

Factor Multivariate Stochastic Volatility Via Wishart Processes

Alexander Philipov
Department of Finance
Kogod School of Business
American University

Mark E. Glickman
Department of Health Services
School of Public Health
Boston University

December 9, 2005

Abstract

This paper proposes a high dimensional factor multivariate stochastic volatility (MSV) model in which factor covariance matrices are driven by Wishart random processes. The framework allows for unrestricted specification of intertemporal sensitivities, which can capture the persistence in volatilities, kurtosis in returns, as well as correlation breakdowns and contagion effects in volatilities. The factor structure allows addressing high dimensional setups used in portfolio analysis and risk management, as well as modeling conditional means and conditional variances within the model framework. Due to the complexity of the model, we perform inference using Markov chain Monte Carlo simulation from the posterior distribution. A simulation study is carried out to demonstrate the efficiency of the estimation algorithm. We illustrate our model on a data set that includes 88 individual equity returns and the two Fama-French size and value factors. With this application, we demonstrate the ability of the model to address high dimensional applications suitable for asset allocation, risk management and asset pricing.

1 Introduction

Most financial applications focus on the risk and return of financial portfolios and are concerned with the correlated movements of their assets. Recent advances in the financial literature have focused on the time-variation of asset covariances and correlations. Efforts to improve covariance forecasts have led to the development of time-varying volatility models, such as multivariate GARCH and Stochastic Volatility (SV). Multivariate models of time-varying volatility, however, have two major drawbacks that have prevented them from enjoying the success of their univariate counterparts. First, multivariate models can be highly parameterized, resulting in potential interpretation difficulties. Second, multivariate models bring a significant increase in the computational complexity of model fitting, in part due to complex model structures, and in part due to the increased number of parameters.

The literature on multivariate volatility models is extensive. Multivariate ARCH models were introduced by Engle and Kraft (1983) and further developed by Bollerslev, Engle, and Wooldridge (1988) (BEW), and by Engle and Kroner (1995) with their very popular BEKK model. Imposing an explicit factor structure in the multivariate ARCH models was first suggested by Engle, Ng, and Rothschild (1990). Factor structure finds motivation in asset pricing theory. It explains the co-movements of asset returns and reduces the dimensionality of the multivariate models. One most characteristic restriction of ARCH models is that all variation in the covariances is driven by return shocks.

Allowing the covariances to have their own stochastic component produces the multivariate stochastic volatility models studied by Harvey, Ruiz, and Shephard (1994), Mahieu

and Schotman (1994), Jacquier, Polson, and Rossi (1995), etc. An alternative approach to these early multivariate SVOL is offered by Uhlig (1997) which can be regarded as an example of a “matrix-variate” random walk. MSV models with factor structure have been proposed by Aguilar and West (2000), Chib, Nardari, and Shephard (2005), and Han (2002). The common feature of these models is that the covariance matrix of the factors is assumed diagonal and the factor variances each follow their independent stochastic process. When the factors are not observable, model estimation requires specific restrictions (Geweke and Zhou (1996), Aguilar and West (2000)). Earlier factor SV specifications are extended in Aguilar and West (2000) by adding a time-varying vector of means. This model deals effectively with issues of overparameterization as well as of identifying the latent factors using the approach suggested by Geweke and Zhou (1996). A detailed survey of early models can be found in Shephard (1994). A survey of recent developments is provided in McAleer (2005).

This paper proposes a factor MSV model that assumes that the factor volatilities follow an unconstrained Wishart random process. Our model has close ties to Philipov and Glickman (2005), which proposes a class of MSV models where the volatilities, unconditional on any factors, follow a Wishart process. Recent work by Gouriéroux, Jasiak, and Sufana (2004) develops an alternative Wishart random process for modeling returns derived from sample second moments on vector-valued Gaussian autoregressive processes. Our model offers several advantages over existing factor multivariate models. First, the evolution of factor covariances is described by a general matrix-variate Wishart process. This modeling approach allows for interaction among the factors, their variances, and covariances, thus

giving more support to the empirical evidence on observed priced factors. Second, our factor model can be seen as a special case of the more general MSV model of Philipov and Glickman (2005). As such, it offers a natural framework for testing the asset-pricing restrictions imposed by factor-based models. We restrict our attention to observable factors, as this case allows the factor covariances to follow an unconstrained Wishart process. We discuss some of the difficulties incorporating unobserved factors later in the paper. Common variation in security returns is assumed to be driven by a set of observable factors whose covariance matrix follows a stochastic Wishart process. This factor MSV model avoids a number of restrictions common to other classes of MSV models, allowing for increased flexibility in modeling returns data.

We adopt a Bayesian framework for model development, and assume inference is carried out using Markov chain Monte Carlo (MCMC) simulation from the posterior distribution. Recent examples of MCMC simulation for SV models include Philipov and Glickman (2005), Aguilar and West (2000), and Jacquier, Polson, and Rossi (1995). Parameter summaries and forecasts are obtained using the empirical distribution of the samples obtained from MCMC posterior simulation.

The remainder of this paper is organized as follows. In Section 2, we introduce the factor multivariate stochastic volatility model and discuss the properties and interpretation of the model. In Section 3, we outline the steps to setting up MCMC posterior simulation. In Section 4, we examine the performance of the MCMC algorithm on simulated returns data, and apply our method to a data example involving 88 individual equity returns and

two factors – the market returns and the Fama-French HML factor. Finally, in Section 5, we discuss the utility of our approach, and limitations and extensions of the basic model presented.

2 The Factor Model

Suppose that we have n asset returns, \mathbf{y}_t , and k factors, \mathbf{f}_t . We assume that the asset returns follow a multivariate normal distribution, where

$$\mathbf{y}_t \mid \mathbf{B}, \mathbf{f}_t, \mathbf{\Omega} \sim N(\mathbf{B}\mathbf{f}_t, \mathbf{\Omega}) \quad (1)$$

$$\mathbf{f}_t \mid \mathbf{V}_t \sim N(\mathbf{0}, \mathbf{V}_t)$$

$$\mathbf{V}_t^{-1} \mid \mathbf{V}_{t-1}^{-1}, \mathbf{A}, \nu, d \sim Wish(\nu, \mathbf{S}_{t-1})$$

where

$$\mathbf{S}_{t-1} = \frac{1}{\nu} \mathbf{A}^{\frac{1}{2}} (\mathbf{V}_{t-1}^{-1})^d \mathbf{A}^{\frac{1}{2}'} \quad (2)$$

is the scale parameter for the Wishart process. In our model specification, \mathbf{V}_t is a matrix of factor volatilities, \mathbf{A} is a symmetric positive definite matrix, d is a scalar persistence parameter, ν is the degrees of freedom parameter for the Wishart distribution of the factor volatilities, \mathbf{B} is a $n \times k$ matrix of factor sensitivities, and $\mathbf{\Omega}$ is a $n \times n$ matrix of idiosyncratic risks for the assets. One feature of our model is that the conditional expectation of \mathbf{V}_t^{-1} is given by

$$E(\mathbf{V}_t^{-1} \mid \mathbf{V}_{t-1}^{-1}) = \nu \mathbf{S}_{t-1} = \left(\mathbf{A}^{\frac{1}{2}} \right) (\mathbf{V}_{t-1}^{-1})^d \left(\mathbf{A}^{\frac{1}{2}} \right)'. \quad (3)$$

In contrast, the conditional mean of the Wishart process in [Gourieroux, Jasiak, and Sufana \(2004\)](#) is an affine function of the previous period volatility matrix. An interesting case of our model results when $d = 1$ and \mathbf{A} is the identity matrix, in which case the model for factor volatilities is a Wishart random walk. In general, we can regard the factor model as a special case of the model proposed by [Philipov and Glickman \(2005\)](#), which imposes no restrictions on the covariance matrix. This model imposes specific restrictions on the covariance matrix of returns, namely it separates unconditional volatility, Σ_t , into systematic and idiosyncratic components,

$$\Sigma_t = \mathbf{B}\mathbf{V}_t\mathbf{B}' + \Omega. \tag{4}$$

Interest often centers on making inferences for the Σ_t , so it is useful to separate the components of variation in the volatility matrix.

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 in large part determines mean reversion characteristics on a multivariate level. The scalar parameter d denotes the overall strength of the intertemporal relationships. This parameter accounts 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 (see [Engle and Patton \(2001\)](#)). The parameter d is theoretically bound between -1 and 1 (see [Philipov and Glickman \(2005\)](#)) to ensure stationarity of the covariance process. For practical purposes, values of d that are applicable to financial data are between 0 and 1, as

we would rule out stochastic processes on the \mathbf{V}_t which alternate between powers of inverses. A value of d 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. Under this model specification, we do not need to impose constraints to ensure positive definiteness (as in early multivariate models) as this property is inherent in the quadratic form in (3) and in the Wishart process.

3 Model Fitting

Inference for our model is based on a standard Bayesian analysis. It is difficult to fit the model in Section 2 using exact methods. Instead, we use MCMC to simulate parameters from the joint posterior distribution. We summarize features of the posterior distribution using the simulated values. We can also forecast future returns using the distribution of simulated parameter values. In order to do so, we create a Gibbs sampler, which consists of drawing simulated values from the conditional posterior densities of the parameters \mathbf{V}^{-1} , \mathbf{A}^{-1} , \mathbf{B} , $\mathbf{\Omega}$, d , and ν . For increased efficiency, we draw simulated values from the conditional posterior distributions of \mathbf{B} and $\mathbf{\Omega}$ simultaneously. In this section we outline the strategy for setting up the MCMC algorithm.

