A Flexible State Space Model and its Applications

Hang Qian[1]

Abstract

The standard state space model treats observations as imprecise measurement of the Markovian states. Our flexible model handles the states and observations symmetrically, which are simultaneously determined by past observations and up to first-lagged states. The only distinction between the states and observations is the observability. When it is applied to the ARMA, dynamic factor and stochastic volatility models, the state space form is both parsimonious and intuitive, for low-dimension states are constructed simply by stacking all the relevant but unobserved components in the structural model.

Keywords: Kalman Filter, ARMA, Factor Model, Stochastic Volatility

1. Introduction

Starting with the path-breaking paper of Kalman (1960), the state space model (SSM) has been widely applied in engineering, statistics and economics. Harvey (1991), Hamilton (1994), Durbin and Koopman (2012) present its theory and applications in time series analysis. Basdevant (2003) surveys macroeconomic applications and Mergner (2009) reviews use cases

[1]I thank the referee for helpful comments to improve the paper. Corresponding author: 3 Apple Hill Drive, Natick, MA 01760. Email: matlabist@gmail.com

in finance. For practitioners, the art consists in the model building, that is, to cast a structural model into its state space form. The representation is not unique, for one can enlarge the state vector but characterize the same process. The major concern is parsimony and intuitiveness. A parsimonious model with the minimum length of the state vector avoids large matrix manipulations, and thus accelerates the Kalman filter. An intuitive form with interpretable states enhances its attractiveness, for predicted and smoothed states bear economic significance.

States usually refer to the unobserved variables with Markovian transition. Observations are imprecise measurement of the states. Since states never echo observations, it is a one-way dependence. Our argument is that the Kalman filter does not necessarily require Markovian states. The recursion is valid as long as no higher than first-lagged states are in the dynamic system, without restrictions on how lagged observations affect current states and observations. That motivates us to bring in more symmetry and two-way dynamics between the states and observations. This feature is most useful when it is combined with the time-varying dimension (TVD) of the state and measurement vectors. The idea of building a flexible SSM is to put all the relevant but unobserved components in the state vector. Since observability of a variable may change over time, the size of the state vector is also dynamic. This often invites lower-dimension states compared with a standard SSM. Furthermore, states in the flexible SSM can always be meaningfully interpreted, for they are simply the unobserved components in the original econometric model.

The rest of the paper is organized as follows. Section 2 sets up the flexible

SSM and Section 3 explains the filtering procedure. Section 4 illustrates the TVD feature by an ARMA model in which the state vector shrinks in the initial periods. Section 5 considers a dynamic factor model with missing data. Our SSM is different from that in the literature and has lower-dimension states. Section 6 discusses a stochastic volatility model in which asymmetric volatility is achieved by putting a non-linear function of past observation in the transition equation, in contrast with the traditional approach that uses correlated disturbances.

2. A flexible state space model

First consider a standard SSM. Let $\boldsymbol{\xi}_t$ be an $m \times 1$ state vector and $\boldsymbol{Y}_t$ be an $n \times 1$ measurement vector. The dynamic system consists of the transition and measurement equation

$$\begin{aligned} \boldsymbol{\xi}_t &= \boldsymbol{c}_t + \boldsymbol{F}_t \boldsymbol{\xi}_{t-1} + \boldsymbol{\varepsilon}_t, \\ \boldsymbol{Y}_t &= \boldsymbol{d}_t + \boldsymbol{H}_t \boldsymbol{\xi}_t + \boldsymbol{u}_t, \end{aligned} \tag{1}$$

where $\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{u}_t \end{pmatrix} \sim N\left[\boldsymbol{0}, \begin{pmatrix} \boldsymbol{Q}_t & \boldsymbol{S}_t \\ \boldsymbol{S}_t' & \boldsymbol{R}_t \end{pmatrix}\right]$. Coefficients $\boldsymbol{c}_t, \boldsymbol{F}_t, \boldsymbol{d}_t, \boldsymbol{H}_t, \boldsymbol{Q}_t, \boldsymbol{R}_t, \boldsymbol{S}_t$ can be time-varying but deterministic. The system starts from time 1 and runs through time T with the observations $\boldsymbol{Y}_1^T \equiv \{\boldsymbol{Y}_1, ..., \boldsymbol{Y}_T\}$, which is the information set at time T. The initial state vector $\boldsymbol{\xi}_0$ has a known distribution.

The flexible SSM is a moderate generalization of the standard model. Let $\boldsymbol{\xi}_t$ be an $m_t \times 1$ state vector and $\boldsymbol{Y}_t$ be an $n_t \times 1$ measurement vector. They are simultaneously determined by lagged observations and up to first-lagged states:

$$\begin{aligned}
\boldsymbol{\xi}_t &= f_t\left(\boldsymbol{Y}_1^{t-1}\right) + \boldsymbol{F}_t\boldsymbol{\xi}_{t-1} + \boldsymbol{\varepsilon}_t, \\
\boldsymbol{Y}_t &= g_t\left(\boldsymbol{Y}_1^{t-1}\right) + \boldsymbol{H}_t\boldsymbol{\xi}_t + \boldsymbol{J}_t\boldsymbol{\xi}_{t-1} + \boldsymbol{u}_t.
\end{aligned} \tag{2}$$

where $f_t(\cdot), g_t(\cdot)$ are two linear or non-linear functions that maps the information set of time $t-1$ into $\mathbb{R}^{m_t}$ and $\mathbb{R}^{n_t}$ respectively. In some applications of the flexible SSM, the contemporaneous correlation between $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{u}_t$ is essential. The flexible model has two features.

First, both the state and measurement vectors can change size over time. The TVD of $\boldsymbol{Y}_t$ is well understood and implemented in practice. For example, if some elements of $\boldsymbol{Y}_t$ are missing, the size of the measurement vector is effectively reduced at time t. If all data are missing in that period, the updating step of the Kalman filter is skipped (see Jones, 1980; Harvey and Pierse, 1984). The TVD of $\boldsymbol{\xi}_t$ had been under-appreciated in the literature until recently. Jungbacker et al. (2011) consider a dynamic factor model with missing data. Common factors and idiosyncratic disturbances corresponding to missing data are put in the state vector. Since the amount of missing data varies over time, the state vector is dynamic in size. Chan et al. (2011) explore the TVD in a different setting. The model switches to a more parsimonious representation at random dates controlled by hidden Markov-switching regimes. This is a dynamic mixture model with stochastic TVD. Our paper is closer to Jungbacker et al. (2011) in that the dimension changes at known dates.

