Efficient Frontier Bounds Under Stochastic Covariances

Alexander Philipov *

Boston College

December 17, 2003

Abstract

This paper examines the implications of multivariate stochastic volatility on efficient portfolio choice. Out-of-sample forecasts and confidence bounds are generated for the efficient frontier, optimal weights, and expected Sharpe ratios, using stochastic time-varying covariance matrix estimates. Covariance forecasts are generated using a multivariate stochastic volatility (SVOL) model introduced by Philipov and Glickman (2002). This new model provides increased flexibility, avoids common estimation constraints, and is feasible in unconstrained form for higher dimensions. Model parameter estimates show time variation both in the stochastic covariances and in the correlations implied from them. Correlations and volatilities move synchronously through time suggesting the effect of common market forces. Optimal weight estimates, however, exhibit limited time variation which can be attributed to the synchronous behavior of volatilities and correlations. Confidence bounds on the stochastic efficient frontiers derived from the multivariate SVOL forecasts appear better behaved than previously derived analytical bounds. Portfolios based on SVOL covariance estimates outperform portfolios using alternative covariance models. Out-of-sample portfolio results are compared using the new analysis tools offered by the SVOL model.

*I gratefully acknowledge the benefit to this paper from comments and suggestions by Eric Jacquier, Alan Marcus, Pierluigi Balduzzi, Wayne Ferson, Cesare Robotti, and seminar participants at Boston College. For comments, suggestions, or further information, please email me at `philipov@bc.edu` or visit *http://www2.bc.edu/~philipov*

1 Introduction

Multivariate stochastic volatility models have emerged as a promising response to the need in portfolio management for accurate time-varying covariance matrix estimates. This paper examines the implications of time-varying stochastic covariances for Markowitz portfolio efficiency. Optimal portfolio weights are obtained using covariance forecasts generated by implementing a multivariate stochastic volatility (SVOL) model developed by Philipov and Glickman (2002). In this model, covariance matrices vary through time driven by Wishart processes. Model parameters are estimated using Bayesian techniques. These estimation methods generate abundant information used here to establish confidence regions for the mean-variance set, as well as to improve and evaluate portoflio performance.

The literature on Markowitz portfolio optimization celebrates an over-forty-year history. It has been readily embraced by academics as the theoretically sound approach to managing financial portfolios, but only slowly and cautiously adopted by practitioners due to a number of reasons. First, mean-variance (MV) optimization tends to maximize the effects of errors in the inputs (Michaud (1989)). Michaud points out that the MV optimizer significantly overweights securities with large estimated returns, negative correlations and small variances. Second, MV optimization is highly sensitive to changes in the mean forecasts (Best and Grauer (1991), Chopra and Ziemba (1993)), and, to a lesser but significant degree, to changes in covariances (Chopra and Ziemba (1993), Pojarliev and Polasek (2001)). Furthemore, the relative importance of mean or variance misspecifications depends on the investor's risk tolerance. Third, MV optimization commonly uses point estimates of means and variances as inputs and produces point estimates of optimal portfolio weights as output. Due to the high sensitivity of the process these point estimates may be unreliable.

Research on portfolio optimization has gone a long way to address these issues. On a practical level, MV optimization has been "tamed" by using constraints and carefully selected benchmarks. Tracking error minimization within narrow bounds of a benchmark is a widespread approach in institutional investing. In a critique of this approach, Roll (1992) shows that managers pursuing tracking error optimization will intentionally fail to produce MV efficient portfolios. Recent studies on the merits of the tracking error optimization approach include Jorion (2002a), Jorion (2002b), Rudolf, Wolter, and Zimmermann (1999), Ammann and Zimmermann (2001) etc. Other ways to reduce MV sensitivity are using Bayes-Stein shrinkage estimators (Jorion (1986)), reducing dimensionality through the application of factor models (Jorion (1991), Jacquier and Marcus (2001)), or using a newly proposed resampled efficiency approach (Michaud (1998), Fletcher and Hillier (2001)). Jagannathan and Ma (2002) re-examine constrained optimization and relate no-short-sale constraints to shrinkages and factor covariance estimators.

Another area of research has taken parameter uncertainty as given and has proposed adjustments to the MV weights that incorporate estimation risk. For example, Barry (1974) and Chen and Brown (1983) evaluate estimation risk in portfolio choice using a Bayesian framework, while Balduzzi and Liu (2001) and Horst, de Roon, and Werker (2002) take a classical utility-based approach in which parameter uncertainty increases utility costs and translates into an adjusted risk aversion parameter in the optimization setup. These approaches are concerned mostly with generating improved estimates of the MV efficient set rather than examining the informational content of the point estimates produced by the MV optimizer.

This paper claims closest association with the volume of research on the informational content of MV optimization, which aims to provide significance intervals and sampling error estimates of optimal portfolio weights. For example, Jobson and Korkie (1989) offer a review of multivariate significance tests of MV efficiency. Britten-Jones (1999) proposes an exact procedure for testing hypotheses about optimal weights. Jobson (1991) constructs analytical confidence regions for the MV efficient set and formulates an F-test to generate sample acceptance regions. Kandel and Stambaugh (1989) link tests of MV efficiency with asset pricing tests. These analytical results, however, apply only to the case of unconstrained optimization. Unconstrained optimization is almost never used by financial managers because it can produce quite unrealistic short sale situations. In the case of constrained optimization, researchers have employed Monte Carlo simulation methods or bootstrapping techniques to infer the sample properties of optimal portfolio weights. The current study makes a contribution to these sampling approaches by introducing conditioning information and stochastic time variation in the covariance parameters.

This paper studies the sample properties of optimal weights with emphasis on constructing confidence regions and testing hypotheses. The sample properties of the optimal weights are related to asymptotic results derived by Jobson and Korkie (1980, 1989) and Jobson (1991) as well as to simulation results based on distributional properties of the returns derived from the data sample. In the current study, model forecasts from the PG SVOL model are used in a dynamic constrained MV optimization to obtain stochastic efficient frontiers. It uses unconditional mean return estimates based on the data sample. Using fixed point estimates of mean returns without conditioning information helps isolate the effect of stochastic covariances. These covariance effects are especially valid for the global minimum variance portfolio whose construction does not rely on mean estimates.

Multivariate models of time varying volatility have two notable drawbacks that have prevented them from enjoying the success of their univariate counterparts. First, multivariate models are extremely unparsimonious. This problem grows exponentially with model dimensionality. Second, multivariate models bring a significant increase in the complexity of the estimation of their parameters which is in large part due to the lack of parsimony. Hence multivariate models are either very tightly constrained or are unmanageable due to the huge number of parameters. For the multivariate SVOL model of Philipov and Glickman (2002), however, the new model formulation and the Bayesian methodology implemented for estimating model parameters manage to avoid common constraints placed solely for estimation purposes (e.g. diagonal covariance parameters or constant correlation structure), making the model feasible for higher dimensions (e.g. a 12 variate model with 740 time observations). In addition, the model provides improved flexibility and a seamless link to univariate space: the PG modeling framework extends naturally from scalar variances to matrix-covariances without changing the general form of the model.

Traditional models of multivariate SVOL have been studied by Harvey, Ruiz, and Shephard (1994), Mahieu and Schotman (1994), Jacquier, Polson, and Rossi (1995), and Shephard (1996). These traditional models have arisen as extensions of successful univariate models, specifying separate log-normal autoregressive processes for the variances. The variance processes in these models are usually tied together by a constant correlations structure, creating models of changing variances rather than changing correlation. In this respect, even though traditional multivariate models allow common trends and cycles in volatility, they do not allow covariances to evolve over time independently of variances. The model implemented in this paper can provide greater flexibility in comparison to traditional multivariate SVOL models, as it offers an unrestricted representation of changing variances, covariances or correlations.

The next section lays out the formulation of the Philipov and Glickman (2002) multivariate SVOL model. Section 3 provides an overview of the analytics of MV optimization and of the efficient frontier confidence regions derived in Jobson (1991). The methodology for estimating the multivariate SVOL model, obtaining the covariance forecasts, and computing the MV efficient sets based on them, is described in section 4, which is followed in sections 5 and 6 by description and analysis of the data and the results.

2 The Covariance Model

To estimate the covariance matrix parameter in the MV optimization, we use a new general model for multivariate stochastic volatilities developed by Philipov and Glickman (2002). This general multivariate model is naturally linked to univariate space in which a special case of the model has a close connection with traditional SVOL models. The general model describes the evolution through time of a collection of k correlated normally distributed asset returns. The covariance structure of this portfolio is also dynamic and is determined by a stochastic process based on the Wishart distribution:

$$\begin{aligned} \mathbf{y}_t \quad | \quad & \mathbf{\Sigma}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_t) \\ \mathbf{\Sigma}_t^{-1} \quad | \quad & \nu, \mathbf{S}_{t-1} \sim Wishart_k(\nu, \mathbf{S}_{t-1}) \end{aligned} \tag{1}$$

where ν and $\mathbf{S}_t$ are the degrees of freedom and the scale parameter of the Wishart distribution. The vector $\mathbf{y}_t$ represents "surprises" over the expected returns for the period, which are represented by returns series pre-whitened via an AR(1) filter (Jacquier, Polson, and Rossi (1994)). Such a formulation, avoiding the estimation of a mean parameter within the model and directing focus solely on volatility, is prevalent in the literature on stochastic volatility (See, for example: Jacquier, Polson, and Rossi (1999, 1994) Harvey, Ruiz, and Shephard (1994), Jiang and van der Sluis (2000), Pitt and Shephard (1999)). It is believed that there is little interaction between the estimation of the mean and the variance term, with the exception of variance-in-the-mean terms in GARCH-M type models. The role of the mean parameter diminishes with increase in data frequency and is non-existent in the limiting continuous time case.

With a time-invariant covariance structure, the above model offers a traditional representation of the behavior of multivariate normal returns. However, we augment this setup by allowing time variation in the scale parameter of the Wishart distribution. Let $\mathbf{A}$ be a positive definite parameter matrix that can be decomposed as $\mathbf{A} = \left(\mathbf{A}^{\frac{1}{2}}\right)\left(\mathbf{A}^{\frac{1}{2}}\right)'$, and d be a scalar parameter. We define the scale parameter of the Wishart distribution in period t as a function of the covariance matrix in that period:

$$\mathbf{S}_t = \frac{1}{\nu}\left(\mathbf{A}^{\frac{1}{2}}\right)\left(\mathbf{\Sigma}_t^{-1}\right)^d\left(\mathbf{A}^{\frac{1}{2}}\right)' \tag{2}$$

The quadratic expression for $\mathbf{S}_t$ ensures that it is symmetric positive definite. Time-variation of the covariances is determined by this quadratic expression which, using the properties of the Wishart distribution (see Gelman et al, 1995), leads to the following conditional expectation of $\mathbf{\Sigma}_t^{-1}$:

$$E(\mathbf{\Sigma}_t^{-1}) = v\,\mathbf{S}_{t-1} = \left(\mathbf{A}^{\frac{1}{2}}\right)\,\left(\mathbf{\Sigma}_{t-1}^{-1}\right)^d\left(\mathbf{A}^{\frac{1}{2}}\right)' \tag{3}$$

We specify the model in terms of the inverse covariance matrix to facilitate deriving the conditional posterior distributions of the parameter estimates in the Gibbs sampler. Covariance matrices themselves follow the inverse-Wishart distribution. We can use the properties of the inverse-Wishart distribution to obtain the conditional expectation of the covariance matrix at time t:

$$E(\mathbf{\Sigma}_t) = (v - k - 1)^{-1}\, \mathbf{S}_{t-1}^{-1} = \frac{\nu}{v - k - 1} \left(\mathbf{A}^{-\frac{1}{2}}\right) (\mathbf{\Sigma}_{t-1})^{d} \left(\mathbf{A}^{-\frac{1}{2}}\right)' \tag{4}$$

where the parameters ν, d, and $\mathbf{A}$ are defined above. The parameters $\mathbf{A}$ and d play an important role in determining the dynamic behavior of the return covariance structure. We would like to study how, on a multivariate level, those two parameters determine stylized types of dynamic volatility behavior well described for univariate series. Stylized volatility characteristics include mean reversion, long memory, asymmetric relation to return innovations, etc. (see Engle and Patton (2001)).

The parameter $\mathbf{A}$ can be interpreted as a measure of intertemporal sensitivity. This matrix parameter reveals how each element of the current period covariance matrix depends on elements of the previous period covariance matrix. It is the parameter that could in a large part determine mean reversion characteristics on a multivariate level. For example, without restrictions on this matrix parameter, each asset variance, σ_{ii}^2, would depend on the previous period variance of this asset's return, as well as on its covariances with all other assets. Thus a change in the volatility of one asset would affect other assets' volatilities. The interpretation of the intertemporal variance relationships would actually be in terms of the inverse $\mathbf{A}$ (see conditional expectation (4)). The elements of $\mathbf{A}^{-1}$ would also reveal the relative importance of variances and covariances. A higher magnitude of the diagonal elements of $\mathbf{A}^{-1}$ would indicate greater influence of past variances compared to past covariances/correlations. We are not able to completely separate the different effects, but we would like to examine curious situations in which low past correlations may have stronger impact on variances than high correlations.

While the matrix parameter $\mathbf{A}$ carries information about the intertemporal covariance relationships elementwise, the scalar parameter d speaks about the overall strength of these relationships. This parameter would in a large degree account for the presence of long memory or persistence, a phenomenon described for univariate series as today's return having a large effect on the forecast variance many periods in the future (Engle and Patton (2001)). Abundant evidence of pronounced long memory of volatility has been observed while investigating univariate SVOL models (Jacquier, Polson, and Rossi (1994, 1999), Shephard (1996)).

