Bayesian Dynamic Factor Models and Variance Matrix Discounting for Portfolio Allocation

Omar Aguilar and Mike West
ISDS, Duke University, Durham, NC 27708-0251

January 1998

Abstract

We discuss the development of dynamic factor models for multivariate financial time series, and the incorporation of stochastic volatility components for latent factor processes. Bayesian inference and computation is developed and explored in a study of the dynamic factor structure of daily spot exchange rates for a selection of international currencies. The models are direct generalisations of univariate stochastic volatility models, and represent specific varieties of models recently discussed in the growing multivariate stochastic volatility literature. We also discuss connections and comparisons with the much simpler method of dynamic variance discounting that, for over a decade, has been a standard approach in applied financial econometrics in the Bayesian forecasting world. We review empirical findings in applying these models to the exchange rate series, including aspects of model performance in dynamic portfolio allocation. We conclude with comments on the potential practical utility of structured factor models and future potential developments and model extensions.

Keywords: Dynamic Factor Analysis; Dynamic Linear Models; Exchange Rates Forecasting; Markov Chain Monte Carlo; Multivariate Stochastic Volatility; Portfolio Selection; Variance Matrix Discounting.

The authors acknowledge useful discussions with Jose M Quintana, Neil Shephard and Hong Chang, and partial support from NSF grants DMS-9704432 and DMS-9707914. This manuscript represents a preliminary draft report subject to revision. Before quoting or referencing, please check the authors' web site for a possible updated version. The latest version will be found as ISDS Discussion Paper 98-03 on the Duke web site, http://www.stat.duke.edu/papers/

1 Introduction

Since the mid-1980s the method of variance/covariance discounting (Quintana and West 1987, 88) has been used as a component of applied Bayesian forecasting models in financial econometric settings (Quintana 1992; Putnam and Quintana 1994, 1995; Quintana and Putnam 1996; Quintana, Chopra and Putnam 1995). In more recent years major developments in structured stochastic volatility (SV) modelling have led to the introduction of various approaches to modelling dependencies in volatility processes that, in principle, may lead to improvements in short-term forecasting of multiple financial and econometric time series. The potential is there for real improvements in practical short-term forecasting relative to the purely "reactive" variance discounting methods that simply allow for volatility changes but do not anticipate the forms of change. Our interest here is to explore dynamic factor models as a context for multivariate stochastic volatility modelling, motivated by:

- questions about the potential for dynamic factor models to provide practical improvements in short-term forecasting, and resulting dynamic portfolio allocations, of international exchange rates and other financial time series;
- issues of model structuring, implementation, Bayesian analysis and computation; and
- questions of comparison with the simpler methods of "tracking" volatility changes based on variance/covariance discounting.

We study these issues in connection with data analysis and portfolio construction using multiple series of international exchange rates.

Variants of the basic method of variance matrix discounting (see above references to Quintana and coauthors) have formal theoretical bases in matrix-variate "random walks" (Uhlig 1994, 97). Further discussion is given below, and more background can be found in West and Harrison (1997, section 16.4.5). The basic discounting methods follow foundational developments for univariate series in Ameen and Harrison (1995) and Harrison and West (1987), and the formal multivariate models are direct generalisations of univariate models of Shephard (1994a). Related discussion appears in West and Harrison (1997, section 10.8.2). In the general multivariate context, the approach leads to the embedding of exponentially smoothed estimates of "local" variance/covariance structure within a Bayesian modelling framework, and so provides for adaptation to stochastic changes as time series data are processed. Modifications to allow for changes in discount rates in order to adapt to varying degrees of change, including marked/abrupt changes in volatility patterns, extend the basic approach. The resulting update equations for sequences of estimated volatility matrices have univariate components that relate closely to variants of ARCH and SV models, and so it is not surprising that they have proven useful in many applications. However, unlike these more formal models, discounting methods do not have real predictive capabilities, simply allowing for and estimating changes rather than anticipating them. Hence the interest in factor models that set out to explicitly describe changes through patterns of time-variation in parameters driving underlying latent processes. This is the key motivating concept underlying interest in SV models generally, and led has to various authors mentioning or developing multivariate SV models with dynamic factor structure. Key references include the initiating work of Harvey, Ruiz and Shephard (1994), and the later developments in the papers of Jacquier, Polson and Rossi (1994, 95), and Kim, Shephard and Chib (1998).

In the following section we detail the basic framework and notation for stochastically time-varying variance matrices, and this leads into the basic factor structure as discussed by the above authors. This is followed by discussion of factor structure where we build on prior work in Bayesian factor analysis *per se*, adopting the structuring used by Geweke and Zhou (1996) in particular, and various extensions of basic factor models. We then discuss model fitting and computation, followed by analyses of international exchange rate time series with comparisons of the dynamic factor models with variance/covariance discounting. The paper concludes with some summary and concluding comments.

2 Time Varying Variance Matrices and Factor Structure

2.1 Bayesian Discount Estimation

To introduce notation, consider a $q-$variate time series $\mathbf{y}_t$, $(t = 1, 2, \ldots,)$ as conditionally independent, Gaussian random vectors with variance matrices $\mathbf{\Sigma}_t$, denoted by $N(\mathbf{y}_t|\mathbf{0}, \mathbf{\Sigma}_t)$ for each t. The variance matrix is stochastically time-varying. Bayesian discounting methods arise from a matrix-variate random walk for the $\mathbf{\Sigma}_t$ process, resulting in simple sequential updating of inverse Wishart posteriors for inference on $\mathbf{\Sigma}_t$ as time evolves. Using the notation of West and Harrison (1997, chapter 16), the time t posterior is of the form $p(\mathbf{\Sigma}_t|D_t) = W_{n_t}^{-1}(\mathbf{\Sigma}_t|\mathbf{S}_t)$ where $D_t = \{D_0, \mathbf{y}_1, \ldots, \mathbf{y}_t\} = \{D_{t-1}, \mathbf{y}_t\}$ is the sequentially updated information set at time t. Here n_t is the degrees of freedom and $\mathbf{S}_t$ a posterior estimate of $\mathbf{\Sigma}_t$, the posterior harmonic mean. The notation $W_r^{-1}(\cdot|\mathbf{S})$ indicates the inverse Wishart distribution with r degrees of freedom and scale matrix $\mathbf{S}$ (see West and Harrison, as referenced). The sequence of estimates $\mathbf{S}_t$ is trivially updated sequentially in time by the forward exponential moving average formula

$$\mathbf{S}_t = (1 - a_t)\mathbf{S}_{t-1} + a_t \mathbf{y}_t \mathbf{y}_t' \tag{1}$$

with weight $a_t = 1/(1 + \delta n_{t-1})$ based on a discount factor δ. This discount factor lies in $(0, 1)$, is typically between 0.9 and 1 and will be very close to unity for data at high sampling rates. Having analysed a fixed stretch of data $t = 1, \ldots, n$, the sequence of estimates $\mathbf{S}_t$ is revised by the related backward smoothing formula to incorporate the data at times $t + 1, \ldots, n$ in inference on $\mathbf{\Sigma}_t$. Denoting the revised estimate of $\mathbf{\Sigma}_t$ by $\mathbf{S}_{t,n}$, the formula is given in terms of inverse variance matrices by the backward recursion

$$\mathbf{S}_{t,n}^{-1} = (1 - \delta)\mathbf{S}_t^{-1} + \delta \mathbf{S}_{t+1,n}^{-1} \tag{2}$$

for each $t = n - 1, n - 2, \ldots, 1$, and starting with $\mathbf{S}_{n,n} = \mathbf{S}_n$. See West and Harrison (1997, pp608-609) for further details, and the various references by Quintana and coauthors listed above for development and application in econometric finance.

In connection with our analyses of exchange rate time series below, extensions to models in which the time series has a non-zero mean, possibly modelled via a dynamic regression model, is straightforward. Then the forward updating formula (1) is modified by replacing the observation $\mathbf{y}_t$ by the appropriate standardised forecast error, following West and

Harrison (as referenced above). The details are standard and a side issue here, though the modification is important in developing portfolio allocations below. Though there are ranges of possible models for a non-zero, stochastic mean function that may be of interest, our analysis is restricted to an assumedly constant mean $\boldsymbol{\theta}$ here. In modelling exchange rate returns, this will be a vector of very small elements and the impact of admitting a non-zero mean is small in resulting inferences on the $\boldsymbol{\Sigma}_t$ sequence. The impact on portfolio allocations that are mean-dependent are not necessarily small, however, so we maintain the more general model. That is, we assume the model

$$\mathbf{y}_t \sim N(\boldsymbol{\theta}, \boldsymbol{\Sigma}_t) \tag{3}$$

and that the resulting filtering and smoothing equations from Bayesian discounting analyses are modified following West and Harrison (1997, pp608-609). Future reports will discuss models with mean functions involving stochastic regressions, such as underpin the applied models of Quintana and Putnam (1996), Quintana, Chopra and Putnam (1995).

2.2 Component and Factor Structure

From very early examples and applications of variance matrix discounting, the principal component or factor structure of the variance matrices $\boldsymbol{\Sigma}_t$ has been the subject of analysis in aiding understanding of the nature of changes in covariance patterns, and the underlying latent mechanisms driving such changes. See, for example, the studies of monthly exchange rate time series in Quintana and West (1987), also reported in West and Harrison (1997, section 16.4.6), where the patterns of change over time in the principal component structure of the estimates $\mathbf{S}_{t,n}$ are explored. A standard principal component decomposition of $\mathbf{S}_{t,n}$ provides insight into the related decomposition of $\boldsymbol{\Sigma}_t$. Often, as in the above examples, this will yield a small number of dominant components representing latent factors contributing measurably to both total variability in the series and the covariance structure, together with additional residual components. This underlies the interest in dynamic factor models to more explicitly represent the latent structure and put the spot-light on inference on factor processes and their parameters directly. In line with Jacquier, Polson and Rossi (1994), Kim, Shephard and Chib (1998), and (extrapolating to the stochastic volatility context here) Geweke and Zhou (1996), a basic $k-$factor dynamic model (with $k < q$,) for $\boldsymbol{\Sigma}_t$ is

$$\boldsymbol{\Sigma}_t = \mathbf{X}_t\mathbf{H}_t\mathbf{X}_t' + \boldsymbol{\Psi}_t = \sum_{j=1}^{k} \mathbf{x}_{tj}\mathbf{x}_{tj}' h_{tj} + \boldsymbol{\Psi}_t \tag{4}$$

where

- $\mathbf{X}_t$ is the $q \times k$ factor loadings matrix at time t, with columns $\mathbf{x}_{tj}$,
- $\mathbf{H}_t = \text{diag}(h_{t1}, \ldots, h_{tk})$ is the diagonal matrix of instantaneous factor variances, and
- $\boldsymbol{\Psi}_t = \text{diag}(\psi_{t1}, \ldots, \psi_{tk})$ is the diagonal matrix of instantaneous, series-specific or "idiosyncratic" variances.

