The Factor Structure of Realized Volatility and its Implications for Option Pricing

Zhi Da† and Ernst Schaumburg‡,*

Nov 2006

VERY PRELIMINARY and INCOMPLETE
Please do not quote or distribute

Abstract

The cross-section of realized return volatilities of US equities between 1965-2004 is well described by a linear factor structure. We show that the identified factor structure has important cross-sectional pricing implications for variance swap contracts. The principal volatility factor accounts for almost 40% of the cross-sectional variation in realized stock volatilities and earns a significant negative risk premium. Importantly, more than one volatility factor is priced in the market and a total of 3-5 factors are needed to explain the cross-section of observed volatility swap rates. Moreover, we find strong evidence for the existence of an aggregate jump risk factor being priced by the market. Our findings are robust to the exclusion of stock index contracts and confirm that index options are expensive relative to individual stock options.

† We would like to thank Torben Andersen, Mikhail Chernov, Robert Korajczyk, Paul Gao, seminar participants at the Risk Management Conference (Mont Tremblant) and especially Ravi Jagannathan for his numerous suggestions and insights.
‡ zda@nd.edu, University of Notre Dame.
‡ e-schaumburg@northwestern.edu, Kellogg School of Management, Northwestern University.
1 Introduction

Stock price volatility plays a crucial role in asset allocation and risk management applications. The measurement and forecasting of volatilities has been the subject of extensive research in the realized volatility literature (e.g. Andersen, Bollerslev, Christoffersen, and Diebold (2005)) as well as the implied volatility literature (e.g. Reisman and Zohar (2004)). While most of these studies have focused on individual stock volatilities, the cross-sectional dynamics of volatility and its implications for asset pricing have received much less attention and will be the topic of the current paper.

In arbitrage-free continuous time models, the (log) stock price process can be represented by the sum of a continuous semi-martingale and a jump process (e.g. Sims and Mareswaran (1993)). It follows that the increment in the quadratic variation of the log price process of stock \( i \) during month \( t \) can be decomposed into a sum of two contributions due to the continuous and jump component respectively. In practice the quadratic variation is of course not observable, but it can be proxied for by the realized return volatility \( RV_{i,t} \) (see e.g. Andersen, Bollerslev, Christoffersen, and Diebold (2005)). Conceptually, the realized volatility admits a similar decomposition:

\[
RV_{i,t} = RV_{i,t}^{C} + RV_{i,t}^{J}.
\]

For pricing purposes it will be convenient to work with log realized volatilities in which case the decomposition can be restated as:

\[
\log RV_{i,t} = \log RV_{i,t}^{C} + J_{i,t},
\]

\[
J_{i,t} \equiv \log RV_{i,t} - \log RV_{i,t}^{C} = \log \left( 1 + \frac{RV_{i,t}^{J}}{RV_{i,t}^{C}} \right).
\]

In this paper, we examine the two components of log volatility separately. This choice can be motivated by the assumption that any systematic factor driving rare large price

\footnote{For a given observation frequency, the realized volatility of the log price process over the time interval \([0;T]\) is defined as the sum of squared returns \( \sum_{t=0}^{T} r_{t}^{2} \). In this paper \( T = 1 \) month and the observation frequency is daily.}

\footnote{The “jump” contribution to realized volatility of stock \( i \) in month \( t \), \( J_{i,t} \), is in this case multiplicative rather than additive. It has the interpretation as a percentage increase in volatility due to the occurrence of jumps.}
moves is separate from the factors driving the day to day volatility of stock prices. We first focus on the cross-sectional dynamics of the continuous component of the volatility \( \log \text{RV}^C_{i,t} \) in the US equity market and demonstrate that it is well described by a simple linear factor model. The three principal factors combined account for about 55% of the cross-sectional variation in \( \log \text{RV}^C_{i,t} \), measured by daily squared returns filtered for extreme price movements (i.e. jumps). This factor structure has remained remarkably robust across various sub-samples during the period 1965-2004.

Controlling for the first 5 principal factors, we find little or no serial correlation in the innovations to the monthly realized volatility \( \log \text{RV}^C_{i,t} \) of individual stocks despite their considerable cross-sectional dependence. The implication is that estimated factor structure effectively summarizes the forecastable component of realized volatilities although it does not capture all contemporaneous dependence. This observation is useful for joint forecasting of realized volatilities as it provides an alternative to the commonly used univariate approaches (see e.g. Andersen, Bollerslev, Christoffersen, and Diebold (2005)). Moreover, we find that forecast errors will tend to co-vary cross-sectionally and this correlation can be estimated ex-ante. This second observation is of particular importance in portfolio risk management applications.

In an ICAPM framework, the existence of a factor structure for volatility has a significant implication: aggregate volatility factors are natural candidate pricing factors. This is broadly consistent with recent empirical evidence showing that market index volatility risk as priced, as in Ang, Hodrick, Xing, and Zhang (2005) using stock data, and Bakshi and Kapadia (2003a), Bakshi and Kapadia (2003b), Bondarenko (2004) and Carr and Wu (2004), using stock option data. In the options market realized index volatility is traded in the form of variance swaps, e.g. the VIX contract. For individual stocks, however, variance swap contracts only trade in the over-the-counter market, rendering the swap rate generally unobservable. Instead, the variance swap contract can be non-parametrically approximated by a basket of put and call options which replicate the swap contract in the absence of jumps, as first pointed out by Neuberger (1994). Under the assumption of no jumps, we construct the implied variance swap rates for a cross section of stocks with sufficient trading volume during each month between 1997 and 2004. Using the constructed synthetic swap rates, we confirm that at least two of the aggregate volatility factors carry a significant risk premium. The primary factor, highly correlated with measures of realized market index volatility, has a negative market price of volatility risk roughly similar
to that found by other authors (using the S&P500 index volatility as a proxy for overall market volatility). The third factor has a small but positive correlation with measures of market index volatility, yet carries a significant positive risk premium. The remaining factors, while statistically important, do not appear to be significantly priced in the market. In terms of the in-sample adjusted R-squared of the factor pricing equation, the market volatility factor explains only about 5.6% of the cross-sectional variation in returns on variance swaps. The one factor model is able to explain 12.3% while the 3 and 5 factor model achieve 19.6% and 43.1%, respectively. Out-of-sample both the market volatility and one factor models fare poorly while the three and five factor models do substantially better in terms of root mean squared pricing error.

We extend our analysis to explicitly account for the presence of jumps within our linear factor pricing framework. Interestingly, an aggregate jump factor constructed based on jumps in market index volatilities is also significantly priced. With the inclusion of the jump factor, the in-sample cross-sectional explanatory power increases significantly: the 3 factor model with jump factor and 5 factor model with jump factor now explain 39% and 51.9%, respectively. Out of sample the impact of the jump factor is negligible, but this may in part be due to the relatively short 6 month sample period.

Across all specifications, the estimated factor models imply that index volatility in some cases is relatively expensive compared to individual stock volatility. Notably, the magnitude of the relative expensiveness shrinks markedly once one accounts for multiple volatility factors and/or jump factors and for some indices there appears to be no systematic mis-pricing.

The identified factor structure in volatility imposes pricing restrictions on the cross-section of variance swap contracts. In particular, any well diversified portfolio of variance swaps constructed to have zero factor loadings on the aggregate volatility factors should incur zero cost, up to the transaction costs. Since the variance swap can be accurately replicated using standard calls and puts, any violation of this restriction will automatically create opportunities for arbitrageurs. As an illustration, we construct a buy-and-hold long-short option strategy which exploits estimated mispricing. We show that this simple strategy earns an annual Sharpe Ratio of around 1.55 during the period from 1997-2004, even after accounting for direct transaction cost (bid-ask spread and margin requirements).

The remainder of the paper is structured as follows. Section 2 summarizes our find-
ings on the factor structure of realized volatility in US equities. Section 3 derives the pricing implications of the volatility structure. Section 4 considers the implications of nondiversifiable jump risk and derives an expression for pricing of jump risk. Section 5 shows how to construct option portfolios that will exploit any potential mispricing in the cross-section. Section 6 concludes.

2 The Factor Structure in Volatility

It is a well-known empirical regularity that stock price volatilities tend to co-move. One can naturally think of this phenomenon as the result of temporal variation in the level of “aggregate” uncertainty in the market and much of the literature has tacitly taken the volatility of the market index to be the single systematic volatility factor. Few studies, however, have explicitly considered the structure of the cross-sectional dependence in volatilities. In this section we present evidence of a simple factor structure in the realized volatilities of US equities between January 1965 and 2004.

