Latent Factor Modeling of Multivariate Conditional Heteroscedasticity

Mike Aguilar†

First Draft: August 24, 2007
This Draft: October 19, 2007

Abstract

This paper examines the joint dynamics of a system of asset returns. I describe and implement a multivariate factor stochastic volatility (MVFSV) model.

I follow closely the work of Doz and Renault (2006), with two important changes. First, I design a sequential testing procedure to determine the dimensions of the appropriate factor structure needed to accommodate the conditional heteroscedasticity among a system of returns. Second, I employ a form of Tikhonov Regularization in order to overcome a near singularity among the moment conditions used for estimation.

Simulation studies suggest that the MVFSV model is able to recover accurately the latent factors that drive the conditional volatility of returns. Moreover, the model estimates can be used to construct conditionally homoscedastic portfolios as linear combinations of the conditionally heteroscedastic assets.

An empirical application to portfolios representing the twelve sectors of the U.S. economy finds that the MVFSV model has important investment implications. Over the period 1993 through 2006, a dynamic asset allocation strategy implied by the MVFSV model is able to outperform the market and perform comparably with a strategy implied by the Dynamic Conditional Correlation model of Engle (2001).

*I would like to thank the members of the UNC Financial Econometric Student Workshop for helpful comments, and particularly the guidance of my dissertation advisor Eric Renault. All errors herein are my own. Comments are welcome.

†UNC at Chapel Hill, Dept. of Economics; maguilar@email.unc.edu
1 Introduction

The literature on modeling univariate volatility processes is well established. GARCH and Stochastic Volatility (SV) models of financial assets have shown to be quite capable in this regard. However, it is clear that the joint estimation of a system of returns is warranted. For instance, Figure 1 illustrates the cross-sector correlation coefficients within the U.S. equity market over the last 16 years. The fact that the correlation is non-zero and that it is time varying call for a complete, joint estimation of returns.

Standard multivariate GARCH and SV models have shown some success in accommodating the joint dynamic of such systems. However, they fail to identify the sources of volatility (and co-volatility) movements. Multivariate factor models, such as O-GARCH or its SV counterparts, allow for these sources of volatility to be identified quite easily. The goal of this paper is to implement a multivariate factor stochastic volatility (MVFSV) model.

The class of MVFSV models was actually introduced in an ARCH context by Diebold and Nerlove (1989). In that paper, the authors formulated a multivariate model of returns centered about a single common factor. The factor was designed to follow ARCH dynamics. However, the factor was also deemed to be latent, and hence the SV characterization.


Latent factor models such as these have several advantages over models that attempt a direct characterization of the variance/covariance matrix of returns. First, the factor representation allows the researcher to capture the time variation in the conditional covariance matrix through the movements of a small number of underlying latent factors. This mitigates the proliferation of parameters often seen in multivariate volatility modeling. Second, by characterizing the common movements in returns and volatilities, these factor models fit nicely into an APT framework. Last, the identification of a limited number of directions of risk has practical advantages to portfolio managers and risk specialists.

The main contribution of this paper is the device of a comprehensive methodology for empirical implementation of the MVFSV model described in Doz and Renault (2006), herein referred to as DR. I illustrate the challenge with estimating this model and attempt to overcome that difficulty via a form of Tikhonov Regularization. Moreover, I refine a testing strategy associated with the model specification by using a sequential procedure to account for a sequence of nested hypotheses.

The MVFSV model is estimated in two phases. Phase 1 estimates the number of latent factors required to accommodate the conditional heteroscedasticity in a system of asset returns. DR accomplish this through an over-identified restrictions test. I modify this approach slightly by introducing a sequential testing procedure,

---

Footnote 1: As described in the Empirical Application section of this paper, I use 12 sectors of the U.S. equity market. For each pair of sectors, I compute the time (t) correlation coefficient over the preceding 3 months. I repeat this process for all pairs of sectors. For each time (t) I then average the correlation coefficients, yielding a time series of cross-sector correlations.
which accounts for the aforementioned series of nested hypotheses. The sequential procedure is shown to offer a very slight power advantage over the one-step DR approach during hypothesis testing against various alternatives. However, the true virtue of the sequential approach is its ability to effectively re-order the conditionally heteroscedastic assets in a way that facilitates the construction of auxiliary portfolios. Akin to the common features work of Engle and Kozicki (1993), these auxiliary portfolios are formed as linear combinations of the conditionally heteroscedastic asset returns and no longer exhibit that common underlying trait; i.e. they are conditionally homoscedastic.

Once the dimensions of the factor structure are revealed in Phase 1, Phase 2 estimates the complete MVFSV model through a well chosen set of moment conditions.

Each phase is estimated via GMM. This has the advantage of overcoming the computational challenge associated with an intractable likelihood function usually encountered when estimating multivariate stochastic volatility models. Moreover, the GMM approach avoids a full parametric specification of the distribution of returns. DR also relaxes certain restrictions, to be made precise in the next section, that are typically placed on the variance/covariance matrix of returns.

Simulation studies indicate that Phase 1 of the MVFSV model is able to identify correctly the number of latent factors in a system of conditionally heteroscedastic asset returns. Moreover, the model identifies linear combinations of the assets that no longer exhibit individual ARCH effects and have constant conditional correlation. In addition, Phase 2 provides parameter estimates sufficient to accurately recover the latent factors that drive the conditional volatility of returns.

An empirical application to portfolios representing the twelve sectors of the U.S. economy finds that the MVFSV model is able to accommodate the individual dynamics of returns quite well throughout the sample period. Joint dynamics are accommodated poorly during the early 1990’s, but much more capably during the mid - 1990’s through 2006.

The MVFSV model also has important out-of-sample investment implications. A dynamic asset allocation strategy implied by the MVFSV model is able to outperform the market portfolio and a strategy implied by the Dynamic Conditional Correlation (DCC) model of Engle (2001) for much of the investment period examined. A considerable level of trading activity, however, can erode the MVFSV strategy’s advantage along several performance metrics when transaction costs are high. This finding is particularly acute during the market boom/bust of the late 1990’s / early 2000’s.

The paper is organized as follows. Section 2 details the model. Section 3 outlines the two-phase estimation procedure. Section 4 describes the results from a simulation study of Phases 1 & 2. Section 5 summarizes the empirical application for 12 sector portfolios of the U.S. equity market.

2 Model

I examine the dynamics of $n$ asset returns $(y_{t+1})$ and their $(n \times n)$ conditional variance matrix $\Sigma_t = V(y_{t+1} | J_t)$. $J_t$ is an information set that contains past values of the returns, $y_{\tau} \forall \tau \leq t$, as well as the past of
some unobserved common factors, \( f_r \forall \tau \leq t \). Notice that \( J_t \) differs from the econometrician’s information set \( I_t \), which contains only past values of returns. Specifically, \( I_t \subset J_t \). For notational convenience I will denote all relevant information sets through a simple time subscript, so that \( \Sigma_t \equiv V_t(y_{t+1}) = V(y_{t+1}|J_t) \).

Consider the decomposition of this variance matrix:

\[
\Sigma_t = \Lambda D_t \Lambda' + \Omega_t
\]

where \( D_t \) is a diagonal matrix of size \( K \) with diagonal coefficients \( \sigma_{kt}^2, k = 1, ..., K \). Consider, for now, \( \Omega_t \) to be a diagonal positive definite matrix. Moreover, the \( \sigma_{kt}^2, k = 1, ..., K \) are positive, stationary stochastic processes with unit expectation and non-zero variance. The unconditional expectation of \( D_t \) is an Identity matrix of size \( K \).

Viewing the \( \sigma_{kt}^2 \) as conditional variances of \( K \) independent common factors, \( \sigma_{kt}^2 = V_t(f_{kt+1}) \), enables the formation of the \( K \)-factor conditional regression representation:

\[
y_{t+1} = \mu + \Lambda f_{t+1} + u_{t+1}
\]

where \( f_{t+1} \) is a \((K \times 1)\) vector of unobserved latent factors, \( u_{t+1} \) is an \((n \times 1)\) vector of idiosyncratic terms, \( \Lambda \) is an \((n \times K)\) matrix of factor loadings, and \( \mu \) is an \((n \times 1)\) vector of constants that may be interpreted as risk premia if \( y_{t+1} \) is taken to be in excess of the risk free rate.

I define the variance/covariance matrices of the errors and factors in equation (2) as \( E_t(u_{t+1}u_r'_{t+1}) = \Omega_t \) and \( E_t(f_{t+1}f_r'_{t+1}) = D_t \), respectively. I also impose the following assumptions: \( E_t(f_{t+1}) = 0, E_t(u_{t+1}) = 0, \) and \( E_t(f_{t+1}u_r'_{t+1}) = 0 \). Notice that this last assumption allows us to interpret the factor loadings as standard conditional regression coefficients of returns on the factors, and \( \Omega_t = Var_t[u_{t+1}] \) as the residual risk.

A factor-analytic structure such as that presented above is not new. A majority of the MVFSV models mentioned earlier use this formulation. However, DR part from the previous literature in two ways. First, the matrix of residual risk is deemed to be time invariant. This assumption implies that the conditional heteroscedasticity of returns can be captured completely through the movements of the underlying common factors. Second, DR remove the somewhat restrictive diagonality assumption on residual risk. The variance/covariance matrix of the returns can then be written as

\[
\Sigma_t = \Lambda D_t \Lambda' + \Omega_t
\]

where \( \Omega \) is a possibly non-diagonal, time invariant matrix.

The price to pay for this specification of \( \Omega \) is that we are unable to identify uniquely all the parameters in the model. Importantly, only the range of \( \Lambda \) is identified. This is due to the non-diagonal structure permitting any constant part of the conditional variance of the factors to be transferred to the variance of the errors by a simple re-scaling of the loading vectors.

\footnote{Note: DR detail the possibility of a time varying risk premium with the specification \( \mu_t = \mu + \Lambda \phi_t \). I leave this for future work.}
Consider $k \leq K$ as a particular number of latent factors being considered for a system of asset returns. Denote $\mathbf{y}_{t+1}$ as the first $k$ components of $y_{t+1}$ and $\mathbf{f}_{t+1}$ as the remaining $(n - K)$ elements of $y_{t+1}$. Consistent with the common features work of Engle and Kozicki (1993), DR show that there exist linear combinations of the asset returns that are conditionally homoscedastic. That is, if there exists a common factor that explains the conditional heteroscedasticity among the assets in this system, there should be linear combinations of those assets that no longer contain conditional heteroscedasticity. Specifically, the conditionally homoscedastic portfolios take the form $(\mathbf{y}_{t+1} - B\mathbf{f}_{t+1})$, where $B$ is a constant matrix of dimension $((n - k) \times k)$.

3 Estimation

I estimate the MVFSV model in two phases, each employing a GMM technique. Phase 1 determines the number of latent factors needed to accommodate adequately the temporal dynamics of the assets under consideration. Phase 2 estimates the complete multivariate factor stochastic volatility model.

