Journal of Financial Econometrics, 2013, Vol. 11, No. 1, 116–153

Dynamic Factor Volatility Modeling: A Bayesian Latent Threshold Approach

JOUCHI NAKAJIMA
Department of Statistical Science, Duke University

MIKE WEST
Department of Statistical Science, Duke University

ABSTRACT
We discuss dynamic factor modeling of financial time series using a latent threshold approach to factor volatility. This approach models time-varying patterns of occurrence of zero elements in factor loadings matrices, providing adaptation to changing relationships over time and dynamic model selection. We summarize Bayesian methods for model fitting and discuss analyses of several FX, commodities, and stock price index time series. Empirical results show that the latent threshold approach can define interpretable, data-driven, dynamic sparsity, leading to reduced estimation uncertainties, improved predictions, and portfolio performance in increasingly high-dimensional dynamic factor models. (*JEL*: C11, C53, C58)

KEYWORDS: Bayesian forecasting, latent threshold dynamic models, multivariate stochastic volatility, portfolio allocation, sparse time-varying loadings, time-varying variable selection

In time series portfolio analysis as in other areas of multivariate dynamic modeling, the increasing dimension of models drives a need for refined prior: model structuring for parsimony in the face of increasingly high-dimensional parameters. In dynamic settings, time-varying model parameter processes may take practically non-negligible values for some epochs, but be effectively zero and practically

The authors thank the editor, the associate editor, and an anonymous referee, for helpful comments and suggestions. This work was supported by the National Science Foundation [DMS-1106516 to M.W.]. Any opinions, findings, and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the National Science Foundation. Address correspondence to Jouchi Nakajima, Department of Statistical Science, Duke University, Box 90251, Durham, NC 27708-0251, USA, or e-mail: jouchi.nakajima@stat.duke.edu.

doi:10.1093/jjfinec/nbs013 Advanced Access publication September 3, 2012

irrelevant at other times. Recently developed Bayesian *sparsity modeling* approaches such as graphical models (e.g., Carvalho and West 2007; Wang and West 2009; Wang 2010) and sparse factor models (e.g., West 2003; Carvalho et al. 2008; Yoshida and West 2010; Carvalho, Lopes, and Aguilar 2011) address *global* sparsity—permitting data-based shrinkage of subsets of model parameters to zero using Bayesian sparsity priors. However, such approaches have not been able to provide a rigorous and satisfactory strategy for models and relationships exhibiting *time-varying sparsity*, where dynamic parameters may be zero for periods of time but non-zero elsewhere. In a recent paper, we introduced a novel and general approach based on *latent threshold modeling*, and demonstrated its efficacy in dynamic regression variable selection as well as time-varying vector autoregressions (Nakajima and West 2010). The current paper extends this to a class of multivariate dynamic factor and regression models for financial time series and evaluates the resulting inferences, predictions, and portfolio decisions against standard approaches in empirical studies.

Dynamic factor models have become standard tools for multivariate time series econometrics and for multivariate stochastic volatility in particular (Aguilar et al. 1999; Pitt and Shephard 1999; Aguilar and West 2000). We adopt a general framework that builds on these prior approaches and extends to time-varying factor loadings matrices (Lopes and Carvalho 2007; Carvalho, Lopes, and Aguilar 2011) as well as integrating short-term dependence structure into latent factor processes. Overlaying the resulting general class of dynamic factor models are the new latent threshold mechanisms that define the ability for time-varying factor loadings to shrink to zero dynamically—i.e., dynamic sparsity models for loadings as well as possibly for other model components including dynamic regression coefficients.

Section 1 outlines the class of dynamic factor and regression models, defines the new approach to dynamic sparsity using latent threshold modeling, and discusses Bayesian analysis and computation for model fitting. A first example in Sections 2 and 3 uses FX data similar to data sets that have been previously analyzed using latent factor volatility models for portfolio studies; we give extensive and detailed evaluation of sets of models including comparisons to the use of standard models, and discuss aspects of model diagnostics. In Section 4, we develop a dynamic factor and regression model analysis of an extended FX and commodity price time series, further demonstrating the practical utility of the approach in short-term forecasting and portfolio decisions, and with additional model comparisons. Section 5 presents a further study of a 40-dimensional series of FX and stock price indices that, in part, anticipates future applications involving higher-dimensional time series. Section 6 provides some summary comments.

Some notation: We use the distributional notation $\boldsymbol{y} \sim N(\boldsymbol{a}, \boldsymbol{A})$, $d \sim U(a, b)$, $p \sim B(a, b)$, $v \sim G(a, b)$, for the normal, uniform, beta, and gamma distributions, respectively. We also use $s:t$ to denote $s, s+1, \ldots, t$ when $s < t$, for succinct subscripting; e.g., $\boldsymbol{y}_{1:T}$ denotes $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T\}$.

1 DYNAMIC FACTOR AND REGRESSION MODELS

1.1 General Model Form

The class of models for a $m\times 1$ vector response time series $\boldsymbol{y}_t$, $(t=1,2,\ldots)$ is

$$\boldsymbol{y}_t=\boldsymbol{c}_t+\boldsymbol{A}_t\boldsymbol{x}_t+\boldsymbol{B}_t\boldsymbol{f}_t+\boldsymbol{\nu}_t, \qquad \boldsymbol{\nu}_t\sim N(\boldsymbol{0},\boldsymbol{\Sigma}_t) \tag{1}$$

$$\boldsymbol{f}_t=\boldsymbol{G}_t\boldsymbol{f}_{t-1}+\boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t\sim N(\boldsymbol{0},\boldsymbol{\Psi}_t) \tag{2}$$

with the following ingredients:

- $\boldsymbol{c}_t$ is a local trend term, or *local mean*;
- $\boldsymbol{x}_t$ is a $q\times 1$ vector of covariates known at time t with an associated $m\times q$ time-dependent parameter matrix $\boldsymbol{A}_t$;
- $\boldsymbol{f}_t$ is a $k\times 1$ vector of latent factors at time t, evolving according to a Vector Autoregression, VAR(1) model with possibly time-varying AR coefficient matrix $\boldsymbol{G}_t$ and also possibly time-varying innovations variance matrix $\boldsymbol{\Psi}_t$ (both $k\times k$);
- $\boldsymbol{B}_t$ is the time-varying loadings matrix of the dynamic latent factor component of the model;
- $\boldsymbol{\nu}_t=(\nu_{1t},\ldots,\nu_{mt})'$ is a residual term with a diagonal residual, time-varying volatility matrix $\boldsymbol{\Sigma}_t$;
- the $\boldsymbol{\nu}_t$ and $\boldsymbol{\varepsilon}_s=(\varepsilon_{1t},\ldots,\varepsilon_{kt})'$ sequences are independent and mutually independent.

Following prior work, factor model identification is imposed by constraining $\boldsymbol{B}_t$ to have an upper right triangle of zeros and upper diagonal elements of unity (Aguilar and West, 2000; Lopes and West, 2004).

There are many possible choices of submodels for the time-varying structures of $\boldsymbol{c}_t$, $\boldsymbol{A}_t$, $\boldsymbol{B}_t$, $\boldsymbol{\Sigma}_t$, and $\boldsymbol{\Psi}_t$. We treat $\{\boldsymbol{A}_t,\boldsymbol{B}_t\}$ using the new latent threshold approach, this being the primary focus of the current work; this is described in Section 1.2 followed by choices of model forms for the remaining components in Section 1.3.

This model class represents a generalization of the standard factor models in a number of respects relevant to financial time series analysis. The elements of the factor loadings matrix (as well as regression parameter matrix) are allowed to change over time (Lopes and Carvalho 2007) while being subject to our novel latent thresholding mechanism to shrink them to zero when data suggest they are irrelevant. Further, the latent factors follow an autoregressive process, so providing opportunity for improvements in short-term forecasting of underlying time-varying common movements in the multiple time series. Adding volatility to this latter component is one further extension of some practical value, discussed below.

1.2 Dynamic Latent Threshold Model Components

We base development on an underlying AR(1) process structure for each of the free elements of $\boldsymbol{B}_t$ as well as for all univariate entries in $\boldsymbol{A}_t$. For the factor loadings, this is related to process models adopted for dynamic factor loadings in Lopes and Carvalho (2007). This basic idea is now adapted using the latent threshold mechanism of Nakajima and West (2010), as follows.

Dynamic regression parameter processes: For $\boldsymbol{A}_t$ we adopt the latent threshold model of Nakajima and West (2010) directly. Write a_{it} for the i-th element of the $mq \times 1$ vector $\boldsymbol{a}_t \equiv \text{vec}(\boldsymbol{A}_t)$. The model assumes the a_{it} independently drawn from processes as follows:

$$a_{it} = \alpha_{it} s_{ait} \quad \text{with} \quad s_{ait} = I(|\alpha_{it}| \geq d_{ai}), \quad i = 1:mq, \tag{3}$$

where α_{it} is a *latent* AR(1) process

$$\alpha_{it} = \mu_{\alpha i} + \phi_{\alpha i}(\alpha_{i,t-1} - \mu_{\alpha i}) + \eta_{\alpha it}, \quad \eta_{\alpha it} \sim N(0, v_{\alpha i}), \tag{4}$$

with $|\phi_{\alpha i}| < 1$ for each $i = 1:mq$.

Dynamic factor loadings processes: We use the same form of construction for the free parameter processes in $\boldsymbol{B}_t$, i.e., those elements below the upper diagonal of the $m \times k$ matrix $\boldsymbol{B}_t$ (n.b. $m > k$). Write $\boldsymbol{b}_t$ for the column stack of those free elements, so that $\boldsymbol{b}_t$ is a $p \times 1$ vector where $p = mk - k(k+1)/2$ with univariate elements b_{it}, $i = 1:p$. The model assumes the b_{it} independently drawn from processes as follows:

$$b_{it} = \beta_{it} s_{bit} \quad \text{with} \quad s_{bit} = I(|\beta_{it}| \geq d_{bi}), \quad i = 1:p, \tag{5}$$

where β_{it} is a *latent* AR(1) process

$$\beta_{it} = \mu_{\beta i} + \phi_{\beta i}(\beta_{i,t-1} - \mu_{\beta i}) + \eta_{\beta it}, \quad \eta_{\beta it} \sim N(0, v_{\beta i}), \tag{6}$$

with $|\phi_{\beta i}| < 1$ for each $i = 1:p$.

The key idea of the latent threshold structure is that the value of each of the dynamic regression and time-varying loading parameters is shrunk to zero when its absolute value falls below a coefficient-specific threshold. On the factor component, the relevance of each factor to the response is time-varying; a factor plays a role in predicting the response only when the corresponding β_{it} is "large enough." For each time series response variable, the factor may have a non-zero loading in some time periods but zero in others, depending on the data and context. This mechanism leads to dynamic model uncertainty in the factor loading matrix by neatly embodying time-varying sparsity/shrinkage and parameter reduction.

1.3 Subsidiary Model Components

Local mean: The time-varying intercept, or local mean $\boldsymbol{c}_t = (c_{1t}, \ldots, c_{mt})'$ evolves based on stationary, independent AR(1) models

$$c_{it} = \mu_{ci} + \phi_{ci}(c_{i,t-1} - \mu_{ci}) + \eta_{cit}, \quad \eta_{cit} \sim N(0, v_{ci}), \tag{7}$$

with $|\phi_{ci}| < 1$ for each $i = 1\!:\!m$.

Residual volatility: With $\boldsymbol{\Sigma}_t = \text{diag}(\sigma_{1t}^2, \ldots, \sigma_{mt}^2)$, we adopt standard stochastic volatility process models (e.g., Jacquier, Polson, and Rossi 1994; Kim, Shephard, and Chib 1998; Aguilar and West 2000; Omori et al. 2007; Prado and West 2010, chap. 7) as follows. The log volatilities $\delta_{it} = \log \sigma_{it}^2$ follow independent, stationary AR(1) models

$$\delta_{it} = \mu_{\delta i} + \phi_{\delta i}(\delta_{i,t-1} - \mu_{\delta i}) + \eta_{\delta it}, \quad \eta_{\delta it} \sim N(0, v_{\delta i}), \tag{8}$$

with $|\phi_{\delta i}| < 1$ for each $i = 1\!:\!m$.

Factor evolution: For the parameters of the factor evolution model in Equation (2) our empirical examples constrain to a constant and diagonal autoregressive parameter matrix $\boldsymbol{G}_t \equiv \boldsymbol{G} = \text{diag}(\gamma_1, \ldots, \gamma_k)$ for all t. We do generally expect—and observe the need for—dynamics in the volatility of factor processes, so take a volatility model for the innovations variance matrix $\boldsymbol{\Psi}_t$ similar to that for $\boldsymbol{\Sigma}_t$. That is, $\boldsymbol{\Psi}_t = \text{diag}(\psi_{1t}^2, \ldots, \psi_{kt}^2)$ and, defining $\lambda_{it} = \log \psi_{it}^2$ for $i = 1\!:\!k$, we have independent, stationary AR(1) models

$$\lambda_{it} = \mu_{\lambda i} + \phi_{\lambda i}(\lambda_{i,t-1} - \mu_{\lambda i}) + \eta_{\lambda it}, \quad \eta_{\lambda it} \sim N(0, v_{\lambda i}), \tag{9}$$

with $|\phi_{\lambda i}| < 1$ for each $i = 1\!:\!k$.

The resulting model class is sufficient for our current purposes but can clearly be expanded with more elaborate submodels if context demands in new applications.

1.4 Bayesian Analysis and Computation

We refer to the framework here as a *Latent Threshold Dynamic Factor Model (LTDFM)*. Bayesian analysis uses Markov chain Monte Carlo (MCMC) methods, building on nowadays standard simulation methods in a set of computational "modules." These include conditional samplers for latent factor models (Aguilar and West 2000; Lopes and West 2004) and for univariate stochastic volatility models (Shephard and Pitt 1997; Kim, Shephard, and Chib 1998; Watanabe and Omori 2004; Omori et al. 2007). These are coupled with additional, direct conditional samplers using the forward filtering, backward sampling strategy for state space dynamic models (e.g., Prado and West 2010) and a set of univariate, conditional Metropolis–Hastings samplers. One key computational novelty is the extension of a Metropolis Hastings algorithm of Nakajima and West (2010) required for the latent threshold component of the LTDFM.