Realized return volatilities can be decomposed into continuous and jump components as in (1). In this section, we explicitly focus on the cross-sectional dependence between the continuous components of realized volatilities, deferring the study of the jump component to section 4. For each month we calculate a log realized volatility measure - $V_{i,t}$ of each stock in CRSP. To ensure no undue bias from illiquidity in the realized volatility calculation, we in a given month filter out stocks whose average trading price fell below 5 dollars in that month as well as stocks with less than 15 active trading days (as defined by positive share volume). We also filter out extreme price moves (i.e. jumps). These jump components will be dealt with separately in the asset pricing tests in section 4. For each stock $i$ and each month $t$, $V_{i,t}$ denotes the (annualized) log realized volatility after filtering out jumps using a 3 robust standard deviation filter:

$$V_{i,t} \equiv \log RV_{i,t}^C = \log \left[ \frac{252}{n_t - \sum \mathbf{1}_{\{|r_{i,t,d}|<3\sigma_{\text{robust}}\}} \sum_{d=1}^{n_t} r_{i,t,d}^2 \mathbf{1}_{\{|r_{i,t,d}|<3\sigma_{\text{robust}}\}} \right]$$

3We work with log volatility, a more natural measure when underlying returns are log-normal.
4In fact, factor extraction using principal component analysis can be quite sensitive to such outliers. Our results remain robust to the choice of filter as long as a handful of extreme observations are removed.
5Section 4 below provides details on the jump identification.
Consider the model

$$V_t = \mu + B' F_t + \varepsilon_t, \quad t = 1, \ldots, T$$

(2)

where $V_t$ is a $N \times 1$ vector of measured log realized volatilities in month $t$ for the subset of CRSP that traded continuously throughout the interval $[1; T]$. The $r \times 1$ vector $F_t$ refers to the set of common components in realized volatilities that have to be determined and $B$ refer to the corresponding factor loadings. $B$ and $F_t$ are clearly not jointly identified since the factors can be pre-multiplied by an invertible $r \times r$ matrix without changing the model. It is therefore customary to impose the additional identifying requirement that the factors are uncorrelated and have unit variance.

The model (2) is an approximate factor model with $N \gg T$ such as considered by (Bai and Ng 2003), and can be solved efficiently for the factors $\hat{F}_t$ and factor loadings $\hat{B}$ using the Asymptotic Principal Component Analysis (APCA) as suggested by Connor and Korajczyk (1986) and Connor and Korajczyk (1988). Conditional on knowing the true number of factors, the APCA approach solves the minimization problem

$$\min_{\mu, B, \{F_t\}} \text{SSE}_r = \sum_t \sum_i (V_{i,t} - \mu_i - B'_i F_t)^2$$

(3)

The solution sets the estimated factors $\hat{F}_t$ equal to the $r$ eigenvectors belonging to the $r$ largest eigenvalues of the $T \times T$ matrix $\hat{V}'\hat{V}$, where $\hat{V} = [V_1 - \hat{\mu}, \ldots, V_T - \hat{\mu}]$. The factor loadings are then estimated as $\hat{B} = \hat{V} \hat{F}'$ where $\hat{F} = [\hat{F}_1, \ldots, \hat{F}_T]$.

The fraction of the variance of the observed realized volatilities explained by the $r-$factor model is given by

$$v(r) = \frac{\lambda_1 + \cdots + \lambda_r}{\lambda_1 + \cdots + \lambda_T}$$

(4)

where $\lambda_1, \ldots, \lambda_T$ are the eigenvalues of the matrix $\hat{V}'\hat{V}$. It can be shown that if the true number of factors is $r$, then the first $r$ eigenvalues of $\hat{V}'\hat{V}$ will diverge as $N, T \to \infty$ and the function $v(r)$ acts as a basis for selecting the correct number of factors. (Bai and Ng 2003) suggest the following information criteria and show that they under general conditions on the idiosyncratic errors $\{\varepsilon_t\}$ lead to consistent estimates of the true number of factors as $N, T \to \infty$:
\[ IC_1 : \log [SSE_r] + r \left( \frac{N+T}{NT} \right) \log \left( \frac{NT}{N+T} \right) \]
\[ IC_2 : \log [SSE_r] + r \left( \frac{N+T}{NT} \right) \log (\min(N,T)) \]

The factor loadings are assumed to remain constant over the estimation interval \([1;T]\). In practice this is not likely the case over very long horizons. We, therefore, as in Connor and Korajczyk (1988), choose to estimate the model (2) over 8 non-overlapping 5 year windows from 1965-2004 rather than over the entire period. This procedure also boasts the advantage of dramatically increasing the size of the cross-section \((N)\) in each window relative to the full sample period since relatively few stocks were continuously traded during the entire 40 year period. Moreover, the survivorship bias introduced by the requirement of non-missing observations is also reduced. Table 1 shows the results of the APCA for each of the 8 windows. The results consistently indicate the presence of 3-4 factors which, combined, explain about 40-60% of the cross-sectional variation in realized volatility measure \(V_{i,t}\). The first factor is clearly dominant and alone accounts for 20-50% of the observed variation. Notably, this factor is highly correlated with the market index volatility during most of the periods, lending support to the common practice of considering market index volatility as the sole source of systematic time variation in volatilities. However, the next two factors contribute an additional 10-15%, which is statistically significant according to the information criteria. The implication is that realized volatilities of individual stocks are well approximated by a 3-factor structure:

\[ FM_3 : \ V_{i,t} = \mu_i + b_{1,i}F^1_t + b_{2,i}F^2_t + b_{3,i}F^3_t + \epsilon_{i,t}^{(3)} \]  

(5)

where \(F^i\) is the \(i^{th}\) principal factor. One would then like to think of \(\epsilon_{i,t}^{(3)}\) as stock \(i\)'s “idiosyncratic” volatility innovations due to, say, firm specific news. However, \(\epsilon_{i,t}^{(3)}\) is not idiosyncratic in the strict factor model sense because they fail to remain cross-sectionally and serially uncorrelated. Figure 1(A) displays the cross-sectional distribution of the \(N\) autocorrelations of order 1-4 for \(\epsilon_{i,t}^{(3)}\) during 2000-2004. There is strong evidence of pos-

---

The APCA analysis requires that there be no missing observations. An alternative, but computationally much more challenging, approach would be to use the Kalman filter to extract the factors. In this case missing observations could be dealt with seamlessly. We do not explore this further here.
itive autocorrelation of order 1 for about 23% of the stocks, not surprising per se since the criterion \((3)\) simply minimizes the sum of squared residuals. Figure 1 (B)-(C) show that adding a fourth or a fifth factor removes most of the autocorrelation, rendering the residuals white. The implication is that in a four- or five-factor model, the factors largely capture the predictable component of realized volatilities. Accordingly, we also consider the four and five factor models in the sequel:

\[
\text{FM}_4 : V_{i,t} = \mu_i + b_{1,i}F_{i1} + b_{2,i}F_{i2} + b_{3,i}F_{i3} + b_{4,i}F_{i4} + \varepsilon_{i,t}^{(4)}
\]

\[
\text{FM}_5 : V_{i,t} = \mu_i + b_{1,i}F_{i1} + b_{2,i}F_{i2} + b_{3,i}F_{i3} + b_{4,i}F_{i4} + b_{5,i}F_{i5} + \varepsilon_{i,t}^{(5)}
\]

Figure 1(D) displays the distribution of cross sectional correlations, \(\text{corr}(\varepsilon_{i,t}^{(5)}, \varepsilon_{j,t}^{(5)})\). Clearly, there is a significant amount of cross-sectional correlation in the unforecastable component of realized volatilities.\(^7\) This contemporaneous cross-sectional dependence persists even with the addition of a large number of factors. Equations (5)-(7) represent approximate rather than strict factor models.\(^8\)

### 2.1 Forecasting Using Approximate Factor Models

As we have seen, the first 5 estimated factors capture the majority of the forecastable component of the cross-section of realized volatilities. It is then pertinent to estimate the joint dynamics of the factors themselves. We estimate three alternative VARMA models based on the monthly factor realizations of the first 5 factors between 2000 and 2004. Table 2 shows that the VAR(1) specification as the preferred model.\(^9\) The factors are, by construction unconditionally uncorrelated. In general, they will be conditionally correlated. This observation is confirmed by the estimated AR parameters in Table 3, which indicate considerable feedback between factors, necessitating joint estimation and forecasting of the factor dynamics. The factors along with the fitted VAR(1) and residuals are displayed in Figure 2 (A)-(B).

---

\(^7\) We are forecasting volatilities but not covariances, which would be needed for, say, stock portfolio risk management (unless one is willing to assume a constant correlation structure). For a portfolio of variance swap contracts, as we will consider below, forecasting volatility is sufficient.

\(^8\) Therefore, the pricing of volatility risk using standard APT arguments is not feasible.

\(^9\) Due to the short time-series (60 monthly observations), we limit our attention to sub-models of VARMA(1,1). Tables 4-5 along with Figure 3 show that the VAR(1) specification successfully captures the factor dynamics.
Based on the estimated VAR(1) model, the factors can be forecasted from samples for the months January 2005 to June 2005 as shown in Figure 2 (A)-(B). Combining the factor forecasts with the estimated factor loadings \( \{\hat{b}_1, \ldots, \hat{b}_5\} \) between 2000-2004 then yields forecasts of individual stock volatilities for the first six months of 2005. Moreover, since the \( N \times 1 \) vector of factor model residuals \( \epsilon_i^{(5)} \) are serially uncorrelated, the covariance matrix of forecast prediction errors \( \text{Var}(\epsilon_i^{(5)}) \) can be consistently estimated. This multivariate factor model provides an alternative to the commonly used univariate approaches in forecasting volatility.