3.1 Phase 1: Search for the Number of Common Factors

The empirical goal of Phase 1 is to determine the number of factors (K). This is a non-trivial task that has been dealt with in a rather ad-hoc fashion in the literature. Broadly speaking, there are three techniques that commonly are used to determine the number of factors in a system of asset returns: 1) Common Features, 2) Principal Components, and 3) Model Selection Criteria. This certainly is not an exhaustive list of options, nor are they necessarily mutually exclusive. However, I use these three to characterize a general class of possibilities for my estimation procedure.

Common Features, as espoused by Engle and Kozicki (1993), appears to be the least utilized of the three major techniques in the field of volatility modeling. Lanne and Saikkonen (2007) offer a rare example of this approach, wherein the authors deal explicitly with the issue of determining the number of latent factors. The authors choose a number of factors a priori and then estimate a factor GARCH model. They then validate this choice post-estimation through an original testing procedure.

Principal Components, on the other hand, is perhaps the most widely used technique to determine the number of factors in the GARCH literature. The O-GARCH model of Alexander (2001) determines the number of factors to be considered pre-estimation. The asset returns are decomposed into the their (K) principal components, accounting for some pre-determined amount total variation among the returns, say 90%. The author then estimates the GARCH model upon the "factors".

Model Selection Criteria are commonly used in determining the dimension of the factor space in multivariate stochastic volatility models. For instance, Connor, Korajczyk, and Linton (2006) choose a range of potential dimensions for the factor space a priori and then estimate their stochastic volatility model over each choice of K. The appropriate number of factors is then chosen post-estimation. Broadly speaking, I categorize this type of analysis as Model Selection because the authors choose the number of factors that best fits their model to the data.
I address the issue of determining the size of the factor space explicitly. Moreover, \( K \) is selected pre-estimation. I say ”pre”-estimation because the dimension of the factor space is determined before the full model is estimated in Phase 2.

DR show that efficient estimation of the stochastic volatility factor structure, and thus the dimension of the factor space, should be conducted via:

\[
E[z_t \otimes \text{vec}\{E_t[(\bar{y}_{t+1} - B\bar{y}_{t+1})y'_{t+1} - D]\} = 0 \quad (4)
\]

The common features work of Engle and Kozicki (1993) is the obvious motivating force for these moment conditions since they are based upon the conditionally homoscedastic portfolios \((\bar{y}_{t+1} - B\bar{y}_{t+1})\). However, the information content in the stochastic volatility factor structure is greater than that of a pure common features model. For instance, if I drop the instrument vector for ease of notation and partition D in an obvious fashion, \((4)\) can be re-written as

\[
E_t[(\bar{y}_{t+1} - B\bar{y}_{t+1})y'_{t+1}] = D_1 \quad (5)
\]

\[
E_t[(\bar{y}_{t+1} - B\bar{y}_{t+1})y''_{t+1}] = D_2 \quad (6)
\]

Meanwhile, the common features model contains information only about linear combinations of the equations above:

\[
E_t[(\bar{y}_{t+1} - B\bar{y}_{t+1})(\bar{y}_{t+1} - B\bar{y}_{t+1})'] = D_2 - BD_1 \quad (7)
\]

The moment conditions in equation \((4)\) pave the way for GMM estimation and inference. The null hypothesis associated with an over-identified restrictions test is that \( k \) factors are sufficient to accommodate the conditional heteroscedasticity among the \( n \) assets that comprise \( y_{t+1} \). An equivalent phrasing of the null is that there exist \((n - k)\) conditionally homoscedastic portfolios of the form \((\bar{y}_{t+1} - B\bar{y}_{t+1})\).

Upon careful inspection it becomes clear that this null nests a series of hypotheses. Accepting the null of \( k \) factors being sufficient to accommodate the conditional heteroscedasticity among the \( n \) assets implies that at most \( k \) factors are needed to accommodate the dynamics of any subset of the \( n \) assets. Failing to test the dimension of the factor structure for subsets of the asset space creates a possible loss of power.

3 Simulation studies suggest a power loss of 3-5 percentage points from ignoring the sequential procedure.

To account for these nested hypotheses, I part from DR slightly by incorporating a sequential testing procedure.

Begin by ordering the assets \( y^{(1)} \) through \( y^{(n)} \) according to their exposure to conditional heteroscedasticity. In the empirical application of this paper, that exposure is determined through simple univariate ARCH tests. I then allocate subsets of the asset space to \( \bar{y} \) and \( \bar{y} \) according to the number of factors being considered.

Consider the possibility of a single common factor. Set \( \bar{y} = y^{(1)} \). The first step of the sequential procedure determines whether a single factor is sufficient to accommodate the conditional heteroscedasticity among the first two assets. I will detail the test statistic shortly. This structure implies \( \bar{y} = y^{(2)} \).

If a single common factor is capable of accommodating the conditional heteroscedasticity among the first
two assets, examine whether this common factor is also sufficient for the dynamics of the first three assets. In this case \( \mathbf{y} = [y^{(2)}, y^{(3)}] \). If a single factor is not sufficient to accommodate the dynamics for these two assets, then examine whether two factors are capable of accommodating the dynamics among the first three assets. A sample of the sequential testing procedure for a single factor is as follows:

\[
H^1_a : \text{1 Factor is sufficient for assets } (y^{(1)}, y^{(2)})
\]
\[
H^1_h : \text{1 Factor is not sufficient for assets } (y^{(1)}, y^{(2)})
\]
If fail to reject, proceed to next step in sequence.
If reject, stop this sequence and consider two factors.

\[
H^2_a : \text{1 Factor is sufficient for assets } (y^{(1)}, y^{(2)}, y^{(3)})
\]
\[
H^2_h : \text{1 Factor is not sufficient for assets } (y^{(1)}, y^{(2)}, y^{(3)})
\]
If fail to reject, proceed to next step in sequence.
If reject, stop this sequence and consider two factors.

\[\vdots\]

\[
H^{n-1}_a : \text{1 Factor is sufficient for assets } (y^{(1)}, y^{(2)}, \ldots, y^{(n)})
\]
\[
H^{n-1}_h : \text{1 Factor is not sufficient for assets } (y^{(1)}, y^{(2)}, \ldots, y^{(n)})
\]
If fail to reject, stop.
If reject, stop this sequence and consider two factors.

If I fail to reject the last null in the sequence above, then I consider one factor to be sufficient to accommodate the conditional heteroscedasticity among all \( n \) assets. As a result, there exists \( n - 1 \) conditionally homoscedastic portfolios formed as linear combinations of the \( n \) conditionally heteroscedastic primitive assets.

If I reject the null hypothesis at any stage of the sequence, I consider the possibility of two common factors. For instance, assume \( H^2_h \) is rejected in the sequence above. This implies that \( y^{(3)} \) introduced new dynamics to the system that could not be accommodated by the single factor that was sufficient for \( y^{(1)} \) and \( y^{(2)} \). Therefore, consider a new sequence, the first null hypothesis of which suggests that two common factors are sufficient for assets \( y^{(1)}, y^{(3)}, y^{(2)} \), where \( \mathbf{y} = [y^{(1)}, y^{(3)}] \) and \( \mathbf{y} = y^{(2)} \). The sequence then proceeds in a similar fashion as that outlined above. In this sense, the sequential procedure re-orders the assets in a way that identifies suitable candidates for the \( \mathbf{y} \) vector.

As is well documented in the literature, care must be taken when evaluating a sequence of nested hypotheses. Particular attention must be given to the construction of the test statistics, the degrees of freedom, and the size of the test.

The test statistic at any stage of the sequence is the \( J \)-statistic from an over-identified restrictions test. Consider the typical GMM objective function \( J = Tg'Wg \), where \( g \) is the sample mean of the moments, the weighting matrix \( (W) \) is set optimally to the inverse of the variance matrix of the moments \( (S) \), and \( T \) is the sample size. The degrees of freedom is equal to \( p - q \), where \( p \) is the number of moment conditions and \( q \) is the number of parameters in the model. This is all standard. Unfortunately, naively using this test statistic and degrees of freedom at any given step of the sequence ignores the fact that I have accepted each of the
preceding steps. The J-statistic and the degrees of freedom must be adjusted.

The motivation for the adjustment comes from Eichenbaum, Hansen, and Singleton (1998). Using their framework, I view each step of the sequence as a test of a subset of the moment conditions. To illustrate in a general setting, consider an H-dimensional instrument vector \( z \). Moreover consider a set of moment conditions that can be partitioned into two non-overlapping subsets \( a \) and \( b \) as follows:

\[
m^a(\theta) = E \left[ z \otimes \begin{bmatrix} (y(2) - B_1 y^{(1)})(y^{(1)})' - D_1 \\ (y(2) - B_1 y_1)(y^{(2)})' - D_2 \end{bmatrix} \right]
\]

where \( q_a \) is the number of parameters and \( p_a \) is the number of moment conditions;

\[
m^b(\theta) = E \left[ z \otimes \begin{bmatrix} (y(2) - B y^{(1)})(y^{(3)})' - D_3 \\ (y(3) - B y^{(1)})(y^{(1)})' - D_4 \\ (y(3) - B y^{(1)})(y^{(2)})' - D_5 \\ (y(3) - B y^{(1)})(y^{(3)})' - D_6 \end{bmatrix} \right]
\]

where \( q_b \) is the number of parameters and \( p_b \) is the number of moment conditions. Let \( \theta \) be a parameter vector containing the relevant elements of \( B \& D \).

Now consider the hypothesis where subsets \( a \) and \( b \) hold under the null, but only subset \( a \) holds under the alternative.

\[
\begin{align*}
H^E_0 &: E[m^a(\theta)] = 0 \quad \& \quad E[m^b(\theta)] = 0 \\
H^E_\Lambda &: E[m^a(\theta)] = 0 \quad \& \quad E[m^b(\theta)] \neq 0
\end{align*}
\]

Eichenbaum, Hansen, and Singleton (1998) offer a simple test statistic, which is the difference of the J-statistics from the over-identified restrictions test. Denote \( J^E_a = T g^a(\hat{\theta})' W g^a(\hat{\theta}) \) as the test statistic pertaining only to the subset \( a \) of moment conditions, and that which is true under both the null and the alternative. Moreover, denote \( J^E_{a+b} = T g^{a+b}(\hat{\theta})' W g^{a+b}(\hat{\theta}) \) as the test statistic using the entire set of moment conditions. The test statistic accompanying the joint hypothesis in (10) is then \( \xi^E = J^E_{a+b} - J^E_a \). Under the proper regularity conditions, this test statistic is chi-square distributed with \( q_b - p_b \) degrees of freedom.