In terms of the time series $\mathbf{y}_t$ in (3), this is equivalent to the representation

$$\mathbf{y}_t = \boldsymbol{\theta} + \mathbf{X}_t\mathbf{f}_t + \boldsymbol{\epsilon}_t \tag{5}$$

where

- $\mathbf{f}_t \sim N(\mathbf{f}_t|\mathbf{0}, \mathbf{H}_t)$ are conditionally independent realisations of the $k-$vector latent factor process,
- $\boldsymbol{\epsilon}_t \sim N(\boldsymbol{\epsilon}_t|\mathbf{0}, \boldsymbol{\Psi}_t)$ are conditionally independent and series-specific quantities, and
- $\boldsymbol{\epsilon}_t$ and $\mathbf{f}_s$ are mutually independent for all t, s.

Principal component decompositions of estimated variance matrix sequences provide insight into the relevant numbers of "important" factors k, the factor loadings and variance components, and their changes through time. Note that the general factor model above does not, however, necessarily involve the notation of orthogonal factor structure key to principal component methods.

2.3 Factor Model Constraints

From here on we restrict attention to models with constant factor loadings and idiosyncratic variances, so that $\mathbf{X}_t = \mathbf{X}$ and $\boldsymbol{\Psi}_t = \boldsymbol{\Psi}$ for all t. This provides a framework very similar to those mooted by the earlier authors, as referenced above, in which we aim to investigate the issues and difficulties in model implementation. We have, however, implemented models that do allow for changes in time, particularly in the factor loadings, and this will be necessary in some application; more on that in a future article.

Well-known questions of model identification and parametrisation arise immediately. First, concerning the number of parameters in the factor model, and second concerning the invariance under invertible linear transformations of the factor vectors. There are many approaches to constraining the model for identification, each raising its own questions of interpretation of the resulting factor structure (Press 1985, chapter 10; Press and Shigemasu 1989). Our preference here is to adopt a variant of the "hierarchical" constraints used, for example, in Geweke and Zhou (1996). Specifically, the loadings matrix is given in the form

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ x_{2,1} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ x_{k,1} & x_{k,2} & x_{k,3} & \cdots & 1 \\ x_{k+1,1} & x_{k+1,2} & x_{k+1,3} & \cdots & x_{k+1,k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{q,1} & x_{q,2} & x_{q,3} & \cdots & x_{q,k} \end{pmatrix}. \tag{6}$$

One immediate implication is that the chosen order of the univariate time series in the $\mathbf{y}_t$ vector is viewed as defining the factors: the first series is the first factor plus a "noise" term, and so forth. This focuses attention on the choice of ordering in model specification, and provides interpretation. A second implication relates to the constraint on the number of factors allowed. If all the idiosyncratic variances ψ_j are non-zero, then, simply by counting the number of free parameters in the implied variance matrix $\boldsymbol{\Sigma}_t$, we deduce an upper bound on k implied by solutions to the quadratic inequality $q(q+1) - 2(qk+q) - k(k+1) \geq 0$.

For example, with $q = 6$ or 7 we have $k \leq 3$, with $q = 15$ or 16 we have $k \leq 10$, while with $q = 30$ we have $k \leq 22$. For realistic values of q this bound is unlikely to be problematic, as practical interest will be in models with smaller numbers of factors.

2.4 Stochastic Volatility for Dynamic Factors

Multivariate generalisations of univariate SV models through dynamic factor models are mentioned by various authors, including Harvey, Ruiz and Shephard (1994), Shephard (1996), Kim, Shephard and Chib (1998), and have been investigated by Jacquier, Polson and Rossi (1995). The basic model of the latter authors assumes that the univariate factor series f_{ti} follow standard univariate SV models, but discuss possible extensions also mentioned in Kim, Shephard and Chib (1998). We adopt such an extension here, one in which the log volatilities of the factors follows a vector autoregression with possibly correlated innovations.

Before proceeding, we note that related developments of multivariate models appear in the multivariate ARCH and related areas, in studies of both SV issues and common components models. Key contributions include Engle (1982), Bollerslev, Engle and Wooldridge (1988), Harvey and Stock (1988), Diebold and Nerlove (1989), Engle, Ng, and Rothschild (1990), Bollerslev and Engle (1993), Müller and Pole (1994), King, Sentana and Wadhwani (1994). Much of what is developed below, in terms of model analysis and inference, would easily transfer to alternative and related models in the areas represented in these various references.

For each $i = 1, \ldots, k$ define $\lambda_{ti} = \log(h_{ti})$, and write $\boldsymbol{\lambda}_t = (\lambda_{t1}, \ldots, \lambda_{tk})$. We assume a stationary vector autoregression of order one, VAR(1), centered around a mean $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)'$ and with individual AR parameters ϕ_i in the matrix $\boldsymbol{\Phi} = \text{diag}(\phi_1, \ldots, \phi_k)$. That is, for $t = 1, 2 \ldots,$ we have

$$\boldsymbol{\lambda}_t = \boldsymbol{\mu} + \Phi(\boldsymbol{\lambda}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\omega}_t \tag{7}$$

with independent innovations

$$\boldsymbol{\omega}_t \sim N(\boldsymbol{\omega}_t | \mathbf{0}, \mathbf{U}) \tag{8}$$

for some innovations variance matrix $\mathbf{U}$. The implied marginal distribution for each $\boldsymbol{\lambda}_t$ is then

$$\boldsymbol{\lambda}_t \sim N(\boldsymbol{\lambda}_t | \boldsymbol{\mu}, \mathbf{W}) \tag{9}$$

where $\mathbf{W}$ satisfies $\mathbf{W} = \boldsymbol{\Phi}\mathbf{W}\boldsymbol{\Phi} + \mathbf{U}$. Note that this marginal variance matrix has elements $W_{ij} = U_{ij}/((1-\phi_i)(1-\phi_j))$. The model allows dependencies across volatility series through non-zero off-diagonal entries in $\mathbf{U}$ and $\mathbf{W}$. Also, this marginal distribution defines the initial distribution for $\boldsymbol{\lambda}_1$.

3 Bayesian Inference and Computation

The model as specified so far comprises the basic factor structure (5) with supporting assumptions of conditional normality and independence, combined with the SV model (7) and its supporting assumptions. Model completion for Bayesian analysis requires prior distributions for the full set of parameters $\{\boldsymbol{\theta}, \mathbf{X}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Phi}, \mathbf{W}\}$. Bayesian inference for any specified prior requires the computation and summarisation of the implied posteriors for these model parameters, together with inferences on the factor processes $\{\mathbf{f}_t, t = 1, 2, \dots, n\}$ and the log-volatility sequence $\{\boldsymbol{\lambda}_t, t = 1, 2, \dots, n\}$ over the time window of n consecutive observations comprising the information set D_n. Computation via MCMC simulation builds on both the work of previous authors in the SV and factor modelling literature, and previous work in quite different models with related technical structure (Aguilar and West 1998; West and Aguilar 1997).

To complete the model specification, we assume a prior specified in terms of conditionally independent components

$$p(\boldsymbol{\theta})p(\mathbf{X})p(\boldsymbol{\Psi})p(\boldsymbol{\mu})p(\boldsymbol{\Phi})p(\mathbf{U}) \tag{10}$$

where the chosen marginal priors are either standard reference priors or proper priors that are chosen to be conditionally conjugate, as discussed below. The outlook here is to explore the use of reference priors to the extent possible to provide an initial analysis framework. Our prior specifications reflect this view, though these models do require the use of informative, proper priors for some model components due to identification issues, as we will discuss. Further, specific applications may use alternative prior specifications, both in terms of informative priors on model components and in terms of prior dependencies between parameters, though we do not discuss other prior specifications here. First, we assume standard reference priors for the univariate entries in the conditional mean, the factor loading matrix and the idiosyncratic variance matrix, so that

$$p(\boldsymbol{\theta})p(\mathbf{X})p(\boldsymbol{\Psi}) \propto \prod_{j=1}^{q} \psi_j^{-1}.$$

Note that the prior for $\mathbf{X}$ is, of course, subject to the specified 0/1 constraints on values in the the upper triangle and diagonal in (6), so the constant prior density applies only to the remaining, uncertain elements. Second, we use independent normal priors for the univariate elements of $\boldsymbol{\mu}$ and the diagonal elements of $\boldsymbol{\Phi}$. This allows for both reference priors, by setting the prior precisions to zero, and restriction of the values of each ϕ_j by adapting the prior to be truncated to $(0, 1)$. Finally, we use an informative inverse Wishart prior for the VAR(1) innovations variance matrix $\mathbf{U}$. This will often be specified with hyperparameters based on prior data analysis, as we illustrate below. Notice that an improper reference prior on $\mathbf{U}$, together with that so specified for $\boldsymbol{\Psi}$, is simply inappropriate, as the two determine separate sources of variability in the data that are confounded in the model. This point, rather critical to model implementation and resulting data analysis, is almost implicit in the prior work of Kim, Shephard and Chib (1998). These authors use informative proper priors for innovations variances that parallel our assumptions in their univariate SV models; though they present these priors without further discussion, the propriety is critical in overcoming otherwise potentially problematic confounding issues. Hence initial analysis of previous data, or some other prior elicitation activity, is needed. In our applied development below we explore the use of Bayesian variance discounting analyses in providing easy preliminary analysis of a reserved initial section of data as input to this. For now, the key point is that the prior for $\mathbf{U}$ is both proper and has the conditionally conjugate inverse Wishart form.

Iterative posterior simulation uses an MCMC strategy that extends those in existing SV models (Jacquier, Polson and Rossi 1995; Kim, Shephard and Chib 1998) to the multivariate case, introduces elements of MCMC algorithms for Bayesian factor analysis as in Geweke and Zhou (1996), and adds novel components derived from models with latent VAR components developed in a quite different context (Aguilar and West 1998; West and Aguilar 1997). We iteratively simulate values of all model parameters together with the full set of values of the latent processes $\mathbf{f}_t$ and $\boldsymbol{\lambda}_t$ by sequencing through the a set of conditional distributions detailed in the Appendix. At some stages we have direct conditional simulations, at others we require the introduction of novel Metropolis-Hastings accept/reject steps. We note in passing that, from an algorithmic viewpoint, there are various possible extensions and alternative methods for components of the MCMC analysis, such as in utilising some of the ideas from Shephard and Pitt (1997) for example, though we have not explored such variants yet. Beyond the appendix material here, further technical details are available on request from the authors, as is Fortran software for this implementation.