3 Pricing of Volatility Factors

The identified factor structure on volatilities implies that the volatility factors \( F \) are natural state variables which describe the investor’s investment opportunity set. In Merton’s ICAPM framework, covariance with \( F \) will therefore be priced\(^\text{10}\). We can write the SDF as

\[
m = \delta(\lambda_0 - \lambda_1 R_m - \lambda_2 \Delta F),
\]

where,

\[
\delta = 1/R_f, \\
\lambda_1 = -J_{WW}^W/J_W \equiv \gamma, \\
\lambda_2 = -J_{WF}/J_W.
\]

and \( J, W \) are the investor’s value function and the level of wealth respectively. Consider the case where the market volatility presents itself as the only volatility factor. The sign of the variance risk premium \( \lambda_2 \) is likely to be negative, since a market downturn usually coincides with high market volatility so that assets with high sensitivity to market volatility serve as a good hedge \((J_{WF} > 0)\). Alternatively, the marginal value of a dollar in high volatility states is higher.

The SDF specification leads to the pricing equation:

\(^\text{10}\)This can be shown using an ICAPM type of derivation, as done in Cochrane (2001), p 167.
\[ E[MR] = 1 \implies E[R] - R_f = \lambda_1 \text{cov}(R, R_m) + \lambda_2 \text{cov}(R, \Delta F). \] (9)

The same intuition holds in the return log-linearization framework of Campbell (1993). Using his setup (with Epstein and Zin utility) and notation, the SDF with time-varying volatility is (up to lognormal adjustment and a constant):

\[ m_{t+1} = \gamma E_t \Delta r_{m,t+1} + (1 - \gamma)(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} \]
\[ - \frac{\theta}{2\sigma} (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \mu_{m,t+j} \]
\[ \mu_{m,t} = \text{var}_t [\Delta c_{t+1} - \sigma r_{m,t+1}], \] (10)

where \( r_m \) and \( \Delta c \) represent (log) market return and aggregate consumption growth, respectively. The conditional variance \( \mu_{m,t} \) enters directly in the pricing kernel. Applying Campbell’s argument, any factor that can predict future market return (as in second term) or predict future volatility (as in the third term) is a good state variable that will be priced. Since aggregate consumption growth is less volatile empirically (\( \text{var}_t [\Delta c_{t+1}] \) is small) and does not much vary with market return (\( \text{cov}_t [\Delta c_{t+1}, r_{m,t+1}] \) is small), \( \mu_{m,t} \approx \sigma^2 \text{var}_t [r_{m,t+1}], \) one may argue for the volatility of a stock market index (denoted as \( V_m \)) as a good candidate for such a pricing factor. This choice of volatility factor will be subject to the usual misspecification critique — the stock market return is only a proxy for the return on the “true” market portfolio \( r_m \). With that consideration, the statistical volatility factors \( F \), extracted using the APCA, could be better predictors for \( \mu_{m,t} \). Whether this is indeed the case presents an interesting empirical question for future research.

The empirical results of several recent papers can easily fall under the interpretive framework described above. Although most of these papers considered market index volatility \( V_m \), the effect is analogous if we replace \( V_m \) with the first statistical factor \( F_1 \) as the two are highly correlated. We therefore use \( V_m \) and \( F_1 \) as interchangeable for the remainder of this particular section.

\[ \text{(9) implies that high } \text{cov}(R, \Delta V_m) \text{ leads to low return as shown in } \text{Ang, Hodrick, Xing}. \]
and Zhang (2005). Adrian and Rosenberg (2006) decompose the market volatility into long-run and short-run component and show that the return covariance with each component is priced. For the market return itself, (9) becomes:

\[ E[R_m] - R_f = \lambda_1 \text{var}(R_m) + \lambda_2 \text{cov}(R_m, \Delta V_m). \] (11)

Ghysels, Santa-Clara, and Valkanov (2005) focus on the first term and find a positive relation between risk and return when \( \text{var}(R_m) \) is estimated using squared daily returns. The risk aversion parameter (\( \lambda_1 = \gamma \)) is estimated as approximately around 2.6. Guo and Whitelaw (2005) explicitly model both the risk component (the first term) and the hedging component (the second term in our context). They find the two terms to be negatively correlated in data; therefore ignoring the hedging component leads to an under-estimation of the risk aversion parameter. Consistent with this pattern, they estimate a higher risk aversion parameter of around 4.93.\footnote{Neither paper explicitly addresses the usual misspecification error — that the stock market return is only a proxy of the return on the wealth portfolio (\( r_m \)).}

A negative variance risk premium is also found using option data. The identification strategy constructs a market-neutral option portfolio sensitive only to volatility risk. Two examples are: a delta hedged option, as in Bakshi and Kapadia (2003a) and Bakshi and Kapadia (2003b), and a synthetic variance swap, as in Bondarenko (2004) and Carr and Wu (2004). For these portfolios, using equation (9) and Stein’s Lemma, we have:

\[
\lambda_1 \text{cov}(R_o, R_m) = \lambda_1 E \left[ \frac{\partial R_o}{\partial V_i} \right] \text{cov}(V_i, R_m),
\]

\[
\lambda_2 \text{cov}(R_o, \Delta V_m) = \lambda_2 E \left[ \frac{\partial R_o}{\partial V_i} \right] \text{cov}(V_i, \Delta V_m)
\]

\[= \lambda_2 \beta_{1,i} E \left[ \frac{\partial R_o}{\partial V_i} \right] \text{var}(\Delta V_m), \]

\[\beta_{1,i} = \text{cov}(V_i, \Delta V_m) / \text{var}(\Delta V_m), \]

Other partial derivatives are effectively zero by portfolio construction. Then (9) be-
comes:

\[ E[R_o] - R_f = \lambda_1 E \left[ \frac{\partial R_o}{\partial V_i} \right] \text{cov}(V_i, R_m) + \lambda_2 \beta_{1,i} E \left[ \frac{\partial R_o}{\partial V_i} \right] \text{var}(\Delta V_m). \] 

(12)

Most of the empirical findings on variance risk in option market can be understood using the pricing equation (12). First, using index option data, the market price of aggregate variance risk is shown to be negative as in Bakshi and Kapadia (2003a), Bondarenko (2004) and Carr and Wu (2004). Second, for the index option, \( V_i = V_m \) and \( \beta_i = 1 \). The first term \( \lambda_1 E \left[ \frac{\partial R_o}{\partial V_i} \right] \text{cov}(V_m, R_m) \) is usually estimated to be negative since down markets tend to associate with above-average levels of volatility although the magnitude is empirically small. Therefore, if the option portfolios have very large negative returns on average, as found in the aforementioned papers, it must be the case that \( \lambda_2 \) is negative. Third, the excess return on the option portfolio could be positive for individual stock option, as found in both Bakshi and Kapadia (2003b) and Carr and Wu (2004). This is consistent with (12) provided that \( \beta_{1,i} \) is negative, which means that the underlying stock tends to have lower than average volatility when aggregate volatility is high. Finally, in a cross sectional regression, the expected excess return on the option portfolio should decrease in \( \beta_{1,i} \). Carr and Wu (2004) first documented this using variance-swap-replicating option portfolios in a small sample of 40 stocks and stock indices. Arguably, the relationship documented in Carr and Wu (2004) comes largely from index options. In the next subsection, we extend their analysis to our volatility factor model using a larger sample.

### 3.1 Pricing tests using replicated variance swaps

Let \( SW_t \) denote the swap rate at time \( t \) of a contract that pays an amount \( RV_{t+1} \) at time \( t + 1 \), which is equal to the realized variance from \( t \) to \( t + 1 \).

\[
\begin{align*}
SW_t &= E_t[M_{t+1}RV_{t+1}] \implies \\
SW_t &= E_t[\exp(m_{t+1} + \log RV_{t+1})] \\
\log SW_t &= E_t[m_{t+1}] + E_t[\log RV_{t+1}] + \frac{1}{2} \text{var}_t[m_{t+1}] + \frac{1}{2} \text{var}_t[\log RV_{t+1}] + \text{cov}_t[m_{t+1}, \log RV_{t+1}].
\end{align*}
\]

(13)
Denote one period’s gross risk free return as $R_{f, t+1}$ and $r_{f, t+1} = \log R_{f, t+1}$, we have,

$$
1 = E_t[M_{t+1} R_{f, t}] \implies 
0 = E_t[m_{t+1}] + r_{f, t} + \frac{1}{2} \text{var}_t[m_{t+1}].
$$

Combining (13) and (14), we obtain:

$$
E_t \left[ \log \left( \frac{RV_{t+1}}{SW_t} \right) \right] - r_{f, t} + \frac{1}{2} \text{var}_t[\log RV_{t+1}] = -\text{cov}_t[m_{t+1}, \log RV_{t+1}].
$$

If we consider a linear beta pricing model where $m_{t+1} = a_t - b_t' F_{t+1}$ and ignore the jump component in the volatility\footnote{The jump component will be dealt with explicitly in the next section.}, then the excess return of the variance swap (after the convexity adjustment term) is linear in volatility factor betas:

$$
E_t \left[ \log \left( \frac{RV_{t+1}}{SW_t} \right) \right] - r_{f, t} + \frac{1}{2} \text{var}_t[\log RV_{t+1}] = \beta' \lambda_t,
$$

\hspace{1cm}

\begin{align*}
\beta_i &= \frac{\text{cov}(F_{i,t+1}, \log RV_{t+1})}{\text{var}(F_{i,t+1})}, \\
\lambda_{i,t} &= \text{var}(F_{i,t+1}) b_{i,t}.
\end{align*}