In the context of observing data $\{\boldsymbol{y}_{1:T}, \boldsymbol{x}_{1:T}\}$ over a time interval $1:T$, the full set of latent process state parameters and model hyper-parameters for inclusion in the posterior analysis are as follows:

- The local trend and latent factor process states $\boldsymbol{c}_{0:T}$ and $\boldsymbol{f}_{0:T}$ including their uncertain initial values at $t=0$;
- The log volatility processes $\delta_{1:m,1:T}$ and $\lambda_{1:k,1:T}$;
- Hyper-parameters defining each of the component univariate AR(1) submodels, namely
  - $\gamma_{1:k}$, the AR(1) coefficients of the factor evolution matrix $\boldsymbol{G}_t \equiv \boldsymbol{G} = \text{diag}(\gamma_{1:k})$,
  - $\boldsymbol{\theta}_c \equiv \{\mu_{ci}, \phi_{ci}, v_{ci}; i=1:m\}$,
  - $\boldsymbol{\theta}_\alpha \equiv \{\mu_{\alpha i}, \phi_{\alpha i}, v_{\beta i}; i=1:mq\}$,
  - $\boldsymbol{\theta}_\beta \equiv \{\mu_{\beta i}, \phi_{\beta i}, v_{\beta i}; i=1:p\}$,
  - $\boldsymbol{\theta}_\delta \equiv \{\mu_{\delta i}, \phi_{\delta i}, v_{\delta i}; i=1:m\}$,
  - $\boldsymbol{\theta}_\lambda \equiv \{\mu_{\lambda i}, \phi_{\lambda i}, v_{\lambda i}; i=1:k\}$;
- The dynamic regression and factor loading parameter process states $\alpha_{i,0:T}$, $(i=1:mq)$, and $\beta_{j,0:T}$, $(j=1:p)$, including their values at $t=0$;
- $\boldsymbol{d} = \{d_{ai}, i=1:mq; d_{bj}, j=1:p\}$, the sets of thresholds in Equation (3) and (5).

The key analysis components for each of the sets of parameters and states are as follows. In each, we simply note the states or parameters being sampled, implicitly conditional on all other states and parameters if not explicitly qualified.

Trend and latent factor process states $\boldsymbol{c}_{0:T}$ and $\boldsymbol{f}_{0:T}$: Conditional on model hyper-parameters, volatility process states and the data, the LTDFM of Equation (1) and (2) reduces to a conditionally linear, Gaussian dynamic model for these states. Resampling the full sets of states is then easily obtained by a direct application of the standard forward filtering, backward sampling (FFBS) algorithm (e.g., Prado and West 2010). Importantly, this is an efficient algorithm that regenerates full trajectories of these latent states over $0:T$ at each iterate of the overall MCMC. These new values are then conditioned upon for resampling of other components.

Log volatility processes: We sample the conditional posteriors for each of the latent log volatility processes $\delta_{i,1:T}$, $(i=1:m)$, and $\lambda_{j,1:T}$, $(j=1:k)$ using the standard MCMC technique for univariate stochastic volatility models (Shephard and Pitt 1997; Kim, Shephard, and Chib 1998; Watanabe and Omori 2004; Omori et al. 2007). Across i,j, these processes are conditionally independent in the posteriors given in all the other model quantities, so this applies in parallel defining efficient resampling of full trajectories at each stage of the overall iterative MCMC. Given values of these volatility trajectories, the diagonal volatility matrices of the LTDFM are instantiated directly, viz $\boldsymbol{\Sigma}_t = \mathrm{diag}(\sigma_{1t}^2, \ldots, \sigma_{mt}^2)$ where each $\sigma_{it}^2 = \exp(\delta_{it})$, and $\boldsymbol{\Psi}_t = \mathrm{diag}(\psi_{1t}^2, \ldots, \psi_{kt}^2)$ where each $\psi_{it}^2 = \exp(\lambda_{it})$.

AR hyper-parameters $\boldsymbol{G} = \mathrm{diag}(\gamma_1, \ldots, \gamma_k)$: Assume independent priors for the γ_i using traditional forms—either truncated normal or shifted beta priors to constrain to the stationary range. We then sample the conditional posteriors either directly or via Metropolis–Hastings accept/reject steps.

AR hyper-parameters $\boldsymbol{\theta}_*$: For each $* \in \{c, \alpha, \beta, \delta, \lambda\}$, we assume prior independence across $*$ and across i for each $*$, with traditional forms of priors for the AR model parameters $(\mu_{*i}, \phi_{*i}, v_{*i})$. That is, we use normal or log-gamma priors for μ_*, truncated normal or shifted beta priors for ϕ_* and inverse gamma priors for v_*. Conditional posteriors can be sampled directly or via Metropolis-Hastings accept/reject steps.

Latent thresholded dynamic regression and factor loadings: Sampling conditional posteriors of all elements of the $\alpha_{i,0:T}$, $(i = 1:mq)$, and $\beta_{j,0:T}$, $(j = 1:p)$, is an extension to the LTDFM of the algorithm introduced in Nakajima and West (2010). We note details for the factor loadings matrix process only here, as that for the dynamic regression parameters is essentially the same. Recall that $\boldsymbol{b}_t$ is the $p \times 1$ vector representing the column stack of the free elements $\boldsymbol{B}_t$; denote by $\boldsymbol{\beta}_t$ the corresponding vector of latent AR(1) processes underlying $\boldsymbol{b}_t$ as defined by Equation (5). At each stage of the overall MCMC, new values are sampled by sequencing through sets of conditional distributions for each $\boldsymbol{\beta}_t$ given $\boldsymbol{\beta}_{-t} = \boldsymbol{\beta}_{0:T} \backslash \boldsymbol{\beta}_t$ and all other parameters. In this Metropolis-within-Gibbs sampling strategy, proposed vectors $\boldsymbol{\beta}_t$ are drawn trivially from the underlying *non-threshold* model obtained by setting each $s_{bit} = 1$, i.e., assuming no sparsity in the loadings matrix at that time t as a Metropolis proposal. Details of the resulting accept/reject probability computation then follow as in dynamic regression applications detailed in Section 2.3 of Nakajima and West (2010). Each MCMC iterate sequences through this process over $t = 0:T$ to resample the entire sequence of $\boldsymbol{\beta}_t$ vectors. Given these trajectories and threshold parameters $\boldsymbol{d}$, the indicators s_{bit} are instantiated, so directly generating values of the full sequence of factor loadings matrices $\boldsymbol{B}_{0:T}$. The same strategy then also applies to resample $\boldsymbol{A}_{0:T}$, as already mentioned.

Threshold parameters $\boldsymbol{d}$: Again following Nakajima and West (2010), we adopt priors for the thresholds that assume the d_{*i} to be independently uniform, for each of $* \in \{a, b\}$ and all i. We detail this for the thresholds of the $\boldsymbol{B}_t$ model, with those of the $\boldsymbol{A}_t$ model being the same but for notation. Specifically, $d_{bj} \sim U(0, |\mu_{\beta j}| + K u_{\beta j}^{1/2})$ where $u_{\beta j} = v_{\beta j}/(1 - \phi_{\beta j}^2)$ is the variance of the stationary marginal distribution of the univariate AR(1) process for β_{jt}. For $K = 3$ or so as adopted and argued by Nakajima and West (2010), these priors cover the range of the β_{jt} as reflected in their marginal distributions, uniformly supporting a relevant range of values of the threshold. Sampling new thresholds from the implied conditional posteriors when given values of the $\boldsymbol{B}_t$ and all other model parameters uses a direct Metropolis–Hastings independence chain step with candidate values for each d_{bj} drawn from the uniform prior.

1.5 Identification

Some comments on general questions of identification are in order. First, the overall latent threshold model strategy inherently promotes data-based shrinkage—completely to zero—of multiple factor loadings when it matters. Coupled with the assumed upper triangular form of the factor loadings matrix $\boldsymbol{B}_t$ to "name" and order factors according to the chosen first k time series variables, the results are inferences on, typically, quite sparse structures. The triangular form ensures mathematical identification, in terms of lack of an ability to "rotate" factors coupled with a restricted number of free elements in $\boldsymbol{B}_t$ (Aguilar and West 2000; Lopes and West 2004). Thresholding then typically reduces the number of non-zero elements more dramatically; in this sense, the model is automatically self-identifying.

One specific additional consideration is the use of models in which both factor processes and factor loadings are time-varying, and whether this introduces the potential for new identification questions. There are two points to note. First, again, thresholding matters: a thresholded value of a factor loading wipes out the contribution of the factor itself over the period of time it is below threshold; so the actual values of the underlying latent factor process are irrelevant over that period. Second, and most importantly, careful specification of informative priors over the parameters of the time series models for the factor loadings—the latent AR(1) process models of Equation (6)—provides *critical* control over where the resulting posteriors place mass. Models with rather vague priors that would allow for very widely fluctuating patterns of change over time in the factor loadings are irrelevant and dangerous, as they could engender models with wild variation in loadings and relatively insignificant factor processes themselves, distort inferences and certainly lead to weak identification problems. The same general comment applies to all hyper-parameters, of course, but is most particularly key in connection with the factor loadings model components. In reflection of this, our priors favor highly persistent stationary AR(1) processes for β_{it}, strongly limiting the global fluctuations in loadings relative to potential fluctuation in the factors. This is attained with priors for the $\phi_{\beta i}$ of Equation (6) that concentrate close to one, and with priors for the innovations variances $v_{\beta i}$ of that equation that strongly favor small values, again in the context of expected ranges of variation in the factor processes as specified in corresponding priors for their hyper-parameters.

Across our examples here, and a range of additional empirical studies, we have observed posterior inferences on factors and loadings that are evidently interpretable, and supporting the notion of identified models. Monitoring MCMC streams has not, in any of our examples reported, raised any concerns about the MCMC switching between different regions in the space of factor loadings and

Table 1 International currencies relative to the U.S. dollar in exchange rate data

1	GBP	British Pound Sterling
2	EUR	Euro
3	JPY	Japanese Yen
4	CAD	Canadian Dollar
5	AUD	Australian Dollar
6	CHF	Swiss Franc

factor values, nor of steady "drifts," either of which could suggest identification issues if encountered. Again, these analyses are based on models with carefully evaluated informative priors that, in part, exert control to rule out such potential problems.

2 A STUDY OF EXCHANGE RATE RETURNS

The first study applies the LTDFM to a series of daily foreign exchange (FX) rate returns. Our focus here is particularly on how the model, data match engenders shrinkage to zero of time-varying factor loadings, coupled with the follow-on implications for the dynamic relationships underlying the FX series. We note the connections with previous work on FX time series using factor models (Aguilar and West 2000; Lopes and West 2004; Lopes and Carvalho 2007).

2.1 Data and Model Setup

The data are $m=6$ daily international currency exchange rates relative to U.S. dollar over a time period of 980 business days beginning in January 2006 and ending in December 2009. The returns are computed as $y_{it}=100(p_{it}/p_{i,t-1}-1)$, where p_{it} denotes the daily closing spot rate. The series are listed in Table 1; the exchange rates and returns are plotted in Figure 1. The most striking feature in these sample periods is a quite volatile and turbulent movement triggered by the financial crisis around 2008; there are some superficial correlated flows and shifts among the series. Initial exploratory analysis using the univariate stochastic volatility fit to each series separately is summarized in Figure 2. There exist marked region-specific patterns. The peaks of stochastic volatility seem to coincide in 2008, but there exist several spikes common only to two or three currencies. From these exploratory analysis, we chose a LTDFM specification in which the currencies are ordered in the $\boldsymbol{y}_t$ vectors as in Table 1.

Our analysis uses a special case of the general LTDFM presented above in which there are no covariates, i.e., $\boldsymbol{x}_t=\mathbf{0}, \boldsymbol{A}_t=\mathbf{0}$ for all t. The following priors are used in our first example: $1/v_{ci}\sim G(40, 0.005)$; $1/v_{\beta i}\sim G(40, 0.02)$; $1/v_{\delta i}\sim G(20, 0.001)$;

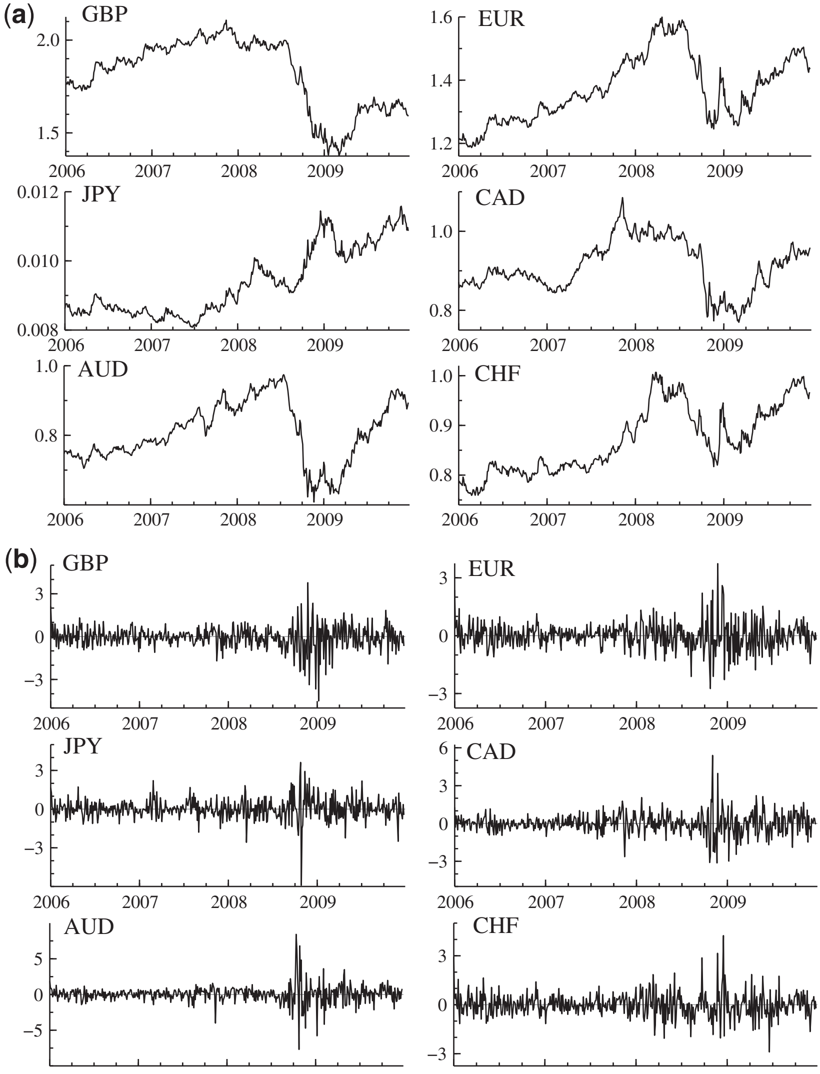

Figure 1 Daily (a) exchange rates and (b) returns.