4 Studies of International Exchange Rates

4.1 Data and Initial Discounting Analyses

Figure 1 displays time series graphs of the weekday closing spot exchange rates of several currencies relative to the US dollar during the period 10/09/86 to 08/09/96, a total of 2567 data points in each series. The currencies are, in order, the Deutschmark/Mark (DEM), Japanese Yen (JPY), Canadian Dollar (CAD), French Franc (FRF), British Pound (GBP) and Spanish Peseta (ESP). We analyse one-day-ahead returns $y_{ti} = s_{ti}/s_{t-1,i} - 1$ for currency $i = 1, \ldots, q = 6$, as graphed in Figure 2. Initial analyses using variance matrix discounting are summarised in Figures 3 and 4. Three separate analyses were run, differing only through the value of the discount factor, specified as $\delta = 0.9, 0.95, 0.99$ for the three cases. The initial prior distribution in each case is very vague, namely $W_1^{-1}(\mathbf{\Sigma}|\mathbf{I})$. In each of the three analyses, principal components decompositions were made of each of the posterior estimates $\mathbf{S}_{t,n}$ over $t = 1, \ldots, n = 2567$. In each analysis and essentially uniformly over the time period, this yields three dominant components with fairly stable time trajectories for the corresponding eigenvectors representing the dynamic factor loadings. Figure 3 displays the time trajectories of the diagonal elements of $\mathbf{S}_{t,n}$, i.e., the sequence of posterior point estimates of the conditional variances of the six currencies. Figure 4 displays the corresponding related trajectories of the estimates of the variances of the underlying latent factors, namely the eigenvalues of the $\mathbf{S}_{t,n}$ matrices over time.

The greater adaptivity induced by lower discount factors is apparent in these graphs. The very low $\delta = 0.9$ is over-adaptive, responding very markedly to small changes in realised volatilities. In contrast, the higher discount factor $\delta = 0.99$ induces a much greater degree of smoothing of the volatility trajectories, and is likely under-adaptive in time of really marked change, such as towards the end of 1992 when Britain withdrew from the EU exchange rate agreement, resulting in marked swings and increased volatility in the European currencies across the board. The impact of this event is evident in the estimated trajectories of both the marginal variances of currencies and in the corresponding variances of the factors arising from the direct principal components decompositions in Figure 4. Notice that the end-1992 volatility changes impact across all factors, highlighting the apparent dependencies in factor trajectories across the entire time period. This indicates the need for dependence structure in modelling latent volatility processes in dynamic factor analyses, as is allowed in the theoretical framework described above and investigated in factor model data analyses below. We view such dynamic principal component analyses as providing informal, exploratory views of possible latent factor structure, albeit conditioned on the mathematically convenient but practically questionable orthogonality constraints. It appears that at most three factors are necessary, which is anticipated as the currencies represent three distinct trading blocs: Canada, Japan and the EU. The trajectories of the three minor eigenvalues remain at consistently negligible levels across the time frame here, so that a model with three factors plus currency-specific random effects is indicated. We now adopt such a model.

4.2 Dynamic Factor Analysis

Taking $q = 6$ and $k = 3$ in the dynamic factor model (5) provides a maximal specification: under the assumed structure of the factor loadings matrix (6), and assuming each of the ψ_j to be non-zero, the number of factors must necessarily be no greater than three. Hence, in addition to being suggested by the discounting analyses, this serves here as an encompassing model; if fewer than three factors are supported by the data, that fact will be reflected in posterior inferences about factor loadings and variances.

We base an appropriate, informative prior for the key matrix $\mathbf{U}$ in the volatility model on a pre-analysis of the initial 200 observations on the time series, reserving these few observations for this alone and then analysing the remaining data with the factor model. From the Bayesian discount model with a discount factor of 0.9, we extracted the point estimates of the three dominant eigenvalues of each $\mathbf{S}_{t,n}$ and used these as ad-hoc estimates of the factor volatilities h_{tj}, for each j, t. Three separate AR(1) models where then fitted to the log-volatilities so computed, using standard reference Bayesian analyses. This provided posteriors for the AR parameters and innovations variances, in each volatility series marginally, that we take as "ball-park" initial estimates to be used to specify an informative prior for $\mathbf{U}$ prior to analysis of the remaining data. This preliminary analysis gave approximate prior means of the three innovations variances around 0.001–0.002. With this in mind, we chose the prior for $\mathbf{U}$ in the factor model analysis to be inverse Wishart $W_{r_0}^{-1}(\mathbf{U}|\mathbf{R}_0)$ with $r_0 = 100$ degrees of freedom (half the prior sample size in the ad-hoc analysis) and $\mathbf{R}_0 = 0.0015\mathbf{I}$, appropriately "centering" the prior for $\mathbf{U}$. Note that the prior does not anticipate correlations across volatility processes, though this could easily be done.

The MCMC analysis of this factor model involved a range of experiments with Monte Carlo sample sizes and starting values, and MCMC diagnostics. Our summary numerical and graphical inferences are based on over 20,000 simulations of posteriors, generated following a 5,000 burn-in period. We subsample a set of 1,000 spaced 20 apart so as to break correlations and record resulting samples for graphical display purposes. Summary graphs appear in Figures 5 to 9 inclusive. First, Figure 5 graphs estimated trajectories of conditional variances of the currencies – the posterior means of the diagonal elements of $\mathbf{\Sigma}_t = \mathbf{X}\mathbf{H}_t\mathbf{X} + \mathbf{\Psi}$. Note the similarity with the trajectories from the more adaptive of the discount analyses in Figure 3, as is to be expected. Next, Figure 7 provides histogram approximations to marginal posteriors for the elements of $\mathbf{X}$. Note that the first column gives positive weight to all but CAD, representing the relative strength of the US dollar to the currencies of the EU and Japan. The CAD has almost no weight here, as to be expected as its value in international markets is most strongly determined by the US dollar alone, and the relative values of the weights on the EU countries naturally reflect their relative strengths. The loadings on the second factor are very small and, if non-negligible, negative, apart from Japan with the fixed unit weight. This is therefore largely the Japan:US factor, with some residual contrast between Japan and the rest of the currencies. Similar comments applies to the third factor which essentially represents the Canadian:US rates, and in which the main residual contrast is that reflecting the differential status of Britain to the rest of the EU, presumably driven in part by the departure of Britain from the exchange rate control system. The graphs in Figure 6 display the trajectories of approximate posterior means for the three factor processes f_{tj} and their conditional standard deviations $\sqrt{h_{tj}}$. The main points to note here are the appearance of peaks in the volatility processes consistent with positive correlations in volatility across the three factors, and the relative scales of volatility: all three factors are evidently contributing measurably to overall variability in the multiple series, though the factors appear to be roughly ordered in terms of decreasing overall levels. Again, this is consistent with expectations from a substantive viewpoint.

Figures 8 and 9 summarise marginal posterior inferences for key fixed model parameters, all in boxplot form. Figures 8 displays boxplots of posterior margins as follows. the upper left frame displays margins for the elements of the conditional mean $\boldsymbol{\theta}$. The upper right frame displays margins for the diagonal elements $\sqrt{\psi_j}$ of the residual variance matrix $\boldsymbol{\Psi}$. The lower frame displays margins for $100\psi_j/\sigma_{tj}^2$ where σ_{tj}^2 is the j^{th} diagonal element of $\boldsymbol{\Sigma}_t$. These ratios measure percent total variation in each of the currency series that is contributed by the idiosyncratic terms – generally non-negligible, and appreciable for both GBP and ESP.

Similar displays appear in Figure 9 for elements of the parameters $\boldsymbol{\mu}, \boldsymbol{\Phi}$ and $\mathbf{U}$ in the VAR(1) volatility model. Specifically, the upper left frame displays margins for the three parameters $\exp(\mu_j/2)$ where the μ_j are the entries of the stationary mean $\boldsymbol{\mu}$ of the VAR(1) model. Converting from the log-volatility to volatility scales, these scale factors $\exp(\mu_j/2)$ represent the standard deviations of the implied stationary distribution, i.e., base levels of conditional variation in the three factor processes. The rough ordering of factors according to marginal variability is clear here. The upper right frame displays margins for the AR parameters ϕ_j in $\boldsymbol{\phi}$, indicating that all three are obviously very close to, but less than, unity, and so representing high persistence in the volatility processes. The approximate posterior means for the three ϕ_j are 0.97, 0.98 and 0.98, respectively. The lower frame displays margins for the standard deviations and correlations of matrix $\mathbf{W}$, the marginal variance matrix in the VAR(1) volatility model. The earlier noted positive correlations between factor processes are indicated here. Related numerical summaries provide approximate posterior means of the variances in $\mathbf{W}$ as 0.50, 0.86 and 0.82, respectively, while the corresponding posterior means for the variances in the innovations matrix $\mathbf{U}$ are 0.027, 0.034 and 0.025, respectively.

4.3 Comparison of Models with Constrained Portfolio Allocations

Model comparisons are made with explicit focus on one-step forecast accuracy in the context of dynamic portfolio allocations, essentially following the perspective of Quintana (1992), Putnam and Quintana (1994), and Quintana and Putnam (1996). A similar perspective is adopted in Polson and Tew (1997) though with very different models. Our comparisons are based on posterior distributions from the models fitted to the entire time series, so that they do not represent real-time, sequential forecasts, but nevertheless do provide a coherent basis for model comparisons with a utility function directly measuring real-world performance in terms of cumulative financial return. In this section, we adopt traditional portfolios in which the total sum invested at each time point is fixed, focusing mainly on questions about how the discount and factor models differ in terms of resulting cumulative returns.

At each time point $t-1$, we suppose that an existing investment in the various currencies under study may be reallocated according to a portfolio $\mathbf{a}_t$ for the next time point. The elements of $\mathbf{a}_t$ are the $US amounts invested in the corresponding currency. For this comparative analysis, we assume no transaction costs and that we may freely reallocate dollars instantaneously to long or short positions across the currencies, subject only $\mathbf{a}_t'\mathbf{1} = 1$. The realised portfolio return at time t is the $US amount $r_t = \mathbf{a}_t'\mathbf{y}_t$, and models may be compared on the basis of cumulative returns over chosen time intervals. The portfolio allocation decision problem involves the general Markowitz mean-variance optimisation, and we apply this at each time point one-step ahead. In all models, the time t situation is summarised through posterior one-step ahead means and variance matrices for $\mathbf{y}_t$, denoted here by $\mathbf{g}_t$ and $\mathbf{G}_t$. The decision context targets minimisation of the one-step ahead variance of returns subject to specified target means. For a specified target mean return m, the one-step portfolio $\mathbf{a}_t$ will be chosen to minimise the one-step ahead variance of returns $\mathbf{a}_t'\mathbf{G}_t\mathbf{a}_t$ subject to constraints $\mathbf{a}_t'\mathbf{1} = 1$ and $\mathbf{a}_t'\mathbf{g}_t = m$. The well-known solution is

$$\mathbf{a}_t^{(m)} = \mathbf{G}_t^{-1}(a\mathbf{g}_t + b\mathbf{1})$$

where

$$a = \mathbf{1}'\mathbf{G}_t^{-1}\mathbf{e} \quad \text{and} \quad b = -\mathbf{g}_t'\mathbf{G}_t^{-1}\mathbf{e}$$

with

$$\mathbf{e} = (\mathbf{1}m - \mathbf{g}_t)/d \quad \text{and} \quad d = (\mathbf{1}'\mathbf{G}_t^{-1}\mathbf{1})(\mathbf{g}_t'\mathbf{G}_t^{-1}\mathbf{g}_t) - (\mathbf{1}'\mathbf{G}_t^{-1}\mathbf{g}_t)^2.$$