3.1 Choice of Priors

We construct a noninformative prior distribution for the parameter set $(\mathbf{\Omega}, \mathbf{B}, \mathbf{A}, d, \nu)$. that factors into independent densities. We examine the prior component for each parameter

separately. Following Gelman, Carlin, Stern, and Rubin (1995, p.408) and Box and Tiao (1992, p.425), we suggest a prior density for $\mathbf{\Omega}$ proportional to the inverse of the determinant:

$$p(\mathbf{\Omega}) \propto |\mathbf{\Omega}|^{-\frac{1}{2}(n+1)} \quad (5)$$

For a more tractable analysis, we express the prior density for \mathbf{A} in terms of \mathbf{A}^{-1} . A Wishart density for \mathbf{A}^{-1} is assumed with a $k \times k$ scale matrix \mathbf{Q}_0 , where $\mathbf{Q}_0 = \mathbf{I}$, and a degrees of freedom parameter, $\gamma_0 = k+1$. For the parameter d , a uniform density such as $p(d) = U_{[0,1]}(d)$ may be used. For ν a locally noninformative gamma density seems appropriate. Because ν must 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 joint prior density for the parameters of the model would be the product of the densities specified above,

$$p(\mathbf{\Omega}, \mathbf{A}^{-1}, d, \nu) = p(\mathbf{B})p(\mathbf{\Omega})Wishart(\gamma_0, \mathbf{Q}_0)p(d)Gamma(\nu - k) \quad (6)$$

3.2 Posterior Summaries via MCMC

Let $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_T)$ be the collection of N asset returns across all time periods, and $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_T)$ represent the collection of k factors across all time periods. By Bayes theorem, the posterior distribution of the parameter set $(\mathbf{V}^{-1}, \mathbf{A}^{-1}, \mathbf{B}, \mathbf{\Omega}, d, \nu)$, is proportional to the product of the prior and the likelihood and is given by:

$$\begin{aligned}
p(\mathbf{V}^{-1}, \mathbf{A}, \mathbf{B}, \boldsymbol{\Omega}, \nu, d \mid \mathbf{F}, \mathbf{Y}) &\propto p(\mathbf{A}^{-1}, \mathbf{B}, d, \nu) L(\mathbf{V}^{-1}, \mathbf{A}, \mathbf{B}, \nu, d \mid \mathbf{Y}, \mathbf{F}) \\
&= p(\boldsymbol{\Omega}) \text{Wishart}(\mathbf{A}^{-1} \mid \gamma_0, \mathbf{Q}_0) p(d) \text{Gamma}(\nu - k) \\
&\quad \times \prod_{t=1}^T \text{Wishart}_k(\mathbf{V}_t^{-1} \mid \nu, \mathbf{S}_{t-1}) N_n(\mathbf{y}_t \mid \mathbf{B}\mathbf{f}_t, \boldsymbol{\Omega}) N_k(\mathbf{f}_t \mid \mathbf{0}, \mathbf{V}_t) \quad (7)
\end{aligned}$$

Direct sampling from this complex joint posterior distribution is infeasible. Therefore, we use MCMC methods, constructing a Gibbs sampler. The Gibbs sampler sequentially draws each parameter from its conditional posterior distribution (see Gelman, Carlin, Stern, and Rubin (1995)). The Gibbs sampler is described by the following steps:

1. Initialize the parameters at appropriate starting values.
2. Draw \mathbf{B} and $\boldsymbol{\Omega}$ simultaneously from their conditional posterior distribution, $p(\mathbf{B}, \boldsymbol{\Omega} \mid \mathbf{V}^{-1}, \mathbf{A}, \nu, d, \mathbf{Y}, \mathbf{F})$.
3. Draw \mathbf{A}^{-1} from its conditional posterior distribution, which is a Wishart distribution and allows direct sampling.
4. Draw \mathbf{V}_t^{-1} for each time period t . For periods $t = 1 \dots T - 1$ the conditional posterior distribution of \mathbf{V}_t^{-1} is a non-standard distribution. We use the Metropolis algorithm to draw \mathbf{V}_t^{-1} for periods 1 through $T - 1$. For the last time period, the conditional posterior for \mathbf{V}_T^{-1} is a Wishart distribution from which we sample directly.
5. Sample ν from its conditional posterior distribution using a discretized range of values using grid sampling (see Liu (2001)).

6. Sample d from its conditional posterior distribution using grid sampling.

After a sufficient number of iterations, the simulated parameters will be draws from the posterior distribution. Detailed description of the Gibbs sampler is presented in the appendix.

4 Results

We first evaluate our model on simulated data to assess convergence properties. The simulation analysis, which involves two observed factors and five asset returns, reveals that the true parameters are estimated without bias and after speedy convergence. We then apply our model to actual monthly returns of individual stocks using the two Fama and French (1993) factors. The data application reveals that the factors exhibit significant time-varying correlations and volatilities. In addition, inference provided by the MCMC simulation reveals that the value factor is not significant in our analysis based on posterior confidence intervals. Finally, we perform an out-of-sample portfolio application and assess the out-of-sample risk of the global minimum variance portfolios based on different model specifications.

4.1 Simulation Results

The simulation analysis consists of two stages. First, known fixed values are determined for \mathbf{A} , d , ν , $\mathbf{\Omega}$, \mathbf{B} , and for the initial factor covariance matrix, \mathbf{V}_0 . These values represent the true model parameters. Using the true parameter values, covariance matrices are sampled once for each time period, according to the model. Using the true covariance parameters, we

generate the sample of returns data. The second stage of the simulation analysis estimates the parameters using the data generated in stage one. The convergence of the Gibbs sampler is assessed by observing convergence plots of particular parameters. We use Gibbs samples from fifty different trials. In each trial the Gibbs sampler was run over 1000 iterations, and then the values of the last draws of the Gibbs sampler was collected. The estimates are the mean values of the draws over the 50 trials of the simulation experiment. The results from estimating the model parameters in the simulation analysis are presented in Table 1. Figures 1 and 2 present graphical evidence of the accuracy of the estimation algorithm to back out the true model values. We can observe that the drawn values of the log-determinants of the intertemporal sensitivity matrix \mathbf{A} and the idiosyncratic risk matrix $\mathbf{\Omega}$ are centered around the true model values. Their posterior distributions are symmetric and bell-shaped. Figure 3 shows the mean of the sensitivities matrix for each draw for all 50000 draws of the simulation study along with the true model parameter. Figure 4 shows all Gibbs sampler draws from all 50 trials for each of the elements of the factor sensitivities matrix centered around the true model parameter values.

Overall, the simulation study shows that the factor volatility model parameters estimates are unbiased and efficiently estimated. The simulation study reveals sensitivity between the matrix parameter \mathbf{A} and the persistence parameter d . The sensitivity remains after imposing a diagonality constraint on \mathbf{A} , which significantly reduces the number of model parameters. The interaction between \mathbf{A} and d in the Gibbs sampler could prevent them from converging to their true values. However, the bias in one of the parameters is com-

compensated by opposite bias in the other parameter, resulting in an unbiased estimate of the conditional expectation of the covariance matrix in time t . The Metropolis algorithm, which is used to draw the covariance matrix estimates only, takes the most CPU time. Hence, the Metropolis efficiency is very important for the overall Gibbs sampler performance. Since we use the same Metropolis algorithm for drawing the factor covariance matrices as in Philipov and Glickman (2005), we point out that analysis performed in that study documents that acceptance ratios from data application of a five-variate model average 16%. Acceptance ratios decrease fast with the increase in model dimensions. For a 12-variate model, the average acceptance ratio is 0.5%.

4.2 Empirical Results

The factor MSV model was implemented on return series data of 88 individual companies from the S&P500. The data set contains 324 monthly observations, from January 1975 to December 2001. The two observable factors in the model are the Fama French market excess return and the value factors. The value factor represents the difference between growth stocks and value stocks, which are defined as high book-to-value and low book-to-value stocks, respectively.

Figure 5 shows the first 10000 draws of the Gibbs sampler in the burn-in stage for the log-determinants of \mathbf{A} and $\mathbf{\Omega}$ and the mean of \mathbf{B} . The initial values of \mathbf{A} are far from the convergence values which helps observe the convergence of the parameter. Table 2 presents the estimates for the intertemporal sensitivity matrix \mathbf{A} . It is interesting to observe that

the two factor volatilities are strongly positively auto-correlated. However there is very little cross-correlation in the volatilities. Indeed the factors are constructed so that they are uncorrelated. However, diagonality in \mathbf{A} does not restrict the volatility matrix to be diagonal, it only implies that the volatilities are unconditionally uncorrelated.

The simulation analysis revealed that the values of the persistence parameter d , drawn from the derived conditional posterior distribution (31), are sensitive to the draws of \mathbf{A} . In our data application we obtained reasonable estimates of the persistence parameter but concern still remains that it may not be centered on the true value. Such bias may affect the inferences we obtain about the intertemporal sensitivity relationships among the factor volatilities, even though we will obtain accurate volatility and correlation estimates. We performed analysis fixing d at different values between 0 and 1. Based on our analysis, we found appropriate to fix the scalar persistence parameter at 0.5 as a satisfactory way to eliminate the sensitivity between \mathbf{A} and d revealed by the simulation study. This constraint does not affect the estimation of the conditional expectation of the factor covariance matrices. In the discussion section we propose an alternative approach to deal with the sensitivity between \mathbf{A} and d .