The variance swap is mostly traded in an OTC market where price is not easily available\footnote{The market for variance swaps has increased in size dramatically. According to Richard Carson, Deutsche Bank’s London-based global head of structured products trading, the market for variance swaps was more than 1 billion euros in vega in 2005, which represents about 300 billion of equivalent options notional.}. However it can be very accurately replicated using portfolios of calls and puts as discussed in Bondarenko (2004) and Carr and Wu (2004). In particular, equation (49) in Carr and Wu (2004) shows that the variance swap contract can be replicated by a continuum of positions in out-of-the-money (OTM) calls and puts:
\[
SW_i = E^Q[RV_i] = \frac{2e^{rt}}{T-t} \int_{-\infty}^{F} \frac{P(K)}{K^2} dK + \int_{F}^{\infty} \frac{C(K)}{K^2} dK.
\] (16)

The weight on an option with a strike of \(K\) is \(w(K) = \frac{2e^{rt}}{(T-t)K^2}\) and (16) can be estimated accurately by interpolating the implied volatility surface and using numerical integration. In a cross sectional regression, Carr and Wu (2004) test a similar version of (15) using only one volatility factor — the volatility on the S&P 500 index. Working with a small sample of 35 stocks and 5 indices, they document a negative risk premium. In this paper, we test (15) directly using the volatility factors we extracted using the APCA in a larger sample.

We obtain the option data from the OptionMetrics’ Ivy database. From 1997 to 2004, on the Monday after the third Friday in each month\(^{14}\) we keep those options that mature in the next month with at least 2 OTM calls and 2 OTM puts and positive trading volume. The number of individual stocks included in our sample per month increases from 90 in 1997 to more than 300 in 2000 and then drops to 200 in 2004 (see Table 6, Panel A). The number of traded strikes per option for individual stock option averages around 6.3. We also include 10 stock index options from the major index list in the OptionMetrics. The underlying stock indices are listed in Panel B of Table 6. Compared to individual stock options, the range of strikes on which index options are traded is much wider. The average number of traded strikes per option is above 20 for index option. For each option, OptionMetrics provides its implied volatility, adjusted then for dividends and the American exercise feature. Based on these option implied volatilities, we compute the implied variance swap rate \(SW_i\) using (16).

Having an the unbalanced panel, we decide to test the factor pricing model (15) using the Fama-MacBeth regression approach. In each month \(t\) and for stock \(i\) which has variance swap rate \(SW_{i,t}\), we compute the stock’s univariate volatility factor betas by regressing \(V_{i,t+1}\) on each volatility factors in a five-year rolling window\(^{15}\). Given the rolling window, we re-estimate the volatility factors from 1992 to 2004. We also estimate the conditional variance on realized volatility - \(\text{var}_i[\log RV_{i,t+1}]\) as the variance on \(\Delta \log RV_{i,t+1}\)

\(^{14}\)We choose this particular day because option trading volume is much higher.

\(^{15}\)We require a stock to have a minimum of 24 months of data to be included in the rolling window regression.
in the same rolling window. In each month $t$, we then run a cross-sectional regression:

$$\log \left( \frac{RV_{i,t+1}}{SW_{i,t}} \right) - r_{f,t} + \frac{1}{2} var_t[\log RV_{i,t+1}] = \lambda_{0,t} + \hat{\beta}_{i,t}^\prime \lambda_t + u_{i,t}. $$

Finally, we compute the time series average of $\lambda_{0,t}$, $\lambda_t$ and the associated t-values. We also compute the Newey-West t-values, which account for the autocorrelation of the estimates with a lag of 12. Bai and Ng (2003) criteria calls for a three factor model; we show that five factors are needed in order to keep the errors in the factor model serially uncorrelated. A five-factor model is therefore considered as supplementary to the benchmark three-factor model.

The regression results are provided in Table 7, Panel A. For all volatility-factor models, the intercept terms of the regressions do not significantly differ from zero. It follows that we cannot reject the factor models [15]. Both $F_1$ and $F_3$ are significant. In particular, $F_1$ which is highly correlated with the market volatility, also carries a negative risk premium, confirming the findings in previous literature. In contrast, $F_3$ has a low but positive correlation with measures of market index volatility but carries a significant positive risk premium. The second factor, while statistically important, does not appear significantly priced in the market. The regression results can be more intuitively represented using one cross-sectional regression similar to that used in Carr and Wu (2004). From 1997 to 2004, for each stock and index with more than 35 observations across time[16] we compute the time-series averages of the actual excess returns on their variance swaps. We then run a cross sectional regression of these excess returns on their volatility factor betas. The regression results are plotted in Figure 4. A one-factor volatility model with its factor as the realized volatility on the market index gives an Adjusted R-square ($AR^2$) of 0.056. If we use the first principal factor ($F_1$) instead, the adjusted R-square improves to 0.123. The $AR^2$ increases with the inclusion of more volatility factors. A five-factor model has a $AR^2$ of 0.431. The stock indices consistently lie below the 45 degree line most of the time, proving that index options in fact are relatively more expensive compared to individual stock options. However, this relative expensiveness of index options shrinks with the inclusion of more volatility factors in the model. In contrast to Carr and Wu (2004), the pricing results are not purely driven by index options.

---

[15]This leaves us with 128 observations in the cross-section.
Since the first factor, $F_1$, highly correlates with the market volatility throughout our sample (see Table 1), it is interesting to test whether the market volatility provides any additional explanatory power in addition to our benchmark three-factor volatility model. We show that this is not the case. We compute the component of market volatility that is orthogonal to $F_1$ to $F_3$ and find that this component is not significant in the regression once $F_1$ to $F_3$ are included (see Table 7, Panel B). Finally, we augment the three-factor volatility model with the usual return factors, such as the market excess returns and the Fama-French three factors. Although these return factors play an important part in explaining the excess returns on stocks, they are not significant in explaining the excess returns on variance swaps once $F_1$ to $F_3$ are included (see Table 7, Panel C).

Finally, we conduct out-of-sample pricing tests using the swap rates from the first 6 months of 2005. We keep those stocks we can compute the swap rate for each of the 6 months. For each stock in our sample, we then compute a predicted excess return on the variance swap as an inner product between factor betas (computed in the rolling window ending in Dec 2004) and factor risk premia (computed in the previous Fama-MacBeth regression) suggested by the pricing equation (15). We then plot the predicted excess returns on variance swaps against the actual average excess returns (averaging across the first 6 months of 2005) in Figure 5. We also report the Root Mean Square Errors (RMSE). Evidently, the one factor volatility model, whether the one factor is market volatility or the first principal volatility factor, has little explanatory power. The returns on index variance swaps are consistently over-estimated, reflecting the underpricing of index options. The out-of-sample performance increases significantly with 3 or 5 factor models. We see a reduction in RMSEs, and a disappearance of systematic mispricing.

4 The importance of Jump risk premia

So far the focus has fallen on the pricing of volatility risk with the assumption that stock returns follow a continuous semi-martingale. The key advantage of the continuity assumption is in its allowing for the model-free construction of variance swap rates given in (16). In this section we consider the implications of jump discontinuities for the pricing of synthetic variance swaps and in particular whether jump risk is priced at the individual stock level. We choose to address the question of jump risk premia within the linear factor pricing framework of the previous section.
4.1 Synthetic Swap Rates with Jumps

From a technical standpoint, the “model-free” expression for the variance swap rate in terms of a basket of out-of-the-money put and call options must be adjusted to account for the occurrence of jumps as shown by Carr and Wu (2004):

$$SW_i = E_Q[RV_i] = \frac{2e^{r(T-t)}}{T-t} \int_{-\infty}^{F} \frac{P(K)}{K^2} dK + \int_{F}^{\infty} \frac{C(K)}{K^2} dK + \varepsilon$$

(17)

where

$$\varepsilon = -\frac{2}{T-t} E_Q\left[ \int_0^T \int_{\mathbb{R}\setminus\{0\}} \left[ e^x - 1 - x - \frac{x^2}{2} \right] v_s(x) dx ds \right]$$

(18)

and $v_s(x)$ denotes the expected number of jumps (in the underlying stock price process) that fall in the interval $[x; x + dx]$ during the infinitesimal time period $[s; s + ds]$ conditional on the current log price level\(^{17}\). The resulting formulation is no longer “model-free” in the sense that it includes a jump-correction term which will depend on the specific parametrization of the jump process.