Compared to this are two other standard portfolio allocations: the target-independent allocation derived at the boundary of the mean-variance efficient frontier, namely

$$\mathbf{a}_t^{(me)} = (\mathbf{1}'\mathbf{G}_t^{-1}\mathbf{g}_t)^{-1}\mathbf{G}_t^{-1}\mathbf{g}_t,$$

and the strictly risk-averse minimum variance portfolio

$$\mathbf{a}_t^{(mv)} = (\mathbf{1}'\mathbf{G}_t^{-1}\mathbf{1})^{-1}\mathbf{G}_t^{-1}\mathbf{1}.$$

Figure 10 graphs the trajectories of cumulative returns over time based on these four models in the frames in the first two rows. These four frames correspond to four different fixed investment strategies: the two risk-averse strategies $\mathbf{a}_t^{(mv)}$ and $\mathbf{a}_t^{(me)}$ and then two strategies $\mathbf{a}_t^{(m)}$ with daily target returns $m = 0.00016$ and $m = 0.00028$ respectively. These appear in the order (top left) $\mathbf{a}_t^{(mv)}$, (top right) $\mathbf{a}_t^{(me)}$, (center left) $\mathbf{a}_t^{(0.00016)}$ and (center right) $\mathbf{a}_t^{(0.00028)}$. After a period in which the model behave very similarly in the determination of portfolio allocations, the changes in volatility in 1992 are more appropriately captured by the most adaptive discount model ($\delta = 0.9$) and the dynamic factor model. These two models proceed to generally dominate the others in terms of cumulative returns under the practically relevant strategies. The factor model clearly wins across all cases in this time window. The trajectories of optimal portfolio weights for the target-independent case appear in Figure 11. This displays the changing values of $\mathbf{a}_t^{(me)}$ from the three Bayesian discount analyses and also from the dynamic factor model analysis. The weights are graphed as percentages, i.e., simply 100 times their actual values, as they are constrained to sum to unity. It is interesting to note that, though the very adaptive discount model produces cumulative returns that closely shadow the factor model, the weights in the factor model are relatively much more stable over time. The discount model adapts the weights quite widely as it permits very marked patterns of change in the full variance matrix of returns, whereas the factor model assigns changes in observed volatilities to appropriate model components and so induces more stability in weight trajectories.

4.4 Investment Performance with Unconstrained Portfolios

We now move to a further portfolio allocation study that is focused on comparison of portfolio strategies rather than on comparison of models. Here we consider allocations in which the portfolios are completely unconstrained; that is, we remove the unit sum constraint on the allocation vector $\mathbf{a}_t$. This means that we may choose each allocation without regard to resources, permitting arbitrary long or short positions across the currencies. This typifies the practical working context in global investments in large financial institutions, and is in line with recent work with discount models (Quintana and Putnam 1996).

Under an unconstrained strategy, the optimum allocation at time t is given by

$$\mathbf{a}_t^{(*m)} = a\mathbf{G}_t^{-1}\mathbf{g}_t$$

where

$$a = m/\mathbf{g}_t'\mathbf{G}_t^{-1}\mathbf{g}_t.$$

The third row of graphs in Figure 10 displays the cumulative returns trajectories using the optimal portfolio weights $\mathbf{a}_t^{(*m)}$ from this analysis, to be compared to the earlier figures using the constrained portfolios. The two graphs provide displays under portfolios (lower left) $\mathbf{a}_t^{(*0.00016)}$ and (lower right) $\mathbf{a}_t^{(*0.00028)}$. Now we see that the dynamic factor model is very clearly dominant, achieving cumulative returns that are about twice as large as those of the most competitive discount model, and, by the way, also exceeding those of the constrained models. The response following the major structural changes in volatility at the end of 1992 leads to a marked swing in portfolio structure that the unconstrained allocations capitalise on in a major way in the factor model, with a persistent effect on cumulative returns thereafter.

The trajectories of optimal, unconstrained portfolio weights for the case $m = 0.00016$ are graphed in Figure 13. The values plotted are 100 times the actual weights divided by the total $\mathbf{1}'\mathbf{a}_t^{(*m)}$ at each time point, indicating the relative weight of each currency in the portfolio at each time. This provides a direct comparison with the corresponding trajectories of weights from the unit-sum constrained allocation appearing in Figure 12, where the actual and relative values coincide. The impact of the Britain's withdrawal from the EU exchange rate agreement, in late 1992, and the resulting marked increases in volatility and the portfolio's response, are very clear in Figures 13 and 10. The allocations shift swiftly and quite radically in extent to short positions on Sterling and the strongly associate Peseta, while simultaneously adopting radically long positions on the strong Mark and Yen. At the same time, the total investment dropped markedly; the major changes in volatility led to the anticipation of high levels of increased risk, and the total $\mathbf{1}'\mathbf{a}_t^{(*m)}$ invested in the market decreased radically as a result. Figure 14 graphs the time trajectory of the total $\mathbf{1}'\mathbf{a}_t^{(*m)}$, indicating relative stability in the fluctuations around a level of unity, but with marked swings up and down in periods of low and high volatility, respectively.

5 Concluding Comments

Our investigations indicate the feasibility of formal Bayesian analysis of structured dynamic factor models. The analysis is accessible computationally with nowadays moderate computational resources, and our empirical studies suggest that the analysis will be manageable with 20-30 dimensional time series and several factors. We are currently investigating more extensive applications in short-term forecasting and on-line portfolio allocations with higher dimensional models for longer-term exchange rate futures. The example here is suggestive of potential benefits, and supportive of the view that exploiting systematic volatility patterns via factor structuring may yield very substantial improvements in short-term forecasting and decision making in dynamic portfolio allocation, especially in the unconstrained optimisation as illustrated in the final row of graphs in Figure 10. In the case of constrained portfolio optimisations, the over-adaptive discounting method does reasonably well at times, though is clearly eventually dominated in terms of cumulative return trajectories by the factor model. We conjecture that, in studies of forecasting and portfolio allocation with longer term horizons, such as 30-day exchange rate futures, and in extended models that incorporate dynamic regression components, the dominance of the factor modelling approach will be even clearer. This is the subject of current and near-future research.

The dynamic factor models illustrated are amenable to direct implementation using our customised MCMC methods with the minimal/reference prior specifications we have used here. Our use of the variance discounting method on a reserved initial section of the data to provide input to informative priors is important in identifying "ball-park" scales for the $\mathbf{U}$ matrix of the VAR(1) SV model. Though not pursued here, other aspects of such preliminary analyses may be used to determine informative priors for other elements of the factor model. The established discounting methods are, relative to dynamic factor models, trivial to implement in the current context, a fact that is important in using discount methods to specify partial prior structure in the dynamic models. Our empirical findings indicate that, not surprisingly with this kind of data, moderately adaptive discount methods fare well in time of slow change in volatility levels and patterns, but are relatively uncompetitive in cases of more marked structural change. This is to be expected. Looking ahead, models and approaches that attempt to simplify the process of factor modelling, perhaps somehow integrating elements and concepts of variance matrix discounting into a specified factor structure, may be attractive from a computational/implementation viewpoint. We are currently investigating such a synthesis of approaches. In the factor model context per se, there are several relevant technical and modelling issues to be explored. These include questions of choices about the numbers of factors, and about the ordering of time series in the context of the specific structure adopted for factor loadings. Model extensions under investigation relax the assumptions of constancy of the factor loadings, and we anticipate extensions to models in which the conditional mean is a dynamic regression, as is likely very necessary for serious practical application. In the meantime, further study and empirical assessments on time series with larger numbers of univariate components and larger numbers of factors are under investigation. Our experience to date leads us to believe that such investigations will be every fruitful and support the preliminary conclusions reached in this report about the potential utility of factor models.

Appendix

In each of the following subsections we detail the conditional posteriors for various parameters and latent variables in turn. At each step it is implicit that we are conditioning on fixed values (previously simulated values) of all other variables. As noted in the text, further technical details are available on request from the authors, as is Fortran software for this implementation.

Sampling the conditional mean

The basic model (3) and the uniform prior for $\boldsymbol{\theta}$ immediately imply a multivariate normal conditional posterior for $\boldsymbol{\theta}$ given the values of $\boldsymbol{\Sigma}_t$. This is easily sampled.

Sampling latent factors

The full conditional distribution of $\mathbf{f}_t$ is given by

$$N(\mathbf{f}_t|\mathbf{A}_t(\mathbf{y}_t-\boldsymbol{\theta}),\mathbf{H}_t-\mathbf{A}_t\mathbf{Q}_t\mathbf{A}_t')$$

where $\mathbf{Q}_t=\mathbf{X}\mathbf{H}_t\mathbf{X}'+\boldsymbol{\Psi}$ and $\mathbf{A}_t=\mathbf{H}_t\mathbf{X}'\mathbf{Q}_t^{-1}$. The $\mathbf{f}_t$ are conditionally independent and so sample values are drawn independently from this set of normal distributions for $t = 1,2\ldots,n$.

Sampling factor loadings

From the model, the conditional likelihood function for the factor loading matrix $\mathbf{X}$ is $\prod_{t=1}^{n} N(\mathbf{y}_t-\boldsymbol{\theta}|\mathbf{X}\mathbf{f}_t,\boldsymbol{\Psi})$. This is a standard form, log-quadratic in the uncertain elements of $\mathbf{X}$, and so combines with a normal or uniform reference prior to imply a multivariate normal conditional posterior, which is easily sampled.

Sampling idiosyncratic variances

The elements of $\mathbf{\Psi}$ are conditionally independent with inverse gamma distributions. Let $g_i = \sum_{t=1}^{n}(y_{ti} - \theta_i - \mathbf{z}_i'\mathbf{f}_t)^2$ where $\mathbf{z}_i$ is the i^{th} row of $\mathbf{X}$. Then ψ_i has the inverse gamma distribution with shape $n/2$ and scale $g_i/2$, denoted by $G(\psi_i^{-1}|n/2, g_i/2)$.

Sampling the mean log-volatilities

Under a normal prior $N(\boldsymbol{\mu}|\mathbf{m}_0, \mathbf{M}_0)$, the conditional posterior is normal $N(\boldsymbol{\mu}|\mathbf{m}, \mathbf{M})$ with

$$\mathbf{M}^{-1} = \mathbf{M}_0^{-1} + \mathbf{W}^{-1} + (n-1)(\mathbf{I} - \Phi)\mathbf{U}^{-1}(\mathbf{I} - \Phi)$$

and

$$\mathbf{M}\mathbf{m} = \mathbf{M}_0^{-1}\mathbf{m}_0 + \mathbf{W}^{-1}\boldsymbol{\lambda}_1 + (\mathbf{I} - \Phi)\mathbf{U}^{-1}\sum_{t=2}^{n}(\boldsymbol{\lambda}_t - \Phi\boldsymbol{\lambda}_{t-1}).$$

The case of a uniform reference prior is recovered by setting $\mathbf{M}_0^{-1} = \mathbf{0}$.

