for SPY options by maturity category. The results are shown in 6 maturity categories ranging from less than 30 days to more than 360 days.

Maturity (days)	Overall	< 30	30–90	90–180	180–360	> 360	< 180	> 180
Number of observations	6,009,931	734,838	1,205,049	1,120,688	1,795,911	1,153,445	3,068,684	2,941,247
Mean number of strikes	113	90	117	143	136	86	116	111
Median number of strikes	111	88	116	144	131	82	115	106
Mean volume	70,883	245,358	132,547	30,782	7576	3002	137,031	5238
Median volume	8180	108,247	58,757	18,158	2710	1192	45,814	1685
Mean open interest	664,944	1,147,062	1,202,816	792,976	334,255	208,665	1,062,912	270,000
Median open interest	257,048	387,017	715,724	431,278	145,703	123,217	514,502	132,040

option contracts across strike prices each day is 113. The maturity categories with less than 30-day maturity and 90–180 days to maturity have the largest strike numbers and trading volume, meaning these two maturity categories are the most liquid among all the maturity categories. In terms of mean open interest, the maturity with less than 30 days and 30–60 day to maturity are the two most popular categories. As most options, such as like SPX options (Bakshi and Kapadia (2003) and Broadie, Chernov, and Johannes (2009)) the SPY options with shorter time to maturity are more heavily traded.

3. Methodology

This section describes the methodology, proposed Zhang and Xiang (2008) to quantify the SPY IV curves. Further, we present a predictive regression analysis using the factors extracted from the IV curves to predict SPY returns. Lastly, we present our methodology for further exploring the predictive power of the IV factors using asset allocation tests.

SPY returns

Following Fama and French (1993), Chang, Christoffersen, and Jacobs (2013), Fama and French (2015) and Ruan (2019), we calculate the monthly excess return at t as:

$$R_t = \ln S_t - \ln S_{t-1} - r_t, \quad (1)$$

where S_t is the SPY ETF mid-quote price at the end of month t adjusted by dividends, r_t is the one-month risk-free interest rate proxied by the one-month U.S. Treasury bill yield. The one-month interest rate listed on the U.S. Treasury website is

the annualized number, the monthly interest rate is the number divided by 12, which is throughout our sample close to zero.

Quantifying IV curve

For a certain maturity date, Zhang and Xiang (2008) apply a second-order polynomial to quantify the IV across all strike prices, as given by,

$$IV(\xi) = \alpha_0 + \alpha_1 \xi + \alpha_2 \xi^2, \quad (2)$$

where α_0 , α_1 , α_2 , are the quantified IV curve coefficients. These can be transformed to the dimensionless factors to obtain the following quantified IV curve::

$$IV(\xi) = \gamma_0(1 + \gamma_1 \xi + \gamma_2 \xi^2), \quad (3)$$

where γ_0 , γ_1 , γ_2 , are the level, slope and curvature factors of the IV curve. Here, moneyness, ξ , is defined as,

$$\xi \equiv \frac{\ln(K/F_0)}{\bar{\sigma}\sqrt{\tau}}, \quad (4)$$

where F_0 is the forward price obtained from the put-call parity, K is the strike price, $\bar{\sigma}$ is the standard deviation, which is proxied by the VIX index. Lastly, τ is the annualized time period to maturity of the given contract. Here the definition of moneyness is adopted by Zhang and Xiang (2008) from Carr and Liuren (2003).

We employ two types of fitting methods, one is volume-weighted least squares, another is ordinary least squares. For the volume-weighted least square method, the coefficients can be solved by minimizing the volume-weighted mean square errors, given by:

$$f_{WLS} = \frac{\sum_{\xi} Vol(\xi)(IV_{MKT}(\xi) - IV_{FIT}(\xi))^2}{\sum_{\xi} Vol(\xi)}. \quad (5)$$

where $Vol(\xi)$ is the trading volume, $IV_{MKT}(\xi)$ is the market IV, $IV_{FIT}(\xi)$ is the fitted IV, for a given moneyness, by estimating equation (3). The contracts with higher volume should exhibit more reliable information, and should therefore be given more weights, so volume-weighted least square (WLS) is the main method we apply to fit equation (3), following Zhang and Xiang (2008). For some IV curves, the volume-weighted curve is not very good because of spotty trading. In those instances, the ordinary least square (OLS) method is more reliable. To determine in which situation ordinary least square should replace the volume-weighted least square method, the increase in the ordinary R^2 is examined.

Since the R^2 of WLS is always smaller than that for its corresponding OLS regression, we use a threshold approach to decide which regression to run to fit the IV smirks. When the R^2 for the WLS plus a threshold is larger than R^2 for the OLS, a WLS regression is applied to fit the second-order polynomial. The threshold number is determined by a trade-off between the proportion of the WLS regression and the fitting improvement. After a test on the threshold, we set the threshold number to be 10%.

Predictive regression analysis

In this section, we outline our predictive regression framework to test the explanatory ability of the factors calculated from IV curves. We test the in-sample and out-of-sample performance of each factor.

In-sample tests

A predictive time series model is run for SPY ETF returns under the standard framework as in Welch and Goyal (2008) and Rapach, Ringgenberg, and Zhou (2016),

$$R_{t:t+1} = \alpha + \beta x_t + \varepsilon_{t:t+1} \quad (6)$$

where $t = 1, 2, \dots, T - 1$, $R_{t:t+1}$ is the excess returns on the SPY ETF on the next month, x_t is the time series of predictor. As an example of the factor predictive power test, we run non-

overlapping regressions of SPY ETF returns on the independent factor, and the test for other

$$r_{t:t+1} = \alpha + \beta x_t + \varepsilon_{t:t+1}$$

forecasting horizons can be conducted in a same way. We examine the significance of the Newey and West (1987) adjusted t-statistic of β_{x_t} and adjusted R^2 to decide whether the predictor is significant in predicting the future SPY ETF returns.

Out-of-sample tests

As suggested in Welch and Goyal (2008), in-sample tests seem to produce robust results. They find that all predictors exhibiting predictability in in-sample tests show poor results in out-of-sample tests. The out-of-sample test is an efficient instrument to further test the results of in-sample tests, and provides some guidance for market investors. We set the last third of the sample period as the out-of-sample. Applying the methodology of Rapach, Ringgenberg, and Zhou (2016), we mainly test the magnitude of out-of-sample R_{OS}^2 as a measure of performance in predicting the excess SPY ETF returns.

For each of the predictors, the predicted value for the out-of-sample regression expression at time $t + h$ conditional on time t is,

$$\hat{R}_{t,t+h} = \hat{\alpha} + \hat{\beta} x_t, \quad (7)$$

where α and β are the coefficients obtained from the regression estimates for factor x on the period from the beginning through month t . As the equation (7) indicates, the predicted value for the return on $t + h$ follows the trend obtained from the beginning of the sample period through t . In general, the movement of the return conditional on time t follows a martingale as in Bakshi, Kapadia, and Madan (2003).

After obtaining all the estimated values for the out-of-sample period, we calculate the R_{OS}^2 with a similar method as the in-sample analysis. The R_{OS}^2 for the out-of-sample period starting from time t_0 is,

$$R_{OS}^2 = 1 - \frac{\sum_{t=t_0}^{t_n} (R_{t+1} - \hat{R}_{t+1|t})^2}{\sum_{i=t_0}^{t_n} (R_{t+1} - \bar{R}_{t+1|t})^2}, \quad (8)$$

where R_t is the real return at time t , \hat{R}_t is the prediction of the return at time t , and $\bar{R}_{t+1|t} =$



$\frac{1}{t} \sum_{i=1}^t R_i$ is the mean of the excess returns from the beginning through time t .

The R^2 for in-sample tests is the adjusted R^2 in the Newey-West regression with six lags. The out-of-sample R^2 measures the performance of the predictor with the prevailing mean forecast as a benchmark. In line with Clark and West (2007), the MSPE statistic for the prevailing mean regression model and the predictor regression model is defined as,

$$f_{t+1} = (R_{t+1} - \bar{R}_{t+1|t})^2 - ((R_{t+1} - \hat{R}_{t+1|t})^2 - (\bar{R}_{t+1} - \hat{R}_{t+1|t})^2)^2, \quad (9)$$

where $\bar{R}_{t+1|t}$ is the forecasting value from the prevailing mean forecast model, $\hat{R}_{t+1|t}$ is the forecasting value from the specified factor model. We present the significance of the t -statistic by regressing the f_t on a constant. The positive t statistic implies the factor predictive model performs better than the historically averaging forecast model.

Asset allocation analysis

To assess the significance of different factors to predict the future excess returns, we construct a mean-variance portfolio consisting of the SPY ETF and a risk-free asset. At the end of month t , we calculate the weight on equity for the next month as follows:

$$\omega_t = \frac{1}{\gamma} \frac{\hat{R}_{t+1}}{\hat{\sigma}_{t+1}^2}, \quad (10)$$

where γ is the relative risk aversion coefficient, we set this to 3 in line with Rapach, Ringgenberg, and Zhou (2016), \hat{R}_{t+1} denotes the excess return for t through $t + 1$ using the predictive regression, and $\hat{\sigma}_s^2$ is an estimate of the equity excess return variance using a five-year-moving window. To make the portfolios realistic, the weight on equity is assumed to be between 0.5 and 1.5, as in Rapach, Ringgenberg, and Zhou (2016). When the calculated weight is beyond the range, the weight is set to be the upper bound or the lower bound. At the end of month t , we hold the portfolio consisting of the

SPY ETF and a risk-free asset with the weight given in equation (10) from t through $t + 1$. At the end of $t + 1$, we rebalance the weights on the portfolio for next month. This procedure is repeated until the end of the sample.