To see how this framework can be applied to Phase 1 of the MVFSV model, consider the first two steps of a sequence in a simple three asset case. I will drop time subscripts for notational convenience. The first step can be written as:

\[
H^1_0 : \text{1 Factor is sufficient for assets } (y^{(1)}, y^{(2)})
\]

\[
H^1_a : \text{1 Factor is not sufficient for assets } (y^{(1)}, y^{(2)})
\]

\[
m^1(\theta) = E \left[ z \otimes \begin{bmatrix} (y(2) - B_1 y^{(1)})(y^{(1)})' - D_1 \\ (y(2) - B_1 y_1)(y^{(2)})' - D_2 \end{bmatrix} \right]
\]

The second step can be rewritten as

\[
H^2_0 : \text{1 Factor is sufficient for assets } (y^{(1)}, y^{(2)}, y^{(3)})
\]

\[
H^2_a : \text{1 Factor is not sufficient for assets } (y^{(1)}, y^{(2)}, y^{(3)})
\]
Moment conditions $m^1(\theta)$ are analogous to $m^a(\theta)$ in the general example, and $m^2(\theta)$ are analogous to $m^{a+b}(\theta)$. Therefore, in order to construct a proper test statistic, I need only subtract the J-statistics and degrees of freedom along each step of the sequence. For example, the adjusted test statistic for step 2 of the sequence is $\xi^2 = J^2 - J^1$. Moreover, the adjusted degrees of freedom is merely the difference in the degrees of freedom from each step along the sequence. Specifically, $J^1 \sim \chi^2_{q_1-p_1}$, $J^2 \sim \chi^2_{q_2-p_2}$, and the adjusted test statistic in the second step is $\xi^2 = J^2 - J^1 \sim \chi^2_{q_2-p_2-(q_1-p_1)}$.

Care must also be taken when denoting the nominal size of each step along the sequence. As Gourieroux and Monfort (1995) illustrate nicely, the nominal size of the test compounds as I progress along the sequence, thereby reducing the confidence I can draw from each inference.

The sequence of tests is designed as a descending step-wise procedure. The null in step one is rather broad. The single factor identified need account only for the conditional heteroscedasticity in two assets. If I fail to reject this null, I proceed to the next step of the sequence by adding on the additional requirement that the factor be able to accommodate the conditional heteroscedasticity among three assets. Through the descending step-wise approach, I am able to control for the nominal size of the test along every step of the sequence.

Denote $\alpha_k$ as the level of the test at step $k$ of the sequence. The probability of a Type I error at step $k$ can be seen as a function of the levels for all the preceding steps of the sequence. Specifically, $\Pr[\text{reject } H^{k}_0 | H^{k}_0 \text{ true}] = 1 - \prod_{j=1}^{k} (1 - \alpha_j)$. Setting $\alpha_j = \alpha \forall j$, I can induce the significance level for testing $H^{k}_0$ as $1 - (1 - \alpha)^k$.

Let us revisit the small-scale example of three assets. The last null in the sequence, and ultimately the one of interest, is $H^3_0$, which I want to test at the 10% level. Solving for $\alpha; 1 - (1 - \alpha)^2 = 0.10 \rightarrow \alpha = 0.0513$. I can now induce the levels of significance at each step of the sequence: $H^1_0$ is tested at 0.0513, and $H^2_0$ is tested at 0.10, as desired.

With the sequential testing procedure in hand, I should be able to identify $K$, the number of factors in the system of assets. Unfortunately, a near singularity among the moment conditions hinders Phase 1 estimation. I attempt to overcome this difficulty through a form of Tikhonov Regularization.

### 3.2 Tikhonov Regularization

I attempt to overcome the aforementioned estimation difficulty through a form Tikhonov Regularization. The technique introduces small perturbations to the diagonal elements of the variance matrix $S$. These perturbations, or Tikhonov Factors, must be large enough to alleviate the ill-posed problem, yet the smallest
possible in order to consider that we still have approximately reach efficient GMM. I regularize $S$ as follows:

$$S^* = \frac{1}{T} \sum_{t=1}^{T} m_t m_t' + \alpha^* I_{(n-K)n}$$

where $\alpha^*$ is the Tikhonov Factor. I then use the following GMM objective function for estimation: $J^* = T g^T W^* g$, where $W^* = (S^*)^{-1}$.

The challenge in utilizing regularization techniques is calibrating the size of the Tikhonov Factor. Picking $\alpha^*$ too small will not avoid the near singularity. Picking $\alpha^*$ too large will cause $S^*$ to grow large, sending the GMM objective function (J) toward zero. This contaminates the parameter estimates and precludes reliable inference. Mindful of this tradeoff, I offer here a somewhat crude, but effective means of choosing $\alpha^*$.

Begin by choosing a large $\alpha^*$ that allows us to avoid the near singularity issue. Conduct GMM estimation with the weighting matrix equal to $(S^*)^{-1}$. Gather the parameter estimates into vector $\theta(1)$. Pick another $\alpha^*$ by decreasing the previous choice. Re-estimate and gather the parameter vector $\theta(2)$. Compute the element-wise percentage change in the parameter vectors and take its norm; $d(2) = \| \theta(2) - \theta(1) \|$. This $d(2)$ is a single point in Figure 5. Pick another $\alpha^*$ by incrementing the previous choice. Re-estimate and gather the parameter vector $\theta(3)$. Compute the element-wise percentage change of the parameter vectors and take its norm; $d(3) = \| \theta(3) - \theta(2) \|$. Repeat this process until $d(i)$ rises above a given tolerance; $\| \theta(i+1) - \theta(i) \| > tol$.

3.3 Phase 2: Full Model Estimation

Once the number of factors is identified in Phase 1, I can estimate the complete K-factors conditional regression representation defined in equation (2).

Estimation requires a specification for the dynamics of the factor volatilities, for which I assume an affine mean reverting structure:

$$E_{t-1}(\sigma^2_{kt}) = 1 - \gamma_k + \gamma_k \sigma^2_{kt-1}$$

Unfortunately, the latent nature of the factors and the non-diagonal residual risk ($\Omega$) prevents unique identification of the loading vector $\Lambda$. In fact, only the range of $\Lambda$ can be identified (see Fiorentini and Sentana (2001)). This lack of identification in $\Lambda$ carries over to the other parameters of interest in the model. However, DR show that the stochastic volatility factor structure permits estimation of functions of the parameters that are invariant to scale changes in $\Lambda$.

Under the constant risk premium specification, DR show that the parameters $B, \gamma_k, k = 1, ..., K, \mu, (\Lambda' + \Omega)$, and $(\Omega_2 - B\Omega_1)$ can be uniquely characterized by the following set of conditional moment restrictions:

---

4I can also form a "regularized" Newey-West variance / covariance matrix as follows: $S^*_{NW} = \hat{\Gamma}(0) + \sum_{j=1}^{T-1} \kappa(j,b)\hat{\Gamma}(j)' + \alpha^* I_{(n-K)n}$, where $\kappa$ is the kernel and $b$ is the bandwidth.

\[ E_t[y_{t+1}] = \mu \]
\[ vecE_t(\vec{y}_{t+1}) = vec(\vec{\mu} - B\vec{\mu}) + (\Omega_2 - B\Omega_1) \]
\[ vechE_{t-K} [\prod_{k=1}^{K}(1 - \gamma_k L)^{(y_{t+1}y'_{t+1})}] = vech(\mu\mu' + \Lambda\Lambda' + \Omega) \prod_{k=1}^{K}(1 - \gamma_k) \]  

once a proper set of instruments is chosen.

I refer the reader to DR for proofs illustrating the ability of these moment conditions to identify the stated parameters. However, the intuition is rather straightforward. The first set of moment conditions in equation (15) obviously identifies the unconditional mean of returns. The second set of moment conditions are akin to those used in Phase 1, estimating \( B \) and \( \Omega_2 - B\Omega_1 \). The third set of moment conditions are in the spirit of the multi-period moments conditions of Meddahi and Renault (2004), where the \( (1 - \gamma L) \) filter annihilates the dynamics of \( \sigma_t^2 \). This set of moments identifies \( \gamma \) and \( \Lambda\Lambda' + \Omega \).

Moreover, I follow DR by partitioning and re-parameterizing the loading vector \( \Lambda \) such that \( \Lambda = [\Lambda' \Lambda']' = [I B']\bar{\Lambda} \), with \( \bar{\Lambda} \) an invertible matrix and \( B = \bar{\Lambda}\Lambda^{-1} \).

4 Simulation Study

In this section I conduct several simulation exercises to evaluate the efficacy of the MVFSV model. Specifically, I aim to present evidence of the presence of univariate ARCH effects as well as multivariate dynamic condition correlation (DCC) effects in the simulated returns. I then estimate Phase 1 of the estimation strategy using the Regularization method. With the estimates of \( B \) in hand, I form auxiliary portfolios and test for the presence of conditional homoscedasticity, in both a univariate and multivariate sense. In addition, I use estimates from Phase2 to extract the latent factors and compare these to the actual simulated factors.

The test I use for examining univariate conditional heteroscedasticity is the standard ARCH LM test due to Engle (1982). Akin to the Breusch-Godfrey test for autocorrelation, the ARCH LM test regresses squared residuals upon their own past. The test statistic is \( TR^2 \), and is distributed \( \chi_q^2 \) where \( q \) is the order of the ARCH process.

The test for DCC effects is due to Engle and Sheppard (2001). Consider the \((n \times 1)\) vector of returns \( y_t | J_{t-1} \sim N(0, H_t) \), where \( H_t \equiv D_t R_t D_t \). \( D_t \) is an \((n \times n)\) diagonal matrix of time varying standard deviations from univariate GARCH models. \( R_t \) is the potentially time varying correlation matrix. The test of constant conditional correlation is as follows:

\[ H_a : R_t = \bar{R} \forall t \in T \]
\[ H_a : vech(R_t) = vech(R_t) + \beta_1 vech(R_{t-1}) + \ldots + \beta_p vech(R_{t-p}) \]

Test test centers around an accompanying artificial vector auto-regression:

\[ Y_t = \alpha + \beta_1 Y_{t-1} + \ldots + \beta_L Y_{t-L} + \eta_t \]
where \( Y_t = \text{vech} u [(R^{-1/2} D_t^{-1} \xi_t)(R^{-1/2} D_t^{-1} \xi_t)'] - I_g \), with \( (R^{-1/2} D_t^{-1} \xi_t) \) representing a \((g \times 1)\) vector of residuals jointly standardized under the null and \( \text{vech} \) is an operator that selects the elements above the diagonal of a matrix.

Under the null of a constant conditional correlation, the constant and all the lagged parameters in the auxiliary regression should be zero. The accompanying test is conducted via a seemingly unrelated regression with a resulting test statistic of \( \hat{\delta} \hat{XX}' \hat{\delta} \hat{\sigma}^2 \), which is asymptotically distributed \( \chi^2_{L+1} \). The \( \hat{\delta} \) are the estimated regression parameters, \( X \) contains all the regressors including the constant, and \( \hat{\sigma}^2 \) is the estimated variance of the error term.

### 4.1 Simulating Return Paths

All of the simulation exercises I will undertake are based on equation (2), which I re-write here for convenience:

\[
y_{t+1} = \mu + \Lambda f_{t+1} + u_{t+1}
\]

Define \( y_{t+1} \) as an \((n \times 1)\) vector of asset returns. Unless otherwise stated, the risk premium, \( \mu \), is assumed to be zero. The error term, \( u_{t+1} \), is distributed multivariate normal such that \( u_{t+1} \sim N_n(0, I) \). In a single factor case, \( f_{t+1} \) will take the GARCH(1,1) form:

\[
\begin{align*}
f_{t+1} &= \sigma_t \varepsilon_{t+1} \\
\sigma_t^2 &= \omega + \alpha f_t^2 + \beta \sigma_{t-1}^2
\end{align*}
\]

where \( \theta = [\omega, \alpha, \beta] \). This volatility specification is consistent with the SR-SARV(1) structure put forth in equation (14), where \( \gamma = \alpha + \beta \).