Sampling the VAR coefficients

The structure of the conditional posterior for $\mathbf{\Phi}$, and the resulting Metropolis-Hastings strategy for simulation, is precisely as developed for component VAR models in a quite different context in West and Aguilar (1998) and Aguilar and West (1997). Writing $\boldsymbol{\gamma}_t = \boldsymbol{\lambda}_t - \boldsymbol{\mu}$, note that the full conditional posterior density for $\mathbf{\Phi}$ is proportional to

$$p(\mathbf{\Phi})N(\boldsymbol{\gamma}_1|\mathbf{0}, \mathbf{W})\prod_{t=2}^{n}N(\boldsymbol{\gamma}_t|\mathbf{\Phi}\boldsymbol{\gamma}_{t-1}, \mathbf{U})$$

where $\mathbf{W} = \mathbf{\Phi}\mathbf{W}\mathbf{\Phi} + \mathbf{U}$ is easily evaluated as a function of $\mathbf{\Phi}$ and $\mathbf{U}$. Write $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_k)'$ for the diagonal of $\mathbf{\Phi}$, and $\mathbf{E} = \mathrm{diag}(\boldsymbol{\gamma}_{t-1})$. Then the conditional posterior may be written as proportional to

$$p(\mathbf{\Phi})c(\mathbf{\Phi})N(\boldsymbol{\phi}|\mathbf{b}, \mathbf{B})$$

where

$$\mathbf{B}^{-1} = \sum_{t=2}^{n}\mathbf{E}'\mathbf{U}^{-1}\mathbf{E} \quad \text{and} \quad \mathbf{B}\mathbf{b} = \sum_{t=2}^{n}\mathbf{E}'\mathbf{U}^{-1}\boldsymbol{\gamma}_t$$

and

$$c(\mathbf{\Phi}) = |\mathbf{W}|^{-I/2}\exp(-\mathrm{trace}(\mathbf{W}^{-1}\boldsymbol{\gamma}_1\boldsymbol{\gamma}_1')/2).$$

Under independent normal or uniform priors for the ϕ_j, the full conditional posterior distribution for $\mathbf{\Phi}$ is the multivariate normal $N(\boldsymbol{\phi}|\mathbf{b}, \mathbf{B})$ truncated to the $(0, 1)$ regions in each dimension, and then multiplied by the factor $c(\mathbf{\Phi})$. This may be sampled by several methods. We use a Metropolis Hastings algorithm that takes the truncated multivariate normal component as proposal distribution. That is, given a "current" value of $\boldsymbol{\phi}$, with corresponding matrices $\mathbf{\Phi}$ and $\mathbf{W}$, we sample a "candidate" vector $\boldsymbol{\phi}^*$ from this truncated normal, compute the corresponding diagonal matrix $\mathbf{\Phi}^*$ and variance matrix $\mathbf{W}^*$ such that $\mathbf{W}^* = \mathbf{\Phi}^*\mathbf{W}^*\mathbf{\Phi}^* + \mathbf{U}$, then accept this new $\boldsymbol{\phi}$ vector with probability

$$\min\{1, c(\mathbf{\Phi}^*)/c(\mathbf{\Phi})\}.$$

Sampling the VAR innovations variance matrix

Again, the structure of the conditional posterior for the innovations variance matrix $\mathbf{U}$ of the VAR(1) volatility model is closely related to developments in a quite different context in West and Aguilar (1998) and Aguilar and West (1997). Again using centred volatilities $\gamma_t = \boldsymbol{\lambda}_t - \boldsymbol{\mu}$ for each t, we have a full conditional posterior for $\mathbf{U}$ proportional to

$$p(\mathbf{U})a(\mathbf{U})|\mathbf{U}|^{-(n-1)/2}\exp(-\text{trace}(\mathbf{U}^{-1}\mathbf{G}))$$

where

$$\mathbf{G} = \sum_{t=2}^{n}(\boldsymbol{\gamma}_t - \boldsymbol{\Phi}\boldsymbol{\gamma}_{t-1})(\boldsymbol{\gamma}_t - \boldsymbol{\Phi}\boldsymbol{\gamma}_{t-1})'$$

and

$$a(\mathbf{U}) = |\mathbf{W}|^{-1/2}\exp(-\text{trace}(\mathbf{W}^{-1}\gamma_1\gamma_1')/2)$$

with $\mathbf{W} = \boldsymbol{\Phi}\mathbf{W}\boldsymbol{\Phi} + \mathbf{U}$. Under a specified inverse Wishart prior $W_{r_0}^{-1}(\mathbf{U}|\mathbf{R}_0)$, this posterior density is proportional to

$$a(\mathbf{U})W_r^{-1}(\mathbf{U}|\mathbf{R})$$

where $r = r_0 + n - 1$ and $r\mathbf{R} = r_0\mathbf{R}_0 + (n-1)\mathbf{G}$. We use this inverse Wishart distribution as a proposal distribution in the Metropolis-Hastings algorithm. That is, given a "current" value of $\mathbf{U}$ and corresponding $\mathbf{W}$, we sample a "candidate" value $\mathbf{U}^*$ from $W_r^{-1}(\mathbf{U}|\mathbf{R})$, and accept it with probability

$$\min\{1, a(\mathbf{U}^*)/a(\mathbf{U})\}$$

where $\mathbf{W}^* = \boldsymbol{\Phi}\mathbf{W}^*\boldsymbol{\Phi} + \mathbf{U}^*$.

Sampling the actual latent log-volatilities

Given currently imputed values for the full factor process $\mathbf{f}_t$ over time t, we follow Kim, Shephard and Chib (1998) in transforming each of the univariate time series of factor elements into a non-Gaussian linear model form. Specifically, for each factor i and time t define $z_{ti} = \log(f_{ti}^2)$ and note that

$$z_{ti} = \lambda_{ti} + \nu_{ti}$$

where the ν_{ti} terms are independent and distributed as $\log{-\chi_1^2}$. In vector form for all k factor series, we then have

$$\mathbf{z}_t = \boldsymbol{\lambda}_t + \boldsymbol{\nu}_t$$

where $\boldsymbol{\nu}_t = (\nu_{t1}, \ldots, \nu_{tk})'$. Based on the current values of the $\mathbf{z}_t$, for each t, this provides the set observation equations for the dynamic linear model with state equations

$$\boldsymbol{\lambda}_t = \boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{\lambda}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\omega}_t$$

for each t. This is a direct multivariate extension of the univariate approach in Kim, Shephard and Chib (1998), whose ensuing analysis uses the now established approximation to the distribution of the error terms ν_{it} as a specified finite mixture of normals (Shephard 1994b). The multivariate extension is immediate:

- Introduce a set of indicator variables s_{ti} such that that s_{ti} identifies the normal mixture component for ν_{ti}.
- Conditional on these indicators, we have a multivariate dynamic linear model for the sequence of log-volatility vectors. The forward-filtering, backward-sampling algorithm for state space models (Carter and Kohn 1994; Frühwirth-Schnatter 1994; West and Harrison 1997, chapter 15) now applies to directly simulate the full set of vectors $\{\boldsymbol{\lambda}_t, t = 1, \ldots, n\}$ from the implied conditional posterior. (Note that, in more elaborate models for the volatility processes, the alternative sampling method using the simulation smoother of de Jong and Shephard (1995) may have computational advantages not realised in this, the simplest VAR model.)
- Given these sampled values of the $\boldsymbol{\lambda}_t$, the complete conditional multinomial posterior probabilities over values of the indicators s_{it} are easily computed, the indicators being conditionally independent and so easily sampled.

References

Aguilar O. and West M. (1998), "Analysis of Hospital Quality Monitors using Hierarchical Time Series Models," In *Case Studies Bayesian Statistics in Science and Technology: Case Studies 4,* (eds: C. Gatsonis et al), Springer-Verlag, New York (forthcoming).

Ameen, J.R., and Harrison, P.J. (1995), "Normal Discount Bayesian Models," In *Bayesian Statistics 2*, (eds: J.M. Bernardo, M.H. De Groot, D.V. Lindley and A.F.M. Smith), North Holland, Amsterdam, and Valencia University Press.

Bollerslev T., Engle R.F. and Wooldridge J.M. (1988), "A Capital Asset Pricing Model With Time Varying Covariances," *Journal of Political Economy*, **96**, 116-131.

Carter, C., and Kohn, R. (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

Frühwirth-Schnatter, S. (1994), "Data Augmentation and Dynamic Linear Models," *Journal of Time Series Analysis,* **15**, 183–202.

de Jong, P., and Shephard, N. (1995), "The Simulation Smoother for Time Series Models," *Biometrika,* **82**, 339-50.

Diebold F.X. and Nerlove M. (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics*, **4**, 1-22.

Engle R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation," *Econometrica*, **50**, 987-1007.

Engle R.F., Ng V. K. and Rothschild M. (1990), "Asset Pricing with a Factor-ARCH Covariance Structure, Empirical Estimates for Treasury Bills," *Journal of Econometrics*, **45**, 213-237.

Geweke J. and Zhou G. (1996), "Measuring the Pricing Error of the Arbitrage Pricing Theory," *The Review of financial Studies*, Vol 9, No 2, 557-587.

Harrison, P.J., and West, M. (1987), "Practical Bayesian Forecasting," *The Statistician,* **36**, 115-125.

Harvey A.C. and Stock J.H. (1988), "Continuous Time Autoregressive Models with Common Stochastic Trends," *Journal of Economic Dynamics and Control*, **12**, 265-284.

Harvey A.C., Ruiz E. and Shephard N. (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies*, **61**, 247-264.

Harvey A. C. and Shephard N. (1996), "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns," *Journal of Business and Economic Statistics*, **14**, 429-434.

Jacquier E., Polson N.G. and Rossi P. (1994), "Bayesian Analysis of Stochastic Volatility Models (with Discussion)," *Journal of Business and Economic Statistics*, **12**, 371-388.

Jacquier E., Polson N.G. and Rossi P. (1995), "Models and Priors for Multivariate Stochastic Volatility," *Discussion Paper, Graduate School of Business, University of Chicago*

Kim S., Shephard N. and Chib S.(1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies,* forthcoming.

King M., Sentana E. and Wadhwani S. (1994), "Volatility and Links Between National Stock Markets," *Econometrica*, **62**, 901-933.

Müller P. and Pole A. (1994), "Monte Carlo Posterior Integration in GARCH Models," *Sankhya, (Ser B),* to appear.

Ng V.K., Engle R.F. and Rothschild M. (1992), "A Multi-Dynamic-Factor Model for Stock Returns," *Journal of Econometrics*, **52**, 245-266.

