A Forward-Looking Factor Model for Volatility: Estimation and Implications for Predicting Disasters*

Ohad Kadan†, Fang Liu‡, Xiaoxiao Tang§

October 2017

Abstract

We show that any factor structure for stock returns can be naturally translated into a factor structure for return volatility. We use this structure to propose a methodology for estimating forward-looking variances and covariances of both factors and individual assets from option prices at a high frequency. We implement the model empirically and show that our forward-looking volatility estimates provide useful predictions of rare disasters for both factors and individual stocks.

*We thank Phil Dybvig, Andrew Karolyi, and seminar participants at Cornell University, Southwest Institute of Finance and Economics, and Washington University in St. Louis for helpful comments and suggestions.

†Olin Business School, Washington University in St. Louis, kadan@wustl.edu.

‡Cornell SC Johnson College of Business, fang.liu@cornell.edu.

§Olin Business School, Washington University in St. Louis, xtang22@wustl.edu.

1 Introduction

Linear factor structures are serving as a central paradigm for modeling variations in stock returns. The idea is that all stock returns are driven by a small number of systematic factors along with an idiosyncratic component (e.g., Ross (1976)). Empirically, linear factor models have been a dominant workhorse in explaining stock returns.[1] In this paper we show that any given such factor model for stock returns translates naturally into a factor model for return volatility. This new factor model allows us to estimate the implied variances and covariances of factors from option-implied variances of individual assets. This gives rise to forward-looking moments for factors and individual assets even if they do not possess traded options. For example, we obtain forward-looking implied volatility estimates for the size, value, and momentum factors. Empirically, we show that such estimates are informative on the future returns of the factors, in particular when it comes to predicting strong negative outcomes and rare disasters.

We start with an arbitrary factor structure in which asset returns are a linear combination of the factors and an idiosyncratic return. Then, calculating the second moment of this relation directly leads to a new factor structure in which each asset's variance is represented as a linear combination of factor variances and covariances and in which the new factor loadings are simply products of the original factor loadings. Thus, if the original factor model has K factors, the new factor model for volatility consists of $\frac{K(K+1)}{2}$ factors. This approach also allows for a simple decomposition of the forward-looking variances to systematic and idiosyncratic components.

To estimate and exploit this factor structure we develop a modification of the Fama-MacBeth (1973) estimation procedure consisting of three stages. The first stage is identical to the standard Fama-MacBeth approach, in which we regress returns of test assets on the factors over time to estimate factor loadings. In the second stage we cross-sectionally regress option-implied variances of test assets on products of the factor loadings estimated in the first stage. This results in estimates of an implied variance-covariance matrix of the factors. In the third stage we use the factor loadings and the factors' implied variance-covariance matrix estimated in the previous stages along with the underlying factor structure to

[1] See, for example, Chen, Roll, and Ross (1986), Fama and French (1993), Carhart (1997), Hou, Xue, and Zhang (2015), and Fama and French (2016).

calculate the implied forward-looking covariances between any pair of individual assets.

Empirically, we apply our estimation method to all S&P 500 stocks over a sample period spanning January 1996 to December 2014, as this is the period during which option data is available to us. We consider two standard underlying factor models: the CAPM one-factor model and the Fama-French-Carhart (FFC) four-factor model. Time-series tests show that our option-implied volatility factor models explain a large portion of the variation in implied variances of both individual stocks and sectors, with average R-squared being about 55% for the CAPM model and 70% for the FFC four-factor model. Additional tests provide further support for the explanatory power of our model.

We next turn to demonstrating the usefulness of our framework by presenting three applications. For convenience, in all three applications we rely on the FFC four-factor model as a model for stock returns. The first application illustrates the value of using forward-looking implied variances and covariances obtained from our model for optimal diversification. Specifically, we use our estimated moments to draw a forward-looking efficient frontier for nine sectors and compare it to an efficient frontier based on historical moments. For illustrative purposes, we focus on September 2008—the month during which Lehman Brothers collapsed—as this is a natural candidate for a period in which historical moments may diverge from forward-looking ones. Our results show that at that time investors expected significantly higher correlations across sectors going forward than in the recent past. This, in turn, translates to inferior diversification opportunities going forward.

In the second application, we examine whether our estimated implied systematic and idiosyncratic variances contain information on future systematic and idiosyncratic jumps in stock prices. Intuitively, higher implied variances are expected to predict a higher likelihood of jumps. We define a systematic jump as a jump in the stock price concurrent with a jump in at least one of the factors, and an idiosyncratic jump as a jump in the stock price without a concurrent jump in any of the factors. We find that higher implied systematic variance predicts a higher likelihood of both systematic and idiosyncratic jumps in the cross section, whereas higher implied idiosyncratic variance predicts higher likelihood of idiosyncratic jumps only. Our results hold for downward jumps and upward jumps separately as well as when we consider them together.

In the third application, we consider the implied-volatility slope based on our model.

The slope has been studied extensively in the literature (e.g., Xing, Zhang, and Zhao (2010) and Yan (2011)) and is often considered a measure of forward-looking downside risk or negative sentiment. Our methodology offers a way to estimate the slope for factors and assets that do not possess traded options (such as the momentum factor). We estimate the implied-volatility slope for the FFC four factors by applying our modified Fama-MacBeth estimation approach to option-implied volatilities at different moneyness levels. This is particularly useful for the size, value, and momentum factors given that there do not exist options actively traded on these factor portfolios. We test whether our estimated slopes predict future downward jumps (disasters) and find confirming results for the market and momentum factors, and to some extent also for the value factor.

Our paper is related to several strands of literature. First, it is related to the large body of research on the factor structure of asset returns originated by Ross (1976) and Merton (1973b). Empirically, researchers have proposed a variety of factor models to explain variations in stock return (e.g., Chen, Roll, and Ross (1986), Fama and French (1993), Carhart (1997), Hou, Xue, and Zhang (2015), and Fama and French (2016)). Our paper helps make a connection between the factor structure of stocks returns and a corresponding factor structure for volatility. Since option prices reflect investors' expectations on future asset prices, the implied volatility factor structure allows us to gauge forward-looking information on individual stock as well as the underlying factors.

The paper also contributes to the recent growing literature on factor structures for option prices. Bakshi, Kapadia and Madan (2003) derive a skew law based on option prices and decompose individual return skewness into a systematic component and an idiosyncratic component. Duan and Wei (2009) show that systematic risk affects the implied volatility level and slope of individual stock options. Serban, Lehoczky, and Seppi (2008) and Christoffersen, Fournier, and Jacobs (2017) develop option valuation models that capture the relative pricing of individual asset returns and systematic factors, and empirically estimate their models using a single market factor. These papers mainly focus on a one-factor framework. Instead, our approach starts from a given factor model for stock returns (with any number of factors) and translates it to a factor model for volatility. Thus, we accommodate popular equity factor models proposed in the literature and show how they are related to factors predicting option prices.

The paper also adds to the literature on estimating jump risk using option prices. Cox and Ross (1976) and Merton (1976a,b) extend the Black and Scholes (1973) model by incorporating jumps in option valuation. A number of papers then show that incorporating jumps helps explain the observed option prices (e.g., Ball and Torous (1985), Naik and Lee (1990), Amin and Ng (1993), Bakshi, Cao, and Chen (1997), Bates (2000), Duffie, Pan, and Singleton (2000), Anderson, Benzoni, and Lund (2002), Pan (2002), and Eraker, Johannes, and Polson (2003)). Xing, Zhang, and Zhao (2010) and Yan (2011) show that option-implied jump risk predicts stock returns. Our factor structure allows us to examine systematic and idiosyncratic jump risk of individual stocks and to estimate the jump risk for factors and individual assets that do not possess traded options.

The rest of the paper proceeds as follows. Section 2 develops a factor structure for implied return volatility and proposes the modified Fama-MacBeth approach for estimation. In Section 3 we conduct the estimation exercise using S&P 500 constituent stocks base on the CAPM one-factor model and the FFC four-factor model. Section 4 examines the time-series performance of our implied-volatility factor structure. Section 5 presents the applications, where we apply our framework to study optimal diversification and to predict jumps for both factors and individual stocks. Section 6 concludes. A technical proof is delegated to the Appendix.

2 The Factor Structure: Theoretical Framework and Estimation Methodology

In this section we derive the factor structure for option-implied volatility and explain how it can be estimated using a modification of the Fama-MacBeth (1973) approach.

2.1 Theoretical Framework

Consider an economy with N assets $n = 1, 2, ..., N$. Assume that the returns of all assets follow a linear K-factor structure[2]

$$\mathbf{r} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{f} + \boldsymbol{\varepsilon}, \tag{1}$$

[2] Hereafter we use boldface to designate vectors and matrices.

where $\mathbf{r} = (r_1, ..., r_N)'$ is a column vector of asset returns, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)'$ is a vector of intercepts, $\mathbf{f} = (f_1, ..., f_K)'$ is column vector of factors, $\boldsymbol{\beta} = (\beta_{n,k})$ is an $N \times K$ matrix of factor loadings, and $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_N)'$ is a column vector of idiosyncratic returns. We assume that $\mathbf{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, $\mathbf{E}(\mathbf{f}\boldsymbol{\varepsilon}') = \mathbf{0}$, and that $\mathbf{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')$ is a diagonal matrix.

Taking the variance on both sides of (1) yields

$$\boldsymbol{\Sigma}^r = \boldsymbol{\beta}\boldsymbol{\Sigma}^f\boldsymbol{\beta}' + \boldsymbol{\Sigma}^\varepsilon, \tag{2}$$

where $\boldsymbol{\Sigma}^r, \boldsymbol{\Sigma}^f$, and $\boldsymbol{\Sigma}^\varepsilon$ are the variance-covariance matrices of the asset returns, factors, and idiosyncratic returns, respectively. Therefore, if the equity returns follow the linear factor structure (1), then the variances and covariances of equity returns also follow a factor structure in which the variances and covariances of the return factors serve as new factors and the variance of the idiosyncratic returns serve as an intercept.[3]

To illustrate this new factor structure consider the diagonal of $\boldsymbol{\Sigma}^r$, which consists of the variances of the N assets, $\boldsymbol{\Sigma}^r_{n,n}$. Then, (2) implies that for each $n = 1, ..., N$, the variance of asset n is given by the new factor structure

$$\Sigma^r_{n,n} = \sum_{k=1}^{K}\sum_{l=1}^{K} \beta_{n,k}\beta_{n,l}\Sigma^f_{k,l} + \Sigma^\varepsilon_{n,n}. \tag{3}$$

Namely, if we start with a given factor structure for returns consisting of K factors, then we obtain a new factor structure for variances of returns consisting of $K(K+1)/2$ factors, of which K are factor variances and $\frac{K(K-1)}{2}$ are factor covariances. Moreover, the factor loadings in the new model are products of the original factor loadings, and the new intercept is the variance of the idiosyncratic return.

Similarly, if we consider the off-diagonal terms in (2) we obtain that for any two assets m and n, the covariance between the returns, $\boldsymbol{\Sigma}^r_{m,n}$, follows the new factor structure

$$\Sigma^r_{m,n} = \sum_{k=1}^{K}\sum_{l=1}^{K} \beta_{m,k}\beta_{n,l}\Sigma^f_{k,l}, \tag{4}$$

with the same $K(K+1)/2$ factors as above but different loadings and no intercept.

[3] Note that the assumptions above imply that $\boldsymbol{\Sigma}^\varepsilon$ is diagonal, and so only the variances play a role in the intercept.

The relationship in (2) should hold for both historical and future return variances and covariances, as long as returns (both past and future) are governed by the factor structure (1). A conventional way to estimate future volatilities is by assuming some option-pricing model (such as Black-Scholes) and then applying this model to derive implied volatilities. Suppose we had options that would allow us to estimate implied variances and covariances for the N assets, the K factors, and for the idiosyncratic returns, then (2) would imply that

$$\mathbf{V}^r = \boldsymbol{\beta}\mathbf{V}^f\boldsymbol{\beta}' + \mathbf{V}^\varepsilon, \tag{5}$$

where $\mathbf{V}^r$, $\mathbf{V}^f$, and $\mathbf{V}^\varepsilon$ are the option-implied variance-covariance matrices for the asset returns, factors, and idiosyncratic returns, respectively.

Equation (5) is the fundamental factor relation we explore in this paper. The challenge in estimating and applying (5) is that in reality most of its ingredients are not easy to calculate from available option prices. For example, while implied volatilities can be estimated from options on individual stocks, there is no easy way to estimate implied covariances between assets. Furthermore, most conventional factors (such as the Fama-French-Carhart factors) and idiosyncratic returns do not have associated options, and even if we had such options, it is not clear how one would calculate the implied covariances. We next present a methodology for estimating these latent implied variances and covariances.

2.2 Estimation Methodology: A Modified Fama-MacBeth Approach

The standard Fama-MacBeth (1973) approach starts from a given factor model and follows two stages to estimate the risk premia associated with the factors. In the first stage, returns of each test asset (or portfolio) are regressed on the factors over time to estimate factor loadings (betas). In the second stage, test-asset (or portfolio) returns are regressed on the factor loadings to estimate risk premia. We argue that a modification of the Fama-MacBeth approach consisting of three stages can be used to estimate all of the ingredients in the new factor structure presented in (5).

The data required is a time series of the factor returns as well as a (possibly unbalanced) panel of the returns of N test assets where $N >> K(K+1)/2$. We require that each of the test assets has traded options so that we have data on its implied volatility – the diagonal elements of the matrix $\mathbf{V}^r$. Practically, in our empirical implementation we use the constituents of the S&P 500 index as test assets.

The first stage is identical to the original Fama-MacBeth approach, assuming that factor loadings are fixed over some period of time $t = 1, ..., \tau$. We use a time-series regression in which we regress the returns of the test assets on the factors over the period $t = 1, ..., \tau$ to obtain estimates of the factor loadings $N \times K$ matrix $\hat{\boldsymbol{\beta}}_\tau$.