The persistence parameter d is theoretically bound between 0 and 1 (see Philipov and Glickman (2002)). A low d which is close to zero would indicate a weak overall effect of current volatility on future values, i.e. shocks in returns are "forgotten" fast - within a few subsequent periods. Values close to 1 indicate high persistence. The case of $d = 0$ implies constant volatility. In this case the conditional expectation of the volatility in time t is equal to the same value:

$$\begin{aligned} E(\mathbf{\Sigma}_t) &= (v-k-1)^{-1}\,\mathbf{S}_{t-1}^{-1} = \frac{\nu}{v-k-1}\left(\mathbf{A}^{-\frac{1}{2}}\right)\,(\mathbf{\Sigma}_{t-1})^{0}\left(\mathbf{A}^{-\frac{1}{2}}\right)' && (5) \\ &= \frac{\nu}{v-k-1}\left(\mathbf{A}^{-1}\right) && (6) \end{aligned}$$

A case in which $d \geq 1$ implies a non-stationary volatility structure. The special case when $d = 1$ and $\mathbf{A} = \mathbf{I}$ corresponds to a simple matrix-variate random walk on the inverse covariance matrix. In this case we obtain a scale parameter for the Wishart process, $\mathbf{S}_t$, equal to:

$$\mathbf{S}_t = \frac{1}{v}\mathbf{\Sigma}_t^{-1}$$

which leads to a conditional mean equal to the inverse covariance matrix at time t:

$$E(\mathbf{\Sigma}_t^{-1}) = v\,\mathbf{S}_{t-1} = \mathbf{\Sigma}_{t-1}^{-1}$$

As long as the parameter $\mathbf{A}$ is a symmetric positive definite matrix, the parameter d is bounded between 0 and 1, and the shape (degrees of freedom) parameter, ν, of the Wishart distribution, is greater than the number of variables in the model, we have a well defined autoregressive stochastic process for the covariance matrices.

3 Markowitz MV Optimization

This section reviews the estimation of efficient frontiers in absolute space following the classical Markowitz setup (Markowitz (1959)). The $N \times 1$ vector of optimal weights $\mathbf{w}^*$ for a portfolio of N risky returns is the result from the quadratic optimization problem:

$$\min_{\mathbf{w}} \mathbf{w}'\boldsymbol{\Omega}\mathbf{w} \tag{7}$$

subject to

$$\begin{aligned} \mathbf{w}'\boldsymbol{\mu} &= E \\ \mathbf{w}'\boldsymbol{\iota} &= 1 \end{aligned}$$

where $\mathbf{w}$ are portfolio weights, $\boldsymbol{\Omega}$ is the covariance matrix of asset returns, $\boldsymbol{\mu}$ is vector of means, and $\boldsymbol{\iota}$ is a vector of ones. Merton (1972) introduced the familiar efficient set constants:

$$a = \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}, \qquad b = \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\iota}, \qquad \text{and} \qquad c = \boldsymbol{\iota}'\boldsymbol{\Omega}^{-1}\boldsymbol{\iota} \tag{8}$$

which define the equations of the set of MV efficient portfolios:

$$\sigma_p^2 = \frac{a - 2b\mu_p + c\mu_p^2}{ac - b^2}, \qquad \mu_p \geq \frac{b}{c}, \qquad \text{and} \qquad \sigma_p > 0 \tag{9}$$

where μ_p and σ_p are the optimal portfolio mean and variance. The above set of equations define the upper right half of the hyperbola associated with the efficient frontier. The equation for the whole hyperbola can be written as (following Jobson (1991)):

$$H = \frac{\sigma_p^2 - \frac{1}{c}}{a - \frac{b^2}{c}} - (\mu_p^2 - \tfrac{b}{c})^2 = 0 \tag{10}$$

In computing the efficient frontier, we use sample estimates of the mean return vector and covariance matrix which form the estimates (unadjusted for sample bias) of the efficient set constants in eq. (8):

$$\hat{a} = \bar{\mathbf{r}}'\mathbf{S}^{-1}\bar{\mathbf{r}}, \qquad \hat{b} = \bar{\mathbf{r}}'\mathbf{S}^{-1}\boldsymbol{\iota}, \qquad \text{and} \qquad \hat{c} = \boldsymbol{\iota}'\mathbf{S}^{-1}\boldsymbol{\iota} \tag{11}$$

Using the assumption that returns are multivariate normal and adjusting for small sample bias, Jobson (1991) derives the unbiased maximum likelihood estimates of the efficient set constants (eq. 8):

$$\tilde{a} = \tfrac{T}{T-1}\hat{a}, \qquad \tilde{b} = \tfrac{T}{T-1}\hat{b}, \qquad \text{and} \qquad \tilde{c} = \tfrac{T}{T-1}\hat{c} \tag{12}$$

where T is the number of time observations. Furthermore, Jobson (1991) reviews the distributional properties of the three statistics in the above equation and establishes that they are asymptotically independent. These results are used to derive an approximate unbiased estimate of the hyperbola (eq 10):

$$\hat{H} \;=\; \left(\sigma_p^2 - \tfrac{T-1}{\hat{c}(T-N)}\right)\left(\tfrac{T-N-1}{T}(\hat{a} - \tfrac{\hat{b}^2}{\hat{c}}) - \tfrac{N-1}{T}\right) - \left(\mu_p - \tfrac{\hat{b}}{\hat{c}}\right)^2 + \tfrac{1+\hat{a}-\frac{\hat{b}^2}{\hat{c}}}{\hat{c}(T-N)} = 0 \tag{13}$$

Through simulation Jobson (1991) is able to conclude that $\hat{H}$ is normally distributed in large samples with estimate of the variance (up to an approximation and based on the independence assumption of the three statistics (12)):

$$\hat{V}_H = \frac{2(\hat{a}-\frac{\hat{b}^2}{\hat{c}})^2}{\hat{c}^2(T-N)} + \frac{4(\mu_p-\frac{\hat{b}}{\hat{c}})^4}{T(\hat{a}-\frac{\hat{b}^2}{\hat{c}})} + \frac{2}{T-N}(\mu_p - \frac{\hat{b}}{\hat{c}})^4 + \frac{4(\mu_p-\frac{\hat{b}}{\hat{c}})^2}{\hat{c}(T-N)} + \frac{4(\mu_p-\frac{\hat{b}}{\hat{c}})^2(\hat{a}-\frac{\hat{b}^2}{\hat{c}})}{\hat{c}(T-N)} \tag{14}$$

where N is the number of data series. Using (13) and (14) we obtain equations for the lower and upper bounds of the confidence interval for the hyperbola:

$$\hat{LB}_H = \hat{H} - z_{\alpha/2}\sqrt{\hat{V}_H} = 0 \tag{15}$$

$$\hat{UB}_H = \hat{H} + z_{\alpha/2}\sqrt{\hat{V}_H} = 0 \tag{16}$$

Solving equations (10), (15), and (16) provides estimates of the optimal portfolio variance, σ_p^2, along with confidence regions, conditional on specified values for mean portfolio return, μ_p.

It should be noted that the above analytical results apply to MV optimization which places no constraints on the optimal weights. Solutions to problems with no short sale constraints and asset group bounds on optimal weights are obtained through numerical quadratic optimization methods.

4 The Methodology

4.1 Estimating the parameters of the covariance model

All multivariate stochastic volatility models share a common feature: intractability due to a very high degree of parameterization. Many of the traditional models have constraints such as diagonal covariance matrices and constant correlations to deal with the problem of estimating model parameters. Even with these constraints, the literature does not present applications of unconstrained SVOL models of high dimensionality (e.g. five- and higher-dimensional models). The parameters of PG SVOL model are estimated without the need for imposing constraints using a straightforward Bayesian setup where the joint posterior distribution of the parameters, conditional on the observed data, is proportional to the product of the joint prior distribution and the likelihood function.

This section formulates the likelihood function, specifies the joint prior distribution of the parameters, and derives the joint posterior distribution. We then develop MCMC algorithms within this Bayesian framework for sampling from the joint posterior distribution. Markov Chain Monte Carlo Simulation allows sampling from the full posterior distribution of all parameters. A special case of MCMC simulation are the Gibbs sampler and the Metropolis-Hastings algorithm which we use to estimate model parameters. A review of these standard MCMC methods can be found in Gelfand and Smith (1990) and in Gelman, Carlin, Stern, and Rubin (1995). For a detailed discussion of the methodology for the multivariate stochastic volatility model see Philipov and Glickman (2002).

4.1.1 Likelihood Function

The data, $\mathbf{y}_t$, represent k filtered stock returns for a single period t following a multivariate-k normal distribution with a mean vector of zeros and a covariance matrix $\mathbf{\Sigma}_t$. The normal density function for the data is:

$$p(\mathbf{y}_t \mid \mathbf{\Sigma}_t) = \frac{1}{(2\pi)^{\frac{k}{2}} \left|\mathbf{\Sigma}_t\right|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\mathbf{y}_t{}'\mathbf{\Sigma}_t^{-1}\mathbf{y}_t)\right) \tag{17}$$

The inverse of the covariance matrix follows a Wishart distribution with parameters ν and $\mathbf{S}_{t-1}$, whose density function is:

$$p(\mathbf{\Sigma}_t^{-1}|\mathbf{S}_{t-1},\nu) = \frac{\left|\mathbf{S}_{t-1}\right|^{-\frac{\nu}{2}} \left|\mathbf{\Sigma}_t^{-1}\right|^{\frac{\nu-k-1}{2}}}{2^{\frac{\nu k}{2}} \pi^{\frac{k(k-1)}{4}} \prod_{j=1}^{k} \Gamma\left(\frac{\nu+j-1}{2}\right)} \exp\left(-\frac{1}{2} tr\left[\left(\mathbf{S}_{t-1}^{-\frac{1}{2}}\right) \mathbf{\Sigma}_t^{-1} \left(\mathbf{S}_{t-1}^{-\frac{1}{2}}\right)'\right]\right) \quad (18)$$

where $\mathbf{S}_t = \frac{1}{\nu} \left(\mathbf{A}^{\frac{1}{2}}\right) \left(\mathbf{\Sigma}_t^{-1}\right)^d \left(\mathbf{A}^{\frac{1}{2}}\right)'$. This quadratic expression emphasizes the positive definiteness of the covariance matrix $\mathbf{\Sigma}_t$.

Based on the specified distributions for the data and the volatility, an extended multivariate SVOL model (1) with k stock returns and T time periods has the following likelihood function:

$$\begin{aligned} L(\mathbf{\Sigma}^{-1}, \mathbf{A}, \nu, d \mid \mathbf{y}) &= \prod_{t=1}^{T} Wish_k(\mathbf{\Sigma}_t^{-1} \mid \nu, \mathbf{S}_{t-1}) \, N(\mathbf{y}_t \mid \mathbf{0}, \mathbf{\Sigma}_t) = \qquad (19) \\ &= \frac{\left|\frac{1}{\nu} A\right|^{-\frac{\nu}{2}}}{2^{\frac{\nu k}{2}} \pi^{\frac{k(k-1)}{4}} \prod_{j=1}^{k} \Gamma\left(\frac{\nu+j-1}{2}\right)} \prod_{t=1}^{T} \left\{ \left|\left(\mathbf{\Sigma}_{t-1}^{-1}\right)^d\right|^{-\frac{\nu}{2}} \left|\mathbf{\Sigma}_t^{-1}\right|^{\frac{\nu-k-1}{2}} \times \right. \\ &\times \left. \exp\left(-\frac{\nu}{2} tr\left[\mathbf{A}^{-1} \left(\mathbf{\Sigma}_{t-1}^{-1}\right)^{-d} \mathbf{\Sigma}_t^{-1}\right]\right) \left|\mathbf{\Sigma}_t\right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \mathbf{y}_t{}' \Sigma_t^{-1} \mathbf{y}_t\right) \right\} \end{aligned}$$

The term $\mathbf{\Sigma}^{-1}$ without a time subscript represents the collection of all inverse covariance matrices, $\mathbf{\Sigma}_t^{-1}$, for time periods $t = 1, ..., T$. It can be regarded as a 3-D array of T elements of size $k \times k$.

4.1.2 Choice of Prior Distributions

The prior distribution is the product of independent densities. It is necessary to specify a prior distribution for the parameter set $(\mathbf{A}, d, \nu)$. The following densities are suggested for the three parameters, as part of their joint prior distribution:

- The densities in the prior and the posterior for the parameter matrix $\mathbf{A}$ are determined in terms of its inverse for greater convenience. Therefore, for $\mathbf{A}^{-1}$ a Wishart density is used with a $k \times k$ scale matrix $\mathbf{Q}_0$ which is a positive definite symmetric matrix, and can be expressed: $\mathbf{Q}_0 = \left(\mathbf{Q}_0^{1/2}\right)\left(\mathbf{Q}_0^{1/2}\right)'$. The value assigned to $\mathbf{Q}_0$ will depend on expectations about the data. We suggest that $\mathbf{A}^{-1}$ is centered on the identity matrix ($\mathbf{Q}_0 = I$), which by equation (3) in the extended multivariate model results in a mean of $\mathbf{\Sigma}_t^{-1}$ equal to the previous period's covariance matrix, $\left(\mathbf{\Sigma}_{t-1}^{-1}\right)^d$. We also suggest a degrees of freedom parameter $\gamma_0 = k + 1$.
- For d a diffused density such as $p(d) \propto 1$ may be used.
- For ν a gamma diffused density is suggested. Since ν has to be greater than k, the dimension of the covariance matrix, the gamma density is shifted by k. Therefore, it would be appropriate to specify that $(\nu - k)$ has a gamma density.

The prior for the parameters of the model would be product of the densities specified above:

$$p(\mathbf{A}^{-1}, d, \nu) = Wish(\gamma_0, \mathbf{Q}_0) \times p(d) \times Gam(\nu - k) \tag{20}$$

4.1.3 Posterior Sampling - Gibbs Sampler

By Bayes theorem, the joint posterior distribution of the parameter set $(\mathbf{\Sigma}^{-1}, \mathbf{A}^{-1}, d, \nu)$, conditional on the data, is proportional to the product of the prior (20) and the likelihood (20):

$$\begin{aligned} p(\mathbf{\Sigma}^{-1}, \mathbf{A}, \nu, d \mid \mathbf{y}) &\propto p(\mathbf{A}^{-1}, d, \nu) \times L(\mathbf{\Sigma}^{-1}, \mathbf{A}, \nu, d \mid \mathbf{y}) = \\ &= Wish(\gamma_0, \mathbf{Q}_0) p(d) Gam(\nu - k) \prod_{t=1}^{T} Wish(\mathbf{\Sigma}_t^{-1} \mid \nu, \mathbf{S}_{t-1})\, N(\mathbf{y}_t \mid \mathbf{0}, \mathbf{\Sigma}_t) = \\ &= Wish(\gamma_0, \mathbf{Q}_0) p(d) Gam(\nu - k) \times \\ &\times\ c^T \prod_{t=1}^{T} \left| \left(\mathbf{\Sigma}_{t-1}^{-1}\right)^d \right|^{-\frac{\nu}{2}} \left|\mathbf{\Sigma}_t^{-1}\right|^{\frac{\nu-k}{2}} \exp\left(-\frac{\nu}{2} tr[\mathbf{S}^{-1}\mathbf{\Sigma}_t^{-1}]\right) \exp\left(-\frac{1}{2}\mathbf{y}_t{}' \Sigma_t^{-1} \mathbf{y}_t\right) \end{aligned} \tag{21}$$

To estimate the parameters of the model, it is necessary to sample all parameters at once from this posterior distribution. Direct sampling from it, however, is not feasible, in which case MCMC methods, specifically, the Gibbs sampler and the Metropolis algorithm, are used to draw each parameter. The Gibbs sampler iteratively draws, in sequence, each of the parameters over a large number of iterations from their conditional posterior distribution, which, under broad regularity conditions, is equivalent to drawing from the joint posterior distribution.