We can obtain a certainty equivalent return (CER) from a series of returns generated by the strategy given in equation (10):

$$\text{CER} = \hat{R}_p - \frac{\gamma}{2} \hat{\sigma}_p^2, \quad (11)$$

where γ is the relative risk aversion coefficient, \hat{R}_p and $\hat{\sigma}_p^2$ are the mean and variance of the portfolio returns for the whole estimation period. Using the prevailing mean forecast as a benchmark, the CER gain in our paper is the excess return difference between the portfolio returns generated by the factor allocation strategy and the prevailing mean strategy. In other words, the CER gain is the return that the factor prediction generates superior to the prevailing mean forecast. When the CER gain is negative, it means the IV factor forecast portfolio underperforms that of the prevailing mean forecast portfolio. The Sharpe ratios for the portfolio returns based on each of the factor strategies are also reported in our paper.

4. Empirical results

In this section, we present the summary results for the average quantified IV curves, the IV factor performance in predicting, in-sample and out-of-sample, and the returns to the factor-based asset allocation tests.

Quantified IV curve

This section presents the dynamics of the IV curves for SPY options. In Table 4, we summarize the quantified IV curve coefficients ($\alpha_0, \alpha_1, \alpha_2$), the quantified IV curve factors ($\gamma_0, \gamma_1, \gamma_2$), the coefficient of determination (R^2) and the mean time to maturity by category.

The level factor ($\hat{\gamma}_0$), an estimate of the exact ATM IV,¹ for SPY options is 0.1809 on average across the whole sample. The mean level factor increases monotonically from 0.1414 to 0.2002 for

¹ $a_0=\gamma_0$ gives an estimate of the exact ATM IV, that is the IV where $\xi=0$, while ATM IV is usually the closest one to $\xi=0$.

Table 4. Summary of IV function estimation for SPY options. This table summarizes the results of quantifying the SPY IV curves, as given by: $IV(\xi) = a_0 + a_1\xi + a_2\xi^2$, where ξ is moneyness. Here, a_0 , a_1 , a_2 are the coefficients of the regression, and γ_0 , γ_1 and γ_2 are the dimensionless level, slope and curvature factors as defined in equation (2). The coefficient mean standard error, mean maturity and mean daily R^2 are calculated overall and by maturity category. The percentage of significant factors for each maturity category are the percentage that are significant at 5%.

	Overall		By Maturity (days)					
	< 30	30 – 90	90 – 180	180 – 360	> 360	≤ 180	> 180	
Mean								
\hat{a}_0	0.1809	0.1414	0.1673	0.1831	0.1929	0.2002	0.1648	0.1965
\hat{a}_1	-0.0511	-0.0355	-0.0472	-0.0541	-0.0564	-0.0556	-0.0460	-0.0560
\hat{a}_2	-0.0013	0.0019	-0.0004	-0.0018	-0.0025	-0.0023	-0.0002	-0.0024
γ_0	0.1809	0.1414	0.1673	0.1831	0.1929	0.2002	0.1648	0.1965
γ_1	-0.2818	-0.2709	-0.2886	-0.2945	-0.2865	-0.2696	-0.2854	-0.2782
γ_2	-0.0051	0.0193	-0.0006	-0.0098	-0.0131	-0.0111	0.0022	-0.0122
Standard Errors of Estimates								
\hat{a}_0	0.0005	0.0007	0.0004	0.0004	0.0004	0.0008	0.0005	0.0006
\hat{a}_1	0.0008	0.0011	0.0007	0.0006	0.0007	0.0011	0.0008	0.0009
\hat{a}_2	0.0007	0.0009	0.0005	0.0005	0.0008	0.0010	0.0006	0.0009
Percentage of Significance (%)								
\hat{a}_0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
\hat{a}_1	99.89	99.48	99.99	99.98	99.96	99.93	99.48	99.99
\hat{a}_2	81.77	74.81	78.15	83.21	86.44	82.84	74.81	78.15
Mean Maturity								
mean maturity	262.09	17.42	55.18	131.55	271.13	634.93	67.82	449.78
Mean and median R^2								
mean R^2	0.9886	0.9894	0.9922	0.9914	0.9880	0.9843	0.9912	0.9862
median R^2	0.9938	0.9949	0.9957	0.9950	0.9928	0.9911	0.9953	0.9920

the less than 30-day maturity to the more than 360-day maturity, respectively. It is consistent with our intuition that the IV of the options rises as the maturity time increases, because the long-term horizon contains more uncertainty than the short-term one. As expected, the level coefficient is significant at 5% for almost all of the quantified IV curves. The standard deviation of the IVs across all maturities seem negligible compared to the magnitude of them. The standard deviation across the whole sample maturity categories is only 0.0005, less than 0.3% of its coefficient magnitude.

When we examine the slope factor ($\hat{\gamma}_1$), we can see that on average the IV curves are downward sloping. The slope factor on average for SPY options is -0.2818. The average magnitude of the slope factor increases from -0.2709 to -0.2945 for < 30 maturity and 90–180 day maturity, respectively, and then decreases after the maturity is greater than 180 days. This shows that the IV curves become steeper as maturity increases. The proportion of the slope factor significant at the 5% level of significance is smaller than that for the level factor, but it is still a very large proportion for all maturity groups ranging from 99.49% to 99.99% of slope factors being significant. The standard deviation of the slope is larger, relative to the coefficient magnitude, than that of the level

factor., showing that the slope factor is harder to estimate or agree upon by option traders.

The last quantified IV curve factor is the curvature factor ($\hat{\gamma}_2$). The average curvature factor over the whole sample is -0.0051 for SPY options. The curvature factor for SPY options is positive for < 30 day maturity, and then turns negative for all maturities greater than 30 days, although the magnitude of the curvature is not remarkable. The proportion of the curvature factor with significant estimates is smaller than the proportion for slope factor estimates, with overall significance proportion at 5% level being 81.77%. The proportion of the significant curvature coefficients at 5% tends to decrease, after a fluctuation, for longer maturities.

The proportion of significant factors decreases going from the level to slope and slope to curvature factor, which is consistent with our intuition, that higher dimensional information is more difficult to estimate by option traders and their views of will be less consistent.

Figure 3, 4, 5 show the fitted lines of the IV, trading volume of each contract and the market IVs for SPY options on 1 October 2007, 1 October 2012 and 30 September 2015, respectively. Figure 3 shows that the IV curves are all downward sloping. The curves become more flat

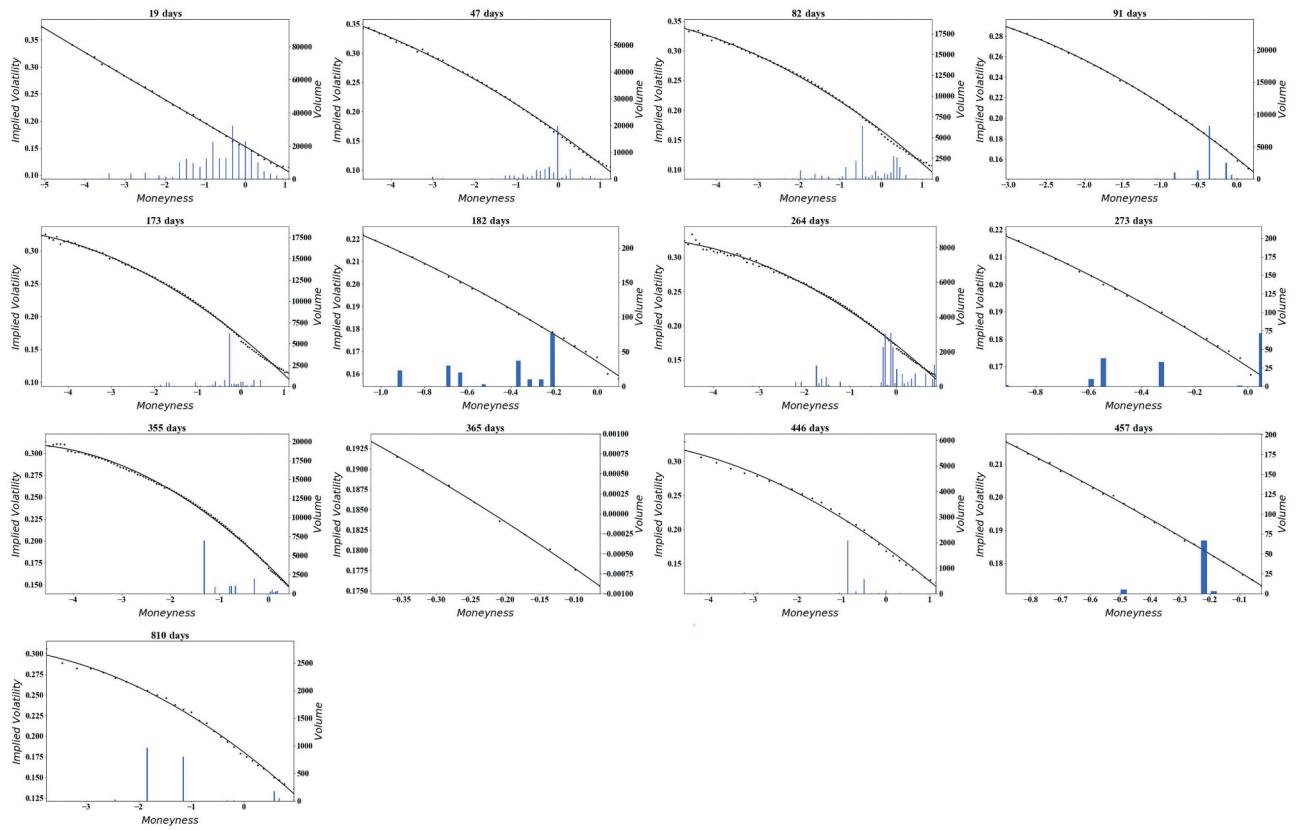


Figure 3. SPY IV against moneyness on 1 October 2007. This figure plots the IV of SPY options against moneyness for different maturity option contract on 1 October 2007. The black dots are the IV. The curve is the fitted IV function, equation (3), and the bars represent the training volume. We apply criteria as described in section (4) to determine whether to use OLS or WOLS.

and less convex as the maturity grows, which is consistent with the above results. As shown in Table 4, the curvature of the IV curves with maturity greater than 60 days mostly turns negative on the three selected days.