The estimated factor sensitivities and the factor variance parameters can best be summarized graphically due to their large numbers. Figures 6 and 7 show the mean estimates of the posterior samples for the factor sensitivities of all assets to the two factors, with 5 and 95% bounds. Our MCMC approach to estimating the factor sensitivities enables exact finite sample inference compared to the asymptotic approximations usually employed in

alternative estimation methods. We can observe in Figures 6 and 7 that the sensitivities to the market factor are significantly different from zero for most of the sample, while the book-to-market factor is not significant.

Figures 8 and 9 show the estimates of the standard deviations for the two factors derived from each draw of the factor covariance matrix in the Gibbs sampler. Figure 10 shows the correlations implied by the estimated covariance matrices \mathbf{V}_t over the whole sample period. Since our model does not restrict covariances or correlations, we can observe significant time variation in these parameters. The average sample correlation is close to zero. However, there is significant time variation in the factor correlations. They vary from -0.7 to about 0.6. Such unrestricted variation in the correlations is a feature of our model that is not captured by alternative volatility models. The standard deviations of the two factors also exhibit significant time variation. The standard deviation of the second factor has smaller values on average but exhibits large spikes.

4.3 Out-of-sample Optimal Portfolio Application

MSV models offer an improved representation of volatility dynamics and the potential for significant improvement in volatility forecasts. As mentioned earlier, MSV models are also very difficult to estimate due to high parameterization and complex estimation methodologies. Such difficulties may generate estimation errors which offset the benefits of improved model representation. To assess the forecasting capabilities of our model, we apply the model in an out-of-sample optimization application using a subsample of 30 stocks. The out-of-

sample period covers the period from January 1990 to January 2002. In the out-of-sample application we run our model at the beginning of each year, under two specifications of the covariance structure:

$$\Sigma_t = \mathbf{b}'\mathbf{V}_t\mathbf{b} + \Omega \quad (8)$$

$$\Sigma_t = \mathbf{b}'\mathbf{V}_t\mathbf{b} + \text{diag}(\Omega) \quad (9)$$

Model (8) is different from a unconstrained sample covariance matrix only in the time variation coming from the factor volatilities. Model (9) is consistent with empirical asset pricing models in which all covariation comes from the factors and the stocks' sensitivities to them. All time variation is due to the factor volatilities matrix. Since our factors are observable, we also have covariance effects through the factor covariation.

We compare the above formulation with the following benchmarks:

$$\Sigma = \mathbf{b}'\mathbf{V}\mathbf{b} + \Omega \quad (10)$$

$$\Sigma = \mathbf{b}'\mathbf{V}\mathbf{b} + \text{diag}(\Omega) \quad (11)$$

Model (10) is the unconditional sample covariance matrix. Model (11) is a traditional empirical pricing factor model. We estimate this model using OLS regression¹.

Using the estimated volatilities, we construct global minimum variance portfolios, using unconstrained optimization (see Campbell, Lo, and MacKinlay (1997, Ch. 5)). We rebalance

¹An OLS regression is still a standard tool in asset pricing applications. Popular studies applying OLS for asset pricing are Fama and MacBeth (1973), Fama and French (1993), and recently Brennan, Chordia, and Subrahmanyam (1998).

the optimal portfolios at the beginning of each year. At the end of the 12 year period we compute the realized monthly returns for the optimal minimum variance portfolios. We then compare the out-of-sample standard deviations of the realized optimal portfolio excess returns. Table 3 shows that from the two unconstrained models, 10 and 8, the portfolio returns based on our factor specifications have an out-of-sample standard deviation which is about 10% lower than the benchmark portfolios (0.0341 vs. 0.0382). For the two models with diagonality constraint on the idiosyncratic covariance matrix, the portfolio based on our factor specification has about 15% lower out-of-sample standard deviation than the benchmark portfolio. Figure 11 shows 12 annual standard deviations based on the monthly returns for the corresponding years. The standard deviations for our factor specifications are consistently lower than their corresponding benchmarks.

5 Discussion

The framework for our factor MSV model has some important benefits and a few challenges to make it a worthy paradigm for modeling multivariate sets of returns. In our empirical application of the factor MSV model, we use an empirical pricing framework. Such empirical pricing applications usually use a choice of observed factors that may not meet orthogonality requirements, and may not capture all common variation among assets. Indeed, in our application, we note that while the observed (not latent) factors appear to be unconditionally orthogonal, they have significant time-varying correlations. Our factor model separates the volatility into systematic and idiosyncratic components – a feature in empirical pricing

models² and portfolio management (see Bodie, Kane, and Marcus (2004) and Jacquier and Marcus (2001)).

One weakness in the estimation of the parameters of our model is the apparent partial aliasing between the intertemporal sensitivity parameter and the persistence parameter. Fitting simulated data with MCMC in which all model parameters are free resulted in large drifts across iterations of the simulated values of d . An alternative approach to fitting the model using MCMC is to choose a grid of values for the persistence parameter, and run separate Gibbs samplers conditional on each fixed value of the persistence parameter on the grid corresponding to an appropriate prior distribution. Inferences can then be averaged over the posterior weights on the grid.

While leverage is an important phenomenon to address in these type of models, the specification of our model does not explicitly capture leverage effects. Leverage effects in the individual securities will depend on the factor sensitivities, the volatilities of the factors, and on idiosyncratic volatilities. Since the factor sensitivities and idiosyncratic volatilities are time-invariant, leverage effects in the individual securities solely depend on leverage effects in the factors. We tested for the significance of leverage effects for the two factors (separately) using OLS regression (results not presented here), and found that the relationship between the factors and their volatilities was not statistically significant (though it was negative).

In this paper we examine the case of observable factors exclusively. Many of the popular factor SV models assume latent factors. In the case of unobservable factors, the model needs

²Theoretical pricing models focus exclusively on systematic risk as a determinant of risk-return trade-offs (e.g. the CAPM of Sharpe (1964), Lintner (1965) and APT of Ross (1976)).

to have imposed additional restrictions on the factor variance structure that eliminate some desirable current features of our model. For example, an orthogonality constraint on the latent factors would restrict both unconditional and time-varying correlations among the factors to zero and would limit the model to estimating a vector of variances, as is the case with popular factor SV models.

6 Conclusion

This paper presented the formulation, the estimation methodology, and data application results for a factor MSV model in which the time-varying factor volatilities matrix is driven by a Wishart process. The model is part of a unified framework, in which univariate and MSV model formulations are part of the same model. The factor model presented in this paper derives from a general MSV model based on Wishart processes which offers several advantages. First, it naturally extends scalar variances into covariance matrices rather than vectors of log-variances. Second, the Wishart distribution scale parameter, which underlies the conditional expectation of the asset covariance matrix, allows for both variances and correlations to evolve stochastically over time. Third, the general model allows for the conditional volatility of an asset to depend not only on its past volatility but also on past covariances with other assets, incorporating the observed contagion among asset returns into their covariance structure.

In the context of existing factor multivariate models, the current model offers an improved representation of the behavior of asset pricing factors. The factor structure allows

addressing high dimensional setups used in portfolio analysis and risk management, as well as modeling conditional means and conditional variances within the model framework. We show that it is possible to implement factor MSV models with high dimensionality, making them feasible for practical applications in risk management, asset allocation, and portfolio optimization.

APPENDIX

A.1 Results used in the Multivariate Factor Model

Result A.1.1 *Let $p(\mathbf{Z})$ be a Wishart(q, Θ) density function (Box and Tiao (1992)):*

$$p(\mathbf{Z}) \propto |\mathbf{Z}|^{\frac{1}{2}q-1} \exp(-\frac{1}{2}\text{tr}\mathbf{Z}\Theta) \quad (12)$$

after integrating \mathbf{Z} out we obtain the identity:

$$\int_{\mathbf{Z}>0} |\mathbf{Z}|^{\frac{1}{2}q-1} \exp(-\frac{1}{2}\text{tr}\mathbf{Z}\Theta) d\mathbf{Z} = |\Theta|^{-\frac{1}{2}(q+n-1)} 2^{\frac{1}{2}n(q+n-1)} \Gamma_n\left(\frac{(q+n-1)}{2}\right) \quad (13)$$

where $\Gamma_p(b)$ is the generalized gamma function (Siegel, 1935):

$$\Gamma_p(b) = [\Gamma(\frac{1}{2})]^{\frac{1}{2}p(p-1)} \prod_{\alpha=1}^p \Gamma\left(b + \frac{\alpha-1}{2}\right), \quad b > \frac{p-1}{2} \quad (14)$$

Result A.1.2 *A fundamental matrix identity:*

$$\left| \begin{array}{cc} K \times K & K \times L \\ \mathbf{I} & -\mathbf{P} \mathbf{Q} \end{array} \right| = \left| \begin{array}{cc} L \times L & L \times K \\ \mathbf{I} & -\mathbf{Q} \mathbf{P} \end{array} \right| \quad (15)$$

A.2 The Conditional Posterior Distribution for \mathbf{B} and Ω

To sample draws from the conditional posterior distribution of \mathbf{B} and Ω , we recognize that sampling the inverse, Ω^{-1} , is slightly more convenient,

$$p(\mathbf{B}, \Omega^{-1} \mid \Sigma^{-1}, \mathbf{A}, \nu, d, \mathbf{Y}, \mathbf{F}) \propto |\Omega|^{-\frac{1}{2}(n+1)} \prod_{t=1}^T \frac{1}{|\Omega|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{y}_t - \mathbf{B}\mathbf{f}_t)' \Omega^{-1} (\mathbf{y}_t - \mathbf{B}\mathbf{f}_t)\right) \quad (16)$$