For the purpose of simple empirical implementation, we choose to model jumps in the log-price process by constant intensity Poisson processes with independent increments. Let $(\lambda_+, \mu_+)$ and $(\lambda_-, \mu_-)$ denote the intensity and average size of positive and negative jumps, respectively. Then:

$$v_s(x) = \begin{cases} 
\frac{\lambda_+ e^{-x/\mu_+}}{\mu_+} & \text{if } x > 0 \\
\frac{\lambda_- e^{-x/\mu_-}}{\mu_-} & \text{if } x < 0
\end{cases} \quad \lambda_+, \lambda_-, \mu_+, \mu_- \in \mathbb{R}_+$

(19)

Since the compensator given by (19) is not state dependent, the jump correction term can be calculated directly as a function of the parameters by noting that the risk neutral expectation, $E_Q[\cdot]$, simply corresponds to discounting at the risk-free rate:

$$\varepsilon = -\frac{2e^{-r(T-t)}}{T-t} \left[ \frac{\lambda_+ \mu_+^3}{1 - \mu_+} - \frac{\lambda_- \mu_-^3}{1 + \mu_-} \right]$$

(20)

\(^{17}\)The quantity $v_s(x)$ is also known as the compensator of the jump process. The formula (18) is developed in Carr and Wu (2004) under the assumption of a finite arrival rate of jumps.
4.2 Identification of Jumps

In order to implement (20), we have to first decide on what constitutes a jump as opposed to a large diffusive move in the log price process of a stock. There is a growing literature on jump identification and detection based on high-frequency intraday data in the context of realized volatility estimation. Here we deal with daily data and therefore require a more ad-hoc identification strategy. For each stock $i$ and each year $k$, we compute the 16% and 84% percentiles, $\kappa_{16}, \kappa_{84}$, from the approximately 252 observed daily log returns and define the robust standard deviation as

$$
\sigma_{i,k}^{\text{robust}} \equiv \frac{\kappa_{84} - \kappa_{16}}{2}
$$

This definition is robust to outliers and it accounts for the temporal variations in the level of volatility from year to year. In our empirical implementation, jumps are identified as one-day log returns larger in magnitude than $3\sigma_{i,k}^{\text{robust}}$ Taking the identified jumps to be observations from a Poisson jump process censored at $\pm 3\sigma_{i,k}^{\text{robust}}$ for positive and negative jumps respectively, the jump intensities and average jump sizes can be estimated by maximum likelihood. Table 8 summarizes the distribution of empirical estimates of the parameters of the jump process in our sample. With most stocks, the typical jump size (1.5-4.5%) and the intensity (1-4 per year) are small; the resulting jump correction term in (17) typically contributes less than 0.1% of the swap rate. Only with a few stocks do we find a sizable caused by large and frequent jumps. Table 8 shows positive jumps as more frequent but smaller than negative jumps. Compared to individual stocks, stock indices (as expected) have smaller jump sizes on average and are more likely to experience negative rather than positive jumps.

4.3 The Pricing of Jump Risk

From an economic standpoint, the presence of jumps leads to the question: are jump risks (partly) systematic and thus priced. In the classic framework of Merton (1976), jump risk is purely idiosyncratic and diversified away in large portfolios. Empirically, however, jumps in individual stock returns have a considerable systematic component as evidenced by intermittent, large one-day returns of large and seemingly well diversified

\footnote{Our findings are robust to choices of larger thresholds (more than 3 robust standard deviations) for jump identification.}
We investigate eight such indices, listed in Table 6(B)\textsuperscript{19}, which cover the main US stock market equity categories. For each index we identify jumps as one-day log-returns larger in magnitude than 3 robust standard deviations. Each month $t$ and each index $i$, we then record the contribution, $J_{i,t}$, to the monthly log $RV_{i,t}$ resultant from the identified jumps:

$$J_{i,t} \equiv \log RV_{i,t} - \log RV_{i,t}^C. \quad (21)$$

$$\log RV_{i,t} = \log \left[ \frac{252}{n_t} \sum_{d=1}^{n_t} r_{i,t,d}^2 \right].$$

Figure 6(A) displays the jump component of realized volatilities for the eight indices. The jump contribution is zero during most months and highly correlated across indices. To summarize aggregate jump risk, we construct an aggregate jump risk factor, the principal component of the eight time series of monthly jump contributions. Figure 6(B) displays the identified jump factor $F^J$. For each stock in our sample, we similarly construct a time series of monthly jump contributions using (21). The jump beta of each stock can then be calculated as the slope in a univariate regression on the aggregate jump risk factor.

The log SDF in a situation with aggregate jump risk must be appropriately augmented to include the jump risk factor:

$$m = \delta(\lambda_0 - \lambda_1 R_m - \lambda_2 \Delta F - \lambda_3 \Delta F^J). \quad (22)$$

The pricing restriction is then

$$E_t \left[ \log \left( \frac{RV_{t+1}}{SW_t} \right) \right] - r_{f,t} + \frac{1}{2} \text{var}_t[\log RV_{t+1}] = -\text{cov}_t \left[ m_{t+1}, \log RV_{t+1}^C \right] - \text{cov}_t \left[ m_{t+1}, J_{i,t} \right] = \beta' \lambda_t + \beta^J \lambda^J_t \quad (23)$$

Table 7 presents the asset pricing results after accounting for aggregate jump risk. The jump risk factor carries a small but highly significant negative risk premium. The jump risk premium remains significant even with the presence of the volatility factors, the or-

\textsuperscript{19}We do not include the S&P small cap index (SML) and PSE Wilshire Smallcap Index (WSX) since they are only introduced after 1994.
orthogonal market volatility factor and the Fama-French three factors. Figure 4 shows that, combining the jump risk factor with the three volatility factors and the five volatility factors, the in-sample explanatory Adjusted R-square of the pricing model at once goes up to 39% and 51.9%, respectively. The inclusion of the jump risk factor also helps to improve the out-of-the-sample performance in terms of yielding smaller RMSEs (as shown in Figure 5).

5  Economic Significance of the Mispricing on Variance Swap Contracts

Although variance swaps seem to be fairly priced on average according to our volatility factor model (the intercept terms are close to zero in the Fama-MacBeth regression as in Table 7), there are still considerable mispricings for variance swaps on certain individual stocks or indices as evident in Figure 4. For example, variance swaps on stock indices seem to be consistently overpriced relative to those on individual stocks. This section further investigates the economic significance of such mispricing by asking the question: can such mispricing be profitably exploited by option portfolio strategies even after accounting for transactions costs? As a results, this section also shreds some light on the effectiveness of replicating a variance swap using portfolios of calls and puts.

5.1  Exploiting mispricing in an optimization framework

The variance swap pricing equation implies:

\[ SW_{i,t} = E_t[M_{t+1}RV_{i,t+1}] \]
\[ = E_t[M_{t+1}E_t[RV_{i,t+1}] + \text{cov}_t(RV_{i,t+1}, M_{t+1})] \]
\[ = E_t[M_{t+1}]E_t[RV_{i,t+1}] + \exp(E_t[m_{t+1}])\text{cov}_t(RV_{i,t+1}, m_{t+1}) \]
\[ = E_t[M_{t+1}]\beta_{0,i} - \beta_i'\text{var}_t(F_{t+1})b_t \cdot \exp(E_t[m_{t+1}]) \]
where,

\[
\begin{align*}
\beta_{0,i} &= E_t[RV_{i,t+1}] \\
\beta_i &= \text{var}_t(F_{t+1})^{-1}\text{cov}_t(RV_{i,t+1}, F_{t+1}),
\end{align*}
\]

and \( F \) represents the vector of volatility factors (potentially also includes the aggregate jump risk factor).

If we construct two different portfolios of variance swaps with the same loadings \( \beta_0 \) and \( \beta \), then their initial cost \( SW \) should also be the same. Otherwise, there is mispricing which can be taken advantage of by proper long-short variance swap portfolio strategies.

Consider a mean-variance arbitrageur who wants to invest optimally in these variance swaps. Denote the position in option portfolio \( i \) using \( N_i \). Let \( N, SW, RV \) be \( n \times 1 \) vectors of positions, option-implied swap rate (\( SW_i \)), and actual one month realized volatility (\( RV_i \)), accordingly. Let \( \beta \) be the \( n \times (k+1) \) matrix of \( [\beta_0, \beta] \). Let \( \gamma \) be the risk aversion.

The payoff of such investment in one month (the option maturity) is:

\[
N'RV - N'SW.
\]

The arbitrageur’s optimization problem is:

\[
\max E [N'RV] - N'SW - \frac{1}{2} \gamma N'\Sigma N, \tag{24}
\]

subject to:

\[
\begin{align*}
N'\beta &= 0, \tag{25} \\
N'\iota &= 0, \tag{26}
\end{align*}
\]

where \( \Sigma \) denotes the var-cov matrix of \( RV \) and \( \iota \) is a vector of ones.

Constraints (25) ensure \( E [N'RV] = 0 \). Constraint (26) ensures the same number of variance swap contracts to long and to short. The objective function of (24) then changes
subject to the same constraints.

The optimal $N$ can be expressed in closed form:

$$N^* = \frac{1}{\gamma} V^{-1} \left[ B(B'V^{-1}B)B'V^{-1}SW - SW \right],$$

(28)

$$B = [\beta \ i].$$

In the expression of $N^*$, the risk aversion parameter $\gamma$ serves only as a scaling factor and should not affect the return calculation.