Constant maturity factors

Until now, we have examined the pattern of the IV curves for different maturity groups. However, within each maturity group, there can be several maturities. To have a more accurate presentation of how the IV factors evolve over time, we construct constant maturity IV factors. By doing this, we can precisely see the time series of the IV factors with the same maturity. We can also explore the evolution of the term structure of the IV factors.

To create the constant maturity factors, two adjacent factors with maturities close to our target maturity are interpolated or extrapolated as follows,

$$\gamma_0^\tau = \gamma_0^{\tau_1} \omega + \gamma_0^{\tau_2} (1 - \omega) \quad (12)$$

$$\gamma_1^\tau = \gamma_1^{\tau_1} \omega + \gamma_1^{\tau_2} (1 - \omega) \quad (13)$$

$$\gamma_2^\tau = \gamma_2^{\tau_1} \omega + \gamma_2^{\tau_2} (1 - \omega) \quad (14)$$

$$\omega = \frac{(\tau - \tau_1)}{\tau_2 - \tau_1}, \quad (15)$$

where τ is the target maturity, τ_1 is the first closest below the target maturity, and τ_2 is the first closest above the target maturity. Interpolation is applied, when the target maturity is between the two closest maturities. If the target maturity is beyond the maturity range, the IV factors are set to be the upper bound or the lower bound maturity of the factors.

In Figure 6, we present the evolution of the 30-, 60- and 90-day constant maturity level, slope and curvature ($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2$). The time series of the level, slope and curvature factors are very similar for

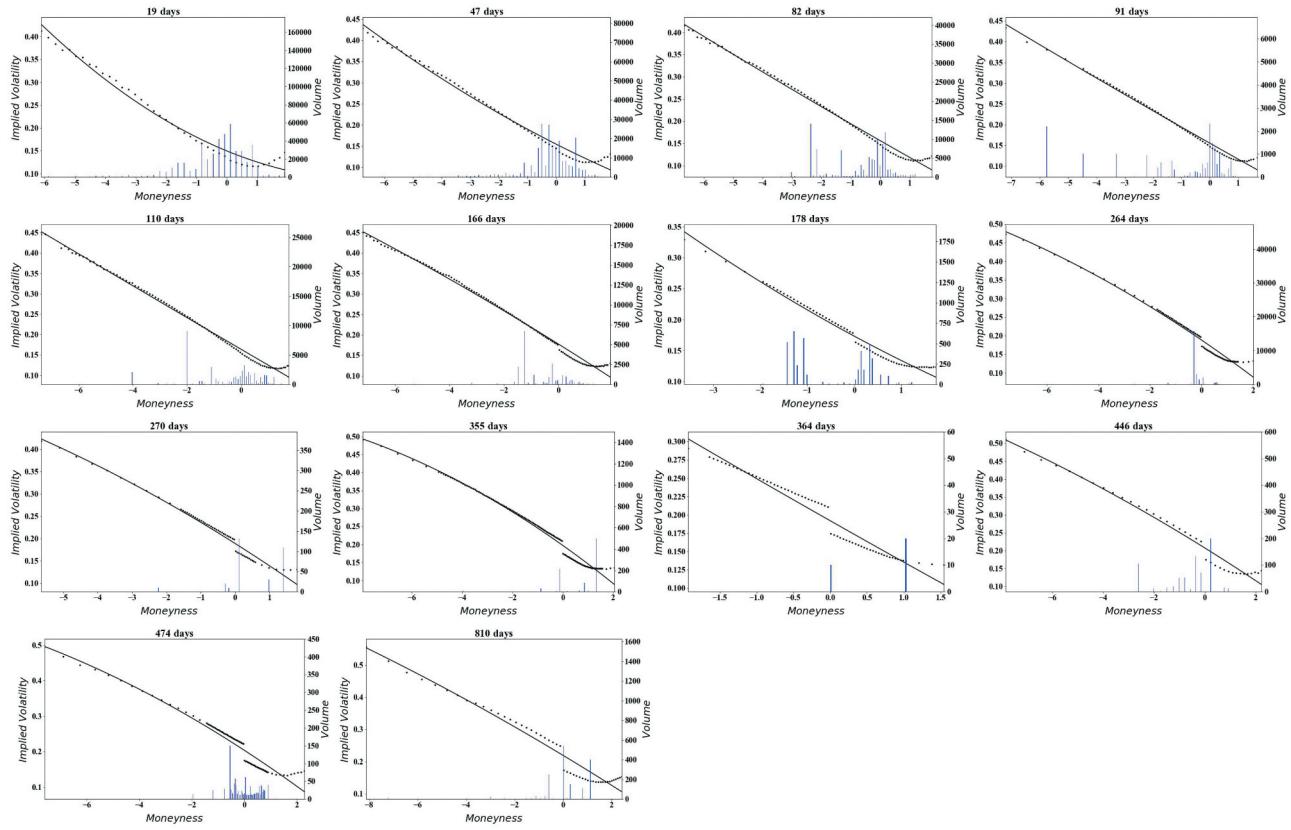


Figure 4. SPY IV against moneyness on 1 October 2012. This figure plots the IV of SPY options against moneyness for different maturity option contract on 1 October 2012. The black dots are the IV. The curve is the fitted IV function, equation (3), and the bars represent the training volume. We apply criteria as described in section (4) to determine whether to use OLS or WOLS.

their three maturities. The level factor is positive throughout the time period, following a mean-reverting pattern, which is a generally accepted pattern of volatility in the literature (Kim, Shephard, and Chib 1998; Alizadeh, Brandt, and Diebold 2002; Yan 2011; Cremers, Halling, and Weinbaum 2015; among others). The volatility of SPY options is small during the normal times before December 2007, followed by a sharp increase lasting for around one year. After that, the SPY ETF volatility reaches its peak of 100% around September 2008, during the peak of the GFC, followed by a sharp fall. That the volatility of SPY options experienced another two spikes in March 2010 and May 2011, which may result from the follow-up effect of the GFC in 2008 and the European sovereign crisis in 2011. We can see the SPY ETF volatility is a good indicator of financial crisis events, meaning it will increase dramatically during the crisis period, and then returns to a normal level after the crisis period.

With respect to the slope factor, it is mostly negative for all of the constant maturity time series. This shows that the IV curves for SPY ETF options are almost always downward sloping. In principle, the time series of the slope factor displays a decreasing trend, albeit associated with highly frequent variation around the main trend line. This shows that over our sample the IV curves are becoming ever more negatively sloped. In line with the spikes found for the level factor, the slope factor becomes steeper briefly during the GFC period and the European sovereign debt crisis period. During these times OTM put options, a form of downside insurance, are more expensive relative to OTM call options.

When examining the curvature time series, it is mostly positive, and occasionally negative for a short period. This is consistent with the above results in Table 4 that the IV curves are mostly linear with minor curvature. The curvature for the IV curves with long time to maturity seems to display higher