$(\phi_{*i}+1)/2 \sim B(20, 1.5)$ for each $* \in \{c, \beta, \delta, \lambda\}$, and $(\gamma_i+1)/2 \sim B(1,1)$. For each $* \in \{c, \beta\}$ we use priors $\mu_{*i} \sim N(0,1)$ whereas for each $* \in \{\delta, \lambda\}$ we take $\exp(-\mu_{*i}) \sim G(3, 0.03)$. The MCMC analysis was run for a burn-in period of 20,000 samples prior to saving the following MCMC sample of size $J = 100{,}000$ for summary posterior inferences. Computations were performed using custom code in Ox (Doornik 2006); the code is available to interested readers.

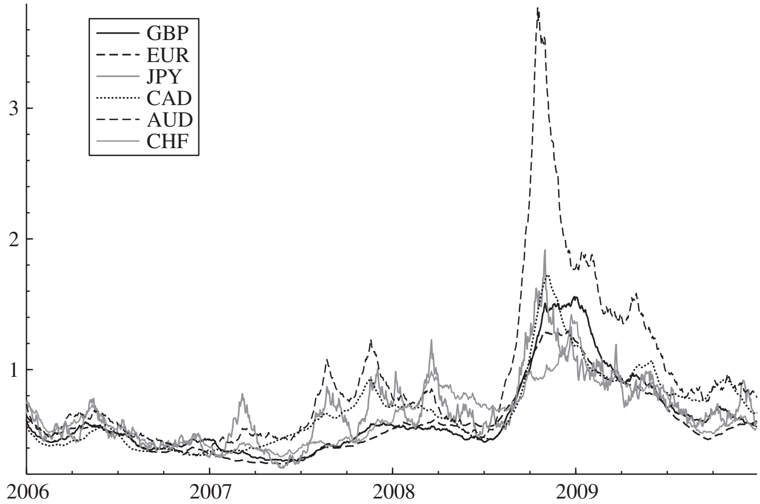

Figure 2 Estimated trajectories of univariate stochastic volatility ($\sigma_{it} = \exp(\delta_{it}/2)$) fit to each series separately.

Determination of the number of factors k has been a challenging issue in the literature (e.g., Lopes and West 2004). For the analysis here, we fit the LTDFM to the FX data and repeated the analysis across models with different number of factors $k = 1, 2, 3$, or 4. We evaluated each of the resulting models by forecasting over the final five business days using the first 930 observations, and repeated this five-step ahead forecasting based on the first 930+5j observations for $j = 1, \ldots, 9$, to obtain the 50-day forecasts. Based on the root mean squared errors (RMSE) for these out-of-sample forecasts, the LTDFM performs substantially better at $k = 3$ factors; this suggests a fourth factor is redundant while three are required, and we focus our analysis summaries on that of the model with $k = 3$. We comment further on this below, where we note some key aspects from the analysis of a model with $k = 4$.

2.2 Summaries of Posterior Inferences

Figure 3 displays posterior estimates of the latent time-varying factor loadings, as well as the posterior probabilities of $s_{bit} = 0$ for the LTDFM model. The estimated latent process exhibits considerable time variation for several loadings; this cannot be captured by a constant-loading factor model. In particular, JPY-Factor1 and AUD-Factor3 are estimated positive in 2006, but become negative after 2007, reflecting a shift of correlated dynamics among the FX series. The trajectories of posterior means of several loadings also evidently change over time. Time-varying sparsity is observed for the loadings of JPY-Factor1, AUD-Factor3, and CHF-Factor3, with estimated values shrinking to zero for some time periods but not others, whereas other loadings such as CAD-Factor2 and CHF-Factor3 are totally selected out over the entire time-frame.

The interesting observation of the shrinkage pattern of JPY-Factor1 in 2008 implies that the JPY returns turned to be negatively correlated with Factor1. The loadings on Factor1 are all positive in the currencies except JPY. In 2008, the financial crisis led investors to reduce their size of U.S. dollar short position and the currencies in our analysis except JPY depreciated against U.S. dollar, whereas the JPY temporarily appreciated.

Figure 4 shows the proportion of variation of the time series explained by each factor; the time variation in the contributions is evident. By construction, Factor1 is a GBP-leading factor, although its explanatory ratio decreases from 2006 to 2009, and there are a few downward spots in the crisis period, which indicates idiosyncratic variation of the GBP. Factor2 is recognized as an European factor that mainly explains the variation of EUR and CHF. The variation of JPY is of interest; Factor1 primarily explains its fluctuation in 2006, but later its role is shifted to Factor3, namely a JPY-leading factor. Clear shrinkage of the contribution in JPY-Factor1, CAD-Factor2, and CAD-Factor3 is detected by the LTDFM structure as observed in Figure 3; this data-induced shrinkage eliminates unnecessary fluctuations in what are negligible but non-zero values in time-varying loadings were we to use a non-thresholded model, and leads to efficient discrimination and contextual interpretation of the factors.

Figure 5 plots the posterior estimates of time trajectories of the factors and their innovation volatilities. The trajectories and peak of the stochastic volatility remarkably differ among the factors. The stochastic volatility of Factor1 has the peak around September 2008 corresponding to the major crash of the market, whereas that of Factor2 has its peak around December 2008, when EUR and CHF returns exhibit relatively higher volatility in reaction to the U.S. Senate's rejection of the financial bailout for the automotive industry. The stochastic volatility of Factor1 exhibits a steeper hike toward the peak and a more moderate diminishing afterwards compared to that of Factor2. We also find that the posteriors for the AR(1) γ_i parameters place substantial mass around zero, indicating lack of predictability in the latent factor processes, i.e., that the factor contributions are effectively unrelated over time, consistent with assumptions of factor volatility modeling of FX returns in prior work.

Figure 6 graphs posterior means of series-specific stochastic volatility, $\sigma_{it} = \exp(\delta_{it}/2)$, that is, volatility elements excluding the common factors, exhibiting quite a different picture from Figure 2. GBP has a unique hike of volatility in the second half of 2008 and CHF has several spikes in 2008 and 2009, which are not explained by the factors. Because most of the fluctuation in EUR returns are explained by Factor1 and Factor2 as can be seen in Figure 4, the remaining series-specific shocks are relatively smaller than the others as shown in Figure 6.

As mentioned above, we determined the number of factors as $k=3$ based on out-of-sample, multi-day ahead forecasting performance. However, we also note

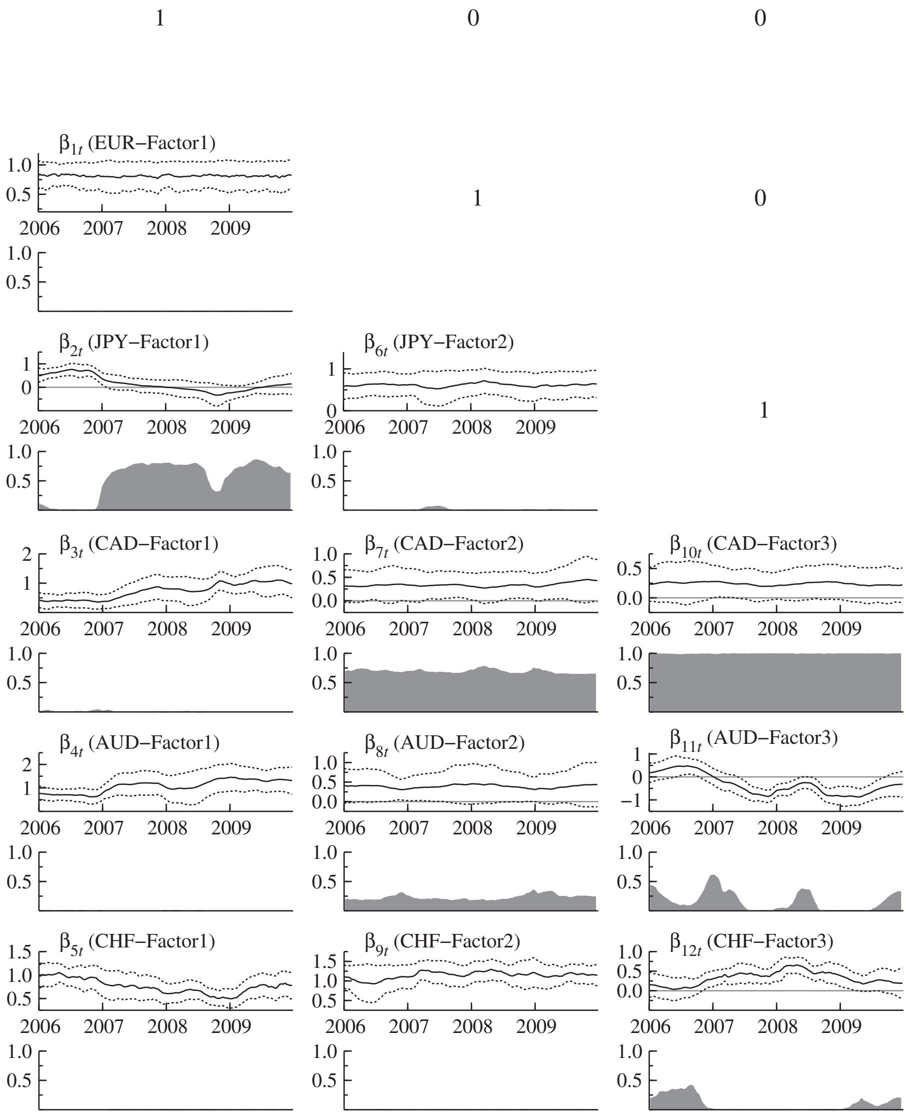

Figure 3 Posterior means (solid) and 95% credible intervals (dotted) of time-varying factor loadings. Posterior probabilities of $s_{bit} = 0$ are plotted below each trajectory.

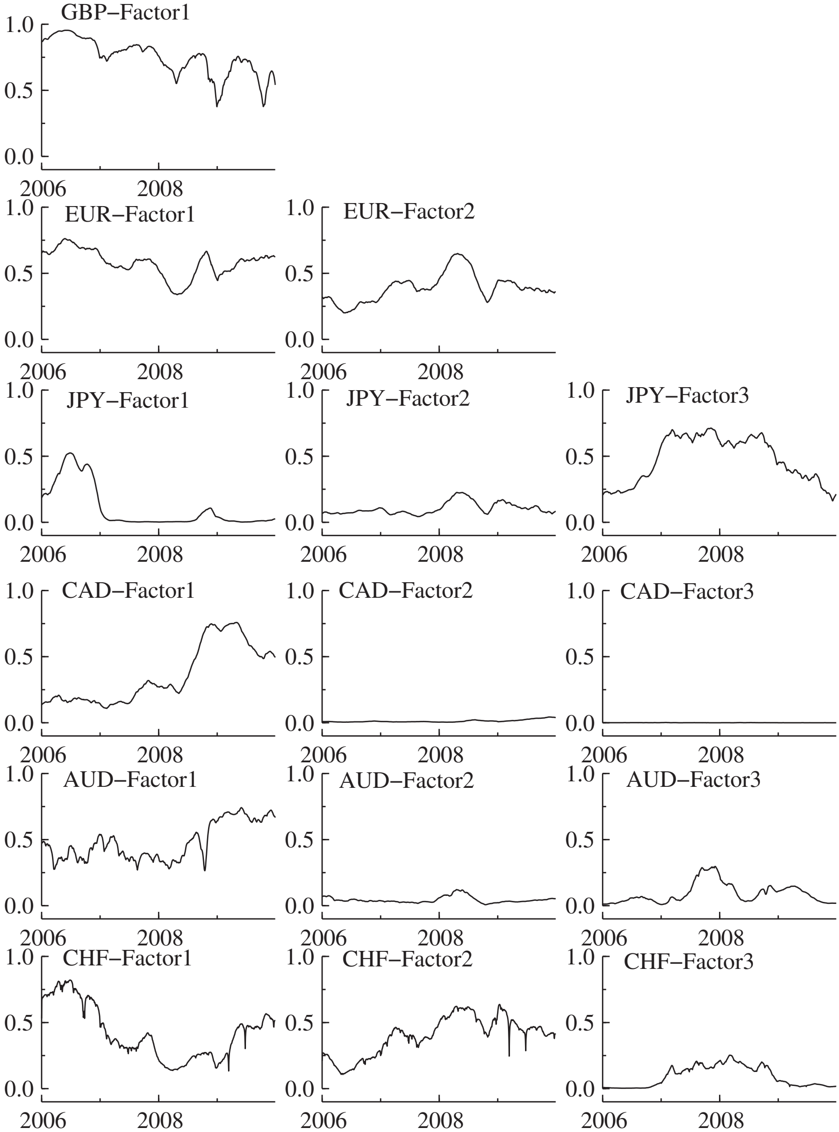

Figure 4 Time series trajectories of proportion of variances explained by factors.