The conditional posterior distribution $\mathbf{\Sigma}_t^{-1}$ in each time period involves the volatility parameters of only the previous, current, and next periods: $\mathbf{\Sigma}_{t-1}^{-1}$, $\mathbf{\Sigma}_t^{-1}$, and $\mathbf{\Sigma}_{t+1}^{-1}$. For periods $t = 1, 2, ..., T-1$ the conditional posterior distribution is proportional to the product:

$$p(\mathbf{\Sigma}_t^{-1} \mid rest) \propto Wish(\mathbf{\Sigma}_t^{-1} \mid \nu, \mathbf{S}_{t-1}) \times N(\mathbf{0}, \mathbf{\Sigma}_t) \times Wish(\mathbf{\Sigma}_{t+1}^{-1} \mid \nu, \mathbf{S}_t) \tag{22}$$

Sampling from this conditional distribution involves use the Metropolis algorithm.

The last period covariance matrix, $\mathbf{\Sigma}_T^{-1}$, does not depend on the next period's volatility, and is found to follow a known $Wishart$ distribution from which it can be sampled directly.

For the parameter $\mathbf{A}$, using the conditional distribution of the inverse leads to a convenient result for sampling. The conditional distribution for $\mathbf{A}^{-1}$ that is a product of a $Wishart$ prior and a $Wishart$ likelihood, and is also a $Wishart$ distribution, allowing direct sampling.

The parameters d and ν are both scalars. They can be sampled by discretizing a range of values and drawing from the discretized conditional posterior mass functions.

4.2 Computing the efficient frontiers under stochastic covariances

To compute the efficient frontiers we draw J covariance forecasts from their posterior distribution, where J is the size of the posterior sample:

$$\hat{\mathbf{\Sigma}}_{T+1}^{(j)} \sim inv\text{-}Wishart(\nu, \mathbf{S}_T^{(j)}), \qquad j = 1, \ldots, J \tag{23}$$

An unconditional sample mean vector, $\boldsymbol{\mu}$, and the covariance forecast, $\hat{\mathbf{\Sigma}}_{T+1}^{(j)}$, are the inputs for a constrained quadratic optimization problem (under no short sales constraints):

$$\min_{\mathbf{w}} \mathbf{w}^{(j)\prime} \hat{\mathbf{\Sigma}}^{(j)}_{T+1} \mathbf{w}^{(j)} \qquad j = 1, \ldots, J \tag{24}$$

subject to

$$\begin{aligned} \mathbf{w}^{(j)\prime} \boldsymbol{\mu} &= E \\ \mathbf{w}^{(j)\prime} \boldsymbol{\iota} &= 1 \\ \mathbf{w}^{(j)} &\geq 0 \end{aligned}$$

Since we are using the same unconditional mean estimate, the differences in the sets of weights, $\mathbf{w}^{(j)}$, stem from the different covariance forecasts and show the marginal covariance effects on the optimal portfolio composition.

4.3 Traditional Monte Carlo approach to MV optimization

Sampling error estimates can be derived using simulation experiments similar to the one described in Jobson (1991) and Jobson and Korkie (1980). Since analytical approaches are not applicable to cases of constrained MV optimization, Monte Carlo simulation methods can be used to shed light on the informational content of optimal portfolio weights.

In a Monte Carlo simulation experiment, the population mean vector, $\boldsymbol{\mu}$, and covariance matrix, $\mathbf{\Sigma}$, are computed for the set of assets in the MV optimization. Under the assumption of multivariate normal returns, a large number J of samples of size T are drawn from the distribution:

$$\boldsymbol{\mu}^{(j)} \sim N_k(\boldsymbol{\mu}, \mathbf{\Sigma}) \qquad j = 1, \ldots, J \tag{25}$$

and are used as inputs in the constrained optimization problem (24), producing J sets of optimal weights. The sample moments of the simulated weights can be used for diagnostics, hypotheses testing, and comparisons with analytical results.

5 Data

The data comprise value-weighted monthly return series of 5 industry portfolios from the 200112 CRSP database. These 5 portfolios are composed of stocks traded on the NYSE, AMEX, and NASDAQ. Each NYSE, AMEX, and NASDAQ stock is assigned to an industry portfolio at the end of June of each year based on its four-digit SIC code. The extracted and prepared data series can be conveniently downloaded from prof. French's web site (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/). The same set of data has been used by Jostova and Philipov (2002) in the analysis of a time-varying stochastic beta CAPM model. Table 1 provides descriptive information about the 5 return series while Figure (1) visualizes the return dynamics over the 20 year in-sample period from Jan. 1972 to Dec. 1991.

The number of portfolios involved in an MV optimization study is an important issue. The selection in this study of a sample of 5 returns series was influenced by several considerations. The current sample of 5 series is not a constraint of the SVOL model used in this study, as it has been tested for larger datasets. A universe of 5 assest is modest by practical standards. Yet it is representative of a common practice of institutional investors to optimize portfolios with respect to a small number of large asset classes. The specific sample size in this study provides a balanced tradeoff between the ability to facilitate the presentation of results and the practical range of the proposed model applications.

6 Results

Several sets of results are presented in this section. The initial estimates of the SVOL model use a data sample from January 1972 to December 1992. The in-sample period is shifted and the Gibbs sampler is re-run to obtain new estimates and forecasts for the next period. The out-of-sample analysis consists of 120 monthly periods from January 1992 to December 2001. This section begins with a summary of in-sample results from estimating the parameters of the multivariate SVOL model. Then follow results from employing out-of-sample SVOL forecasts in a MV optimization problem. These are compared to analytical results in several studies by Jobson and Korkie, as well as to traditional Monte Carlo analysis results. Finally, focus is shifted to global minimum variance portfolios and the effect of stochastic volatility on MV optimization. The questions that this analysis aims to answer are: could time-varying stochastic covariances bring useful information in MV optimal weights estimation, the estimation of optimal portfolio return forecasts and for expected Sharpe ratios; would we be able to construct useful ex-ante bounds on future realized Sharpe ratios, especially, in the case of constrained portfolio optimization in which analytical results may not be derived?

6.1 Covariance Estimation Results

The in-sample SVOL covariance estimates display significant time-variability. Figure 2 illustrates the mean and 5 and 95 percent confidence bounds of standard deviations of the five industries obtained from the Gibbs posterior sample of covariance matrices. Figure 2 also shows the values of the ordinary sample standard deviations, represented by straight horizontal lines. The standard deviations in Figure 2 are ordered by magnitude, from high to low, as follows: (1) Retail, (2) Finance, (3) Manufacturing, (4) Other, and (5) Utilities. Note that even though the parameter estimates vary significantly through time, changing by more than 50% in certain periods, they preserve their magnitude rank, evolving through time in parallel movements that reveal the influence of common driving factors.

In-sample correlation estimates were backed out of the estimated SVOL covariance matrices. Figure 3 displays the 10 pair-wise mean correlations (first section), and the five data series of returns (second section). The pair-wise correlations can be ranked from high to low as follows: (1) Other-Manufacturing. (2) Retail-Manufacturing, (3) Finance-Manufacturing, (4) Other-Finance, (5) Finance-Retail, (6) Other-Retail, (7) Finance-Utilities, (8) Other-Utilities, (9) Utilities-Manufacturing, and (10) Retail-Utilities. Lower values of pairwise correlations display significantly higher time variability than high correlation estimates. We can also observe evidence of the correlations-breakdown phenomenon documented in previous research (Longin and Solnik (2001); Jacquier and Marcus (2001)). Longin and Solnik (2001), for example, take extreme return outcomes and observe highly increased asset correlations in such periods which diminish diversification benefits. In contrast to this extreme outcome approach, the current time-varying volatility model preserves time continuity and provides a new view-angle at correlations behavior in severe market movements. It is also interesting to observe that the correlation pairs preserve their ranking over the 240 in-sample periods, especially for lower correlations. This evidence supports the proposition in Jacquier and Marcus (2001) that correlations are driven by common market forces. Such synchronous correlations behavior could also explain the behavior of the optimal weights discussed later in the section.

6.2 Covariance forecasts

For every period t, the Gibbs sampler provides a posterior sample of J covariance matrices. Using a posterior sample size of $J = 2000$, we simulate a forecast covariance matrix for each of the J Gibbs sampler draws of the last period covariance matrix $\mathbf{\Sigma}_T^{(j)},\ j = 1, \ldots, J$, according to the model:

$$\hat{\mathbf{\Sigma}}_{T+1}^{(j)} \sim inv\text{-}Wish(\nu, \mathbf{S}_T^{(j)}) \qquad j = 1, \ldots, J \tag{26}$$

where

$$\mathbf{S}_T^{(j)} = (\mathbf{A}^{(j)})^{\frac{1}{2}} (\mathbf{\Sigma}_T^{-1\ (j)})^d (\mathbf{A}^{(j)}) \frac{1}{2} \tag{27}$$

Figure 4 shows sample draw sequences of standard deviation forecasts from the predictive posterior distribution for a single period (January 1992). Also shown are histograms of the draw sequences, which exhibits skewness similar to the univariate case in which volatilities are usually described to follow a χ^2 distribution. The white horizontal line represents the predictive posterior mean forecast. The right-hand tail of the predictive posterior density is extended allowing for large volatility values of the forecasts. Such large values can push the bounds on the volatility forecasts upward allowing for events that have been described as extremely low probability (e.g. the concurrence of events that brought down LTCM, see Jorion (1999)) to be considered reasonably probable. The average standard deviation forecasts for all 120 out-of-sample monthly periods from January 1992 to December 2001 are depicted in Figure 5. We can observe that the forecasts preserve their ranking. Exceptions are Manufacturing and Other, whose estimates are close to each other.

6.3 Optimal portfolio results

The optimal weights of the minimum variance portfolio under no-short-sale constraint are computed using the draw sequence of SVOL forecasts for January 1992 (see Figure 4) as inputs in the MV optimizer. Retail and Finance have the lowest weights in the minimum variance portfolio. These two return series have the highest unconditional and conditional variances (see Table 1 and Figure 2). Their correlations with the other series are in the mid-range. Utilities are the most heavily weighted. They have the lowest unconditional volatility and among the lowest pair-wise correlations with the rest of the industries.

The first section in Figure 6 shows the set of all efficient frontiers based on the draw sequence of 2000 covariance forecasts for January 1992. The MV optimizer computes all 2000 frontiers using the same unconditional mean return vector. The solid line represents the mean efficient frontier based on mean standard deviation for each level of return. The dashed lines are the 95th and the 5th percentile bounds. The percentile bounds are computed separately for every level of return and do not necessarily represent portfolios on the same frontier. This point is illustrated in the second section of the graph. The two dashed lines contain minimum variance portfolios which are in the 5th and 95th percentiles. However, other portfolios along these frontiers are far from the two bounds. This graph provides evidence that the stochastic frontiers are not parallel even though optimal portfolios are based on exactly the same set of assets.

In Figure 7, the stochastic frontier bounds for January 1992 are compared to the analytical confidence bounds derived in Jobson (1991). The analytical bounds widen considerably with the increase in expected optimal portfolio returns. The 95th percentile analytical bound is inverted and includes high probability outcomes of portfolios which dominate all others and contradict the efficient set framework. The stochastic frontier bounds are well-behaved in all of the different period samples. For this particular sample the stochastic bounds are $\pm 1\%$ around the mean measured in monthly standard deviation values.

Figure 8 compares the stochastic frontier bounds with bounds obtained through a traditional Monte Carlo simulation approach. This approach involves generating 2000 samples of data from a multivariate normal distribution using the ordinary sample mean vector and covariance matrix as parameters. The covariance parameters of the generated data sets are used as inputs in the MV optimizer. These are estimates of the unconditional covariance matrix. The fixed mean vector used to obtain the SVOL efficient frontiers was also used to produce the MC-simulated frontier results. The unconditional bounds obtained through traditional MC simulation are tighter. The large difference in the sizes of the two sets of bounds speaks of the effect of time-variability of the SVOL forecasts.

Figure 9 shows the five sets of forecasted optimal weights for the minimum variance portfolio over the out-of-sample period from January 1992 to December 2001. The mean weights do not exhibit notable time variation, despite the variation in covariance forecasts. This lack of time variation can be explained by the fact the although variances and correlations vary though time, they preserve their ranking (see Figure 3). The percentile bounds on the weights, however, show significantly greater variability for the assets with smaller weights. In addition, if we look at the samples of optimal weights for a single period, we observe very high sensitivity of the weights with respect to the covariance inputs, which is in line with the results of Best and Grauer (1991) and Chopra and Ziemba (1993). Figure 12 plots the point estimates of minimum-variance weights for the out-of-sample periods based on covariance forecasts from three alternative approaches: (1) the ordinary sample covariance matrix based on a moving sample window (dot-dashed line), and (2) a variant of Engle's dynamic conditional correlation (DCC) model. The alternatively computed sets of weights, especially Engle's DCC model weights, exhibit markedly different behavior. In general, these weights underweight more significantly Retail and Fiance. DCC weights have significantly higher time-variation that can be largely attributed to more independent movement of covariance forecasts.

Figure 10 displays a plot and a histogram of the sample of Sharpe ratio forecasts for January 1992, along with 5 and 95 percent bounds. The low autocorrelation in the obtained sequence of Sharpe ratio forecasts reveals a high degree of efficiency in the estimates. Without loss of generality, the risk free rate is assumed zero. The Sharpe ratios are based on the expected returns and standard deviations of the optimal minimum variance portfolios based on the SVOL forecasts for the period (see Figure 6). The histogram shows a posterior distribution of the Sharpe ratios that is symmetric and bell-shaped.

Figure 11 plots the mean ex-ante Sharpe ratios of the minimum variance portfolio over the 120 out-of-sample periods from January 1992 to December 2001, along with the 5th and 95th percentile. The percentile bounds can provide important information on the uncertainty of the forecasts. The bounds are wider for higher expected Sharpe ratios, conveying that higher performing portfolios should carry higher adjustments for uncertainty.