Second, historical observations $\boldsymbol{Y}_{t-1}, ..., \boldsymbol{Y}_1$ can affect both $\boldsymbol{Y}_t$ and $\boldsymbol{\xi}_t$. The dependence of $\boldsymbol{Y}_t$ on $\boldsymbol{Y}_{t-1}, ..., \boldsymbol{Y}_1$ is well understood. The setup of the

SSM in Hamilton (1994, p.372 - 373) includes an $\boldsymbol{A}'\boldsymbol{x}_t$ term in the measurement equation. Hamilton mentions "$\boldsymbol{x}_t$ could include lagged values of $\boldsymbol{y}$ ...", though no application of such feature is provided in the book. In fact, lagged variables in the measurement equation are most useful when they are used together with the TVD feature. Suppose we write $g_t(\cdot)$ as a function of p lagged values $\boldsymbol{Y}_{t-1}, ..., \boldsymbol{Y}_{t-p}$, we will encounter a problem of handling the presample since $\boldsymbol{Y}_0, ..., \boldsymbol{Y}_{-p+1}$ are not observed. With the TVD feature, unobserved lagged variables can be temporarily put in the state vector and then removed when data become available.

Allowing lagged observations in the transition equation is rarely seen in the literature. Some may argue that the modeling philosophy of the SSM is to keep the state vector Markovian – summarizing the entire history into the states of yesterday. This argument is not entirely relevant for our model, for we never introduce high-order lagged states $\boldsymbol{\xi}_{t-2}, \boldsymbol{\xi}_{t-3}, ...$ in the system, but only allow past observations in the transition equation. In the Kalman filter, $\boldsymbol{\xi}_t$ is predicted and updated conditional on $\boldsymbol{Y}_1^{t-1}$. Technically, introducing $f_t\left(\boldsymbol{Y}_1^{t-1}\right)$ does not change the filter since it is treated as a constant conditional on $\boldsymbol{Y}_1^{t-1}$. However, this feature substantially enriches the dependence structure of the SSM. In the standard SSM, $\boldsymbol{\xi}_t$ evolves regardless of $\boldsymbol{Y}_t$. If we cast a time series model into Eqs. (1), we must ensure the state vector can evolve in a self-sufficient manner. This often entails a larger state vector by including variables that we do observe. However, in the flexible SSM the state vector may temporarily disappear, but reappear later relying on $f_t\left(\boldsymbol{Y}_1^{t-1}\right)$.

If we put the flexible SSM in the control system framework and replace disturbances $\boldsymbol{\varepsilon}_t, \boldsymbol{u}_t$ by the control variable $\boldsymbol{v}_t$, it is interesting to consider the implication for the observability and controllability of the system. For simplicity, consider time-invariant coefficients $\boldsymbol{F}, \boldsymbol{H}, \boldsymbol{J}$. Observability is the inference of initial states based on observations. Since $f_t\left(\boldsymbol{Y}_1^{t-1}\right)$ and $g_t\left(\boldsymbol{Y}_1^{t-1}\right)$ are observed, they play no role in the observability condition, which indeed requires the rank of $\left(\boldsymbol{A}_1', \ldots \boldsymbol{A}_m'\right)', \boldsymbol{A}_j = \boldsymbol{H}\boldsymbol{F}^j + \boldsymbol{J}\boldsymbol{F}^{j-1}$ equals m. Controllability is the determination of future states using the initial states and the control sequence. Unlike the standard SSM where the controllability condition is determined solely by the coefficients of the transition equation, the controllability condition under the flexible SSM involves coefficients of both transition and measurement equations due to the presence of $f_t\left(\boldsymbol{Y}_1^{t-1}\right)$ and $g_t\left(\boldsymbol{Y}_1^{t-1}\right)$. If they are linear in $\boldsymbol{Y}_1^{t-1}$, by backward substitution of Eqs. (2) we can rewrite $\boldsymbol{\xi}_t$ as a linear function of $\boldsymbol{\xi}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_t$. The controllability requires the concatenated matrix of coefficients corresponding to $v_1, \ldots, v_t$ has the rank m.

3. The filtering procedure

The forward recursion consists of prediction and update steps. Assume $\boldsymbol{\xi}_0 \sim N\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0\right)$. Since $\boldsymbol{Y}_1^0$ is empty, $\boldsymbol{\xi}_0 \left| \boldsymbol{Y}_1^0 \sim N\left(\widehat{\boldsymbol{\xi}}_{0|0}, \boldsymbol{P}_{0|0}\right)\right.$, where $\widehat{\boldsymbol{\xi}}_{0|0} = \boldsymbol{\mu}_0$, $\boldsymbol{P}_{0|0} = \boldsymbol{\Sigma}_0$.

At time t ($t = 1, ..., T$), we predict $\boldsymbol{\xi}_t$ and $\boldsymbol{Y}_t$ conditional on the information set of time $t-1$.

$$\left(\begin{array}{c} \boldsymbol{\xi}_t \\ \boldsymbol{Y}_t \end{array}\right) \left| \boldsymbol{Y}_1^{t-1} \sim N\left[\left(\begin{array}{c} \widehat{\boldsymbol{\xi}}_{t|t-1} \\ \widehat{\boldsymbol{Y}}_{t|t-1} \end{array}\right), \left(\begin{array}{cc} \boldsymbol{P}_{t|t-1} & \boldsymbol{L}_{t|t-1} \\ \boldsymbol{L}_{t|t-1}' & \boldsymbol{D}_{t|t-1} \end{array}\right)\right],\right.$$

where

$$
\begin{aligned}
\widehat{\boldsymbol{\xi}}_{t|t-1} &= f_t\left(\boldsymbol{Y}_1^{t-1}\right) + \boldsymbol{F}_t\widehat{\boldsymbol{\xi}}_{t-1|t-1}\,, \\
\widehat{\boldsymbol{Y}}_{t|t-1} &= g_t\left(\boldsymbol{Y}_1^{t-1}\right) + \boldsymbol{H}_t\widehat{\boldsymbol{\xi}}_{t|t-1} + \boldsymbol{J}_t\widehat{\boldsymbol{\xi}}_{t-1|t-1}\,, \\
\boldsymbol{P}_{t|t-1} &= \boldsymbol{F}_t\boldsymbol{P}_{t-1|t-1}\boldsymbol{F}_t^{'} + \boldsymbol{Q}_t, \\
\boldsymbol{D}_{t|t-1} &= \boldsymbol{H}_t\boldsymbol{P}_{t|t-1}\boldsymbol{H}_t^{'} + \boldsymbol{R}_t + \boldsymbol{J}_t\boldsymbol{P}_{t-1|t-1}\boldsymbol{J}_t^{'} + \boldsymbol{H}_t\boldsymbol{F}_t\boldsymbol{P}_{t-1|t-1}\boldsymbol{J}_t^{'} \\
&\quad + \boldsymbol{J}_t\boldsymbol{P}_{t-1|t-1}\boldsymbol{F}_t^{'}\boldsymbol{H}_t^{'} + \boldsymbol{H}_t\boldsymbol{S}_t + \boldsymbol{S}_t^{'}\boldsymbol{H}_t^{'}, \\
\boldsymbol{L}_{t|t-1} &= \boldsymbol{P}_{t|t-1}\boldsymbol{H}_t^{'} + \boldsymbol{F}_t\boldsymbol{P}_{t-1|t-1}\boldsymbol{J}_t^{'} + \boldsymbol{S}_t.
\end{aligned}
$$