In the second stage we estimate cross-sectional regressions of the implied variances of the test assets as of time τ, $V^r_{n,n,\tau}$, on products of the estimated factor loadings from the first stage. Specifically, the second stage model estimates the cross-sectional model

$$V^r_{n,n,\tau} = \lambda_\tau + \sum_{k=1}^{K}\sum_{l=1}^{K} V^f_{k,l,\tau}\hat{\beta}_{n,k,\tau}\hat{\beta}_{n,l,\tau} + \eta_{n,n,\tau}, \tag{6}$$

subject to the constraint that the matrix $\hat{\mathbf{V}}^f_\tau = \left(\hat{V}^f_{k,l,\tau}\right)_{k,l=1,...,K}$ is positive semidefinite. We assume that the regressors are of full rank for regularity.

Thus, for each time τ we obtain an estimate $\hat{V}^f_{k,l,\tau}$ of $V^f_{k,l,\tau}$, the implied covariance between factor k and factor l at time τ. The intercept in this regression $\hat{\lambda}_\tau$ is an estimate of the cross-sectional average of implied idiosyncratic variances as of time τ. The underlying assumption in (6) is that the factor loadings are uncorrelated with the error term $\eta_{n,n,\tau}$, which reflects unobserved determinants of idiosyncratic variance. The reason for introducing the constraint is that a variance-covariance matrix must be positive semidefinite.[4] We prove in the Appendix that this constrained regression has a unique solution.

In the third stage we use the estimated factor loadings $\hat{\boldsymbol{\beta}}_\tau$ and implied factor variance-covariance matrix $\hat{\mathbf{V}}^f_\tau$ to estimate the off-diagonal elements of $\mathbf{V}^r_\tau$ by applying the relation in (5), i.e.,

$$\hat{V}^r_{m,n,\tau} = \sum_{k=1}^{K}\sum_{l=1}^{K} \hat{V}^f_{k,l,\tau}\hat{\beta}_{m,k,\tau}\hat{\beta}_{n,l,\tau} \text{ for } m \neq n. \tag{7}$$

The three stages are then repeated for different times τ, using a standard "rolling window" approach, yielding a time series of implied variance-covariance matrix for the factors $\hat{\mathbf{V}}^f_\tau$ and the individual assets $\hat{\mathbf{V}}^r_\tau$ as well as a time series of the average cross-sectional idiosyncratic variances $\hat{\lambda}_\tau$.

[4] Practically, we impose the positive semidefiniteness constraint by requiring all eigenvalues of $\hat{\mathbf{V}}^f_\tau$ to be nonnegative.

A key by-product of this approach is that it also allows us to decompose the implied variance of each asset into a systematic component and an idiosyncratic component. Based on (6), the estimated implied systematic variance of asset n at time τ is

$$\hat{V}^s_{n,n,\tau} = \sum_{k=1}^{K}\sum_{l=1}^{K} \hat{V}^f_{k,l,\tau}\hat{\beta}_{n,k,\tau}\hat{\beta}_{n,l,\tau}, \tag{8}$$

and the estimate of the implied idiosyncratic variance of asset n at time τ is given by the total implied variance less the estimate of the systematic part, i.e.,

$$\hat{V}^\varepsilon_{n,n,\tau} = V^r_{n,n,\tau} - \sum_{k=1}^{K}\sum_{l=1}^{K} \hat{V}^f_{k,l,\tau}\hat{\beta}_{n,k,\tau}\hat{\beta}_{n,l,\tau}. \tag{9}$$

To summarize, we have shown that any given factor structure on returns translates naturally to a factor structure on variances and covariances of individual assets' returns. Our modified Fama-MacBeth method exploits cross-sectional variations in the factor loadings and in the implied volatility of the test assets to recover the implied variance-covariance matrix of the factors themselves along with the implied variance-covariance matrix of the individual assets. Moreover, this method allows us to decompose the implied volatility of each asset into systematic and idiosyncratic components. With these estimates at hand, we can go back to (5) and apply the new factor structure to any asset, including assets with no traded options.

3 Model Estimation

In this section we implement the framework discussed above. We begin with conventional factor models for the equity market and use our modified Fama-MacBeth approach to estimate the parameters of a factor model for implied volatility.

3.1 Data and Descriptive Statistics

Our option data is drawn from OptionMetrics for the period January 1996 to December 2014. As test assets we use all S&P 500 constituent stocks within the sample period with options traded on them. This leaves us with a total of 975 stocks. We obtain the implied volatility from the volatility surface file in OptionMetrics, which contains the Black-Merton-Scholes (BMS) implied volatility (Black and Scholes (1973), Merton (1973a)) for European options and the Cox-Ross-Rubinstein (CRR) implied volatility (Cox, Ross and Rubinstein (1979)) for American options. We take the average implied volatility of the at-the-money call and put options (delta equal to 0.5) that mature in 30 days (similar to Yan (2011), An et al. (2014), and Christoffersen, Fournier, and Jacobs (2017)). We use the square of this average implied volatility as the option-implied variance of equity returns. We obtain Fama-French factor data from Kenneth French's website. To estimate the return factor betas we use daily stock returns from CRSP.

We begin by exploring the implied volatility of our test assets – the S&P 500 constituents between 1996 and 2014. Figure 1 plots the cross-sectional average of option-implied variances of S&P 500 stocks over time as well as the option implied variance of the S&P500 (obtained from the SPX ETF). The average of option-implied variances of individual stocks fluctuates considerably during the 19-year sample period, ranging from below 0.1 to over 1. It has experienced two periods of substantial increases, 1998–2002 and 2008–2009, respectively. The former is relatively mild, corresponding to the Asian crisis and the subsequent internet bubble, whereas the latter is more dramatic, corresponding to the most recent financial crisis. The fact that the option-implied variances of individual stocks and the S&P 500 index follow similar trends suggests the market volatility as a primary factor for individual stock option-implied volatility.

Table 1 presents summary statistics of option-implied variances of the test assets by year. The number of stocks included in our analyses in each year varies from 610 to 745. Within each year, we compute the average of daily option-implied variances for each stock. The table then reports the cross-sectional mean, median, standard deviation, minimum and maximum of the average option-implied variances of all stocks for that year. Consistent with Figure 1, the cross-sectional mean and median of the option-implied variance are substantially higher during the 1998–2002 and 2008–2009 periods. The mean option-implied variance is consistently higher than the median, suggesting that the option-implied variance is right-skewed cross-sectionally. In addition, the cross-sectional standard deviation of the option-implied volatility follows the same time trend as the mean and median, i.e., it increases substantially during 1998–2002 and 2008–2009 and stays low in the rest of the sample. This suggests that when the market goes through crises, not only do investors consider all stocks to be more risky, they also perceive a higher level of variation in terms of how risky each individual stock is.

3.2 Modified Fama-MacBeth Estimation

Our approach enables one to transform a given factor model for the stock returns into a factor model for implied variances and covariances of returns. We apply this approach to two standard factor models for stock returns. First is the traditional CAPM one-factor model, where the unique factor is the market excess return ($mktrf$), and second is the Fama-French-Carhart four-factor model (Fama and French (1993), Carhart (1997), hereafter FFC), in which the factors are $mktrf$, as well as size (smb), value (hml), and momentum (umd).[5] The CAPM one-factor model transforms into a one-factor model for implied variances-covariances, whereas the FFC model transform into ten-factor models for implied variances-covariances.

3.2.1 First Stage Estimation

For the first stage we estimate the factor loadings using a rolling-window approach. Specifically, at the end of each month we estimate the factor loadings by regressing the returns of the test assets on daily factors over the preceding one-year window (250 trading days). The factor loadings estimated from these regressions are then used for all second stage cross-sectional regressions in the following month.

Table 2 reports summary statistics for the time-averages of the factor loadings estimated in the first stage for the test assets as well as all the relevant cross products. The table reports the cross-sectional mean, median, standard deviation, minimum and maximum of these time averages for each of the factor models. In all cases we see that the average market beta is close to 1, while other betas can be either positive or negative as expected. Note that when calculating the factor loading associated with cross-products, we multiply each product by 2 to comply with the symmetry in (5).

[5] We also checked the robustness of our results using the four-factor model recently introduced by Hou, Xue and Zhang (2015), and found similar results to those obtained using the FFC model. These results are available upon request.

3.2.2 Second Stage Estimation

For the second stage we estimate daily constrained cross-sectional regressions as in (6). Our dependent variable is the daily implied variance of the test assets, and the independent variables are products of the factor-loadings estimated in the first stage. Each such daily regression results in estimates of the requisite implied variances and covariances of the factors, which are the new daily factors for implied volatility.

Table 3 reports averages of the implied factor variances and covariances as estimated in the second-stage regressions for each of the factor models. To account for potential autocorrelation, we use Newey-West standard errors with 12 lags.[6] Column (1) corresponds to the CAPM one-factor model. The estimated implied variance of the market excess return is strictly positive for all dates. The time average of this estimated variance is 0.0778 over our sample period, which is significantly positive at the 1% level. The intercept, which represents the cross-sectional mean of the implied idiosyncratic variance, has an average of 0.0740 over time, also significantly positive at the 1% level.

Column (2) presents results for our ten-factor model based on the FFC four-factor model. Out of the 4802 days in our sample period, our estimated implied variance-covariance matrix of the factors is positive definite for 4733 days and positive semidefinite with a zero determinant for 69 dates.[7] The time averages of the estimated implied variances of the market, size, value and momentum factors are 0.0440, 0.0478, 0.0218 and 0.0713, respectively, all significantly positive at the 1% level. The implied covariances between two different factors, however, take different signs. For example, the estimated implied covariance between the market and size factors is 0.0152 on average, which is positive and strongly significant. By contrast, the implied covariance between the market and value factors has a negative average value of -0.0100, also strongly significant. The cross-sectional mean of the implied idiosyncratic variance is significantly positive and has an average value

[6] Following Stock and Watson (2011) page 599, we choose the number of lags for the Newey-West test based on the rule of thumb:

$$L = 0.75T^{1/3},$$

where L is the number of lags used and T is the number of observations in the time series. We have 4802 days in our sample, which leads to the use of 12 lags.

[7] A positive semidefinite variance-covariance matrix of the factors with a zero determinant means that the factors are linearly dependent, i.e., we can express one factor as the linear combination of the others.

of 0.0218.

3.2.3 Third Stage Estimation

In the first stage we have estimated at each time point τ, KN factor loadings, whereas in the second stage we have estimated $K(K+1)/2$ factor variances and covariances, for a total $K(2N+K+1)/2$ estimated parameters. In the third stage we use these estimates along with the factor structure to calculate the $N(N-1)/2$ implied covariances between the returns of all possible pairs of individual stocks based on (7). To illustrate, we have calculated these implied covariances (and correlations) for the 30 Dow Jones Industrial Average (DJIA) stocks (using index composition as of April 2016) based on the FFC model. We find an average (median) implied correlation of 0.39 (0.35) and a standard deviation of 0.33.[8] About 92% of the implied correlations are positive.

3.2.4 Variance Decomposition

Estimates of the factor loadings and the factor variance-covariance matrix allow us to decompose the implied total variance of each stock into the systematic and idiosyncratic components based on (8) and (9). Over our sample period and for all test assets, the average ratio of the implied systematic (idiosyncratic) variance to implied total variance is 60% (40%) under the CAPM model and 86% (14%) under the FFC model (last row of Table 3). The increased proportion of implied systematic variance under FFC compared to CAPM suggests that the additional factors indeed account for a substantial amount of the implied total variances of the test assets.

The cross-sectional average of the implied idiosyncratic variance is estimated as the intercept in the second-stage regressions. Figure 2 plots this average implied idiosyncratic variance obtained from the FFC four-factor model along with the average implied total variance over time. The figure shows that the average implied idiosyncratic variance is substantially lower than the average implied total variance and yet they follow similar time-series patterns. Herskovic, Kelly, Lustig, and Van Nieuwerburgh (2016) document similar patterns for realized volatilities.

[8] Our estimation approach does not guarantee that implied correlations lie between -1 and 1, but in practice over 90% of our correlation estimates indeed fall in this range. For the summary statistics we winsorize the estimates at -1 and 1.

3.2.5 Comparison to Realized Moments

Before proceeding to testing and using our estimates we explore whether the implied moments (variances and covariances) estimated using our modified Fama-MacBeth approach correlate with the corresponding historical moments. Note that our implied moments are forward looking, and thereby we should not expect them to be perfectly correlated with historical moments. Nevertheless, it is useful to compare the two to evaluate whether they share similar time-series characteristics.

We first plot the estimated implied factor variances and covariances against their realized counterparts over time at a monthly frequency.[9] Figure 3 considers the CAPM one-factor model. The top graph compares the implied market variance against the realized market variance. The two curves match closely, with our estimated implied variance being slightly higher than the realized variance during most of the sample period. In particular, the estimated implied variance captures both the mild spikes during 1998–2002 and the sharp increase during the 2008–2009 financial crisis. This suggests that our estimated implied variance shares some time-series properties with historical return volatility. In the bottom graph, we also compare our daily estimates of the implied variance for the market factor against the option-implied variance of the S&P 500 index (taken from the SPX ETF). Both the trend and magnitude of the two curves match closely.

Figure 4 presents similar plots based on the FFC four-factor model. The implied variances for all four factors are positive at all times as a result of the positive semidefiniteness constraint. Additionally, the implied market variance is quite similar to the realized variance. The implied variances of the other three factors also seem to capture the general trends in the corresponding realized variance despite larger discrepancies. The figure further shows that the six pairwise estimated implied covariances fluctuate between positive and negative numbers. These fluctuations are quite similar to the underlying realized covariances, with the market-value pair and the market-momentum pair having the closest match.

In Figure 5 we plot the estimated implied covariance between two individual stocks,

[9] Specifically, at the end of each month, we estimate realized factor variances and covariances based on daily factor returns within the month. For consistency, we construct implied variances and covariances at monthly frequency by averaging daily estimates from the second-stage modified Fama-MacBeth regressions within the month.

