Recovering Implied Volatility*

Ohad Kadan[†], Fang Liu[‡], Xiaoxiao Tang[§]

September 2019

Abstract

We propose a methodology for estimating option-implied forward-looking variances and covariances of assets and portfolios, which may not possess actively-traded options. Our approach relies on the observation that any factor structure for stock returns induces a factor structure for return volatility. We implement the methodology empirically and show that our forward-looking moment estimates provide useful implications for the prediction of jumps and for portfolio choice.

Keywords: Implied volatility, options, factor model.

*We thank Phil Dybvig, Andrew Karolyi, Matthew Linn, Dmitry Orlov, Cisil Sarisoy, Georgios Skoulakis, Xiwei Tang, Liuren Wu, seminar participants at Cornell University, Southwestern University of Finance and Economics, the University of Rochester, Washington University in St. Louis, as well as participants at the 29th Annual Conference on Financial Economics & Accounting, the 4th Annual Young Scholars Finance Consortium, the 2018 Conference on Frontiers of Factor Investing, the 71th European Meeting of the Econometric Society, and the 2018 Northern Finance Association Conference for helpful comments and suggestions.

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

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

[§]Naveen Jindal School of Management, University of Texas at Dallas, xiaoxiao.tang@utdallas.edu.

1 Introduction

Options serve as important tools for investors to either bet on or hedge against future events and price movements. Given this, researchers have long argued that option prices reflect investors' expectations and thus contain forward-looking information on underlying asset prices (e.g., Ross (1976a) and Breeden and Litzenberger (1978)). The recent recovery literature ignited by Ross (2015) has focused on eliciting information from option prices on forward-looking expected returns (Martin (2017), Tang (2018), Jensen, Lando, and Pedersen (2019), Kadan and Tang (2019), and Martin and Wagner (2019)). A limitation, however, is that many key assets and portfolios do not possess actively-traded options. For example, commonly-used systematic factors such as size, value, and momentum do not have traded options. Small stocks and a variety of alternative assets also rarely have active option trading. This paper develops a new methodology to relax this limitation and recover option-implied forward-looking variances and covariances of assets and portfolios, even if they do not possess traded options. Empirically, we show that such recovered estimates provide useful insights on the prediction of jumps for systematic factors and individual stocks and on optimal portfolio choice.

Our methodology relies on the conventional linear factor representation of asset returns (Merton (1973b), Ross (1976b)). Specifically, we start with an arbitrary factor structure in which asset returns are linear combinations of the factors and idiosyncratic returns. Calculating the second moment of this relation induces a new factor structure in which each asset's variance is a linear combination of factor variances and covariances, and the new factor loadings are cross 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 new factor structure applies not only to historical variances, but also to forward-looking (implied) ones — typically obtained from option prices. As usual, these implied volatilities are calculated under the risk-neutral measure as opposed to the physical (true) measure. Our first key result (Theorem 1) establishes that, under mild assumptions, the volatility factor structure is preserved under the risk-neutral measure, whereby the factor loadings can still be calculated under the physical measure. The resulting hybrid factor structure consisting of factors calculated under the risk-neutral measure and factor loadings calculated under the physical measure facilitates the model's estimation.

To estimate and exploit this new 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 assets on factors over time to estimate factor loadings. In the second stage, we restrict attention to assets that do have traded options, and we cross-sectionally regress option-implied variances of these assets on cross products of the factor loadings estimated in the first stage. This results in an estimated risk-neutral implied variance-covariance matrix of the factors. Our second main result (Theorem 2) guarantees that the estimation procedure in this stage yields a unique solution. In the third stage, we use outputs from the previous two stages along with the underlying factor structure to calculate: (i) the implied covariance between any pair of individual assets — even those without options; and (ii) the implied systematic volatility of all assets — even those without options. This approach also yields a decomposition of implied variances into systematic and idiosyncratic components — for assets that possess traded options.

Empirically, we apply our estimation method to all CRSP common stocks over a sample period spanning January 1996 to December 2017, where the subset of stocks that possess traded options is taken to be constituents of the S&P 500 index during this period. While our approach can be applied to any linear factor structure for returns, for concreteness, we base our estimations on the Fama-French-Carhart (FFC) four-factor model (Fama and French (1993), Carhart (1997)). To demonstrate the usefulness of our framework, we present three applications in which our methodology is used to recover forward-looking moments for assets, factors, and portfolios that do not have actively-traded options.

In the first application, we consider the implied-volatility slope, the difference between at-the-money and out-of-the-money implied volatilities. Prior research has shown that the slope is useful in predicting downward jumps of individual stocks (Xing, Zhang, and Zhao (2010) and Yan (2011)). More generally, there is a voluminous literature showing that option prices reflect the jump risk associated with the underlying assets.[1] This literature is limited to assets with actively-traded options. We use our methodology to estimate the implied-volatility slope of the FFC factors and use the estimated slope to predict downward jumps of the factors. This is particularly useful for the size, value, and momentum factors given that there do not exist options actively traded on these factor portfolios.

In the second application, we consider the volatility spread, the difference between the forward-looking (implied) risk-neutral volatility and the historical volatility calculated under the physical measure. Bali and Hovakimian (2009) show that this spread has predictive power on future stock returns. This insight, however, can only be applied to stocks that possess actively-traded options. We first show that the effect of the spread is primarily attributed to its systematic component. We then use this observation and our methodology to extend Bali and Hovakimian's result to small stocks, which often do not have traded options. We find that a higher systematic volatility spread is associated with a higher likelihood of return jumps.

The third application illustrates the value of using the forward-looking variance-covariance matrix obtained from our modified Fama-MacBeth approach for optimal portfolio choice. Specifically, we consider a mean-variance investor whose investment opportunity set consists of sector exchange-traded funds (ETFs) and a risk-free asset. The investor chooses portfolio weights to maximize her utility and rebalances her portfolio on a monthly basis. We show that optimal portfolios constructed based on our estimated forward-looking variance-covariance matrix feature higher average returns, higher Sharpe ratios, more right skewness, and less heavy tails relative to those constructed based on historical moments.[2] We also show that our model outperforms the multivariate GARCH model (Engle and Sheppard (2001)), which is an

[1] See, for example, Cox and Ross (1976), Merton (1976a,b), 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).

[2] Our forward-looking moments are calculated under the risk-neutral measure, introducing a potential bias in the estimation (which should be under the physical measure). This bias, however, is dwarfed by the information content of our estimates, as evidenced by the superior performance they induce.

alternative method for estimating forward-looking variances and covariances.

Our paper is related to the growing recovery literature. Ross (2015) proposes that option prices can be used to uniquely recover the entire physical distribution of market returns. Borovička, Hansen, and Scheinkman (2016) criticize Ross's approach for restricting the dynamics of the stochastic discount factor in an unrealistic manner. Subsequent papers focus on recovering the first moment of asset returns from option prices based on milder assumptions (e.g., Martin (2017), Kadan and Manela (2018), Tang (2018), Jensen, Lando, and Pedersen (2019), Kadan and Tang (2019), and Martin and Wagner (2019)). Instead, our paper concentrates on the recovery of second moments of assets, factors, and portfolios that do not have traded options, building on a linear factor structure for the underlying asset returns.

Several prior papers have considered option-implied correlations, restricting attention to assets that do have traded options and making specific assumptions on the correlation structure. Skinzi and Refenes (2005), Driessen, Maenhout, and Vilkov (2009), and Buss, Schönleber, and Vilkov (2018) propose methods for estimating average option-implied correlation, assuming equal correlations between all asset pairs. Buss and Vilkov (2012) calculate option-implied betas by relying on the realized correlations between assets adjusted for a parametric premium reflected in options. Our modified Fama-MacBeth approach offers a method for recovering implied variances and covariances for a variety of assets (whether or not they have options) without resorting to an adjustment and not assuming equal correlations between asset pairs. Liu (2018) estimates an option-implied correlation between the disaster risk of individual stocks and that of the market index. Our approach yields pairwise correlation estimates for arbitrary pairs of factors and stocks without relying on strong parametric assumptions.

Our paper is related to the large body of research on the factor structure of asset returns originated by Merton (1973b) and Ross (1976b). Empirically, researchers have proposed a variety of factor models to explain variations in stock returns (e.g., Chen, Roll, and Ross (1986), Fama and French (1993, 2016), Carhart (1997), and Hou, Xue, and Zhang (2015)). Our paper helps connecting the factor structure of stock returns to a corresponding factor structure for option-implied volatility.

The paper is also related to the growing literature on the factor structure of option prices (e.g., Serban, Lehoczky, and Seppi (2008), Christoffersen, Fournier, and Jacobs (2018)).

Section 2 introduces our theoretical framework and estimation method. In Section 3, we conduct estimation based on the FFC four-factor model. Section 4 presents applications and Section 5 concludes. A technical proof is presented in the Appendix.

2 Theoretical Framework and Estimation Methodology

In this section we derive a factor structure for option-implied volatility and explain how it can be used to estimate forward-looking implied variances and covariances for a variety of assets and portfolios that may not have actively-traded options.

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[3]

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

where $\mathbf{r} = (r_1, ..., r_N)'$ is an $N \times 1$ vector of random asset returns, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)'$ is an $N \times 1$ vector of intercepts, $\mathbf{f} = (f_1, ..., f_K)'$ is a $K \times 1$ 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 an $N \times 1$ vector of idiosyncratic returns. We make the following two assumptions:

A1. $\mathbf{E}(\boldsymbol{\varepsilon}|\mathbf{f}) = \mathbf{0}$

A2. $\mathbf{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{f})$ is a diagonal matrix.