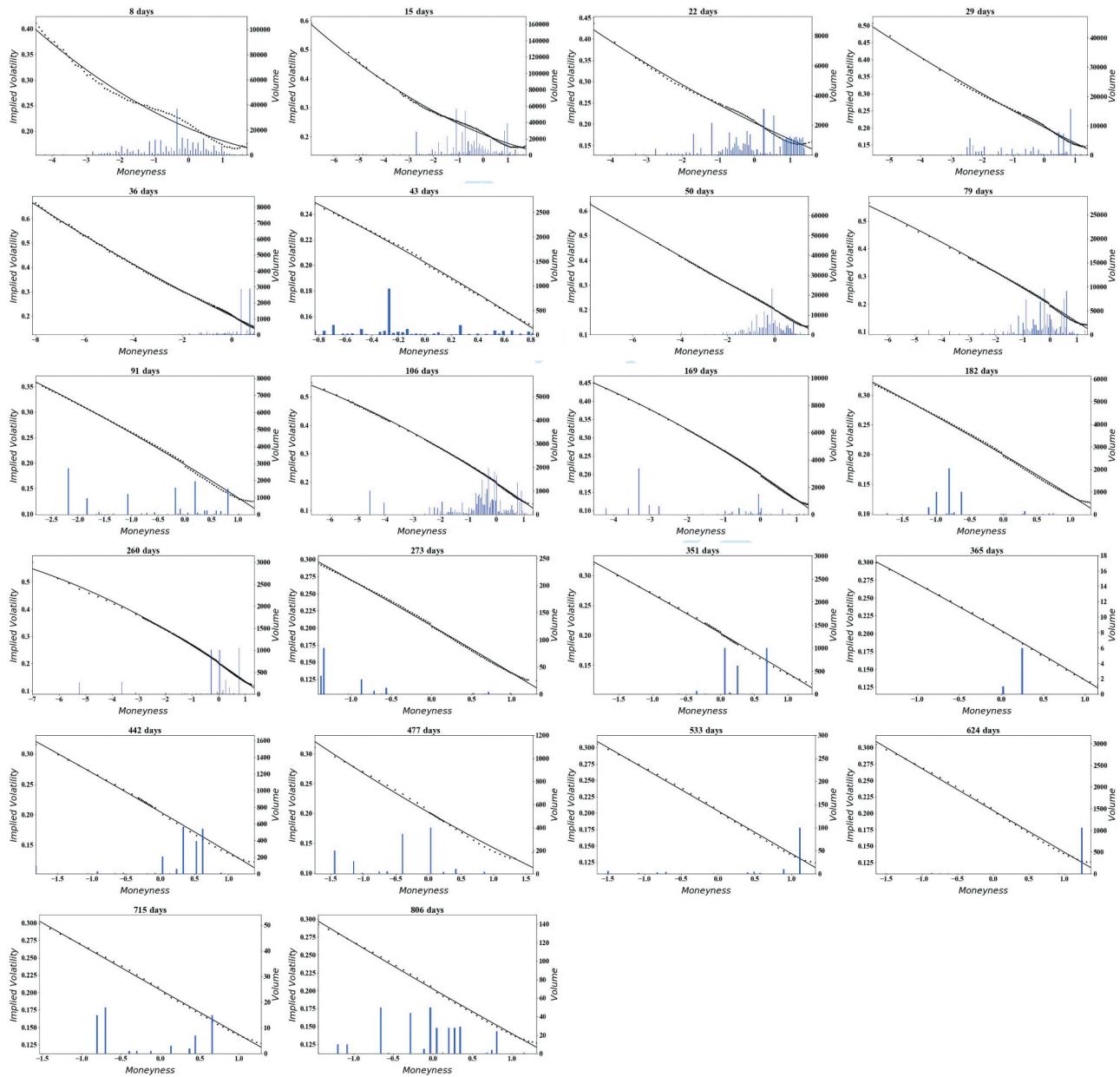


Figure 5. SPY IV against moneyness on 30 September 2015. This figure plots the IV of SPY options against moneyness for different maturity option contract on 1 October 2015. The black dots are the IV. The curve is the fitted IV function, equation (3), and the bars represent the traing volume. We apply criteria as described in section (4) to determine whether to use OLS or WOLS.

spikes than that for short maturity IV curves, during the GFC period. Meanwhile, its time series contains many fluctuations rather than a few spikes in the financial crisis periods, as found for the other factors.

Figure 7 shows the average constant maturity curves, determined by fitting a line based on the average constant maturity factors. As the maturity increases, the IV curves of SPY options become more negatively sloped, and the curvature of SPY options increases.

Subsample analysis for the GFC

To explore the effect of the GFC more closely, a subdivision of the sample is made. Referring to the GFC period set in Lins, Servaes, and Tamayo (2017), the time period between January 2005 to December 2017 is divided into three period, January 2005 – July 2008 (Before the GFC), August 2008 – March 2009 (GFC), and April 2009 – December 2017 (After the GFC).

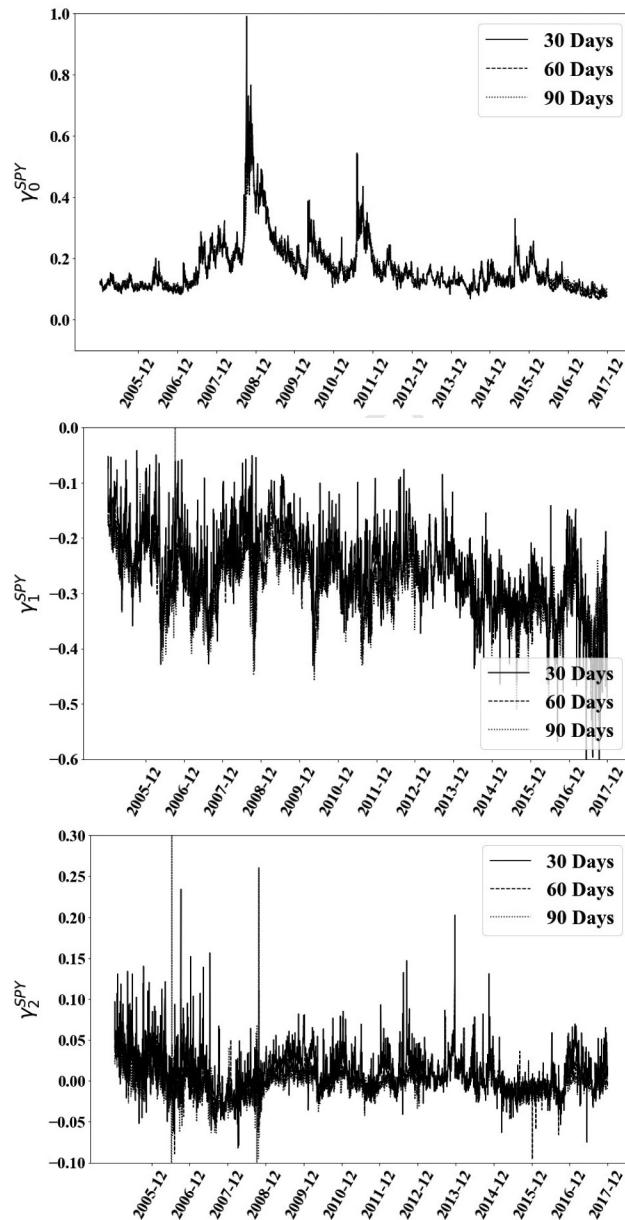


Figure 6. Evolution of Coefficients of SPY Options over Time. This figure show the evolution of coefficients of SPY Options over time at 30-, 60-, 90-day constant maturity range from 1 January 2005 to 31 December 2017. The three coefficients are level, slope and curvature of the fitted IV curve. The first column plots the level, slope and curvature for SPX with constant 30, 60 and 90 day maturity. The second column plots the level, slope and curvature for SPY with constant 30, 60 and 90 day maturity.

The fitted IV curves resulting from the constant maturity average factors within each sub-period are plotted in Figure 8. We can see that the IV curves within the GFC are very different from those before and after the GFC. We can see that, for our sample period, we do not find a ‘smirk’ pattern outside of the GFC period. This leads us to believe that the well accepted smirk shape (Pan 2002; Jiang 2002; Carverhill, Cheuk, and Dyrting 2009, among

others) in the index options markets has dissipated in recent years or is driven predominantly by crisis observations, as is the case for our sample. The IV slope during the GFC becomes more negatively skewed, and the curvature becomes more positively curved compared to normal periods.

Following the theory of Zhang, Zhao, and Chang (2012) and Zhang, Chang, and Zhao (2020), we provide a theory to explain why the IV curve

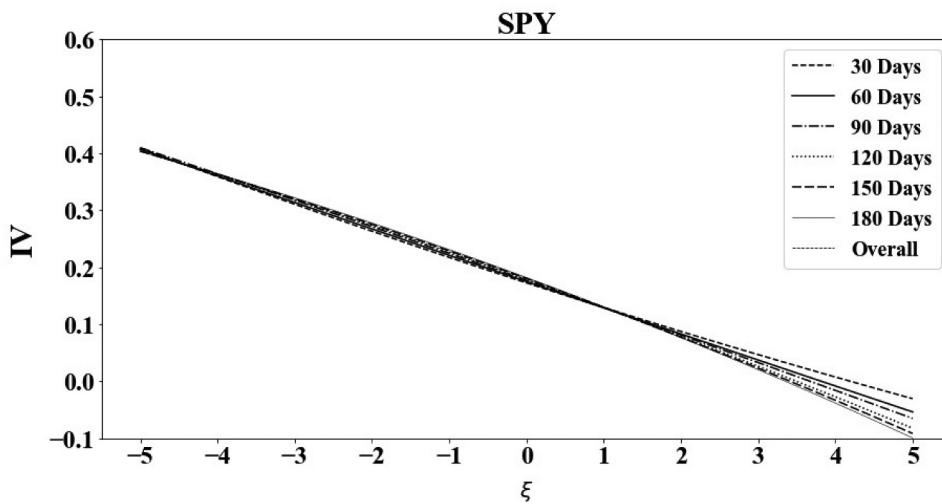


Figure 7. IV of SPY Options. This figure depicts the IV curves of SPY options for constant maturities over the horizon of 30-, 60-, 90-, 120-, 150-, 180-day and overall sample with coefficients averaged on all contracts.

becomes more negatively skewed and positively curved during the GFC period. In the appendix, we build an equilibrium asset pricing model in an production economy. We assume the price of the asset follows a jump-diffusion process with stochastic volatility and jump intensity. By solving the optimal problem of a representative investor with constant relative risk aversion, we obtain the optimal condition, and build relationships between variance, third and forth cumulants, with jump. We can see that the IV smirk becomes more significant during the GFC period due to the increase in jump intensity and a more negative jump size. In the appendix, Equation A.4, we show that the third and fourth cumulants are associated with jump intensity and jump size. When jump intensity becomes more positive and jump size becomes more negative, the third cumulant becomes more negative and the fourth cumulant becomes more positive, thus the IV curve becomes more negatively skewed and positively curved. Bollerslev and Todorov (2011) show that jump intensity increases and jump size becomes more negative during the GFC period, which is consistent with the results presented here.