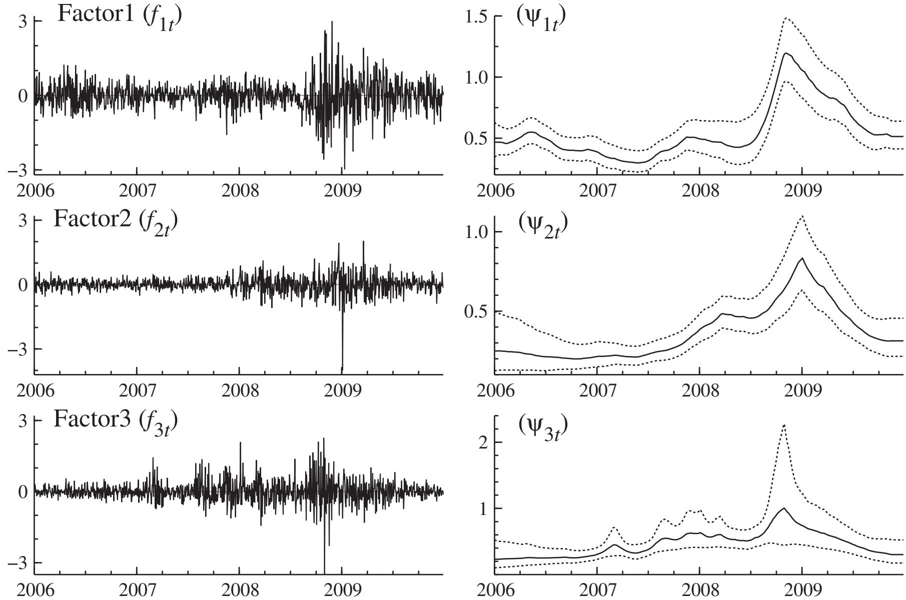

Figure 5 Left panels: posterior means of factors f_{it}. Right panels: posterior means (solid) and 95% credible intervals (dotted) of factor stochastic volatility $\psi_{it} = \exp(\lambda_{it}/2)$.

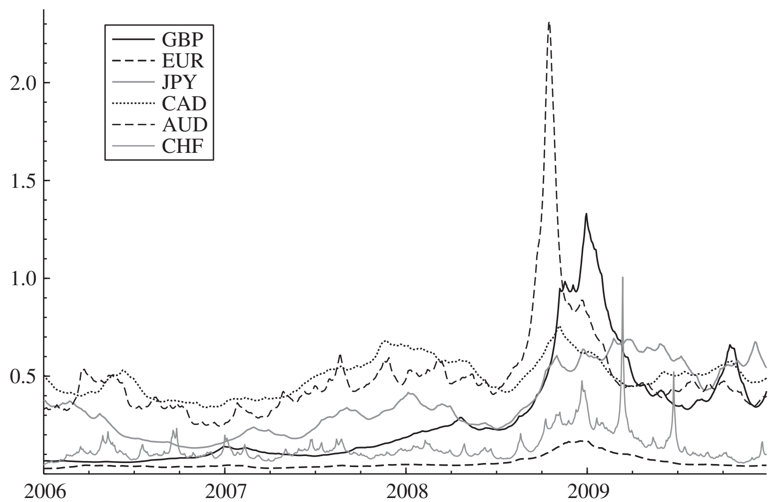

Figure 6 Posterior means of series-specific stochastic volatility ($\sigma_{it} = \exp(\delta_{it}/2)$).

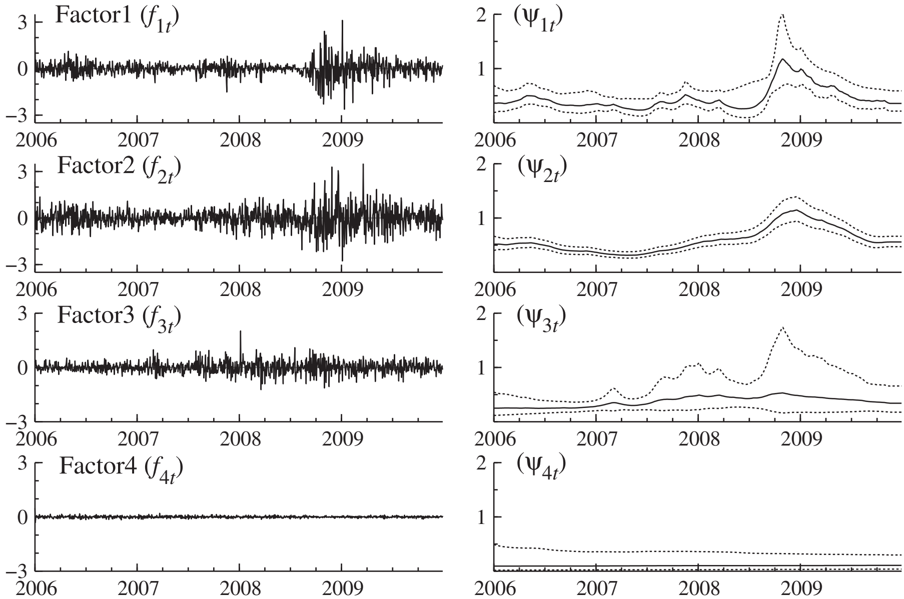

Figure 7 Results of FX analysis in model with $k=4$ factors. Left panels: posterior means of factors f_{it}. Right panels: posterior means (solid) and 95% credible intervals (dotted) of factor stochastic volatility $\psi_{it}=\exp(\lambda_{it}/2)$.

some aspects of analysis results using $k=4$ factors. As can be seen in Figure 7, the estimated trajectory of Factor4 is clearly shrunk with significantly smaller stochastic volatility than the other three relevant factors. Figure 8 shows results of time-varying loadings and sparsity, indicating that the loadings of Factor4 are mostly shrunk in the entire periods. This finding implies that the LTDFM structure can facilitate data-driven model search and selection of factor models via shrinkage of time-varying loadings.

Our experiences with various orderings of the FX series in $\boldsymbol{y}_t$ vectors are that the ordering reported here yields more reasonable factor estimates with plausible econometric interpretations than other ordering. We continue with the model as specified here in the following portfolio analysis.

3 PORTFOLIO ALLOCATION ANALYSIS

This section explores the forecasting performance of the LTDFM in the context of dynamic portfolio allocations. This portfolio analysis follows previous Bayesian model evaluation using sequential portfolio decisions (Quintana 1992; Putnum and Quintana 1994; Quintana and Putnum 1996; Aguilar and West 2000; Carvalho

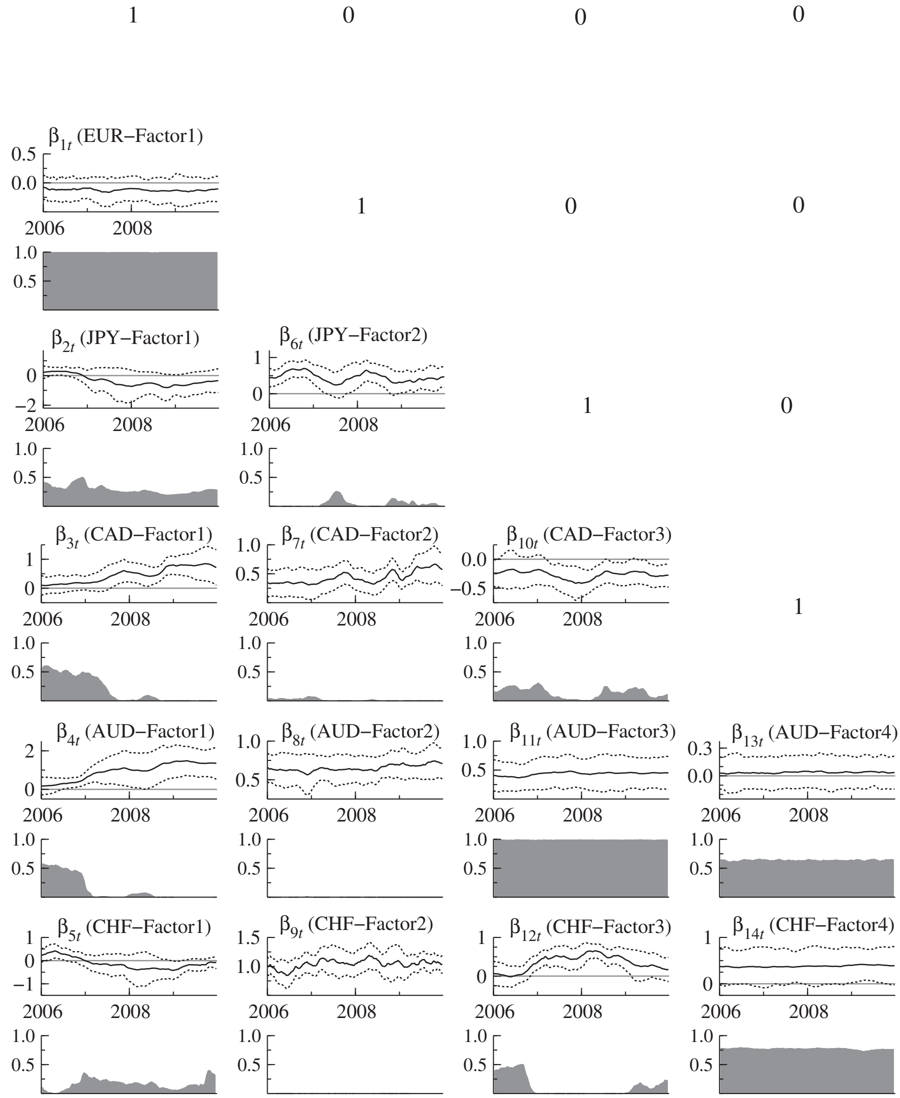

Figure 8 Results of FX analysis in model with $k=4$ factors. Posterior means (solid) and 95% credible intervals (dotted) of time-varying factor loadings. Posterior probabilities of $s_{bit}=0$ are plotted below each trajectory.

and West 2007; Carvalho, Lopes, and Aguilar 2011; Wang and West 2009). Our main focus is the impact of the LTDFM structure on forecast accuracy relative to standard dynamic factor models, considering how the dynamic sparsity/shrinkage in time-varying factor loadings plays a relevant role in investment experiments. We compare six competing models by implementing sequential portfolio allocations based on forecast means and variances of the FX series.

3.1 Portfolio Allocation Experiments

We reanalyze the FX time series restricting to only the first $T = 880$ observations. The remaining 100 business days are then used for forecasting, portfolio reallocations and model comparisons. After observing the closing rates on the business day $t-1$ ($\geq T$), the one-step ahead forecast mean vector and variance matrix of $\boldsymbol{y}_t$ are denoted by $\boldsymbol{g}_t$ and $\boldsymbol{Q}_t$; the implied precision matrix is then $\boldsymbol{K}_t = \boldsymbol{Q}_t^{-1}$. These are computed via the MCMC based on draws from one-step ahead predictive posterior distributions using the available data, $\boldsymbol{y}_{1:t-1}$. Investments are reallocated according to a vector of portfolio weights $\boldsymbol{w}_t$ optimized by a specific allocation rule, described below. The realized portfolio return at time t is $r_t = \boldsymbol{w}_t' \boldsymbol{y}_t$. The portfolio is reallocated on each business day based on one-step ahead forecasting computed via the MCMC given updated data. We fix the total sum invested on each business day by restricting $\boldsymbol{w}_t' \mathbf{1} = 1$.

We use traditional Bayesian mean-variance optimization (Markowitz 1959) subject to constraints. Given a scalar daily return target m, we optimize the portfolio weights $\boldsymbol{w}_t$, by minimizing the one-step ahead variance of returns among the restricted portfolios whose one-step ahead expectation is equal to m. Specifically, at time t, we minimize an *ex-ante* portfolio variance $\boldsymbol{w}_t' \boldsymbol{Q}_t \boldsymbol{w}_t$, subject to $\boldsymbol{w}_t' \boldsymbol{g}_t = m$, and $\boldsymbol{w}_t' \mathbf{1} = 1$. The solution is $\boldsymbol{w}_t^{(m)} = \boldsymbol{K}_t (a_t \boldsymbol{g}_t + b_t \mathbf{1})$, where $a_t = \mathbf{1}' \boldsymbol{K}_t \boldsymbol{e}$, and $b_t = -\boldsymbol{g}_t' \boldsymbol{K}_t \boldsymbol{e}$, where $\boldsymbol{e} = (\mathbf{1} m - \boldsymbol{g}_t)/d$, and $d = (\mathbf{1}' \boldsymbol{K}_t \mathbf{1})(\boldsymbol{g}_t' \boldsymbol{K}_t \boldsymbol{g}_t) - (\mathbf{1}' \boldsymbol{K}_t \boldsymbol{g}_t)^2$. We also consider the target-free minimum-variance portfolio given by $\boldsymbol{w}_t^* = \boldsymbol{K}_t \mathbf{1} / (\mathbf{1}' \boldsymbol{K}_t \mathbf{1})$. We implicitly assume that we can freely reallocate the resources to arbitrary long or short positions across the currencies without any transaction cost.

In addition to one-step ahead prediction, we also examine the same analysis strategy but now focused on five-step ahead forecasting and portfolio revisions. Every five business days, the posterior predictive distribution of five-step horizons, $(\boldsymbol{y}_t, \boldsymbol{y}_{t+1}, \ldots, \boldsymbol{y}_{t+4})$ is computed via MCMC based on the available data $\boldsymbol{y}_{1:t-1}$. This experiment assumes a possible situation that investors allocate their resource every business day based on weekly updated forecasts.

3.2 Model Evaluations and Comparisons

We consider the following three models from the class of LTDFMs:

- **LM-AF**: Local autoregressive means $\boldsymbol{c}_t$ and autoregressive factors $\boldsymbol{f}_t$.
- **LM-IF**: Local autoregressive means $\boldsymbol{c}_t$ and time-independent factors $\boldsymbol{f}_t$; i.e., $\boldsymbol{G} = \boldsymbol{O}$.

Table 2 Cumulative returns (%) over 100 business days

	LM-AF		LM-IF		CM-IF	
	NT	LT	NT	LT	NT	LT
One-step ahead forecasts						
Target return:						
$m = 0.02$	−1.08	1.25	1.18	3.11	−3.03	2.06
0.04	−2.20	1.52	−0.33	2.55	−4.85	1.61
0.06	−3.31	1.78	−1.85	1.99	−6.67	1.15
Target-free	−1.42	2.11	−1.26	0.98	−1.12	−0.04
Five-step ahead forecasts						
Target return:						
$m = 0.02$	−1.26	2.09	−1.54	0.95	−0.55	1.20
0.04	−0.96	1.88	−1.33	1.03	−3.12	0.82
0.06	−0.66	1.68	−1.12	1.11	−5.69	0.45
Target-free	−0.98	0.52	−0.55	1.25	−0.53	0.50

- **CM-IF**: Constant levels $\boldsymbol{c}_t \equiv \boldsymbol{c}$ and time-independent factors.