6.4 Out-of-sample performance

The above results show abilities of the SVOL model to expand the informational content of the MV optimizer. An optimization study like this, however, carries the indispensable question: does the model improve portfolio performance? The performance of stochastic covariances is benchmarked to three alternative approaches to estimating covariance parameters: (1) forecasts based on the ordinary sample covariance matrix of a moving sample window of monthly observations, (2) a variant of the Dynamic Conditional Correlations model of Engle (2002), and (3) an index model for time-varying covariances of the type proposed in Jacquier and Marcus (2001).

The dynamic conditional correlations (DCC) model was introduced by Engle (2002) and classified by him as a new class of multivariate GARCH estimators which can be best viewed as a generalization of the constant conditional correlation estimator of Bollerslev (1990) . The model is formulated as:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t \qquad \mathbf{D}_t = diag(\sqrt{h_{i,t}}) \tag{28}$$

The model uses a covariance structure that combines time-varying variances and time-varying correlations structure. Time-varying volatilities are modeled using univariate GARCH(1,1) processes, while time-varying correlations are modeled using exponentially smoothed standardized GARCH(1,1) residuals.

The index model decomposes the covariance matrix into systematic and idiosyncratic components. The systematic component is a quadratic form product of factor sensitivies and the factor volatility.

$$\mathbf{\Sigma}_t = \mathbf{b}\sigma^2_{m,t}\mathbf{b}' + \mathbf{\Omega} \tag{29}$$

All time variation in this index model is driven by the scalar factor volatility. The constant factor sensitivities and the idiosyncratic risk are estimated as parameters of an OLS regression model using the full sample of high frequency data. The market volatility is estimated using high frequency data of very recent periods (in this study, the last 60 days). The covariance estimate then is adjusted to reflect monthly volatilities.

The out-of-sample minimum variance weights for the three alternative approaches are shown in Figure 12. With the exception of the moving average results, the weights from the DCC and the factor models are significantly time-varying. In the case of the DCC model, this time variation can be attributed to the fact that asset variances do not preserve their rank, since they are modeled as separate GARCH(1,1) processes producing independent forecasts. In the case of the factor model, variances preserve rank. However, pronounced differences in betas and wide variation of market volatility set variance forecasts so much apart as to force the optimizer to load all weight on one asset, ignoring the correlation structure. For example, the Utilities portfolio has a beta of 0.6, much smaller than other sensitivities. The small beta appears to lead to a volatility forecast much lower than the rest. We observe that almost in half of the out-of-sample periods the Utilities portfolio has unit weight.

Figure 14 illustrates a particular way of representing out-of-sample Sharpe ratios in the context of the forecasts produced by the SVOL model. For each month of the second half of the out-of-sample period, Figure 14 displays out-of-sample Sharpe ratios based on a moving window of the preceding 60 months. We observe that the SVOL forecast with their 5% and 95% bounds are able to capture the realized out-of-sample Sharpe ratios. Another interesting observation is the significant month-to-month variation in those Sharpe ratios, not typical for moving average series. Such variation speaks of large swings in actual monthly returns.

The overall performance of minimum variance portfolios over the whole out-of-sample period is reported in Table 2. This table compares realized Sharpe ratios for the full out-of-sample period generated using the four alternative covariance models: (1) the ordinary sample covariance matrix based on a moving sample window, (2) a variant of Engle's dynamic conditional correlation (DCC) model, (3) an index model, and (4) the multivariate SVOL model. The reported mean returns, standard deviations and realized Sharpe ratios of actual returns cover the out-of-sample period from January 1992 to December 2001. We notice that the SVOL model generates the minimum variance portfolio with smallest standard deviation and highest out-of-sample realized Sharpe ratio. Table 2 also reports the average of the 5% and 95% ex-ante bounds on the monthly Sharpe ratio forecasts. The realized out-of-sample Sharpe ratios fall within these bounds. In addition, Table 2 provides standard errors for all performance statistics. These standard errors have been computed using GMM and the delta method. In the estimation of sample means and variances, the GMM methodology provides their covariance matrix as standard output. The delta method is used to compute the standard error of the Sharpe ratio, which is a non-linear function of the two sample statistics. The delta method specifies that if we have a vector of estimated parameters, $\hat{\theta}$, which are asymptotically normally distributed:

$$\sqrt(T)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \sim N(\mathbf{0}, \mathbf{V}_\theta) \tag{30}$$

then a function of $\boldsymbol{\theta}$ follows a normal distibution with parameters:

$$\sqrt(T)(f(\hat{\boldsymbol{\theta}}) - f(\boldsymbol{\theta}_0)) \sim N(\mathbf{0}, \mathbf{V}_f) \tag{31}$$

where

$$\mathbf{V}_f = \frac{\partial f'}{\partial \boldsymbol{\theta}} \mathbf{V}_\theta \frac{\partial f}{\partial \boldsymbol{\theta}} \tag{32}$$

Using the above results, we obtain the Sharpe ratio's standard error:

$$\begin{aligned} \boldsymbol{\theta} &= \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}, \qquad f(\boldsymbol{\theta}) = \frac{\mu}{\sqrt{\sigma^2}}, \qquad \frac{\partial f}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{1}{\sigma} \\ -\frac{\mu}{2\sigma^3} \end{bmatrix} \\ SE_{SR} &= \begin{bmatrix} \frac{1}{\sigma} & -\frac{\mu}{2\sigma^3} \end{bmatrix} \times \begin{bmatrix} v_\mu^2 & v_{\mu,\sigma^2} \\ v_{\mu,\sigma^2} & v_{\sigma^2}^2 \end{bmatrix} \times \begin{bmatrix} \frac{1}{\sigma} \\ -\frac{\mu}{2\sigma^3} \end{bmatrix} \end{aligned}$$

Table 2 also shows confidence intervals for the SVOL statistics. In each of the months of the out-of-sample period, the optimization process generates 2000 sets of optimal weights based on the SVOL forecasts. The actual investment strategy is the mean vector of optimal weights. All other sets of weights are treated as potential investment paths. By summarizing those potential investment strategies at the end of the out-of-sample period, we obtain the 5% and 95% bounds to form the confidence intervals for the statistics. These SVOL confidence intervals are tighter than the ones obtained by the GMM and the delta method.

7 Conclusion

This paper presents an approach to examining the informational content of Markowitz efficient portfolios using a general multivariate stochastic volatility model. The paper focuses exclusively on the effect of stochastic covariances on MV optimization. Even though the results in this paper were obtained using a very general specification of a multivariate SVOL model, they provided useful insights into the behavior of volatilities and correlations, and the way this behavior can affect optimal portfolio weights. For example, keeping expected means constant, a well observed time variation in correlations and volatilities may not necessarily lead to markedly time-varying optimal weights. A common factor driving volatilities and correlations can preserve their ranking and reduce time-variation in optimal weights.

The multivariate SVOL model has a flexible formulation which allows easy impositions of constraints. One set of useful constraints can impose a factor structure for the model. Such factor structure gives opportunities for providing richer conditioning information in estimating the model parameters and for testing a wider range of financial hypotheses. Second, a factor structure in the SVOL model will incorporate conditioning information about expected returns, a piece which could complement current analysis. Third, a factor structure allows easier estimation of higher-dimensional models as it effectively reduces the dimensionality of the time-varying stochastic component.

References

Ammann, M., and H. Zimmermann, 2001, Tracking Error and Tactical Asset Allocation, *Financial Analysts Journal* 57(2), 32–43.

Balduzzi, P., and Ludan Liu, 2001, Parameter Uncertainty and International Investment in a Multi-Period Setting, Unpublished Paper, Department of Finance, Boston College.

Barry, C.B., 1974, Portfolio Analysis under Uncertain Means, Variances, and Covariances, *Journal of Finance* 29, 515–522.

Best, M.J., and R.R. Grauer, 1991, On The Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results, *The Review of Financial Studies* 4(2), 315–342.

Bollerslev, T., 1990, Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model, *Review of Economics and Statistics* 72, 498–505.

Britten-Jones, M., 1999, The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights, *The Journal of Finance* 54(2), 655–671.

Chen, S., and S.J. Brown, 1983, Estimation Risk and Simple Rules for Optimal Portfolio Selection, *Journal of Finance* 38, 1087–1093.

Chopra, V.K., and W.T. Ziemba, 1993, The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice, *The Journal of Portfolio Management* 19(2), 6–11.

Engle, R., 2002, Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics* 20(3), 122–150.

Engle, R., and A. Patton, 2001, What Good Is A Volatility Model, Unpublished Paper, Department of Economics, University of California, San Diego.

Fletcher, J., and J. Hillier, 2001, An Examination of Resampled Portfolio Efficiency, *Financial Analysts Journal* 57(5), 66–74.

Gelfand, A.E., and A.F.M. Smith, 1990, Sampling-based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association* 85, 398–409.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin, 1995, *Bayesian Data Analysis*. (Chapman & Hall, London) 1st edn.

Harvey, A., E. Ruiz, and N. Shephard, 1994, Multivariate Stochastic Variance Models, *Review of Economic Studies* 61, 247–264.

Horst, J.R., F.A. de Roon, and B.J.M. Werker, 2002, Incorporating Estimation Risk in Portfolio Choice, Unpublished Paper, Department of Finance and CentER, Tilburg University.

Jacquier, E., and A.J. Marcus, 2001, Asset Allocation Models and Market Volatility, *Financial Analysts Journal* 57(2), 16–30.

Jacquier, E., N. G. Polson, and P.E. Rossi, 1994, Bayesian Analysis of Stochastic Volatility Models (with discussion), *Journal of Business and Economic Statistics* 12, 281–300.

Jacquier, E., N. G. Polson, and P.E. Rossi, 1995, Models and Prior Distributions for Multivariate Stochastic Volatility, Unpublished Paper, Graduate School of Business, University of Chicago.

Jacquier, E., N. G. Polson, and P.E. Rossi, 1999, Stochastic Volatility: Univariate and Multivariate Extensions, Unpublished Paper, CIRANO Working Paper Series, Paper 99s-26, Montreal.

Jagannathan, R., and T. Ma, 2002, Risk Reduction in Large Portfolios: Why Imposing the Wrong Consraints Helps, Unpublished Paper, Department of Finance, Kellog School of Management, Northwestern University.

Jiang, G., and P. van der Sluis, 2000, Index Option Pricing Models with Stochastic Volatility and Stochastic Interest Rates, Unpublished Paper, Schulich School of Management, York University.

Jobson, J.D., 1991, Confidence Regions for the Mean-Variance Efficient Set: An Alternative Approach to Estimation Risk, *Review of Quantitative Finance and Accounting* 1, 235–257.

Jobson, J.D., and B. Korkie, 1980, Estimation for Markowitz Efficient Portfolios, *Journal of the American Statistical Association* 75(371), 544–554.

Jobson, J.D., and B. Korkie, 1989, A Performance Interpretation of Multivariate Tests of Asset Set Intersection, Spanning, and Mean-Variance Efficiency, *Journal of Financial and Quantitative Analysis* 24(2), 185–204.

Jorion, P., 1986, Bayes-Stein Estimation for Portfolio Analysis, *Journal of Financial and Quantitative Analysis* 21(3), 279–292.

Jorion, P., 1991, Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection, *Journal of Financial Economics* 22, 717–727.

Jorion, P., 1999, Risk Management Lessons from Long-Term Capital Management, Unpublished Paper, Graduate School of Management, University of California at Irvine.

Jorion, P., 2002a, Enhanced Index Funds and Tracking Error Optimization, Unpublished Paper, Graduate School of Management, University of California at Irvine.

Jorion, P., 2002b, Portfolio Optimization with Constraints on Tracking Error, Unpublished Paper, Graduate School of Management, University of California at Irvine.

Jostova, G., and A. Philipov, 2002, Bayesian Analysis of Stochastic Betas, Unpublished Paper, Department of Finance, Boston College.

Kandel, S., and R.F. Stambaugh, 1989, A Mean-Variance Framework for Tests of Asset Pricing Models, *Review of Financial Studies* 2, 125–156.

Longin, F., and B. Solnik, 2001, Extreme Correlations of International Equity Markets, *Journal of Finance* 56, 649–676.

Mahieu, R., and P. Schotman, 1994, Stochastic Volatility and the Distribution of Exchange Rate News, Unpublished Paper, Department of Finance, University of Limburg.

Markowitz, H. M., 1959, *Portfolio Selection: Efficient Diversification of Investments.* (Wiley and Sons New York).

Merton, R., 1972, An Analytical Derivation of the Efficient Portfolio Frontier, *Journal of Financial and Quantitative Analysis* 7, 1851–1872.

Michaud, R.O., 1989, The Markowitz Optimization Enigma: Is 'Optimized' Optimal?, *Financial Analysts Journal* 45(1), 31–42.

Michaud, R.O., 1998, *Efficient Asset Management.* (Harvard Business Scholl Press Boston, MA).

Philipov, A., and M.E. Glickman, 2002, Multivariate Stochastic Volatility via Wishart Random Walks, Unpublished Paper, Department of Finance, Boston College.

Pitt, M., and N. Shephard, 1999, Time Varying Covariances: A Factor Stochastic Volatility Approach, in J.M. Bernardo, J.O. Berger, A.P.David, and A.F.M. Smith, eds.: *Bayesian Statistics, Volume 4* (Oxford University Press, Oxford, UK).

Pojarliev, M., and W. Polasek, 2001, Portfolio Construction by Volatility Forecasts: Does the Covariance Structure Matter?, Unpublished Paper, Institute Of Statistics and Econometrics, University of Basel.

Roll, R., 1992, A Mean/Variance Analysis of Tracking Error, *The Journal of Portfolio Management* 18(4), 13–22.

Rudolf, M., H.J. Wolter, and H. Zimmermann, 1999, A Linear Model for Tracking Error Minimization, *Journal of Banking and Finance* 23, 85–103.

Shephard, N., 1996, Statistical Aspects of ARCH and Stochastic Volatility, in D.R. Cox, D.V. Hinkley, and O.E. Barndorff-Nielson, eds.: *Time Series Models: In Econometrics, Finance and Other Fields* (CRC Press, London).

Table 1
Descriptive Statistics of Industry Portfolios

The table defines the company composition in the five industry portfolios used in the study and describes their returns. Each industry portfolio contains all companies listed on CRSP whose four-digit SIC code falls in the corresponding broad industry group. We use the monthly weighted average returns for each industry.