The update step follows from the fact that $\boldsymbol{\xi}_t \left| \boldsymbol{Y}_1^t \right. \sim N\left(\widehat{\boldsymbol{\xi}}_{t|t}\,, \boldsymbol{P}_{t|t}\right)$, where

$$
\begin{aligned}
\widehat{\boldsymbol{\xi}}_{t|t} &= \widehat{\boldsymbol{\xi}}_{t|t-1} + \boldsymbol{L}_{t|t-1}\left(\boldsymbol{D}_{t|t-1}\right)^{-1}\left(\boldsymbol{Y}_t - \widehat{\boldsymbol{Y}}_{t|t-1}\right), \\
\boldsymbol{P}_{t|t} &= \boldsymbol{P}_{t|t-1} - \boldsymbol{L}_{t|t-1}\left(\boldsymbol{D}_{t|t-1}\right)^{-1}\boldsymbol{L}_{t|t-1}^{'}\,.
\end{aligned}
$$

This completes a recursion cycle and the filter proceeds to the next period.

The TVD feature is embodied in the time-varying size of the matrices, while the recursion formulae do not change. It is possible that in some period we have no state or measurement vector, which can be interpreted as a zero-dimension column vector (i.e., a 0×1 vector). As long as a programming platform adopts the conformable matrix algebra for empty matrices[2], the

[2]An $m \times n$ matrix is said to be empty if either $m = 0$ or $n = 0$ (or both). The algebra for empty matrixes satisfies: i) a $0 \times m$ matrix times an $m \times n$ matrix yields a $0 \times n$ matrix. ii) a $m \times 0$ matrix times a $0 \times n$ matrix yields a $m \times n$ matrix of zeros; iii) the summation of two $0 \times m$ matrixes yields a $0 \times m$ matrix. For example, let $\boldsymbol{\xi}_{t-1}, \boldsymbol{\xi}_t, \boldsymbol{\varepsilon}_t$ be $m \times 1, 0 \times 1, 0 \times 1$ vectors, $\boldsymbol{F}_t$ be a $0 \times m$ matrix. It follows that $\boldsymbol{F}_t\boldsymbol{\xi}_{t-1} + \boldsymbol{\varepsilon}_t$ is a 0×1 vector, which is conformable with $\boldsymbol{\xi}_t$. Further assume $\boldsymbol{Y}_t$ is an $n \times 1$ vector and $\boldsymbol{H}_t$ is a $n \times 0$ matrix. It follows that $\boldsymbol{H}_t\boldsymbol{\xi}_t$ is a $n \times 1$ vector of zeros, whose size is conformable with $\boldsymbol{Y}_t$.

above formulae still apply, though they can be simplified as follows:

If $\boldsymbol{\xi}_t$ is empty, $\widehat{\boldsymbol{\xi}}_{t|t-1}$, $\boldsymbol{P}_{t|t-1}$, $\boldsymbol{L}_{t|t-1}$, $\widehat{\boldsymbol{\xi}}_{t|t}$, $\boldsymbol{P}_{t|t}$ are empty while $\widehat{\boldsymbol{Y}}_{t|t-1} = g_t\left(\boldsymbol{Y}_1^{t-1}\right) + \boldsymbol{J}_t\widehat{\boldsymbol{\xi}}_{t-1|t-1}$ and $\boldsymbol{D}_{t|t-1} = \boldsymbol{R}_t + \boldsymbol{J}_t\boldsymbol{P}_{t-1|t-1}\boldsymbol{J}_t'$. In other words, the prediction and update on $\boldsymbol{\xi}_t$ are skipped. Note that in the next period, states can reappear relying on past observations. That is, $\boldsymbol{\xi}_{t+1} = f_{t+1}\left(\boldsymbol{Y}_1^t\right) + \boldsymbol{\varepsilon}_{t+1}$.

If $\boldsymbol{Y}_t$ is empty, $\widehat{\boldsymbol{Y}}_{t|t-1}$, $\boldsymbol{D}_{t|t-1}$, $\boldsymbol{L}_{t|t-1}$ are empty while $\widehat{\boldsymbol{\xi}}_{t|t} = \widehat{\boldsymbol{\xi}}_{t|t-1}$ and $\boldsymbol{P}_{t|t} = \boldsymbol{P}_{t|t-1}$. In other words, we update the states by making a one-period-ahead prediction. In addition, empty $\boldsymbol{Y}_t$ does not contribute to the likelihood evaluation.

4. The state space form of ARMA

Let $\{Z_t\}$ be a univariate $ARMA\,(p,q)$ process

$$Z_t = c + \sum_{i=1}^{p}\phi_i Z_{t-i} + \varepsilon_t + \sum_{i=1}^{q}\theta_i\varepsilon_{t-i},$$

where $\varepsilon_t \sim iidN\left(0,\sigma^2\right)$. There are various ways to write an ARMA model into its state space form. In Akaike (1973, 1974) and Jones (1980), the state vector is chosen as the projection of $Z_t, ..., Z_{t+r-1}$ on the information set of time t, where $r \equiv \max\left(p, q+1\right)$. Hamilton (1994) explores the fact that the lagged sum of an AR process is an ARMA process. The state vector keeps track of r recent values of a latent $AR\,(p)$ process. de Jong and Penzer (2004) extend the idea of Pearlman (1980) and propose a canonical form in which the length of the state vector is reduced to $\max\left(p,q\right)$.