Taking variance on both sides of (1) and noting that A1 implies that $\boldsymbol{\varepsilon}$ and $\mathbf{f}$ are uncorrelated yield

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

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

where $\mathbf{\Sigma}^r, \mathbf{\Sigma}^f$, and $\mathbf{\Sigma}^\varepsilon$ are the variance-covariance matrices of the asset returns, factors, and idiosyncratic returns, respectively. Therefore, if the asset returns follow the linear factor structure (1), the variances and covariances of asset 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 is the intercept.

To illustrate this new factor structure, consider the diagonal of $\mathbf{\Sigma}^r$, which consists of the variances of the N assets, $\Sigma^r_{n,n}$. 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 cross products of the original factor loadings, and the new intercept is the variance of the idiosyncratic return.

Similarly, considering the off-diagonal terms in (2), for any two assets m and n, the covariance between the returns, $\Sigma^r_{m,n}$, 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. These off-diagonal factor structures have no intercept owing to A1 and A2.

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 (Black and Scholes (1973), Merton (1973a)), and then applying this model to derive implied volatilities from options. To this end, we divide the universe of N assets into two subsets. Assets $1, ..., L$ are assumed to have traded options ($L < N$), whereas assets $L + 1, ..., N$ are assumed not to have traded options.

A complication in using the option-implied volatilities of assets $1, ..., L$ is that these implied volatilities are calculated under the risk-neutral measure, whereas assumptions A1 and A2 along with the factor structures have been stated under the physical (true) measure of the return distribution. To facilitate the use of these implied volatilities, we will now show that under our rather standard assumptions, the same factor structure (2) also holds under the risk-neutral measure. Intuitively, this follows because in a world in which all returns are governed by a factor structure, any stochastic discount factor (SDF) must be a function of the factors, and thereby orthogonal to the idiosyncratic returns. Moreover, the SDF is (up to a constant) a Radon-Nikodym derivative, facilitating a change of measure between the physical and risk-neutral measures. Thus, assumptions A1 and A2, which are stated under the physical measure, imply that, under the risk-neutral measure, $\boldsymbol{\varepsilon}$ and $\mathbf{f}$ must be uncorrelated, and the elements of $\boldsymbol{\varepsilon}$ must also be uncorrelated.

To see this point formally, let $\mathbf{V}^r$, $\mathbf{V}^f$, and $\mathbf{V}^\varepsilon$ denote the variance-covariance matrices for the asset returns, factors, and idiosyncratic returns, respectively, under the risk-neutral measure (of dimensions $N \times N$, $K \times K$, and $N \times N$). Let $\mathbf{E}^*(\cdot)$ denote the expectation operator under the risk-neutral measure. And, let $m(\mathbf{f}) > 0$ denote a stochastic discount factor (a function of the return factors). Then,

$$\begin{aligned}
\mathbf{E}^*(\boldsymbol{\varepsilon}) &= \mathbf{E}\left(\frac{m(\mathbf{f})}{\mathbf{E}(m(\mathbf{f}))}\boldsymbol{\varepsilon}\right) \\
&= \frac{\mathbf{E}(\mathbf{E}(m(\mathbf{f})\boldsymbol{\varepsilon}|\mathbf{f}))}{\mathbf{E}(m(\mathbf{f}))} \\
&= \frac{\mathbf{E}(m(\mathbf{f})\mathbf{E}(\boldsymbol{\varepsilon}|\mathbf{f}))}{\mathbf{E}(m(\mathbf{f}))} \\
&= \mathbf{0},
\end{aligned}$$

where the first equality follows since $\frac{m(\mathbf{f})}{\mathbf{E}(m(\mathbf{f}))}$ is a Radon-Nikodym derivative, the second by iterated expectations, the third from the fact that $m(\mathbf{f})$ is measurable with respect to $\mathbf{f}$, and the last by A1. We conclude that all idiosyncratic returns have zero mean under the risk-neutral measure. An analogous argument shows that the covariance between the factors and the idiosyncratic returns under the risk-neutral measure is also zero: $\mathbf{E}^*(\mathbf{f}\boldsymbol{\varepsilon}') = \mathbf{0}$. And, a similar application of A2 implies that

$\mathbf{V}^{\varepsilon} = \mathbf{E}^{*}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')$ is diagonal.

We have thus established the following three risk-neutral implications of assumptions A1 and A2:

A0*. $\mathbf{E}^{*}(\boldsymbol{\varepsilon}) = \mathbf{0}$

A1*. $\mathbf{E}^{*}(\mathbf{f}\boldsymbol{\varepsilon}') = \mathbf{0}$

A2*. $\mathbf{E}^{*}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')$ is a diagonal matrix.

Now, taking variance on both sides of (1) under the risk-neutral measure and applying A0*, A1*, and A2*, we obtain the following result establishing a risk-neutral analogue to the factor structure in (2).

Theorem 1 *The following linear factor structure holds under the risk-neutral measure*

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

where $\mathbf{V}^{\varepsilon}$ is a diagonal matrix.

Equation (5) is the fundamental factor relation exploited in this paper. Note that this equation is a hybrid in the sense that while all variances and covariances are calculated under the risk-neutral measure, $\boldsymbol{\beta}$, the factor loadings in (5), can still be calculated under the physical measure, and are identical to the factor loadings used in the original return factor structure (1). The hybrid nature of this relation is key to our estimation as it facilitates the use of implied-volatility data from options in combination with factor loadings calculated from historical returns.

There are several challenges in estimating and applying (5). First, only a subset of $L < N$ assets have traded options. Moreover, there is no easy way to estimate implied covariances between assets. Thus, for the the implied variance-covariance matrix $\mathbf{V}^{r}$, only a subset of the diagonal elements are available from data, and none of the off-diagonal elements is observed. Similarly, $\mathbf{V}^{f}$ and $\mathbf{V}^{\varepsilon}$, the implied variance-covariance matrices of the factors and idiosyncratic returns, are not readily available. Indeed, most conventional factors do not have traded options, and certainly idiosyncratic returns have no options. 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 cross-sectionally to estimate risk premia. We will show that a modification of the Fama-MacBeth approach consisting of three stages can be used to estimate the ingredients in (5).

The data required are time series of the K factors, a panel of the returns of the N assets of which L assets have traded options, and a panel of the implied volatilities associated with these L assets (derived from options). To facilitate the estimation, we require that $L >> K(K+1)/2$, namely the number of assets with traded options is at least an order of magnitude higher than K^2.

The first stage is identical to the original Fama-MacBeth approach. We use a time-series regression in which we regress the returns of the N assets on the factors over the period $t = \tau_0, ..., \tau$ to obtain estimates of the factor loadings as of time τ, denoted by the $N \times K$ matrix $\hat{\boldsymbol{\beta}}_\tau$.

In the second stage, we rely on Theorem 1 and estimate cross-sectional regressions of the implied (risk-neutral) variances of the L assets with option trading as of time τ, $V^r_{n,n,\tau}$ $(n = 1, ..., L)$, on cross products of the estimated factor loadings from the first stage. Specifically, the second stage 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 symmetric and positive semidefinite (i.e., a variance-covariance matrix).[4]

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 (the implied factor variance if $k = l$). The intercept in

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

this regression $\hat{\lambda}_\tau$ is an estimate of the cross-sectional average of implied idiosyncratic variances as of time τ for the L assets with options. 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 the idiosyncratic variance.

Let $\hat{\mathbf{B}}_\tau$ be a matrix consisting of the regressors in (6). That is, $\hat{\mathbf{B}}_\tau$ is the matrix of all cross 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 $L \times \left(\frac{K(K+1)}{2} + 1\right)$ matrix with a generic form (omitting time subscripts inside the matrix for brevity)

$$\hat{\mathbf{B}}_\tau = \begin{pmatrix} \hat{\beta}_{1,1}^2 & 2\hat{\beta}_{1,1}\hat{\beta}_{1,2} & \cdots & 2\hat{\beta}_{1,1}\hat{\beta}_{1,K} & \hat{\beta}_{1,2}^2 & 2\hat{\beta}_{1,2}\hat{\beta}_{1,3} & \cdots & 2\hat{\beta}_{1,2}\hat{\beta}_{1,K} & \cdots & \hat{\beta}_{1,K}^2 & 1 \\ \hat{\beta}_{2,1}^2 & 2\hat{\beta}_{2,1}\hat{\beta}_{2,2} & \cdots & 2\hat{\beta}_{2,1}\hat{\beta}_{2,K} & \hat{\beta}_{2,2}^2 & 2\hat{\beta}_{2,2}\hat{\beta}_{2,3} & \cdots & 2\hat{\beta}_{2,2}\hat{\beta}_{2,K} & \cdots & \hat{\beta}_{2,K}^2 & 1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ \hat{\beta}_{n,1}^2 & 2\hat{\beta}_{n,1}\hat{\beta}_{n,2} & \cdots & 2\hat{\beta}_{n,1}\hat{\beta}_{n,K} & \hat{\beta}_{n,2}^2 & 2\hat{\beta}_{n,2}\hat{\beta}_{n,3} & \cdots & 2\hat{\beta}_{n,2}\hat{\beta}_{n,K} & \cdots & \hat{\beta}_{n,K}^2 & 1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ \hat{\beta}_{L,1}^2 & 2\hat{\beta}_{L,1}\hat{\beta}_{L,2} & \cdots & 2\hat{\beta}_{L,1}\hat{\beta}_{L,K} & \hat{\beta}_{L,2}^2 & 2\hat{\beta}_{L,2}\hat{\beta}_{L,3} & \cdots & 2\hat{\beta}_{L,2}\hat{\beta}_{L,K} & \cdots & \hat{\beta}_{L,K}^2 & 1 \end{pmatrix}. \tag{7}$$

The next theorem facilitates our second-stage estimation by establishing that the constrained least-squares regression in (6) has a unique solution. The theorem requires that the matrix $\hat{\mathbf{B}}_\tau'\hat{\mathbf{B}}_\tau$ be of full rank, which is equivalent to $\hat{\mathbf{B}}_\tau$ being of full column rank. This requirement is natural as it essentially says that cross products of factor loadings are not linearly related, which intuitively corresponds to the factors being sufficiently different from each other.