The choice of the loading vector \( \Lambda \) and the parameters \( \theta \) for the factor(s) \( f_{t+1} \) must be guided by four key properties. 1) The parameter vector \( \theta \) must be realistic. 2) Moments of at least order four must exist for the asset returns. 3) The parameter vector \( \theta \) must be chosen such that the ARCH LM test can detect the presence of conditional heteroscedasticity within the factor. 4) The loading vector \( \Lambda \) must be chosen such that the ARCH LM test will also be able to detect the presence of conditional heteroscedasticity in the return series. Properties three and four ensure that the ARCH LM test is a useful diagnostic tool.

Choosing \( \theta \) values that are realistic is straightforward. My application in this paper focuses on daily stock returns. Fitting GARCH(1,1) specifications to a variety of daily U.S. equities yields a range of \( \theta \) values depicted by the enclosed box in Figure 2.

The need for finite fourth moments comes from the structure of the moment conditions. Equation (4) indicates that the first moment condition could be written as \( (y_t^{(1)})^2 (y_{t+1}^{(2)} - By_{t+1}^{(1)})y_{t+1}^{(1)} - D_1) \). Expanding out this multiplication reveals something like \( y^4 \). The existence conditions detailed by Bollerslev (1986) can be used as a tool to determine the existence of the fourth moments of the simulated return series. All points to the left of the frontier in Figure 2 represent \((\alpha, \beta)\) combinations that yield finite 4th moments.
The third key property of the simulations is that \( \theta \) is chosen such that the ARCH LM test is able to detect the conditional heteroscedasticity present in the factor(s). I investigate the power of the ARCH(1) LM test at detecting the conditional heteroscedasticity in a GARCH(1,1) process across a variety of \( \alpha \) and \( \beta \) combinations. The \( \omega \) is set equal to \( 1 - \alpha - \beta \) so as to generate unit (unconditional) variance. The graph on the left of Figure 2, depicts a power surface of the LM test for ARCH(1) given a GARCH(1,1) alternative for a sample size of 750 and a simulation size of 100. The nominal size for the tests is 10%. I examine only those \((\alpha, \beta)\) combinations that are both realistic and correspond to finite fourth order moments. The right hand graph in Figure 3 depicts a power surface of the LM test for ARCH(2) errors. In either case, power appears to be highest at low levels of \( \alpha \).

The final consideration for the simulations is that the loading vector \( \Lambda \) be set such that the ARCH LM test be able to detect the conditional heteroscedasticity within the return series. Recall the variance decomposition detailed in equation (1). For asset one this equation reduces to
\[
\sigma^2_{1t} = \lambda_1^2 \text{Var}(f) + \text{Var}(u_t) .
\]
Recall as well that \( \text{Var}(f) = \text{Var}(u) = 1 \) by design. So, the proportion of the total variance \( \sigma^2_{1t} \) accounted for by the factor is \( \frac{\lambda_1^2}{\sigma^2_{1t}} \). As \( \lambda_1 \) increases so does that amount of variation in the asset accounted for by the conditionally heteroscedastic factor.

I run a simple Monte Carlo experiment to determine the appropriate \( \Lambda \) in the simulations. I build an ARCH(1) factor as follows:
\[
\begin{align*}
 f_{t+1} &= \sigma_t \varepsilon_{t+1} \\
 \sigma_t^2 &= \omega + \alpha f_{t-1}^2
\end{align*}
\]
where \( \varepsilon \sim N(0, 1) \). I choose \( \omega \) & \( \alpha \) by modeling the sector returns used in the empirical application section of this paper as ARCH(1) processes. Setting \( \omega = 0.01 \) and \( \alpha = 0.50 \) seem to be reasonable approximations of this data, which I will detail later. I then construct a series \( Y_t = \lambda_t f_t + u_t \) of length \( T = 25,000 \), where \( u_t \) is a standard normal random variable and \( \lambda_t \epsilon \Lambda = \{0, 0.1, 0.2, ..., 5\} \). For each \( \lambda_t \) I conduct an LM test for the presence of ARCH(1) effects in the residuals \( \hat{u}_t \). Figure 4 plots the p-values from this test against \( \Lambda \). A \( \lambda \) of at least 2.5 is required for the ARCH LM test to detect the conditional heteroscedasticity in \( Y \).

### 4.2 Gauging the Near Singularity

To illustrate the aforementioned near singularity in the variance/covariance matrix of the moments, consider a small-scale five asset example. I generate 100 paths simulated from equations (16) & (17), where \( \theta = (0.03, 0.06, 0.91) \) and \( \Lambda = [10, 9, 8, 7, 6]' \). Estimating the first step of the sequential testing procedure, that which is associated with \( H^1_o \), is straightforward. However, proceeding to the second step, that which is associated with \( H^2_o \), is challenging.

Define a \((j \times 1)\) vector of instruments \( z_t = [1 \ (y_t^{(1)})^2] \). Form the moments \( m_{t+1} = z_t \otimes \text{vec}(\bar{y}_{t+1} - B\bar{y}_{t+1})y_{t+1}' - D \) and denote their mean over the \( T \) observations as \( g \).

Without loss of generality, I set \( y^{(1)} = \bar{y} \) for this example. Similar results are found for \( \bar{y} = y^{(2)} \) or \( y^{(3)} \). Use the identity matrix as an initial weighting matrix and form the typical GMM objective function \( J = Tg'y'Wy \). Denote \( S = \frac{1}{T} \sum_{t=1}^{T} m_t m_t' \) as the empirical counterpart of the asymptotic variance matrix of the moments. As
per Hansen (1982), the optimal weighting matrix is the inverse of the asymptotic variance matrix, \( W = S^{-1} \).

A singularity among the moments clearly would impede the use of the optimal weighting matrix. Table 1 illustrates the condition number of \( S \) averaged over 100 sample paths. The Table is separated into three panels, each corresponding to a different sample size \( T = \{500, 1000, 2500\} \). The first column indicates the number of GMM iterations used in calculating \( S \). The columns labeled "Cond(S)" capture the condition number of \( S \). Please note that I disregard all paths for which the condition number is infinite in computing this average. The columns labeled "% Fail" capture the number of the 100 sample paths for which \( S \) is singular, and thereby precludes GMM estimation.

Consistent, albeit inefficient, estimation of the model is possible via a single GMM iteration in all of the simulated paths. However, attempting to proceed beyond one GMM iteration is difficult. For two GMM iterations, the average condition number of the variance matrix is quite large for the case where the sample size is 500, \((1.05e+19)\). Moreover, approximately 64% of the simulated paths generate variance matrices \( (S) \) that are singular. Notice, that this does not appear to be a small sample issue, since large condition numbers persist even as I increase the sample size. In addition, my research suggests that the magnitude of the problem increases as I consider \( H_3^0 \) and \( H_4^0 \).

In Figure 5 I calibrate the Tikhonov Factor for \( H_3^0 \) for a given sample path from the example above. The solid red line is the chosen tolerance level of 0.01. The starred blue line is the norm of the difference in the parameter vector. In this example, an \( \alpha^* \) of about 1.5 is appropriate.

### 4.3 Determining the Number of Factors

I generate \( H \) sample paths of size \( T \) for the return series \( y_t \) as given in equation (16), where \( n = 5 \). The parameters for the variance process detailed in equation (17) are \( (\omega, \alpha, \beta) = (0.03, 0.06, 0.91) \).

I use a "regularized" HAC variance matrix and an instrument vector \( z_t = [1 \ y_{1,t}^2]' \). The loading vectors used vary according to whether size or power of the test is the object of interest.

Determining the empirical size is straightforward. I generate one factor and consider the following loading vector: \( \Lambda_S^{(a)} = (10 \ 9 \ 8 \ 7 \ 6)' \) In this way, the assets differ primarily by their weighting on the factors. My sequence of hypotheses are as follows:

\[
H_1^0: \text{1 factor is sufficient for assets } (y^{(1)}, y^{(2)}) \\
H_1^a: \text{1 factor is not sufficient for assets } (y^{(1)}, y^{(2)}) \\
\vdots \\
H_4^0: \text{1 factor is sufficient for assets } (y^{(1)}, y^{(2)}, \ldots, y^{(5)}) \\
H_4^a: \text{1 factor is not sufficient for assets } (y^{(1)}, y^{(2)}, \ldots, y^{(5)})
\]

If I reject the null at any stage along the sequence, it is considered a failure. I then count the number of

---

*Condition Number of matrix \( A \) is computed as \( \text{cond}(A) = \| A^{-1} \| \cdot \| A \| \), where I use the frobenius norm.*
failures as a proportion of the number of simulated paths to yield a measure of empirical size.

I test power against several different alternatives, each containing a different number of GARCH factors. Initially, I consider an alternative where the returns are built from two GARCH(1,1) factors. The first factor is as detailed above. The second factor takes a similar form, except that \((\omega, \alpha, \beta) = (0.05, 0.10, 0.85)\). The factors have zero correlation by design. The loading vector \(\Lambda_a = [\Lambda_S^a, \Lambda_S^b] \odot [I_2, \iota_{2,3}]\), where \(I_2\) is a \((2 \times 2)\) identity matrix and \(\iota_{2,3}\) is an \((2 \times 3)\) matrix of ones.\(^7\)

The shape of this loading vector is useful for identification purposes, as mentioned in section 3.3. Also, as discussed in section 4.1, the magnitude of the elements of the loading vector correspond to the amount of variation the factors are able to capture.

As when testing empirical size, a rejection of the null anywhere along the sequence of tests is considered a failure. The probability of rejection is then measured as the number of failures as a proportion of the number of simulated paths.

The base case for the balance of this exercise involves simulating return paths via equation (2) with the following: (number of simulated paths) \(H = 250\), (sample size) \(T = 1000\), \(z_t = y_{1,t}^2\), \(\bar{y} = y^{(1)}\), loading vector \(\Lambda^a\), and weighting matrix equal to the inverse of the regularized HAC variance matrix. For a single path I conduct an ARCH(1) LM test and gather the p-values for each portfolio. I repeat this test for all 250 sample paths. The first column of Table 3 records the p-values averaged across the 250 paths. The small p-values suggest I can reject the null of No ARCH effects for all 5 portfolios. In addition, Figure 6 illustrates the size and power of \(H_0^4\) versus a two factor alternative. Notice that the test has only slight size distortions and has excellent levels of power.

These findings are robust to various characterizations of the model. Table 2 details the empirical size and power of \(H_0^4\) as I alter the model relative to the base case. All calculations are shown for a 10% nominal size. I summarize the results here.

The first panel depicts a reasonably sized test when I use \(\bar{y} = \{y^{(1)}, y^{(2)}, y^{(3)}\}\). However, there are size distortions when using \(\bar{y} = \{y^{(4)}, y^{(5)}\}\). In all cases, the test is powerful.

The second panel of Table 2 suggests that there are slight size distortions when the sample grows large, yet the test is quite powerful in all cases considered.