Our state space representation of the ARMA model is different from the well-known SSMs in three aspects. First, it is more parsimonious. The length of the state vector is q except for the initial p periods when states have dynamic dimensions. Second, it is more general. The well-known SSMs are most suitable for a stationary ARMA process with the initial states coming from the stationary distribution. Our representation can conveniently handle other types of initial distributions and time-varying parameters. Third, it is more intuitive. States simply consist of the disturbance terms and unobserved presample values in the ARMA model.

Let $\boldsymbol{W}_t = (Z_t, ..., Z_{t-p+1}, \varepsilon_t, ..., \varepsilon_{t-q+1})'$, $t = 0, ..., T$. Assume $\boldsymbol{W}_0 \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The ARMA literature distinguishes the exact likelihood and the conditional likelihood. The exact likelihood approach assumes $\boldsymbol{W}_0$ is conformable with the stationary distribution of the ARMA process. The conditional likelihood method treats either $\boldsymbol{W}_0$ or $\boldsymbol{W}_p$ as deterministic. The well-known SSMs are all suitable for exact likelihood evaluation, but is inconvenient for handling the conditional likelihood since the states are not expressed in terms of Z_t or ε_t. The flexible SSM accommodates both exact and conditional likelihood by properly specifying the initial states.

The state vector in our flexible SSM is

$$\boldsymbol{\xi}_t = (Z_0, Z_{-1}..., Z_{t-p+1}, \varepsilon_t, ..., \varepsilon_{t-q+1})'.$$

By definition, $\boldsymbol{\xi}_0 = \boldsymbol{W}_0$. Note that the state vector shrinks in the initial p periods. Starting from time p, the state vector only contains disturbances $(\varepsilon_t, ..., \varepsilon_{t-q+1})'$.

Let $\boldsymbol{\Phi} = (\phi_1, ..., \phi_p)$, $\boldsymbol{\Theta} = (\theta_1, ..., \theta_q)$, $\boldsymbol{E}_i = \begin{pmatrix} \boldsymbol{I}_i & \boldsymbol{0}_{i,1} \end{pmatrix}$, where $\boldsymbol{I}_i$ and $\boldsymbol{0}_{i,1}$ are identity and zero matrices respectively, whose subscripts indicate matrix dimension. The transition equation in time $t = 1, ..., p$ is given by

$$\boldsymbol{\xi}_t = \begin{pmatrix} \boldsymbol{E}_{p-t} & \mathbf{0}_{p-t,q} \\ \mathbf{0}_{1,p-t+1} & \mathbf{0}_{1,q} \\ \mathbf{0}_{q-1,p-t+1} & \boldsymbol{E}_{q-1} \end{pmatrix} \boldsymbol{\xi}_{t-1} + \begin{pmatrix} \mathbf{0}_{p-t,1} \\ \varepsilon_t \\ \mathbf{0}_{q-1,1} \end{pmatrix},$$

and the measurement equation is

$$Z_t = c + \sum_{i=1}^{t-1} \phi_i Z_{t-i} + \begin{pmatrix} \mathbf{0}_{1,p-t} & 1 & \mathbf{0}_{1,q-1} \end{pmatrix} \boldsymbol{\xi}_t + (\phi_t, ..., \phi_p, \boldsymbol{\Theta})\, \boldsymbol{\xi}_{t-1}.$$

Note that at time $t = p$, $\boldsymbol{E}_{p-t}, \mathbf{0}_{p-t,q}, \mathbf{0}_{p-t,1}, \mathbf{0}_{1,p-t}$ are empty, but the formulae still apply.

For time $t = p+1, ..., T$, the state space form has time-invariant parameters and dimensions:

$$\boldsymbol{\xi}_t = \begin{pmatrix} \mathbf{0}_{1,q} \\ \boldsymbol{E}_{q-1} \end{pmatrix} \boldsymbol{\xi}_{t-1} + \begin{pmatrix} \varepsilon_t \\ \mathbf{0}_{q-1,1} \end{pmatrix},$$

$$Z_t = c + \sum_{i=1}^{p} \phi_i Z_{t-i} + \begin{pmatrix} 1 & \mathbf{0}_{1,q-1} \end{pmatrix} \boldsymbol{\xi}_t + \boldsymbol{\Theta} \boldsymbol{\xi}_{t-1}.$$

Suppose the distribution of $\boldsymbol{W}_0$ is explicitly specified (as in the case of the conditional likelihood), we can immediately apply the flexible SSM. However, we often do not articulate the initial distribution but require $\boldsymbol{W}_0$ being generated from the stationary distribution. Note that the transition matrix at time 1 is not square, so we cannot invert it to generate a stationary distribution. However, we can make it square by temporarily expanding $\boldsymbol{\xi}_1 = \boldsymbol{W}_1$ so that $\boldsymbol{\xi}_1 = \boldsymbol{c}_1 + \boldsymbol{F}_1 \boldsymbol{\xi}_0 + \widetilde{\varepsilon}_1$, where

$$\boldsymbol{c}_1 = \begin{pmatrix} c \\ \mathbf{0}_{p+q-1,1} \end{pmatrix}, \boldsymbol{F}_1 = \begin{pmatrix} \boldsymbol{\Phi} & \boldsymbol{\Theta} \\ \boldsymbol{E}_{p-1} & \mathbf{0}_{p-1,q} \\ \mathbf{0}_{1,p} & \mathbf{0}_{1,q} \\ \mathbf{0}_{q-1,p} & \boldsymbol{E}_{q-1} \end{pmatrix}, \widetilde{\varepsilon}_1 = \begin{pmatrix} \varepsilon_t \\ \mathbf{0}_{p-1,1} \\ \varepsilon_t \\ \mathbf{0}_{q-1,1} \end{pmatrix}.$$

Then the stationary distribution can be generated by

$$E(\boldsymbol{\xi}_0) = \left(\boldsymbol{I}_{(p+q)} - \boldsymbol{F}_1\right)^{-1} \boldsymbol{c}_1,$$
$$vec\left[Var(\boldsymbol{\xi}_0)\right] = \left(\boldsymbol{I}_{(p+q)^2} - \boldsymbol{F}_1 \otimes \boldsymbol{F}_1\right)^{-1} vec(\boldsymbol{Q}_1).$$

where $\boldsymbol{Q}_1$ is the covariance matrix of $\widetilde{\varepsilon}_1$, that is, a $(p+q)\times(p+q)$ matrix of zeros except for $(1,1)$, $(1,p+1)$, $(p+1,1)$,$(p+1,p+1)$ elements being σ^2.

In summary, our flexible SSM employs TVD states to handle the initial distribution, but from time $p+1$ to T, the fixed-length state vector only tracks the MA part of the series. The AR part is predetermined and thus put in the measurement equation.