Theorem 2 *Suppose that $\hat{\mathbf{B}}_\tau'\hat{\mathbf{B}}_\tau$ is of full rank. Then, the problem of finding $\hat{\mathbf{V}}_\tau^f$ that minimizes the sum of squared residuals associated with (6) under the constraint that $\hat{\mathbf{V}}_\tau^f$ is symmetric and positive semidefinite has a unique solution.*

Proof: See the Appendix.

In the third stage, we use the estimated factor loadings $\hat{\boldsymbol{\beta}}_\tau$ and implied factor variance-covariance matrix $\hat{\mathbf{V}}_\tau^f$ to estimate the implied covariances between all assets

— including those with no options — as the off-diagonal elements of $\mathbf{V}_\tau^r$, by applying the relation in (5), i.e.,

$$\hat{V}_{m,n,\tau}^r = \sum_{k=1}^{K} \sum_{l=1}^{K} \hat{V}_{k,l,\tau}^f \hat{\beta}_{m,k,\tau} \hat{\beta}_{n,l,\tau} \text{ for } m,n = 1,...,N,\ m \neq n.^5 \tag{8}$$

The three stages are then repeated for different times τ, using a rolling window approach, yielding a time series of the implied variance-covariance matrix for the factors $\hat{\mathbf{V}}_\tau^f$. The procedure also yields a time series of the implied covariances between all N assets as well as the cross-sectional average idiosyncratic variances $\hat{\lambda}_\tau$ for the first L assets. Finally, for each asset $n = 1,...,N$ (including assets with no options) we can calculate the implied systematic variance at time τ, $\hat{V}_{n,n,\tau}^s$, i.e., the part of the implied variance driven by systematic factors. Specifically, based on (6) we have

$$\hat{V}_{n,n,\tau}^s = \sum_{k=1}^{K} \sum_{l=1}^{K} \hat{V}_{k,l,\tau}^f \hat{\beta}_{n,k,\tau} \hat{\beta}_{n,l,\tau}. \tag{9}$$

For assets with options ($n = 1,...,L$) we additionally calculate the implied idiosyncratic variance as the total implied variance less the estimated systematic part,

$$\hat{V}_{n,n,\tau}^\varepsilon = V_{n,n,\tau}^r - \sum_{k=1}^{K} \sum_{l=1}^{K} \hat{V}_{k,l,\tau}^f \hat{\beta}_{n,k,\tau} \hat{\beta}_{n,l,\tau}. \tag{10}$$

That is, for assets with options we can decompose the implied volatility to its systematic and idiosyncratic components.

To summarize, our modified Fama-MacBeth method exploits cross-sectional variations in the factor loadings and in the implied volatility of the assets with options to recover the implied variance-covariance matrix of the factors themselves — whether or not they have traded options — along with the implied covariances and systematic variances of all assets — including those without options.

[5] The sub-matrix of $\hat{\mathbf{V}}_\tau^r$ restricted to the first L assets with options is not guaranteed to be positive semidefinite, because the diagonal elements of $\hat{\mathbf{V}}_\tau^r$ are directly observed from option prices instead of being calculated from the factor structure. When this sub-matrix is not positive semidefinite, we estimate its nearest symmetric positive semidefinite matrix following Higham (1988).

With these results at hand, we can also calculate the implied covariances and systematic variances of any portfolio of assets. For diversified portfolios with a large number of assets, the implied idiosyncratic volatility should be small (owing to A2*), and the implied systematic variance will effectively capture the total implied variance.

2.3 Discussion

Before proceeding to the empirical analysis, we pause to discuss some of the model's features and assumptions. To gain better intuition, consider the CRSP universe consisting of about 14000 common stocks (using our model's notation, $N \approx 14000$). Out of these, roughly about 1000 stocks have active option trading and implied volatility estimates ($L \approx 1000$). The cross-sectional regressions in Stage 2 of our methodology are restricted to these L assets and thus will each have about 1000 observations. The regressors in Stage 2 are the cross products of the factor betas calculated in Stage 1. Thus, assuming the Fama-French-Carhart (Fama and French (1993), Carhart (1997), hereafter FFC) four-factor model ($K = 4$), we will have 1000 observations to estimate $\frac{K(K+1)}{2} + 1 = 11$ parameters for the variances and covariances of the factors and an intercept. This example demonstrates that the assumption $L >> \frac{K(K+1)}{2}$ is essential to ensure that the cross-sectional regressions in Stage 2 have a sufficient number of degrees of freedom.

A feature of Stage 3 is that it allows us to estimate a large number of latent parameters using a relatively small dataset of implied variances. The key point here is that Stage 3 does not actually estimate any new parameters but rather applies the estimation in Stages 1 and 2 and the factor structure to conduct direct calculations. Indeed, the linear factor structure (1) implies that (8) and (9) must hold for all N assets, regardless of whether they have options. We use these equations and the parameters estimated in Stages 1 and 2 to calculate the covariances of all the 14000 CRSP assets plus their systematic variances, amounting to roughly 98 million latent parameters. This, however, should not be surprising, as the essence of linear factor models is to allow the estimation of a large number of asset expected returns from a small number of factor premia along with factor loadings. The innovation here is in looking at the implied second (rather than first) moments of factors and individual assets with no options, which are imputed from the implied volatilities of a subset of assets with options.

Of course, for these estimations to be of any practical value, it must be that the factor structure is useful in capturing variations in stock returns. Accordingly, our two assumptions, A1 and A2, are effectively assumptions on the quality of the factor structure in capturing stock return variation. In essence, these assumptions require that there are no omitted factors in the model in the sense that the idiosyncratic return is uncorrelated with the factors (A1) and uncorrelated cross-sectionally (A2). Similar assumptions have traditionally been used to justify the arbitrage pricing theory (Ross (1976b)) as they imply that idiosyncratic risk vanishes in large portfolios.

In reality, it is hard to expect any factor model to precisely satisfy A1 and A2, and the failure of these assumptions would lead to less accurate estimates from our model. Nevertheless, the empirical analysis below shows that applying our methodology to a standard factor model yields a satisfactory representation of stock returns as evidenced by the useful forward-looking insights we generate.

3 Empirical Estimation

In this section we implement the framework discussed above using CRSP stocks as assets. While our approach would work with any factor structure for stock returns, for concreteness we rely on the Fama-French-Carhart (FFC) four-factor model.

3.1 Data

Our data consist of all CRSP stocks with share codes 10 and 11 during January 1996 to December 2017 (13904 stocks). As the subset of stocks with traded options we use all S&P 500 constituent stocks during the sample period with options traded on them (1038 stocks). Our option data are drawn from OptionMetrics. We obtain the implied volatility from the implied-volatility surface file in OptionMetrics, which contains the Black-Scholes implied volatility for European options and the Cox-Ross-Rubinstein implied volatility (Cox, Ross and Rubinstein (1979)) for American options. We take the average implied volatility of the at-the-money (ATM) 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 (2018)). We use the square of this average implied volatility as the option-implied variance of stock returns. We obtain the FFC factor data from WRDS.

We begin by exploring the implied volatility of the S&P 500 constituents between 1996 and 2017. 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&P 500 index. The average of option-implied variances of individual stocks fluctuates considerably during the sample period. It has experienced two periods of substantial increases, 1998–2002 and 2008–2009, respectively. The former is relatively mild, corresponding to the Asian financial crisis and the subsequent internet bubble, whereas the latter is more dramatic, corresponding to the 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.

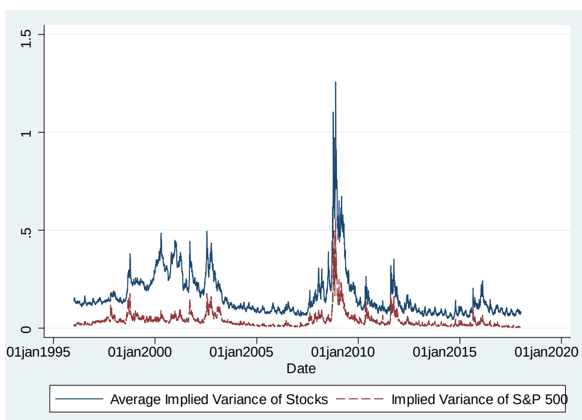

Figure 1: Option-Implied Variances of S&P 500 Stocks and S&P 500 Index

3.2 Modified Fama-MacBeth Estimation

Our methodology translates a given factor model for stock returns into a factor model for implied variances and covariances of returns. We apply this approach to the FFC model in which the factors are the market excess return ($mktrf$), size (smb), value (hml), and momentum (umd).[6] The FFC model transforms into a ten-factor model for implied variances.

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 daily excess returns of stocks on daily factors over the preceding 12-month window. Table 1 reports cross-sectional summary statistics for the time averages of the factor loadings for all CRSP common stocks (Panel A) and S&P 500 constituents (Panel B). We see that the average market beta is 0.82 for CRSP stocks and is close to 1 for S&P 500 stocks. As expected, other betas can be either positive or negative. For example, average momentum betas are negative for both CRSP and S&P 500 stocks.

3.2.2 Second-Stage Estimation

For the second stage we estimate daily constrained cross-sectional regressions as in (6) for the subset of S&P 500 stocks. Our dependent variable is the daily implied variance, and the independent variables are cross 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.

Table 2 reports summary statistics of the estimated implied factor variances and covariances. The time averages of the implied variances of the market, size, value, and momentum factors are 0.0402, 0.0476, 0.0219, and 0.0667, respectively. For all factors, the mean of the estimated implied variance is higher than the median. In addition, the minimum values of all implied variances are positive, which is guaranteed by our positive semidefiniteness regression constraint. The implied covariances and correlations between two different factors, however, take different signs.