Polson, N.G. and Tew, B.V. (1997), "Bayesian Porfolio Selection: An Empirical Analysis of the SP500 Index 1970-1996," Technical Report, Graduate School of Business, University of CHicago.

Press, S.J. (1985), *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference,* California: Krieger.

Press S.J. and Shigemasu K. (1989), "Bayesian Inference in Factor Analysis," In contributions to Probability and Statistics: *Essays in Honor of Ingram Olkin.* eds L.J Gleser, M.D. Perlman, S.J. Press and A.R. Sampson, 271-287. New York: Springer Verlag.

Putnam, B. H. and Quintana, J. M. (1994), "New Bayesian Statistical Approaches to Estimating and Evaluating Models of Exchange Rates Determination," In *Proceedings of the ASA Section on Bayesian Statistical Science, 1994 Joint Statistical Meetings,* American Statistical Association

Putnam, B. H. and Quintana, J. M. (1994), "Shrink Wrap for Currencies," *Asset Allocation: Applying Quantitative Discipline to Asset Allocation,* (ed: Putnam, B.H.), London: Global Investor, Euromoney Publications, 175-181.

Putnam, B. H. and Quintana, J. M. (1995), "The Evolution of Bayesian Forecasting Models," *Asset Allocation: Applying Quantitative Discipline to Asset Allocation,* (ed: Putnam, B.H.), London: Global Investor, Euromoney Publications, 139-146.

Quintana, J.M. and West, M. (1987), "An Analysis of International Exchange Rates using Multivariate DLMS," *The Statistician,* **36**, 275-281.

Quintana J.M. and West, M. (1988), "Time Series Analysis of Compositional Data," In *Bayesian Statistics 3*, (eds: J.M. Bernardo, M.H. De Groot, D.V. Lindley and A.F.M. Smith), Oxford University Press.

Quintana, J.M. (1992), Optimal Portfolios of Forward Currency Contracts. In *Bayesian Statistics 4*, (eds: J.O. Berger, J.M. Bernardo, A.P. Dawid and A.F.M. Smith), Oxford University Press.

Quintana J.M., Chopra V.K. and Putnam B.H. (1995), "Global Asset Allocation: Stretching Returns by Shrinking Forecasts," In *Proceedings of the ASA Section on Bayesian Statistical Science, 1995, Joint Statistical Meetings*, American Statistical Association.

Quintana, J. M., and Putnam, B. H. (1996), "Debating Currency Markets Efficiency using Multiple-Factor Models," In *Proceedings of the ASA Section on Bayesian Statistical Science, 1996 Joint Statistical Meetings,* American Statistical Association

Shephard N. (1994a), "Local Scale Models: State Space Alternatives to Integrated GARCH Processes," *Journal of Econometrics,* **6**, 333-364.

Shephard N. (1994b), "Partial Non-Gaussian State Space," *Biometrika,* **81**, 115-132.

Shephard N. (1996), "Statistical Aspects of ARCH and Stochastic Volatility," In D.R. Cox, O.E. Barndorff-Nielson and D.V. Hinkley eds. *Time Series Models in Econometrics, finance and Other Fields*, pp 1-67. London: Chapman & Hall.

Shephard, N., and Pitt, M.K. (1997), "Likelihood analysis of non-Gaussian measurement time series," *Biometrika,* **84**, 653-67.

Uhlig, H. (1994), "On singular Wishart and singular multivariate beta distributions," *Ann. Statist.*, **22**, 395-405.

Uhlig, H. (1997), "Bayesian vector-autoregressions with stochastic volatility," *Econometrica*, **65**, 59-73.

West M. and Aguilar O. (1997), "Studies of Quality Monitor Time Series: The V.A. Hospital System". Report for the VA Management Science Group, Bedford, MA., Institute of Statistics and Decision Sciences, Discussion Paper 96-02.

West M., and Harrison, P.J. (1997), *Bayesian Forecasting and Dynamic Models*, (2nd Edn.), New York: Springer Verlag.

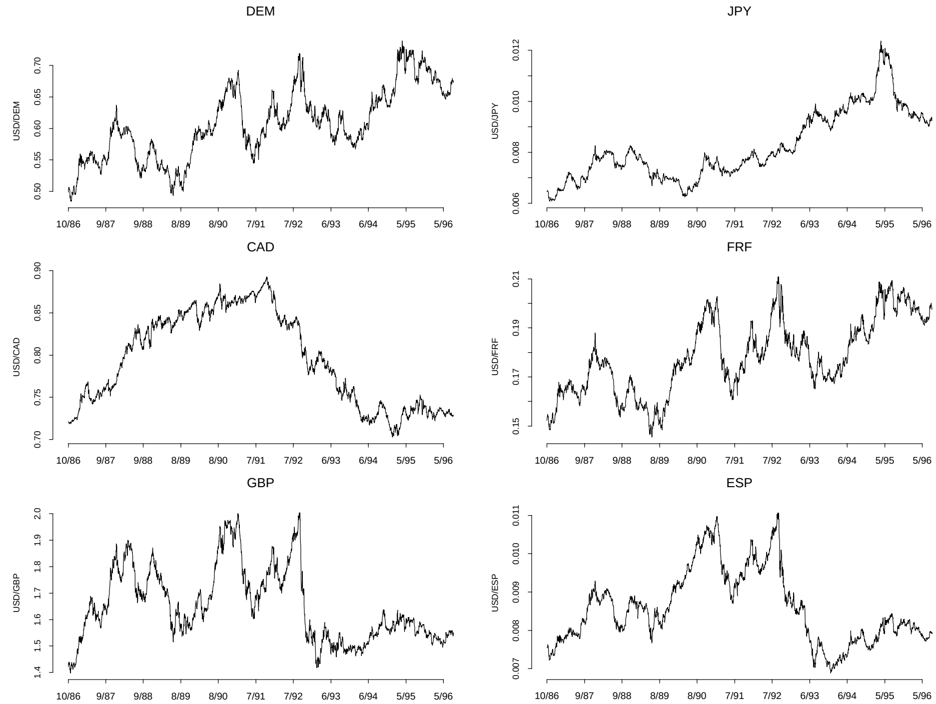

Figure 1: Exchange rate time series.

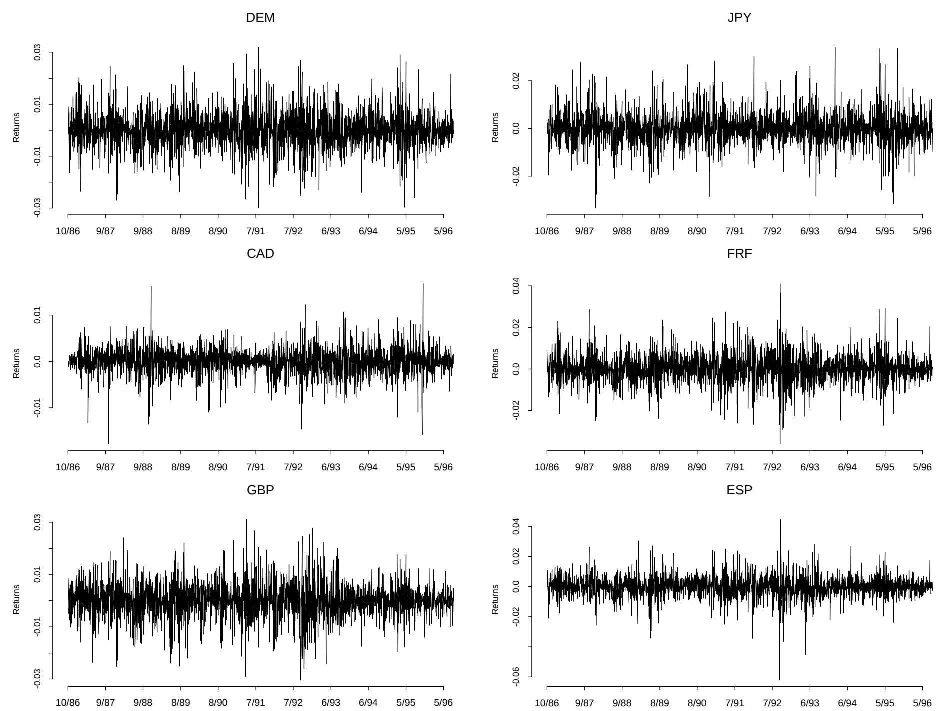

Figure 2: Exchange rate returns.

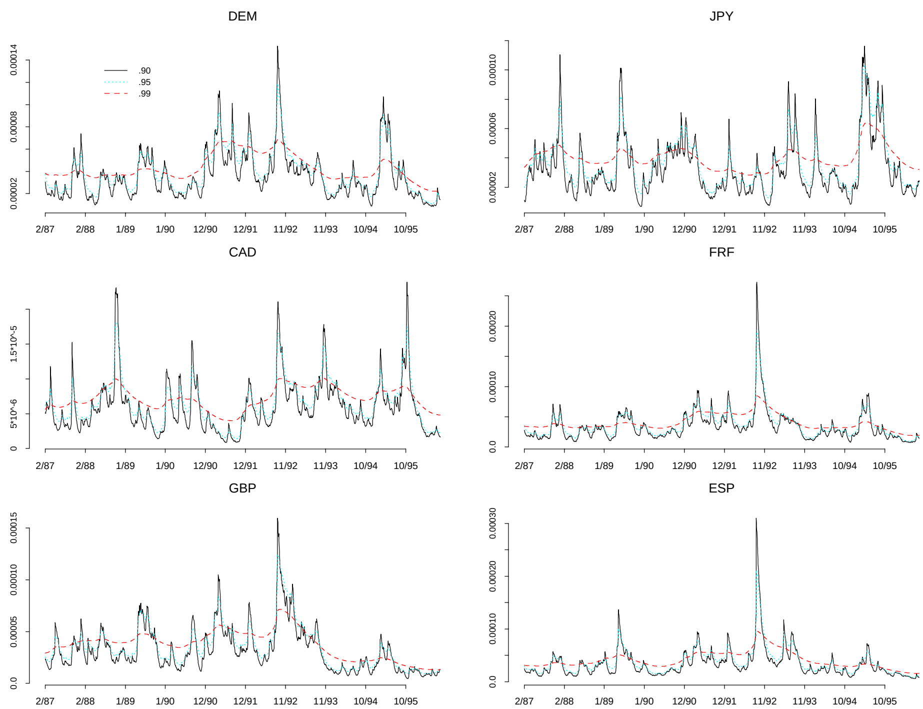

Figure 3: Conditional variances from discount analyses.