We can carry out such strategies using calls and puts as in (16). In practice, this option portfolio strategy is complicated considerably by various empirical considerations. First, in reality, options are only traded at discrete strikes. Denote the traded strikes as $K_i$, $i = 1, ..., N$ with $K_1 < \cdots < K_N$; in practice, the weight on option with a strike of $K_i$ is:

$$w(K_i) = \begin{cases} \frac{1}{(T-t)K_i^2}(K_2 - K_1) & \text{if } i = 1 \\ \frac{1}{(T-t)K_i^2}(K_{i+1} - K_{i-1}) & \text{if } 1 < i < N \\ \frac{1}{(T-t)K_i^2}(K_N - K_{N-1}) & \text{if } i = N \end{cases},$$

(29)

and the cost of the actual replication strategy is:

$$SW^*_i = \sum_{i=1}^{N} w(K_i) \left[ I_{(F \geq K_i)} P(K_i) + I_{(F < K_i)} C(K_i) \right].$$

(30)

$e^{r(T-t)}SW^*_i$ can be viewed as an approximation of $SW_i$. We define the relative approximation error of such replication strategy as $err = (SW_i - e^{r(T-t)}SW^*_i) / SW_i$. The actual payoff of the replicating portfolio at maturity is:
\[ X_i = \sum_{i=1}^{N} w(K_i) \left[ I_{(F \geq K_i)} \max(K_i - S_T, 0) + I_{(F < K_i)} \max(S_T - K_i, 0) \right]. \]

Second, one needs to post margins when shorting options. Specifically, the margin requirement for short naked call or put is\(^{20}\)

\[
\text{margin} = \begin{cases} 
C + \max(0.2 \times S - \max(K - S, 0), 0.1 \times S, 2.5) & \text{for Call} \\
P + \max(0.2 \times S - \max(S - K, 0), 0.1 \times K, 2.5) & \text{for Put.} 
\end{cases}
\]

Third, one incurs transaction costs when trading options. We capture the direct transaction cost by \(TC = (\text{Ask} - \text{bid})/2\)\(^{21}\). Assume he will lend or borrow at a risk free rate at \(t = 0\) so the initial net cash flow is zero. Let \(R_f\) be the gross risk free rate (\(R_f = e^{r(T-t)}\)). The excess return on the quasi-arbitrage is therefore:

\[
R^e = \frac{N'X - N'SW_A \cdot R_f - |N'| TC \cdot R_f + \max(-N', 0) \cdot \text{margin} \cdot R_f}{\max(-N', 0) \cdot \text{margin}} - R_f, \tag{31}
\]

where \(X\) denotes the payoff of the options at maturity.

### 5.2 Empirical results

From 1997 to 2004, on the Monday after the third Friday in each month, we keep those options that mature in the next month with at least 2 OTM calls and 2 OTM puts and positive trading volume. To ensure diversification in the cross-section, we only keep a cross-section if the number of eligible underlying stocks or indices exceed 70. Based on these option implied volatilities, we compute the implied variance swap rate \(SW_i\) using\(^{16}\) and the associated replicating portfolio weights \(w(K_i)\) using\(^{29}\).

The option portfolio is constructed as follows. Each month, we first compute \(\beta\) for

\(^{20}\)This is a conservative assumption on margin. In practice, the margin requirement is smaller when you take positions on a portfolio of options.

\(^{21}\)Since we hold each option to maturity, we only suffer half of the bid-ask spread at portfolio construction.
each stock by regressing $RV$ on $F$ using the usual five-year rolling window. We then compute the optimal $N_t$ using (28) and the monthly excess return (in excess of risk-free rate) using (31). These excess returns measure the extent to which relative mispricing of variance swaps can be profitably exploited even after accounting for direct transaction costs. Since the computation of optimal $N_t$ requires no future information, the strategy is tradable in practice. Figure 7 plots these excess returns over time from 1997 to Dec 2004. The mispricing turns out to be significant economically. The average excess return stands at 2.65% per month after direct transaction cost and margin requirement, with an annual Sharpe ratio of 1.55. This compares favorably against the return of the market during the same time period, which is 0.26% per month with an annual Sharpe ratio of 0.18. As expected, the excess return does not load on the market return factor. The CAPM beta is 0.21 with a $t$-value of 1.61 and the CAPM risk-adjusted excess return is 2.60% with a $t$-value of 4.61. To analyze the impact of the approximation error, we sort the replicated variance swaps in our sample according to $err$ to five portfolios each month. We then report various characteristics for each portfolio as (see Table 9). In general, $SW^A$ underestimates $SW$, resulting in a positive approximation error – $err$. $err$ is smaller (in absolute term) when the number of contracts are large. In addition, there does not seem to be a systematic relationship between $err$ and return. To conclude, although variance swaps seem to be fairly priced on average, the relative mispricing in a cross section still exist and the magnitude of such mispricing is significant both statistically and economically.

6 Conclusion

We have demonstrated that the realized volatilities of US equities can be represented using a simple factor model. The first three principal factors account for about 55% of the cross-sectional variation the realized monthly volatilities. The estimated factor structure also largely captures the forecastable component of realized volatilities. After controlling for the first 4 to 5 principal factors, the innovations to the realized volatilities become serially uncorrelated.

In Merton’s ICAPM, the factor structure on volatilities implies that the aggregate volatilities factors are natural state variables and are likely to be priced. We test this implication using the cross-section of variance swaps replicated using standard calls and puts. Consistent with previous literature, we find that the first principal factor, one highly cor-
related with the volatility on a market index, carries a significant negative risk premium. Compared to the volatility on a market index, the first principal factor has a slight advantage in capturing the cross-sectional variations in returns on variance swaps. We find the third principal factor also priced. A multiple volatility factor model improves the cross-sectional fit significantly. Moreover, we find strong evidence of the market pricing an aggregate jump risk factor. Interestingly, index options are found to be more expensive relative to individual stock options after accounting for the risk due to the first five volatility factors. This is consistent with previous findings by (Bakshi and Kapadia (2003b)) and others.

The factor structure on volatilities also imposes a restriction on the price of an option portfolio: any well diversified variance-swap-replicating option portfolio constructed to have zero volatility factor loadings should have approximate zero cost. We construct a buy-and-hold option trading strategy to exploit empirical violations of such restriction. This strategy produces a CAPM risk adjusted return of 2.60% per month with an annual Sharpe ratio of around 1.55, even after accounting for direct transaction cost.
References


Table 1: Results of the Asymptotic Principal Component Analysis (APCA)

From 1965 to 2004, we run APCA in each of the eight five-year windows. We report: the number of stocks in each period; the optimal number of factors according to Bai & Ng (2003) criteria; the cumulative percentage variance explained using one to five factors; and the correlation between the first principal factor and the market volatility.

<table>
<thead>
<tr>
<th>Period</th>
<th># of stocks</th>
<th>Optimal # of factors (IC1)</th>
<th>Optimal # of factors (IC2)</th>
<th>% of var explained - 1 factor</th>
<th>% of var explained - 2 factors</th>
<th>% of var explained - 3 factors</th>
<th>% of var explained - 4 factors</th>
<th>% of var explained - 5 factors</th>
<th>Correlation between 1st factor and mkt vol</th>
</tr>
</thead>
<tbody>
<tr>
<td>65-69</td>
<td>1061</td>
<td>3</td>
<td>3</td>
<td>20.3%</td>
<td>30.5%</td>
<td>36.2%</td>
<td>39.4%</td>
<td>42.0%</td>
<td>0.722</td>
</tr>
<tr>
<td>70-74</td>
<td>961</td>
<td>3</td>
<td>3</td>
<td>40.9%</td>
<td>48.8%</td>
<td>53.0%</td>
<td>55.4%</td>
<td>57.4%</td>
<td>0.934</td>
</tr>
<tr>
<td>75-79</td>
<td>1126</td>
<td>3</td>
<td>2</td>
<td>35.2%</td>
<td>42.3%</td>
<td>46.3%</td>
<td>48.9%</td>
<td>50.9%</td>
<td>0.898</td>
</tr>
<tr>
<td>80-84</td>
<td>1205</td>
<td>3</td>
<td>3</td>
<td>23.1%</td>
<td>30.5%</td>
<td>35.9%</td>
<td>39.8%</td>
<td>42.5%</td>
<td>0.798</td>
</tr>
<tr>
<td>85-89</td>
<td>1631</td>
<td>2</td>
<td>2</td>
<td>28.4%</td>
<td>38.7%</td>
<td>42.6%</td>
<td>46.5%</td>
<td>49.2%</td>
<td>0.907</td>
</tr>
<tr>
<td>90-94</td>
<td>1871</td>
<td>3</td>
<td>3</td>
<td>22.1%</td>
<td>30.4%</td>
<td>36.8%</td>
<td>41.1%</td>
<td>44.4%</td>
<td>0.756</td>
</tr>
<tr>
<td>95-99</td>
<td>2556</td>
<td>4</td>
<td>4</td>
<td>30.0%</td>
<td>37.7%</td>
<td>43.1%</td>
<td>47.3%</td>
<td>50.3%</td>
<td>0.850</td>
</tr>
<tr>
<td>00-04</td>
<td>2520</td>
<td>3</td>
<td>3</td>
<td>49.4%</td>
<td>55.0%</td>
<td>59.3%</td>
<td>61.9%</td>
<td>63.8%</td>
<td>0.730</td>
</tr>
</tbody>
</table>
Table 2: Information criteria for three alternative specifications of the five-factor dynamics

The Corrected Akaike Information Criterion (AICC), the Akaike Information Criterion (AIC) and the Schwartz Information Criterion (SBC). All three criteria select the VAR(1) specification.