Predictive regression estimation

As the IV curve factors are proportionally related to the risk neutral moments (Zhang and Xiang

2008), we use these proxies for the risk neural moments and test their ability to forecast SPY ETF returns.

Predictor variables

We apply the level, slope and curvature of the IV curves as predictors to predict the next month excess returns. The excess returns and IV curve factors are measured at the end of the month.

Considering the higher moment information and lag effect in time series regression, we create the first difference factors from the IV curve factors. In line with Xing, Zhang, and Zhao (2010) and Ruan and Zhang (2018), we also create third, $TC = \gamma_1 \times \gamma_0^3$, and fourth order, $FC = \gamma_2 \times \gamma_0^4$, cumulants as predictors. When considering the lag effect of the factors in predicting the SPY ETF excess returns, we include the first difference of the level (γ_0), slope (γ_1), curvature (γ_2), the third cumulant (TC) and the fourth cumulant (FC) in line Chang, Christoffersen, and Jacobs (2013) and Conrad, Dittmar, and Ghysels (2013). 10 potential predictors are created to forecast the one-month ahead SPY ETF excess returns.

Table 5 presents the summary statistics for 10 predictors from January 2005 to December 2017. The mean of the level, slope and curvature is 0.1677, -0.2529 and 0.0123, respectively. The third cumulant is only -0.0021, and the fourth cumulant is close to zero. The standard deviation of

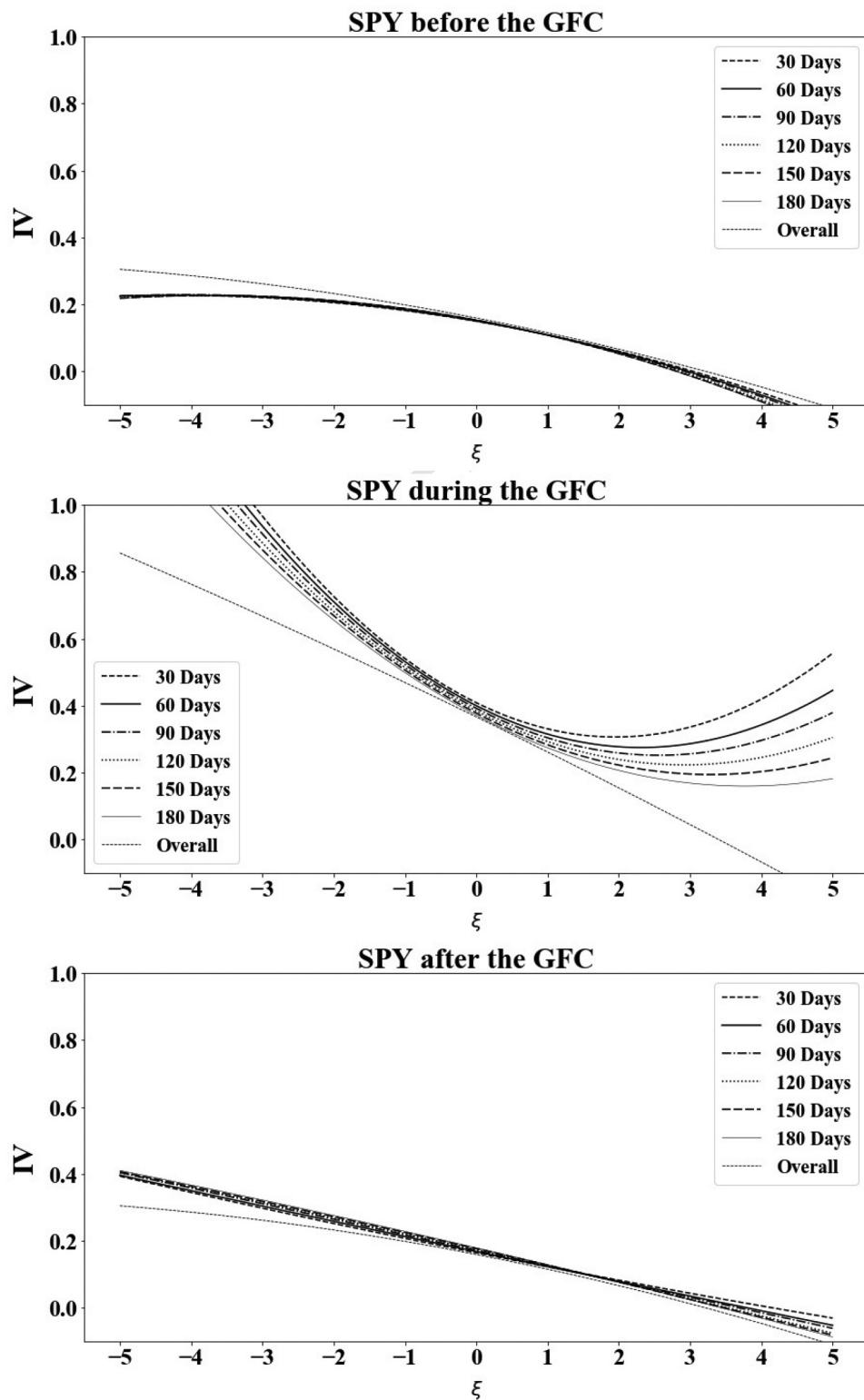


Figure 8. IV of SPY Options and the GFC. This figure depicts the IV curves of SPY options for constant maturities over the horizons at 30-, 60-, 90-, 120-, 150-, 180-day and overall sample for three sub-divided periods, namely before, during and after the GFC with coefficients averaged on these subperiod contracts.



Table 5. Summary statistics for predictors. This table displays summary statistics for 10 predictors used in predicting SPY ETF returns from January 2005 to December 2017. The statistics include mean, standard deviation (sd), skewness (skew), the 10th percentile (p10), the 50th percentile (p50) and the 90th percentile (p90) of the predictors. γ_0 , γ_1 and γ_2 is the level, slope and curvature of IV smirks with 30-day constant maturity, respectively. TC and FC are the risk-neutral third and fourth cumulants calculated by $TC = \gamma_1 * \gamma_0^3$, and $FC = \gamma_2 * \gamma_0^4$. $\Delta\gamma_0$, $\Delta\gamma_1$, $\Delta\gamma_2$, ΔTC and ΔFC are the first difference of γ_0 , γ_1 , γ_2 , TC and FC , respectively.

	mean	sd	skew	p10	p50	p90
γ_0	0.1673	0.0834	2.0786	0.0969	0.1426	0.2525
γ_1	-0.2523	0.0718	-0.5076	-0.3311	-0.2503	-0.1662
γ_2	0.0125	0.0252	0.4648	-0.0156	0.0116	0.0427
TC	-0.0021	0.0047	-5.5971	-0.0040	-0.0008	-0.0002
$FC(\times 10)$	0.0002	0.0012	17.0228	-0.0001	0.0000	0.0006
$\Delta\gamma_0$	-0.0002	0.0456	0.5804	-0.0501	-0.0026	0.0455
$\Delta\gamma_1$	-0.0021	0.0659	-0.4504	-0.0769	-0.0016	0.0693
$\Delta\gamma_2$	-0.0002	0.0258	-0.0369	-0.0284	-0.0013	0.0288
$\Delta TC(\times 10^5)$	0.0009	368.7100	-177,036.2000	-116.9600	1.0500	142.4700
$\Delta FC(\times 10^5)$	-0.0037	13.9400	77,114.1200	-3.8600	-0.0680	2.4400

Table 6. Correlation for factors. This table shows correlations among the 10 predictors calculated from SPY IV curves with constant 30-day to maturity from January 2005 through December 2017. All the factors are monthly data using the data at the end of each month. γ_0 , γ_1 and γ_2 is the level, slope and curvature of IV smirks with 30-day constant maturity, respectively. TC and FC are the risk-neutral third and fourth cumulants calculated by $TC = \gamma_1 * \gamma_0^3$, and $FC = \gamma_2 * \gamma_0^4$. $\Delta\gamma_0$, $\Delta\gamma_1$, $\Delta\gamma_2$, ΔTC and ΔFC are the first difference of γ_0 , γ_1 , γ_2 , TC and FC , respectively.

	γ_0	γ_1	γ_2	TC	FC	$\Delta\gamma_0$	$\Delta\gamma_1$	$\Delta\gamma_2$	ΔTC	ΔFC
γ_0	1.00									
γ_1	0.31	1.00								
γ_2	-0.15	0.56	1.00							
TC	-0.85	-0.11	0.14	1.00						
FC	0.26	0.26	0.25	-0.05	1.00					
$\Delta\gamma_0$	0.28	-0.06	-0.22	-0.27	-0.01	1.00				
$\Delta\gamma_1$	0.06	0.49	0.29	0.05	0.04	-0.01	1.00			
$\Delta\gamma_2$	0.00	0.34	0.51	0.05	0.13	-0.22	0.66	1.00		
ΔTC	-0.19	0.09	0.12	0.39	0.04	-0.69	0.19	0.18	1.00	
ΔFC	-0.07	0.10	0.12	0.29	0.58	-0.14	0.14	0.23	0.41	1.00

Table 7. In-sample estimation results. This table displays the coefficients and adjusted R^2 statistic adjusted by Newey and West (1987) with six lags, for the in-sample prediction model: $r_{t:t+1} = a + \beta x_t + \varepsilon_{t:t+1}$ where $t = 1, \dots, T-1$, r_t is the continuous excess returns for SPY ETF at month t , x_t is the predictor at month t . The sample period is from January 2005 to December 2017. *, ** and *** indicate significance at 10%, 5% and 1% level, respectively.