Microsoft and P&G, over time. We have chosen these two arbitrarily, and other pairs behave quite similarly. It can be seen that again the estimated implied covariance closely matches both the level and the trend of the realized covariance between the two stocks.

We have also calculated the correlations between the estimated moments and the realized ones. In almost all cases these correlations are highly positive and statistically significant (not reported for brevity).

4 Time Series Tests of the Model

So far we have estimated the parameters of our newly suggested factor model. In this section we use the estimated model to test its performance in explaining the time-series variations of implied-volatility of different assets. We use two sets of assets for this purpose. First are the nine SPDR sector ETFs, all of which have traded options. Second are the 30 DJIA constituent stocks as of April 2016.

4.1 Basic Tests

As a preliminary step we use standard tests to examine the performance of the CAPM and FFC factor models in explaining the returns (rather than implied variance) of the test assets. We regress (but do not tabulate) daily returns of each sector ETF on the return factors over time. This regression yields an average adjusted R-squared of 60% under the CAPM model and 66% under the FFC model. For the DJIA stocks, the average adjusted R-squared is 36% under CAPM and 39% under FFC. Thus, these factor models appear to do a descent job in explaining the time-series variation of our test-assets returns, although the explanatory power of the multi-factor model is not dramatically higher than that of the one-factor model.

We now turn to examining the time-series explanatory power of our implied-variance factor structure. Table 4 shows the adjusted R-squared from regressing the implied variance of each sector ETF on the estimated implied variances and covariances based on the CAPM one-factor model (Column (1)) and the FFC four-factor model (Column (2)). Under the CAPM model, the average adjusted R-squared is 56%. The highest and lowest R-squared values are 76% and 39% for the financial (XLF) and utility (XLU) sectors, respectively.

The FFC four-factor model induces a ten-factor model for implied variance consisting of four implied factor variances and six pairwise implied covariances. The average adjusted R-squared under the FFC model is 69%. The highest value is 89% for financials (XLF), indicating that our estimated implied variances and covariances of the FFC four factors together explain about 89% of the total time-series variation in the option-implied variance of the financial sector ETF. The lowest R-squared under the FFC model is 43% for utilities (XLU). Columns (1)–(2) of Table 5 report a similar analysis for DJIA stocks. Here the average adjusted R-squared is 56% under CAPM, and it rises to 72% under FFC.

Overall, these results suggest that our proposed volatility-implied factor model explains on average 56%–69% of the time-series variation of ETF implied variance, where the multi-factor model of FFC does a significantly better job than the one-factor model. For individual stocks the explanatory power of our model is similar ranging from 56% to 72%. Here again, the explanatory power of the multi-factor model is about 16% higher compared to the one-factor CAPM model.

4.2 Additional Tests

We next modify the model in various directions to assess its sensitivity to various assumptions we have made. First, we ask whether the implied covariances of different return factors contribute to the explanatory power of our factor structure, or whether they have just a trivial contribution and thus can be dropped to obtain a more parsimonious model. Namely, we explore whether the $K(K+1)/2$ factor model can be reduced into a model with just K factors without significantly reducing the time-series explanatory power.

To address this question, for the FFC model we eliminate the implied covariance terms in (6) and regress the option-implied variance of each sector ETF on the four implied variances only. The results are reported in Column (3) of Table 4. For most sectors, the adjusted R-squared obtained without the covariance terms is only slightly lower than that obtained with the covariance terms. However, for financials (XLF) and technology (XLK), the adjusted R-squared is substantially reduced after dropping the covariances. We also repeat the test for the DJIA stocks and see that the average adjusted R-squared drops from 72% to 60% when the covariance terms are excluded (see Column (3) of Table 5). These results suggest that the implied covariance terms indeed play an important role in explaining the time-series variations of the implied variance of the sector ETFs as well as individual stocks and hence should not be dropped.

The fundamental reason leading us to estimate the implied moments for the return factors using our modified Fama-MacBeth approach is that most return factors (such as SMB and HML) do not have options traded on them. This, however, is not the case for the market factor, for which a proxy for the implied variance can be directly obtained from S&P 500 options. To check the robustness of our results, we consider an alternative approach in which we replace our estimated market variance (obtained in Stage 2) with the implied variance from S&P 500 options. We repeat all our tests in Columns (1)–(3) of Tables 4 and 5, replacing our estimated implied market variance by the option-implied variance of the S&P 500 index obtained from OptionMetrics. The results are reported in Columns (4)–(6) of Table 4 for the sector ETFs and Columns (4)–(6) of Table 5 for the DJIA stocks.

When using the S&P 500 implied variance, the average adjusted R-squared for the sector ETFs is 71% under the CAPM model. Including the additional factors estimated using our approach increases the adjusted R-squared to 78% under the FFC model. For the DJIA stocks, the average adjusted R-squared values are 61% and 80% under CAPM and FFC, respectively. This suggests that the additional implied variances and covariances estimated using our modified Fama-MacBeth approach do indeed improve the time-series performance of the option factor model even when market variance itself is not based on our estimation approach.

As a final test we compare the implied systematic variances of sector ETFs estimated from our proposed factor model with the option-implied total variances obtained directly from OptionMetrics. The idea is that each sector ETF consists of a large number of stocks and thus can be viewed as a well-diversified portfolio with little idiosyncratic risk. Thus, if one assumes that the implied idiosyncratic variance of each sector ETF is low, then the implied systematic variance of the sector ETF should be close to the implied total variance.

Figure 6 plots the implied systematic variance of each sector ETF based on the FFC model against the corresponding implied total variance obtained directly from the ETF options over time. For all nine sectors, the two curves match in terms of both trend and magnitude, with systematic variance lying slightly below total variance most of the time.

This serves as additional supporting evidence for the performance of our option-implied volatility factor structure.

5 Applications

We next provide three applications for the model. We first show how the model can be used to construct a forward-looking Markowitz-style mean-variance efficient frontier. We use this to illustrate that the forward-looking correlation structure may contain useful information that is not reflected in correlation estimated based on historical data. Next, we apply the model to predict systematic and idiosyncratic jumps for individual stocks. Finally, we show how the model can be used to derive an implied volatility slope (smile) for the different factors (even those that do not have options written on them), and that this slope is often indicative of future downward jumps.

5.1 Optimal Diversification with Forward-Looking Moments

Consider a Markowitz mean-variance framework in which an investor's utility increases with the expected return and decreases with the variance of his portfolio and all investors agree on the first two moments of the asset returns. Then, the efficient frontier consists of all portfolios that minimize variance given any level of expected return, μ. Formally, a point on the efficient frontier is an $N \times 1$ column vector of portfolio weights $\mathbf{w}$ that minimize portfolio variance $\mathbf{w}'\mathbf{\Sigma}^r\mathbf{w}$ subject to $\mathbf{w}'\mathbf{E}(\mathbf{r}) = \mu$,and $\mathbf{w}'\mathbf{e} = 1$, where $\mathbf{\Sigma}^r$ is an $N \times N$ variance-covariance matrix of returns, $\mathbf{E}(\mathbf{r})$ is an $N \times 1$ column vector of expected returns, and $\mathbf{e}$ is an $N \times 1$ vector of ones.

A standard approach to solving for the efficient frontier is to use a historical variance-covariance matrix of asset returns. Instead, our modified Fama-MacBeth approach yields a forward-looking implied variance-covariance matrix of all assets, $\mathbf{V}^r$, and we can use it to build a forward-looking efficient frontier. Such an efficient frontier can be calculated at high frequency (e.g., daily or even intra daily) from option prices. This is in contrast to efficient frontiers based on historical variance-covariance matrices, which do not change much over short periods of time.

To illustrate, we use the modified Fama-MacBeth approach applied to the FFC model to construct an implied variance-covariance matrix associated with the nine SPDR sector ETFs as of September 2nd, 2008. This is the first trading day of the month during which Lehman Brothers collapsed, and it seems like a good choice for an illustration of the differences between forward-looking and historical information. For comparison purposes, we also compute the historical variance-covariance matrix based on daily equity returns of the most recent 20 trading days.

Table 6 reports the forward-looking and historical variance-covariance matrices in Panels A and B, respectively. The average value of forward-looking variances for the nine sector ETFs (diagonal elements of the variance-covariance matrix) is 0.0797, which is slightly lower than the average historical variance 0.0815. However, the forward-looking covariances (off-diagonal elements) are in general higher than the historical covariances, with averages of 0.0390 vs. 0.0259. This reflects that different sectors are expected to comove to a larger extent going forward than realized in the recent past. Such asset comovements are rather typical during crisis periods. Panels C and D report the forward-looking and historical correlation matrices of the various sectors. The results show that the forward-looking cross-sector correlation is on average 0.5736, which is 34.74% higher than the average historical correlation 0.4275.[10]

To construct the efficient frontier of sector ETFs we use the Arbitrage Pricing Theory (APT) to obtain the vector of expected returns using the FFC factor model:

$$\mathbf{E}(\mathbf{r}) = r_f \mathbf{e} + \boldsymbol{\beta}' \mathbf{E}(\mathbf{f}),$$

where r_f is the risk-free rate, and $\mathbf{E}(\mathbf{f})$ is the expectation of the FFC four factors. Empirically, for $\mathbf{E}(\mathbf{f})$ we use the annualized average returns of the FFC factor portfolios over the period July 1926 (the earliest time the FFC factor returns are available) through August 2008. For the risk-free rate we use the 30-day T-bill rate as of September 2nd, 2008.

Figure 7 depicts the two efficient frontiers derived based on our forward-looking and the historical variance-covariance matrices, respectively. We keep the sector expected returns

[10] The forward-looking variances are directly obtained from option-implied volatilities of the sector ETFs, whereas the forward-looking covariances between sectors are computed in the third stage of our modified Fama-MacBeth approach. There is intrinsically no guarantee that the resulting forward-looking correlation coefficients lie between -1 and 1. In fact, there are two pairs of sectors with the estimated forward-looking correlation slightly greater than 1. This is technically not feasible and could be due to noises in the option prices or estimation errors. We leave the numbers as they are in our illustration.

the same for both efficient frontiers, and hence any discrepancy between the two curves come from forward-looking versus historical information in the variance-covariance matrix. As discussed above, the forward-looking information reflects higher correlations between sectors, which in turn curtails investors' ability to diversify investment risk. Accordingly, the forward-looking efficient frontier lies strictly within the bounds of the historical one.

5.2 Predicting Systematic and Idiosyncratic Jumps

Our modified Fama-MacBeth approach allows us to decompose the implied total variance of an asset into a systematic component (8) and an idiosyncratic component (9), which intuitively should contain forward-looking information on future systematic and idiosyncratic volatilities, respectively.

The literature has considered two main sources of volatility. First is "diffusion" volatility, which is typically associated with a continuous process such as a Brownian motion. Second is "jump" volatility, which is associated with discrete changes in price. There is a vast literature dedicated to forecasting continuous volatility (see for example Poon and Granger (2003)), but there is very little in the literature about forecasting jumps. Below we test the usefulness of our implied systematic and idiosyncratic variances for the prediction of future systematic and idiosyncratic jumps of stocks.

We consider both upward and downward jumps. For the four FFC factors as well as the individual stocks, we define a downward (upward) jump as a monthly return below the 5th percentile (above the 95th percentile) of the corresponding historical return distribution over our sample period. Based on this definition, the downward jump cutoffs for the market, size, value and moment factors are -7.26%, -4.17%, -4.27% and -7.58%, and the upward jump cutoffs for the four factors are 7.13%, 4.92%, 4.79% and 6.79%, respectively. The downward and upward jump cutoffs for all S&P 500 stocks in our sample pooled together over our sample period are -15.99% and 18.46%, respectively. We further define a systematic jump of a stock as a jump in the stock price concurrent with a jump in at least one of the four factors, and we define an idiosyncratic jump as a jump in the stock price without a concurrent jump in any of the four factors. We represent the systematic (idiosyncratic) jumps by a dummy variable $J^s_{n,\tau}$ ($J^\varepsilon_{n,\tau}$) which equals one if a systematic (idiosyncratic) jump occurs in stock n at time τ, and zero otherwise.

To gauge the predictive power of the forward-looking systematic and idiosyncratic variances on future systematic jumps we estimate the following cross-sectional regressions at each month τ

$$J^s_{n,\tau+1} = \delta^s_{0,\tau} + \delta^s_{s,\tau}\left(V^s_{n,n,\tau} - \Sigma^s_{n,n,\tau}\right) + \delta^s_{\varepsilon,\tau}\left(V^\varepsilon_{n,n,\tau} - \Sigma^\varepsilon_{n,n,\tau}\right) + controls + \phi^s_{n,\tau+1},$$

where $V^s_{n,n,\tau}$ and $V^\varepsilon_{n,n,\tau}$ are the implied systematic and idiosyncratic variances of asset n in month τ, and $\Sigma^s_{n,n,\tau}$ and $\Sigma^\varepsilon_{n,n,\tau}$ are the realized systematic and idiosyncratic variances computed from daily equity returns based on the FFC model.[11] We subtract the realized variances from the implied ones because they are highly correlated. We thus use the variance spreads in order to focus on the predictive power of the implied variances beyond the historical realized quantities (similar to Bali and Hovakimian (2009)). We then conduct a t-test on the estimated $\delta^s_{s,\tau}$ and $\delta^s_{\varepsilon,\tau}$ over time to see if they are significantly different from zero. We repeat our test for downward jumps, upward jumps as well as downward and upward jumps mixed together. The results are shown in Table 7.

We start by using the implied total variance spread as the explanatory variable without decomposing it into the systematic and idiosyncratic parts. As shown in Panel A, the implied total variance spread does not predict future systematic jumps in the cross section regardless of whether we use downward, upward or mixed jumps. Controlling for the CAPM beta, firm size, book-to-market ratio, and lagged stock return does not change the results.

We then decompose the implied total variance spread into the systematic and idiosyncratic parts and include both components in our cross-sectional regressions. Panel B shows that the implied systematic variance spread positively predicts systematic jumps in the cross section, whereas the implied idiosyncratic variance spread does not have a significant effect. These results remain qualitatively the same for different types of jumps and are not affected by including controls in the regression.