For each model of these three specifications, our study analyzed two variants: the LT (latent threshold) and NT (non-threshold) versions, giving a total of six competing models analyzed via the forecasting and portfolio strategy over the final 100 business days. For the CM-IF model, an additional prior is assumed: $\boldsymbol{c} \sim N(\boldsymbol{0}, \boldsymbol{I})$. The other prior specifications and simulation size are as in the previous analysis and detailed in Section 2.1. We assessed each model repeatedly using portfolios based on a range of daily target returns of $m = 0.02$, 0.04, and 0.06 percent, corresponding to monthly (20 day) return of approximately 0.4, 0.8 and 1.2 percent, respectively.

Table 2 reports cumulative returns of the portfolios resulting from the sequential investment over 100 business days. It is evident that the use of LT structure remarkably dominates the NT models regardless of other aspects of model specification or the portfolio allocation rules. Figure 9 displays the cumulative returns across time periods from one-step ahead forecasting. LT models clearly outperform NT models and their difference is larger in the mean-target strategy of $m = 0.06\%$ than in the target-free minimum-variance portfolio. A sudden loss in the CM-IF-NT model in the late periods of the mean-target analysis indicates weakness of the constant mean, independent factor model; it cannot fully react to a rapid change of market circumstances. In contrast, it is remarkable that CM-IF-LT model avoids such a significant drop due to reasonably flexible shrinkage in the loadings matrix that also reduces uncertainty about the time-varying parameters, resulting in reduced forecast variances as well as improved forecast performance from short-term adaptability.

Additional insights into comparative performance come from examining levels of portfolio risk. From one-step ahead forecasting using the LM-AF models,

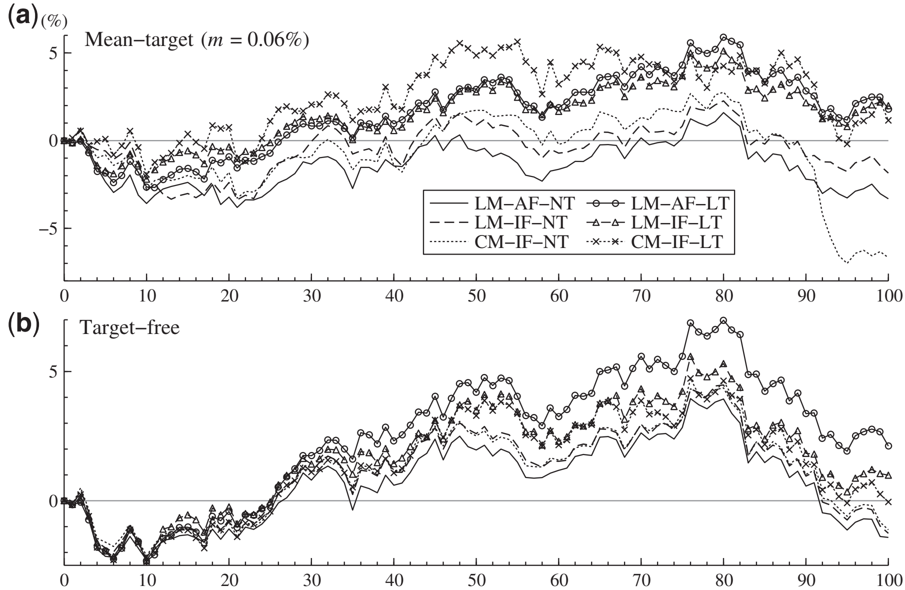

Figure 9 Cumulative returns over 100 business days based on one-step ahead forecasting for (a) portfolio under the target returns of 0.06% and (b) target-free minimum-variance portfolio.

Figure 10 plots *ex ante* portfolio standard deviations $(\boldsymbol{w}_t'\boldsymbol{Q}_t\boldsymbol{w}_t)^{1/2}$ at the optimizing weights $\boldsymbol{w}_t$. LT models clearly yield smaller variances of optimizing portfolios than NT models. The risk ratios based on these portfolio standard deviations for the NT model relative to the LT model are displayed in Figure 10; LT models exhibit roughly 20 percent lower risk in their portfolios. Table 3 reports the average risk ratios from the full analyses, indicating that the portfolio variances are likely smaller in using LT models, and that the ratios tend to be larger in rather more flexible specifications, LM-AF and LM-IF models, than in CM-IF models. Overall, this implies that the LTDFM structure plays a considerably beneficial role in reducing predicted investment risk in the portfolio allocations.

One further aspect of interest is the trajectories of optimized portfolio weights based on one-step ahead and five-step ahead forecasting, displayed in Figures 11 and 12, respectively. In the NT model, the allocation is mainly led by a long position on EUR and a short position on CHF. This reflects quite a high correlation between these two currencies in the estimated forecast variance matrix. By contrast, LT models provide milder degrees of correlation between those two currencies, yielding a smaller proportion of resources allocated to the entire currencies. The lower degree of variation in portfolio weights resulting from the LT models leads to more stable portfolios, which would also reduce transaction costs in reality.

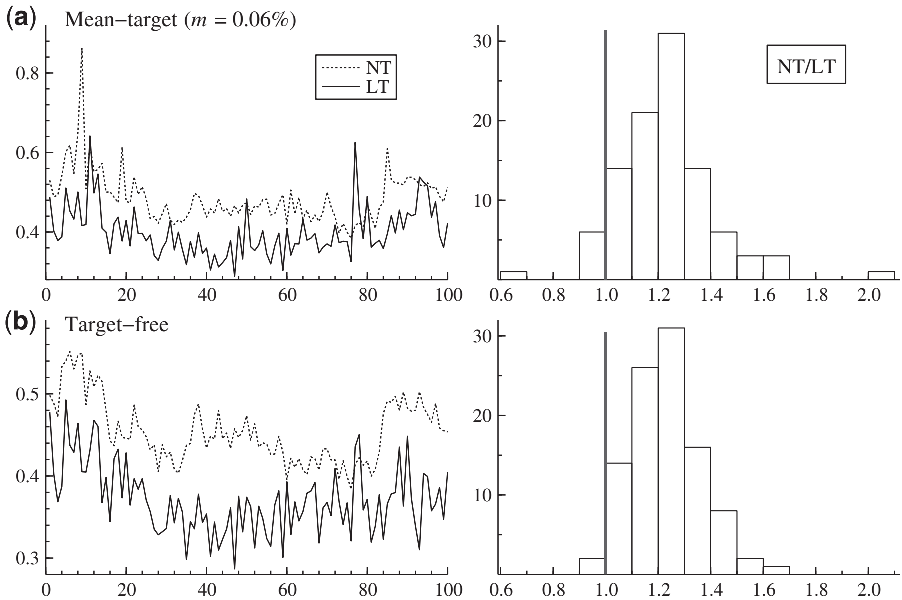

Figure 10 Portfolio risk: *ex ante* portfolio standard deviation, $(w_t' Q_t w_t)^{1/2}$ based on one-step ahead forecasting of LM-AF models for (a) the portfolio under the target return of $m = 0.06\%$ (top), and (b) the target-free minimum-variance portfolio (bottom). Time trajectory (left panels) and histogram of the risk ratio for NT model relative to LT model (right panels). The vertical line in the histogram refers to the ratio of one.

Table 3 Portfolio risk ratio: *ex ante* ratio of portfolio standard deviation under the NT models relative to the LT models, averaged across 100 business days

	LM-AF	LM-IF	CM-IF
One-step ahead forecasts			
Target return:			
$m = 0.02$	1.033	1.007	1.185
0.04	1.149	1.187	1.062
0.06	1.232	1.407	0.969
Target-free	1.234	1.205	1.193
Five-step ahead forecasts			
Target return:			
$m = 0.02$	1.067	1.027	1.178
0.04	1.158	1.102	1.058
0.06	1.223	1.150	0.979
Target-free	1.249	1.224	1.172

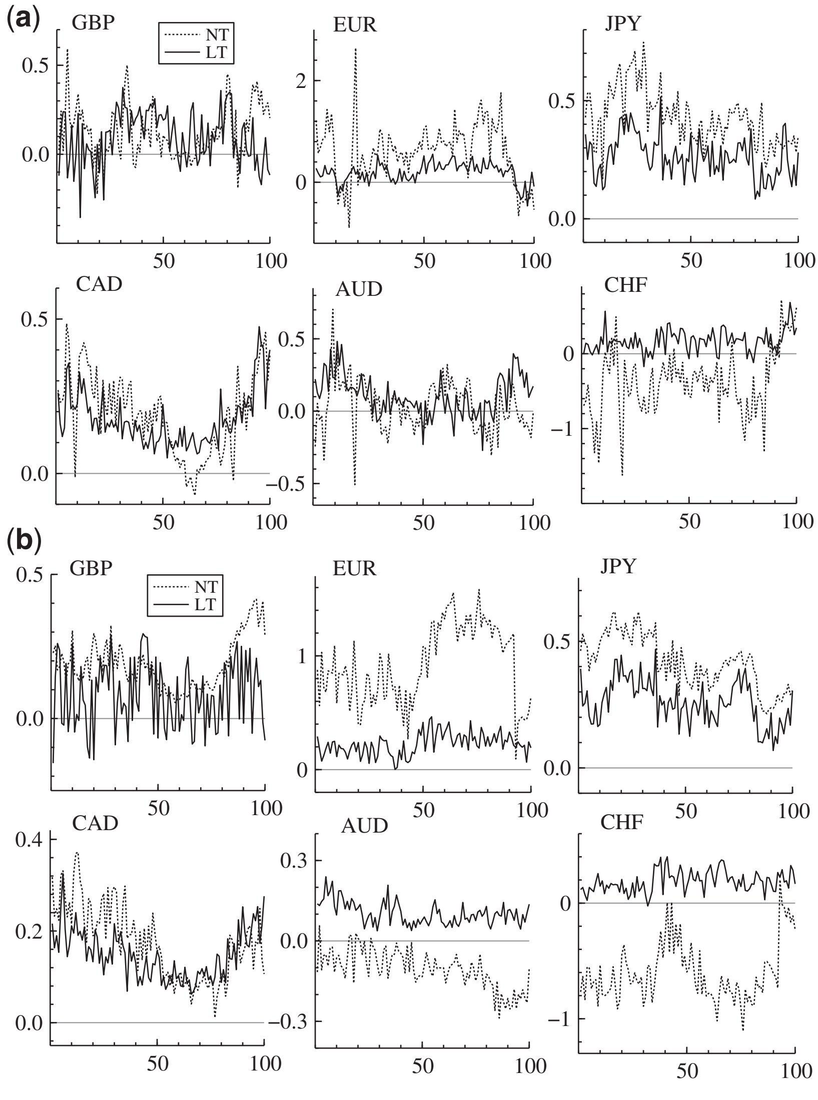

Figure 11 Trajectories of portfolio weights based on one-step ahead forecasts of the LM-AF models for (a) the portfolio under the target return of $m = 0.06\%$ (upper), and (b) the target-free minimum-variance portfolio (lower).

3.3 Diagnostics

The model analysis and mean-variance portfolio strategy rely on a number of modeling assumptions, among which is the assumed conditional normality of the error terms—that for the observation errors $\boldsymbol{\nu}_t$ and for the latent factor innovations $\boldsymbol{\varepsilon}_t$

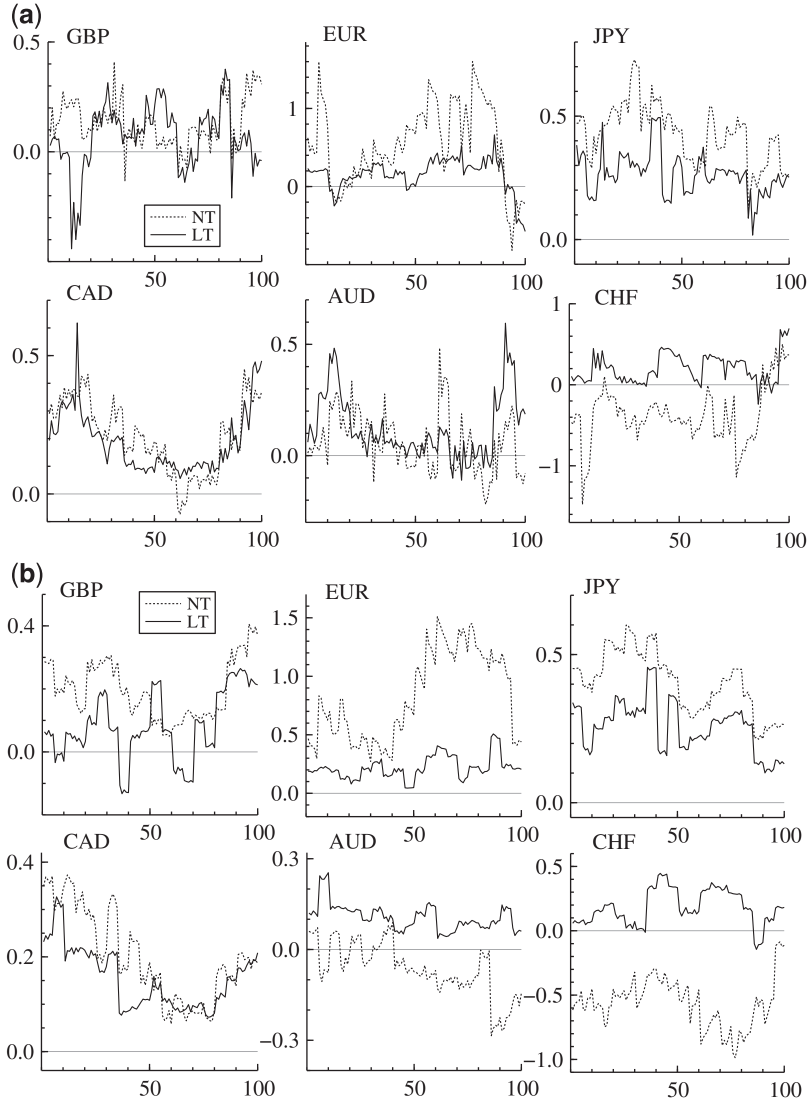

Figure 12 Trajectories of portfolio weights based on five-step ahead forecasts of the LM-AF models for (a) the portfolio under the target return of $m = 0.06\%$ (upper), and (b) the target-free minimum-variance portfolio (lower).

in Equations (1) and (2), respectively. Non-normality would suggest consideration of other forms of utility function in the portfolio strategy, such as power utility functions that go beyond the mean-variance based forms to weigh risks in terms of other distributional aspects.