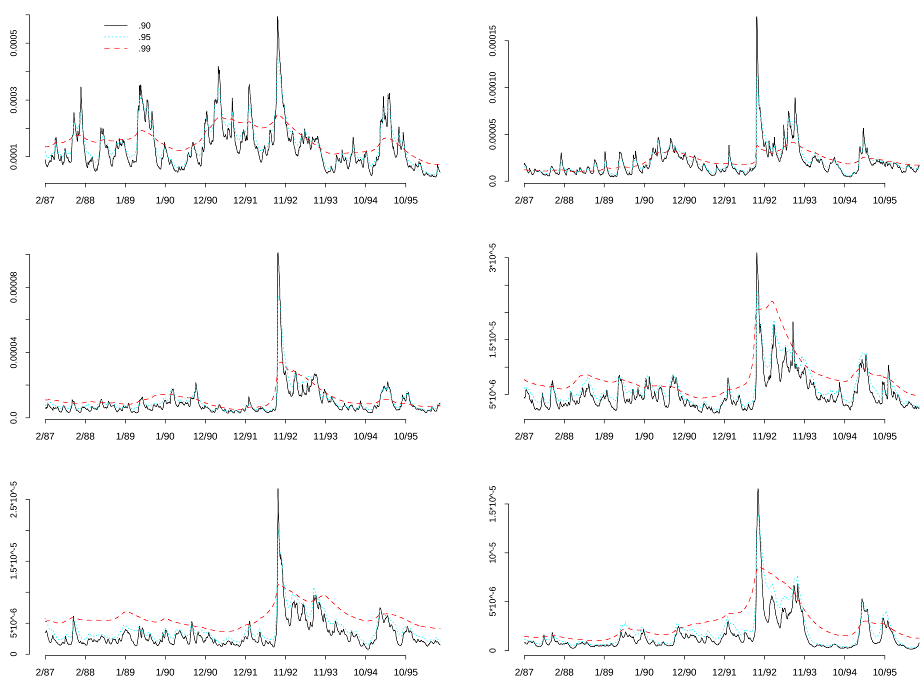

Figure 4: Factor variances from discount analyses.

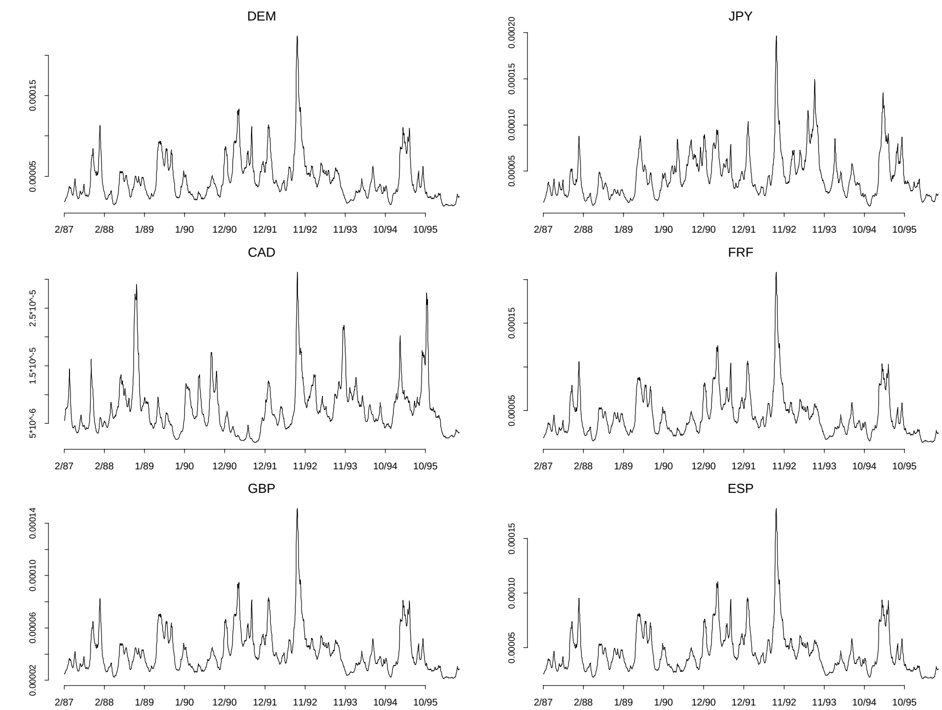

Figure 5: Conditional variances from dynamic factor model.

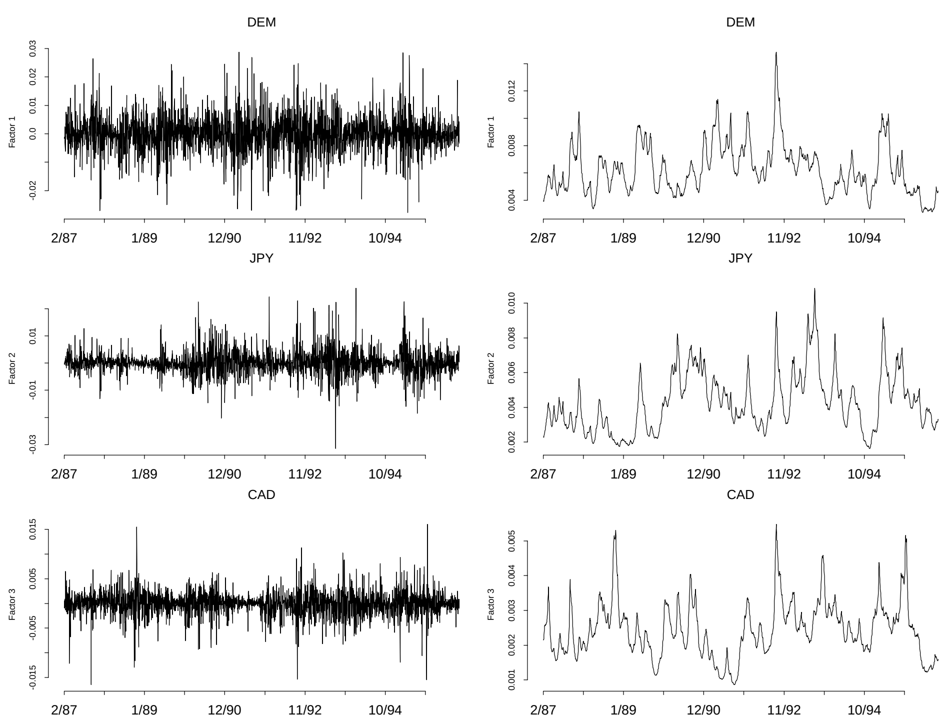

Figure 6: Dynamic factor processes and factor standard deviations.

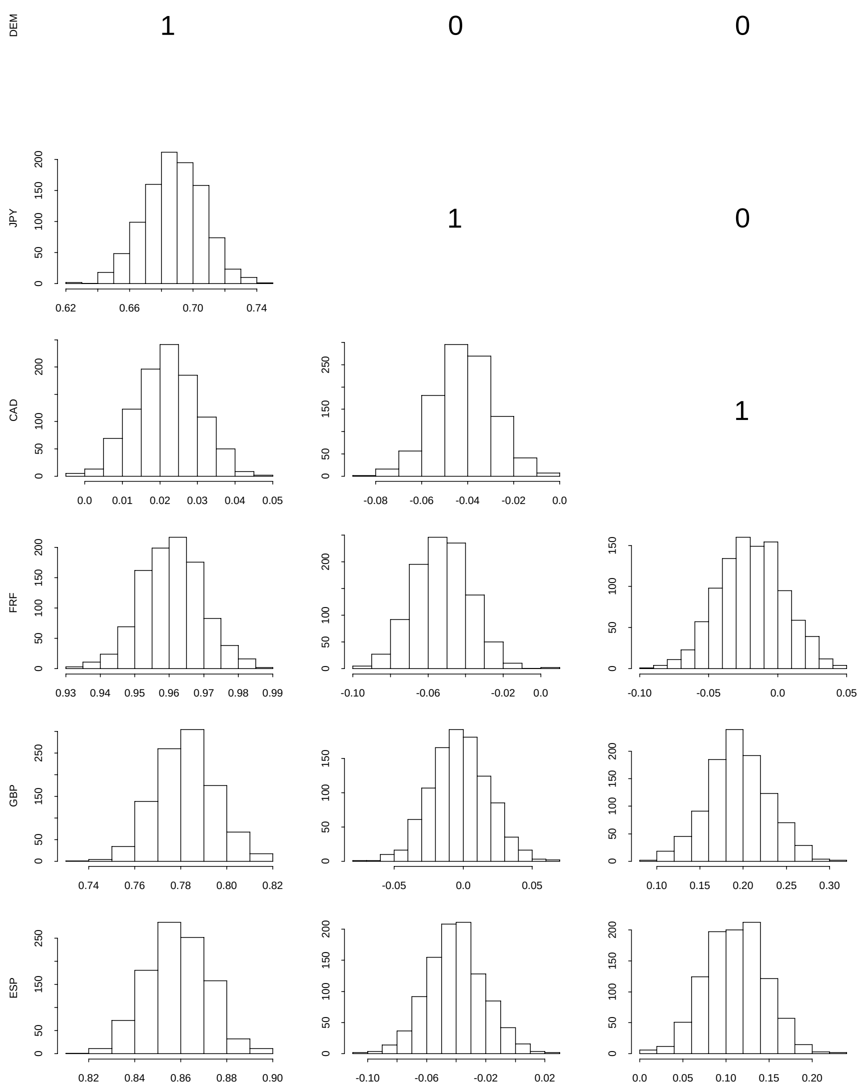

Figure 7: Posterior summaries for factor loadings $\mathbf{X}$.

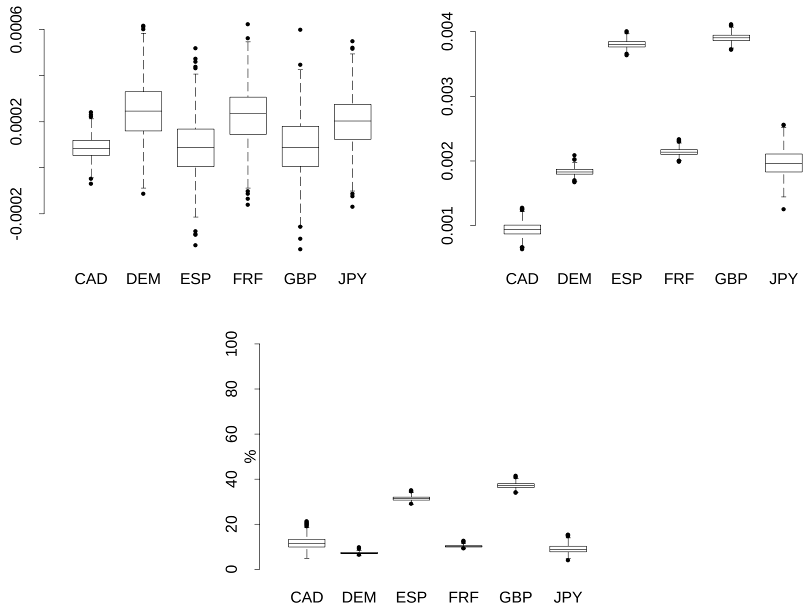

Figure 8: Posterior summaries for $\boldsymbol{\theta}$ and $\boldsymbol{\Psi}$.