First, we show that \mathbf{B} can be sampled from a distribution which is a matrix-variate extension of the multivariate Student t -distribution. Using the identity in Result 1 in the

appendix, and setting q equal to $(T - n + 1)$, we integrate out $\boldsymbol{\Omega}^{-1}$ in the conditional posterior density. We obtain the conditional posterior distribution of \mathbf{B} ,

$$p(\mathbf{B} \mid \mathbf{Y}, \mathbf{F}) = \int_{\boldsymbol{\Omega}^{-1} > 0} p(\mathbf{B}, \boldsymbol{\Omega}^{-1} \mid \mathbf{Y}, \mathbf{F}) d\boldsymbol{\Omega}^{-1} \propto |\mathbf{R}|^{-\frac{T}{2}} \quad (17)$$

where each element of the matrix \mathbf{R} is:

$$\mathbf{R}_{ij} = (\mathbf{y}_i - \hat{\mathbf{b}}_i \mathbf{F})' (\mathbf{y}_j - \hat{\mathbf{b}}_j \mathbf{F}) + (\mathbf{b}_i - \hat{\mathbf{b}}_i)' \mathbf{F}' \mathbf{F} (\mathbf{b}_j - \hat{\mathbf{b}}_j) \quad (18)$$

where $\hat{\mathbf{b}}_i = (\mathbf{F}\mathbf{F}')^{-1} \mathbf{F}\mathbf{y}'_i$ is a $k \times 1$ vector with the least squares estimates of \mathbf{b}_i for asset i .

Let:

$$\mathbf{D} = (\mathbf{Y} - \hat{\mathbf{B}}\mathbf{F})(\mathbf{Y} - \hat{\mathbf{B}}\mathbf{F})' \quad (19)$$

From the general result in the appendix, we obtain the posterior distribution of \mathbf{B}

$$p(\mathbf{B} \mid \mathbf{Y}, \mathbf{F}) \propto \left| \mathbf{D} + (\mathbf{B} - \hat{\mathbf{B}})\mathbf{F}\mathbf{F}'(\mathbf{B} - \hat{\mathbf{B}})' \right|^{-\frac{T}{2}} \quad (20)$$

We can represent each row of \mathbf{B} as a vector following a k -variate t -distribution. Thus, we can regard the general case (20) as a matrix-variate extension of the multivariate t -distribution:

$$p(\mathbf{B} \mid \mathbf{Y}, \mathbf{F}) \propto t_{nk}(\hat{\mathbf{B}}, \mathbf{D}^{-1}, (\mathbf{F}\mathbf{F}')^{-1}, T - (n + k) + 1) \quad (21)$$

The density function for the matrix-variate extension (21) is:

$$p(\mathbf{B} \mid \mathbf{Y}, \mathbf{F}) = C(n, k, \nu) |\mathbf{F}\mathbf{F}'|^{\frac{T}{2}} |\mathbf{D}^{-1}|^{\frac{k}{2}} \left| \mathbf{I}_n + \mathbf{D}^{-1}(\mathbf{B} - \hat{\mathbf{B}})\mathbf{F}\mathbf{F}'(\mathbf{B} - \hat{\mathbf{B}})' \right|^{-\frac{1}{2}(\nu + k + n - 1)} \quad (22)$$

where

$\nu = T - (n + k) + 1$ is the degrees of freedom parameter, and

$C(n, k, \nu) = \left[\Gamma\left(\frac{1}{2}\right)^{nk} \frac{\Gamma_n\left(\frac{1}{2}(\nu+n-1)\right)}{\Gamma_n\left(\frac{1}{2}(\nu+k+m-1)\right)} \right]^{-1}$, where $\Gamma_n(x)$ is the generalized Gamma function.

We next represent the distribution of \mathbf{B} as a product of conditional densities. Create two partitions of \mathbf{B} with n_1 and n_2 rows, $n_1 + n_2 = n$.

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \text{---} \\ \mathbf{B}_2 \end{pmatrix} \quad \hat{\mathbf{B}} = \begin{pmatrix} \hat{\mathbf{B}}_1 \\ \text{---} \\ \hat{\mathbf{B}}_2 \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \text{---} & \text{---} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{pmatrix} \quad (23)$$

We obtain the distribution of \mathbf{B}_2 conditional on \mathbf{B}_1 :

$$\begin{aligned} p(\mathbf{B}_2 \mid \mathbf{B}_1, \mathbf{Y}) &\propto C(n_2, k, \nu + n_1) |\mathbf{H}|^{\frac{n_1}{2}} |\mathbf{D}_{22}^*|^{\frac{k}{2}} \\ &\times \left| \mathbf{I}_{n_2} + \mathbf{D}_{22}^{-1}(\mathbf{B}_2 - \tilde{\mathbf{b}}_2)\mathbf{H}(\mathbf{B}_2 - \tilde{\mathbf{B}}_2)' \right|^{-\frac{1}{2}[(\nu+n_1)+k+n_2-1]} \end{aligned} \quad (24)$$

where $\mathbf{H}^{-1} = (\mathbf{F}\mathbf{F}')^{-1} + (\mathbf{B}_1 - \hat{\mathbf{B}}_1)'\mathbf{D}_{11}^{-1}(\mathbf{B}_1 - \hat{\mathbf{B}}_1)$ We obtain the marginal density of \mathbf{B}_1 :

$$p(\mathbf{B}_1 \mid \mathbf{Y}) \propto \left| \mathbf{I}_{n_1} + \mathbf{D}_{11}^{-1}(\mathbf{B}_1 - \hat{\mathbf{B}}_1)(\mathbf{F}\mathbf{F}')(\mathbf{B}_1 - \hat{\mathbf{B}}_1)' \right|^{-\frac{1}{2}[\nu+n_1+k-1]} \quad (25)$$

Alternatively, if we set $n_1 = n - 1$ and $n_2 = 1$, then the conditional distribution of $\mathbf{b}_2 = \mathbf{b}_n$ conditional on $\mathbf{B}'_1 = [\mathbf{b}_1, \dots, \mathbf{b}_{n-1}]$ is:

$$\mathbf{b}_n \sim t_k \left(\tilde{\mathbf{b}}_n, (\nu + n - 1)^{-1} d_{nn}^* \mathbf{H}^{-1}, \nu + n - 1 \right) \quad (26)$$

Therefore, if we partition \mathbf{B} ($n - 1$) times we obtain the result that the distribution of \mathbf{B} is a product of multivariate t -distributions,

$$p(\mathbf{B} \mid \mathbf{Y}) = p(\mathbf{b}_1 \mid \mathbf{Y})p(\mathbf{b}_2 \mid \mathbf{b}_1, \mathbf{Y}) \dots p(\mathbf{b}_n \mid \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-1}, \mathbf{Y}) \quad (27)$$

Given \mathbf{B} , sampling $\mathbf{\Omega}^{-1}$ is straightforward. From (??), we obtain

$$\mathbf{\Omega}^{-1} \sim p(\mathbf{\Omega}^{-1} \mid \mathbf{B}, \mathbf{Y}, \mathbf{F}) = \text{Wishart}_n(T, \mathbf{R}^{-1}) \quad (28)$$

where the degrees of freedom parameter is equal to the number of time observations, T , and the scale parameter \mathbf{R}^{-1} is a matrix function of \mathbf{B} which is defined in (??).

A.3 Conditional Posterior Distribution of \mathbf{V}_t

For periods $t = 1, 2, \dots, T - 1$ the conditional posterior distribution is proportional to:

$$\begin{aligned}
p(\mathbf{V}_t^{-1} \mid rest) &\propto \text{Wish}(\mathbf{V}_t^{-1} \mid \nu, \mathbf{S}_{t-1}) \times N(\mathbf{0}, \mathbf{V}_t) \times \text{Wish}(\mathbf{V}_{t+1}^{-1} \mid \nu, \mathbf{S}_t) \\
&\propto \exp(-\frac{1}{2}\text{tr}[\mathbf{S}_{t-1}^{-1}\mathbf{V}_t^{-1}]) |\mathbf{V}_t^{-1}|^{\frac{\nu(1-d)-k}{2}} \exp(-\frac{1}{2}\mathbf{y}'_t\mathbf{V}_t^{-1}\mathbf{y}_t) \exp(-\frac{1}{2}\text{tr}[\mathbf{S}_t^{-1}\mathbf{V}_{t+1}^{-1}]) \\
&\propto \text{Wish}(\mathbf{V}_t^{-1} \mid \tilde{\nu}, \tilde{\mathbf{S}}_{t-1}) \times f(\mathbf{V}_t^{-1})
\end{aligned} \tag{29}$$

where, $\tilde{\nu} = \nu(1 - d) + 1$, $\tilde{\mathbf{S}}_{t-1} = \mathbf{S}_{t-1}^{-1} + \mathbf{y}_t\mathbf{y}'_t$, and $f(\mathbf{V}_t^{-1}) = |\mathbf{V}_t^{-1}|^{\frac{1-\nu d}{2}} \exp(-\frac{1}{2}\text{tr}[\mathbf{S}_t^{-1}, \mathbf{V}_{t+1}^{-1}])$ is a residual term. We assume a known initial period covariance matrix \mathbf{V}_0^{-1} .