[6] We also perform estimation based on the CAPM one-factor model and the four-factor model of Hou, Xue and Zhang (2015). These results are available upon request.

3.2.3 Third-Stage Estimation

In the first stage, we have estimated for each day KN factor loadings, whereas in the second stage, we have estimated for each day $K(K+1)/2$ factor variances and covariances, for a total of $K(2N+K+1)/2$ estimated parameters. In the third stage, we use these estimates and the factor structure to calculate for each day (i) the $N(N-1)/2$ implied covariances between the returns of all possible pairs of individual stocks based on (8); (ii) the N implied systematic variances of all stocks based on (9); and (iii) the L implied idiosyncratic variances for all S&P 500 stocks based on (10). Thus, our approach recovers daily implied covariances and systematic variances for all stocks, whether they have active option trading or not.

To illustrate the results of the third stage, we have first calculated these implied covariances and correlations for the Dow Jones Industrial Average (DJIA) stocks (using index composition as of September 2018).[7] The estimated pairwise implied correlation of the DJIA stocks has a mean value of 0.5871, a median of 0.7573, and a standard deviation of 0.4323. About 88% of the implied correlations are positive.

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. Over our sample period, the average implied systematic and idiosyncratic variances across all DJIA stocks are 0.0519 and 0.0451, and the average ratios of the implied systematic and idiosyncratic variances to the implied total variance are 53.5% and 46.5%, respectively.

As a second illustration, we calculate the implied covariances and systematic variances for 30 stocks that are, in terms of their market capitalization as of September 2018, immediately below the S&P 500 threshold, not resorting to whether these stocks have actively-traded options or not. The mean estimated pairwise implied covariance of these stocks is 0.0341, and only 51% of the covariances are positive. In addition, the average implied systematic variance of these stocks is 0.0739.

[7] There are 30 DJIA constituent stocks. However, one constituent, DowDuPont Inc., was formed after the merger of Dow Chemical and DuPont in August 2017, and hence does not have available data. As a result, we only compute the summary statistics for 29 constituent stocks (406 pairs).

3.2.4 Comparison to Realized Moments

Before proceeding to apply our estimates, we explore whether the implied variances and covariances estimated using our modified Fama-MacBeth approach correlate with the corresponding realized moments. Note that our implied moments are forward looking, and thereby we should not expect them to be perfectly correlated with realized 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.[8] Figure 2 presents the factor variances. The implied and realized market variances match closely. In particular, the estimated implied variance captures the spikes during 1998–2002 and 2008–2009. The implied variances of the other three factors also seem to capture the general trends in the corresponding realized variances despite larger disparities.

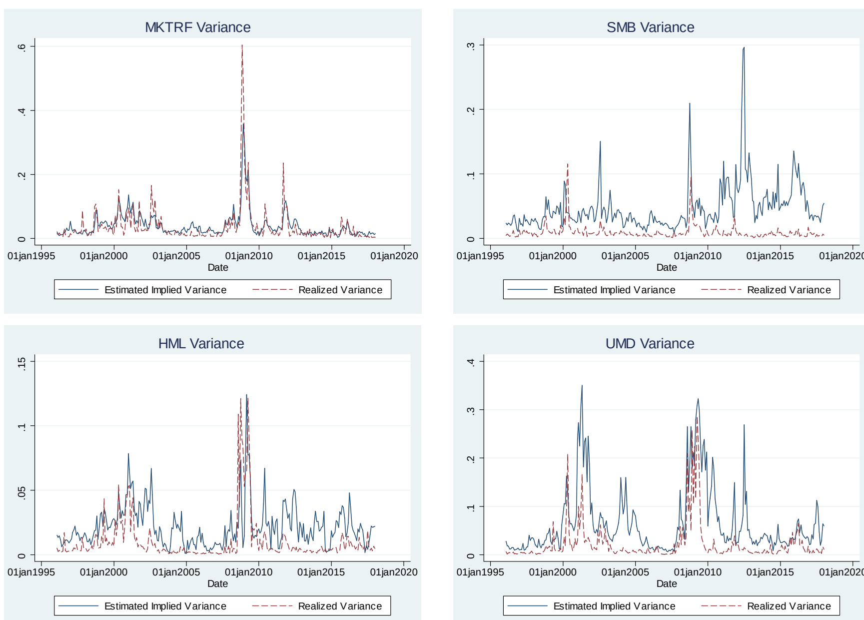

Figure 2: Implied vs. Realized Factor Variances

[8] 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.

Figure 3 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.

Figure 3: Implied vs. Realized Factor Covariances

In Figure 4 we plot the estimated implied covariance between two individual stocks, Microsoft and P&G, over time. This is obtained from the third stage of our modified Fama-MacBeth procedure. We have chosen these two stocks arbitrarily, and other pairs behave quite similarly. As before, it can be seen that the estimated implied covariance matches the trend of the realized covariance in general, despite some discrepancies.

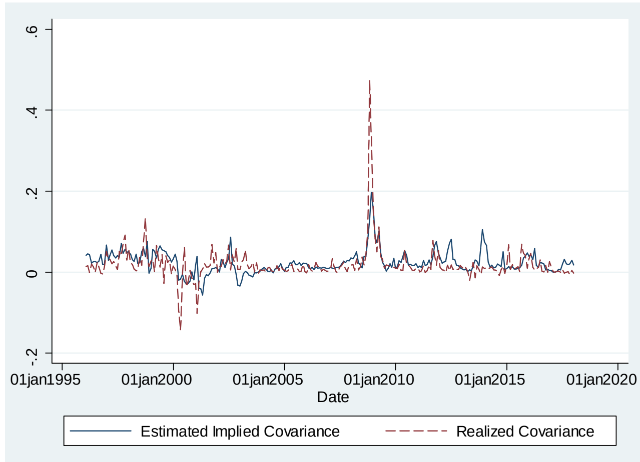

Figure 4: Implied vs. Realized Covariance Between Microsoft and P&G

To sum up, while our estimated implied moments and the corresponding realized moments in general feature similar time trends, we can also observe apparent deviations between the two. These discrepancies reflect that our estimated implied moments potentially contain differential information on the riskiness of the associated factors and stocks.

4 Applications

We next provide three applications for our framework. First, we apply the model to derive an implied-volatility slope (smile) for factors — even those that do not have options written on them, and show that this slope is often indicative of future downward jumps. Next, we construct a synthetic systematic volatility spread to predict returns, systematic jumps, and idiosyncratic jumps for individual stocks, which may not have actively-traded options. Finally, we show how the model can be used to construct an optimal investment portfolio for a mean-variance investor based on forward-looking information. A common theme of these applications is that the recovered implied volatility and correlation contain useful forward-looking information not captured by historical data.

4.1 Factor Implied-Volatility Slope and Jump Risk

While the Black-Scholes 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 (smile) for both index options and individual stock options (e.g., Bates (1991, 2003), Pan (2002), Bollen, and Whaley (2004), and Gârleanu, Pedersen, and Poteshman (2009)). A conventional explanation for this negative slope is that it is associated with downside jump risk (e.g., Bates (1991), Pan (2002), Xing, Zhang, and Zhao (2010), and Yan (2011)). In this section we use our methodology to estimate an implied-volatility slope for factors, even if they do not possess traded options, and we show that this slope is associated with future downward jumps.

Following Xing, Zhang, and Zhao (2010), we define the implied-volatility slope (S) as the difference between implied volatilities of at-the-money call ($\sqrt{V^{ATMC}}$) and out-of-the-money put ($\sqrt{V^{OTMP}}$) options with 30 days to maturity,

$$S = \sqrt{V^{ATMC}} - \sqrt{V^{OTMP}}.$$

Intuitively, when the market expects prices to drop, put (call) options become more (less) expensive, resulting in a higher (lower) implied volatility for puts (calls), and thus a more negative slope. We estimate the implied-volatility slope for each of the FFC factors by applying our modified Fama-MacBeth approach at different moneyness levels. We continue to use January 1996 to December 2017 as our sample period and all S&P 500 constituents within the period as the L assets with traded options. We begin our analysis by considering the slope for the market factor, for which the estimation is more straightforward than for the other factors, and for which we have existing options allowing for a direct comparison. We then turn to studying the slope for the remaining three factors for which traded options are not in existence at this time.

4.1.1 Implied-Volatility Slope of the Market Factor

We estimate V^{ATMC} for the market portfolio using our modified Fama-MacBeth approach applied to ATM call options of S&P 500 constituents. Similarly, we estimate

V^{OTMP} using OTM put options on S&P 500 constituents. Following Xing, Zhang, and Zhao (2010), we choose a delta value of -0.2 for OTM put options, corresponding to a strike price that is roughly 5% lower than the current 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.

Column (2) of Table 3 reports summary statistics for the implied-volatility slope of the market factor. The mean (median) is -0.0075 (-0.0084), consistent with the negative slope documented in the literature. The standard deviation of our estimates is 0.0237.

We next test whether our estimated market implied-volatility slope can predict future downward jumps of the market factor. We define a downward jump as a monthly return below the 5th percentile of the historical return distribution over our sample period. To this end, we define a downward jump dummy DJ_τ that equals one if the market excess return over month τ is lower than -7.59% and zero otherwise. We regress this dummy on the market implied-volatility slope estimated as of the end of the previous month over time, i.e.,

$$DJ_\tau = \delta_0 + \delta_1 S_{\tau-1} + \phi_\tau. \tag{11}$$

As reported in Column (2) of Table 3, the regression coefficient is negative and statistically significant at the 5% level according to Newey-West standard errors with 5 lags.[9] 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 estimated slope is associated with roughly a 5.92% 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 readily available from S&P 500 index options. As reported in Column

[9]Following Stock and Watson (2011, p. 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.