The third panel indicates that both empirical size and power increase as the magnitude of the loading vector \(\Lambda\) increases.

The fourth panel of Table 2 suggests that the power increases from an already high level as the alternative hypothesis considered expands from a system constructed from 2 factors to 3 or 4 factors.

The last panel varies the choice and number of instruments used in estimation. The first five rows of the panel suggest that the size and power results are relatively robust to instrument choice. However, increasing the dimension of the instrument vector to include more assets causes the model to reject too often. The last row of the panel captures the case where the instrument vector includes a vector of ones appended by the

---

\(^7\)The symbol \(\odot\) is used to represent the Hadamard product.
square of the one period lagged value for each of the five assets. Empirical size is 1.0, implying that the null of a single factor is rejected in each and every simulated path. Clearly, an instrument vector with a single asset is preferable.

For each simulated path I form the time series associated with four auxiliary portfolios of the form \((\bar{y}_{t+1} - B\bar{y}_{t+1})\). As suggested earlier, these portfolios should be conditionally homoscedastic. I confirm this hypothesis by searching for ARCH(1) effects via an LM test. The p-values from the test are averaged over the 250 sample paths for each of the portfolios. The first column of Table 3 captures the p-values from the ARCH LM test for the five base assets, while the second column captures the p-values for the four auxiliary portfolios. The typical p-value rises from about 0.10 in the base assets to about 0.50 among the auxiliary portfolios, indicating that no discernible ARCH(1) effects remain in the auxiliary portfolios.

In addition, I examine each of the auxiliary portfolios for DCC effects. The first column of Table 4 captures the p-values from the DCC test for the five base assets. The second column of the Table captures the p-value from the DCC test for the four auxiliary portfolios. The first row of the Table pertains to the full sample of 250 simulations. The p-values listed are averaged over the 250 simulated paths. For the full sample, the p-value for the auxiliary portfolios is about 0.09 percentage points higher (0.733 vs. 0.644) for the auxiliary portfolios than for the base assets, which implies that the auxiliary portfolios are able to alleviate some of the time variation in the correlation matrix of the assets. However, this is not a difficult feat since there wasn’t strong evidence of DCC effects in the base assets. To gauge better the ability of this procedure to accommodate the dynamics of the five asset system, I examine only those simulated paths for which DCC effects are strong. The second row of Table 4 pertains only to the 14 simulated paths for which the p-value is less than 0.15. The average p-value of 0.094 suggests strong evidence of DCC effects among the base assets. The auxiliary portfolios pertaining to these 14 paths has an average p-value of 0.641, thereby eliminating the DCC effects. When DCC effects are present, Phase1 is able to accommodate for the time varying correlation matrix.

I conduct a second exercise to determine how well the model can identify the number of latent factors in a system of asset returns. Once again consider the base case simulation. Each row of Table 5 represents the first element used to define \(\bar{y}\). Each column indicates the number of factors determined by the sequential testing procedure. The entries indicate the frequency of simulated paths for which Phase 1 determined a particular number of factors. For instance, row one / column one indicates that only 83 out of the 100 simulated paths suggest that one factor is sufficient to accommodate the conditional heteroscedasticity present in the five asset system while using \(y^{(1)}\) as the first column of \(\bar{y}_{t+1}\). Table 6 is identical in structure except that the true simulated series is built from two factors rather than one. Specifically, the assets are built from two GARCH(1,1) factors with parameters \((\omega, \alpha, \beta) = (0.03, 0.06, 0.91)\) and \((0.05, 0.10, 0.85)\), where I use a loading vector \(\Lambda = [\Lambda^a \Lambda^b] \odot [I_2 \ i_{2,3}]^\prime\).

Table 5 indicates that a vast majority of the simulations identify correctly the number of factors in the one-factor system. In addition, the model rarely misidentifies this one factor system for two factors. However, three to four factors are chosen mistakenly in (15 – 20%) of the simulations. Table 6 suggests that very seldom does Phase 1 underestimate the number of factors in the two-factor system. For example, when \(\bar{s} = s_1\), only 2% of the simulated paths chose a single factor as sufficient. Setting \(\bar{y} = s^{(1)}\) or \(y^{(2)}\), the two
most conditionally heteroscedastic assets by definition, seems to work best in the simulations. 98% (97%) of the simulations correctly pick a single factor when setting $\pi = s^{(1)}(y^{(2)})$. However, as the choice of $\pi$ moves to less conditionally heteroscedastic assets, the number of factors tends to be overestimated. For example, when $\pi = y^{(3)}$, 34% of the simulations incorrectly pick three factors and 2% pick four factors. These findings are robust to instrument choice. Although not shown here, as I vary the instrument vector, the accuracy of Phase 1 remains consistent with the results reported. Once again, using too many instruments is detrimental. When I use an instrument vector that includes lagged squared values of all of the five assets, Phase 1 incorrectly chooses 5 factors for about 90% of the simulations.

4.4 Full Model Estimation

This section examines the efficacy of Phase 2 of the estimation procedure through a Monte Carlo - type analysis. I compare the parameter estimates of the unconditional mean and variance of returns, $\mu$ and $\Lambda'\Lambda + \Omega$ respectively, to the true simulated values over several paths.

I also extract the latent factors implied by the model and compare these to the true simulated factors. With the factors now ‘observed’, I can model them via conventional time-varying volatility techniques such as GARCH or SV models, and compute conditional forecasts. This opens the door to a Markowitz-style portfolio optimization, where I create tracking portfolios of the dynamic allocation strategy implied by the MVFSV model. I utilize this tracking portfolio approach in the Empirical Application section of this paper.

I begin the evaluation by simulating a system of five asset returns corresponding to our base case from the previous section. The persistence parameters associated with the single common factor are $(\omega, \alpha, \beta) = (0.03, 0.06, 0.91)$, with loading vector $\Lambda^a = [10; 9; 8; 7; 6]$.

In order to judge Phase 2 fairly, I isolate it from any potential errors that may arise during Phase 1. I accomplish this by assuming that the number of factors is given, thereby avoiding Phase 1 entirely.

This single factor structure yields the following set of moment conditions for a model with a constant risk premium:

$$
E_t[y_{t+1}] = \mu \\
vec E_t[(\bar{y}_{t+1} - B\bar{y}_{t+1})y'_{t+1}] = \text{vec}((\bar{y} - B\bar{y})\mu' + (\Omega_2 - B\Omega_1)) \\
vech E_{t-1}[(1 - \gamma_1 L)(y_{t+1}y'_{t+1})] = \text{vech}(\mu\mu' + \Lambda'\Lambda + \Omega)(1 - \gamma_1)
$$

A Tikhonov Regularization technique once again is required to overcome a near singularity among the moment conditions.

Evaluating the accuracy of the unconditional mean and variance of returns, $\mu$ and $\Lambda'\Lambda + \Omega$, is relatively straightforward since these values are set explicitly in the simulation design. The unconditional mean of each series is zero. Moreover, with the error terms having unit variance and the loading vector being $[10; 9; 8; 7; 6]$, the diagonal elements of the unconditional variance/covariance matrix $\Lambda'\Lambda + \Omega$ are $[101; 82; 65; 50; 37]$.

Panels A & B of Table 7 detail the estimates of $\mu$ and $\Lambda'\Lambda + \Omega$ averaged over 250 simulated paths for a sample size of $T = 500$. The column labeled ‘MVFSV’ captures estimates from the MVFSV model, and the
column labeled ‘Simulated’ captures those values actually simulated. The average values of \( \hat{\mu} \) are somewhat negatively biased, but still quite close to the true value of zero. Similarly, the average values of \( \Lambda \hat{\Lambda} + \Omega \) are very close to the true value, with the sample estimates taking values [101.2; 83.3; 65.9; 50.3; 37.1]. Panels A & B of Table 8 list the results from the same exercise undertaken with a sample size of \( T = 1,000 \). The estimates of the mean improve slightly, while the variance estimates are basically unchanged.\(^8\)

Evaluating the accuracy of parameters associated with the loading vectors (\( B \)), the factor volatilities (\( \gamma \)), and the factors themselves (\( f_t \)), requires some work.

Recall the conditional variance specification:

\[
\Sigma_t = \Lambda D_t \Lambda' + \Omega
\]

(19)

and the fact that we can characterize the loading vector as \( \Lambda = [I \ B]' \Xi \). Also recall that while \( B \) is identified by the model, \( \Lambda \) is not. In fact, only the range of \( \Lambda \) is identified. This is due primarily to the possibility of a non-diagonal variance of the errors, \( \Omega \). This non-diagonality implies that any constant part of the conditional variance of the factors can be transferred to the variance of the errors by a simple rescaling of the loading vectors.

For the purposes of evaluating the model, I take advantage of this re-scaling property by normalizing \( \Xi \), without any identifiable impact on \( \Sigma_t \). The normalization scheme I use exploits the link between the factors and their loading vectors as well as the fact that the factors have unit unconditional variance by design.

The normalized value of the loading vector is computed as \( \hat{\Lambda} = [I \hat{B}]' \hat{\Xi} \), where \( \hat{B} \) is the GMM estimate. Moreover, the normalized error variance is then easily computed as \( \hat{\Omega} = \hat{\Lambda} \hat{\Lambda}' + \Omega - \hat{\Lambda} \hat{\Lambda}' \).

For ease of exposition, I set \( \Xi = I_k * a \), where \( a \) is some constant. For a given \( a \), we can normalize the loading vector \( \hat{\Lambda}(a) \) and extract the latent factors via simple OLS regressions of \( y_t - \hat{\mu} \) on \( \hat{\Lambda}(a) \) for each period \( t = 1, \ldots, T \). The factor takes the usual form \( \hat{f}(a)_t = (\hat{\Lambda}(a)' \hat{\Lambda}(a))^{-1} \hat{\Lambda}(a)'(y_t - \hat{\mu}) \). I choose \( a \) to minimize the distance between the sample variance of the factors and the population expectation of 1. In other words, \( a = \text{argmin}_{a \in A} \| (\frac{1}{T} \sum_{t=1}^{T} \hat{f}(a)_t - \hat{f}(a))' \|^2 \).

Panel C of Tables 7 and 8 illustrate the estimated value of the loading vector averaged over 250 paths for \( T = 500 \) and \( T = 1,000 \), respectively. In either case, the estimates of the loading vector are quite accurate, both in terms of order of magnitude and shape; i.e. the first element is the largest and the last element is the smallest.

Panel D of Tables 7 and 8 compares the unconditional moments of the simulated factors with the extracted factors after normalization. Regardless of sample size, the first four sample moments of the extracted factors match remarkably well those of the simulated factors. For instance, the kurtosis of the extracted factor is 3.250 averaged over the 250 paths, which matches precisely the kurtosis of the simulated factors. Moreover, the correlation between the extracted and simulated factors is above 0.95 in both sample sizes.