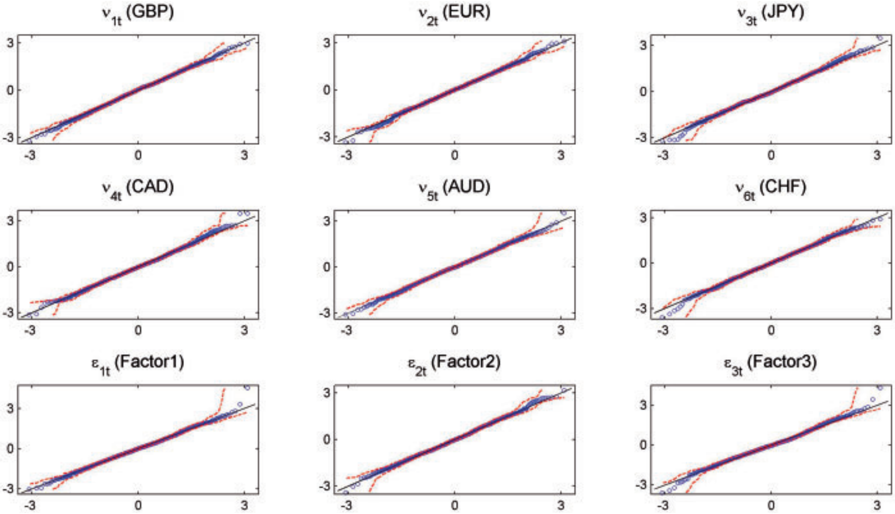

Figure 13 qq-plots for standardized errors $\tilde{\nu}_{it}$ and $\tilde{\varepsilon}_{jt}$ against the standard normal (theoretical quantiles on the horizontal axis) with approximate posterior 99% credible intervals (dashed lines). Here, the residuals used represent just one random selection from the full posterior sample of residuals, as an example of a typical case.

We assess consistency of the data with the assumed normal forms using a number of traditional methods applied to posterior samples of the standardized errors: $\tilde{\nu}_{it} = \nu_{it}/\sigma_{it}$, and $\tilde{\varepsilon}_{jt} = \varepsilon_{jt}/\psi_{jt}$ for $i = 1:m$, $j = 1:k$. At each MCMC iteration we save the current sampled values of each $\tilde{\nu}_{it}$ and $\tilde{\varepsilon}_{jt}$, building up a full posterior sample of all $(m+k) \times T$ error terms. We can then explore and evaluate these using any of a number of normality diagnostics. For this study, we utilized qq-plots and two traditional nonparametric and parametric tests. Figure 13 graphs qq-plots of the standardized errors against the standard normal, with approximate posterior 99 percent credible intervals, using some of the residual values at one randomly selected MCMC iterate. This is very typical; viewing many such plots across the MCMC iterates, we see close concordance with normality as a general rule, with only a few posterior samples observed showing slight deviations from the theoretical quantiles in the tails. In simply viewing hundreds of repeat plots, in no cases do we see significant deviation of sampled residuals outside the intervals. These findings provide initial, exploratory graphical support for the assumed normality.

We make this more formal and quantify the assessment of normality using traditional test statistics, namely the nonparametric Pearson's chi-square goodness-of-fit test and the Shapiro–Wilk parametric test. For each of the two sets of residuals sampled from their posterior at each MCMC iterate, we compute notional p-values for each of these statistics. This yields a posterior distribution for the p-values as one global numerical summary of extent to which the data support the

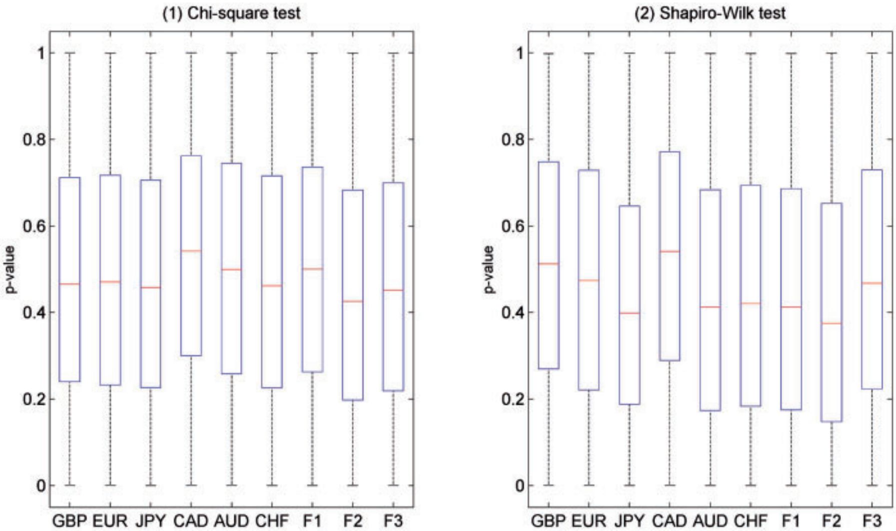

Figure 14 Box-plots for posterior distributions of p-values in the normality tests using full posterior samples of standardized errors $\tilde{\nu}_{it}$ and $\tilde{\varepsilon}_{jt}$: (1) Pearson's chi-square normality test, and (2) Shapiro–Wilk parametric normality test. Fi refers to Factor i.

normality assumed. Figure 14 shows box plots of these posteriors for each series; it is clear that all the distributions are spread and favor larger values, strongly confirming lack of evidence against normality. Importantly, this applies to both the observation residual/error terms ν_{it} and the innovation terms ε_{jt} impacting on the latent factor processes.

This concordance with the normality assumption is not, perhaps, so surprising given that the overall LTDFM allows for volatility patterns overlaying the residuals, and the models have dynamic adaptability in the latent factor components. We comment further on these general questions in summary comments in the final Section 6, simply concluding here that the model assumptions checked seem to be quite reasonable based on these detailed diagnostics.

4 A STUDY OF FX AND COMMODITY PRICE TIME SERIES

A second study extends the FX time series above with three additional currencies as listed in Table 4 and explores analysis now including a dynamic regression component based on commodity prices, oil, and gold. This analysis now directly models the logged prices of currencies in U.S. dollars, rather than the returns analyzed above; that is, $\boldsymbol{y}_t$ is the vector of (natural logs of) exchange rates at 4:30 pm on day t. As covariates, we use prices of Brent Crude Oil futures and Comex Gold

Table 4 Additional international currencies relative to the U.S. dollar in exchange rate data

7	NOK	Norwegian Krone
8	NZD	New Zealand Dollar
9	ZAR	South African Rand

futures. We define $\boldsymbol{x}_t$ as the 2−vector of logged commodity prices on day $t-1$, and y_t as the 9−vector of logged exchange rates on day t. Thus the model has $m=9$ daily currency exchange rates, $q=2$ covariates and again we specify $k=3$ latent factor processes. The time period is $T=125$ business days from the beginning of January to the end of June in 2009, a turbulent period for the financial markets.

We employ the LTDFM with dynamic regression and autoregressive latent factors, taking a special case of constant means ($\boldsymbol{c}_t \equiv \boldsymbol{c}$) in view of the predictive component of the model. Note that both the factor loadings and dynamic regression coefficients are governed by latent threshold VAR(1) process. We use priors as specified in the initial analysis, detailed in Section 2.1 above, now augmented by required priors on the dynamic regression AR hyper-parameters: $1/v_{\alpha i} \sim G(40, 0.005)$, $(\phi_{\alpha i}+1)/2 \sim B(20, 1.5)$, and $\mu_{\alpha i} \sim N(0,1)$ for $i=1:18$.

Figure 15 plots the trajectories of posterior probabilities of $s_{bit}=0$ for the factor loadings. Both global and local sparsity patterns found across the currencies indicate adequately distinguished factors oriented by GBP, EUR, and JPY. Figure 16 shows the posterior means of trajectories of the factors and posterior distributions of the latent factor AR parameters γ_i. The latter indicate values consistent with highly persistent AR(1) processes for the factors. This suggests that modeling exchange rates themselves can yield an utility in short-term forecasting; taking returns from the prices may lose some information about short-term changes in patterns of covariation. Figure 17 graphs posterior estimates of the dynamic regression coefficients $\boldsymbol{\alpha}_t$ and the posterior probabilities of $s_{ait}=0$. For several currencies, the posterior means of the α_{it} are positive throughout the sample period; consistent with context, positive movements in oil and gold tend to predict marginal appreciation of these currencies against the U.S. dollar, all other things being equal. The inferences further indicate effective whole-sequence sparsity for the major currencies GBP, EUR, and JPY; the dynamic regressions on commodities are playing no role in short-term prediction of these prices, with inferred coefficients zero or close to zero across the entire time frame. This indicates that there is little room for commodity prices to explain meaningful residual changes in these currencies in the context of the dynamic factors estimated, which is further interpreted by examining inferences on the factors themselves. Figure 16 plots the trajectories of the commodity prices as well as estimated factors. We see that the posterior mean trajectories of Factor1 and Factor2 are primarily correlated with the fluctuations of oil price; in fact these correlations are high as 0.92 for Factor1 and 0.84 for Factor2. The correlation between the factors and gold prices is relatively mild, around 0.2.

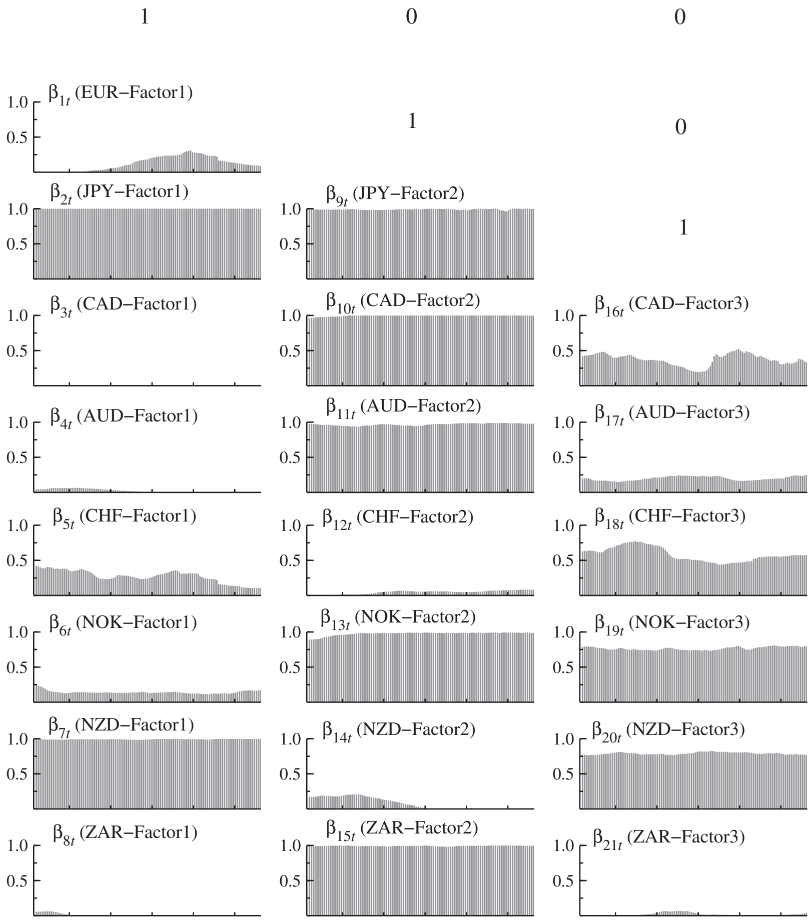

Figure 15 Posterior probability of $s_{bit}=0$ associated with the loadings for exchange rate and commodity data.

The GBP and EUR leading factors essentially capture the predictive information in oil price returns and, as a result, the latent threshold mechanism shrinks dynamic regression effects very substantially.

Finally, we provide some comments on a further, detailed comparison using portfolio analysis, focused on comparing NT and LT models. The portfolio allocation rules and estimation strategy are as in Section 3. One-day ahead forecasts and the optimal portfolios are computed over 25 business days beginning with the first $T=100$ observations. We assume that the commodity returns in $\boldsymbol{x}_{n+1}$ are available when we forecast the one-day ahead exchange rate $\boldsymbol{y}_{n+1}$. Figure 18 graphs the resulting cumulative returns from this using both the NT and LT models; here

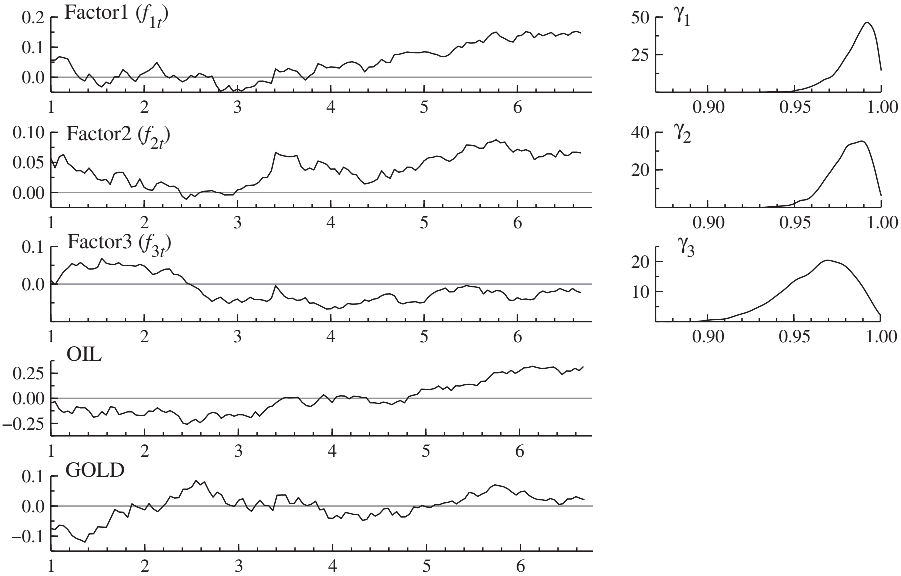

Figure 16 Left panels: posterior means of factors f_{it} and log of commodity prices (x-axis refers to the month in 2009). Right panels: estimated posterior distribution of the γ_i.

we set a target return of $m=0.06$ percent and compared also based on the target-free minimum-variance strategy. It is evident that the LT model dominates the NT model for both the strategies. The most striking evidence is that the optimal weights allocated to the 10 FX currencies are considerably shrunk towards zero in the LT model as can be seen in Figure 19. This effective shrinkage yields higher performance of the portfolios under the LTDFM structure.