(1) of Table 3, the mean, median, and standard deviation of the slope are -0.0460, -0.0421, and 0.0227, respectively. Surprisingly, when we regress the market downward jump dummy on the lagged S&P 500 implied-volatility slope, the coefficient is very small in magnitude and statistically insignificant. This suggests that our estimated market implied-volatility slope performs better than the S&P 500 implied-volatility slope in predicting future market downward jumps. Our findings confirm that options written on individual stocks can be aggregated to generate market-wide information, which is not necessarily reflected in options written on an index comprised of these individual stocks. This may be the case because, unlike with equities, there is no mechanical aggregation of individual stock option prices into the price of an option written on the market as a whole.

4.1.2 Implied-Volatility Slope of the Size, Value, and Momentum Factors

Similar to the case of the market factor, the ATM call implied variance, V^{ATMC}, of the size, value, and momentum factors can be estimated by applying the modified Fama-MacBeth procedure to ATM call implied volatilities of S&P 500 constituents. The estimation of the OTM put implied variance, V^{OTMP}, is trickier, and we discuss it below. We use the momentum factor to illustrate the main idea.

The momentum factor is a long-short portfolio with long positions in recent winner firms and short positions in recent loser firms. To estimate the OTM put implied volatility of the momentum factor, we are no longer able to use OTM put implied volatilities of S&P 500 stocks as we did for the market factor, because winner and loser firms typically have inverse correlation patterns with the momentum factor. This is in contrast to the market factor, in which case 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 winner firms at a delta of -0.2 (corresponding to a monthly return of around -5%) and OTM call implied volatilities of all losers firms at a delta of 0.2 (corresponding to a monthly return of around 5%). The idea is that if all winner firms lose 5% and all loser firms gain 5% at the same time, then overall the momentum portfolio would experience a downward jump. To apply this procedure we need to differentiate between "winners" and "losers". We do this based on the stock return correlation with the momentum factor. Namely, we define a firm as a "winner" if its momentum beta is positive and define it as a "loser" otherwise.

Columns (3)–(5) of Table 3 report summary statistics for the implied-volatility slope of the size, value, and momentum factors. Also reported in the table are the estimated slope coefficients $\hat{\delta}_1$ from regressing downward jump dummies of the three factors on their corresponding estimated implied-volatility slopes over time (see (11)). As before, the downward jump dummy is equal to one if the monthly return of a factor drops below the 5th percentile of the historical return distribution over our sample period. Specifically, the downward jump thresholds for the size, value, and moment factors are -4.35%, -4.18%, and -7.97%, respectively. The resulting regression coefficients for both the value and momentum factors are negative and strongly significant, indicating that our estimated implied-volatility slope predicts future downward jumps of these two factors. In particular, a one-standard-deviation increase in the value (momentum) implied-volatility slope is associated with roughly a 2.93% (5.75%) drop in the probability of a future downward jump in the value (momentum) factor. The coefficient for the size factor is also negative, but not statistically significant.[10]

4.2 Volatility Spread and Jump Risk

Bakshi and Kapadia (2003a,b) show that both the market index and individual stocks exhibit a negative variance risk premium. That is, higher exposure to volatility risk is associated with lower returns. This, in turn, implies that the stochastic discount factor (SDF) takes larger values when variance is high. Recall that the risk-neutral measure assigns higher probabilities to states in which the SDF is high compared to the physical measure. Thus, the negative variance risk premium gives rise to a positive spread between option-implied risk-neutral volatility and physical return volatility, known as the volatility spread. Bali and Hovakimian (2009) find that the volatility spread predicts stock returns in the cross section.

A limitation of this literature is that the volatility spread can only be calculated

[10] The reason that we do not have significant result for the size (small-minus-big) factor might be that our estimation is based on big stocks only (the S&P 500 constituents).

for stocks that have active option trading. Therefore, one cannot use this spread to obtain insights on future returns of smaller stocks, which often do not possess actively-traded options. For example, during December 2017 about 76% of all option trading volume is concentrated in options written on S&P 500 stocks.

Our methodology can be used to address this limitation. For any asset n (whether it has active option trading or not), we can calculate a synthetic systematic volatility spread $\sqrt{V^s_{n,n}} - \sqrt{\Sigma^s_{n,n}}$, whereby $V^s_{n,n}$ is given by (9) and $\Sigma^s_{n,n}$ is the asset's historical systematic variance. For assets with active option trading we can also calculate an idiosyncratic volatility spread $\sqrt{V^\varepsilon_{n,n}} - \sqrt{\Sigma^\varepsilon_{n,n}}$, where $V^\varepsilon_{n,n}$ is given by (10), and $\Sigma^\varepsilon_{n,n}$ is the historical idiosyncratic variance associated with the factor model.

4.2.1 Motivation: Volatility Spread and Stock Returns for S&P 500 Stocks

To motivate our analysis, we will first show that the lion's share of the effect of the spread on future returns comes from the systematic component. For this purpose we will apply our methodology to S&P 500 stocks, which typically do possess actively-traded options. For each month τ, we run the following cross-sectional regression

$$r_{n,\tau} = \delta^r_{0,\tau} + \delta^r_{s,\tau}\left(\sqrt{V^s_{n,n,\tau-1}} - \sqrt{\Sigma^s_{n,n,\tau-1}}\right) + \delta^r_{\varepsilon,\tau}\left(\sqrt{V^\varepsilon_{n,n,\tau-1}} - \sqrt{\Sigma^\varepsilon_{n,n,\tau-1}}\right) + controls_{n,\tau-1} + \phi^r_{n,\tau}, \tag{12}$$

where the control variables include the CAPM beta, firm size, book-to-market ratio, and lagged stock return.[11] We then conduct t-tests with Newey-West standard errors on the estimated $\hat{\delta}^r_{s,\tau}$ and $\hat{\delta}^r_{\varepsilon,\tau}$ over time to see if they are significantly different from zero. Column (1) in Panel A of Table 4 reports the result. While both $\hat{\delta}^r_{s,\tau}$ and $\hat{\delta}^r_{\varepsilon,\tau}$ are positive and statistically significant, the economic magnitude of $\hat{\delta}^r_{s,\tau}$ is much larger than that of $\hat{\delta}^r_{\varepsilon,\tau}$. Increasing the systematic and idiosyncratic volatility spreads by one standard deviation raises the monthly stock return by 0.34% and 0.12%,

[11] We estimate CAPM beta by regressing daily stock excess returns on daily market excess returns over a rolling window of 12 months prior to the month of interest. We estimate firm size as the log of the month-end market capitalization, and book-to-market is the log of the ratio of the book value of equity to market value of equity at the end of each month. Finally, the lagged stock return is obtained from the 12 months preceding the month of interest.

respectively. Thus, the systematic volatility spread has an effect that is almost three times larger than that of the idiosyncratic spread in the prediction of future stock returns.

Similarly, we estimate the effect of the spreads on jumps in stock returns using

$$J_{n,\tau} = \delta^J_{0,\tau} + \delta^J_{s,\tau}\left(\sqrt{V^s_{n,n,\tau-1}} - \sqrt{\Sigma^s_{n,n,\tau-1}}\right) + \delta^J_{\varepsilon,\tau}\left(\sqrt{V^\varepsilon_{n,n,\tau-1}} - \sqrt{\Sigma^\varepsilon_{n,n,\tau-1}}\right) + controls_{n,\tau-1} + \phi^J_{n,\tau}, \tag{13}$$

where $J_{n,\tau}$ is a dummy variable capturing jumps (similar to Section 4.1). We consider both upward and downward jumps. For the FFC factors as well as 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 thresholds for the factors are the same as in Section 4.1, and the upward jump thresholds for the market, size, value, and moment factors are 7.04%, 4.98%, 5.52%, and 7.26%, respectively. The downward and upward jump thresholds for all S&P 500 stocks pooled together are -15.52% and 18.42%, 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 test (13) for both systematic and idiosyncratic jumps and for downward, upward, and mixed jumps.

Columns (2)-(4) in Panel A of Table 4 show that the systematic volatility spread positively predicts all types of systematic jumps in the cross section, whereas the idiosyncratic volatility spread has significant predictive power only for upward and mixed systematic jumps. For all three jump types, the economic effect of the systematic volatility spread is much larger than that of the idiosyncratic component. For example, increasing the systematic volatility spread by one standard deviation is associated with an increase in the probability of future systematic jumps (downward and upward pooled together) by 1.57%, as opposed to 0.17% for the idiosyncratic spread. Columns (5)-(7) show similar dominance of the systematic spread for idiosyncratic jumps.

In conclusion, the effect of the volatility spread on stock returns and on return jumps is largely captured by the systematic component of the spread. We next use this result to apply our methodology to smaller stocks, which rarely have actively-traded options.

4.2.2 Volatility Spread and Jump Risk for Non S&P 500 Stocks

As noted above, our methodology can be used to calculate the systematic spread, $\sqrt{V^s_{n,n}} - \sqrt{\Sigma^s_{n,n}}$, for all stocks, even if they do not have actively-traded options. And, as we have seen, the systematic spread is capturing the lion's share of the predictive power of the total spread for returns and jumps. Thus, we repeat our regressions (12) and (13) without the idiosyncratic spread applied to all non S&P 500 stocks in each month during our sample period.

Panel B of Table 4 reports the average regression coefficients. The results in Column (2)-(7) show that the synthetic systematic spread is useful in predicting both systematic and idiosyncratic jumps of all types. For example, a one-standard-deviation increase in the systematic volatility spread is associated with 1.00% and 1.90% increases in the probabilities of future systematic and idiosyncratic jumps, respectively. These are large increases as the unconditional probability of jumps is (by our definition) just 10%. In contrast to the results in Panel A, the systematic spread does not help in predicting non-jump returns as reflected in the insignificant coefficient in Column (1) (although adding the spread to the regression does increase the adjusted R-squared from 3% to 4%).

