STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

ASSESSING THE TRANSMISSION OF MONETARY POLICYSHOCKS USING DYNAMIC FACTOR MODELS

BY

DIMITRIS KOROBILIS

NO. 09-14

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

Assessing the Transmission of Monetary Policy Shocks Using Dynamic Factor Models

Dimitris Korobilis*

Abstract

The evolution of monetary policy in the U.S. is examined based on structural dynamic factor models. I extend the current literature which questions the stability of the monetary transmission mechanism, by proposing and studying time-varying parameters factor-augmented vector autoregressions (TVP-FAVAR), which allow for fast and efficient inference based on hundreds of explanatory variables. Different specifications are compared where the factor loadings, VAR coefficients and error covariances, or combinations of those, may change gradually in every period or be subject to small breaks. The model is applied to 157 post-World War II U.S. quarterly macroeconomic variables. The results clearly suggest that the propagation of the monetary and non-monetary (exogenous) shocks has altered its behavior, and specifically in a fashion which supports smooth evolution rather than abrupt change. The most notable changes were in the responses of real activity measures, prices and monetary aggregates, while other key indicators of the economy remained relatively unaffected.

Keywords: Structural FAVAR, time varying parameter model, monetary policy

JEL Classification: C11, C32, E52

*Department of Economics, University of Strathclyde, Glasgow G4 0GE, United Kingdom, Tel: (+44) 141 548 4340, e-mail: gcb07101@strath.ac.uk & The Rimini Center for Economic Analysis, Via Patara 3, 47900 Rimini (RN), Italy

1 Introduction

A challenge of great importance in modern macroeconomics is to identify whether the effect of monetary policy on the economy has changed over the years and to what extent. Understanding the evolution of the economy over the last 50 years and recognizing the degree of effectiveness of monetary policy nowadays, is of immediate interest to policy makers. Two classes of econometric models have emerged as the dominant approaches for determining the impacts of (mainly U.S.) monetary policy over time. The first one is based on estimating identified vector autoregressive (VAR) models and either comparing the impulse responses on several sub-samples of post-WW II data, or testing for structural change. In this context, Boivin and Giannoni (2006b) and Stock and Watson (2002) provide evidence of a more aggressive and stabilizing monetary policy over the recent past. A second approach is based on structural VAR or DSGE models with time-varying parameters (regression coefficients and/or volatilities), which has the implication that the mechanism that generates the shocks is also time-varying. Cogley and Sargent (2001, 2005), Primiceri (2005), Sims and Zha (2006) and Koop et al. (2009) are all studies that allow the parameters and shocks to vary either on every time period or to be subject to structural change in some periods.

Common place of these papers is that they attempt to model the effects of monetary policy in the economy as a whole by using only a restricted set of variables. While the early VAR literature relied on usually three fundamental quantities (as suggested by small theoretical models) it is currently recognized that modeling using an extended information set has crucial implications. As Stock and Watson (2005) and Bernanke et al. (2005) point out, when extracting the structural shocks from the innovations of a VAR it is important to make sure that there is no omitted variable bias. Since during the decision process there are hundreds of variables available to economic agents and policy makers, especially Central Banks (Bernanke and Boivin (2003)), it is expected that the innovations of a VAR with just three variables will not span the space of structural disturbances. This lack of information has been identified as the source of the 'price puzzle' (Sims, 1992), which for example lead Boivin and Giannoni (2006b) to consider commodity price inflation as an additional variable in their VAR, even though it was not justified by the theoretical model.

This paper adopts the structural dynamic factor framework of Stock and Watson (2005) and Bernanke et al. (2005) as the starting point, however, for the purpose of modeling the evolution of the monetary policy in the US, all model parameters are evolving over time as well. This assumption subsequently implies that the transmission of monetary and non-monetary shocks also varies in each time period. In essence, the dynamic factor model is a means of summarizing information in a large data-set - in the order of some hundreds of variables - using just few - usually less than 10 - latent variables called factors. These factors usually are the first few principal components of the large data-set, but also different methods for estimating latent factors have been proposed and used successfully the last ten years. Among the vast literature, notable studies include Bai (2003), Boivin and Ng (2005), Giannone et al. (2008)and Boivin and Giannoni (2006a). The recent implementations of Stock and Watson (2005) and Bernanke et al. (2005) have the advantage of treating the dynamic factor model as a direct generalization of large-scale structural VAR's, without though suffering from the curse of dimensionality problem.