5 A HIGHER-DIMENSIONAL STUDY: FX AND STOCK PRICE TIME SERIES

A third data set serves to demonstrate a substantial, higher-dimensional application. We analyze $m=40$ financial variables: daily FX time series of 20 currencies relative to the U.S. dollar and stock price indices from 20 countries; see Table 5. The returns are computed over a time period of 1000 business days beginning in January 2008 and ending in October 2011. This data set includes severely high volatility periods of the financial crisis around 2008 and the European sovereign debt crisis from 2010 to 2011. There exist similar volatility dynamics across countries as well as several regionally or industrially common fluctuations. Exploratory analysis suggests an initial model using $k=7$ factors, ordering the first seven variables in $\boldsymbol{y}_t$ as GBP, U.S.-stock, EUR, JPY, IDR, Germany-stock, and Brazil-stock. In addition to GBP, EUR, and JPY used in the previous analysis, IDR (Indonesian Rupiah)

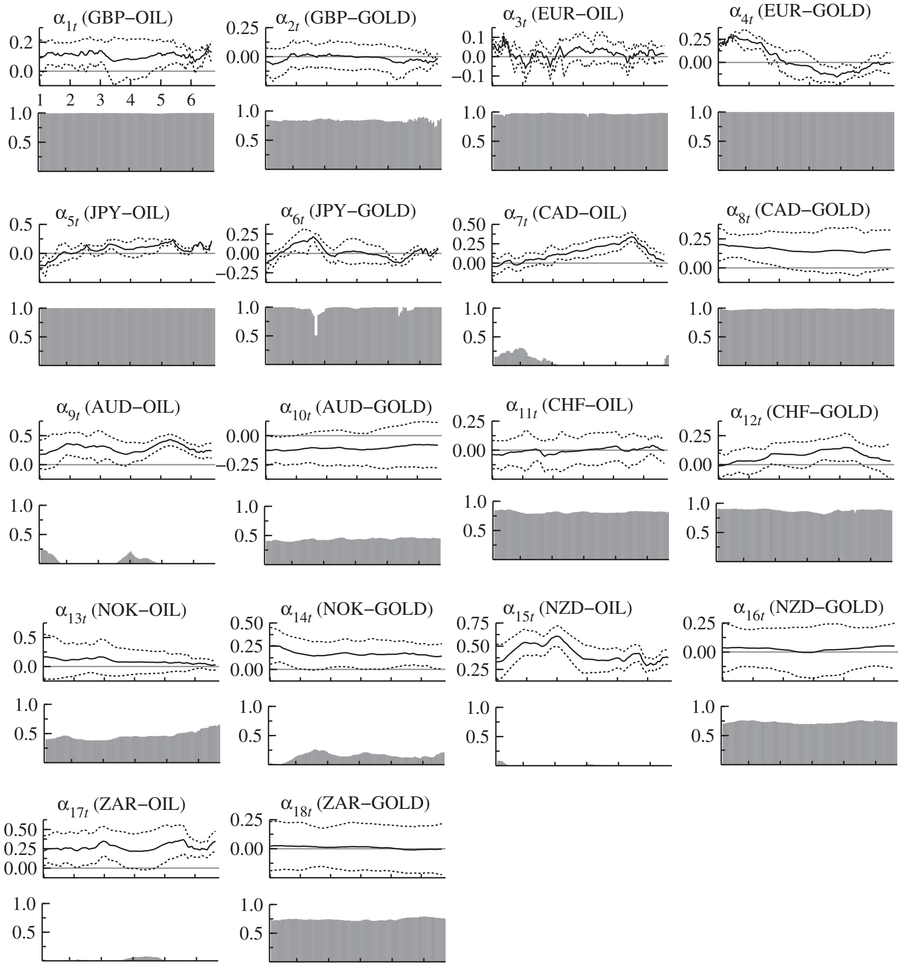

Figure 17 Posterior means (solid) and 95% credible intervals (dotted) of time-varying regression coefficients $\boldsymbol{\alpha}_t$ for exchange rate and commodity data. Posterior probabilities of $s_{ait} = 0$ are plotted below each trajectory.

is selected to capture exchange rate flows of emerging countries against the U.S. dollar. U.S.-stock is intended to represent a global stock price measure, Germany-stock describes common stock movements of European countries, and Brazil-stock traces stock markets in the emerging countries. The analysis uses the same LTDFM specification and priors as in the first example of Section 2.

Figure 20 plots the trajectories of estimated latent factors and their innovation volatilities. Factor2 and Factor6 have levels of volatility higher than other factors as they are led by stock prices of United States and Germany, respectively, and the

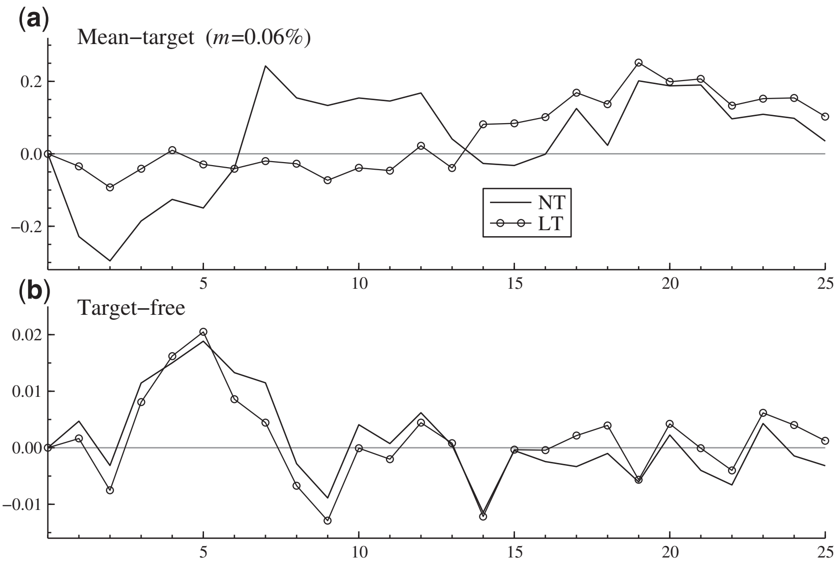

Figure 18 Cumulative returns over 25 business days based on one-step ahead forecasting for (a) portfolio under the target returns of 0.06% and (b) target-free minimum-variance portfolio using exchange rate and commodity data.

latter exhibit higher levels of variability than the exchange rate series. As observed in the first example, all the factor stochastic volatilities have peaks around the last 3 months of 2008 due to the financial crisis. The volatilities of GBP-leading Factor1 and JPY-leading Factor4 decline after that crisis to the end of the sample, while the EUR-leading Factor3 shows considerable increases in volatility around mid-2010 and during 2011 corresponding to the European sovereign debt crisis triggered by the Greek government debt crisis arising in April 2010. At that time, the U.S.-stock-leading Factor2, IDR-leading Factor5, and Germany-stock-leading Factor6 commonly capture a sudden hike in volatility. These factors also exhibit another significantly volatile period beginning in August 2011 when Standard & Poor's cut the long-term U.S. credit rating from triple-A to AA-plus.

Figure 21 shows the trajectories of estimated posterior probabilities of $s_{bit}=0$ for selected loadings. The first row indicates that the U.S. stock market (Factor2) is relevant to describing the Brazilian stock market in addition to the leading-factor itself (Factor7). A drop of shrinkage probability on Brazil-Factor5 from the beginning of 2009 presumably reflects the fact that exchange rate fluctuations in emerging countries led by IDR affect some proportion of Brazilian stock market volatility. Several factors associated with NZD (in the second row of the figure) show clearly time-varying shrinkage patterns; in contrast, the model suggests the Spain stock price index (in the third row) is primarily explained by the Germany-stock-leading Factor6. The shrinkage pattern of the India stock price index (in

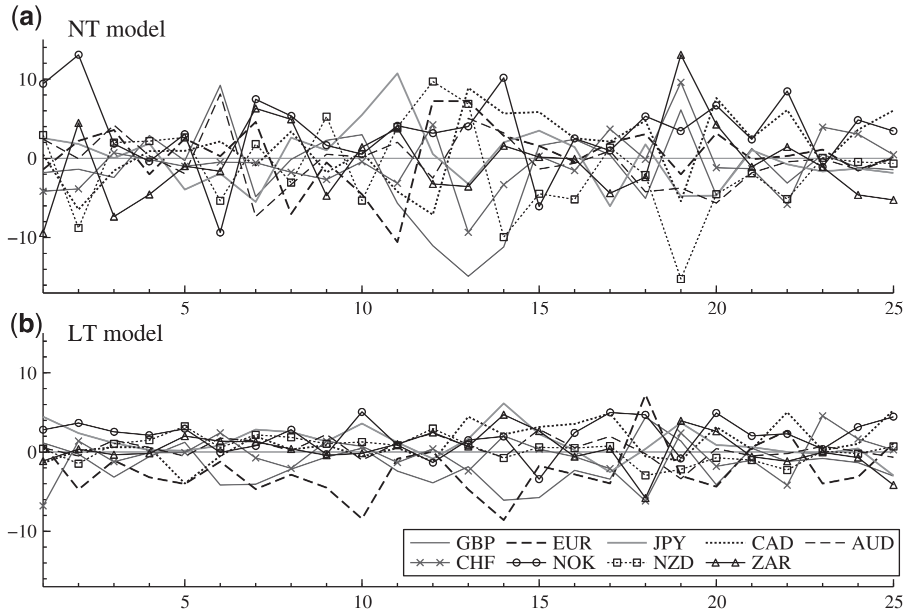

Figure 19 Trajectories of portfolio weights based on the portfolio under the target return of $m=0.06\%$ for exchange rate and commodity data.

the fourth row) exhibits some temporal changes on Factor2 and Factor5, while German-stock Factor6 and Brazil-stock Factor7 are evidently relevant in explaining the volatility dynamics. A variance decomposition analysis (not graphed here) also indicates that each factor evidently well identifies the characteristics of commonly observed fluctuations across the high-dimensional financial responses. In addition, Figure 22 graphs a heat map of posterior probabilities of $s_{bit}=0$ for all i and t. To make the image clear, the rows corresponding to factor loadings have been ordered to highlight patterns of difference. Red (black, in grey-scale version) areas show high levels of LT shrinkage, while blue (white) areas imply relevant factor loadings beyond the threshold. There exist red (black) bands indicating whole-sequence shrinkage as well as some patterns of temporal shrinkage where the color is clearly changing through time.

In line with the preceding examples, further analysis evaluated the LTDFM models based on forecasting in portfolio allocation experiments. Forecast means and covariance matrices of the 40 dimensional variables from one-step to five-step ahead were computed every five business days with the first subset of data up to time $n=950$. The optimal weights of the portfolio strategy were obtained for target returns of $m=0.02$, 0.04, and 0.06 percent, as well as the target-free minimum variance strategy. Table 6 reports cumulative returns of the portfolios resulting from the sequential investment over 50 business days. Evidently, the LT structure leads to

Table 5 20 international currencies for FX and 20 countries for stock price index

FX			Stock price	
1	GBP	British Pound Sterling	1	United States
2	EUR	Euro	2	Germany
3	JPY	Japanese Yen	3	Brazil
4	IDR	Indonesian Rupiah	4	Canada
5	CAD	Canadian Dollar	5	Mexico
6	AUD	Australian Dollar	6	UK
7	NZD	New Zealand Dollar	7	France
8	CHF	Swiss Franc	8	Spain
9	NOK	Norwegian Krone	9	Italy
10	SEK	Swedish Krona	10	The Netherlands
11	RUB	Russian Ruble	11	Sweden
12	INR	Indian Rupee	12	Swiss
13	PHP	Philippine Peso	13	Japan
14	SGD	Singapore Dollar	14	China
15	KRW	South Korean Won	15	Hong Kong
16	TWD	Taiwanese Dollar	16	Taiwan
17	THB	Thai Baht	17	Korea
18	ZAR	South African Rand	18	India
19	BRL	Brazilian Real	19	Russia
20	CLP	Chilean Peso	20	Australia

outcomes that dominate those of NT models regardless of the portfolio allocation rules. The table also reports average transaction volume in weights per day for each model; this indicates that the LT models entail lower levels of transaction than the NT models. The investor benefit is essentially cumulative return minus transaction cost, so that this is a very clear additional attraction of the threshold model approach for increasingly high-dimensional problems: the sparsity induced by the LT structure more effectively decreases overall uncertainties in the portfolio decision-making process as dimension grows.

6 CONCLUDING REMARKS

We have presented and illustrated the use of latent threshold modeling in dynamic factor volatility models. Our substantive examples assess comparative model fit and sequential portfolio allocation analysis using FX return time series, and reveal considerable utility of latent thresholding as an overlay to time-varying parameter models for latent factor loadings. Full shrinkage to zero of subsets of parameters when supported by the data leads to parameter reduction and model simplification, and can feed through to both improved short-term forecasting as

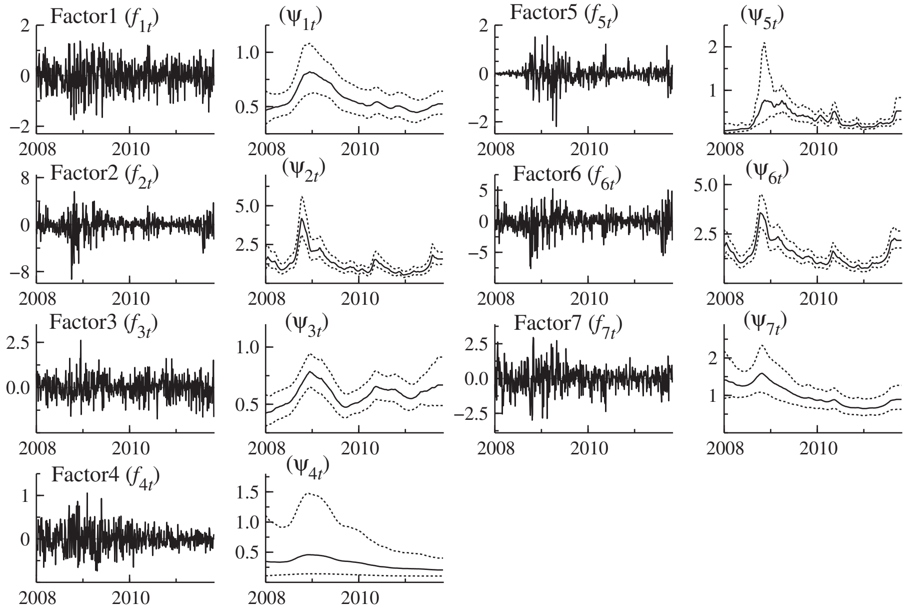

Figure 20 Posterior means of factors f_{it}, posterior means (solid) and 95% credible intervals (dotted) of factor stochastic volatility $\psi_{it} = \exp(\lambda_{it}/2)$ for FX and stock price data.