Overall, the findings in this section show that the predictive power of the volatility spread on future stock returns and jumps mainly comes from the systematic component. Furthermore, our methodology establishes that the systematic volatility spread continues to be informative on the future likelihood of jumps even for stocks with illiquid or no traded options.

4.3 Optimal Portfolio Choice 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 her portfolio.

Specifically, an investor's utility function U takes the form

$$U = \mu_p - \frac{\gamma}{2}\sigma_p^2,$$

where μ_p represents the mean return of the investor's portfolio, σ_p stands for the return volatility of her portfolio, and γ is the risk aversion level of the investor.

Suppose the investment opportunity set consists of N risky assets and a risk-free asset. The expected returns of the N risky assets are given by $\boldsymbol{\mu}^r = (\mu_1, \mu_2, ..., \mu_N)'$, the variance-covariance matrix of the risky assets is $\boldsymbol{\Sigma}^r$, and the risk-free return is r_f. The investor chooses weights for the N risky assets $\mathbf{w} = (w_1, w_2, ..., w_N)'$ and a weight for the risk-free asset w_0 to maximize her utility function. The investor's utility maximization problem is then given by

$$\begin{aligned} &\max_{w_0,\mathbf{w}} \; w_0 r_f + \mathbf{w}'\boldsymbol{\mu}^r - \frac{\gamma}{2}\mathbf{w}'\boldsymbol{\Sigma}^r\mathbf{w}, \\ &s.t. \; w_0 + \mathbf{w}'\mathbf{e} = 1, \end{aligned} \tag{14}$$

where $\mathbf{e}$ is an $N \times 1$ vector of ones.

A standard approach to solving this optimization problem is to use historical moments of asset returns. Instead, our modified Fama-MacBeth approach yields a forward-looking implied variance-covariance matrix, and we can use it to solve for the optimal investment portfolio. Such an optimal portfolio can be calculated at high frequency (e.g., daily or even intradaily) from option prices and reflect real-time changes in investors' expectations. This is in contrast to a strategy based on historical variance-covariance matrices, which features slow-moving portfolio weights over time. Note that our recovered variance-covariance matrix is calculated under the risk-neutral measure, whereas (14) is calculated under the physical measure, introducing a potential bias in our estimation. Nevertheless, our approach offers a unique setting for the estimation of forward-looking moments, and as we show below, this key feature leads to superior performance as compared to historical estimates, despite this potential bias.

To demonstrate the investment value of using our forward-looking moments, we use the nine SPDR sector ETFs as the set of risky assets and a 30-day T-bill as a risk-free asset. Data on the SPDR sector ETFs are available starting from December 1998. We solve (14) based on historical and forward-looking variance-covariance matrices, respectively, at a monthly frequency. We estimate annualized historical variances and covariances at the end of each month based on daily equity returns of the most recent 12-month rolling window. The forward-looking variance-covariance matrix is estimated at the end of each month from the third stage of our modified Fama-MacBeth approach.

Solving (14) also requires estimates of the expected returns of the sector ETFs. To focus on the differential information from the second moments, we employ the same method to estimate $\boldsymbol{\mu}^r$ for both cases. Specifically, we rely on the arbitrage pricing theory based on the FFC four-factor model:

$$\boldsymbol{\mu}^r = r_f \mathbf{e} + \boldsymbol{\beta}\boldsymbol{\mu}^f,$$

where $\boldsymbol{\mu}^f$ is a 4×1 vector of expectations of the factors. Empirically, we use annualized 120-month rolling averages of the factors as estimates of $\boldsymbol{\mu}^f$. In addition, we use conventional risk aversion levels of 3 and 5 as values of γ in (14), similar to Rapach, Strauss and Zhou (2010).

The optimal weights from (14) under both forward-looking and historical information involve rather large short positions. This is due to the high correlations across different sectors. To make our portfolios more feasible, we impose constraints on the shorting weights. Specifically, we consider both the case without short selling ($w_i \geq 0$, $\forall i = 0, 1, ..., N$) and the case allowing moderate short selling ($w_i \geq -1$, $\forall i = 0, 1, ..., N$).

Table 5 reports the performance of the monthly-rebalanced optimal investment portfolios obtained using historical and forward-looking variances and covariances. For both levels of γ and both types of short-selling constraints, the average and median returns are uniformly much higher based on forward-looking information than historical information. On the other hand, the return volatility of portfolios obtained from forward-looking information is not always higher than that based on historical moments. For example, when $\gamma = 3$ and no short selling is allowed, the monthly mean (median) returns of the forward-looking and historical strategies are 0.58% vs.

0.32% (0.86 vs. 0.60%), respectively, while the corresponding volatilities are 3.99% and 4.13%. Accordingly, the Sharpe ratio of the forward-looking strategy is always higher than that of the historical strategy. Their difference is economically large and statistically significant in most cases applying HAC standard errors (Ledoit and Wolf (2008)).[12] Furthermore, our forward-looking strategy is always more right-skewed and less heavy-tailed (reflected in higher skewness and lower kurtosis). Both features are considered attractive to investors. Thus, while our forward-looking moments are potentially biased being estimated under the risk-neutral measure, empirically the bias is dwarfed by their superior information content on future returns.

We next consider an alternative strategy, which could also potentially provide improvement relative to the historical benchmark. This strategy estimates the variances and covariances using the dynamic conditional correlation multivariate GARCH model (Engle and Sheppard (2001)). The results are also reported in Table 5. In all four cases, the GARCH method provides limited improvement compared to historical estimates, with the Sharpe ratio being somewhat higher than the historical Sharpe ratio, but the difference is never statistically significant. Importantly, our forward-looking strategy beats both the historical and GARCH methods in every single case.

We end this section with an extreme example, illustrating the differential information reflected in forward-looking versus historical variance-covariance matrices. We have calculated (but do not report for brevity) the variance-covariance matrices and the corresponding correlation matrices estimated based on historical and forward-looking information as of December 2008. This month was at the height of the financial crisis, right before asset prices reached the trough level in early 2009. We find that the average historical variance across all sectors is 0.0728, corresponding to a volatility of 27%. By contrast, the average forward-looking risk-neutral variance is 0.1976, corresponding to a volatility of 44%. Similarly, the average historical and forward-looking correlations are 0.4026 and 0.9313, respectively.

[12] The HAC standard errors account for potential autocorrelation and heavy tails in the time series of returns. The only case in which we do not have statistical significance is when $\gamma = 3$ and no short selling is allowed. The associated p-value for this case is 0.1039.

It is hard to tell what portion of the increase in volatility and correlation comes from the change in measure, and what part reflects forward-looking information about heightened volatility and higher correlation during the crisis period. To gauge these two components, we compute the realized variance and correlation of the sectors in January 2009 based on daily sector returns and use these quantities as the ex-post realized moments under the physical measure.[13] The average ex-post realized variance across the sectors is 0.2872, corresponding to a volatility of 54%, and the average ex-post realized correlation is 0.8234. These numbers, which are much higher than the historical moments, indicate that our forward-looking moments, while estimated under the risk-neutral measure, contain useful information on future changes of asset return volatilities. Thus, in December 2008, investors indeed expected sector returns to be both more volatile and way more correlated than reflected in historical moments. Such patterns are rather typical during crisis periods (Roll (1988), Schwert (1990), Jorion (2000), Ang and Chen (2002), and Hong, Tu, and Zhou (2007)).

5 Conclusion

In this paper, we present a novel methodology for recovering implied variances and covariances of both individual assets and portfolios that may not have actively-traded options. We start with a linear factor structure of stock returns and derive a corresponding factor model for the return variance, in which the new factors are the variances and covariances of the original factors and the new factor loadings are cross products of the original factor loadings. We propose a three-stage modification of the Fama-MacBeth approach to estimating this new factor model, yielding forward-looking estimates of the implied volatilities and correlations of both factors and stocks. Empirically, we demonstrate the usefulness of our framework by applying it to form optimal investment portfolios and to predict jumps of both the factors and individual stocks.

In addition to the examples provided in this paper, our methodology could po-

[13] To match with the 30-day maturity of options used to obtain our forward-looking estimates, here we use daily returns within one month to compute the ex-post realized variance and correlation.

tentially lend itself to many more applications. One possible direction is to examine the correlation risk premium, as captured by the difference between our estimated implied correlations and the realized ones, as well as its effects on asset prices (e.g., Driessen, Maenhout, and Vilkov (2009) and Buss, Schönleber, and Vilkov (2018)). Another potential path is to back out prices of synthetic options for factors and assets with no existing traded options. These synthetic options may have implications for risk management. Furthermore, a similar factor-based approach to the one presented here may be used for recovering higher forward-looking moments, such as skewness and kurtosis, and for testing their usefulness for investment decisions. We leave these for future research.

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] Ang, Andrew, and Joseph Chen, 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63, 443–494.

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

[6] Bakshi, Gurdip, and Nikunj Kapadia, 2003a. Delta-hedged gains and the negative volatility risk premium. *Review of Financial Studies* 16, 527–566.

[7] Bakshi, Gurdip, and Nikunj Kapadia, 2003b. Volatility risk premiums embedded in individual equity options: Some new insights. *Journal of Derivatives* 11, 45–54.

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

[9] 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.

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

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

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

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

[14] 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.

[15] Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman, 2016. Misspecified recovery. *Journal of Finance* 71, 2493–2544.

[16] Breeden, Douglas T., and Robert H. Litzenberger, 1978. Prices of state-contingent claims implicit in option prices. *Journal of Business* 51, 621–651.