A. Company classification based on SIC four-digit company code.

	Industry	**SIC code**
1	Manufacturing	2000-3999
2	Utilities	4900-4999
3	Shops - Wholesale, Retail, and Some Services	5000-5999 and 7000-7999
4	Money, Finance	6000-6999
5	Other - Agriculture, Mines, Oil, Construction, Transportation, Telecom, Health and Legal Services	all others

B. Correlation Matrix of Monthly Returns

	Manufacturing	Utilities	Retail	Finance	Other
Manufacturing	1.00				
Utilities	0.68	1.00			
Retail	0.89	0.63	1.00		
Finance	0.88	0.77	0.87	1.00	
Other	0.92	0.71	0.82	0.88	1.00

C. Descriptive Statistics of Monthly Return Series

	Manufacturing	Utilities	Retail	Finance	Other
Mean	0.0107	0.0107	0.0113	0.0102	0.0104
Standard deviation	0.0500	0.0414	0.0627	0.0536	0.0485
S.E. of the mean	0.0032	0.0027	0.0040	0.0035	0.0031
Median	0.0095	0.0082	0.0125	0.0079	0.0133
Skewness	-0.3586	0.2505	-0.3164	-0.0746	-0.4458
Kurtosis	5.6994	4.6777	5.6089	4.3248	5.1475
Minimum	-0.2390	-0.1212	-0.2858	-0.2017	-0.2200
Maximum	0.1674	0.1896	0.2646	0.2051	0.1775
Range	0.4064	0.3108	0.5504	0.4068	0.3975

Table 2
Out-of-sample Sharpe Ratios

The table compares the out-of-sample performance of the minimum variance portfolio constructed using four alternative covariance models: (1) the ordinary sample covariance matrix based on a moving sample window, (2) a variant of Engle's dynamic conditional correlation (DCC) model, (3) an index model, and (4) the multivariate SVOL model. The reported mean returns, standard deviations and realized Sharpe ratios of actual returns cover the out-of-sample period from January 1992 to December 2001. Standard errors of all statistics are reported in parentheses. These standard errors were derived using GMM and the delta method. Also reported for the SVOL model are 5% and 95% ex-post bounds (in brackets) using the 2000 sets of optimal weights through the 120 out-of-sample periods as potential investment paths. The last two-rows present the averate of SVOL ex-ante bounds on the Sharpe ratios.

	Mean Return	Standard Deviation	Sharpe Ratio
Moving window	0.0072	0.0342	0.2102
	(0.0031)	(0.0022)	(0.0934)
Single factor	0.0113	0.0581	0.1945
	(0.0053)	(0.0043)	(0.0967)
Dynamic Correlations	0.0074	0.0344	0.2151
	(0.0032)	(0.0021)	(0.0943)
SVOL	0.0077	0.0329	0.2337
	(0.0030)	(0.0021)	(0.0945)
	[0.0062 0.0089]	[0.0321 0.0369]	[0.1801 0.2734]
Average 5% (ex-ante) bound of SVOL Sharpe ratio forecasts			0.1981
Average 95% (ex-ante) bound of SVOL Sharpe ratio forecasts			0.3825

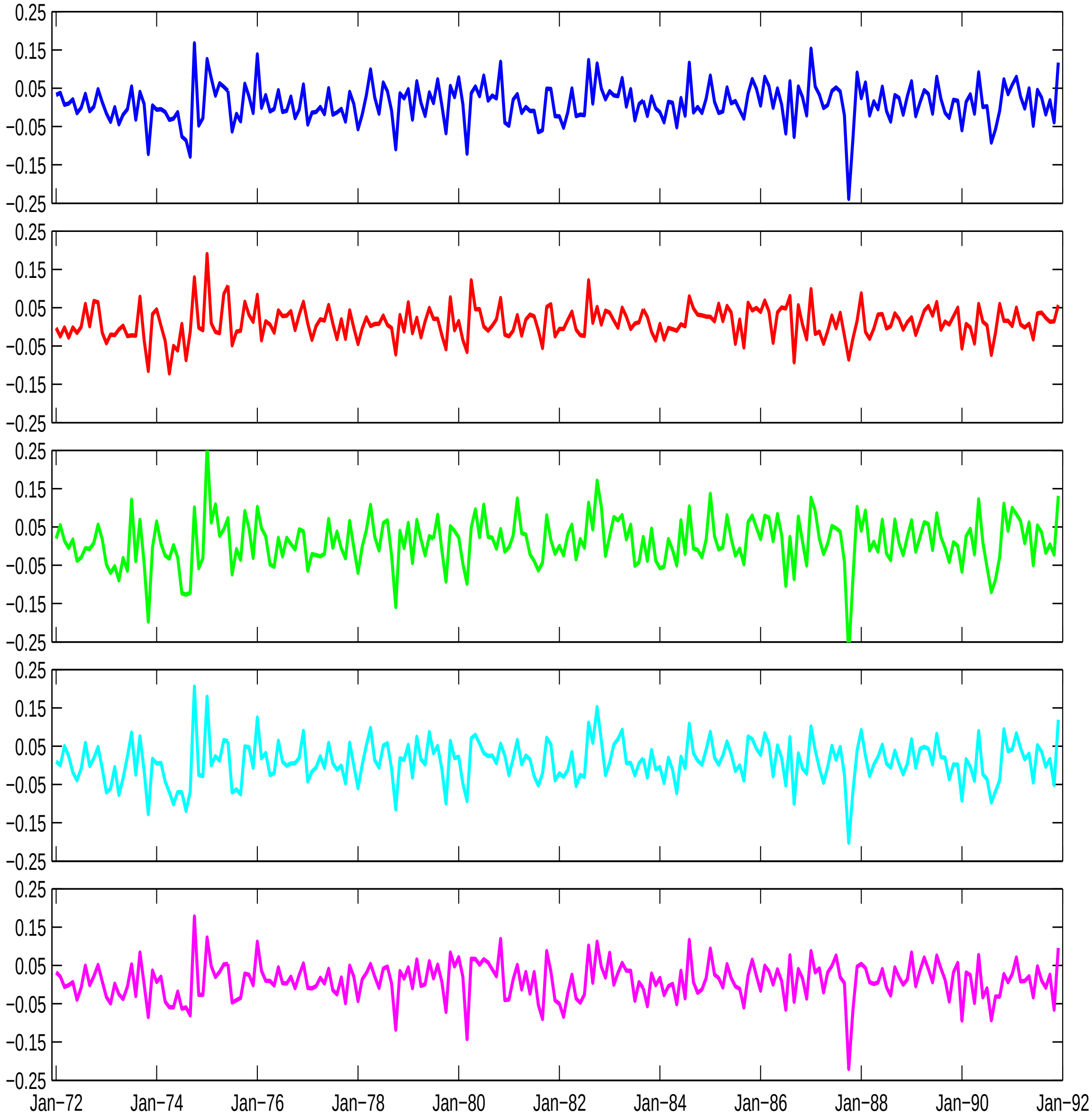

Figure 1. Plots of the data series for the five industries, Manufacturing, Utilities, Retail and Wholesale, Finance, and Other, over 240 monthly periods from January 1972 to December 1991.

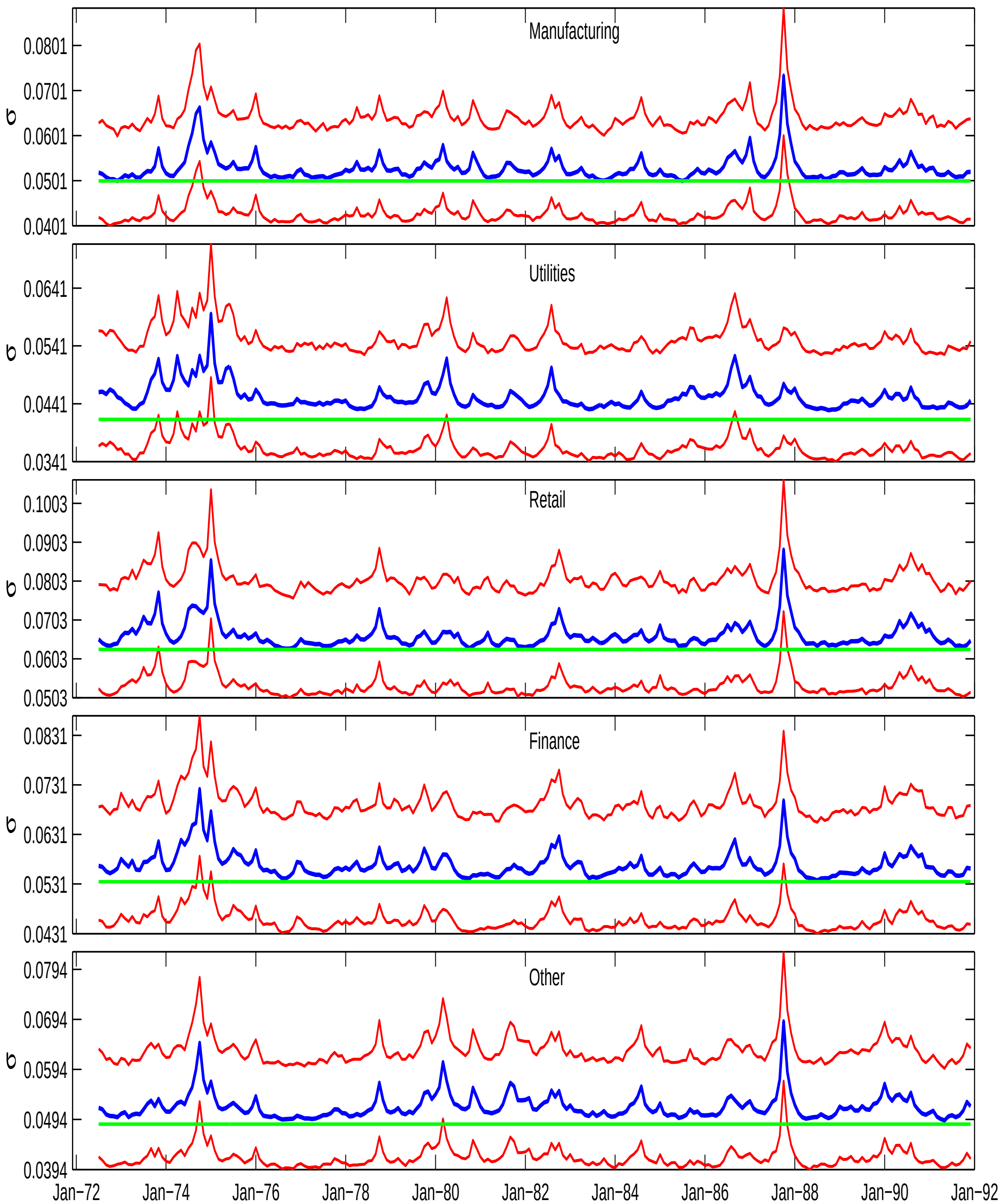

Figure 2. Posterior standard deviations from the SVOL model with the 5th and 95th percentile bounds. The straight horizontal line represents unconditional mean estimate of the sample standard deviation.

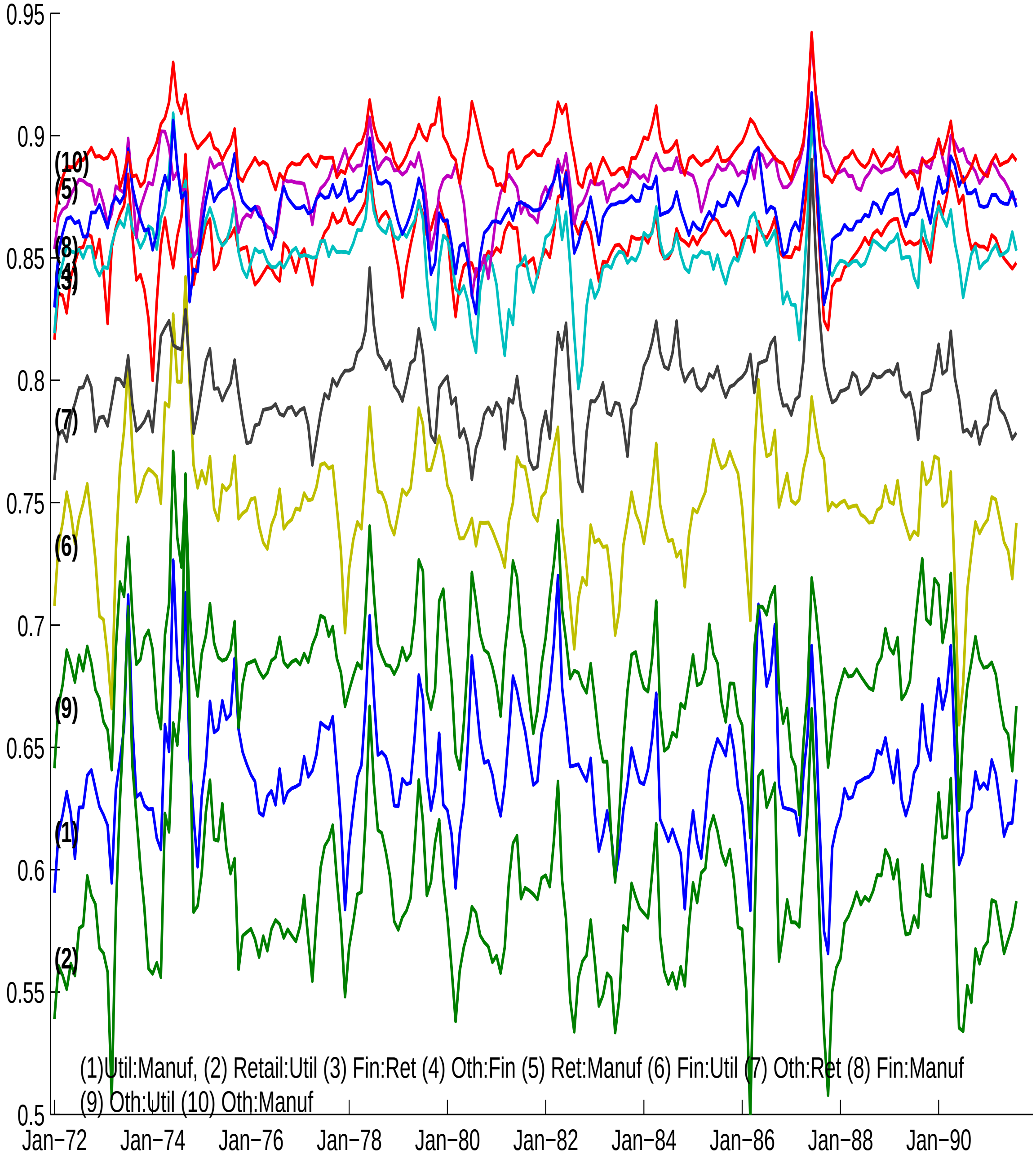

Figure 3. Plots of the 10 pair-wise correlations. The correlation matrices were backed out from the SVOL covariance matrix estimates. The correlation estimates are plotted against the five data series in the second section of the graph.