A.4 Conditional posterior distribution of \mathbf{A}^{-1}

We obtain a conditional distribution for \mathbf{A}^{-1} that is a product of a *Wishart* prior and a *Wishart* likelihood, and is also a *Wishart* distribution:

$$\begin{aligned}
p(\mathbf{A}^{-1} \mid rest) &\propto \text{Wish}(\gamma_0, \mathbf{Q}_0) \times \text{Wish}(\gamma, \mathbf{Q}) \\
&\propto |A^{-1}|^{\frac{\gamma_0-k-1}{2}} \exp(-\frac{1}{2}\text{tr}(\mathbf{Q}_0^{-1}\mathbf{A}^{-1})) |A^{-1}|^{-\frac{T\nu}{2}} \exp(-\frac{1}{2}\text{tr}(\nu\mathbf{Q}^{-1}\mathbf{A}^{-1})) \\
&\propto \text{Wish}(\mathbf{A}^{-1} \mid \tilde{\gamma}, \tilde{\mathbf{Q}})
\end{aligned} \tag{30}$$

where $\tilde{\mathbf{Q}} = \tilde{\mathbf{Q}}(\mathbf{V}^{-1}, d, \nu)$ is the scale parameter, a function of \mathbf{V}^{-1}, d, ν , $\tilde{\mathbf{Q}}^{-1} = \nu\mathbf{Q}^{-1} + \mathbf{Q}_0^{-1}$, $\mathbf{Q}^{-1} = \sum_{t=1}^T (\mathbf{V}_t^{-1})^{-\frac{d}{2}} \mathbf{V}_t^{-1} (\mathbf{V}_{t-1}^{-1})^{-\frac{d}{2}}$, $\tilde{\gamma} = \gamma + \gamma_0 - k - 1 = T\nu + \gamma_0$, $\gamma = T\nu + k + 1$ (we use the

$|A^{-1}|^{-\frac{T\nu}{2}}$ term in the likelihood above to solve for γ). We are able to sample directly from the conditional distribution of \mathbf{A}^{-1} as a step in the Gibbs sampler.

A.5 Conditional posterior distributions of d and ν

The conditional distribution of d is:

$$p(d \mid rest) \propto \exp(d\psi - \frac{1}{2}tr[\nu\mathbf{A}^{-1}\mathbf{Q}(d)^{-1}]) \quad (31)$$

where $\psi = -\frac{\nu}{2} \sum_{t=1}^T \ln(|\boldsymbol{\Sigma}_t^{-1}|)$, and \mathbf{Q}^{-1} is the parameter defined in the previous section.

While d was treated as a fixed parameter in the previous section, now it is a argument of the function \mathbf{Q}^{-1} .

The conditional distribution of ν is:

$$p(\nu \mid rest) \propto e^{(\alpha-1)\log(\nu-k)-\beta(\nu-k)} \left(\frac{|\nu\mathbf{A}^{-1}|^{\frac{\nu}{2}}}{2^{\nu k} \prod_{j=1}^k \Gamma(\frac{\nu+j-1}{2})} \right)^T \quad (32)$$

$$\times \prod_{t=1}^T |\mathbf{Q}_t^{-1}|^{\frac{\nu}{2}} \prod_{t=1}^T \exp(-\frac{\nu}{2}tr[\mathbf{A}^{-1}\mathbf{Q}_t^{-1}])$$

References

- Aguilar, Omar, and Mike West, 2000, Bayesian Dynamic Factor Models and Portfolio Allocation, *Journal of Business and Economic Statistics* 18(3), 338–357.
- Bodie, Zvi, Alex Kane, and Alan J. Marcus, 2004, *Investments*. (Irwin/McGraw-Hill New York) 6th edn.
- Bollerslev, Tim, Robert F. Engle, and Jeffrey M. Wooldridge, 1988, A Capital Asset Pricing Model with Time Varying Covariances, *Journal of Political Economy* 96(1), 116–131.
- Box, George E.P., and George C. Tiao, 1992, *Bayesian Inference in Statistical Analysis*. (John Wiley and Sons, New York) Wiley Classics Library edn.
- Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns, *Journal of Financial Economics* 49, 345–373.
- Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*. (Princeton University Press Princeton, NJ).
- Chib, Siddhartha, Federico Nardari, and Neil Shephard, 2005, Analysis of High Dimensional Multivariate Stochastic Volatility Models, Forthcoming, *Journal of Econometrics*.
- Engle, Robert F., and Dennis F. Kraft, 1983, Multiperiod Forecast Error Variances of Inflation Estimated from ARCH Models, in Zellner, eds.: *Applied Time Series Analysis of Economic Data* (Bureau of the Census, Washington, D.C.).
- Engle, Robert F., and Kenneth F. Kroner, 1995, Multivariate Simultaneous GARCH, *Econometric Theory* 11, 122–150.
- Engle, Robert F., Victor K. Ng, and Michael Rothschild, 1990, Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates For Treasury Bills, *Journal of Econometrics* 45, 213–238.
- Engle, Robert F., and Andrew J. Patton, 2001, What Good Is A Volatility Model, Unpublished Paper, Department of Economics, University of California, San Diego.
- Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3–56.
- Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607–636.
- Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin, 1995, *Bayesian Data Analysis*. (Chapman & Hall, London) 1st edn.

- Geweke, John, and Guofu Zhou, 1996, Measuring the Pricing Error of the Arbitrage Pricing Theory, *Review of Financial Studies* 9, 557–587.
- Gourieroux, Christian, Joann Jasiak, and Razvan D. Sufana, 2004, The Wishart Autoregressive Process of Multivariate Stochastic Volatility, Unpublished Paper, Department of Economics, University of Toronto.
- Han, Y., 2002, The Economic Value of Volatility Modeling: Asset Allocation with a High Dimensional Dynamic Latent Factor Multivariate Stochastic Volatility Model, Unpublished Paper, Olin School of Business, Washington University in St. Louis.
- Harvey, A., E. Ruiz, and N. Shephard, 1994, Multivariate Stochastic Variance Models, *Review of Economic Studies* 61, 247–264.
- Jacquier, Eric, and Alan J. Marcus, 2001, Asset Allocation Models and Market Volatility, *Financial Analysts Journal* 57(2), 16–30.
- Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi, 1995, Models and Prior Distributions for Multivariate Stochastic Volatility, *Srie scientifique = Scientific series*,; no. 95s-18; Montral: CIRANO.
- Lintner, John, 1965, The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics* 47, 13–37.
- Liu, Jun S., 2001, *Monte Carlo Strategies in Scientific Computing*. (Springer-Verlag New York, LLC New York).
- Mahieu, Ronald, and Peter Schotman, 1994, Stochastic Volatility and the Distribution of Exchange Rate News, Discussion Paper / Institute for Empirical Macroeconomics 96, Federal Reserve Bank of Minneapolis.
- McAleer, Michael, 2005, Automatic Inference and Learning in Modeling Financial Volatility, *Econometric Theory* 21, 232–261.
- Philipov, Alexander, and Mark E. Glickman, 2005, Multivariate Stochastic Volatility via Wishart Processes, Forthcoming, *Journal of Business and Economic Statistics*.
- Ross, Stephen A., 1976, The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory* 13, 341–360.
- Sharpe, William F., 1964, Capital Asset Prices: A Theory of Market Equilibrium, *Journal of Finance* 19, 425–442.
- Shephard, Neil, 1994, Partial Non-Gaussian State Space, *Biometrika* 81(1), 115–131.
- Uhlig, Harald, 1997, Bayesian Vector Autoregression with Stochastic Volatility, *Econometrica* 65, 59–73.

TABLE 1
Simulation Results

The table presents the true predetermined values of the parameters in the model that describe the data generating process, and the estimated parameters using the Bayesian estimation procedures. The estimates are based on Gibbs samples from fifty different trials. In each trial the Gibbs sampler was run over 1000 iterations, and then the values of the last draws of the Gibbs sampler was collected. The estimates are the mean values of the last draws over the 50 trials of the simulation experiment.

A. True pre-determined values for the model used in the generation stage

	$d = 0.5 \quad \nu = 10$				
\mathbf{V}_0	1.20	-0.50			
	-0.50	0.80			
\mathbf{B}_{model}	1.50	-0.50			
	1.00	0.50			
	-1.20	-0.50			
	2.00	0.75			
	3.75	-7.13			
$\mathbf{\Omega}_{model}$	4.83	-1.52	-0.62	2.02	0.22
	-1.52	6.16	-0.26	-2.17	0.38
	-0.62	-0.26	5.08	1.07	-0.59
	2.02	-2.17	1.07	6.68	-1.24
	0.22	0.38	-0.59	-1.24	5.72

B. Estimated values for the model parameters

	d, ν fixed at true values				
\mathbf{V}_0	1.20	-0.50			
	-0.50	0.80			
$\hat{\mathbf{B}}$	1.64	-0.46			
	0.94	0.46			
	-1.18	-0.53			
	2.04	0.82			
	3.72	-7.25			
$\hat{\mathbf{\Omega}}$	5.07	-1.93	-0.61	2.05	0.81
	-1.93	6.21	0.52	-2.21	1.47
	-0.61	0.52	4.63	0.42	-0.60
	2.05	-2.21	0.42	4.71	-1.44
	0.81	1.47	-0.60	-1.44	6.59

TABLE 2
Empirical Results

The table presents the estimate of the intertemporal sensitivity matrix \mathbf{A} for the two-factor covariance structure over 324 time observations.