Next we repeat our test to predict idiosyncratic jumps in the cross section. For each month τ we estimate the following cross-sectional regressions

$$J^\varepsilon_{n,\tau+1} = \delta^\varepsilon_{0,\tau} + \delta^\varepsilon_{s,\tau}\left(V^s_{n,n,\tau} - \Sigma^s_{n,n,\tau}\right) + \delta^\varepsilon_{\varepsilon,\tau}\left(V^\varepsilon_{n,n,\tau} - \Sigma^\varepsilon_{n,n,\tau}\right) + controls + \phi^\varepsilon_{n,\tau+1}.$$

[11] The implied systematic and idiosyncratic variances $V^S_{n,n,\tau}$ and $V^\varepsilon_{n,n,\tau}$ can be estimated on a daily basis. Here we use the values as of the last day of each month τ. Using average values within each month does not change our results.

Table 8 shows the results. The implied total variance spread positively predicts all types of idiosyncratic jumps in the following month. When decomposing the total variance spread into a systematic part and an idiosyncratic part, we see that both components have a strongly positive effect, and the effect from the systematic variance spread indeed has a larger economic magnitude.

Overall, our results show that the implied systematic variance is positively informative on future systematic as well as idiosyncratic jumps. In contrast, the implied idiosyncratic variance is only positively informative on the realization of future idiosyncratic jumps.

5.3 Factor Implied Volatility Slope and Jump Risk

While the Black-Scholes option pricing model assumes that the return volatility of the underlying asset is fixed across all strike prices, empirical evidence has shown that the implied volatility exhibits a negative slope (smirk) with out-of-the-money (OTM) put implied volatility being consistently higher than at-the-money (ATM) call implied volatility for both index options and individual stock options (e.g., Bate (1991, 2003), Pan (2002), Bollen and Whaley (2004), and Gârleanu, Pedersen and Poteshman (2009)). A commonly accepted explanation for the negative implied volatility slope is the downside jump risk (e.g., Bate (1991), Pan (2002), Xing, Zhang, and Zhao (2010), and Yan (2011)). The idea is that the OTM put options are useful investments tools for investors to express their concern on potential downward jumps (disasters) of the underlying assets. When the downward jump risk is high, the prices of OTM put options become more expensive, pushing up OTM put implied volatilities.

In this section, we apply our modified Fama-MacBeth approach to estimate the implied volatility slopes of the factor portfolios base on the FFC model and examine how these estimated implied volatility slopes can inform us on the occurrence of disasters of the corresponding factors. This is apparently interesting for the size, value and momentum factors for which no traded options are available. This is also interesting for the market factor, even through the S&P 500 index is actively traded in the option market as a proxy for the market portfolio. An important difference between equity prices and option prices is in terms of how individual prices aggregate into the market price. For equity prices, this aggregation is mechanical, and so individual stock prices do not provide additional information on the market beyond what is reflected in the price of the market index. This is, however, not the case for option prices. There is no mechanical aggregation of individual stock option prices into the market option price. As a result, there might exist additional information on the market contained in individual option prices not reflected in the option prices of the market index. Our estimation allows us to explore this additional information.

Following Xing, Zhang, and Zhao (2010), we define the implied volatility slope as the difference between implied volatilities of ATM call and OTM put options with one month to maturity, i.e.,

$$Slope = V^{ATMC} - V^{OTMP}.$$

The implied volatilities of OTM put options reflect concerns related to downward jumps, whereas ATM call options are used as a benchmark since they are the most liquid. Below we estimate the implied volatility slope for each of the FFC four factors.

5.3.1 Implied Volatility Slope of the Market Factor

We start with the market factor. The ATM call implied volatility of the market portfolio can be estimated by our modified Fama MacBeth approach using ATM call implied volatilities of all test assets. To estimate the OTM put implied volatility of the market, we repeat the modified Fama-MacBeth approach with OTM put implied volatilities of all test assets. In particular, we choose a delta value of -0.2 for OTM put options, corresponding to a strike price that is roughly 5% lower than the stock price assuming zero risk-free rate and implied volatility of 0.22 (a typical value for the implied volatility). The idea is that if the values of all assets in the market lose about 5% within one month, then the market overall should also lose 5%, representing a downward jump of the market factor. As a result, estimating the market OTM implied volatility using individual stock implied volatilities at a delta of -0.2 is appropriate.

Figure 8 shows daily estimates of the market implied volatility slope exhibiting wide fluctuations over time. Table 9 reports summary statistics for the estimated market implied-volatility slope. The mean (median) is -0.0109 (-0.0102), consistent with the negative slope documented in the literature, known as the volatility smile. The standard deviation of the estimated slope is 0.0239.

Given the discussion in the previous section we next test whether our estimated market implied-volatility slope can predict future downward jumps of the market factor portfolio. As in Section 5.2, we define a downward jump dummy as being equal to one if the monthly return of the market factor portfolio is lower than -7.26% (the 5th percentile of the historical return distribution over our sample period) and zero otherwise. We regress this dummy on the market implied volatility slope estimated from the previous month. Since this slope can be estimated on a daily basis, for robustness we use two ways to measure the monthly slope. First is to use the estimated slope as of the last trading day of each month, and second is to use the average estimated slope within the month. As reported in Table 9, the regression coefficient is negative and statistically significant based on both month-end and month-average slopes. Thus a more negative market implied-volatility slope is associated with a higher likelihood of a market downward jump. In particular, a one standard deviation increase in the month-end slope is associated with roughly 4.5% drop in the probability of a future downward jump in the market.

For comparison, we also compute the implied volatility slope of the S&P 500 index, which is directly available from S&P 500 index options. As reported in Table 9, the mean, median, and standard deviation of the slope are -0.0473, -0.0436, and 0.0197, respectively. Unlike the previous results, when we regress the market downward jump dummy on the lagged S&P 500 implied volatility slope we do not find a significant coefficient. This suggests that our estimated market implied volatility slope performs better than the S&P 500 implied volatility slope in predicting future market disasters. Apparently, individual stock option prices can be aggregated to generate market-wide information, which is not reflected in S&P 500 index options.

5.3.2 Implied Volatility Slope of the Size, Value and Momentum Factors

In the case of the market factor, one can find close proxies such as the S&P500 for which options are actively traded, and so one readily calculate the implied volatility. By contrast, for the size, value and momentum factors, such options are not in existence. Our approach facilitates the calculation of the implied volatility slope in these cases as well as for other factor models. Indeed, the ATM call implied volatility can be estimated in the same way as for the market factor by applying the modified Fama-MacBeth procedure to ATM call implied volatilities of the test assets. The estimation of the OTM put implied volatility is trickier, and we discuss it below. We use the size factor (SMB) to illustrate the main idea, but the estimation for the value and momentum factors is parallel.

The size factor is a long-short portfolio with long positions in small firms and short positions in big firms. To estimate the OTM implied volatility of the size factor portfolio, we cannot use implied volatilities of all test assets at the same delta of -0.2 as we did with the market factor since small and big firms typically have inverse correlation patterns with the SMB factor. This is in contrast to the market factor where nearly all firms have positive betas. To deal with this issue we apply the modified Fama-MacBeth approach to OTM put implied volatilities of all small firms at a delta of -0.2 (corresponding to a monthly return of around -5%) and OTM call implied volatilities of all big firms at a delta of 0.2 (corresponding to a monthly return of around 5%). The idea is that if all small firms lose 5% and all big firms earn 5% at the same time, then overall the SMB portfolio would experience a downward jump. To apply this procedure we need to differentiate between "small" and "big" firms. We do this based on the stock return correlation with the SMB factor. Namely, we define a firm as "small" if its SMB beta is positive and define it as "big" otherwise.

Figure 8 plots daily estimates of the implied volatility slope for the size, value and momentum factors, and Table 9 reports corresponding summary statistics. Also reported in the table are results from regressing downward jump dummies of the three factors on their corresponding estimated implied volatility slopes over time. The regression coefficient for the momentum factor is significantly negative based on both month-end and month-average slopes, indicating that our estimated implied volatility slope predicts future downward jumps of the momentum factor. A one standard deviation increase in the month-end slope is associated with roughly 7.7% drop in the probability of a future disaster in the momentum factor. In addition, the regression coefficient is also significantly negative for the value factor based on the month-average slope. A one standard deviation increase in the slope is associated with roughly 4.4% drop in the probability of a future disaster in the value factor. The results for the size factor are also negative as expected, but they are not statistically significant.

6 Conclusion

To be completed.

Appendix

Let $\mathbf{B}$ be the matrix consisting of the regressors in (6). That is, $\mathbf{B}$ is the matrix of all products of factor loading estimators and a constant. In this matrix, each row represents an asset, and columns represent different combinations of the factor loadings as well as a constant. This is an $N \times \left(\frac{K(K+1)}{2} + 1\right)$ matrix with a generic form (we drop time subscripts for brevity):

$$\mathbf{B} = \begin{pmatrix} \hat{\beta}_{1,1}^2 & \hat{\beta}_{1,1}\hat{\beta}_{1,2} & \cdots & \hat{\beta}_{1,k}\hat{\beta}_{1,l} & \cdots & \hat{\beta}_{1,K}^2 & 1 \\ \hat{\beta}_{2,1}^2 & \hat{\beta}_{2,1}\hat{\beta}_{2,2} & \cdots & \hat{\beta}_{2,k}\hat{\beta}_{2,l} & \cdots & \hat{\beta}_{2,K}^2 & 1 \\ \vdots & & & \vdots & & & \vdots \\ \hat{\beta}_{n,1}^2 & \hat{\beta}_{n,1}\hat{\beta}_{n,2} & \cdots & \hat{\beta}_{n,k}\hat{\beta}_{n,l} & \cdots & \hat{\beta}_{n,K}^2 & 1 \\ \vdots & & & \vdots & & & \vdots \\ \hat{\beta}_{N,1}^2 & \hat{\beta}_{N,1}\hat{\beta}_{N,2} & \cdots & \hat{\beta}_{N,k}\hat{\beta}_{N,l} & \cdots & \hat{\beta}_{N,K}^2 & 1 \end{pmatrix}.$$

Let $\mathbf{V}$ be a symmetric $K \times K$ matrix and let λ be a constant. Then, we can identify with $(\mathbf{V}, \lambda)$ a column vector $\mathbf{p}_{\mathbf{V},\lambda}$ of dimension $\left(\frac{K(K+1)}{2} + 1\right) \times 1$ by simply stacking the rows of the upper right triangle of $\mathbf{V}$ and then adding λ as follows

$$\mathbf{p}_{\mathbf{V},\lambda} = (V_{1,1}, V_{1,2}, ..., V_{k,l}, ..., V_{K,K}, \lambda)'.$$

Let $\mathbb{V}$ denote the set of $K \times K$ positive semi-definite matrices. Given $(\mathbf{V},\lambda) \in \mathbb{R}^{\frac{K(K+1)}{2}+1}$, define the least-squares loss function $L : \mathbb{R}^{\frac{K(K+1)}{2}+1} \to \mathbb{R}$ as

$$L(\mathbf{V}, \lambda) = \left\| Diag(\mathbf{V}^r) - \mathbf{B}\mathbf{p}_{\mathbf{V},\lambda} \right\|^2,$$

where $Diag(\mathbf{V}^r)$ is the $N \times 1$ vector consisting of the diagonal elements of $\mathbf{V}^r$ representing implied variances of the test assets, and $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^N$.

Then, the constrained least squares optimization problem corresponding to the regression model in (6) is

$$\min_{\mathbf{V} \in \mathbb{V}, \lambda \in \mathbb{R}} L(\mathbf{V}, \lambda). \tag{10}$$

Theorem 1 *Suppose that* $\mathbf{B}'\mathbf{B}$ *is of full rank. Then, there exists a unique solution to Problem (10).*

We first prove the following two lemmas.

Lemma 1 $\mathbb{V}\times\mathbb{R}$ *is a convex and closed (in the Euclidean topology) subset of* $\mathbb{R}^{\frac{K(K+1)}{2}+1}$.

Proof of Lemma 1: It is sufficient to show that $\mathbb{V}$ is convex and closed in $\mathbb{R}^{\frac{K(K+1)}{2}}$.

To see the convexity of $\mathbb{V}$, consider $\mathbf{V}^1, \mathbf{V}^2 \in \mathbb{V}$ and a nonzero $K \times 1$ vector $\mathbf{x}$. Let $\alpha \in [0,1]$ and let $\mathbf{V}^* = \alpha\mathbf{V}^1 + (1-\alpha)\mathbf{V}^2$. Then,

$$\mathbf{x}'\mathbf{V}^*\mathbf{x} = \alpha\mathbf{x}'\mathbf{V}^1\mathbf{x} + (1-\alpha)\,\mathbf{x}'\mathbf{V}^2\mathbf{x} \geq 0,$$

where the inequality follows from positive semidefiniteness of $\mathbf{V}^1$ and $\mathbf{V}^2$.

To see the closedness of $\mathbb{V}$, let $\{\mathbf{V}^i\}$ be a sequence of matrices in $\mathbb{V}$ converging to a matrix $\mathbf{V}^*$ in the Euclidean topology. Then, for each i and for any nonzero $K \times 1$ vector $\mathbf{x}$ we have $\mathbf{x}'\mathbf{V}^i\mathbf{x} \geq 0$. Since Euclidean convergence implies convergence of each element of the matrices we have $\mathbf{x}'\mathbf{V}^*\mathbf{x} \geq 0$, as needed.

Lemma 2 *Let* $\mathbf{B}$ *be an* $N \times \left(\frac{K(K+1)}{2}+1\right)$ *matrix such that* $\mathbf{B}'\mathbf{B}$ *is of full rank. Then, there exists an* $\varepsilon > 0$ *such that* $\|\mathbf{Bq}\| > \varepsilon$ *for all* $\mathbf{q} \in \mathbb{R}^{\frac{K(K+1)}{2}+1}$ *with* $\|\mathbf{q}\| = 1$.