---

\(^8\)Note: The column labeled ‘Simulated’ contains the actual sample mean and variance of the simulated assets, which may differ slightly from the true values set in the experiment design due to a finite sample bias during simulation.
Panel E of Tables and 7 and 8 compare the conditional behavior of the extracted factors. Given the SR-SRAV(1) specification of the volatility dynamics, we know \( \gamma = \alpha + \beta \) from a GARCH(1,1) model. The simulated values of \( \alpha = 0.063 \) and \( \beta = 0.762 \) imply a simulated value of \( \gamma = 0.825 \), for the case of \( T = 500 \). The estimates of \( \alpha \) and \( \beta \) are quite good; 0.06 and 0.754, respectively. In addition, these estimates improve as we increase the sample size. \( \gamma \) are quite poor; taking values of 0.082 and 0.097, for \( T = 500 \), and \( T = 1,000 \), respectively.

As mentioned previously, one of the aims of extracting the factors is to engage in a dynamic asset allocation strategy. Under the simple case of a constant risk premium all of the dynamics of our model reside with \( \Sigma_t \). The investor can be considered a volatility-only forecaster. Thus, all we require are forecasts of the conditional variance/covariance matrix of returns; \( \hat{\Sigma}_t = \hat{\Lambda}(a)Var(\hat{f}(a)_t)\hat{\Lambda}(a) + \hat{\Omega} \). Following the procedure just detailed, these forecasts are built around factors extracted thanks to a convenient normalization scheme and an estimate of B. Fortunately, such an estimate is availed from Phase 1 as well as Phase 2.

Simulation evidence suggests that the estimates of \( B \) from Phases 1 & 2 match each other rather well. For example, with a sample size of \( T = 500 \), the average value of \( B \) estimated from Phase 2 is: \( \hat{B}^{(2)} = [0.85, 0.80, 0.70, 0.61]' \). The average difference between these estimates and those from Phase 1 (\( \hat{B}^{(1)} \)) are relatively small. Specifically, the average of the element by element difference in the estimates \( \{\frac{1}{H} \sum_{h=1}^{H} (\hat{B}^{(1)} - \hat{B}^{(2)})\} = [0.055, 0.090, 0.049, 0.088]' \), where \( H \) is the simulation size of 250. This average difference improves as we increase the sample size. For \( T = 1,000 \), the average difference is \( [0.054, 0.032, 0.001, -0.004]' \). These results suggest that we can extract safely the factors using estimates of \( B \) from Phase 1 only, and bypass the potentially time consuming estimation of Phase 2.

5 Empirical Application

In this section I use the MVFSV model to investigate the dynamics of the U.S. equity market. I examine portfolios representing the 12 sectors of the U.S. economy. I determine how many factors are required to accommodate the conditional heteroscedasticity in this system of returns. I then use the MVFSV model to forecast the conditional variance matrix of returns. These forecasts pave the way for a dynamic asset allocation strategy. I will track the investment strategy and compare its performance to a strategy implied by the DCC model.

5.1 Data

I segment the CRSP universe of U.S. equities into twelve value weighted sector portfolios for every trading day from 01/02/1990 through 12/31/2006. The portfolios are formed via a time series of the two-digit SIC codes obtained from Ken French’s website. Table 9 provides summary statistics, illustrating the non-Gaussianity, negative skewness and excess kurtosis typical of daily stock returns.

Figures 7, 8, 9 depict the cumulative returns of each of the 12 sectors over the 17 year period. Each graph illustrates the growth of a single dollar invested on 01/02/1990 and held for the entire period.
The dynamic allocation strategies I will detail allow for the presence of a risk free rate. I follow French, Schwert, and Stambaugh (1987) by using 1-month U.S. Treasury yields as a standard proxy. Specifically, monthly yields are gathered from the average of bid and ask prices for the U.S. government security that matures closest to the end of the month. Daily yields are then computed by dividing the monthly yield by the number of trading days in the month.

5.2 Estimation & Results

The Empirical Application’s design centers around an investor who trusts the dynamic nature of their model. Each month the investor estimates the model and generates one month’s worth (approximately 21 trading days) of conditional variance forecasts. With these forecasts, the investor solves a portfolio optimization problem and re-balances their portfolio every trading day.

Given the simulation evidence seen in the previous section of this paper, the investor believes it is sufficient to estimate only Phase 1 of the MVFSV model. The investor uses three years (approximately 756 trading days) worth of past returns data for estimation.

The investor relies upon the sequential testing procedure of Phase 1 to determine the proper ordering of the sector portfolios. Twelve choices are available to begin the sequence. The investor examines each choice and selects the ordering that achieves the best balance between the size of the factor structure and the model’s ability to accommodate the conditional heteroscedasticity in the returns. Specifically, the investor chooses the smallest factor representation that has auxiliary portfolios that are on average less conditionally heteroscedastic than the original sectors, as measured by individual LM tests, and exhibit less dynamic conditional correlation, as measured by a joint test for constant conditional correlation.

The first estimation period runs from 01/02/1990 through 01/02/1993. I estimate Phase 1, pick the ordering of the sectors that meets the selection rule described above, form the auxiliary portfolios, and test them for ARCH and DCC effects. I then roll the investor’s estimation window forward one month, using returns from 02/01/1990 through 02/01/1993. I re-estimate the model and form new auxiliary portfolios. This process continues for a total of 168 times until the end of the last estimation period is 12/31/2006.

Figure 10 depicts the number of factors selected by the investor at each of the 168 estimation points. As intuition would suggest, the relative tranquility of the economy and market in the mid 1990’s generally required only a single factor. However, the turbulent times of late 90’s and early 2000’s required as many as seven factors. It is in times such as these that a factor analytic approach might not be useful as it removes the key advantage of parsimony in the model.

The auxiliary portfolios formed from these factor structures are able to accommodate the individual ARCH effects among the assets rather well. Fig 11 illustrates the average p-values from ARCH(1) tests. The solid line is the average of the 12 p-values from the LM test on the sector portfolios. The stars are the average of the (n-k) p-values implied by the auxiliary portfolios. The null on the LM test for ARCH(1) effects is that no ARCH is present. Therefore, a low p-value is suggestive of strong ARCH effects, while a large p-value is suggestive of conditional homoscedasticity. The relative tranquil period of the mid-1990’s, when only a
single factor was needed, was able to accommodate the ARCH effects quite well. In 1995 for instance, the average p-value for the sector portfolios was approximately zero. Yet, the average p-value of the auxiliary portfolios was as high as 0.55.

Fig 12 illustrates the auxiliary portfolios’ ability to accommodate the joint dynamics in the data. The solid line represents the p-value from a test for constant conditional correlation averaged over the twelve sector portfolios. The stars again represent the p-value averaged over the (n-k) auxiliary portfolios. The null hypothesis of this test is that the returns exhibit constant conditional correlation. Therefore, a low average p-value is indicative of DCC effects. The MVFSV model is unable to accommodate for the strong DCC effects during the early 1990’s. However, in the mid-1990’s DCC effects are handled quite well. The model’s ability to generate auxiliary portfolios with constant conditional correlation is somewhat erratic during the late 1990s boom. And interestingly, there is little evidence of DCC effects in the sector portfolios during the 2003-2006 period in the sector portfolios.

As described in the Simulation section, the estimates of $B$ allow us to normalize the loading vector, extract the factors, and forecast the conditional variance matrix of returns using Phase 1 of the model. Forecasts of the mean of returns, on the other hand, are computed simply as the sample average over the estimation period. As such, the investor can be characterized as a volatility-only forecaster. With these forecasts, the investor conducts a Markowitz-style portfolio optimization. Their objective function is to minimize volatility of the portfolio subject to a 10% target rate of return. I allow for short sales as well as for a time varying risk free rate, proxied for by the 1-month U.S. Treasury bill. Details of the investor’s objective functions and the analytic solution for the portfolio weights are as per Han (’05) and are found in the Appendix.

I compare the portfolio implied by the MVFSV model against two portfolios. First, is a Buy & Hold benchmark consisting of the return on a value-weighted portfolio of the entire CRSP universe. Second, I consider a portfolio formed from forecasts made by the DCC model. The MVFSV and DCC models are referred to as dynamic strategies since they involve dynamically updating the portfolio each trading day.

Several criteria are used to evaluate the performance of the Buy & Hold and dynamic strategies. $\mu_p$ is the average daily return of the portfolio. $\sigma_p$ is the standard deviation of the daily portfolio returns. $SR = (\mu_p - r_f)/\sigma_p$ is an ex-post Sharpe Ratio of portfolio returns. The Sharpe Ratio, however, is known to underestimate the performance of dynamic strategies because it does not account for time variation in conditional volatility of the portfolio. Therefore, I also consider $M2 = \sigma_b(\text{SR}_p - \text{SR}_b)$, where the $b$ subscript indicates the static Buy & Hold benchmark. $M2$ can be seen as a risk adjusted measure of abnormal return. $IR = \frac{\mu_p - \mu_b}{\sigma_p - \sigma_b}$, is the Information Ratio. A higher Information Ratio is indicative of better performance relative to the benchmark. $Z$ is the percentage of trading days the dynamic strategy has a higher return than the benchmark strategy, providing an indication as to whether out-performance is due to a few large outliers. Finally, $\phi$ is the performance-fee measure of Fleming, Kirby, and Ostdiek (2001). It measures the hypothetical fee that an investor would be willing to pay to switch from the benchmark static strategy to a given dynamic strategy, without being worse off. The performance fee $\phi$ solves:

$$\frac{1}{T} \sum_{t=0}^{T-1} [(R_{d,t+1} - \phi) - \frac{\gamma}{2(1+\gamma)}(R_{d,t+1} - \phi)^2] = \frac{1}{T} \sum_{t=0}^{T-1} [\gamma \frac{2(1+\gamma)}{R^2_{b,t+1}}]$$
where $R_{d,t+1}$ and $R_{b,t+1}$ are the realized gross returns of the dynamic and static benchmark strategies, respectively.

I consider trading costs as per Marquering and Verbeek (2004), where costs are equal to $\tau$ percentage points of the value traded. Total transactions costs are captured by $\tau W_t |\Delta \omega_{t+1}|$, where $W_t$ denotes wealth at time $t$, and $\Delta \omega_{t+1} = \omega_{t+1} - \omega_t$. Returns after transaction costs are measured by $r_{p,t+1} - \tau |\Delta \omega_{t+1}|$. I follow Marquering and Verbeek (2004) by considering low, medium, and high transaction costs of 0.1%, 0.5%, and 1.0% of the value traded. The Buy & Hold benchmark portfolio is assumed to have no transaction costs.

Figures ?? through ?? illustrate the performance of the dynamic allocation strategies for the minimum volatility investor at each of the four levels of trading costs. Four criteria are selected to summarize succinctly the relative performance of the portfolios. The upper left panel in each figure contains $\pi_p$, the daily return averaged over the roughly 252 trading days in each year from 1993 through 2006. Units are in percent, so that 0.1 is 0.1% per day. The top right panel illustrates the average daily standard deviations $\sigma_p$. The bottom left and right illustrate the M2 and $\phi$ measures, respectively.

Focus first on Figure 13. In terms of absolute daily returns, the dynamic strategies generally do well relative to the benchmark. The year 1993 is an exception, where the dynamic strategies lose money on average, with the MVFSV model performing worse than the DCC. The mid-90’s through early 2000 saw the MVFSV model outperform the benchmark as well as the DCC model. The years 2001 through 2006 were turbulent times for both the market and the dynamic strategies. According to the M2 measure, the MVFSV model appears more attractive than the DCC model in just about every year. And the $\phi$ performance fee measure provides somewhat mixed results. In all, the dynamic strategies perform well relative to the benchmark, and the MVFSV performs better than DCC.