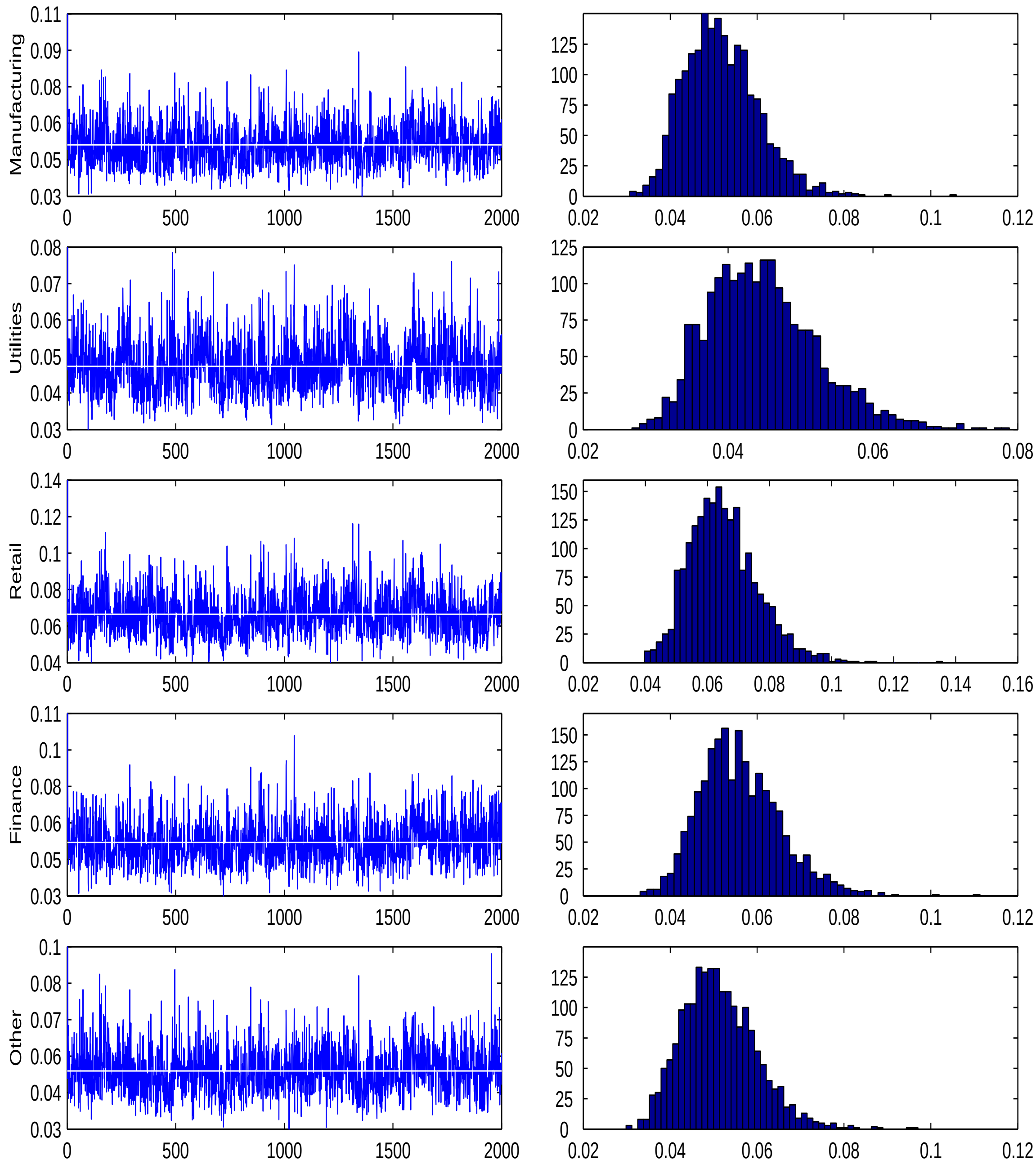

Figure 4. Plots and histograms of the set of draws of one-step-ahead forecasts of individual standard deviations for January 1992 based on the posterior sample of the last period covariance matrix.

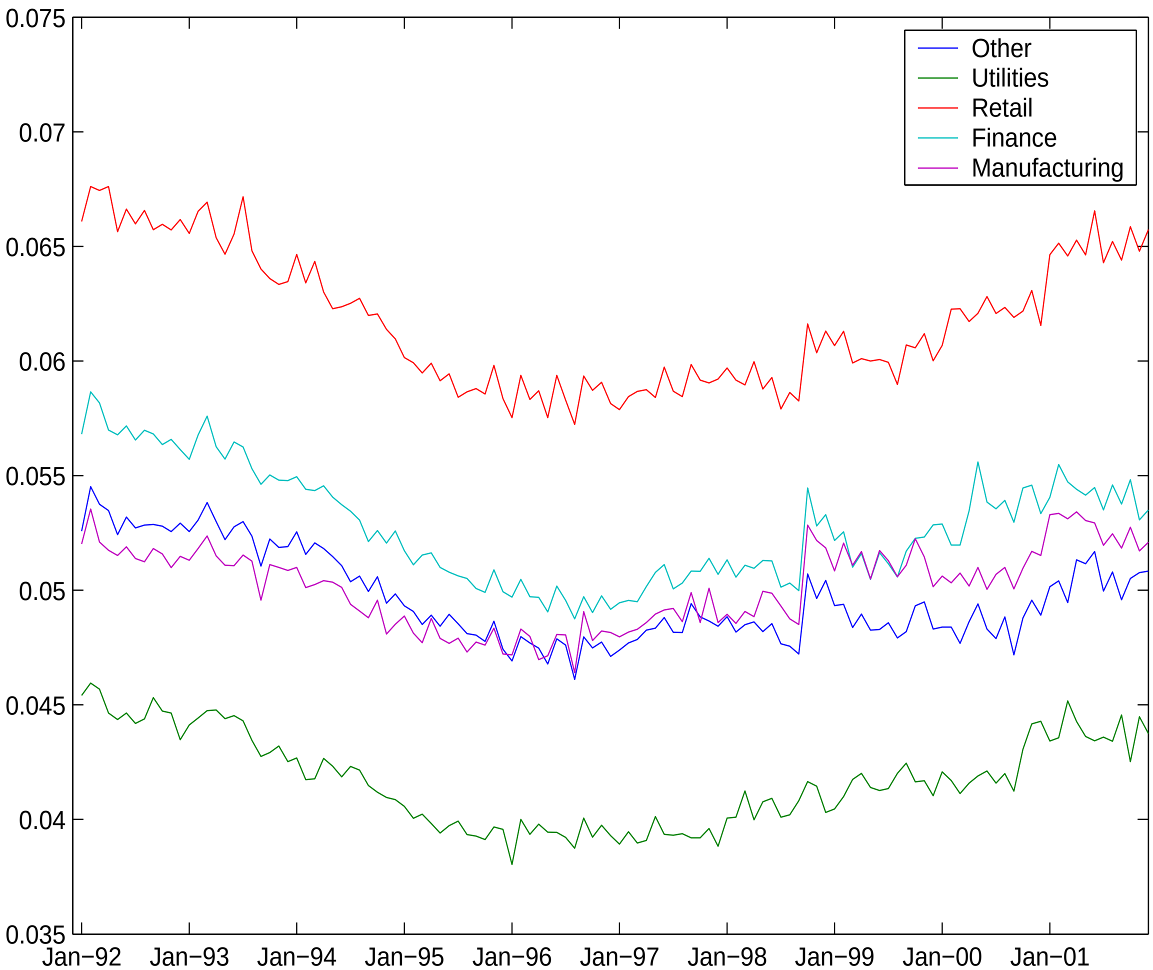

Figure 5. Plot of the out-of-sample one-period-ahead standard deviation forecasts for the period of January 1992 to December 2001. The standard deviation forecasts preserve their rank and are ordered, from high to low: (1) Retail, (2) Finance, (3) Other, (4) Manufacturing, and (5) Utilities.

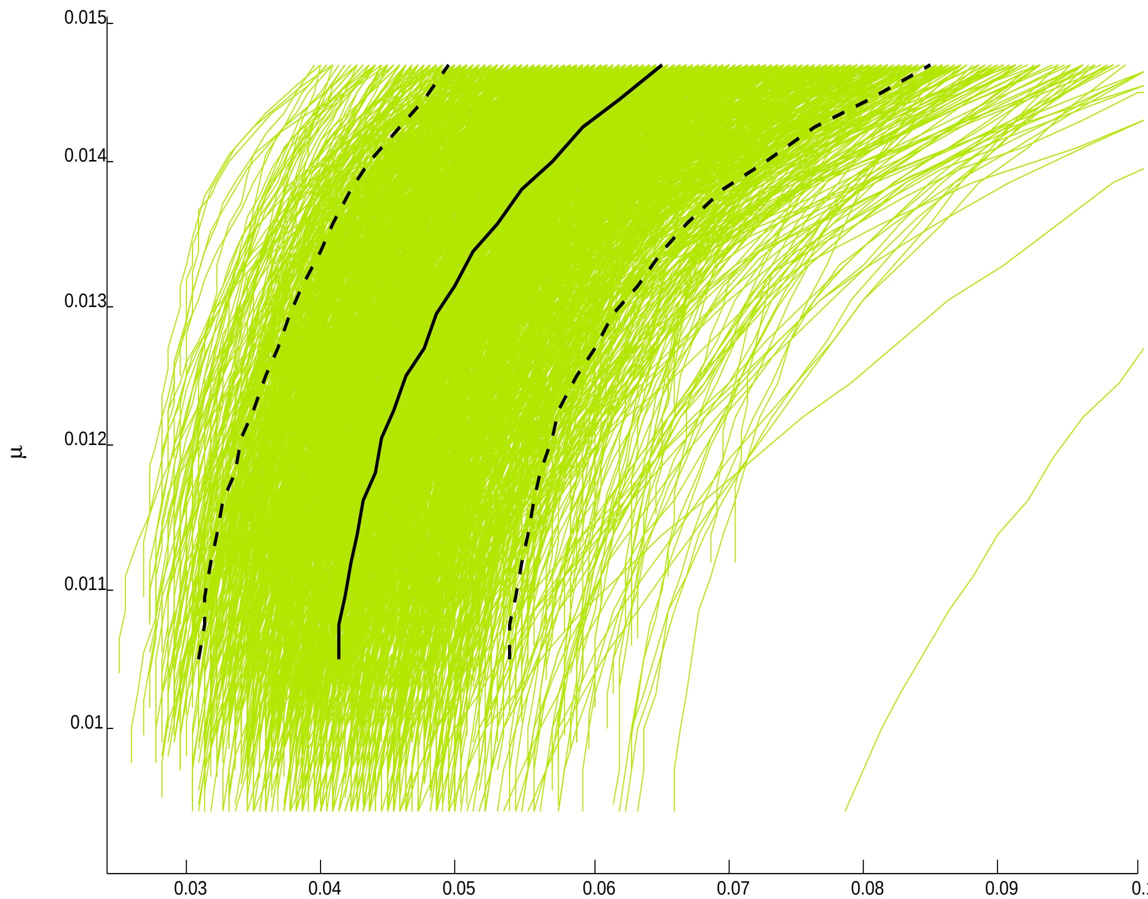

Figure 6. The first graph shows all frontiers based on the sample of SVOL covariance forecasts for Jan 1992, along with the mean frontier and the 95th and 95th percentile bounds.

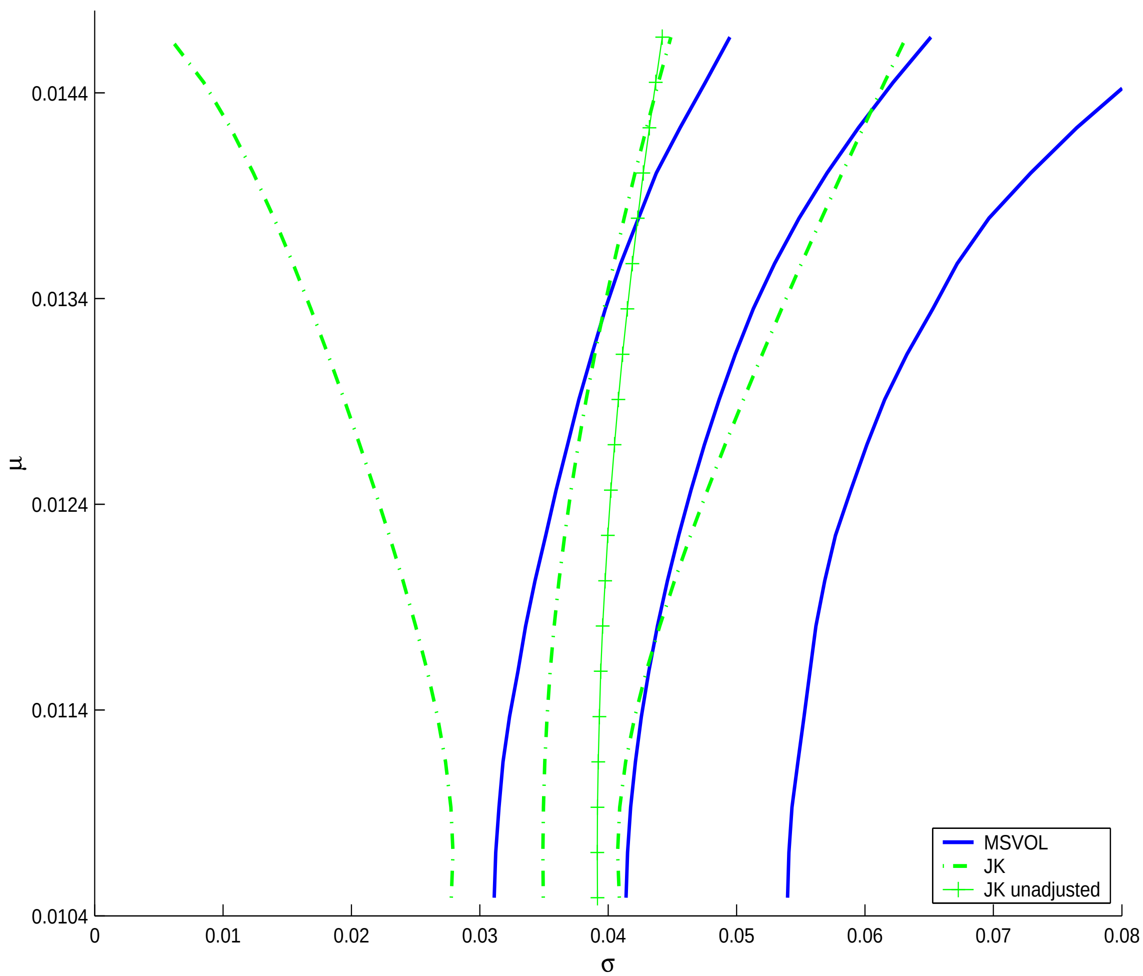

Figure 7. Plots of the stochastic frontiers bounds based on the multivariate SVOL (also shown in figure 6) and the analytical results in Jobson (1991). The stochastic frontier bounds are obtained through constrained MV optimization as in Figure 6, while the analytical bounds are based on unconstrained optimization. The shape of the analytical bounds appears to be data sensitive and in some cases the upper bound may reach infeasible areas.