Table 6 Cumulative returns (%) and average transaction (in weights per day) over 50 business days for FX and stock price data

	NT	LT
Cumulative returns		
Target return:		
$m=0.02$	2.20	2.50
0.04	2.03	2.29
0.06	1.86	2.08
Target-free	2.39	2.73
Average transaction		
Target return: $m=0.02$	0.27	0.18
0.04	0.38	0.25
0.06	0.51	0.36
Target-free	0.16	0.15

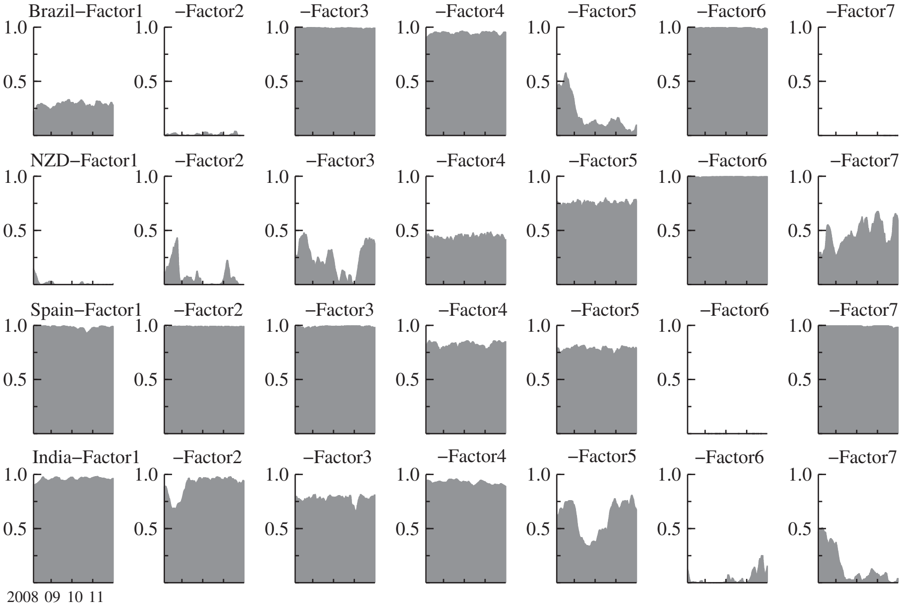

Figure 21 Posterior probability of $s_{bit} = 0$ associated with selected loadings for FX and stock price data. Brazil-stock, New Zealand Dollar FX, Spain-stock, and India-stock.

well as increased precision of both parameter inferences and forecasts. Dynamic shrinkage–that allows parameter processes to take non-zero, time-varying values over some periods of time but shrink to zero completely in others–is a natural extension of traditional Bayesian sparsity modeling to the time series context, and can engender these benefits adaptively over time.

There are a number of methodological and computational areas for further investigation. Among them, we note the potential for more elaborate factor models, such as FAVAR (Factor-Augmented VAR) models (e.g., Bernanke, Boivin, and Eliasz 2005; Baumeister, Liu, and Mumtaz 2010) where the latent thresholding strategy can be expected to be increasingly beneficial as model dimension grows. We are also interested in potential computational strategies for sequential model learning, including sequential particle learning algorithms (e.g., Carvalho et al. 2010) for LTDFMs, as implementation of sequential updating and forecasting is central to investment decision making. In such contexts, the computations involved in repeat MCMC analyses with a moving time window may become prohibitive and sequential Monte Carlo methods become increasingly relevant.

That said, our examples show that scaling of the MCMC to moderately high dimensional series will be accessible. The first example in Section 2 uses $m = 6$ series and $k = 3$ factors (with $p = 12$ loadings), while the $m = 40$ dimensional example in Section 5 has $k = 7$ factors ($p = 252$), with the time length almost equal. Regarding

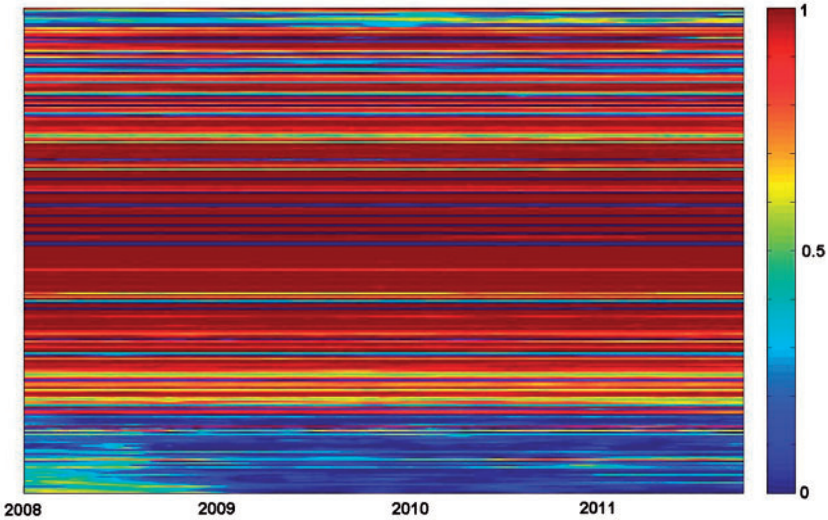

Figure 22 Heat map of posterior probabilities of $s_{bit}=0$ for FX and stock price data. The rows correspond to the full set of factor loadings, each plotted over time; the ordering of the rows has been chosen simply to provide a clear visual layout of those loadings that tend to have high probability over all time compared to those that are negligible for much of the time.

computational time, using a dual-core 3.3GHz CPU computer, the first example takes 0.3 second per one iteration of MCMC, while this high-dimensional setting takes only 3.7 seconds. That is, compute time increases just over 12-fold in moving to the larger mode context, whereas the increase in the number of loadings is 21-fold. We have not formally benchmarked performance in scaling to increasingly high-dimensional settings, but this first experiment in scaling is encouraging. Moving to higher-dimensions (hundreds of series) will require different coding and computational strategies. For both MCMC and particle learning methods, scaling will have to explore parallelization. Distributing core computational tasks within each iterate of the demanding MCMC is one immediate consideration, and most sequential Monte Carlo methods are naturally parallelizable. Future studies will investigate the opportunities for parallelization that can exploit CUDA/GPU (Suchard et al. 2010; Suchard, Holmes, and West 2010) implementations for massively parallel desktop analyses as well as traditional multi-core approaches.

Among a number of potential model extensions and generalizations are models in which the normal distributions of observation error terms and innovations might be better represented as non-normal. With reference to the discussion of Section 3.3, some potential applied contexts may suggest, on a theoretical and/or empirical basis, heavier-tails or some degree of skewness. A very natural way to elaborate the current model to adapt to this is to introduce discrete mixtures of normals in place of the current normal conditional error forms. The Bayesian MCMC analysis is then easily adapted to allow for this (e.g., Frühwirth-Schnatter 2006; Prado and West 2010, ch. 7). That understood, in such contexts it can be argued that the traditional mean-variance optimization strategy for portfolio decisions might be replaced by approaches using utility functions that reflect risk measured in terms of tail behavior as well as variance, as noted in Section 3.3. Thus, exploration of decision analyses with power utilities and/or value-at-risk focused studies will respresent interesting and important further directions. The issues of choice of utility functions and how the resulting expected utilities capture different notions of risk is in fact subtler than this; in any model in which fully Bayesian inference is properly developed, the full posterior predictive distributions will be non-normal. Even with normality of the *conditional* error and innovations distributions, as in the current study, the posterior predictive distributions are not normal. Predictions average over all uncertain factors and parameters, so these distributions are inherently non-normal and heavy-tailed as a matter of routine. This raises the interesting additional considerations for other utility functions as a general matter, beyond the scope of the current study but certainly of interest in future work.

REFERENCES

Aguilar, O., R. Prado, G. Huerta, and M. West. 1999. "Bayesian Inference on Latent Structure in Time Series (with discussion)." In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, (eds.), *Bayesian Statistics 6*. Oxford: Oxford University Press, 3–26.

Aguilar, O., and M. West. 2000. Bayesian Dynamic Factor Models and Portfolio Allocation. *Journal of Business and Economic Statistics* 18: 338–357.

Baumeister, C., P. Liu, and H. Mumtaz. 2010. "Changes in the Transmission of Monetary Policy: Evidence from a Time-Varying Factor-Augmented VAR." Working paper No.401, Bank of England.

Bernanke, B., J. Boivin, and P. Eliasz. 2005. Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *Quarterly Journal of Economics* 120: 387–422.

Carvalho, C. M., J. Chang, J. E. Lucas, J. R. Nevins, Q. Wang, and M. West. 2008. High-Dimensional Sparse Factor Modeling: Applications in Gene Expression Genomics. *Journal of the American Statistical Association* 103: 1438–1456.

Carvalho, C. M., M. S. Johannes, H. F. Lopes, and N. G. Polson. 2010. Particle Learning and Smoothing. *Statistical Science* 25: 88–106.

Carvalho, C. M., H. F. Lopes, and O. Aguilar. 2011. "Dynamic Stock Selection Strategies: A Structured Factor Model Framework (with discussion)." In J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, (eds.), *Bayesian Statistics 9*. Oxford: Oxford University Press, 69–90.

Carvalho, C. M., and M. West. 2007. Dynamic Matrix-variate Graphical Models. *Bayesian Analysis* 2: 69–98.

Doornik, J.A. 2006. *Ox: Object Oriented Matrix Programming*. London: Timberlake Consultants Press.

Frühwirth-Schnatter, S. 2006. *Finite Mixture and Markov Switching Models*. New York: Springer-Verlag.

Jacquier, E., N. G. Polson, and P. E. Rossi. 1994. Bayesian Analysis of Stochastic Volatility Models. *Journal of Business and Economic Statistics* 12: 371–389.

Kim, S., N. Shephard, and S. Chib. 1998. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies* 65: 361–393.

Lopes, H. F., and C. M. Carvalho. 2007. Factor Stochastic Volatility with Time Varying Loadings and Markov Switching Regimes. *Journal of Statistical Planning and Inference* 137: 3082–3091.

Lopes, H. F., and M. West. 2004. Bayesian Model Assessment in Factor Analysis. *Statistica Sinica* 14: 41–67.

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

Nakajima, J., and M. West. 2010. "Bayesian Analysis of Latent Threshold Dynamic Models." Working paper 10-25, Department of Statistical Science, Duke University.

Omori, Y., S. Chib, N. Shephard, and J. Nakajima. 2007. Stochastic Volatility with Leverage: Fast Likelihood Inference. *Journal of Econometrics* 140: 425–449.

Pitt, M., and N. Shephard. 1999. "Time Varying Covariances: A Factor Stochastic Volatility Approach (with discussion)." In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, (eds.), *Bayesian Statistics VI*. Oxford: Oxford University Press, 547–570.

Prado, R., and M. West. 2010. *Time Series Modeling, Computation, and Inference*. New York: Chapman & Hall/CRC.

Putnum, B. H., and J. M. Quintana. 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, 232–237.

Quintana, J. M. 1992. "Optimal Portfolios and Forward Currency Contracts." In J. O. Berger, J. M. Bernardo, A. P. Dawid, and A. F. M. Smith, (eds.), *Bayesian Statistics IV*. Oxford: Oxford University Press, 753–762.

Quintana, J. M., and B. H. Putnum. 1996. "Debating Currency Markets Efficiency Using Dynamic Multiple-factor Models." In *Proceedings of the ASA Section on Bayesian Statistical Science*. Alexandria, VA, 55–60.

Shephard, N., and M. Pitt. 1997. Likelihood Analysis of Non-Gaussian Measurement Time Series. *Biometrika* 84: 653–667.

Suchard, M. A., C. Holmes, and M. West. 2010. Some of the *What?, Why?, How?, Who?* and *Where?* of Graphics Processing Unit Computing for Bayesian Analysis. *Bulletin of the International Society for Bayesian Analysis* 17: 12–16.

Suchard, M. A., Q. Wang, C. Chan, J. Frelinger, A. J. Cron, and M. West. 2010. Understanding GPU Programming for Statistical Computation: Studies in

Massively Parallel Massive Mixtures. *Journal of Computational and Graphical Statistics* 19: 419–438.

Wang, H. 2010. Sparse Seemingly Unrelated Regression Modelling: Applications in Finance and Econometrics. *Computational Statistics and Data Analysis* 54: 2866–2877.

Wang, H., and M. West. 2009. Bayesian Analysis of Matrix Normal Graphical Models. *Biometrika* 96: 821–834.

Watanabe, T., and Y. Omori. 2004. A Multi-move Sampler for Estimating Non-Gaussian Time Series Models: Comments on Shephard & Pitt (1997). *Biometrika* 91: 246–248.

West, M. 2003. "Bayesian Factor Regression Models in the "Large p, Small n" Paradigm." In J.M. Bernardo, M.J. Bayarri, J.O. Berger, A.P. David, D. Heckerman, A.F.M. Smith, and M. West, (eds.), *Bayesian Statistics 7*. Oxford: Oxford University Press, 723–732.

Yoshida, R., and M. West. 2010. Bayesian Learning in Sparse Graphical Factor Models via Annealed Entropy. *Journal of Machine Learning Research* 11: 1771–1798.