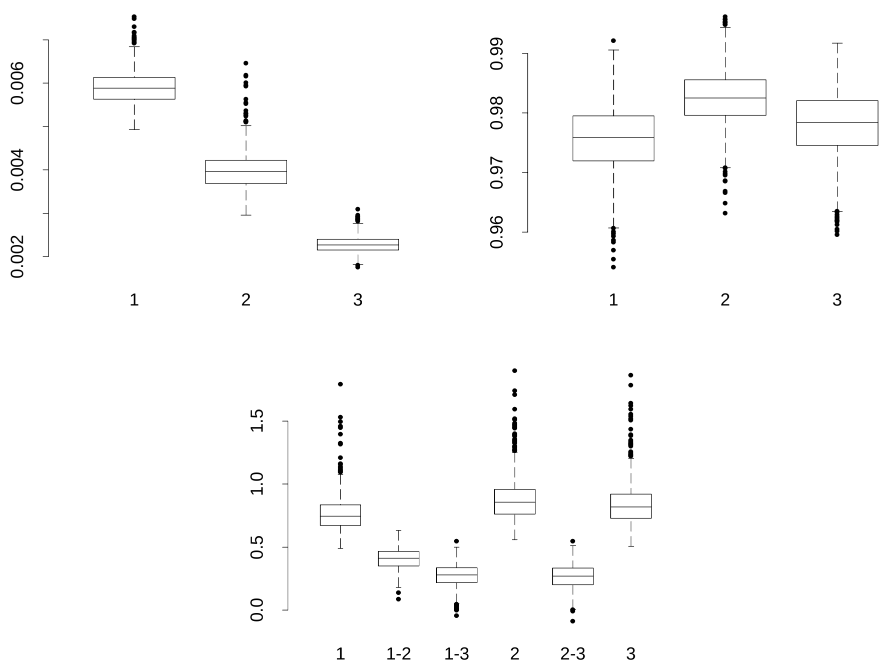

Figure 9: Posterior summaries for $\boldsymbol{\mu}$, $\boldsymbol{\Phi}$ and $\mathbf{U}$.

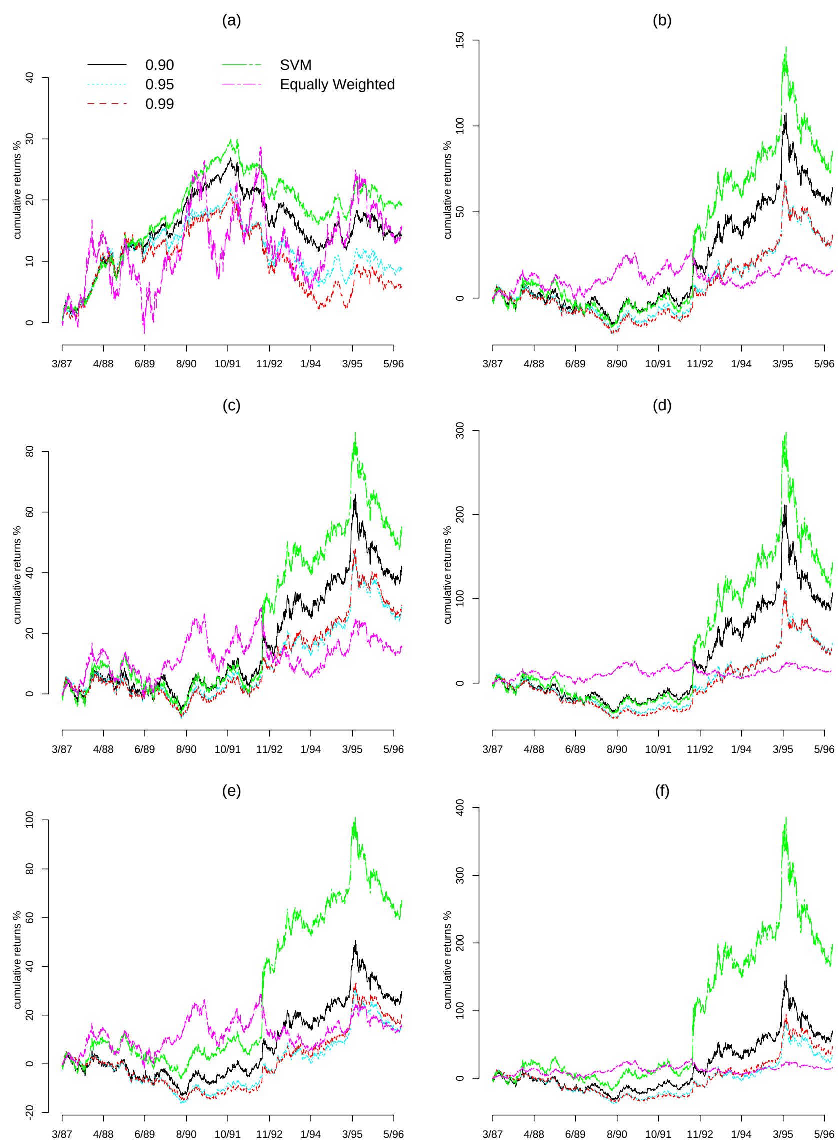

Figure 10: Cumulative returns under different dynamic portfolios.

Figure 11: Dynamic weights $\mathbf{a}_t^{(me)}$ for the mean-efficient portfolio: all models.

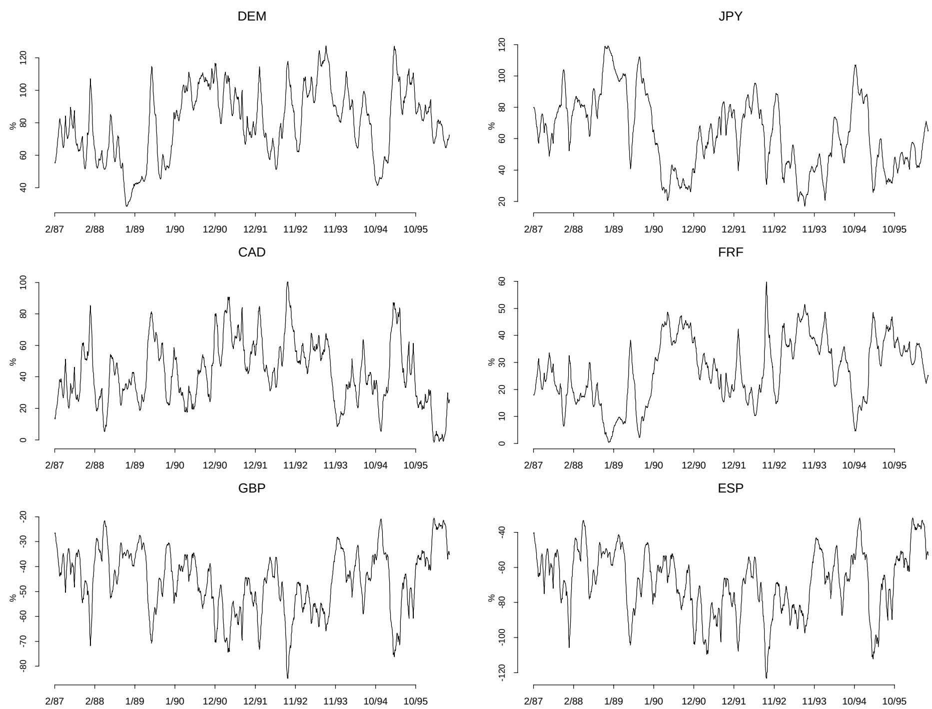

Figure 12: Dynamic weights $\mathbf{a}_t^{(*m)}$ with $m = 0.00016$ in the unit-sum, constrained portfolios: factor model.

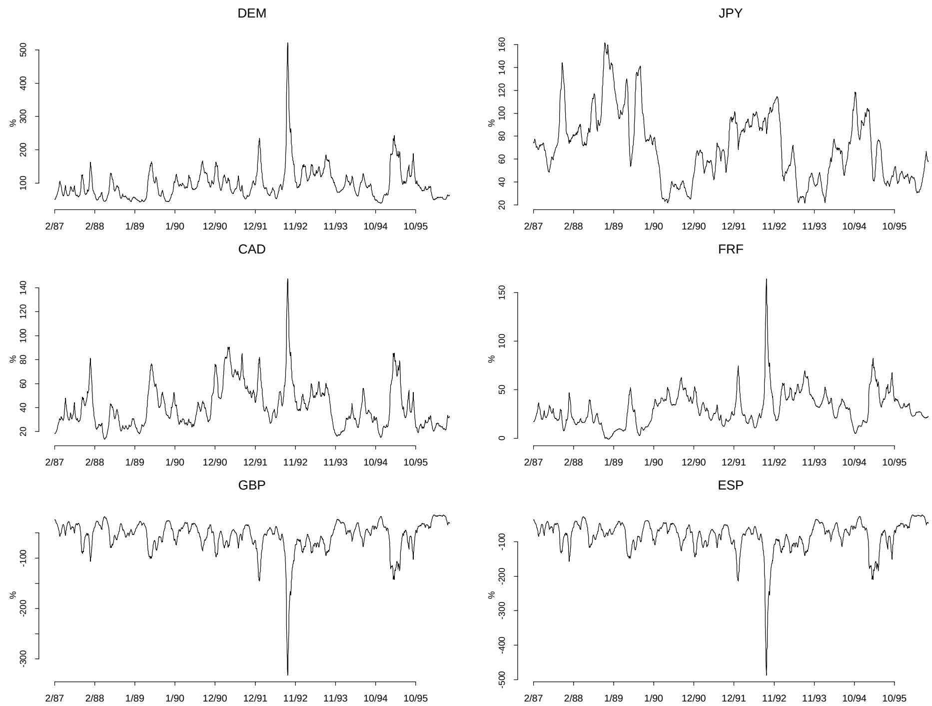

Figure 13: Dynamic weights $\mathbf{a}_t^{(*m)}$ with $m = 0.00016$ in the unconstrained portfolios: factor model. The weights are graphed here as percentages of the totals $\mathbf{1}'\mathbf{a}_t^{(*m)}$ at each time.

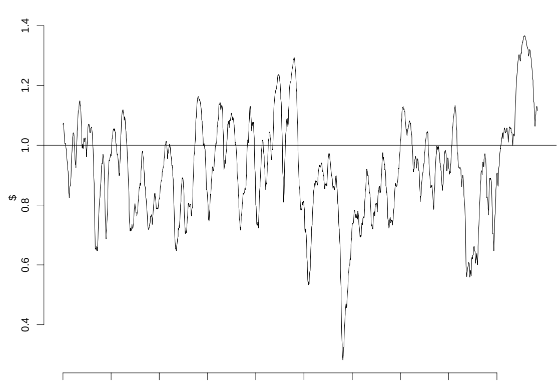

Figure 14: Total investments $\mathbf{1}'\mathbf{a}_t^{(*m)}$ with $m = 0.00016$ in the unconstrained portfolios: factor model.