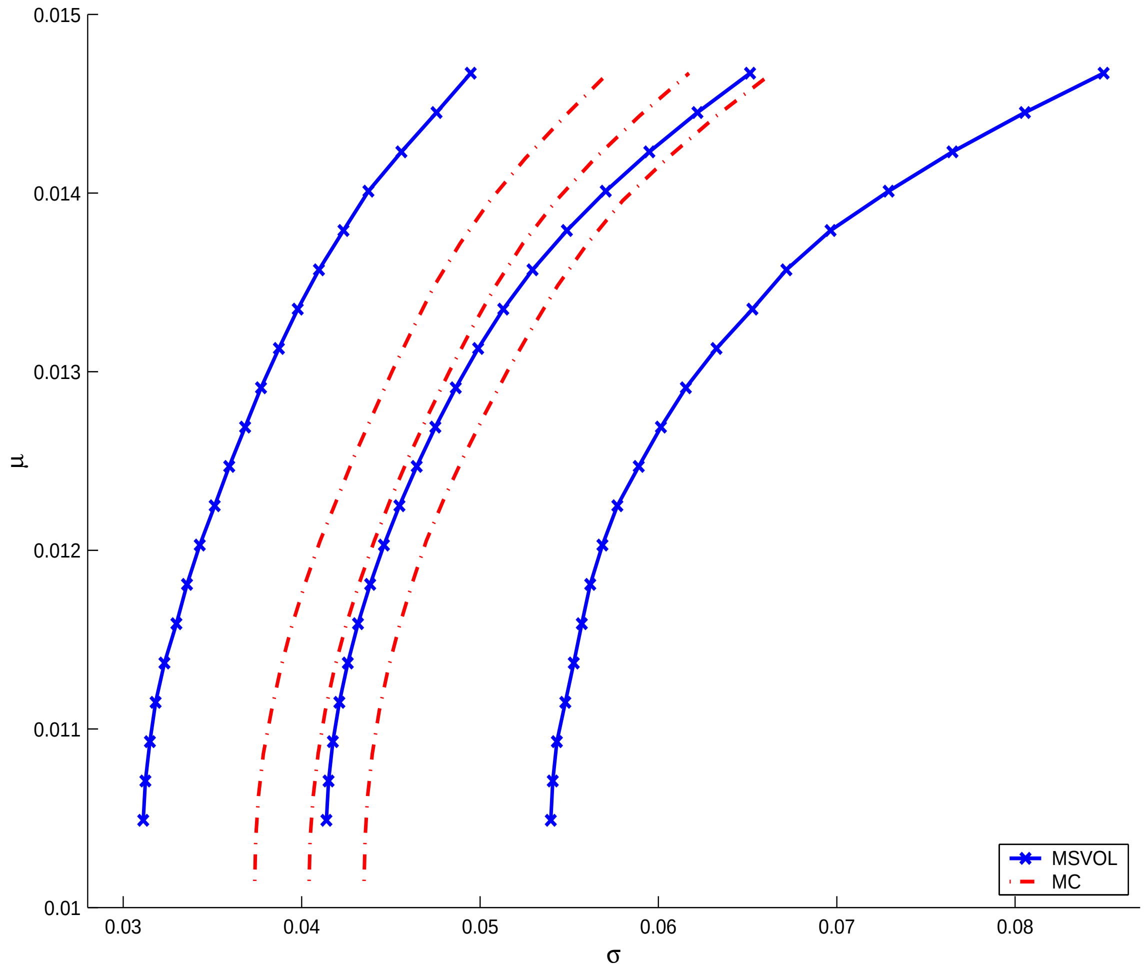

Figure 8. Plots of the multivariate SVOL efficiency bounds and bounds obtained via traditional Monte Carlo simulation. In the simulation exercise, 2000 samples of returns were drawn from a multivariate normal distribution with mean vector and covariance matrix equal to the unconditional data sample means and and covariances. The set of MC-simulated bounds are based only on the covariance matrices from the simulated samples (using the same unconditional mean of the original data). These bounds are directly comparable to the SVOL efficiency bounds. The wider SVOL bounds illustrate the effect of using conditioning information.

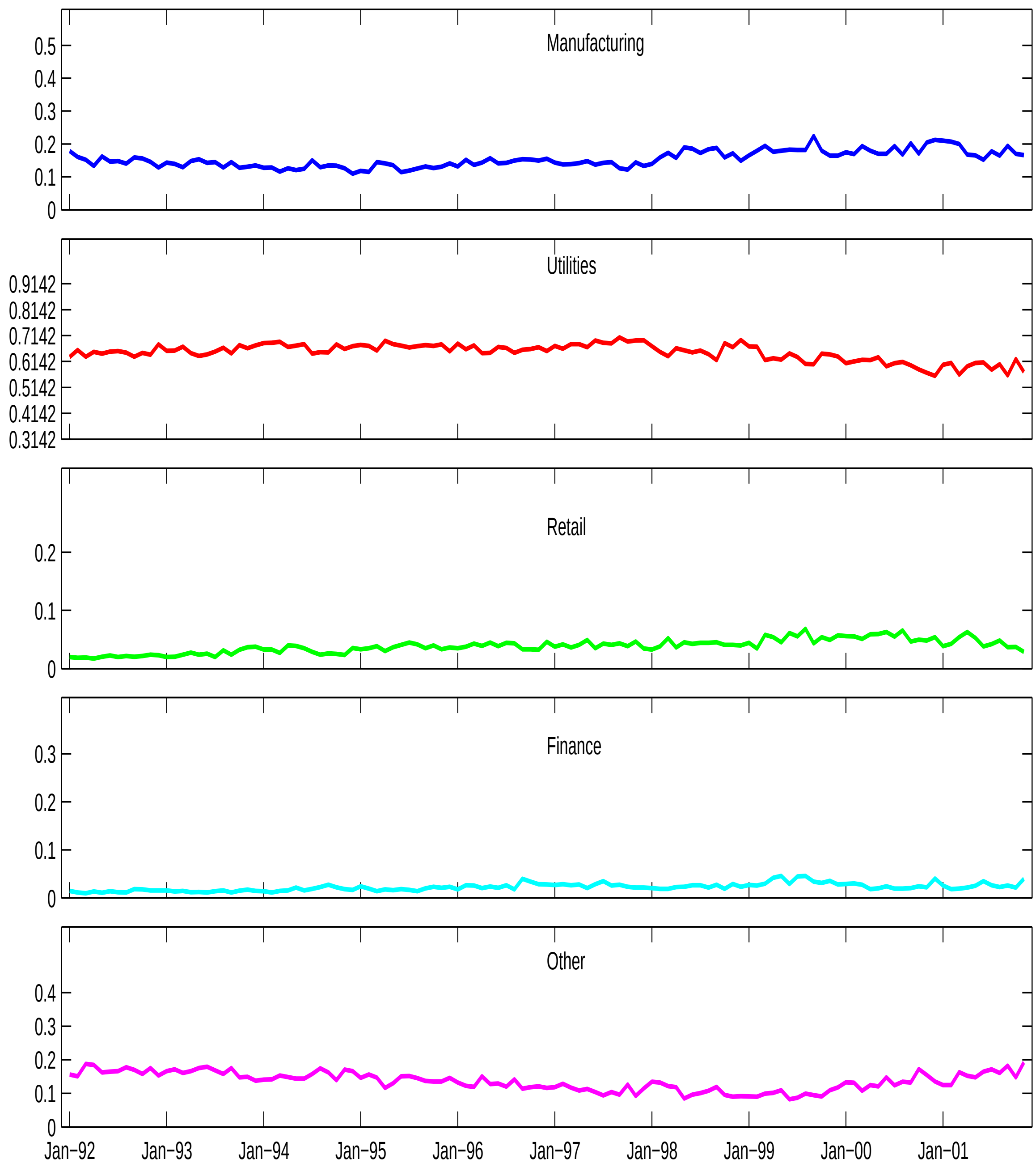

Figure 9. Plots of the optimal weights of the minimum variance portfolios, estimated using the SVOL model, in the out of sample period from January 1992 to December 2001.

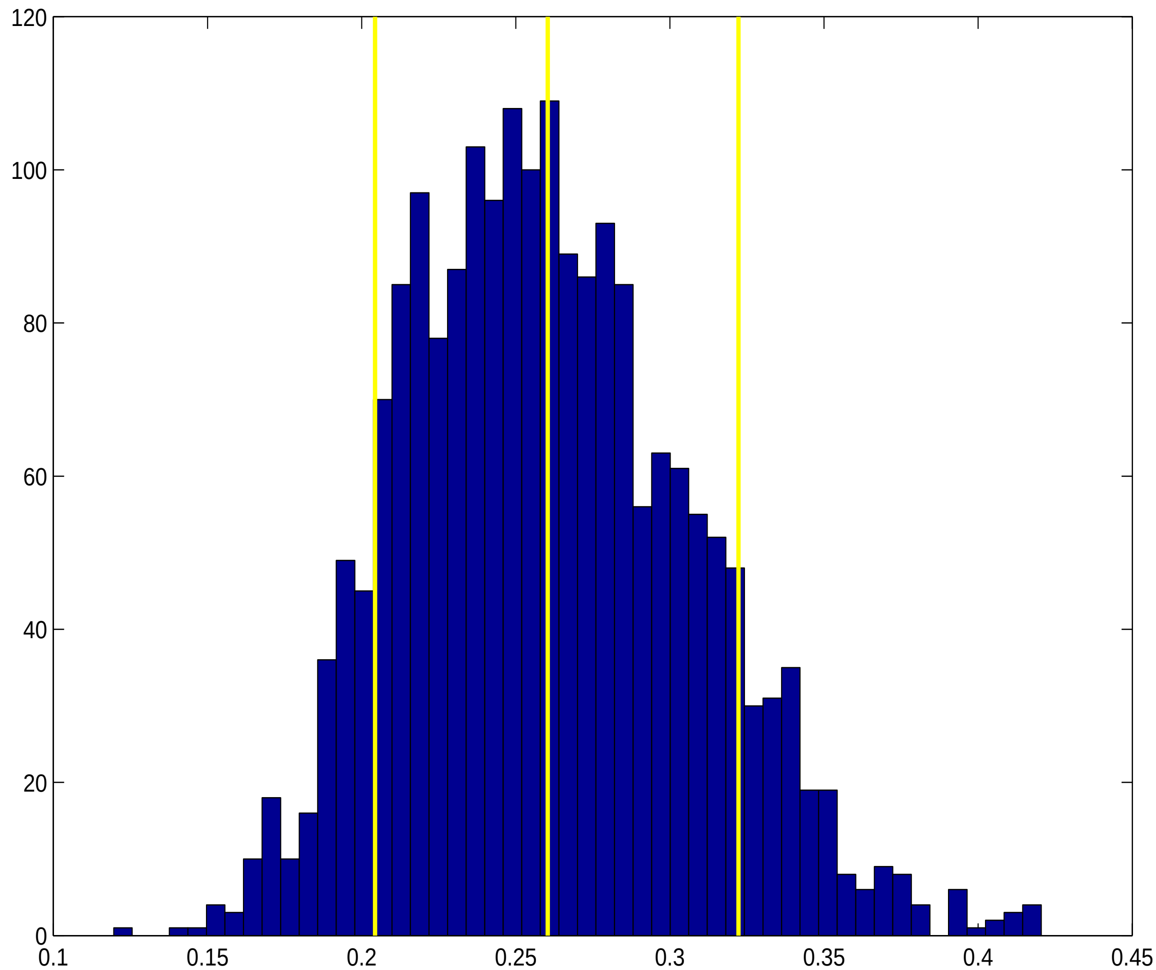

Figure 10. Histogram of the Sharpe ratios of the estimated minimum variance optimal portfolios (the risk free rate is assumed zero) for January 1992. The light color vertical lines are the mean expected Sharpe ratio and the 5th and the 95th percentile.