The MVFSV model’s performance, however, is accompanied by a large amount of turnover in the portfolio. To illustrate, consider a measure of total portfolio activity (TPA) as follows: $TA = \sum_{i=1}^{n} |w_{i,t+1} - w_{i,t}|$, where $n$ is the number of assets in the portfolio, and $w$ are the asset weights. Figure 17 illustrates the total portfolio activity for the MVFSV and DCC models, smoothed over a rolling three month period for illustrative purposes.

During the early 1990’s, the MVFSV model exhibits modestly more activity than the DCC model. Roughly equal levels of activity are seen during the mid 1990’s. However, the boom/bust period of late 90’s / early 2000’s and beyond are plagued with very large amounts of portfolio activity for the MVFSV strategy.

Such high levels of activity clearly will impact portfolio performance when transactions costs are considered. For example, at low and medium levels of transactions costs, as seen in Figures 14 and 15, the dynamic strategies’ absolute performance tends to worsen. Moreover, the gap between the MVFSV and DCC models narrows. Once a high level of transactions costs are considered, as seen in Figure 16, the MVFSV model performs quite poorly along most measures.

In summary, the MVFSV model generates portfolios that outperform the DCC model along several performance criteria. Transaction costs, however, are important. The presence of transaction costs could have

---

9The general theme reported here does not change as we consider the other performance criteria mentioned.
serious effects on absolute performance and cause the DCC model to outperform the MVFSV on a relative basis.

This drastic deterioration of daily returns of the MVFSV based portfolio may be associated with the model’s (in)ability to capture adequately the conditional covariances during the boom/bust of the late 1990’s / early 2000’s. For example, recall Fig 8, which depicts the cumulative returns of several sectors of the economy. In particular, notice the the Business Equipment sector experienced tremendous returns during the late 1990’s, followed by an equally as dramatic fall.

Figures 18 and 19 illustrate the elements of the conditional variance matrix that are associated with the Business Equipment sector. The solid red lines depict the elements of the conditional variance matrix as implied by the MVFSV model, while the dashed green line depicts those implied by the DCC model.

In general, the DCC and MVFSV estimates of volatility coincide quite nicely during the boom/bust period. Notice that conditional covariance between Business Equipment and Shops or Non-Durables are cases in point. However, there are several elements of the conditional variance matrix that exhibit significant differences between the DCC and MVFSV models. For instance, the Business Equipment sector is extremely volatile during the boom/bust period, regardless of the model used. However, the variance implied by the MVFSV model is less volatile than that of the DCC model. Moreover, notice the discrepancy between the estimates of the covariance between the Business Equipment sector and the Manufacturing sector. These differences may contribute to the relative under-performance of the MVFSV model during the late 1990’s.

6 Conclusion

In this paper I illustrate that the multivariate latent factor stochastic volatility model (MVFSV) of Doz and Renault (2006) can be implemented successfully, and yield potentially significant investment implications. It offers a parsimonious way to capture the joint dynamics of a system of asset returns without resorting to a full parameterization of conditional probability distributions. The issue of comparing the statistical fit of this MVFSV model to other factor stochastic volatility models remains unanswered.

An obvious extension of this paper is to include a time varying risk premium in Phase 2, as outlined by Doz and Renault (2006). This may improve the tracking portfolio performance as well as open the door to testing the price of idiosyncratic risk as in King, Sentana, and Wadhwani (1994).

7 Appendix

7.1 Solutions to the Portfolio Allocation Problem

I follow Han (2006) in designing the portfolio allocation problem faced by the investor. Let $\mu_{p,t+1}$ and $\sigma^2_{p,t+1}$ denote the conditional mean and variance of the portfolio returns $r_{p,t+1}$. The objective function is minimum volatility subject to a target rate of return ($\mu^*_p$). The investor allows for short sales and the existence of a
risk free rate \((r_f)\). All other variables are as defined in the body of this paper. The objective function and solution for the portfolio weights \((w)\) are below:

**Minimum Volatility:**

\[
\min_{w_t} \{ \sigma_{p,t+1}^2 = w_t' \Sigma_{t|t+1} w_t \} 
\]

subject to:

\[
w_t' \mu_{t+1|t} + (1 - w_t' 1) r_f = \mu_p^*
\]

Define \(\kappa_t = (\mu_{t+1|t} - 1 r_f)' \Sigma_{t+1|t} (\mu_{t+1|t} - 1 r_f)\). The portfolio weights are:

\[
w_t = \Sigma_{t|t+1}^{-1} \left( \mu_{t+1|t} - 1 r_f \right) \frac{\mu_p^* - r_f}{\kappa_t}
\]
References


8 Tables & Graphs

Figure 1: Equally weighted average of cross-sector correlation coefficients. Computed over preceding three months for 12 sectors of the U.S. economy.
Figure 2: Points to the left of the solid line indicate \((\alpha, \beta)\) combinations for which fourth moments of returns exist. The box indicates \((\alpha, \beta)\) combinations that are realistic given daily U.S. equity returns.

Figure 3: Simulation of 100 paths of a 750 observation GARCH(1,1) process for a range of \(\alpha\) and \(\beta\) values. I conduct an ARCH LM test and compute power as the proportion of the simulations for which the null of no ARCH effects is rejected. The right hand figure tests for ARCH(2), while the left hand side tests for ARCH(1).
Figure 4: P-values from ARCH LM test on $\hat{u}_t$, where $u_t$ are errors in the regression $Y_t = \lambda_i f_t + u_t$. The sample size is set to 25,000, $\lambda_i \in \{0, 0.1, ..., 5\}$ and $f_t$ is an ARCH(1) process with $\omega = 0.01$ and $\alpha = 0.50$ in order to approximate the sectors returns used in the empirical application section of this paper.
Table 1: Cond(S) is the condition number of the variance matrix of moments. %Fail is the proportion of simulated paths for which the GMM optimization fails to converge due to a singularity problem among the moments.

<table>
<thead>
<tr>
<th>GMM It.</th>
<th>Cond(S)</th>
<th>% Fail</th>
<th>Cond(S)</th>
<th>% Fail</th>
<th>Cond(S)</th>
<th>% Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T=500</td>
<td>T=1000</td>
<td>T=2500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.94e+19</td>
<td>0%</td>
<td>1.59e+19</td>
<td>0%</td>
<td>2.84e+19</td>
<td>0%</td>
</tr>
<tr>
<td>2</td>
<td>1.05e+19</td>
<td>64%</td>
<td>7.12e+18</td>
<td>69%</td>
<td>3.67e+18</td>
<td>75%</td>
</tr>
</tbody>
</table>

Figure 5: Calibrating the Tikhonov Factor for $H^2_o$ in a five asset example. The starred line is the normmed % change in the parameter vector from using successively larger Tikhonov Factors during estimation. Choose the Tikhonov Factor such that $\| \Theta(i+1) - \Theta(i) \| > tol$, where the tolerance (tol) is the solid line in the graph.
Figure 6: Size and Power of over-identified restrictions test in Phase 1 to determine the number of common factors. 250 paths of a five asset system from a single GARCH(1,1) factor with loading vector $\Lambda^o = [10, 9, 8, 7, 6]$ are simulated. The proportion of sample paths that reject the null of a single common factor is used to compute the empirical size of the test. The power is calculated similarly using a two factor alternative.
Table 2: Testing the empirical size and power of the overidentified restrictions test associated with the search for the number of common factors. The analysis pertains to a simulated system of five asset returns built from a single GARCH(1,1) factor. The base case for the study is $y = y^{(1)}$, $T=1000$, $\Lambda = \Lambda^a$, and $z_t = (y_t^{(1)})^2$, $H_a = 2$ factors. The top panel alters $y$, the second panel alters sample size, the third panel alters the loading vectors, the fourth panel alters the number of factors in the alternative hypothesis, and the fifth panel alters the instrument. Nominal size is 10%.

<table>
<thead>
<tr>
<th>$y$</th>
<th>Size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y = y^{(1)}$</td>
<td>0.152</td>
<td>0.952</td>
</tr>
<tr>
<td>$y = y^{(2)}$</td>
<td>0.148</td>
<td>0.996</td>
</tr>
<tr>
<td>$y = y^{(3)}$</td>
<td>0.164</td>
<td>0.988</td>
</tr>
<tr>
<td>$y = y^{(4)}$</td>
<td>0.312</td>
<td>0.980</td>
</tr>
<tr>
<td>$y = y^{(5)}$</td>
<td>0.336</td>
<td>0.992</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>Size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>0.124</td>
<td>0.952</td>
</tr>
<tr>
<td>1000</td>
<td>0.152</td>
<td>0.952</td>
</tr>
<tr>
<td>2500</td>
<td>0.252</td>
<td>0.963</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Lambda$</th>
<th>Size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda^b = [5, 4, 3, 2, 1]$</td>
<td>0.112</td>
<td>0.812</td>
</tr>
<tr>
<td>$\Lambda^a = [10, 9, 8, 7, 6]$</td>
<td>0.152</td>
<td>0.952</td>
</tr>
<tr>
<td>$\Lambda^c = [20, 19, 18, 17, 16]$</td>
<td>0.296</td>
<td>0.996</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H_a$</th>
<th>Size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 factors</td>
<td>0.152</td>
<td>0.952</td>
</tr>
<tr>
<td>3 factors</td>
<td>-</td>
<td>1.00</td>
</tr>
<tr>
<td>4 factors</td>
<td>-</td>
<td>0.992</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$z$</th>
<th>Size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(y_t^{(1)})^2$</td>
<td>0.152</td>
<td>0.952</td>
</tr>
<tr>
<td>$(y_t^{(2)})^2$</td>
<td>0.104</td>
<td>1.0</td>
</tr>
<tr>
<td>$(y_t^{(3)})^2$</td>
<td>0.144</td>
<td>0.992</td>
</tr>
<tr>
<td>$(y_t^{(4)})^2$</td>
<td>0.144</td>
<td>0.992</td>
</tr>
<tr>
<td>$(y_t^{(5)})^2$</td>
<td>0.128</td>
<td>0.996</td>
</tr>
<tr>
<td>$(y_t^{(j)})^2 \forall j = 1, \ldots, 5$</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Table 3: P-Values from ARCH(1) LM tests. \( H_0 \): No ARCH Effect. Base Assets are five simulated asset returns \( y \). Auxiliary Portfolio are linear combinations of the Base Assets, which take the form \((\bar{y} - B\bar{y})\). The p-values are averaged over 250 simulated paths.

<table>
<thead>
<tr>
<th>Base Assets</th>
<th>Auxiliary Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.104</td>
</tr>
<tr>
<td>2</td>
<td>0.103</td>
</tr>
<tr>
<td>3</td>
<td>0.106</td>
</tr>
<tr>
<td>4</td>
<td>0.104</td>
</tr>
<tr>
<td>5</td>
<td>0.121</td>
</tr>
</tbody>
</table>

Table 4: P-Values from DCC Test. \( H_0 \): Constant Conditional Correlation. Base Assets are five simulated asset returns \( y \). Auxiliary Portfolios are linear combinations of the Base Assets, which take the form \((\bar{y} - B\bar{y})\). The p-values are averaged over 250 simulated paths. "Strong DCC" refer only to those simulations for which the p-value \( \leq 0.15 \).