(1)	(2)	(3)	(4)
In-sample tests			
Predictor	$\hat{\beta}$	t-statistics	R^2_{IS} (%)
γ_0	-0.0032	-0.48	-0.01
γ_1	0.0005	0.13	-0.64
γ_2	0.0044	1.76*	0.54
TC	0.0047	1.23	0.70
FC	-0.0029	-0.38	-0.1
$\Delta\gamma_0$	-0.0053	-1.08	1.12
$\Delta\gamma_1$	0.0055	1.80*	1.04
$\Delta\gamma_2$	0.0035	1.54	0.10
ΔTC	0.0045	1.30	0.59
ΔFC	-0.0070	-2.33**	2.35

the first difference of the five factors is much larger than its mean number, meaning the variation of the first difference is much larger than its original number, therefore increasing the robustness of the regression.

Table 6 presents the pairwise correlation for all predictors. Among them, the level and the third cumulant have the highest negative correlation of -0.85, which is consistent with results in Ruan and Zhang (2018) and Chang, Christoffersen, and

Jacobs (2013). The strongest correlation among the first difference factors, 0.66, is between the first difference of slope, $\Delta\gamma_1$, and the first difference of curvature, $\Delta\gamma_2$, remarkably weaker than the correlation between γ_0 and TC of -0.85 .

In-sample results

Table 7 reports the OLS results of equation (6) for each of the factors associated with IV curves for in-sample estimations. The table contains the β , t -statistic, adjusted R^2 and p value. To address the autocorrelation and heteroskedasticity in the time series regression, we adjust the standard errors using the Newey and West (1987) method with six lags. The t statistics are estimated on a two side inference to improve the robustness of the estimation. To avoid the bias in the overlapping regression mentioned in Goetzmann and Jorion (1993), Ang and Bekaert (2006) and Cochrane (2007), we only forecast the future excess returns with one-month horizon, that is $h = 1$.

In line with the effect of the predictive factors in Gharghori, Maberly, and Nguyen (2017), most predictors have poor performance for the in-sample tests. In the monthly horizon forecasting estimation, only three factors, γ_2 , $\Delta\gamma_1$, ΔFC , exhibit significant predictive power. ΔFC shows the strongest statistical significance among all the IV factors with the t

statistic of -2.33 , while γ_2 and $\Delta\gamma_1$ show moderate significance. After the standardization for the factors, all the regression coefficients have the comparable values, except the smallest number of γ_1 , 0.0005, which is much smaller than other factor coefficients. The ΔFC coefficient, -0.0070 , is the largest among all the factors, which means a standard deviation increase in ΔFC will contribute to -0.7% decrease in the next month excess return.

The adjusted R^2 is also a measure to estimate the performance of the factors. After considering the autocorrelation and heteroskedasticity in the time series regression, the adjusted R^2 itself is not enough to obtain a reliable result for the factor estimation. As suggested by Campbell and Thompson (2007), a monthly adjusted R^2 of approximately 0.5% associated with the factor regression represents a significant predictive power. In Table 7, all the adjusted R^2 of the significant factors are much larger than this suggested value. The adjusted R^2 of γ_2 , $\Delta\gamma_1$ and ΔFC is 1.18%, 1.69%, and 2.99%, respectively. Moreover, the adjusted R^2 estimates of other factors are larger than 0.5%, even 1.77% for $\Delta\gamma_0$, but the t -statistic is not statistically significant after the New-west adjustment with six lags. Overall, the adjusted t -statistic is a more strict measure for estimating the significance of the predictors in the regression than the adjusted R^2 . Among all the IV factors in our paper, γ_2 , $\Delta\gamma_1$ and

Table 8. Out-of-sample estimation results. This table displays the coefficients and adjusted R^2 statistic adjusted by Newey and West (1987) with six lags, for out-of-sample prediction model. The adjusted mean squared prediction error (MSPE) is calculated as, $f_{t+1} = (r_{t+1} - \bar{r}_{t+1|t})^2 - ((r_{t+1} - \hat{r}_{t+1|t})^2 - (\bar{r}_{t+1} - \hat{r}_{t+1|t})^2)^2$ for $t = 1, T - 1$ in line with Clark and West (2007), where \bar{r}_t is the prediction value from the prevailing mean forecast at month t , \hat{r}_t is the predicted value from the IV predictor regression at month t . The sample period is from January 2005 to December 2017. *, ** and *** indicate significance obtained from the one-side estimation at 10%, 5% and 1% level, respectively.

(1)	(2)	Out-of-sample tests
		R^2_{OS}
Predictor		
γ_0		-0.07
γ_1		-3.33
γ_2		-3.62
TC		1.28
FC		0.65
$\Delta\gamma_0$		-7.12
$\Delta\gamma_1$		3.03*
$\Delta\gamma_2$		1.07
ΔTC		0.22
ΔFC		-0.99



ΔFC show significant power to predict the next month excess returns, in-sample.

Out-of-sample tests

We apply out-of-sample tests to further estimate whether the predicting model outperforms the average history forecasting model. Table 8 reports the estimation results for out-of-sample tests with the out-of-sample period being the one third of the sample period from September 2013 to December 2017. For the out-of-sample tests, $\Delta\gamma_1$ is the only predictor with a statistically significant improvement in MSPE, and the highest adjusted R^2 of 3.03%, implying that the $\Delta\gamma_1$ prediction beats the average history forecast model. In addition, TC, FC, $\Delta\gamma_2$ and ΔTC also show positive improvements in MSPE, but the improvements are not significant. For the remaining predictors, MSPE presents negative values, indicating that these factors fail to beat the prevailing mean forecast model.

Combining the in-sample and out-of-sample tests, we can conclude that $\Delta\gamma_1$ is the best predictor to predict the next-month SPY excess returns among all IV predictors.

Asset allocation

Table 9 reports the CER gains and Sharpe ratios for the portfolios obtained from the predictive regressions from IV factors, as described in section (5.3.1). The CER gain is a measure to estimate the predic-

model. Similar to the measure of the CER gain, the Sharpe ratio is another measure to estimate the overall performance of the portfolio, consisting of an equity and a risk-free bill with weights determined by the factor forecasting regression. In principle, these two measures generate the same comparison results for a set of factors. By combining the comparison results from these two measures, we can obtain a more reliable conclusion.

From columns (2) through (3) in Table 9, $\Delta\gamma_1$ performs best among all the predictors in terms of CER gains. $\Delta\gamma_1$ generates the highest annualized CER gains of 1.65% among all the predictors. Besides, only four predictors, TC, FC, $\Delta\gamma_2$, ΔTC , generate the positive CER gains, but well below that of $\Delta\gamma_1$. The Sharpe ratio of 1.05 for $\Delta\gamma_1$ is the second largest among all the predictors, very close to the largest one of 1.08 generated by $\Delta\gamma_2$. Even the smallest Sharpe ratio for $\Delta\gamma_0$ has 0.60, relatively large compared to its CER gains of -3.61. In general, the differences of the Sharpe ratios among the predictors are much smaller than those for the CER gains, that is because the Sharpe ratio has a relative small variation range compared to that for the CER gains. So the CER gains may be a better measure to estimate the predictive ability of the factors. In terms of the CER gains and Sharpe ratios, we can draw a conclusion that $\Delta\gamma_1$ has the strongest predictive power among all the IV predictors. This is consistent with the results in Neumann and Skiadopoulos

Table 9. Asset allocation. This table reports the annualized certainty equivalent return (CER) and Sharpe ratios for the portfolios created from a mean-variance strategy, which consists of a risk-free asset and an equity with weights obtained from the coefficients of the regressions that run excess SPY ETF returns on each of the IV factors, as described in section (5.3.1). The weights for the portfolio are rebalanced at the end of each month.

(1)	(2)	(3)
Predictor	CER gains	Sharp ratio
γ_0	-0.09	0.92
γ_1	-4.35	0.85
γ_2	-3.79	0.77
TC	0.92	1.00
FC	0.44	1.02
$\Delta\gamma_0$	-3.21	0.60
$\Delta\gamma_1$	1.65	1.05
$\Delta\gamma_2$	1.06	1.08
ΔTC	0.51	1.03
ΔFC	-1.21	0.96

tion power of the factor in the forecast regression

(2013) and Amaya et al. (2015), who show the predictive ability of skewness for equity returns.

Controlling for popular predictors

In this section, we provide robustness checks for the predictability of the first difference of the slope factor, by including some popular predictors as control variables. We employ 15 predictors from Welch and Goval (2008) and Moskowitz, Ooi, and Pedersen (2012), the definitions of these predictors are as follows,

- (1) Log-dividend price ratio (DP): log of a 12-month moving sum of dividends paid on the SPX minus the log of the SPX price.
- (2) Log dividend yield (DY): log of a 12-month moving sum of dividends minus the log of lagged stock prices.
- (3) Log earnings-price ratio (EP): log of a 12-month moving sum of earnings on the SPX minus the log of stock prices.
- (4) Log dividend-payout ratio (DE): log of a 12-month moving sum of dividends minus the log of a 12-month moving sum of earnings.
- (5) Excess stock return volatility (RVOL): 12-month moving standard deviation of SPY returns.²
- (6) Book-to-market ratio (BM): book-to-market value ratio for the Dow Jones Industrial Average.
- (7) Net equity expansion (NTIS): ratio of a 12-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of NYSE stocks.
- (8) Treasury bill rate (TBL): interest rate on a three-month Treasury bill (secondary market).
- (9) Long-term yield (LTY): long-term government bond yield.
- (10) Long-term return (LTR): return on long-term government bonds.
- (11) Term spread (TMS): long-term yield minus the Treasury bill rate.
- (12) Default yield spread (DFY): difference between Moody's BAA- and AAA-rated corporate bond yields.