Proof of Lemma 2: Suppose this is not the case, then there exists a sequence $\{\mathbf{q}^i\} \in \mathbb{R}^{\frac{K(K+1)}{2}+1}$ such that $\|\mathbf{q}^i\| = 1$ for all i and $\|\mathbf{Bq}^i\| \to 0$. Since the set of all $\mathbf{q} \in \mathbb{R}^{\frac{K(K+1)}{2}+1}$ such that $\|\mathbf{q}\| = 1$ is compact in the Euclidean norm, there is a convergent subsequence $\{\mathbf{q}^{i_j}\}$ converging to $\mathbf{q}^*$ such that $\|\mathbf{q}^*\| = 1$. However,

$$\|\mathbf{Bq}^*\| = \lim_{j\to\infty}\left\|\mathbf{Bq}^{i_j}\right\| = \lim_{i\to\infty}\left\|\mathbf{Bq}^i\right\| = 0,$$

where the first equality follows by the continuity of the Euclidean norm and matrix multiplication, and the second and third since $\|\mathbf{Bq}^i\|$ is converging to zero. But, $\|\mathbf{Bq}^*\| = 0$ if and only if each of the elements of $\mathbf{Bq}^*$ is zero. It follows that $(\mathbf{B}'\mathbf{B})\,\mathbf{q}^*$ is a vector of zeros in $\mathbb{R}^{\frac{K(K+1)}{2}+1}$. But, since $\|\mathbf{q}^*\| = 1$, we have that $\mathbf{q}^*$ is not a vector of zeros. This contradicts the assumption that $\mathbf{B}'\mathbf{B}$ is of full rank.

Proof of Theorem 1: L is strictly convex in $\mathbb{R}^{\frac{K(K+1)}{2}+1}$, and by Lemma 1 the domain of minimization is a convex set in $\mathbb{R}^{\frac{K(K+1)}{2}+1}$. Hence, if a solution to (10) exists, it must be unique. Therefore, we are left to prove that a solution exists.

Let

$$\underline{L} = \inf_{\mathbf{V}\in\mathbb{V},\lambda\in\mathbb{R}} L\left(\mathbf{V},\lambda\right).$$

Then, there is a sequence $\left\{L^i\right\}$ converging to $\underline{L}$, and a sequence of $\left\{\mathbf{V}^i,\lambda^i\right\}$, where $\mathbf{V}^i \in \mathbb{V}$, such that $L^i = L\left(\mathbf{V}^i,\lambda^i\right)$. Let $\left\{\mathbf{p}_{\mathbf{V}^i,\lambda^i}\right\}$ denote the corresponding sequence of vectors in $\mathbb{R}^{\frac{K(K+1)}{2}+1}$. To prove existence of a solution of (10), it is sufficient to show that $\left\{\mathbf{p}_{\mathbf{V}^i,\lambda^i}\right\}$ has a bounded subsequence in the Euclidean topology. Indeed, if $\left\{\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\}$ is such a subsequence, then it has a convergent subsequence to a limit, which we denote by $\mathbf{p}^*$. Let $(\mathbf{V}^*,\lambda^*)$ be the corresponding matrix and constant. Since $\mathbb{V}$ is closed (by Lemma 1), we have that $\mathbf{V}^* \in \mathbb{V}$, and by the continuity of L, $\underline{L} = L\left(\mathbf{V}^*,\lambda^*\right)$. Therefore, $(\mathbf{V}^*,\lambda^*)$ is a solution to (10).

Thus, it remains to show that $\left\{\mathbf{p}_{\mathbf{V}^i,\lambda^i}\right\}$ has at least one bounded subsequence. Suppose to the contrary that all subsequences $\left\{\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\}$ are unbounded, i.e., $\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\| \to +\infty$. Consider the ratio

$$\begin{aligned}\frac{\sqrt{L\left(\mathbf{V}^{i_j},\lambda^{i_j}\right)}}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|} &= \frac{\left\|Diag\left(\mathbf{V}^r\right)-\mathbf{B}\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|} \\ &= \left\|\frac{Diag\left(\mathbf{V}^r\right)}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}-\frac{\mathbf{B}\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\|.\end{aligned}$$

Note that

$$\left\|\frac{\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\| = 1.$$

By Lemma 2, there is an $\varepsilon > 0$ such that for all i_j,

$$\left\|\frac{\mathbf{B}\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\| > \varepsilon.$$

It follows that

$$\begin{aligned}\left\|\frac{Diag\left(\mathbf{V}^r\right)}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}-\frac{\mathbf{B}\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\| &\geq \left\|\frac{\mathbf{B}\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\|-\left\|\frac{Diag\left(\mathbf{V}^r\right)}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\| \\ &> \varepsilon-\left\|\frac{Diag\left(\mathbf{V}^r\right)}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\|.\end{aligned}$$

When $j \to \infty$, $\frac{Diag(\mathbf{V}^r)}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}$ converges to $\mathbf{0}$ since $\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\| \to +\infty$. Thus, there exists a j_0 sufficiently large such that for all $j > j_0$,

$$\left\|\frac{Diag\left(\mathbf{V}^r\right)}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|}\right\| < \frac{\varepsilon}{2}.$$

We conclude that for $j > j_0$,

$$\frac{\sqrt{L\left(\mathbf{V}^{i_j},\lambda^{i_j}\right)}}{\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\|} > \frac{\varepsilon}{2}.$$

Since $\left\|\mathbf{p}_{\mathbf{V}^{i_j},\lambda^{i_j}}\right\| \to +\infty$, this implies that $L\left(\mathbf{V}^{i_j},\lambda^{i_j}\right) \to +\infty$. However, this contradicts that $L\left(\mathbf{V}^{i_j},\lambda^{i_j}\right) \to \underline{L} < +\infty$. Thus, $\left\{\mathbf{p}_{\mathbf{V}^i,\lambda^i}\right\}$ has at least one bounded subsequence. This completes the proof.

References

[1] Amin, Kaushik I., Victor K. Ng, 1993. Option valuation with systematic stochastic volatility. *Journal of Finance* 48, 881–910.

[2] 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.

[3] Anderson, Torben G., Luca Benzoni, Jesper Lund, 2002. An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239–1284.

[4] Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997. Empirical performance of alternative option pricing models. *Journal of Finance* 52, 589–667.

[5] Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003. Stock return characteristics, skew laws, and differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

[6] Bali, Turan G., and Armen Hovakimian, 2009. Volatility spreads and expected stock returns. *Management Science* 55,1797–1812.

[7] Ball, Clifford A., and Walter N. Torous, 1985. On jumps in common stock prices and their impact on call option pricing. *Journal of Finance* 40, 155–173.

[8] Bates, David S., 1991. The crash of '87: was it expected? The evidence from options markets. *Journal of Finance* 46, 1009–1044.

[9] Bates, David S., 2000. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics* 94, 181–238.

[10] Bates, David S., 2003. Empirical option pricing: a retrospection. *Journal of Econometrics* 116, 387–404.

[11] Black, Fischer, and Myron Scholes, 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.

[12] Bollen, Nicolas P. B., and Robert E. Whaley, 2004. Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* 59, 711–753.

[13] Carhart, Mark M., 1997. On persistence in mutual fund performance. *Journal of Finance* 52, 57–82.

[14] Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986. Economic forces and the stock market. *Journal of Business* 59, 383–403.

[15] Christoffersen, Peter, Mathieu Fournier, and Kris Jacobs, 2017. The factor structure in equity options. *Review of Financial Studies*, forthcoming.

[16] Cox, John C., Stephen A. Ross, 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3, 145–166.

[17] Cox, John C., Stephen A. Ross, and Mark Rubinstein, 1979. Option pricing: a simplified approach. *Journal of Financial Economics* 7, 229–263.

[18] Duan, Jin-Chuan, and Jason Wei, 2009. Systematic risk and the price structure of individual equity options. *Review of Financial Studies* 22, 1981–2006.

[19] Duffie, Darrell, Jun Pan, Kenneth Singleton, 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.

[20] Eraker, Bjørn, Michael Johannes, Nicholas Polson, 2003. The impact of jumps in returns and volatility. *Journal of Finance* 58, 1269–1300.

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

[22] Fama, Eugene F., and Kenneth R. French, 2016. Dissecting anomalies with a five-factor model. *Review of Financial Studies* 29, 69–103.

[23] Fama, Eugene F., and James D. MacBeth, 1973. Risk, return, and equilibrium: empirical tests. *Journal of Political Economy* 81, 607–636.

[24] Gârleanu, Nicolae, Lasse Heje Pedersen, and Allen M. Poteshman, 2009. Demand-based option pricing. *Review of Financial Studies* 22, 4259–4299.

[25] 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.

[26] Hou, Kewei, Chen Xue, and Lu Zhang, 2015. Digesting anomalies: an investment approach. *Review of Financial Studies* 28, 650–705.

[27] Kadan, Ohad, and Xiaoxiao Tang, 2016. A bound on expected stock returns. Working paper, Washington University in St. Louis.

[28] Martin, Ian, and Christian Wagner, 2016. What is the expected return on a stock? Working paper, London School of Economics and Copenhagen Business School.

[29] Merton, Robert C., 1973a. Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4, 141–183.

[30] Merton, Robert C, 1973b. An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society,* 867-887.

[31] Merton, Robert C., 1976a. Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics* 4, 125–144.

[32] Merton, Robert C., 1976b. The impact on option pricing of specification error in the underlying stock price returns. *Journal of Finance* 31, 333–350.

[33] Naik, Vasanttilak, and Moon Lee, 1990. General equilibrium pricing of options on the market portfolio with discontinuous returns. *Review of Financial Studies* 3, 493–521.

[34] Pan, Jun, 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.

[35] Poon, Ser-Huang, and Clive W.J. Granger, 2003. Forecasting volatility in financial markets: a review. *Journal of Economic Literature* 41, 478–539.

[36] Ross, Stephen A., 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13, 341–360.

[37] Serban, Mihaela, John Lehoczky, and Duane Seppi, 2008. Cross-sectional stock option pricing and factor models of returns. Working paper, Morgan Stanley and Carnegie Mellon University.

[38] Stock, James H., and Mark W. Watson, 2011. *Introduction to Econometrics (3rd edition)*. Addison Wesley Longman.

[39] Xing, Yuhang, Xiaoyan Zhang, and Rui Zhao, 2010. What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis* 45, 641–662.

[40] Yan, Shu, 2011. Jump risk, stock returns, and slope of implied volatility smile. *Journal of Financial Economics* 99, 216–233.