[17] Buss, Adrian, and Grigory Vilkov, 2012. Measuring equity risk with option-implied correlations. *Review of Financial Studies* 25, 3113–3140.

[18] Buss, Adrian, Lorenzo Schönleber, and Grigory Vilkov, 2018. Expected correlation and future market returns. Working paper, INSEAD and Frankfurt School of Finance & Management.

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

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

[21] Christoffersen, Peter, Mathieu Fournier, and Kris Jacobs, 2018. The factor structure in equity options. *Review of Financial Studies* 31, 595–637.

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

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

[24] Driessen, Joost, Pascal J. Maenhout, and Grigory Vilkov, 2009. The price of correlation risk: Evidence from equity options. *Journal of Finance* 64, 1377–1406.

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

[26] Engle, Robert F., and Kevin Sheppard, 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. No. w8554. *National Bureau of Economic Research.*

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

[28] 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.

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

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

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

[32] Higham, Nicholas J., 1988. Computing a nearest symmetric positive semidefinite matrix. *Linear algebra and its applications* 103, 103–118.

[33] Hong, Yongmiao, Jun Tu, and Guofu Zhou, 2007. Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies* 20, 1547–1581.

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

[35] Jensen, Christian Skov, David Lando, and Lasse Heje Pedersen, 2019. Generalized recovery. *Journal of Financial Economics* 133, 154–174.

[36] Jorion, Philippe, 2000. Risk management lessons from long-term capital management. *European Financial Management* 6, 277–300.

[37] Kadan, Ohad, and Asaf Manela, 2018. Estimating the value of information. *Review of Financial Studies* 32, 951–991.

[38] Kadan, Ohad, and Xiaoxiao Tang, 2019. A bound on expected stock returns. *Review of Financial Studies*, forthcoming.

[39] Ledoit, Oliver, and Michael Wolf, 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance* 15, 850–859.

[40] Liu, Fang, 2018. Option-implied systematic disaster concern. Working paper, Cornell University.

[41] Martin, Ian, 2017. What is the expected return on the market? *Quarterly Journal of Economics* 132, 367–433.

[42] Martin, Ian, and Christian Wagner, 2019. What is the expected return on a stock? *Journal of Finance*, forthcoming.

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

[44] Merton, Robert C., 1973b. An intertemporal capital asset pricing model. *Econometrica* 41, 867–887.

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

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

[47] 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.

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

[49] Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies* 23: 821-862.

[50] Roll, Richard, 1988. The international crash of October 1987. *Financial Analysts Journal* 44, 19–35.

[51] Ross, Stephen A., 1976a. Options and efficiency. *Quarterly Journal of Economics* 90, 75–89.

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

[53] Ross, Stephen A., 2015. The recovery theorem. *Journal of Finance* 70, 615–648.

[54] Schwert, G. William, 1990. Stock volatility and the crash of '87. *Review of Financial Studies* 3, 77–102.

[55] 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.

[56] Skintzi, Vasiliki D., and Apostolos-Paul N. Refenes, 2005. Implied correlation index: A new measure of diversification. *Journal of Futures Markets: Futures, Options, and Other Derivative Products* 25, 171–197.

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

[58] Tang, Xiaoxiao, 2018. Variance asymmetry managed portfolios. Working paper, University of Texas at Dallas.

[59] 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.

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

Appendix: Proof of Theorem 2

In the proof we omit all time subscripts for brevity. Let $\mathbf{V}$ be a symmetric $K \times K$ matrix and let λ be a constant. 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

$$\mathbf{p}_{\mathbf{V},\lambda} = \left(V_{1,1}, V_{1,2}, ..., V_{1,K}, V_{2,2}, V_{2,3}, ..., V_{2,K}, ..., V_{K,K}, \lambda\right)'.$$

Given $(\mathbf{V},\lambda)$, define the least-squares loss function $\mathcal{L} : \mathbb{R}^{\frac{K(K+1)}{2}+1} \to \mathbb{R}$

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

where $Diag\left(\mathbf{V}^r\right)$ is the $L \times 1$ vector consisting of the diagonal elements of $\mathbf{V}^r$ representing implied variances of the L assets with options, $\hat{\mathbf{B}}$ is given by (7), and $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^L$.

Let $\mathbb{V}$ denote the set of $K \times K$ positive semidefinite matrices. The constrained least squares optimization problem corresponding to the regression model in (6) is

$$\min_{\mathbf{V}\in\mathbb{V},\lambda\in\mathbb{R}} \mathcal{L}\left(\mathbf{V},\lambda\right). \tag{15}$$

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 $\mathbf{V}^2$.

The closedness of $\mathbb{V}$ follows because Euclidean convergence implies convergence of each element of the matrices and by continuity of matrix multiplication. ■

Lemma 2 *Let* $\mathbf{B}$ *be an* $L\times\left(\frac{K(K+1)}{2}+1\right)$ *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 $\left\|\mathbf{Bq}^i\right\| \to 0$. 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 $\|\mathbf{q}^*\| = 1$. Now,

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

where the first equality follows by the continuity of the Euclidean norm and matrix multiplication, and the second and third since $\left\|\mathbf{Bq}^i\right\|$ 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 $\mathbb{R}^{\frac{K(K+1)}{2}+1}$. 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 2: $\mathcal{L}$ is a strictly convex function, 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 (15) exists, it must be unique. Therefore, we are left to prove that a solution exists.

Let

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

Then, there is a sequence $\{\mathcal{L}^i\}$ converging to $\underline{\mathcal{L}}$, and a sequence of $\{\mathbf{V}^i,\lambda^i\}$, where $\mathbf{V}^i \in \mathbb{V}$, such that $\mathcal{L}^i = \mathcal{L}\left(\mathbf{V}^i,\lambda^i\right)$. Let $\{\mathbf{p}_{\mathbf{V}^i,\lambda^i}\}$ denote the corresponding sequence of vectors in $\mathbb{R}^{\frac{K(K+1)}{2}+1}$. To prove existence of a solution of (15), it is sufficient to show that $\{\mathbf{p}_{\mathbf{V}^i,\lambda^i}\}$ 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 $\mathbf{p}^* \in \mathbb{R}^{\frac{K(K+1)}{2}+1}$. 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 $\mathcal{L}$, $\underline{\mathcal{L}} = \mathcal{L}(\mathbf{V}^*,\lambda^*)$. Therefore, $(\mathbf{V}^*,\lambda^*)$ is a solution to (15).

Thus, it remains to show that $\{\mathbf{p}_{\mathbf{V}^i,\lambda^i}\}$ 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

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

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,

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

It follows that

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

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 $j > j_0$,

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

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

Table 1: First-Stage Factor Loadings

This table reports summary statistics for factor loadings estimated in the first stage of the modified Fama-MacBeth procedure based on the Fama-French-Carhart four-factor model. The sample period is January 1996 to December 2017. Panel A considers all CRSP common stocks (share code 10 or 11) and Panel B considers all S&P 500 stocks during the sample period. At the end of each month, we estimate the betas of each stock by regressing daily stock excess returns on daily factor returns over 12-month rolling windows. For each stock, we compute the average of each factor loading over time. The table reports the cross-sectional summary statistics of these time averages.

Panel A: All CRSP Common Stocks

	Obs	Mean	Median	S.D.	Min	Max
β_{mktrf}	13,904	0.8233	0.8420	0.5707	-7.5030	9.0293
β_{smb}	13,904	0.7195	0.6896	0.6920	-6.7965	10.2517
β_{hml}	13,904	0.1604	0.2175	0.7937	-10.7894	11.8929
β_{umd}	13,904	-0.1299	-0.0775	0.5808	-17.0211	5.7203

Panel B: All S&P 500 Stocks

	Obs	Mean	Median	S.D.	Min	Max
β_{mktrf}	1,038	1.0617	1.0652	0.2599	0.1030	2.0785
β_{smb}	1,038	0.1920	0.1665	0.3411	-0.8759	1.6330
β_{hml}	1,038	0.1533	0.2091	0.5513	-3.5318	3.0138
β_{umd}	1,038	-0.0931	-0.0637	0.2335	-1.3157	0.9393

Table 2: Second-Stage Implied Factor Variances and Covariances

This table reports summary statistics for the estimated implied variances, covariances, and correlations of the Fama-French-Carhart four factors. These estimates are obtained from the second-stage modified Fama-MacBeth regressions (6) based on factor loadings summarized in Table 1. The sample period is January 1996 to December 2017 and the data consist of all S&P 500 constituent stocks during this period.

	Mean	Median	S.D.	Min	Max
$V^f_{mktrf,mktrf}$	0.0402	0.0270	0.0424	0.0005	0.5623
$V^f_{smb,smb}$	0.0476	0.0355	0.0423	0.0040	0.4732
$V^f_{hml,hml}$	0.0219	0.0183	0.0175	0.0000	0.1872
$V^f_{umd,umd}$	0.0667	0.0401	0.0723	0.0010	0.5410
$V^f_{mktrf,smb}$	0.0147	0.0158	0.0164	-0.1202	0.0824
$V^f_{mktrf,hml}$	-0.0086	-0.0069	0.0193	-0.1136	0.1336
$V^f_{mktrf,umd}$	-0.0054	0.0003	0.0344	-0.2600	0.1199
$V^f_{smb,hml}$	-0.0021	-0.0028	0.0178	-0.0888	0.1582
$V^f_{smb,umd}$	-0.0047	-0.0010	0.0324	-0.4107	0.1358
$V^f_{hml,umd}$	-0.0049	-0.0033	0.0232	-0.1501	0.1163
$\rho^f_{mktrf,smb}$	0.4926	0.5668	0.3388	-0.8891	0.9999
$\rho^f_{mktrf,hml}$	-0.3534	-0.4091	0.4526	-0.9999	0.9998
$\rho^f_{mktrf,umd}$	-0.0578	0.0031	0.4751	-0.9999	0.9866
$\rho^f_{smb,hml}$	-0.0914	-0.1208	0.4333	-0.9998	0.9999
$\rho^f_{smb,umd}$	-0.0682	-0.0178	0.4507	-0.9999	0.9997
$\rho^f_{hml,umd}$	-0.1572	-0.1866	0.4846	-0.9999	1.0000

Table 3: Factor Implied-Volatility Slope and Downward Jump Risk

This table reports summary statistics for the estimated implied-volatility slopes of the market, size, value, and momentum factors. The implied-volatility slope for each factor is estimated by applying the modified Fama-MacBeth procedure. The sample period is January 1996 to December 2017, and the data consist of all S&P 500 constituents within this period. We also report the slope coefficient from regressing the downward jump dummy of each factor on the corresponding estimated implied-volatility slope available at the end of the previous month (see (11)). For comparison purposes, we also include summary statistics and the estimated regression coefficient for the market implied-volatility slope directly obtained from S&P 500 index options. 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.