5. Dynamic factor model with missing data

Factor models have wide applications in macroeconomic forecasting (e.g., Stock and Watson, 2002; Forni et al., 2003; Schumacher, 2007), monetary policy analysis (Bernanke et al., 2005; Stock and Watson, 2005) and business cycle transmission study (Eickmeier, 2007). We consider a factor model with randomly missing data similar to Jungbacker et al. (2011), but propose a more parsimonious state space representation.

Let $\boldsymbol{Y}_t$ be an $n\times 1$ vector of observations, determined by an $m\times 1$ vector of common factors $\boldsymbol{f}_t$ and an $n\times 1$ vector of idiosyncratic terms $\boldsymbol{v}_t$ such that

$$\boldsymbol{Y}_t = \boldsymbol{\Lambda}\boldsymbol{f}_t + \boldsymbol{v}_t. \tag{3}$$

Both factors and idiosyncratic components follow autoregressive processes

$$\boldsymbol{f}_t = \boldsymbol{F}\boldsymbol{f}_{t-1} + \boldsymbol{\varepsilon}_t,$$
$$\boldsymbol{v}_t = \boldsymbol{\Phi}\boldsymbol{v}_{t-1} + \boldsymbol{u}_t,$$

where $\boldsymbol{\varepsilon}_t \sim iidN(\mathbf{0}, \boldsymbol{Q})$ and $\boldsymbol{u}_t \sim iidN(\mathbf{0}, \boldsymbol{R})$.

If complete data of $\boldsymbol{Y}_t$ were available, the model could be cast in two state space forms. The first formulation takes partial difference of Eq. (3) and obtain

$$\boldsymbol{Y}_t = \boldsymbol{\Phi}\boldsymbol{Y}_{t-1} + \boldsymbol{\Lambda}\boldsymbol{f}_t - \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{f}_{t-1} + \boldsymbol{u}_t. \tag{4}$$

The second formulation puts both $\boldsymbol{f}_t$ and $\boldsymbol{v}_t$ in the state vector. Note that the dimension of the observation vector is typically much larger than the number of factors. The second formulation induces a high-dimension state vector. However, in the presence of missing data, the second formulation is still valid, while the first one is not, due to lack of full observability of the lagged term $\boldsymbol{Y}_{t-1}$.

We follow the notations of Jungbacker et al. (2011) in handling missing data in $\boldsymbol{Y}_t$. Consider an $n \times 1$ vector $\boldsymbol{Z}_t$. The vector $\boldsymbol{Z}_t(\boldsymbol{o}_s)$ contains all elements of $\boldsymbol{Z}_t$ that correspond to observed entries in $\boldsymbol{Y}_s$ (the subscripts t, s may be distinct). In other words, $\boldsymbol{o}_s$ is a logical index indicating the observed entries in $\boldsymbol{Y}_s$ and we use $\boldsymbol{o}_s$ to select corresponding elements in $\boldsymbol{Z}_t$. Similarly, $\boldsymbol{Z}_t(\boldsymbol{m}_s)$ contains all elements of $\boldsymbol{Z}_t$ that correspond to missing entries in $\boldsymbol{Y}_s$. We can also use logical indexes to extract corresponding rows and columns of an $n \times n$ matrix $\boldsymbol{A}$. For example, $\boldsymbol{A}(\boldsymbol{o}_s, :)$ denotes a row selection, $\boldsymbol{A}(:, \boldsymbol{o}_s)$ is a column selection, and $\boldsymbol{A}(\boldsymbol{m}_s, \boldsymbol{o}_s)$ is a both row and column selection. In addition, we may take logical AND of two indexes. For example, $\boldsymbol{Z}_t(\boldsymbol{o}_s\boldsymbol{o}_r)$ selects all elements of $\boldsymbol{Z}_t$ observed in both $\boldsymbol{Y}_s$ and $\boldsymbol{Y}_r$. This double index is needed for the SSM of Jungbacker et al. (2011), but not ours.

In the SSM of Jungbacker et al. (2011), the state vector $\boldsymbol{\xi}_t$ consists of $\boldsymbol{f}_t, \boldsymbol{f}_{t-1}, \boldsymbol{v}_t(\boldsymbol{o}_t\boldsymbol{m}_{t-1}), \boldsymbol{v}_t(\boldsymbol{m}_t\boldsymbol{m}_{t-1}), \boldsymbol{v}_t(\boldsymbol{m}_t\boldsymbol{o}_{t-1})$ stacked as a column vector.

The observations in period t consist of $\boldsymbol{Y}_t(\boldsymbol{o}_t\boldsymbol{o}_{t-1})$ and $\boldsymbol{Y}_t(\boldsymbol{o}_t\boldsymbol{m}_{t-1})$. Jungbacker et al. (2011) assume that $\boldsymbol{\Phi}$ is diagonal so that $\boldsymbol{Y}_t(\boldsymbol{o}_t\boldsymbol{o}_{t-1})$ only depends on $\boldsymbol{Y}_{t-1}(\boldsymbol{o}_t\boldsymbol{o}_{t-1})$ in Eq. (4), which enables us to write the observation equation as

$$\begin{bmatrix} \boldsymbol{Y}_t(\boldsymbol{o}_t\boldsymbol{o}_{t-1}) \\ \boldsymbol{Y}_t(\boldsymbol{o}_t\boldsymbol{m}_{t-1}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}(\boldsymbol{o}_t\boldsymbol{o}_{t-1}, \boldsymbol{o}_t\boldsymbol{o}_{t-1})\boldsymbol{Y}_{t-1}(\boldsymbol{o}_t\boldsymbol{o}_{t-1}) \\ \boldsymbol{0} \end{bmatrix} +$$
$$\begin{bmatrix} \boldsymbol{\Lambda}(\boldsymbol{o}_t\boldsymbol{o}_{t-1},:) & \widetilde{\boldsymbol{\Phi}}(\boldsymbol{o}_t\boldsymbol{o}_{t-1},:) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{\Lambda}(\boldsymbol{o}_t\boldsymbol{m}_{t-1},:) & \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{\xi}_t + \begin{bmatrix} \boldsymbol{u}_t(\boldsymbol{o}_t\boldsymbol{o}_{t-1}) \\ \boldsymbol{0} \end{bmatrix},$$

where $\widetilde{\boldsymbol{\Phi}} = -\boldsymbol{\Phi}\boldsymbol{\Lambda}$.