<table>
<thead>
<tr>
<th>Information Criteria (2000-2004)</th>
<th>VMA(1)</th>
<th>VAR(1)</th>
<th>VARMA(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AICC</strong></td>
<td>-3.89</td>
<td>-8.94</td>
<td>-8.72</td>
</tr>
<tr>
<td><strong>AIC</strong></td>
<td>-3.93</td>
<td>-8.98</td>
<td>-8.89</td>
</tr>
<tr>
<td><strong>SBC</strong></td>
<td>-3.36</td>
<td>-8.41</td>
<td>-7.77</td>
</tr>
</tbody>
</table>

Table 3: VAR(1) coefficient estimates for the five factor model

Bold numbers are significant at the 5% level.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor</strong></td>
<td>F1</td>
<td>F2</td>
<td>F3</td>
<td>F4</td>
<td>F5</td>
</tr>
<tr>
<td>F1</td>
<td>0.89631</td>
<td>0.04992</td>
<td><strong>-0.12537</strong></td>
<td>-0.10185</td>
<td>-0.10223</td>
</tr>
<tr>
<td>F2</td>
<td>-0.00886</td>
<td>0.95104</td>
<td>0.09178</td>
<td>-0.03866</td>
<td>0.03457</td>
</tr>
<tr>
<td>F3</td>
<td><strong>-0.17818</strong></td>
<td>-0.03071</td>
<td><strong>0.76495</strong></td>
<td><strong>-0.16380</strong></td>
<td><strong>-0.19975</strong></td>
</tr>
<tr>
<td>F4</td>
<td>-0.14403</td>
<td>0.10028</td>
<td><strong>-0.20260</strong></td>
<td><strong>0.78863</strong></td>
<td>-0.13743</td>
</tr>
<tr>
<td>F5</td>
<td>0.06601</td>
<td>-0.08527</td>
<td>-0.04561</td>
<td><strong>-0.18289</strong></td>
<td>0.83908</td>
</tr>
</tbody>
</table>
Table 4: Estimated residual cross-correlations

The row labeled “F1” shows the cross-correlation between the innovations to the first factor and the innovations to the remaining factors at various lags. “+” indicates a significant positive correlation; “-” a significant negative correlation; “?” an insignificant correlation. For instance, the innovation to the first factor has a positive autocorrelation at lag 3 but no significant correlation with the 3 month lagged innovations to the remaining factors.

<table>
<thead>
<tr>
<th>Factor/Lag</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>+++++</td>
<td>....</td>
<td>....</td>
<td>++..</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>F2</td>
<td>-+.</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>F3</td>
<td>+++++</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>F4</td>
<td>+++.</td>
<td>....</td>
<td>....</td>
<td>+++.</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>F5</td>
<td>+++.</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>

Table 5: Univariate model white noise diagnostics for the factor innovations

The Durbin-Watson test statistic value of 2 indicates that there is no serial autocorrelation. The Chi square test detects too many large squared residuals to be consistent with a normal distribution of the residuals, except for the second factor equation. The ARCH test fails to find serial correlation in the squared residuals.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Chi-Square</td>
<td>Pr &gt; ChiSq</td>
</tr>
<tr>
<td>F1</td>
<td>2.2005</td>
<td>10.69</td>
<td>0.0048</td>
</tr>
<tr>
<td>F2</td>
<td>2.2135</td>
<td>0.43</td>
<td>0.8055</td>
</tr>
<tr>
<td>F3</td>
<td>2.0470</td>
<td>26.94</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>F4</td>
<td>1.9048</td>
<td>28.57</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>F5</td>
<td>1.7620</td>
<td>10.08</td>
<td>0.0065</td>
</tr>
</tbody>
</table>
Table 6: Option sample description

Panel A lists the average number of stocks and indices in our sample, and the average number of strikes associated with each stock/index. Panel B lists the 10 stock indices in the sample taken from the list of major indices in OptionMetrics.

Panel A: Sample description

<table>
<thead>
<tr>
<th>year</th>
<th>Individual Stock</th>
<th>Stock Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg # of obs per month</td>
<td>Avg # of obs per month</td>
</tr>
<tr>
<td>1997</td>
<td>90.3</td>
<td>5.6</td>
</tr>
<tr>
<td>1998</td>
<td>111.6</td>
<td>5.7</td>
</tr>
<tr>
<td>1999</td>
<td>201.4</td>
<td>7.3</td>
</tr>
<tr>
<td>2000</td>
<td>342.3</td>
<td>8.8</td>
</tr>
<tr>
<td>2001</td>
<td>242.9</td>
<td>6.3</td>
</tr>
<tr>
<td>2002</td>
<td>209.4</td>
<td>5.7</td>
</tr>
<tr>
<td>2003</td>
<td>192.8</td>
<td>5.5</td>
</tr>
<tr>
<td>2004</td>
<td>207.5</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Panel B: Stock index options included

<table>
<thead>
<tr>
<th>Index Name</th>
<th>Ticker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dow Jones Industrial Average</td>
<td>DJX</td>
</tr>
<tr>
<td>NASDAQ 100 Index</td>
<td>NDX</td>
</tr>
<tr>
<td>CBOE Mini</td>
<td>MNX</td>
</tr>
<tr>
<td>AMEX Major Market Index</td>
<td>XMI</td>
</tr>
<tr>
<td>S&amp;P 500 Index</td>
<td>SPX</td>
</tr>
<tr>
<td>S&amp;P 100 Index</td>
<td>OEX</td>
</tr>
<tr>
<td>S&amp;P Midcap 400 Index</td>
<td>MID</td>
</tr>
<tr>
<td>S&amp;P Smallcap 600 Index</td>
<td>SML</td>
</tr>
<tr>
<td>Russell 2000 Index</td>
<td>RUT</td>
</tr>
<tr>
<td>PSE Wilshire Smallcap Index</td>
<td>WSX</td>
</tr>
</tbody>
</table>
Table 7: Cross-sectional regression results

$F_1$ to $F_5$ are the first five volatility factors extracted using Asymptotic Principal Component Analysis (APCA). $F_J$ is the jump factor. $V_{m,orth}$ is the component of the market volatility that is orthogonal to $F_1$ to $F_5$. MKT, SMB and HML are the three Fama-French factors. The cross-sectional regressions are estimated from 1997 to 2004. We report the Newey-West t-values (FMNW-t) which account for the autocorrelation of the estimates with a lag of 12.

Panel A: Volatility-Factor Models

<table>
<thead>
<tr>
<th></th>
<th>intercept</th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>$F_3$</th>
<th>$F_4$</th>
<th>$F_5$</th>
<th>$F_J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM Coef</td>
<td>-0.0007</td>
<td>-0.2818</td>
<td>-0.0252</td>
<td>0.2523</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>-0.01</td>
<td>-2.70</td>
<td>-0.34</td>
<td>2.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM Coef</td>
<td>0.0286</td>
<td>-0.2064</td>
<td>-0.0167</td>
<td>0.2198</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>0.48</td>
<td>-2.16</td>
<td>-0.26</td>
<td>3.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM Coef</td>
<td>0.0120</td>
<td>-0.3084</td>
<td>-0.0338</td>
<td>0.2857</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>0.42</td>
<td>-2.28</td>
<td>-0.30</td>
<td>3.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM Coef</td>
<td>0.0336</td>
<td>-0.2425</td>
<td>-0.0242</td>
<td>0.2489</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>0.61</td>
<td>-1.85</td>
<td>-0.29</td>
<td>2.89</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: With market volatility

<table>
<thead>
<tr>
<th></th>
<th>intercept</th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>$F_3$</th>
<th>$F_4$</th>
<th>$F_5$</th>
<th>$F_J$</th>
<th>$V_{m,orth}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM Coef</td>
<td>-0.0076</td>
<td>-0.3894</td>
<td>-0.0308</td>
<td>0.2045</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>-0.13</td>
<td>-3.47</td>
<td>-0.43</td>
<td>2.53</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM Coef</td>
<td>0.0234</td>
<td>-0.2908</td>
<td>-0.0242</td>
<td>0.1845</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>0.40</td>
<td>-2.90</td>
<td>-0.38</td>
<td>2.66</td>
<td>-5.67</td>
<td>1.48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: With return factors

<table>
<thead>
<tr>
<th></th>
<th>intercept</th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>$F_3$</th>
<th>$F_4$</th>
<th>$F_5$</th>
<th>$F_J$</th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM Coef</td>
<td>-0.0329</td>
<td>-0.4322</td>
<td>-0.1000</td>
<td>0.2266</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>-0.59</td>
<td>-3.34</td>
<td>-1.00</td>
<td>3.79</td>
<td>-1.87</td>
<td>0.30</td>
<td>1.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM Coef</td>
<td>-0.0080</td>
<td>-0.3495</td>
<td>-0.0784</td>
<td>0.2096</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMNW-t</td>
<td>-0.14</td>
<td>-3.03</td>
<td>-0.98</td>
<td>3.54</td>
<td>-3.95</td>
<td>-1.58</td>
<td>-0.04</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8: Descriptive statistics for jumps in the log price processes for individual stocks and indices from 1997 to 2004

We report the mean, first quartile (Q1), median and third quartile (Q3) for the jump correction term ($\varepsilon$), the swap rate (SW), the percentage jump correction term (jump/SW), negative jump intensity per year ($\lambda^-$), average magnitude of negative jumps ($\mu^-$), positive jump intensity per year ($\lambda^+$) and average magnitude of positive jumps ($\mu^+$).