	(1)	(2)	(3)	(4)	(5)
	S&P 500	$mktrf$	smb	hml	umd
Mean	-0.0460	-0.0075	-0.0154	0.0022	-0.0126
Median	-0.0421	-0.0084	-0.0097	0.0043	-0.0078
S.D.	0.0227	0.0237	0.0531	0.0280	0.0377
Min	-0.2362	-0.1659	-0.5509	-0.2302	-0.2785
Max	0.0556	0.1195	0.4030	0.2148	0.2804
Reg Coef	-0.2492	-2.4994	-0.1617	-1.0465	-1.5259
	(0.2287)	(1.1846)**	(0.2017)	(0.4553)**	(0.6586)**

Table 4: Predictive Power of Systematic and Idiosyncratic Volatility Spreads

This table reports results of predicting monthly returns, systematic jumps, and idiosyncratic jumps using volatility spreads. The sample period is January 1996 to December 2017. Panel A predicts returns and jumps for all S&P 500 stocks during the sample period using the systematic volatility spread ($\sqrt{V^s_{n,n}} - \sqrt{\Sigma^s_{n,n}}$) and the idiosyncratic volatility spread ($\sqrt{V^{\varepsilon}_{n,n}} - \sqrt{\Sigma^{\varepsilon}_{n,n}}$) based on the Fama-French-Carhart four-factor model. Panel B includes all CRSP non S&P 500 common stocks (share code 10 or 11) using only the systematic volatility spread. We define a systematic jump as a jump in the stock price concurrent with a jump in at least one of the four factors and an idiosyncratic jump as a jump in the stock price without a concurrent jump in any of the factors. We perform regressions (12) and (13) monthly and report time averages of the estimated slope coefficients in the table. The risk-neutral systematic and idiosyncratic variances, $V^s_{n,n}$ and $V^{\varepsilon}_{n,n}$, are estimated using our modified Fama-MacBeth approach ((9) and (10)), and the physical systematic and idiosyncratic variances, $\Sigma^s_{n,n}$ and $\Sigma^{\varepsilon}_{n,n}$, are estimated using daily equity returns over each month. The control variables include the CAPM beta, β, firm size, $Ln(Size)$, book-to-market ratio, $Ln(B/M)$, and lagged stock return, *LagRet*. We estimate beta by regressing daily stock excess returns on daily market excess returns over a rolling window of 12 months prior to the month of interest. We estimate firm size as the log of the month-end market capitalization, and book-to-market is the log of the ratio of the book value of equity to market value of equity at the end of each month. The lagged stock return is obtained from the 12 months preceding the month of interest. Newey-West standard errors with 5 lags are reported in the parentheses. Asterisks denote significance at the 1% (***), 5% (**) and 10% (*) levels.

Panel A: Predictive Power of Systematic and Idiosyncratic Volatility Spreads for S&P 500 Stocks

	Return	Systematic Jump			Idiosyncratic Jump		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Downward	Upward	Mixed	Downward	Upward	Mixed
$\sqrt{V^s_{n,n}} - \sqrt{\Sigma^s_{n,n}}$	0.0316	0.0337	0.0719	0.1462	0.2096	0.2253	0.4680
	(0.0124)**	(0.0146)**	(0.0156)***	(0.0270)***	(0.0218)***	(0.0209)***	(0.0448)***
$\sqrt{V^{\varepsilon}_{n,n}} - \sqrt{\Sigma^{\varepsilon}_{n,n}}$	0.0072	-0.0017	0.0091	0.0103	0.0189	0.0324	0.0437
	(0.0032)**	(0.0029)	(0.0028)***	(0.0037)***	(0.0056)***	(0.0060)***	(0.0103)***
β	-0.0001	0.0160	0.0140	0.0386	0.0328	0.0295	0.0388
	(0.0039)	(0.0053)***	(0.0036)***	(0.0081)***	(0.0053)***	(0.0042)***	(0.0055)***
$Ln(Size)$	-0.0028	-0.0014	-0.0030	-0.0052	-0.0034	-0.0072	-0.0111
	(0.0006)***	(0.0006)**	(0.0009)***	(0.0013)***	(0.0005)***	(0.0008)	(0.0010)***
$Ln(B/M)$	0.0003	0.0016	0.0005	0.0024	-0.0004	0.0003	-0.0012
	(0.0007)	(0.0007)**	(0.0009)	(0.0012)**	(0.0008)	(0.0008)	-0.0010
LagRet	0.0027	-0.0046	-0.0047	-0.0140	-0.0152	-0.0052	-0.0142
	(0.0035)	(0.0020)**	(0.0035)	(0.0057)**	(0.0042)***	(0.0045)	(0.0056)**

Panel B: Predictive Power of Systematic Volatility Spread for Non S&P 500 Stocks

	Return	Systematic Jump			Idiosyncratic Jump		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Downward	Upward	Mixed	Downward	Upward	Mixed
$\sqrt{V^s_{n,n}} - \sqrt{\Sigma^s_{n,n}}$	0.0056	0.0224	0.0251	0.0706	0.0844	0.0725	0.1338
	(0.0087)	(0.0066)***	(0.0065)***	(0.0143)***	(0.0088)***	(0.0081)***	(0.0132)***
β	0.0014	0.0120	0.0115	0.0298	0.0211	0.0182	0.0331
	(0.0026)	(0.0036)***	(0.0026)***	(0.0066)***	(0.0026)***	(0.0022)***	(0.0026)***
$Ln(Size)$	-0.0021	-0.0035	-0.0036	-0.0100	-0.0125	-0.0162	-0.0258
	(0.0009)**	(0.0008)***	(0.0008)***	(0.0019)***	(0.0007)***	(0.0010)***	(0.0012)***
$Ln(B/M)$	0.0052	-0.0025	-0.0009	-0.0044	-0.0102	-0.0053	-0.0145
	(0.0012)***	(0.0008)***	(0.0008)	(0.0014)***	(0.0008)***	(0.0007)***	(0.0012)***
LagRet	0.0011	-0.0035	-0.0020	0.0089	-0.0099	-0.0033	-0.0099
	(0.0021)	(0.0016)**	(0.0018)	(0.0033)***	(0.0019)***	(0.0019)*	(0.0028)***

Table 5: Optimal Portfolio Choice

This table reports the performance of monthly-rebalanced optimal investment portfolios obtained based on three methods of variance and covariance estimation: historical, forward-looking, and GARCH. The historical moments are computed based on daily equity returns over the past 12-month window, the forward-looking moments are estimated using our modified Fama-MacBeth approach based on the Fama-French-Carhart four-factor model, and the GARCH moments are estimated using the dynamic conditional correlation multivariate GARCH model. The optimal portfolios are obtained by solving (14) using the nine SPDR sector ETFs as risky assets applying a 30-day T-bill as a risk-free asset. The sample period is December 1998 to December 2017. We consider cases with no short selling and with limited short selling for risk aversion levels of 3 and 5. We also report the improvement in the Sharpe ratio of the forward-looking and GARCH methods relative to the historical strategy. HAC standard errors are displayed in parentheses below the corresponding estimates. Asterisks denote statistical significance at the 1% (***), 5% (**) and 10% (*) levels.

Panel A: No Short Selling						
	$\gamma = 3$			$\gamma = 5$		
	Historical	Forward-Looking	GARCH	Historical	Forward-Looking	GARCH
Mean	0.0032	0.0058	0.0038	0.0020	0.0052	0.0037
Median	0.0060	0.0086	0.0055	0.0038	0.0052	0.0032
Vol	0.0413	0.0399	0.0397	0.0321	0.0356	0.0329
Skew	-0.7129	-0.0108	-0.3668	-0.4469	-0.0702	-0.2597
Kurt	5.9162	4.2637	4.0652	5.3590	4.0449	4.0120
SR	0.2685	0.5024	0.3313	0.2125	0.5091	0.3925
ΔSR		0.2339	0.0628		0.2966	0.1800
		(0.1438)	(0.1209)		(0.1393)**	(0.1181)
Panel B: Limited Short Selling						
	$\gamma = 3$			$\gamma = 5$		
	Historical	Forward-Looking	GARCH	Historical	Forward-Looking	GARCH
Mean	0.0024	0.0332	0.0057	0.0011	0.0312	0.0042
Median	0.0059	0.0372	0.0056	0.0035	0.0257	0.0055
Vol	0.0638	0.1994	0.0744	0.0408	0.1880	0.0514
Skew	-0.1605	0.1279	0.1091	-0.2397	0.3724	0.1118
Kurt	4.8457	3.9806	3.5734	5.0895	4.3895	3.6012
SR	0.1289	0.5773	0.2632	0.0930	0.5756	0.2840
ΔSR		0.4484	0.1343		0.4825	0.1910
		(0.2421)*	(0.1424)		(0.2508)*	(0.1579)