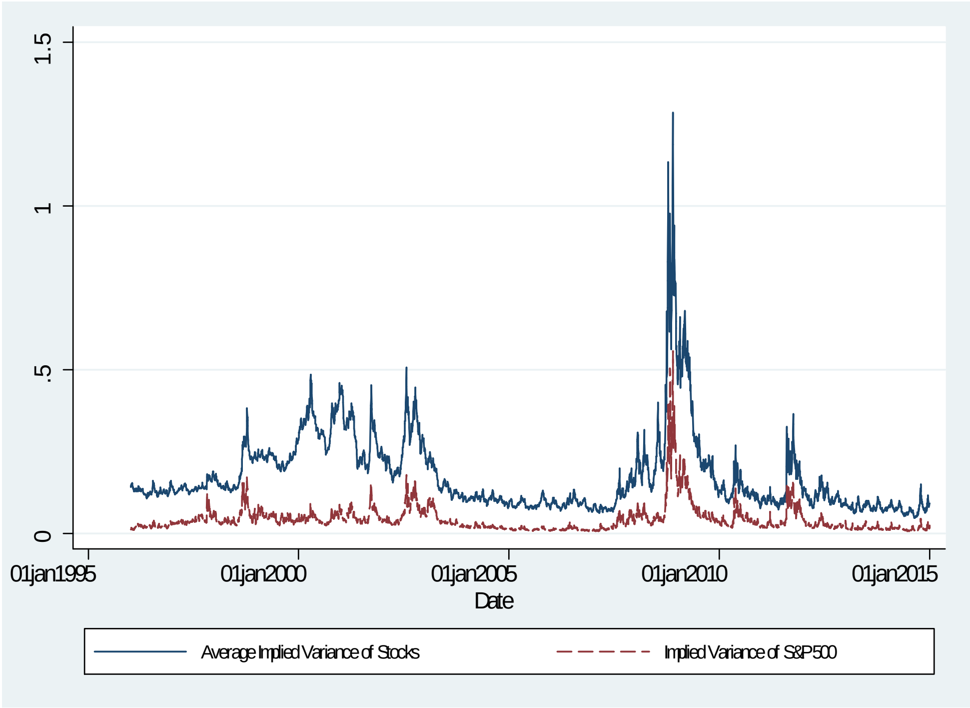


Figure 1: Option-Implied Variances of Stocks and S&P 500 Index

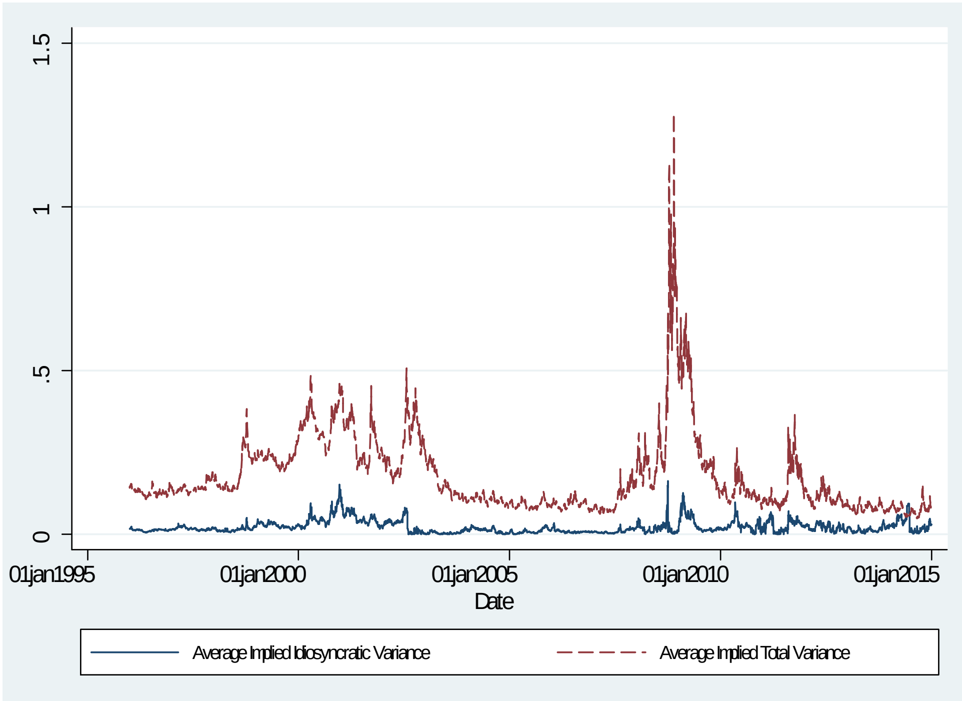
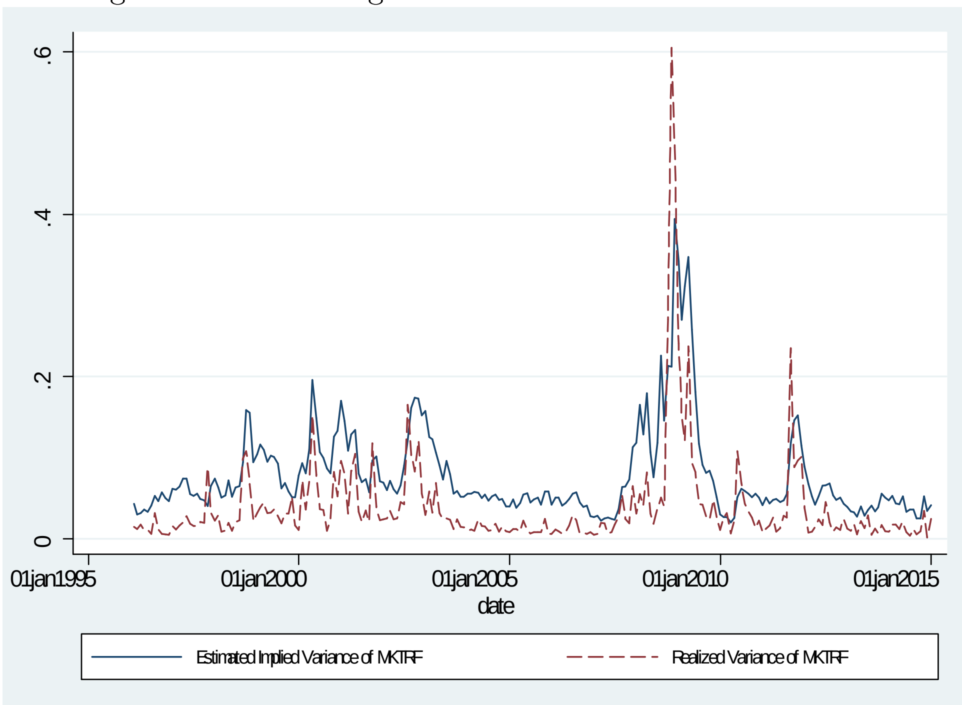


Figure 2: Average Implied Idiosyncratic vs. Total Variances Based on FFC Model

Figure 3: Second-Stage Estimation Based on CAPM Model



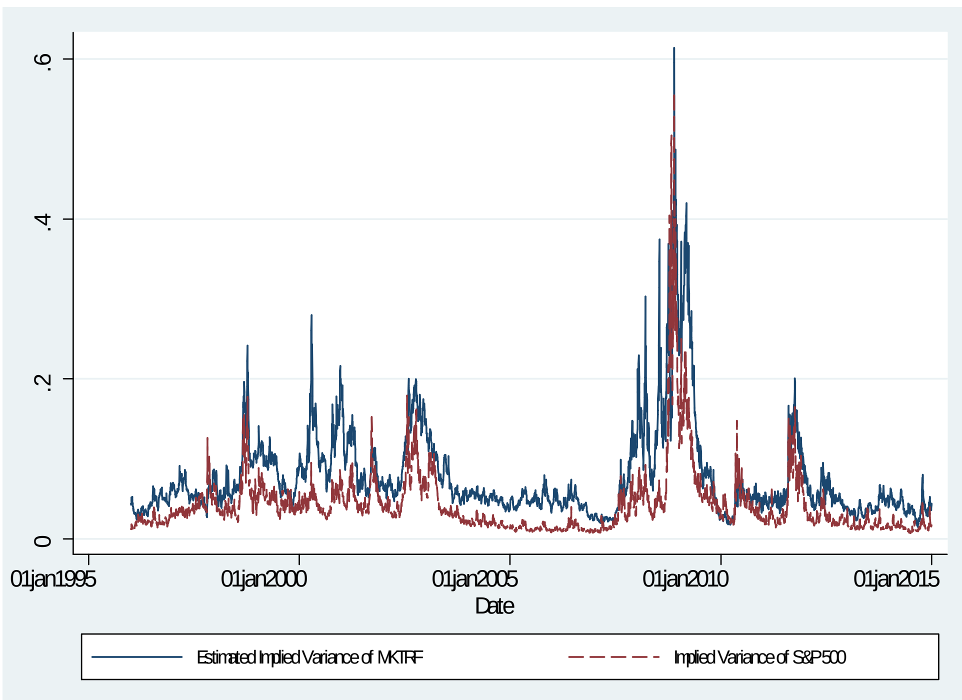


Figure 4: Second-Stage Estimation Based on FFC Model

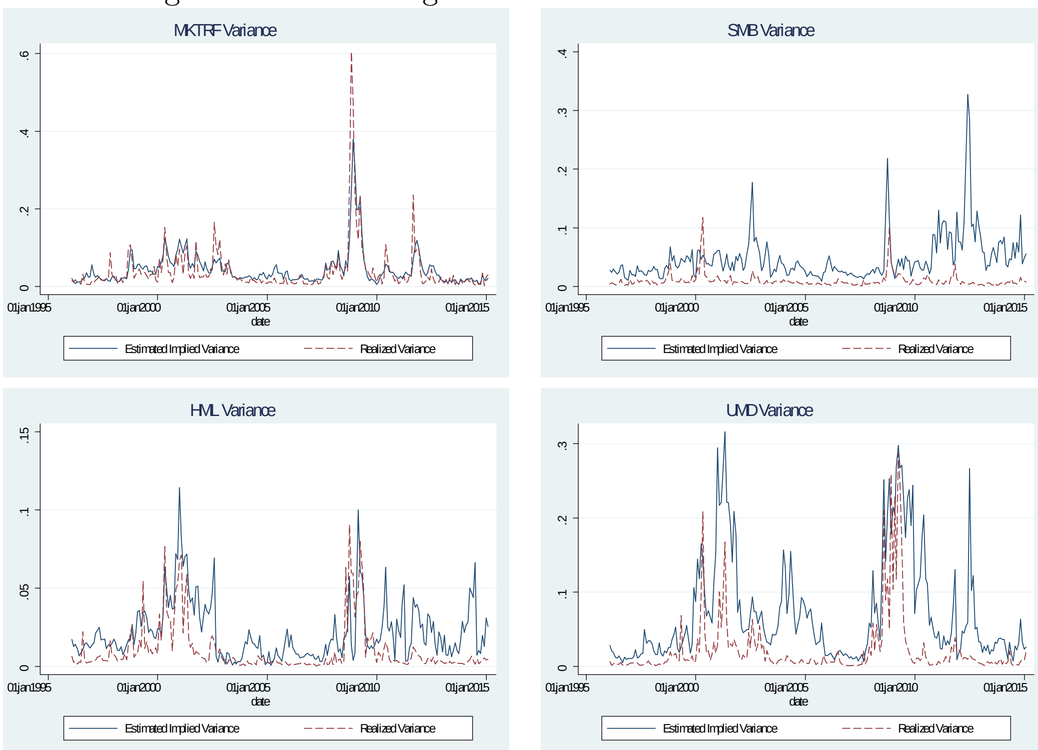


Figure 4: Second-Stage Estimation Based on FFC Model (Continued)

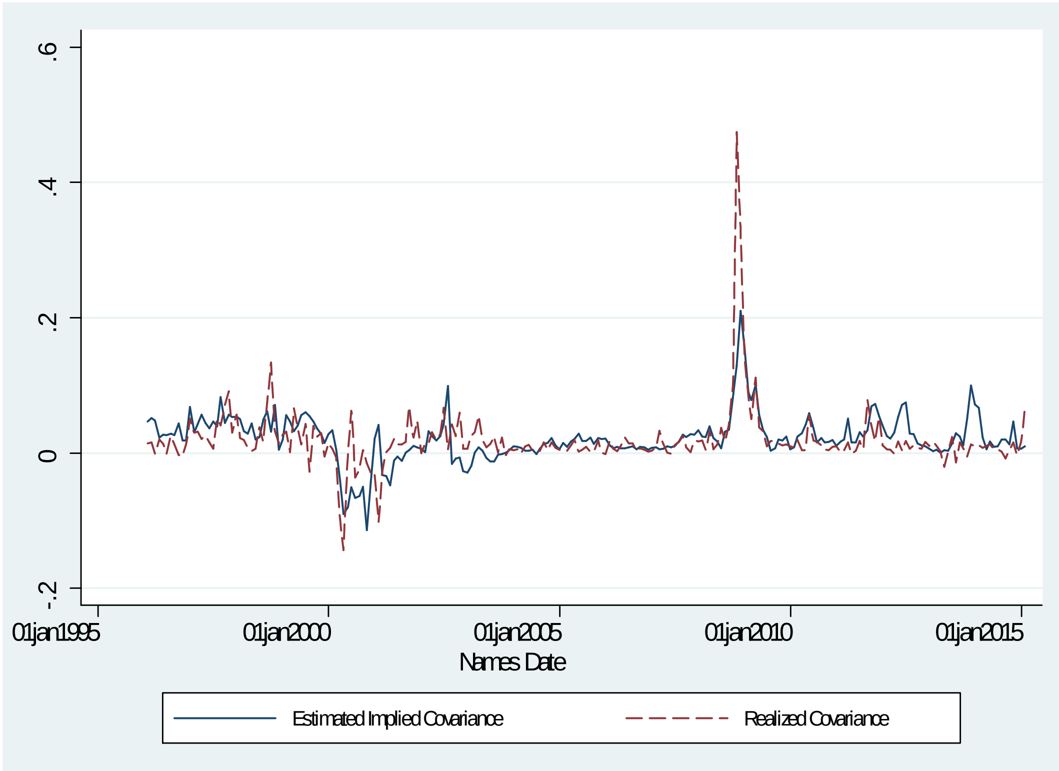
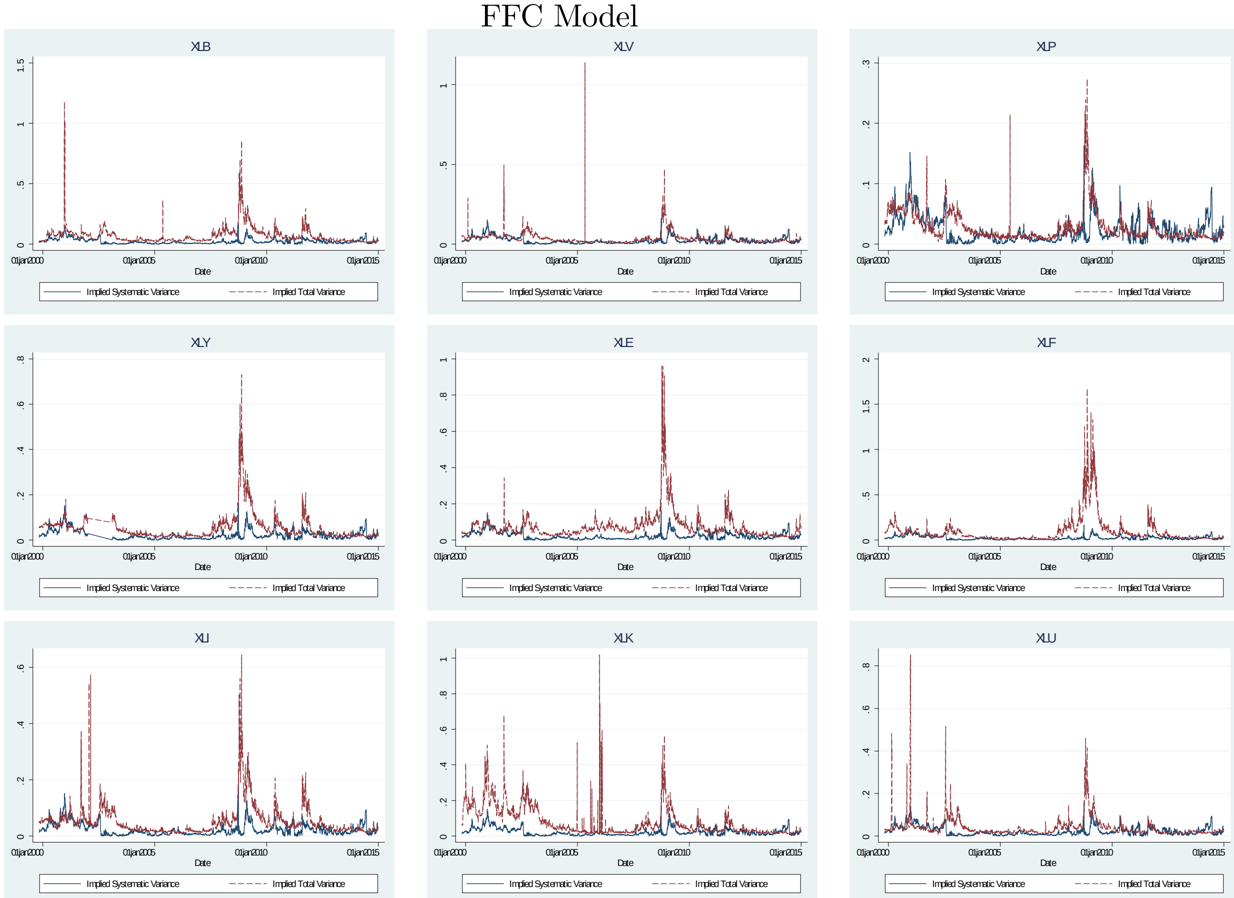