<table>
<thead>
<tr>
<th></th>
<th>Jump ($\varepsilon$) (%)</th>
<th>SW (%)</th>
<th>Jump/SW (%)</th>
<th>$\lambda^-$</th>
<th>$\mu^-$</th>
<th>$\lambda^+$</th>
<th>$\mu^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Stock Options (#obs = 16402)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0250</td>
<td>35.7591</td>
<td>0.0967</td>
<td>2.2906</td>
<td>0.0350</td>
<td>3.2128</td>
<td>0.0254</td>
</tr>
<tr>
<td>Q1</td>
<td>-0.0030</td>
<td>12.1172</td>
<td>-0.0143</td>
<td>1.5975</td>
<td>0.0168</td>
<td>2.1948</td>
<td>0.0146</td>
</tr>
<tr>
<td>Median</td>
<td>0.0008</td>
<td>23.4793</td>
<td>0.0061</td>
<td>2.2053</td>
<td>0.0283</td>
<td>3.0049</td>
<td>0.0221</td>
</tr>
<tr>
<td>Q3</td>
<td>0.0196</td>
<td>47.9410</td>
<td>0.0717</td>
<td>2.8636</td>
<td>0.0451</td>
<td>4.0127</td>
<td>0.0327</td>
</tr>
<tr>
<td>Individual Stock Options (#obs = 15887)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0258</td>
<td>36.6572</td>
<td>0.0997</td>
<td>2.2954</td>
<td>0.0358</td>
<td>3.2570</td>
<td>0.0259</td>
</tr>
<tr>
<td>Q1</td>
<td>-0.0033</td>
<td>12.7753</td>
<td>-0.0151</td>
<td>1.5987</td>
<td>0.0177</td>
<td>2.2053</td>
<td>0.0151</td>
</tr>
<tr>
<td>Median</td>
<td>0.0011</td>
<td>24.3589</td>
<td>0.0066</td>
<td>2.2053</td>
<td>0.0292</td>
<td>3.0096</td>
<td>0.0226</td>
</tr>
<tr>
<td>Q3</td>
<td>0.0211</td>
<td>48.9772</td>
<td>0.0767</td>
<td>2.9013</td>
<td>0.0461</td>
<td>4.0727</td>
<td>0.0333</td>
</tr>
<tr>
<td>Stock Index options (#obs = 515)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0000</td>
<td>8.0562</td>
<td>0.0016</td>
<td>2.1416</td>
<td>0.0098</td>
<td>1.8498</td>
<td>0.0090</td>
</tr>
<tr>
<td>Q1</td>
<td>0.0000</td>
<td>3.7347</td>
<td>0.0006</td>
<td>1.3978</td>
<td>0.0074</td>
<td>1.4045</td>
<td>0.0054</td>
</tr>
<tr>
<td>Median</td>
<td>0.0002</td>
<td>5.1595</td>
<td>0.0035</td>
<td>2.2070</td>
<td>0.0104</td>
<td>1.7957</td>
<td>0.0077</td>
</tr>
<tr>
<td>Q3</td>
<td>0.0005</td>
<td>8.4334</td>
<td>0.0088</td>
<td>2.6062</td>
<td>0.0122</td>
<td>2.2053</td>
<td>0.0097</td>
</tr>
</tbody>
</table>
Table 9: Variance swap rate approximation error and returns

For each underlying stock, we can compute an implied variance swap rate (SW) using interpolation and numerical integration on volatility smile. We can compute an approximate variance swap rate (SW$_{approx}$) as the cost of constructing a replication portfolio using the available OTM calls and puts. We define the variable $diff = (SW - e^{(T-t)SW_{approx}})/SW$ as a measure of the replicating strategy’s approximation error. Each month, when the size of the cross-section exceeds 70, we sort all variance swaps according to the approximation error variable - diff into 5 portfolios, then compute the average characteristics for each portfolio. % of zero payoff is defined as the percentage of options that expired out-of-money.

<table>
<thead>
<tr>
<th>portid</th>
<th>diff</th>
<th>return</th>
<th>% of zero payoff</th>
<th>N_contract</th>
<th>sw</th>
<th>sw$_{approx}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.5%</td>
<td>-0.0923</td>
<td>37.6%</td>
<td>5.12</td>
<td>0.262</td>
<td>0.235</td>
</tr>
<tr>
<td>2</td>
<td>6.7%</td>
<td>-0.1324</td>
<td>26.3%</td>
<td>6.18</td>
<td>0.324</td>
<td>0.304</td>
</tr>
<tr>
<td>3</td>
<td>4.3%</td>
<td>-0.1426</td>
<td>23.3%</td>
<td>7.00</td>
<td>0.377</td>
<td>0.362</td>
</tr>
<tr>
<td>4</td>
<td>1.9%</td>
<td>-0.0391</td>
<td>23.4%</td>
<td>7.06</td>
<td>0.392</td>
<td>0.385</td>
</tr>
<tr>
<td>5</td>
<td>-2.7%</td>
<td>-0.0628</td>
<td>25.7%</td>
<td>5.51</td>
<td>0.333</td>
<td>0.340</td>
</tr>
</tbody>
</table>
Figure 1: Histogram of autocorrelations and contemporaneous correlations of volatility factor model innovations from 2000-2004)

Panel A plots the autocorrelations of order 1 thru 4 for the residuals of the 3-factor factor model for each individual stock in the cross-section. The vertical bars denote the 95% confidence interval. Panel B and C plot similar autocorrelations for residuals of the 4- and 5-factor model, respectively. Panel D plots the contemporaneous correlations for the residuals of the 5-factor model for each individual stock in the cross-section.

Panel A

![Autocorrelation plots for 3-factor model](image-url)
Panel B

Figure 1 Panel B: The 4 factor model, 2000-2004. 
Autocorrelations of order 1 thru 4 for the residuals of the factor model 
for each individual stock in the cross-section. The vertical bars denote 
the 95% confidence interval.

Panel C

Figure 1 Panel C: The 5 factor model, 2000-2004. 
Autocorrelations of order 1 thru 4 for the residuals of the factor model 
for each individual stock in the cross-section. The vertical bars denote 
the 95% confidence interval.
Panel D

Figure 1 Panel D: The 5 factor model, 2000-2004. Contemporaneous correlations for the residuals of the factor model for each individual stock in the cross-section.
Figure 2: Results of VAR estimation

In Panel A, the left column displays the estimated first three factors, the fitted VAR(1) series and corresponding 95% confidence bands. The right column shows the VAR(1) residuals along one and two standard deviation bands. In Panel B, the left column displays the estimated fourth and fifth factors, the fitted VAR(1) series and corresponding 95% confidence bands. The right column shows the VAR(1) residuals along one and two standard deviation bands.

Panel A
Panel B

Figure 2: The left column displays the estimated fourth and fifth factors, the fitted VAR(1) series and corresponding 95% confidence bands. The right column shows the VAR(1) residuals along with one and two standard deviation bands.
Figure 3: Autocorrelation functions and partial autocorrelation functions for the 5-factor VAR(1) residuals.
Figure 4: Average predicted and actual excess returns on variance swaps across stocks

From 1997 to 2004, for each stock and index with more than 35 observations, we compute the time-series averages of the actual excess returns on its variance swap. We then run a cross sectional regression of these excess returns on their volatility factor betas and the jump factor beta. We plot the factor-model-predicted excess returns against the actual excess returns on variance swaps separately for the market volatility, one-factor, the three-factor, the five-factor, the three-factor with jump and the five-factor with jump volatility models. The adjusted R-squares are also reported. The stock indices are represented using (*) and the individual stocks are represented using (.). There are 128 observations in total.
Figure 5: Average predicted and actual excess returns on variance swaps across stocks out of sample

For each stock and index with 6 complete monthly observations in the first half of 2005, we compute the time-series averages of the actual excess returns on its variance swap. We then compute the factor-model-predicted excess returns using volatility factor betas and the jump factor beta estimated at the end of 2004. We plot the predicted excess returns against the actual excess returns on variance swaps separately for the market volatility, one-factor, the three-factor, the five-factor, the three-factor with jump and the five-factor with jump volatility models. The adjusted R-squares are also reported. The stock indices are represented using (*) and the individual stocks are represented using (.). There are 69 observations in total.
Panel A plots the (demeaned) jump components in volatility for 8 stock indices. The descriptions of these indices can be found in Table 6B. We then extract the aggregate jump factor as the first principal component and plot it in Panel B.
Figure 7: Monthly excess return on the volatility arbitrage portfolio and on the market portfolio

From 1997 to 2004, on the Monday after the third Friday of each month, we keep those options that mature in the next month with at least 2 OTM calls and 2 OTM puts and positive trading volume. We only keep a cross-section of more than 70 eligible underlying stocks. Based on these option-implied volatilities, we computed the implied variance swap rate and the replicating option portfolios for each stock. In a constrained portfolio optimization framework, we compute the optimal weights on each replicating portfolios. The monthly arbitrage portfolio returns are then computed using these optimal weights and option payoffs at maturity, after accounting for margin requirements and bid-ask spread. The excess returns (in excess of risk free rate) on the arbitrage portfolio and the market portfolio are plotted.