Factor 1	0.0382	
	[0.0355 0.0410]	
Factor 2	-0.0001	0.0179
	[-0.0015 0.0013]	[0.0166 0.0193]

TABLE 3
Out-of-sample portfolio application

This table summarizes the results from the out-of-sample optimal portfolio application. Four different portfolios were formed based on the specifications below:

$$\Sigma_t = \mathbf{b}'\mathbf{V}_t\mathbf{b} + \Omega \tag{1}$$

$$\Sigma_t = \mathbf{b}'\mathbf{V}_t\mathbf{b} + \text{diag}(\Omega) \tag{2}$$

$$\Sigma = \mathbf{b}'\mathbf{V}\mathbf{b} + \Omega \tag{3}$$

$$\Sigma = \mathbf{b}'\mathbf{V}\mathbf{b} + \text{diag}(\Omega) \tag{4}$$

The table shows the sample standard deviations of the realized monthly returns of the global minimum variance portfolios over the out-of-sample period under the four different model specifications.

Model No.	Out-of-sample standard deviation
Model 1	0.0341
Model 2	0.0389
Model 3	0.0382
Model 4	0.0451

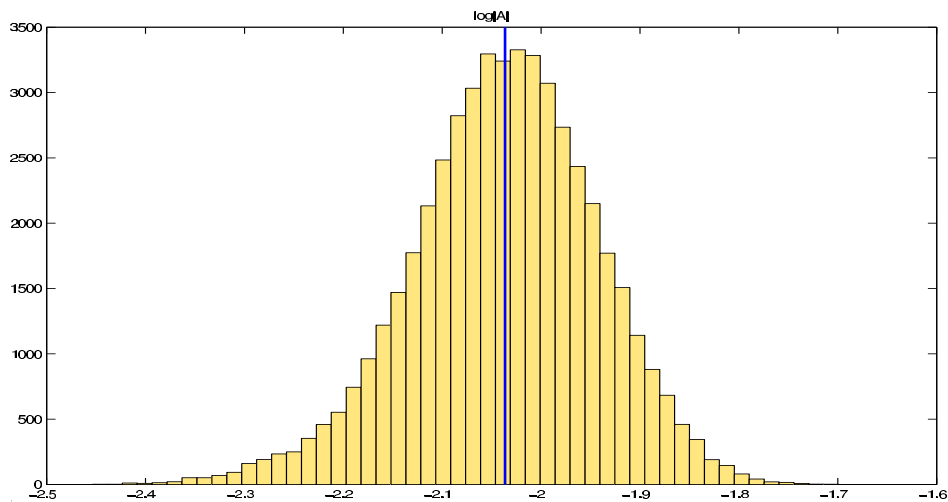


Figure 1. Histogram of the log of the determinant of \mathbf{A} over 50 000 draws in a simulation exercise. The vertical dark line represents the true model value for the log-determinant of \mathbf{A} . In the simulation exercise true model values were specified for the parameters, base on which the covariance matrices and data were generated. Using the generated covariance matrices and data, the model parameters were estimated and the estimates were compared to the true values to assess the precision and efficiency of the MCMC algorithm.

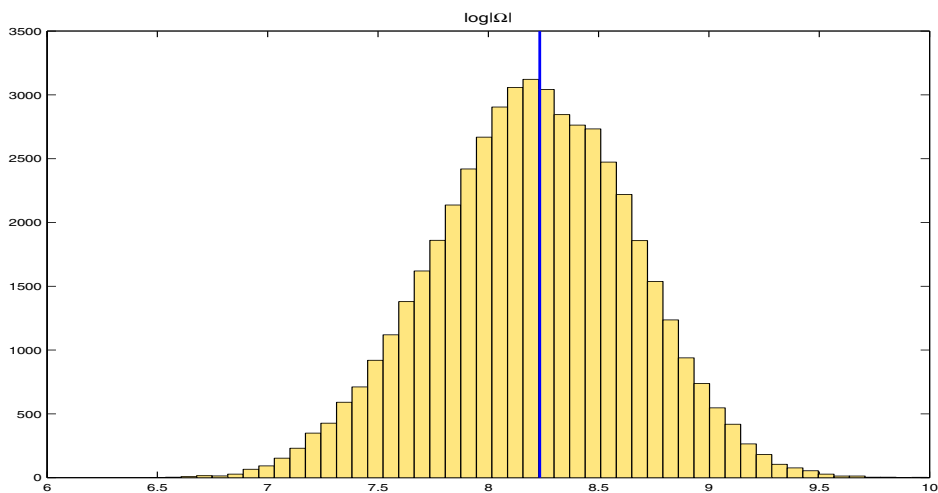


Figure 2. Histogram of the log of the determinant of $\mathbf{\Omega}$ over 50 000 draws in a simulation exercise. The vertical dark line represents the true model value for the log-determinant of $\mathbf{\Omega}$. In the simulation exercise true model values were specified for the parameters, base on which the covariance matrices and data were generated. Using the generated covariance matrices and data, the model parameters were estimated and the estimates were compared to the true values to assess the precision and efficiency of the MCMC algorithm..

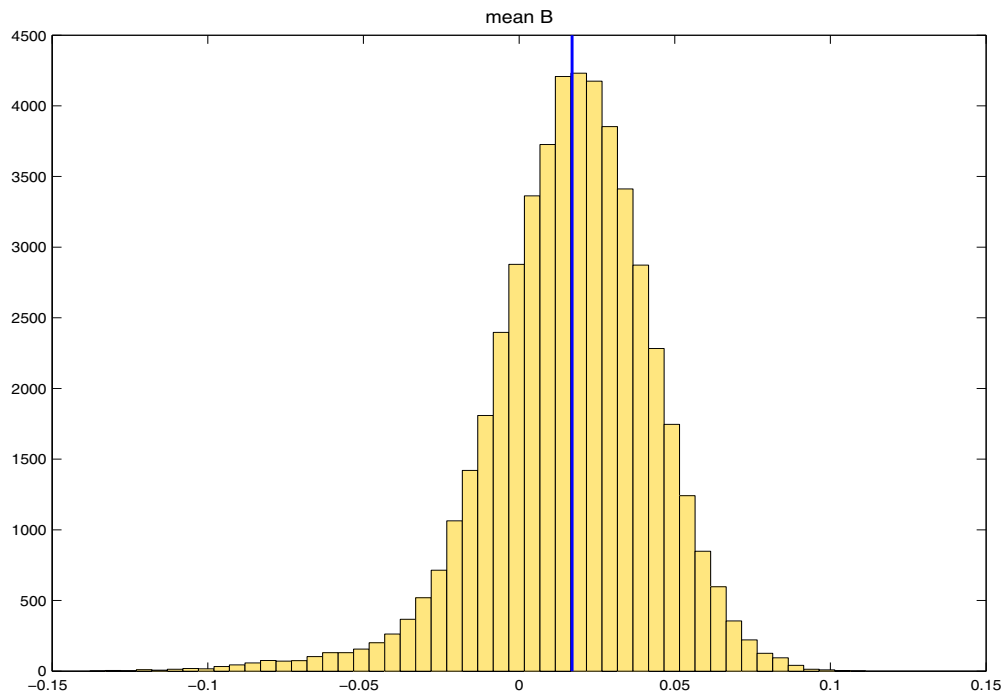


Figure 3. Histogram of the mean of all elements of the beta matrix \mathbf{B} over 50 000 draws in a simulation exercise. The vertical dark line represents the true model value for the mean \mathbf{B} . In the simulation exercise true model values were specified for the parameters, base on which the covariance matrices and data were generated. Using the generated covariance matrices and data, the model parameters were estimated and the estimates were compared to the true values to assess the precision and efficiency of the MCMC algorithm..