Figure 5: Implied Covariance of Microsoft and P&G

Figure 6: Implied Systematic vs. Total Variances of SPDR Sector ETFs Based on FFC Model



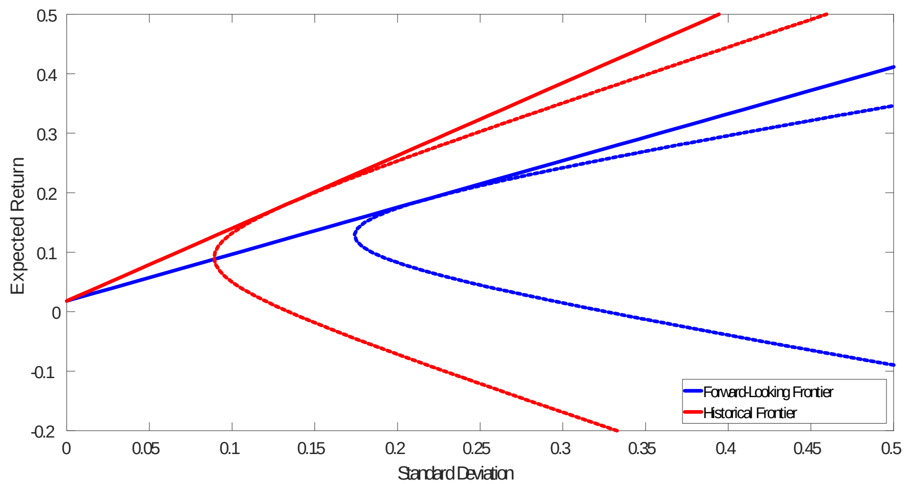


Figure 7: Forward-Looking vs. Historical Efficient Frontiers (9/2/2008)

Figure 8: Implied Volatility Slope of FFC Factor Portfolios

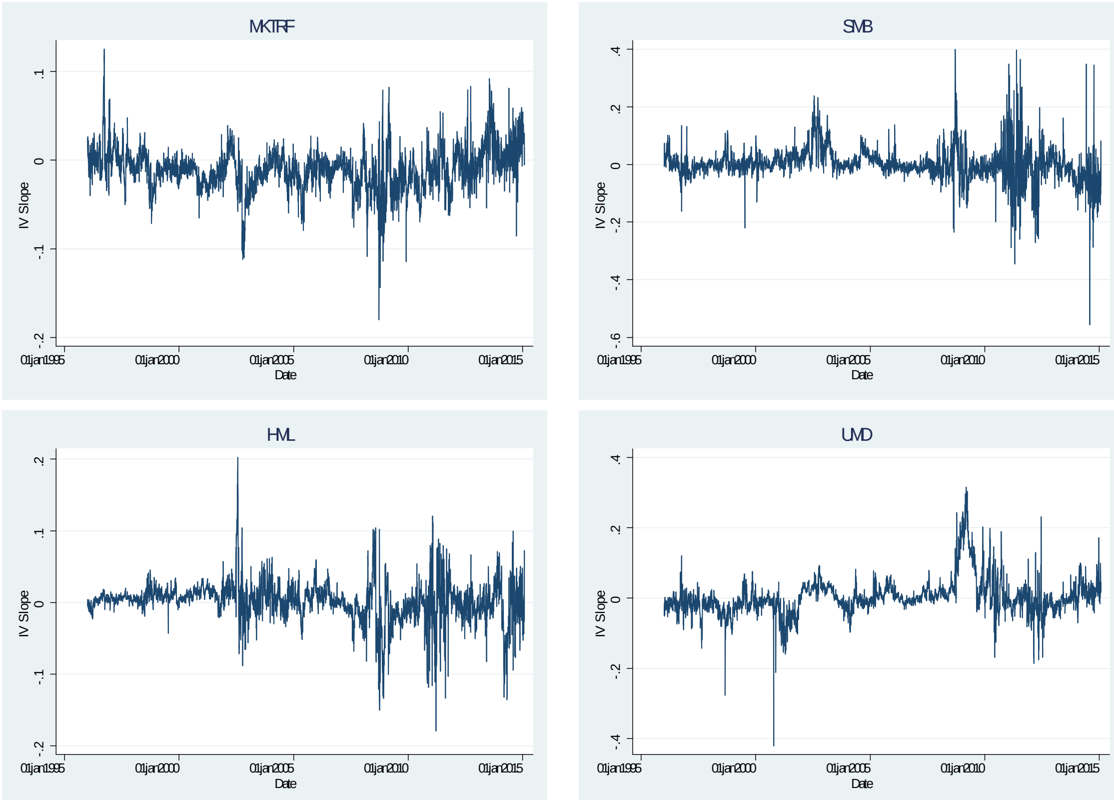


Table 1: Summary Statistics of Option-Implied Variances

This table reports summary statistics of option-implied return variances of all S&P 500 constituent stocks. Each year, we compute the average of daily option-implied variances for each stock. The table then reports the cross-sectional mean, median, standard deviation, minimum and maximum of this average option-implied variance by the year.

Year	Number of Stocks	Mean	Median	S.D.	Min	Max
1996	685	0.1320	0.0871	0.1181	0.0237	0.7841
1997	722	0.1503	0.0989	0.1858	0.0236	3.0220
1998	745	0.2034	0.1445	0.2817	0.0285	6.1896
1999	737	0.2343	0.1777	0.1836	0.0361	1.9525
2000	708	0.3386	0.2441	0.2568	0.0202	1.5579
2001	675	0.2951	0.1922	0.2721	0.0418	1.7791
2002	694	0.2778	0.1791	0.2663	0.0380	1.6042
2003	692	0.1559	0.1085	0.1432	0.0245	1.4683
2004	698	0.1064	0.0740	0.1056	0.0144	1.1918
2005	695	0.0932	0.0674	0.0926	0.0120	1.1272
2006	695	0.0925	0.0709	0.0724	0.0111	0.6782
2007	685	0.1059	0.0885	0.0698	0.0019	0.6004
2008	658	0.3619	0.2930	0.2761	0.0318	2.4714
2009	644	0.3364	0.2465	0.2997	0.0008	2.2244
2010	640	0.1347	0.1062	0.1229	0.0072	1.5028
2011	629	0.1472	0.1206	0.1336	0.0049	1.4820
2012	623	0.1148	0.0805	0.1530	0.0044	2.6541
2013	618	0.0809	0.0579	0.0834	0.0118	1.2612
2014	610	0.0755	0.0535	0.0831	0.0040	1.3157

Table 2: Summary Statistics of First-Stage Factor Loadings

This table reports summary statistics of factor loadings estimated in the first stage of the modified Fama-MacBeth procedure as well as all the relevant cross products under the CAPM one-factor model (Panel A) and the FFC four-factor model (Panel B). For each stock, we compute the average of each factor loading over time. The table reports the cross-sectional mean, median, standard deviation, minimum and maximum of these averages.

Panel A: CAPM Model					
	Mean	Median	S.D.	Min	Max
β_{mktrf}	1.0292	0.9926	0.3590	-0.0768	2.3396
β^2_{mktrf}	1.3074	1.0977	0.9065	0.0155	5.8187

Panel B: FFC Model					
	Mean	Median	S.D.	Min	Max
β_{mktrf}	1.0764	1.0748	0.2543	0.0265	2.0662
β_{smb}	0.1918	0.1670	0.3445	-0.8199	1.7685
β_{hml}	0.1729	0.2529	0.6120	-3.6693	2.7848
β_{umd}	-0.0836	-0.0688	0.2408	-1.4730	1.1197
β^2_{mktrf}	1.3186	1.2513	0.5980	0.0770	4.2698
β^2_{smb}	0.3189	0.1907	0.3713	0.0098	3.1955
β^2_{hml}	0.8052	0.4948	1.0645	0.0144	14.4982
β^2_{umd}	0.3528	0.2049	0.3873	0.0078	3.5822
$2\beta_{mktrf}\beta_{smb}$	0.5580	0.3755	0.9056	-1.4827	5.2218
$2\beta_{mktrf}\beta_{hml}$	0.5546	0.5857	1.3755	-4.4739	10.2483
$2\beta_{mktrf}\beta_{umd}$	-0.2253	-0.1750	0.6025	-3.9971	3.0061
$2\beta_{smb}\beta_{hml}$	0.2019	0.1168	0.6057	-2.8109	5.0804
$2\beta_{smb}\beta_{umd}$	-0.0696	-0.0222	0.3001	-2.7415	0.9653
$2\beta_{hml}\beta_{umd}$	-0.0043	-0.0274	0.6592	-9.2461	6.4819

Table 3: Second-Stage Modified Fama-MacBeth Estimation

This table reports time averages of the estimated implied variances and covariance of the factors from the second-stage modified Fama-MacBeth procedure based on the CAPM one-factor model (Column (1)) and the FFC four-factor model (Column (2)). The Newey-West standard errors with 12 lags are reported in the parentheses below the corresponding estimates. Asterisks denote statistical significance at the 1% (***), 5% (**) and 10% (*) levels. The table also shows the average ratio of implied systematic variance over implied total variance of the test assets based on each model.

	CAPM	FFC
$V^f_{mktrf,mktrf}$	0.0778	0.0440
	(0.0030)***	(0.0024)***
$V^f_{smb,smb}$		0.0478
		(0.0021)***
$V^f_{hml,hml}$		0.0218
		(0.0010)***
$V^f_{umd,umd}$		0.0713
		(0.0037)***
$V^f_{mktrf,smb}$		0.0152
		(0.0008)***
$V^f_{mktrf,hml}$		-0.0100
		(0.0011)***
$V^f_{mktrf,umd}$		-0.0052
		(0.0019)***
$V^f_{smb,hml}$		-0.0041
		(0.0008)**
$V^f_{smb,umd}$		-0.0038
		(0.0015)***
$V^f_{hml,umd}$		-0.0033
		(0.0010)***
λ	0.0740	0.0218
	(0.0034)***	(0.0010)***
V^s/V^r	0.5995	0.8637

Table 4: Time Series Adjusted R-Squared for SPDR Sector ETFs

This table reports the adjusted R-squared from regressing the option-implied variances of the nine SPDR sector ETFs on the estimated implied variances and covariances of the factors over time based on the CAPM one-factor model (Column (1)) and the FFC four-factor model (Column (2)). Columns (3) shows results from regressing the option-implied variances of the sector ETFs on the implied variances of the FFC four factors only. Columns (4)–(6) repeat all tests in Columns (1)–(3), replacing the estimated implied variance of the market factor by the option-implied variance of the S&P 500 index.

	Using Estimated Implied Mkt Var			Using Implied Var of S&P 500		
	CAPM	FFC		CAPM	FFC	
	(1)	(2)	(3)	(4)	(5)	(6)
XLB	0.5531	0.6692	0.6605	0.7122	0.7392	0.7311
XLV	0.4142	0.5125	0.4902	0.6168	0.6269	0.6153
XLP	0.6245	0.7123	0.6834	0.7539	0.8454	0.7770
XLY	0.7216	0.8342	0.8032	0.9456	0.9515	0.9466
XLE	0.5422	0.7688	0.7415	0.8110	0.8297	0.8191
XLF	0.7559	0.8925	0.7560	0.7387	0.8616	0.7635
XLI	0.6180	0.7676	0.7194	0.8996	0.9089	0.9038
XLK	0.3881	0.6414	0.4682	0.4160	0.7231	0.5149
XLU	0.3865	0.4260	0.4029	0.5231	0.5565	0.5272
Average	0.5560	0.6916	0.6362	0.7130	0.7825	0.7332

Table 5: Time Series Adjusted R-Squared for DJIA Stocks

This table reports the adjusted R-squared from regressing the option-implied variances of the 30 DJIA component stocks as of April 2016 on the estimated implied variances and covariances of the factors over time based on the CAPM one-factor model (Column (1)) and the FFC four-factor model (Column (2)). Columns (3) shows results from regressing the option-implied variances of the stocks on the implied variances of the FFC four factors only. Columns (4)–(6) repeat all tests in Columns (1)–(3), replacing the estimated implied variance of the market factor by the option-implied variance of the S&P 500 index.