The transition equation of $\boldsymbol{\xi}_t$ needs careful specification. The transitions of $\boldsymbol{f}_t$ and $\boldsymbol{f}_{t-1}$ are straightforward. As for $\boldsymbol{v}_t(\boldsymbol{o}_t\boldsymbol{m}_{t-1})$ and $\boldsymbol{v}_t(\boldsymbol{m}_t\boldsymbol{m}_{t-1})$, note that their union is $\boldsymbol{v}_t(\boldsymbol{m}_{t-1})$. Under the assumption that $\boldsymbol{\Phi}$ is diagonal, $\boldsymbol{v}_t(\boldsymbol{m}_{t-1})$ only depends on its own autoregressive term $\boldsymbol{v}_{t-1}(\boldsymbol{m}_{t-1})$, which can be decomposed into $\boldsymbol{v}_{t-1}(\boldsymbol{m}_{t-1}\boldsymbol{m}_{t-2})$ and $\boldsymbol{v}_{t-1}(\boldsymbol{m}_{t-1}\boldsymbol{o}_{t-2})$, namely the fourth and fifth components of $\boldsymbol{\xi}_{t-1}$. The only complication is the variable order of $\boldsymbol{v}_t(\boldsymbol{o}_t\boldsymbol{m}_{t-1})$, $\boldsymbol{v}_t(\boldsymbol{m}_t\boldsymbol{m}_{t-1})$ may be different from $\boldsymbol{v}_{t-1}(\boldsymbol{m}_{t-1}\boldsymbol{m}_{t-2})$, $\boldsymbol{v}_{t-1}(\boldsymbol{m}_{t-1}\boldsymbol{o}_{t-2})$. Therefore, by reshuffling $\boldsymbol{\Phi}$, we can obtain the transition of $\boldsymbol{v}_t(\boldsymbol{o}_t\boldsymbol{m}_{t-1})$ and $\boldsymbol{v}_t(\boldsymbol{m}_t\boldsymbol{m}_{t-1})$ in an autoregressive form. Lastly, the transition of $\boldsymbol{v}_t(\boldsymbol{m}_t\boldsymbol{o}_{t-1})$ depends on $\boldsymbol{v}_{t-1}(\boldsymbol{m}_t\boldsymbol{o}_{t-1})$, which can be represented by a linear combination of $\boldsymbol{Y}_{t-1}(\boldsymbol{m}_t\boldsymbol{o}_{t-1})$, $\boldsymbol{f}_{t-1}$ and $\boldsymbol{u}_t(\boldsymbol{m}_t\boldsymbol{o}_{t-1})$.

Our flexible SSM does not rely on $\boldsymbol{v}_t$ in the state vector, which only consists of factors $\boldsymbol{f}_t$ and missing data $\boldsymbol{Y}_t(\boldsymbol{m}_t)$. First rewrite Eq. (4) as

$$\boldsymbol{Y}_t = \boldsymbol{\Phi}\boldsymbol{Y}_{t-1} + \boldsymbol{G}\boldsymbol{f}_{t-1} + \boldsymbol{w}_t, \tag{5}$$

where $\boldsymbol{G}=\boldsymbol{\Lambda F}-\boldsymbol{\Phi\Lambda}$, $\boldsymbol{w}_t=\boldsymbol{\Lambda}\boldsymbol{\varepsilon}_t+\boldsymbol{u}_t$, $\begin{pmatrix}\boldsymbol{\varepsilon}_t\\ \boldsymbol{w}_t\end{pmatrix}\sim N\left[\boldsymbol{0},\begin{pmatrix}\boldsymbol{Q} & \boldsymbol{Q\Lambda}'\\ \boldsymbol{\Lambda Q} & \boldsymbol{\Lambda Q\Lambda}'+\boldsymbol{R}\end{pmatrix}\right]$.

Note that $\boldsymbol{Y}_{t-1}$ can be decomposed into $\boldsymbol{Y}_{t-1}(\boldsymbol{o}_{t-1})$ and $\boldsymbol{Y}_{t-1}(\boldsymbol{m}_{t-1})$. Eq. (5) implies that $\boldsymbol{Y}_t$ is determined by $\boldsymbol{Y}_{t-1}(\boldsymbol{o}_{t-1})$, $\boldsymbol{Y}_{t-1}(\boldsymbol{m}_{t-1})$ and $\boldsymbol{f}_{t-1}$. The first one is predetermined, while the last two are exactly the state vector of time $t-1$. It follows that the transition equation is given by

$$\begin{aligned}\begin{pmatrix}\boldsymbol{f}_t\\ \boldsymbol{Y}_t(\boldsymbol{m}_t)\end{pmatrix}=&\begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{\Phi}(\boldsymbol{m}_t,\boldsymbol{o}_{t-1})\boldsymbol{Y}_{t-1}(\boldsymbol{o}_{t-1})\end{pmatrix}\\ &+\begin{bmatrix}\boldsymbol{F} & \boldsymbol{0}\\ \boldsymbol{G}(\boldsymbol{m}_t,:) & \boldsymbol{\Phi}(\boldsymbol{m}_t,\boldsymbol{m}_{t-1})\end{bmatrix}\begin{bmatrix}\boldsymbol{f}_{t-1}\\ \boldsymbol{Y}_{t-1}(\boldsymbol{m}_{t-1})\end{bmatrix}+\begin{bmatrix}\boldsymbol{\varepsilon}_t\\ \boldsymbol{w}_t(\boldsymbol{m}_t)\end{bmatrix},\end{aligned}$$

and the measurement equation is

$$\begin{aligned}\boldsymbol{Y}_t(\boldsymbol{o}_t)=&\,\boldsymbol{\Phi}(\boldsymbol{o}_t,\boldsymbol{o}_{t-1})\boldsymbol{Y}_{t-1}(\boldsymbol{o}_{t-1})\\ &+\begin{bmatrix}\boldsymbol{G}(\boldsymbol{o}_t,:) & \boldsymbol{\Phi}(\boldsymbol{o}_t,\boldsymbol{m}_{t-1})\end{bmatrix}\begin{bmatrix}\boldsymbol{f}_{t-1}\\ \boldsymbol{Y}_{t-1}(\boldsymbol{m}_{t-1})\end{bmatrix}+\boldsymbol{w}_t(\boldsymbol{o}_t).\end{aligned}$$

If we are willing to take the initial observation as fixed, we have already obtained the desired state space form. However, if we intend to i) handle presample values, ii) initialize states by the stationary distribution, and iii) obtain exactly the same filtering result of the SSM that tracks both $\boldsymbol{f}_t$ and $\boldsymbol{v}_t$, we may temporarily enlarge the state equation in period 1 such that

$$\begin{pmatrix}\boldsymbol{f}_t\\ \boldsymbol{Y}_t\end{pmatrix}=\begin{pmatrix}\boldsymbol{F} & \boldsymbol{0}\\ \boldsymbol{G} & \boldsymbol{\Phi}\end{pmatrix}\begin{pmatrix}\boldsymbol{f}_{t-1}\\ \boldsymbol{Y}_{t-1}\end{pmatrix}+\begin{pmatrix}\boldsymbol{\varepsilon}_t\\ \boldsymbol{w}_t\end{pmatrix},$$

so as to remove the unobserved lagged term in the observation equation in period 1.