- (13) Default return spread (DFR): long-term corporate bond return minus the long-term government bond return.
- (14) Inflation (INFL): calculated from the Consumer Price Index (CPI) for all urban consumers.
- (15) Time series momentum (TSM): 12-month moving average of SPY returns.

Table 10 presents the prediction results for the first difference of the slope along with each of the popular factors in the in-sample tests. The first difference of the slope factor exhibits intermediate significance for predicting the next-month SPY returns in all regressions. The first difference of the slope, is around 0.08 in these regressions, meaning a standard deviation increase in the first difference of the slope, is associated with a 8 basis point increase in next-month's excess SPY return. This verifies that the first difference of the slope factor is a robust predictor for the next-month SPY returns.

5. Conclusions

In this paper, we apply the method of Zhang and Xiang (2008) to quantify the IV curves for SPY options. We calculate the level, slope and curvature factors from the IV curves, and obtain the third and fourth cumulants as predictors for the next-month SPY excess returns. Moreover, the in-sample and out-of-sample tests show that the first difference of the slope factor is a significant predictor of SPY excess returns at the monthly horizon. Furthermore, the mean-variance portfolio gains created based on the factor's predictive results verify its predictability.

Our results show that the well-accepted IV smirk for options only appear for SPY options during the GFC period, and the IV curves are more like negative-sloped line outside the GFC period. This is different from the pattern of index options IV curves documented by (Jiang 2002; Carverhill, Cheuk, and Dyrting 2009), indicate the IV curves pattern of options may be changing over time. Following the theory of Zhang, Zhao, and Chang (2012) and Zhang, Chang, and Zhao (2020) we

²Following the definition of Rapach, Ringgenberg, and Zhou (2016), we use excess stock return volatility to measure stock return volatility. This definition avoids producing a severe outlier and yields more plausible results.



Table 10. Predictability of SPY returns. This table reports β for the time-series prediction regression for SPY returns. The sample period is from January 2005 to December 2017. ΔY_1 is the first difference of the slope factor. TSM is the momentum defined as in Moskowitz, Ooi, and Pedersen (2012) and the other 14 popular predictors are given in Welch and Goyal (2008). This table presents the slope, β , and Newey-West adjusted t -statistics. The Newey-West standard errors are computed with six lags. *, **, and *** indicate significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	
ΔY_1	0.082*	0.080*	0.083*	0.083*	0.083*	0.083*	0.083*	0.083*	0.083*	0.080*	0.081*	0.083*	0.084*	0.080*	0.087*	
DP	0.010	(0.21)													0.083*	
DY	0.025	(0.61)													(1.82)	
EP					-0.003 (-0.22)		-0.001 (-0.10)									
DE																
RVOL						0.0151 (0.64)										
BM							0.066 (0.76)									
NTIS								0.375 (1.25)								
TBL									-0.219 (-1.54)							
LTY										-0.665** (-2.20)						
LTR											-0.116 (1.57)					
TMS											-0.016 (-0.06)					
DFY												-0.772 (-0.62)				
DFR													-0.236 (0.89)			
INFL														0.976 (1.17)		
TSM																
Constant	0.0435	0.103 (0.622)	-0.00381 (-0.109)	0.000399 (0.0106)	0.00882*** (3.320)	0.00734** (-0.576)	-0.0164 (2.564)	0.00685 (1.593)	0.0281*** (3.211)	0.00364 (0.900)	0.00464 (0.989)	0.0128 (1.143)	0.00415 (1.096)	0.00262 (0.557)	0.00405 (0.698)	0.044 (0.09)

believe this phenomenon is driven by the abnormal jumps during the GFC period, and can explain the IV smirk. The IV curve becomes more negatively skewed and positively curved, because jump intensity becomes more positive and jump size becomes more negative during the GFC period.

Due to the inherent difficulty in predicting future returns, most IV factors show poor performance in forecasting the future excess returns. However, we find that the first difference of the slope factor ($\Delta\gamma_1$) is significant for predicting one month ahead SPY excess returns. It exhibits strong predictability both in the in-sample and out-of-sample tests. The certainty equivalent returns and Sharpe ratios for the portfolios based on this further confirm our results. We indeed find that can predict the next-month SPY excess returns using the quantified IV curve slope factor, even when controlling for 15 popular equity market predictors from Welch and Goyal (2008) and Moskowitz, Ooi, and Pedersen (2012).

6. Appendix: a theory for the IV smirk

Applying the theory in Zhang, Zhao, and Chang (2012) and Zhang, Chang, and Zhao (2020), we can explain why SPY IV curves become more negatively skewed and positively curved during the GFC period.

Assuming that the stock index, S_t , follows a stochastic process with a stochastic variance, V_t , and jump intensity, λ_t , as follows,

$$\frac{dS_t}{S_t} = \mu_t dt + \sqrt{V_t} dB_t^S + (e^\chi - 1)dN_t - \lambda_t E(e^\chi - 1)dt \quad (\text{A.1})$$

$$dv_t = \kappa_V(\theta_V - V_t)dt + \sigma_V \sqrt{V_t} dB_t^V \quad (\text{A.2})$$

$$d\lambda_t = \kappa_\lambda(\theta_\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dB_t^\lambda \quad (\text{A.3})$$

where μ_t is the instantaneous expected return, B_t^S , B_t^V and B_t^λ are three Brownian motions, N_t is a Possion motion, with jump intensity λ_t and jump size χ .

After some derivations, we build the links between the model parameters and the total variance, the third and fourth cumulants, as follows,

$$\begin{aligned} TV_t &= V_t + \lambda_t E(\chi^2), & TC_t &= \lambda_t E(\chi^3), \\ FC_t &= \lambda_t E(\chi^4). \end{aligned} \quad (\text{A.4})$$

From the above results, we can see that the third cumulant, TC_t , and the fourth cumulant, FC_t , are associated with jump intensity, λ_t , and jump size, χ . During the crisis period, the jump intensity becomes more positive and jump size becomes more negative, thus making the third cumulant more negative, and the fourth cumulant more positive. Accordingly, the skewness becomes more negative, and the kurtosis becomes more positive. Based on Zhang and Xiang (2008) theory, the IV curve becomes more negatively skewed and more positively curved. Finally, we can conclude that SPY IV curves become more negatively skewed and positively curved during the GFC period because of being more positive in jump intensity and being more negative in jump size.

Acknowledgement

Wei Guo appreciates being awarded the University of Otago Doctoral Scholarship. Jin E. Zhang has been supported by an establishment grant from the University of Otago and the National Natural Science Foundation of China (Project No. 71771199).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Wei Guo  <http://orcid.org/0000-0002-7216-7956>
 Sebastian A. Gehricke  <http://orcid.org/0000-0002-3251-9275>
 Xinfeng Ruan  <http://orcid.org/0000-0002-1447-6603>
 Jin E. Zhang  <http://orcid.org/0000-0002-8556-474X>

References

- Alizadeh, S., M. W. Brandt, and F. X. Diebold. 2002. “Range-based Estimation of Stochastic Volatility Models.” *Journal of Finance* 57: 1047–1091. doi:[10.1111/1540-6261.00454](https://doi.org/10.1111/1540-6261.00454).
 Amaya, D., P. Christoffersen, K. Jacobs, and A. Vasquez. 2015. “Does Realized Skewness Predict the Cross-section of Equity Returns?” *Journal of Financial Economics* 118: 135–167. doi:[10.1016/j.jfineco.2015.02.009](https://doi.org/10.1016/j.jfineco.2015.02.009).