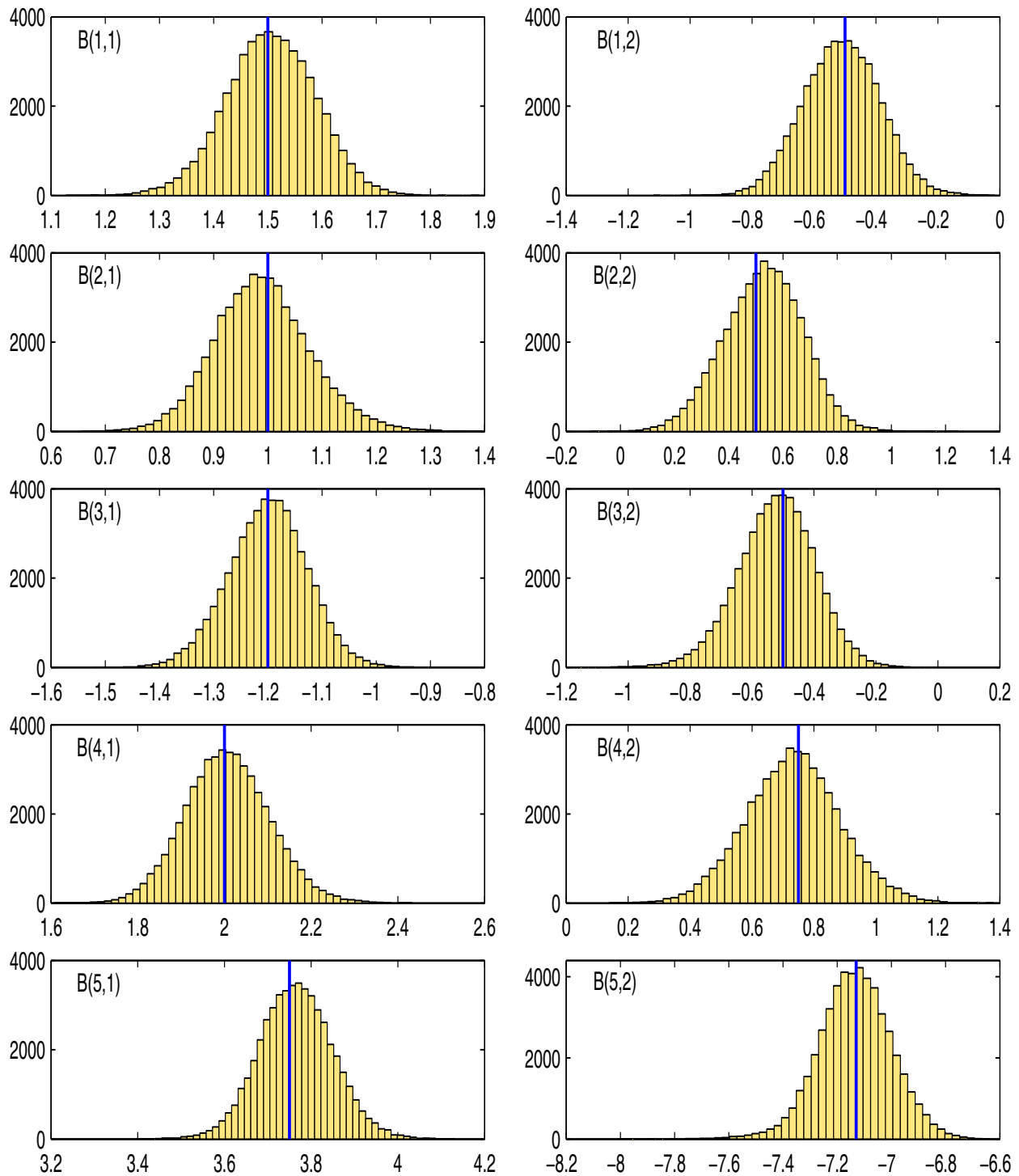


Figure 4. Histogram plots of the elements of the factor sensitivities matrix estimated in the simulation exercise, against the true model values, represented by straight horizontal lines.

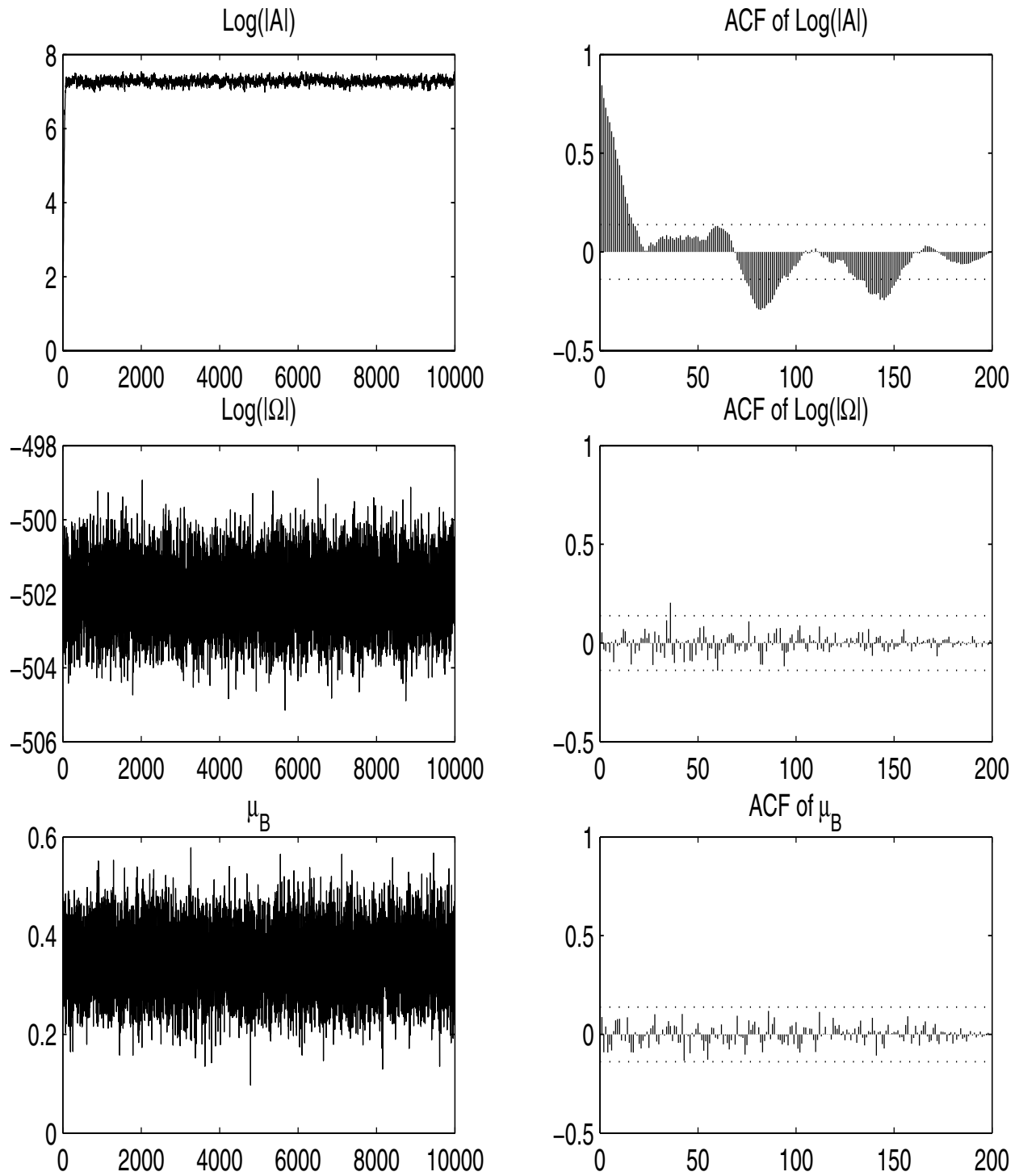


Figure 5. Plot of the log-determinant of the intertemporal sensitivity matrix \mathbf{A} , the idiosyncratic covariance matrix, Ω , and the mean of all betas, over 10000 iterations of the burn-in stage of the Gibbs sampler. Also plotted are ACF graphs of a subsample of subsequent draws of the Gibbs sampler.

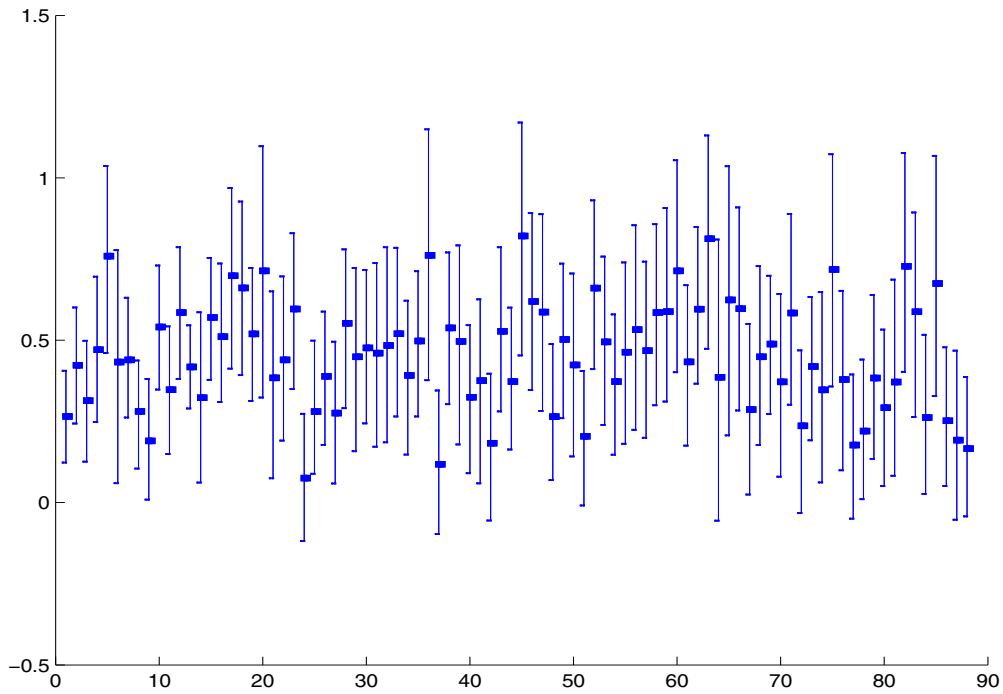


Figure 6. Box plot of the mean of the first factor beta for the 88 securities in the sample, with 5 and 95% bounds represented by vertical lines.

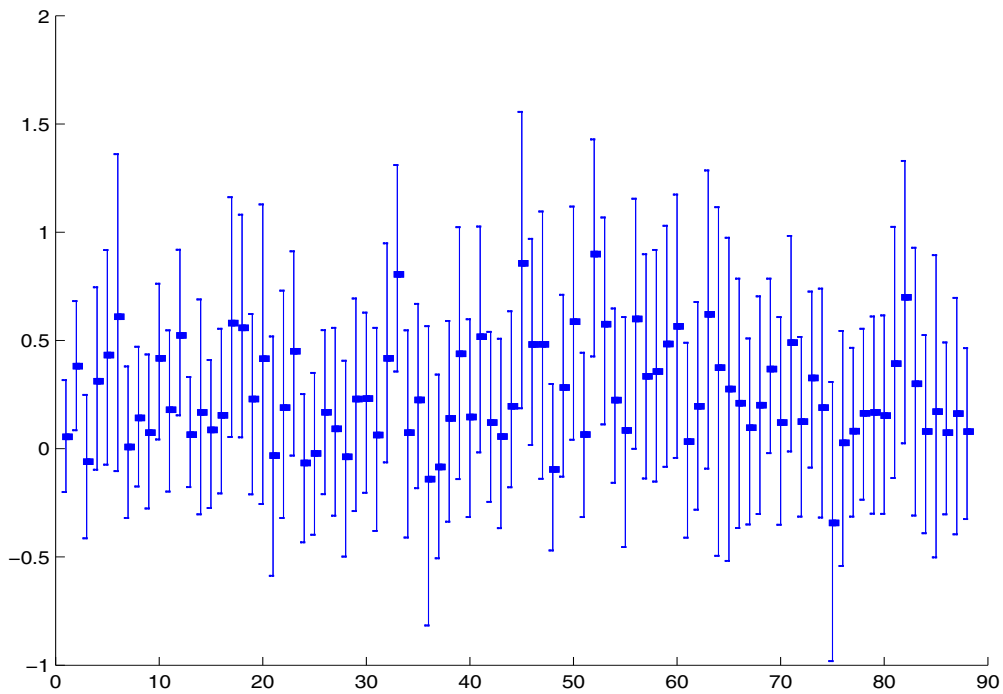


Figure 7. Box plot of the mean of the second factor beta for the 88 securities in the sample, with 5 and 95% bounds represented by vertical lines.