Compared with the state space representation of Jungbacker et al. (2011), our flexible SSM represents the same process but has some advantages. First, the state vector is shorter. Second, no diagonal restrictions are put on $\mathbf{\Phi}$. Third, states need not to be reshuffled. Fourth, the representation is intuitive. The elements in $\boldsymbol{Y}_t$, no matter as the states or observations, always get access to $\boldsymbol{Y}_{t-1}$ partially from the past observations and partially from the previous states.

We conduct a simulation exercise to compare the Jungbacker et al. (2011) form (SSM2) and the flexible form (SSM3). Since the high-dimension form (SSM1) by stacking both $\boldsymbol{f}_t$ and $\boldsymbol{v}_t$ is also valid in the presence of missing data, we include it to benchmark the results. Note that we are working on time-varying coefficients, it needs time to configure the coefficient matrices as well as run the Kalman filter, so we report both the coefficient building time as well as the total computation time, which is the sum of the building and filtering time. The computation time of the benchmark model is normalized to 1, and the speed relative to SSM1 is reported. As for the size of the factor model, the number of factors is fixed at 10 and the sample size equals 1000, while the dimension of the observation vector varies among 10, 50 and 100. Three scenarios with 1%, 10% and 30% randomly missing observations are experimented for 100 times with different pseudo-datasets. Table 1 presents the average computation time and number of states with standard deviations in parentheses. The average number of states is not an integer because it is averaged across the T periods. When the dimension of the observation vector is a low as 10, the time-invariant SSM1 runs faster than the two time-varying models, for it saves the model building time. The flexible SSM is

1.73, 1.77, 1.81 times faster than the traditional one for the three missing data scenarios. As the dimension of the observation vector rises, the advantage of time-varying SSMs becomes more apparent, since the states of the time-varying model grow slower than the time-invariant model. For example, in the case of $n = 50$ and 1% missing data, the SSM2 needs 78% of the total time of SSM1, while SSM3 further reduces it to 53%. The relative speed gains of the flexible SSM is largest when both the observation vector is huge and missing data portion is large. For $n = 100$ and 30% missing data, SSM3 is 2.58 times faster than SSM2.

6. Asymmetric stochastic volatility

Conditional heteroskedasticity models are widely used in financial volatility forecasting. Empirical evidence from Kim et al. (1998), Danielsson (1998), Yu (2002) suggests that stochastic volatility (SV) models often outperform GARCH models in characterizing the stylized fact of volatility clustering. Another stylized feature that both models attempt to capture is the asymmetric volatility due to the leverage effect. EGARCH by Nelson (1991) and GJR-GARCH by Glosten et al. (1993) allow (a non-linear function of) signed returns affect the conditional variance. However, SV models typically accommodate asymmetry by negatively correlated disturbances. See, among others, Harvey and Shephard (1996), Jacquier et al. (2004), Kirby (2006). By the Cholesky decomposition, two correlated variables can be represented by linear combinations of two uncorrelated variables. Therefore, SV models with correlated innovations allow a specific form that signed returns affect volatility, usually in a linear manner. See Eq. (7) in Harvey and Shephard

Table 1: Comparison of State Space Forms of the Dynamic Factor Model

1% missing	n = 10			n = 50		n = 100	
	SSM 1	SSM 2	SSM 3	SSM 2	SSM 3	SSM 2	SSM 3
Total Time	1	2.005	1.160	0.780	0.528	0.334	0.307
		(0.021)	(0.012)	(0.032)	(0.028)	(0.047)	(0.050)
Build Time	1	381.0	72.5	312.2	79.4	309.5	155.4
		(21.80)	(4.02)	(24.58)	(7.16)	(32.31)	(25.42)
Dimension	20	20.20	10.11	21.03	10.55	22.09	11.10
		(0.019)	(0.010)	(0.047)	(0.024)	(0.064)	(0.032)
10% missing							
Total Time	1	2.087	1.176	0.812	0.525	0.315	0.249
		(0.021)	(0.013)	(0.029)	(0.023)	(0.041)	(0.035)
Build Time	1	417.7	78.7	351.4	92.1	306.7	145.0
		(41.42)	(7.18)	(22.54)	(7.20)	(35.08)	(18.20)
Dimension	20	21.90	11.01	29.55	15.05	39.06	20.08
		(0.059)	(0.031)	(0.111)	(0.062)	(0.161)	(0.086)
30% missing							
Total Time	1	2.138	1.180	0.929	0.567	0.560	0.217
		(0.065)	(0.052)	(0.041)	(0.032)	(0.061)	(0.036)
Build Time	1	447.5	84.9	410.2	100.0	457.4	125.8
		(46.73)	(9.06)	(38.99)	(10.94)	(52.46)	(16.57)
Dimension	20	25.09	13.01	45.53	25.04	71.03	40.06
		(0.073)	(0.049)	(0.151)	(0.106)	(0.237)	(0.165)

SSM1 is the large state space form by stacking $\boldsymbol{f}_t$ and $\boldsymbol{v}_t$. SSM2 is the Jungbacker et al. (2011) form and SSM3 is flexible form (SSM3). Experiments are conducted for observation dimensions of 10, 50 and 100 and missing data fraction 1%, 10% and 30%. The computation time of SSM1 is always normalized to 1, and the time relative to SSM1 is reported with standard deviations in parentheses. Build Time means the time spent on constructing SSM coefficient matrices and Total Time is the sum of building and filtering time. SSM2 and SSM3 have time-varying dimension states, and Dimension reflects the average number of states across T periods.

(1996) and Eq. (2.3) in Yu (2005).