- Ang, A., and G. Bekaert. 2006. "Stock Return Predictability: Is It There?" *Review of Financial Studies* 20: 651–707. doi:[10.1093/rfs/hhl021](https://doi.org/10.1093/rfs/hhl021).
- Ang, A., and G. Bekaert. 2007. "Stock Return Predictability: Is It There?" *Review of Financial Studies* 20: 651–707.
- Aschakulporn, P., and J. E. Zhang. 2020. "New Zealand Whole Milk Powder Options." *Accounting and Finance*. doi:[10.1111/acfi.12660](https://doi.org/10.1111/acfi.12660).
- Bakshi, G., and N. Kapadia. 2003. "Delta-hedged Gains and the Negative Market Volatility Risk Premium." *Review of Financial Studies* 16: 527–566. doi:[10.1093/rfs/hhg002](https://doi.org/10.1093/rfs/hhg002).
- Bakshi, G., N. Kapadia, and D. Madan. 2003. "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options." *Review of Financial Studies* 16: 101–143. doi:[10.1093/rfs/16.1.0101](https://doi.org/10.1093/rfs/16.1.0101).
- Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein. 2011. "Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash." *Journal of Financial Economics* 101: 552–573. doi:[10.1016/j.jfineco.2011.01.008](https://doi.org/10.1016/j.jfineco.2011.01.008).
- Bollerslev, T., G. Tauchen, and H. Zhou. 2009. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies* 22: 4463–4492. doi:[10.1093/rfs/hhp008](https://doi.org/10.1093/rfs/hhp008).
- Bollerslev, T., and V. Todorov. 2011. "Estimation of Jump Tails." *Econometrica* 79: 1727–1783.
- Broadie, M., M. Chernov, and M. Johannes. 2009. "Understanding Index Option Returns." *Review of Financial Studies* 22: 4493–4529. doi:[10.1093/rfs/hhp032](https://doi.org/10.1093/rfs/hhp032).
- Campbell, J. Y., and S. B. Thompson. 2007. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies* 21: 1509–1531. doi:[10.1093/rfshhm055](https://doi.org/10.1093/rfshhm055).
- Campbell, J. Y., and S. B. Thompson. 2008. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies* 21: 1509–1531.
- Carr, P., and W. Liuren. 2003. "The Finite Moment Log Stable Process and Option Pricing." *Journal of Finance* 58: 753–777. doi:[10.1111/1540-6261.00544](https://doi.org/10.1111/1540-6261.00544).
- Carverhill, A., T. H. F. Cheuk, and S. Dyrting. 2009. "The Smirk in the S&P500 Futures Options Prices: A Linearized Factor Analysis." *Review of Derivatives Research* 12: 109–139. doi:[10.1007/s11147-009-9037-2](https://doi.org/10.1007/s11147-009-9037-2).
- Chang, B. Y., P. Christoffersen, and K. Jacobs. 2013. "Market Skewness Risk and the Cross Section of Stock Returns." *Journal of Financial Economics* 107: 46–68. doi:[10.1016/j.jfineco.2012.07.002](https://doi.org/10.1016/j.jfineco.2012.07.002).
- Chen, J. H., and C. Y. Huang. 2010. "An Analysis of the Spillover Effects of Exchange-traded Funds." *Applied Economics* 42: 1155–1168. doi:[10.1080/00036840701721182](https://doi.org/10.1080/00036840701721182).
- Clark, T. E., and K. D. West. 2007. "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics* 138: 291–311. doi:[10.1016/j.jeconom.2006.05.023](https://doi.org/10.1016/j.jeconom.2006.05.023).
- Cochrane, J. H. 2007. "The Dog that Did Not Bark: A Defense of Return Predictability." *Review of Financial Studies* 21: 1533–1575. doi:[10.1093/rfs/hhm046](https://doi.org/10.1093/rfs/hhm046).
- Cochrane, J. H. 2011. "Presidential Address: Discount Rates." *Journal of Finance* 66: 1047–1108. doi:[10.1111/j.1540-6261.2011.01671.x](https://doi.org/10.1111/j.1540-6261.2011.01671.x).
- Conrad, J., R. F. Dittmar, and E. Ghysels. 2013. "Ex Ante Skewness and Expected Stock Returns." *Journal of Finance* 68: 85–124. doi:[10.1111/j.1540-6261.2012.01795.x](https://doi.org/10.1111/j.1540-6261.2012.01795.x).
- Cremers, M., M. Halling, and D. Weinbaum. 2015. "Aggregate Jump and Volatility Risk in the Cross-section of Stock Returns." *Journal of Finance* 70: 577–614. doi:[10.1111/jofi.12220](https://doi.org/10.1111/jofi.12220).
- Fama, E. F., and K. R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33: 3–56. doi:[10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5).
- Fama, E. F., and K. R. French. 2015. "A Five-factor Asset Pricing Model." *Journal of Financial Economics* 116: 1–22. doi:[10.1016/j.jfineco.2014.10.010](https://doi.org/10.1016/j.jfineco.2014.10.010).
- Gehrke, Sebastian A., and Jin E. Zhang. 2020. "The implied volatility smirk in the VXX options market." *Applied Economics* 52: 769–788.
- Gharghori, P., E. D. Maberly, and A. Nguyen. 2017. "Informed Trading around Stock Split Announcements: Evidence from the Option Market." *Journal of Financial and Quantitative Analysis* 52: 705–735. doi:[10.1017/S0022109017000023](https://doi.org/10.1017/S0022109017000023).
- Goetzmann, W. N., and P. Jorion. 1993. "Testing the Predictive Power of Dividend Yields." *Journal of Finance* 48: 663–679. doi:[10.1111/j.1540-6261.1993.tb04732.x](https://doi.org/10.1111/j.1540-6261.1993.tb04732.x).
- Gourio, F. 2012. "Disaster Risk and Business Cycles." *American Economic Review* 102: 2734–2766. doi:[10.1257/aer.102.6.2734](https://doi.org/10.1257/aer.102.6.2734).
- Huang, Z., C. Tong, and T. Y. Wang. 2020. "Which Volatility Model for Option Valuation in China? Empirical Evidence from SSE 50 ETF Options." *Applied Economics* 52: 1866–1880. doi:[10.1080/00036846.2019.1679348](https://doi.org/10.1080/00036846.2019.1679348).
- Jiang, G. J. 2002. "What Could Potentially Aggravate Implicit Volatility 'Smile' and 'Asymmetry'?-a Note." *Applied Economics Letters* 9: 75–80. doi:[10.1080/13504850110052799](https://doi.org/10.1080/13504850110052799).
- Kim, S., N. Shephard, and S. Chib. 1998. "Stochastic Volatility: Likelihood Inference and Comparison with Arch Models." *Review of Economic Studies* 65: 361–393. doi:[10.1111/1467-937X.00050](https://doi.org/10.1111/1467-937X.00050).
- Li, J., S. A. Gehrke, and J. E. Zhang. 2019. "How Do US Options Traders 'Smirk' on China? Evidence from FXI Options." *Journal of Futures Markets* 39 (11): 1450–1470. doi:[10.1002/fut.22005](https://doi.org/10.1002/fut.22005).
- Lin, Q. 2018. "Technical Analysis and Stock Return Predictability: An Aligned Approach." *Journal of Financial Markets* 38: 103–123. doi:[10.1016/j.finmar.2017.09.003](https://doi.org/10.1016/j.finmar.2017.09.003).
- Lins, K. V., H. Servaes, and A. Tamayo. 2017. "Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility during the Financial Crisis." *Journal of Finance* 72: 1785–1824. doi:[10.1111/jofi.12505](https://doi.org/10.1111/jofi.12505).
- McLean, R. D., and J. Pontiff. 2016. "Does Academic Research Destroy Stock Return Predictability?" *Journal of Finance* 71: 5–32. doi:[10.1111/jofi.12365](https://doi.org/10.1111/jofi.12365).

- Moskowitz, T. J., Y. H. Ooi, and L. H. Pedersen. 2012. "Time Series Momentum." *Journal of Financial Economics* 104: 228–250. doi:[10.1016/j.jfineco.2011.11.003](https://doi.org/10.1016/j.jfineco.2011.11.003).
- Neely, C. J., D. E. Rapach, T. Jun, and G. Zhou. 2014. "Forecasting the Equity Risk Premium: The Role of Technical Indicators." *Management Science* 60: 1772–1791.
- Neumann, M., and G. Skiadopoulos. 2013. "Predictable Dynamics in Higher-order Risk-neutral Moments: Evidence from the S&p 500 Options." *Journal of Financial and Quantitative Analysis* 48: 947–977. doi:[10.1017/S002210901300032X](https://doi.org/10.1017/S002210901300032X).
- Newey, W. K., and K. D. West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55: 703–708. doi:[10.2307/1913610](https://doi.org/10.2307/1913610).
- Pan, J. 2002. "The Jump-risk Premia Implicit in Options: Evidence from an Integrated Time-series Study." *Journal of Financial Economics* 63: 3–50. doi:[10.1016/S0304-405X\(01\)00088-5](https://doi.org/10.1016/S0304-405X(01)00088-5).
- Rapach, D. E., M. C. Ringgenberg, and G. Zhou. 2016. "Short Interest and Aggregate Stock Returns." *Journal of Financial Economics* 121: 46–65. doi:[10.1016/j.jfineco.2016.03.004](https://doi.org/10.1016/j.jfineco.2016.03.004).
- Ruan, X. 2019. "Volatility-of-volatility and the Cross-section of Option Returns." *Journal of Financial Markets* 48: 100492.
- Ruan, X. F., and J. E. Zhang. 2018. "Risk-neutral Moments in the Crude Oil Market." *Energy Economics* 72: 583–600. doi:[10.1016/j.eneco.2018.04.026](https://doi.org/10.1016/j.eneco.2018.04.026).
- Welch, I., and A. Goyal. 2008. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21: 1455–1508. doi:[10.1093/rfs/hhm014](https://doi.org/10.1093/rfs/hhm014).
- Xing, Y., X. Zhang, and R. Zhao. 2010. "What Does the Individual Option Volatility Smirk Tell Us about Future Equity Returns?" *Journal of Financial and Quantitative Analysis* 45: 641–662. doi:[10.1017/S0022109010000220](https://doi.org/10.1017/S0022109010000220).
- Yan, S. 2011. "Jump Risk, Stock Returns, and Slope of Implied Volatility Smile." *Journal of Financial Economics* 99: 216–233. doi:[10.1016/j.jfineco.2010.08.011](https://doi.org/10.1016/j.jfineco.2010.08.011).
- Zhang, J. E., E. C. Chang, and H. Zhao. 2020. "Market Excess Returns, Variance and the Third Cumulant." *International Review of Finance* 20: 605–637.
- Zhang, J. E., H. Zhao, and E. C. Chang. 2012. "Equilibrium Asset and Option Pricing under Jump Diffusion." *Mathematical Finance* 22: 538–568. doi:[10.1111/j.1467-965.2010.00468.x](https://doi.org/10.1111/j.1467-965.2010.00468.x).
- Zhang, J. E., and Y. Xiang. 2008. "The Implied Volatility Smirk." *Quantitative Finance* 8: 263–284. doi:[10.1080/14697680601173444](https://doi.org/10.1080/14697680601173444).