<table>
<thead>
<tr>
<th>Base Assets</th>
<th>Auxiliary Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Sample</td>
<td>0.644</td>
</tr>
<tr>
<td>Strong DCC</td>
<td>0.094</td>
</tr>
</tbody>
</table>

Table 5: Simulate 100 sample paths of length 1000 for a five asset system. Built from a single GARCH(1,1) factor using loading vector \( \Lambda = \Lambda_\Lambda \), and estimated with \( z_t = (y_t^{(1)})^2 \). Use Phase 1 sequential procedure to determine \# of latent factors. Each row indicates the first element of \( \bar{y}_{t+1} \). Each column indicates signifies the number of factors suggested. For instance, entry (1,1) suggests that if we set the first element of \( \bar{y}_{t+1} \) equal to \( y_t^{(1)} \), then only 83 simulated paths suggest a single common factor is sufficient.

<table>
<thead>
<tr>
<th>( y_t^{(1)} )</th>
<th># of Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
<td>0</td>
</tr>
<tr>
<td>84</td>
<td>1</td>
</tr>
<tr>
<td>78</td>
<td>0</td>
</tr>
<tr>
<td>76</td>
<td>3</td>
</tr>
<tr>
<td>67</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 6: Simulate 100 sample paths of length 1000 for a five asset system. Built from two GARCH(1,1) factors using loading vector \( \Lambda = [\Lambda_\Lambda \ \Lambda_\Lambda] \odot [I_2 \ \I_2] \), and estimated with \( z_t = (y_t^{(1)})^2 \). Use Phase 1 sequential procedure to determine \# of latent factors. Each row indicates the first element of \( \bar{y}_{t+1} \). Each column indicates signifies the number of factors suggested. For instance, entry (1,1) suggests that if we set the first element of \( \bar{y}_{t+1} \) equal to \( y_t^{(1)} \), then only 2 simulated paths suggest a single common factor is sufficient.

<table>
<thead>
<tr>
<th>( y_t^{(1)} )</th>
<th># of Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>98</td>
</tr>
<tr>
<td>1</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>62</td>
</tr>
<tr>
<td>0</td>
<td>79</td>
</tr>
<tr>
<td>1</td>
<td>82</td>
</tr>
</tbody>
</table>
Table 7: Base Case simulation with 250 paths, each of length 500. Column labeled 'Model' contains the parameter estimates from Phase 2 of MVFSV. Column labeled 'Simulated' are the parameter values actually simulated. Statistics reported are averaged over all 250 paths. Panels A & B detail the unconditional mean and variance of returns, respectively. Panel C describes the estimated loading vector. Panel D illustrates the ARCH(\(\alpha\)) and GARCH(\(\beta\)) parameters from a GARCH(1,1) model of the factors.

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>Simulated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: (\mu)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(y^{(1)})</td>
<td>-0.016</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(2)})</td>
<td>-0.014</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(3)})</td>
<td>-0.016</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(4)})</td>
<td>-0.028</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(5)})</td>
<td>-0.014</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Panel B: (\Lambda\Lambda' + \Omega)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(y^{(1)})</td>
<td>101.240</td>
<td>103.387</td>
</tr>
<tr>
<td>(y^{(2)})</td>
<td>83.292</td>
<td>83.943</td>
</tr>
<tr>
<td>(y^{(3)})</td>
<td>65.854</td>
<td>66.504</td>
</tr>
<tr>
<td>(y^{(4)})</td>
<td>50.313</td>
<td>51.178</td>
</tr>
<tr>
<td>(y^{(5)})</td>
<td>37.019</td>
<td>37.835</td>
</tr>
<tr>
<td><strong>Panel C: (\Lambda)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(y^{(1)})</td>
<td>9.562</td>
<td>10.0</td>
</tr>
<tr>
<td>(y^{(2)})</td>
<td>7.810</td>
<td>9.0</td>
</tr>
<tr>
<td>(y^{(3)})</td>
<td>7.619</td>
<td>8.0</td>
</tr>
<tr>
<td>(y^{(4)})</td>
<td>6.522</td>
<td>7.0</td>
</tr>
<tr>
<td>(y^{(5)})</td>
<td>5.668</td>
<td>6.0</td>
</tr>
<tr>
<td><strong>Panel D: (f_t) - Unconditional</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Variance</td>
<td>1.000</td>
<td>1.024</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.020</td>
<td>-0.020</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.250</td>
<td>3.250</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.957</td>
<td>-</td>
</tr>
<tr>
<td><strong>Panel E: (f_t) - Conditional</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.060</td>
<td>0.063</td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.754</td>
<td>0.762</td>
</tr>
</tbody>
</table>
Table 8: Base Case simulation with 250 paths, each of length 1000. Column labeled 'Model' contains the parameter estimates from Phase 2 of MVFSV. Column labeled 'Simulated' are the parameter values actually simulated. Statistics reported are averaged over all 250 paths. Panels A & B detail the unconditional mean and variance of returns, respectively. Panel C describes the estimated loading vector. Panel D illustrates the ARCH(\(\alpha\)) and GARCH(\(\beta\)) parameters from a GARCH(1,1) model of the factors.

<table>
<thead>
<tr>
<th>Panel A: (\mu)</th>
<th>Model</th>
<th>Simulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>(y^{(1)})</td>
<td>0.018</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(2)})</td>
<td>0.010</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(3)})</td>
<td>0.015</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(4)})</td>
<td>-0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>(y^{(5)})</td>
<td>0.015</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: (\Lambda' + \Omega)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(y^{(1)})</td>
<td>101.240</td>
<td>101.619</td>
</tr>
<tr>
<td>(y^{(2)})</td>
<td>82.209</td>
<td>82.526</td>
</tr>
<tr>
<td>(y^{(3)})</td>
<td>64.987</td>
<td>65.411</td>
</tr>
<tr>
<td>(y^{(4)})</td>
<td>49.287</td>
<td>50.302</td>
</tr>
<tr>
<td>(y^{(5)})</td>
<td>36.443</td>
<td>37.254</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: (\Lambda)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(y^{(1)})</td>
<td>9.987</td>
</tr>
<tr>
<td>(y^{(2)})</td>
<td>8.443</td>
</tr>
<tr>
<td>(y^{(3)})</td>
<td>7.912</td>
</tr>
<tr>
<td>(y^{(4)})</td>
<td>6.815</td>
</tr>
<tr>
<td>(y^{(5)})</td>
<td>6.030</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: (f_t) - Unconditional</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Variance</td>
<td>1.000</td>
<td>1.006</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.016</td>
<td>-0.015</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.296</td>
<td>3.299</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.998</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel E: (f_t) - Conditional</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>0.060</td>
<td>0.061</td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.867</td>
<td>0.863</td>
</tr>
</tbody>
</table>
Table 9: Descriptive Statistics: Daily Returns of 12 sectors in the U.S. market over the period (01/02/1990 - 12/29/2006). Mean and Standard Deviation are annualized and presented in %; i.e. a mean return of 11.83 indicates a 11.83% annualized return. JB-Pval is the p-value from the Jarque-Bera test for normality, the null hypothesis of which is a Gaussian process.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>JB-Pval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Non-Durables</td>
<td>11.83</td>
<td>13.73</td>
<td>-0.06</td>
<td>6.58</td>
<td>0.00</td>
</tr>
<tr>
<td>2 Durables</td>
<td>9.81</td>
<td>19.98</td>
<td>-0.07</td>
<td>6.68</td>
<td>0.00</td>
</tr>
<tr>
<td>3 Manufacturing</td>
<td>13.71</td>
<td>16.25</td>
<td>-0.12</td>
<td>6.18</td>
<td>0.00</td>
</tr>
<tr>
<td>4 Energy</td>
<td>14.40</td>
<td>19.68</td>
<td>0.06</td>
<td>5.25</td>
<td>0.00</td>
</tr>
<tr>
<td>5 Chemicals</td>
<td>11.92</td>
<td>16.02</td>
<td>-0.47</td>
<td>13.83</td>
<td>0.00</td>
</tr>
<tr>
<td>6 Business Equipment</td>
<td>14.94</td>
<td>27.62</td>
<td>0.30</td>
<td>8.24</td>
<td>0.00</td>
</tr>
<tr>
<td>7 Telecommunications</td>
<td>7.95</td>
<td>19.13</td>
<td>-0.03</td>
<td>6.87</td>
<td>0.00</td>
</tr>
<tr>
<td>8 Utilities</td>
<td>10.79</td>
<td>13.97</td>
<td>-0.30</td>
<td>11.39</td>
<td>0.00</td>
</tr>
<tr>
<td>9 Shops</td>
<td>12.26</td>
<td>18.07</td>
<td>0.03</td>
<td>6.98</td>
<td>0.00</td>
</tr>
<tr>
<td>10 Health</td>
<td>13.12</td>
<td>18.48</td>
<td>-0.19</td>
<td>6.47</td>
<td>0.00</td>
</tr>
<tr>
<td>11 Money</td>
<td>15.75</td>
<td>17.81</td>
<td>0.13</td>
<td>7.04</td>
<td>0.00</td>
</tr>
<tr>
<td>12 Other</td>
<td>8.09</td>
<td>16.83</td>
<td>-0.23</td>
<td>9.33</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure 7: Cumulative returns of value weighted sector portfolios of the U.S. equity market. Growth of one dollar invested at 01/02/1990 and held through 12/29/2006.
Figure 8: Cumulative returns of value weighted sector portfolios of the U.S. equity market. Growth of one dollar invested at 01/02/1990 and held through 12/29/2006.

Figure 9: Cumulative returns of value weighted sector portfolios of the U.S. equity market. Growth of one dollar invested at 01/02/1990 and held through 12/29/2006.
Figure 10: # of factors implied by Phase 1 sequential testing procedure. Computed for a three year estimation period beginning at the dates marked.

Figure 11: P-values from ARCH(1) test computed over three period beginning at the dates marked. The solid blue line are the p-values at each date averaged over the 12 sector portfolios. The red stars are the p-values at each date averaged over the $(n - k)$ auxiliary portfolios.
Figure 12: P-values from test for Constant Conditional Correlation computed over three period beginning at the dates marked. The solid blue line are the p-values at each date for the 12 sector portfolios. The red stars are the p-values at each date for the \((n - k)\) auxiliary portfolios.
Figure 17: Total Portfolio Activity: $TPA = \sum_{i=1}^{n} |w_{i,t+1} - w_{i,t}|$, where $n$ is number of assets and $w$ are the asset weights.
Figure 18: Elements of the Conditional Covariance Matrix associated with the Business Equipment Sector. Solid Red line: MVFSV estimate. Dashed Green line: DCC estimate.
Figure 19: Elements of the Conditional Covariance Matrix associated with the Business Equipment Sector. Solid Red line is MVFSV estimate. Dashed Green line is DCC estimate.