The flexible SSM supports any function of past observations in the transition equation. Therefore, our version of the SV model can characterize asymmetric volatility as flexibly as in the GARCH model. Suppose the observed returns $\{r_t\}$ are generated by

$$r_t = \sigma_t \cdot z_t, \tag{6}$$

$$\ln \sigma_t^2 = f(r_{t-1}) + \gamma \ln \sigma_{t-1}^2 + u_t, \tag{7}$$

where $z_t \sim iidN(0,1)$, $u_t \sim iidN(0,\lambda^2)$ and z_t, u_t are mutually independent. $f(r_{t-1})$ can be any non-linear function of past returns. For example, similar to a GJR-GARCH model we can set $f(r_{t-1}) = \alpha r_{t-1}^2 + \beta r_{t-1}^2 I(r_{t-1} < 0)$, where $I(\cdot)$ is an indicator function.

The insights of Harvey and Shephard (1996)'s state space method is the inference conditional on the sign of observations. We also explore this idea. Imagine that z_t is generated by two steps: first a Bernoulli draw to determine its sign, and then an independent draw from the half normal to determine its magnitude. Similarly, an observed r_t can also be interpreted as two observations, its sign s_t and its monotone transformed magnitude $y_t \equiv \ln r_t^2$. The former is generated by an independent Bernoulli draw, and the latter is generated according to Eq. (6)

$$y_t = \ln \sigma_t^2 + \ln z_t^2.$$

That is, y_t is a draw from the log of $\chi^2(1)$, shifted by $\ln \sigma_t^2$. The normality approximation of this system is proposed by Harvey et al. (1994),

$$y_t = -1.27 + \ln \sigma_t^2 + v_t, \tag{8}$$

where $v_t \sim iidN\left(0, \frac{\pi^2}{2}\right)$.

For the purpose of parameter estimation by quasi-maximum likelihood, the likelihood of $\{y_t, s_t\}$ is proportional to that of $\{y_t\}$ conditional on realized $\{s_t\}$. In other words, $\{s_t\}$ can be treated as if an exogenous series. If we let $\ln \sigma_t^2$ be the state variable, Eq. (8) serves as the measurement equation. The transition equation can be obtained from Eq. (7),

$$\ln \sigma_t^2 = f\left(s_{t-1}\sqrt{e^{y_{t-1}}}\right) + \gamma \ln \sigma_{t-1}^2 + u_t.$$

Compared with the existing state space approaches to the asymmetric SV model, our formulation is both general and straightforward. Since asymmetry is not embodied in the correlation between disturbances, there is no need to compute the disturbance moments conditional on the sign s_t, which is usually cumbersome as in Harvey and Shephard (1996).

7. Conclusion

In the standard SSM, the states are detached from observations due to its own autoregression. The observations are noise-ridden representation of the states. In this paper, the SSM is examined from a new angle. Our SSM is flexible due to the symmetry and two-way dynamics between the states and observations. This feature merits concise translation from a structure model to its state space form. In additional, the translation is straightforward. Relevant but unobserved components are placed in the state vector while all observables are in the measurement vector. The number of unobserved (observed) variables in the model may vary over time, so the length of the state (measurement) vector is also dynamic.

Despite the differences in interpreting the system dynamics, the same Kalman filter can be applied to both the standard and flexible SSM. The latter is more concise and thus the Kalman filter is expected to run faster.

Akaike, H., 1973. Maximum likelihood identification of gaussian autoregressive moving average models. Biometrika 60 (2), 255–265.

Akaike, H., 1974. Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. Annals of the Institute of Statistical Mathematics 26, 363–387.

Basdevant, O., 2003. On applications of state-space modelling in macroeconomics. Reserve Bank of New Zealand Discussion Paper Series.

Bernanke, B., Boivin, J., Eliasz, P. S., 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. The Quarterly Journal of Economics 120 (1), 387–422.

Chan, J., Koop, G., Leon-Gonzalez, R., Strachan, R., 2011. Time varying dimension models. Working papers, University of Strathclyde Business School, Department of Economics.

Danielsson, J., 1998. Multivariate stochastic volatility models: Estimation and a comparison with vgarch models. Journal of Empirical Finance 5 (2), 155–173.

de Jong, P., Penzer, J., 2004. The arma model in state space form. Statistics and Probability Letters 70 (1), 119–125.

Durbin, J., Koopman, S. J., 2012. Time Series Analysis by State Space Methods: Second Edition. Oxford University Press.

Eickmeier, S., 2007. Business cycle transmission from the us to germany–a structural factor approach. European Economic Review 51 (3), 521–551.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2003. Do financial variables help forecasting inflation and real activity in the euro area? Journal of Monetary Economics 50 (6), 1243–1255.

Glosten, L. R., Jagannathan, R., Runkle, D. E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48 (5), 1779–1801.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press: Princeton.

Harvey, A. C., 1991. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Harvey, A. C., Pierse, R. G., 1984. Estimating missing observations in economic time series. Journal of the American Statistical Association 79 (385), 125–131.

Harvey, A. C., Shephard, N., 1996. Estimation of an asymmetric stochastic volatility model for asset returns. Journal of Business & Economic Statistics 14 (4), 429–34.

Jacquier, E., Polson, N. G., Rossi, P. E., 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. Journal of Econometrics 122 (1), 185–212.

Jones, R. H., 1980. Maximum likelihood fitting of arma models to time series with missing observations. Technometrics 22 (3), 389–395.

Jungbacker, B., Koopman, S., van der Wel, M., 2011. Maximum likelihood estimation for dynamic factor models with missing data. Journal of Economic Dynamics and Control 35 (8), 1358–1368.

Kalman, R. E., 1960. A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME,Journal of Basic Engineering, 35–45.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: Likelihood inference and comparison with arch models. Review of Economic Studies 65 (3), 361–93.

Kirby, C., 2006. Linear filtering for asymmetric stochastic volatility models. Economics Letters 92 (2), 284–292.

Mergner, S., 2009. Applications of State Space Models in Finance. Universitatsverlang Gottingen.

Nelson, D. B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59 (2), 347–70.

Pearlman, J. G., 1980. An algorithm for the exact likelihood of a high-order autoregressive-moving average process. Biometrika 67 (1), 232–233.

Schumacher, C., 2007. Forecasting german gdp using alternative factor models based on large datasets. Journal of Forecasting 26 (4), 271–302.

Stock, J. H., Watson, M. W., 2002. Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics 20 (2), 147–62.

Stock, J. H., Watson, M. W., 2005. Implications of dynamic factor models for var analysis. NBER Working Papers 11467, National Bureau of Economic Research, Inc.

Yu, J., 2002. Forecasting volatility in the new zealand stock market. Applied Financial Economics 12 (3), 193–202.

Yu, J., 2005. On leverage in a stochastic volatility model. Journal of Econometrics 127 (2), 165–178.