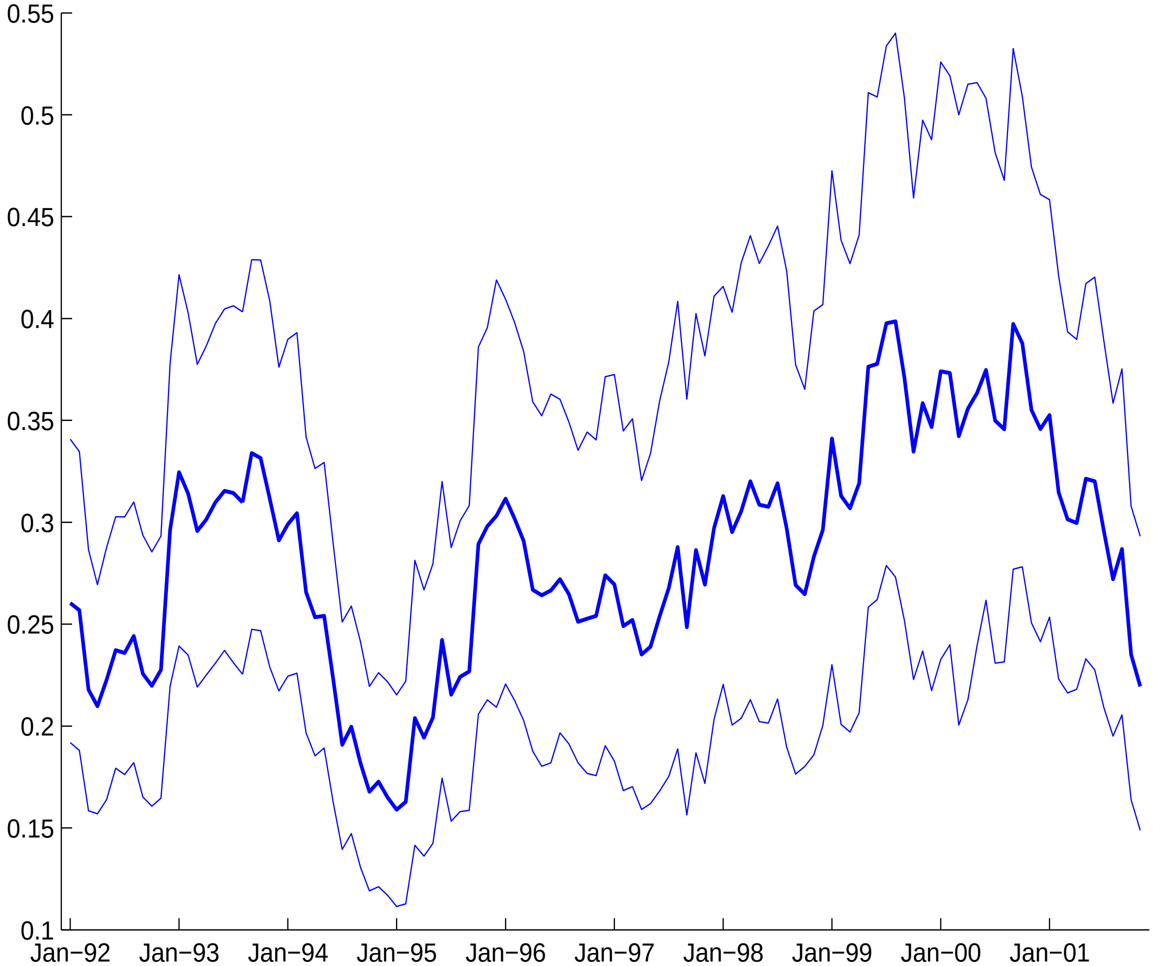

Figure 11. Plot of the expected Sharpe ratios for the posterior mean of the minimum variance portfolio for the periods from January 1992 to December 2001, along with the the 5th and 95th percentiles. The forecasts are based on expected optimal optimal portfolio returns (using the ordinary sample mean as return forecast) and the expected covariances (using the multivariate SVOL model forecasts).

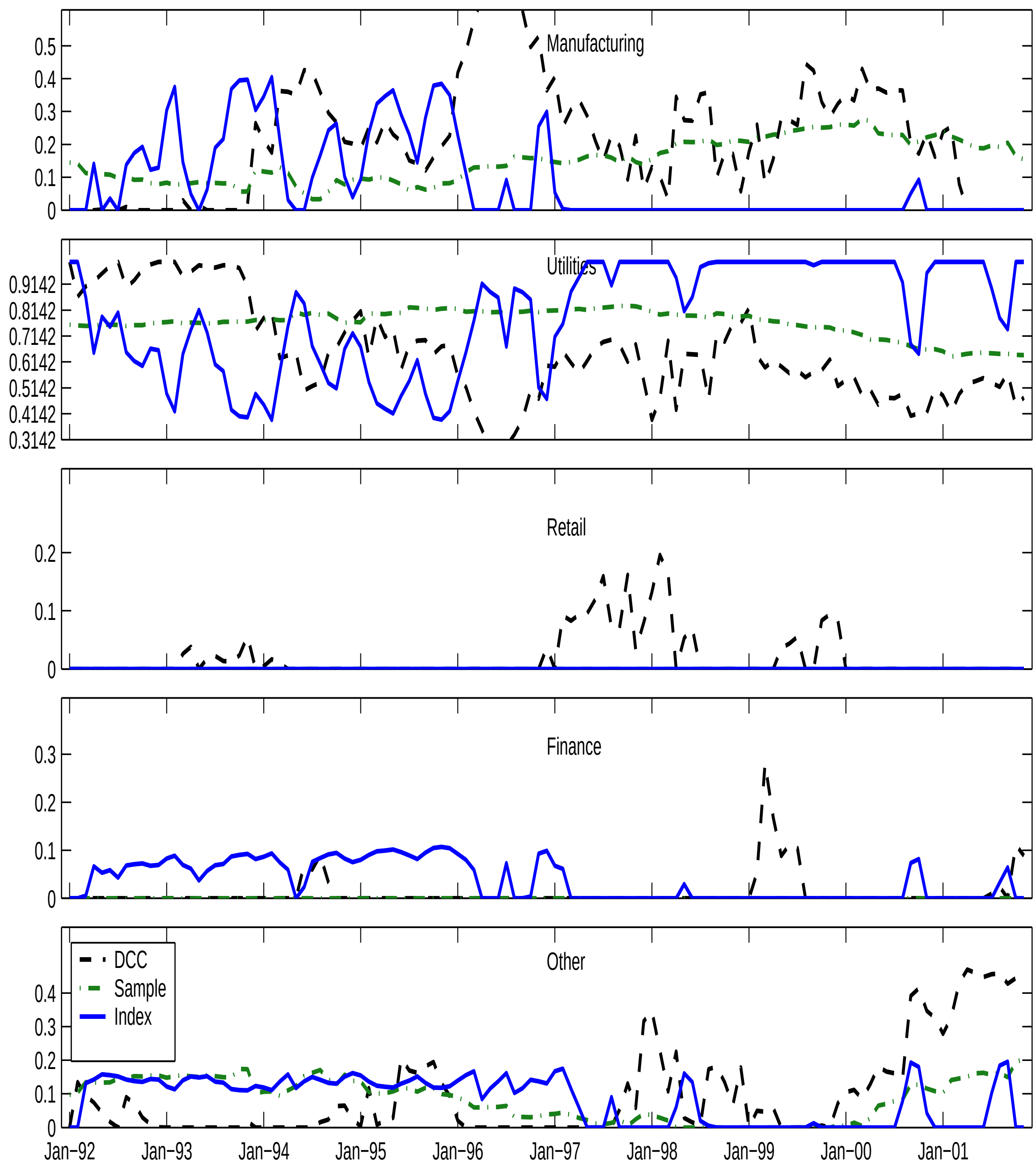

Figure 12. Plots of the global minimum variance portfolio weights for the out-of-sample periods based on covariance forecasts from three alternative approaches: (1) an index model (solid line), (2) the ordinary sample covariance matrix based on a moving sample window (dashdot line), and (3) Engle's dynamic conditional correlation model (dashed line).

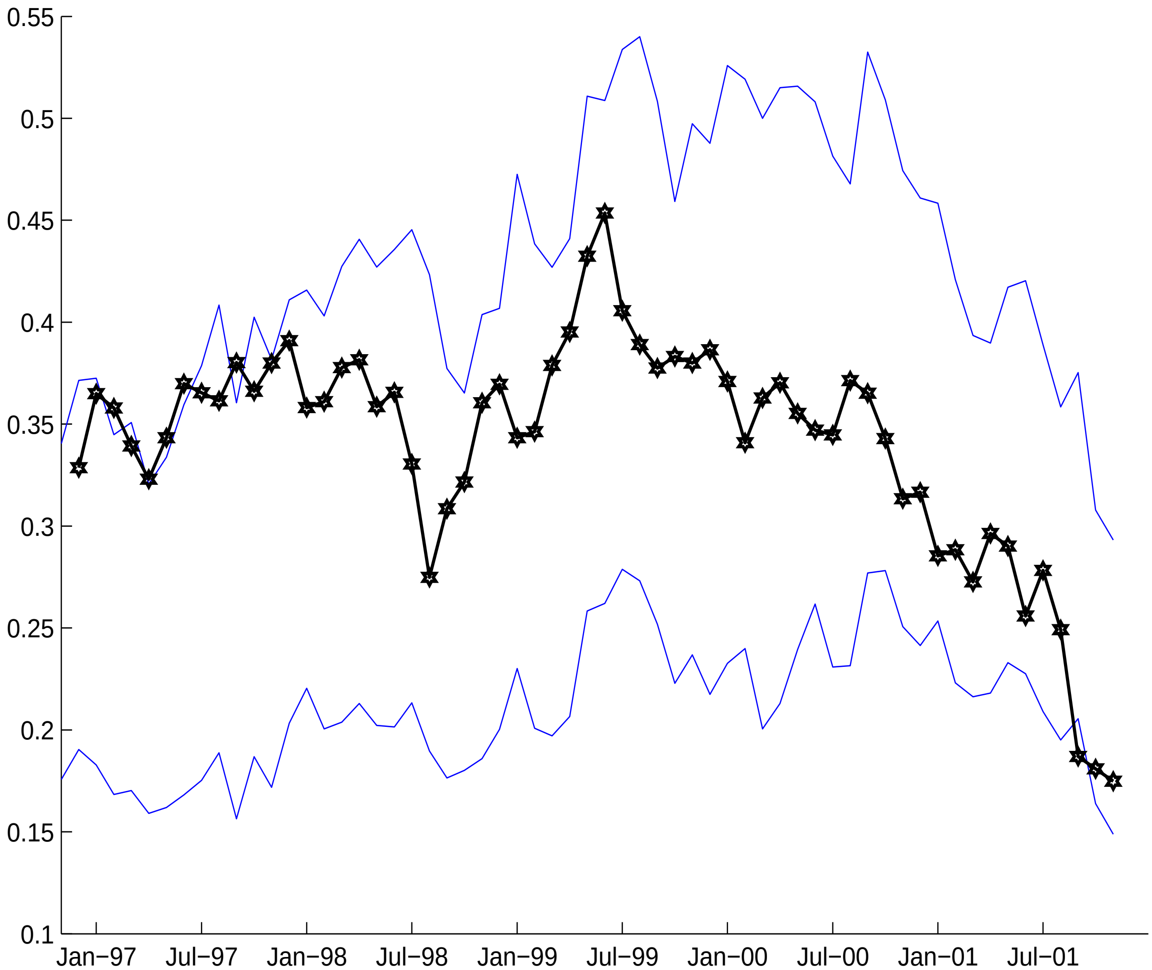

Figure 13. Plots of the 5 and 95% bounds for the monthly Sharpe ratio forecasts based on the SVOL model, also shown in Figure 9. Added to the plot were the point estimates of out-of-sample realized Sharpe ratios based on the preceding 60 months for the minimum variance portfolio constructed using the SVOL model covariance forecasts.

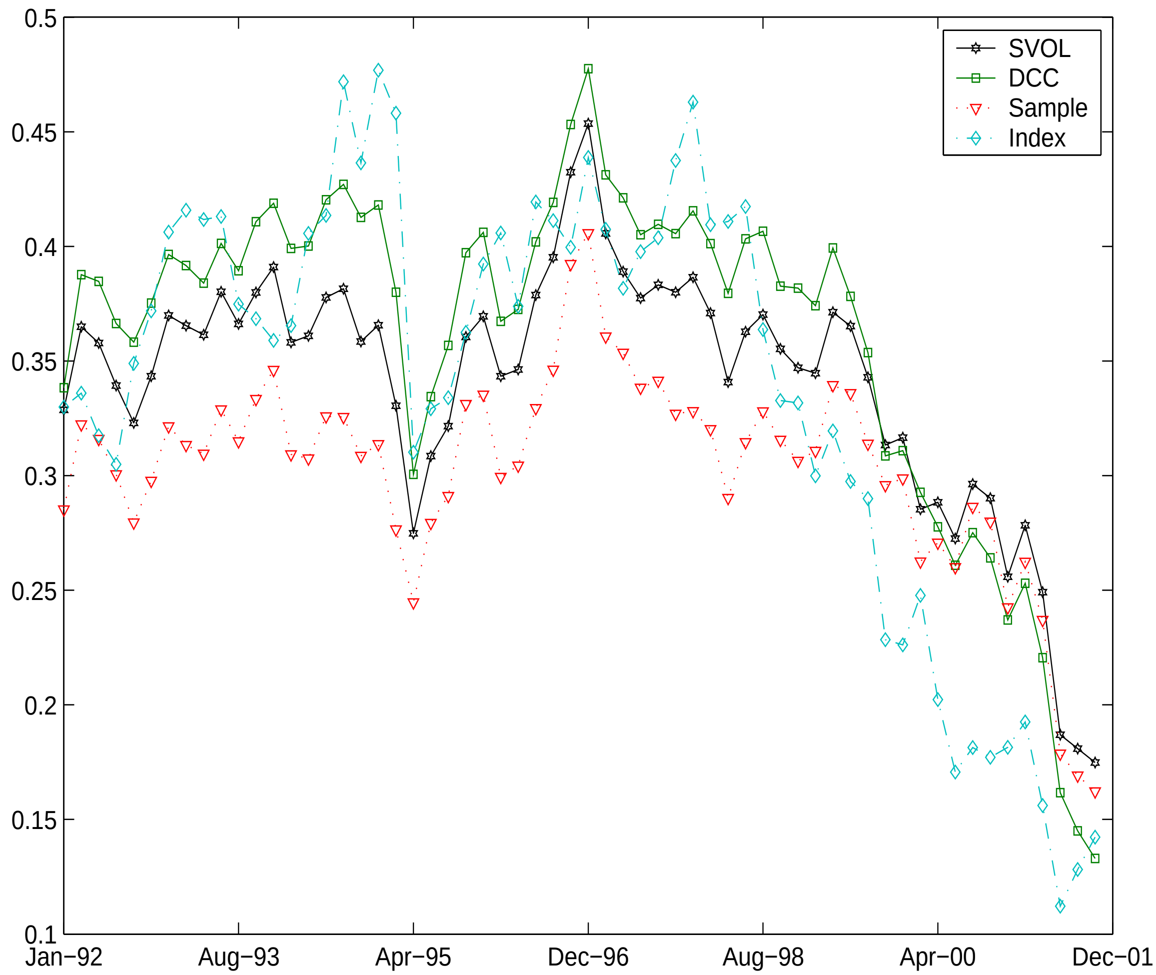

Figure 14. Point estimates of out-of-sample realized Sharpe ratios based on the preceding 60 months. The four sets of realized Sharpe ratios based on (1) SVOL model parameters (SVOL) (2) Engle's dynamic conditional correlations model (DCC), (3) unconditional sample parameters (Sample), and (4) an index model (Index).