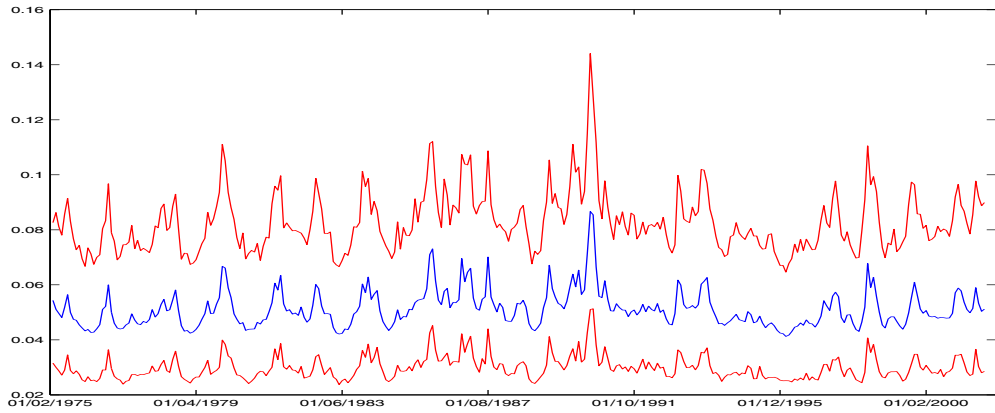


Figure 8. Time series plot of the posterior distribution of the standard deviations of the first for all sample periods. We show 5th percentile, mean, and 95th percentile.

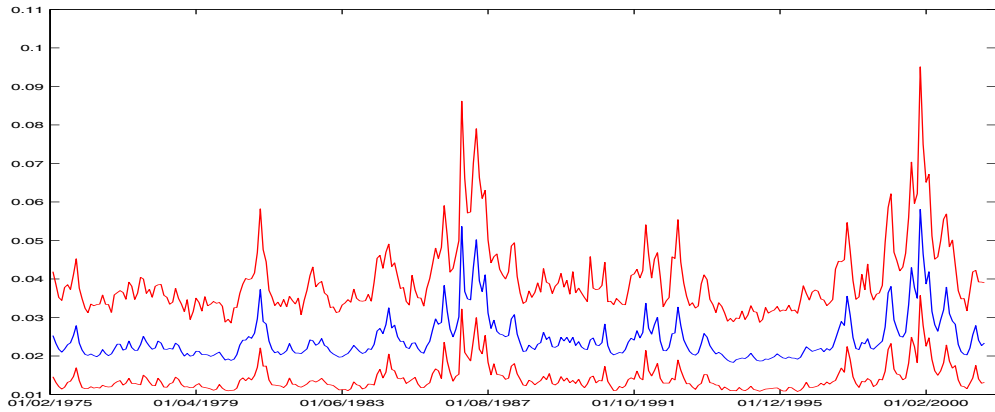


Figure 9. Time series plot of the posterior distribution of the standard deviations of second factor for all sample periods. We show 5th percentile, mean, and 95th percentile.

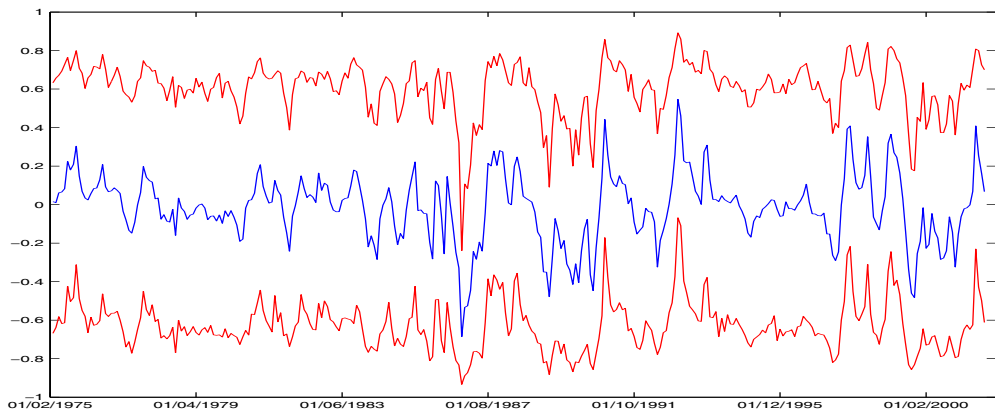


Figure 10. Plot of the correlations between the two factors for all periods in the sample. We show 5th percentile, mean, and 95th percentile.

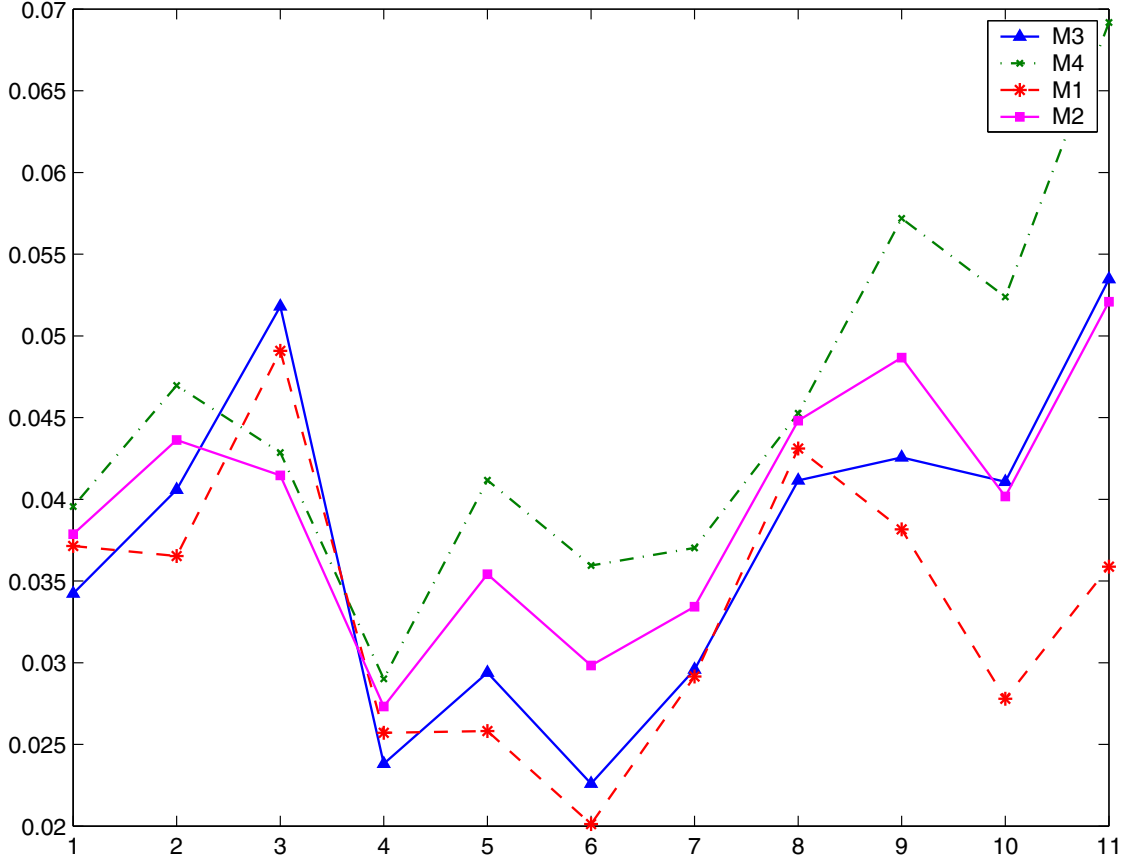


Figure 11. Plot of the 12 annual standard deviations in the years of the out-of-sample period. These standard deviations are computed as the sample statistics for the monthly returns of the global minimum variance portfolios in the corresponding years for the four model specifications. The four models are as follows:

$$\Sigma_t = \mathbf{b}'\mathbf{V}_t\mathbf{b} + \Omega \quad (1)$$

$$\Sigma_t = \mathbf{b}'\mathbf{V}_t\mathbf{b} + \text{diag}(\Omega) \quad (2)$$

$$\Sigma = \mathbf{b}'\mathbf{V}\mathbf{b} + \Omega \quad (3)$$

$$\Sigma = \mathbf{b}'\mathbf{V}\mathbf{b} + \text{diag}(\Omega) \quad (4)$$

Models (1) and (2) are the unconstrained and constrained specifications of our factor specification. Models (3) and (4) are the corresponding benchmarks.