	Using Estimated Implied Mkt Var			Using Implied Var of S&P 500		
	CAPM	FFC		CAPM	FFC	
	(1)	(2)	(3)	(4)	(5)	(6)
AXP	0.7607	0.8888	0.8095	0.7881	0.8804	0.8253
AAPL	0.2398	0.6681	0.3305	0.1867	0.6824	0.3152
BA	0.6348	0.7670	0.7119	0.7954	0.8910	0.8362
CAT	0.6517	0.7625	0.7372	0.8093	0.8632	0.8494
CHV	0.5784	0.7526	0.7463	0.8325	0.8491	0.8358
CSCO	0.2552	0.6988	0.3919	0.2263	0.7390	0.4102
KO	0.4932	0.6614	0.4874	0.5369	0.8120	0.5813
DIS	0.6239	0.7464	0.6582	0.7722	0.8703	0.7920
DD	0.6908	0.7792	0.7474	0.8161	0.8894	0.8567
XOM	0.5707	0.6928	0.6724	0.8228	0.8532	0.8242
GE	0.6613	0.8074	0.6788	0.6577	0.8231	0.7291
GS	0.6899	0.8490	0.8377	0.7256	0.7765	0.7570
HD	0.6763	0.7564	0.6715	0.7246	0.8660	0.7581
INTC	0.4237	0.5962	0.5115	0.2708	0.6363	0.5225
IBM	0.4442	0.7011	0.4762	0.4548	0.8083	0.5425
JPM	0.7982	0.8220	0.7049	0.7272	0.8414	0.7614
JNJ	0.3735	0.5277	0.3443	0.4276	0.6872	0.4593
MCD	0.4725	0.5580	0.3935	0.4622	0.6766	0.4760
MRK	0.6520	0.7454	0.6944	0.7343	0.8179	0.7569
MSFT	0.4834	0.7201	0.5314	0.4853	0.8158	0.6004
JPM	0.6050	0.8130	0.6448	0.6176	0.7968	0.6946
NKE	0.4587	0.6633	0.4963	0.4084	0.7122	0.5080
PFE	0.4708	0.5533	0.4498	0.5391	0.7060	0.5671
PG	0.4566	0.6641	0.4891	0.4733	0.7723	0.5463
MMM	0.6152	0.7952	0.7401	0.7444	0.8927	0.8126
STA	0.6454	0.7549	0.7048	0.7206	0.7784	0.7376
UTX	0.6085	0.7318	0.6510	0.7388	0.8515	0.7626
VZ	0.5327	0.5837	0.4624	0.5767	0.7174	0.5857
V	0.7261	0.7883	0.7535	0.7628	0.8148	0.7804
WMT	0.3919	0.6875	0.4173	0.3618	0.7679	0.4456
Average	0.5562	0.7179	0.5982	0.6067	0.7963	0.6643

Table 6: Forward-Look vs. Historical Covariances of SPDR Sector ETFs

This table reports the forward-looking covariance (Panel A), historical covariance (Panel B), forward-looking correlation (Panel C) and historical correlation (Panel D) matrices of the nine SPDR sector ETFs estimated as of September 2, 2008.

Panel A: Forward-Looking Covariance									
	XLB	XLV	XLP	XLY	XLE	XLF	XLI	XLK	XLU
XLB	0.0796	0.0222	0.0164	0.0193	0.0855	0.0043	0.0337	0.0433	0.0318
XLV	0.0222	0.0314	0.0169	0.0335	0.0240	0.0546	0.0261	0.0275	0.0173
XLP	0.0164	0.0169	0.0275	0.0360	0.0239	0.0681	0.0269	0.0259	0.0206
XLY	0.0193	0.0335	0.0360	0.0853	0.0016	0.1497	0.0509	0.0496	0.0226
XLE	0.0855	0.0240	0.0239	0.0016	0.1328	-0.0010	0.0421	0.0477	0.0726
XLF	0.0043	0.0546	0.0681	0.1497	-0.0010	0.2064	0.0866	0.0740	0.0505
XLI	0.0337	0.0261	0.0269	0.0509	0.0421	0.0866	0.0643	0.0420	0.0296
XLK	0.0433	0.0275	0.0259	0.0496	0.0477	0.0740	0.0420	0.0518	0.0286
XLU	0.0318	0.0173	0.0206	0.0226	0.0726	0.0505	0.0296	0.0286	0.0386

Panel B: Historical Covariance									
	XLB	XLV	XLP	XLY	XLE	XLF	XLI	XLK	XLU
XLB	0.0366	0.0103	0.0039	0.0053	0.0483	0.0165	0.0135	0.0147	0.0120
XLV	0.0103	0.0327	0.0279	0.0510	-0.0066	0.0632	0.0401	0.0261	0.0093
XLP	0.0039	0.0279	0.0358	0.0530	-0.0257	0.0681	0.0411	0.0278	0.0043
XLY	0.0053	0.0510	0.0530	0.1219	-0.0454	0.1448	0.0758	0.0456	0.0154
XLE	0.0483	-0.0066	-0.0257	-0.0454	0.1389	-0.0457	-0.0119	-0.0060	0.0242
XLF	0.0165	0.0632	0.0681	0.1448	-0.0457	0.2501	0.1007	0.0566	0.0176
XLI	0.0135	0.0401	0.0411	0.0758	-0.0119	0.1007	0.0680	0.0363	0.0120
XLK	0.0147	0.0261	0.0278	0.0456	-0.0060	0.0566	0.0363	0.0338	0.0096
XLU	0.0120	0.0093	0.0043	0.0154	0.0242	0.0176	0.0120	0.0096	0.0160

Panel C: Forward-Looking Correlation									
	XLB	XLV	XLP	XLY	XLE	XLF	XLI	XLK	XLU
XLB	1.0000	0.4433	0.3498	0.2340	0.8321	0.0338	0.4715	0.6751	0.5742
XLV	0.4433	1.0000	0.5764	0.6470	0.3723	0.6779	0.5802	0.6812	0.4961
XLP	0.3498	0.5764	1.0000	0.7441	0.3964	0.9045	0.6395	0.6877	0.6313
XLY	0.2340	0.6470	0.7441	1.0000	0.0155	1.1280	0.6872	0.7467	0.3944
XLE	0.8321	0.3723	0.3964	0.0155	1.0000	-0.0060	0.4550	0.5750	1.0131
XLF	0.0338	0.6779	0.9045	1.1280	-0.0060	1.0000	0.7514	0.7157	0.5657
XLI	0.4715	0.5802	0.6395	0.6872	0.4550	0.7514	1.0000	0.7273	0.5944
XLK	0.6751	0.6812	0.6877	0.7467	0.5750	0.7157	0.7273	1.0000	0.6394
XLU	0.5742	0.4961	0.6313	0.3944	1.0131	0.5657	0.5944	0.6394	1.0000

Panel D: Historical Correlation									
	XLB	XLV	XLP	XLY	XLE	XLF	XLI	XLK	XLU
XLB	1.0000	0.2988	0.1087	0.0788	0.6778	0.1725	0.2713	0.4169	0.4944
XLV	0.2988	1.0000	0.8152	0.8078	-0.0974	0.6987	0.8513	0.7870	0.4060
XLP	0.1087	0.8152	1.0000	0.8031	-0.3648	0.7200	0.8340	0.7993	0.1777
XLY	0.0788	0.8078	0.8031	1.0000	-0.3491	0.8292	0.8322	0.7112	0.3495
XLE	0.6778	-0.0974	-0.3648	-0.3491	1.0000	-0.2451	-0.1222	-0.0883	0.5139
XLF	0.1725	0.6987	0.7200	0.8292	-0.2451	1.0000	0.7721	0.6158	0.2774
XLI	0.2713	0.8513	0.8340	0.8322	-0.1222	0.7721	1.0000	0.7574	0.3649
XLK	0.4169	0.7870	0.7993	0.7112	-0.0883	0.6158	0.7574	1.0000	0.4142
XLU	0.4944	0.4060	0.1777	0.3495	0.5139	0.2774	0.3649	0.4142	1.0000

Table 7: Predicting Systematic Jump

This table reports test results of predicting systematic jumps using the implied total variance spread (Panel A) and using implied systematic and idiosyncratic variance spreads separately (Panel B). In both panels, we look at downward jumps (Columns (1)–(2)), upward jumps (Columns (3)–(4)) as well as downward and upward jumps mixed together (Columns (5)–(6)), where the jumps are defined based on monthly equity returns and denoted by dummy variables that equal 1 when a jump occurs and 0 otherwise. We perform estimation monthly and report time averages of our estimates in the table. Columns (1), (3) and (5) do not include control variables, and Columns (2), (4) and (6) control for the CAPM beta, firm size, book-to-market ratio, and lagged return. The Newey-West standard errors with 5 lags are reported in the parentheses below the corresponding estimates. Asterisks denote statistical significance at the 1% (***), 5% (**) and 10% (*) levels.

Panel A: Using Total Variance Spread

	Downward Jump		Upward Jump		Mixed Jump	
	(1)	(2)	(3)	(4)	(5)	(6)
$V^r_{n,n} - \Sigma^r_{n,n}$	-0.0076	-0.0016	0.0032	0.0035	-0.0074	0.0022
	(0.0054)	(0.0023)	(0.0035)	(0.0028)	(0.0068)	(0.0042)
β		0.0201		0.0208		0.0550
		(0.0056)***		(0.0053)***		(0.0115)***
$Ln(Size)$		-0.0028		-0.0052		-0.0096
		(0.0010)***		(0.0015)***		(0.0022)***
$Ln(B/M)$		0.0019		-0.0008		0.0006
		(0.0012)		(0.0010)		(0.0012)
$LagRet$		-0.0080		-0.0091		-0.0249
		(0.0047)*		(0.0055)		(0.0102)**

Panel B: Using Systematic and Idiosyncratic Variance Spreads

	Downward Jump		Upward Jump		Mixed Jump	
	(1)	(2)	(3)	(4)	(5)	(6)
$V^s_{n,n} - \Sigma^s_{n,n}$	0.0669	0.0305	0.1031	0.0739	0.2402	0.1491
	(0.0231)***	(0.0140)**	(0.0227)***	(0.0192)***	(0.0439)***	(0.0278)***
$V^\varepsilon_{n,n} - \Sigma^\varepsilon_{n,n}$	-0.0061	-0.0019	0.0002	0.0030	-0.0074	0.0022
	(0.0042)	(0.0024)	(0.0021)	(0.0025)	(0.0050)	(0.0038)
β		0.0187		0.0124		0.0402
		(0.0059)***		(0.0039)***		(0.0094)***
$Ln(Size)$		-0.0023		-0.0041		-0.0075
		(0.0009)**		(0.0013)***		(0.0019)***
$Ln(B/M)$		0.0018		-0.0005		0.0008
		(0.0011)		(0.0010)		(0.0012)
$LagRet$		-0.0081		-0.0094		-0.0252
		(0.0046)*		(0.0053)*		(0.0099)**

Table 8: Predicting Idiosyncratic Jump

This table reports test results of predicting idiosyncratic jumps using the implied total variance spread (Panel A) and using implied systematic and idiosyncratic variance spreads separately (Panel B). In both panels, we look at downward jumps (Columns (1)–(2)), upward jumps (Columns (3)–(4)) as well as downward and upward jumps mixed together (Columns (5)–(6)), where the jumps are defined based on monthly equity returns and denoted by dummy variables that equal 1 when a jump occurs and 0 otherwise. We perform estimation monthly and report time averages of our estimates in the table. Columns (1), (3) and (5) do not include control variables, and Columns (2), (4) and (6) control for the CAPM beta, firm size, book-to-market ratio, and lagged return. The Newey-West standard errors with 5 lags are reported in the parentheses below the corresponding estimates. Asterisks denote statistical significance at the 1% (***), 5% (**) and 10% (*) levels.

Panel A: Using Total Variance

	Downward Jump		Upward Jump		Mixed Jump	
	(1)	(2)	(3)	(4)	(5)	(6)
$V^r_{n,n} - \Sigma^r_{n,n}$	0.0218	0.0223	0.0310	0.0372	0.0558	0.0592
	(0.0097)**	(0.0079)***	(0.0097)***	(0.0098)***	(0.0159)***	(0.0145)***
β		0.0328		0.0296		0.0484
		(0.0044)***		(0.0041)***		(0.0052)***
$Ln(Size)$		-0.0055		-0.0095		-0.0133
		(0.0006)***		(0.0010)***		(0.0011)***
$Ln(B/M)$		-0.0011		-0.0007		-0.0013
		(0.0010)		(0.0007)		(0.0013)
$LagRet$		-0.0130		-0.0059		-0.0111
		(0.0045)***		(0.0056)		(0.0065)*

Panel B: Using Systematic and Idiosyncratic Variances

	Downward Jump		Upward Jump		Mixed Jump	
	(1)	(2)	(3)	(4)	(5)	(6)
$V^s_{n,n} - \Sigma^s_{n,n}$	0.3118	0.2222	0.3469	0.2393	0.5886	0.4168
	(0.0367)***	(0.0363)***	(0.0335)***	(0.0323)***	(0.0546)***	(0.0519)***
$V^\varepsilon_{n,n} - \Sigma^\varepsilon_{n,n}$	0.0133	0.0198	0.0190	0.0325	0.0339	0.0511
	(0.0075)*	(0.0075)***	(0.0076)**	(0.0092)***	(0.0122)***	(0.0137)***
β		0.0195		0.0147		0.0250
		(0.0036)***		(0.0034)***		(0.0044)***
$Ln(Size)$		-0.0035		-0.0073		-0.0097
		(0.0006)***		(0.0007)***		(0.0009)***
$Ln(B/M)$		-0.0009		-0.0010		-0.0014
		(0.0009)		(0.0007)		(0.0012)
$LagRet$		-0.0119		-0.0051		-0.0093
		(0.0039)***		(0.0052)		(0.0055)*

Table 9: Predicting Factor Downward Jumps by Implied Volatility Slope

This table reports summary statistics of the estimated implied volatility slopes of the FFC four factors. For each factor, we also report the coefficient from regressing a downward jump dummy on the estimated implied volatility slope available at the beginning of the month. The downward-jump dummy is set to one in months during which the factor return falls below the 5th percentile over the sample period and zero otherwise. For comparison purposes, we also report statistics for the market factor by replacing our estimated implied variance for the market by the option-implied variance of the S&P 500 index. Newey-West standard errors with 5 lags are displayed in parentheses below the corresponding estimates. Asterisks denote statistical significance at the 1% (***), 5% (**) and 10% (*) levels.

	S&P 500	$mktrf$	smb	hml	umd
Mean	-0.0473	-0.0109	0.0011	0.0015	0.0028
Median	-0.0436	-0.0102	-0.0022	0.0042	-0.0039
S.D.	0.0197	0.0239	0.0555	0.0273	0.0517
Min	-0.1409	-0.1800	-0.5557	-0.1791	-0.4025
Max	-0.0103	0.1252	0.3990	0.2025	0.3146
Predicting Downward Jump	-1.0356	-1.9021	-0.1485	-0.5947	-1.4906
(Using month-end IV slope)	(0.8763)	(0.7451)**	(0.1436)	(0.5564)	(0.6705)**
Predicting Downward Jump	-1.2052	-1.9804	-0.1496	-1.6001	-1.0406
(Using month-avg IV slope)	(1.0149)	(1.2010)*	(0.2049)	(0.7706)**	(0.6201)*