Del Negro and Otrock (2008) is the first modeling attempt to use factors in a time-varying parameters setting. They assume that the latent factors have to be estimated from the data, using simulation methods to approximate their generating distribution. In that case, inferences are based on their full posterior density and not on point estimates that are prone to sampling error. However, in a structural setting with hundreds of macroeconomic variables, likelihood-based approaches raise several identification issues. The common solution is to place arbitrary identifying restrictions (i.e. of purely statistical nature) on some of the parameter matrices of the dynamic factor model, resulting in factors that lack interpretability and impulse responses that may not comply with economic theory. Following Del Negro and Otrock (2008), Felices and Wieladek (2009) estimate common factors of key fundamentals driving sovereign debt crises in 28 emerging market economies, and examine the evolution of the link between the common factors and the fundamentals.

In this study, while the time-varying parameters are estimated in a Bayesian context using the Gibbs sampler, the factors are replaced by the first principal components (PC) coming from the singular value decomposition of the data matrix, and consequently are treated as observed. That way the parameters can be estimated at a second step, conditional on the observed factors. The principal components estimates have economic meaning and approximate asymptotically the true factors in the case of constant loadings. However if the parameters are changing in each time period, the PC estimator may be completely different than the true factors implied by the new model. In order to alleviate any problems that might occur due to bad model fit, the following strategy is proposed: First a typical random walk evolution is defined for all drifting mean and variance equation parameters of the DFM, see for example Del Negro and Otrock (2008) and Primiceri (2005), which simplifies computations by using standard state-space methods. At a second step the random walk evolution is augmented using the flexible mixture innovation specification of Giordani and Kohn (2008). By defining time-varying parameters with stochastic innovations that are mixtures of normals, it is possible to define endogenously whether these parameters vary in every time period or they are constant in every period, plus all the possible combinations between those two (i.e. parameters which vary only in some periods).

Having established the advantage of accounting for omitted variable bias, this study adds to an expanding recent literature (Blanchard and Simon (2001), Cogley and Sargent (2001, 2005), Stock and Watson (2002), Gambetti et al. (2008), Primiceri (2005), Sims and Zha (2006), to name but a few) which tries to explain whether the Great Moderation[1] in the U.S. has occurred due to a change in Feds' reaction function (change in the propagation mechanism of the shocks, 'good policy') or due to a decline in the volatility of exogenous shocks ('good luck').It is of paramount importance to have a complete model for the economy to enable us to track how changes in the interest rate affect target variables like GDP growth, unemployment and inflation. To that end, the potential contribution of the TVP-FAVAR approach is that we are able to better understand the true behavioral source of the shocks held in the economy. By expanding the standard three-equation New-Keynesian model with the information contained in 157 U.S. quarterly macroeconomic variables we can get closer to answering whether there where any exogenous sources to the U.S. economy that resulted in the Great Moderation, or was it the value of good policy.

The remainder of the paper is as follows. Section 2 specifies the dynamic factor model as a time-varying parameters VAR model on latent factors and the monetary policy variable. Section 3 describes the data, model fit and model selection issues, Section 4 the empirical results from the new model, and Section 5 concludes.

2 Methodology

2.1 The model

The standard approach to examine the effects of monetary policy on the economy is to estimate a structural VAR on some key variables. Models of this form have the following structure

$$y_t = b_1 y_{t-1} + \ldots + b_p y_{t-p} + v_t \tag{1}$$

[1]i.e. the reduction in the volatility of output and inflation empirically observed in the post-1984 period.

where $y_t^{'} = [x_t^{'}, r_t]$, x_t is a $(n \times 1)$ vector of variables provide a representation of the economy (like output, prices, interest rates, monetary aggregates and so on), and r_t is the monetary policy instrument, i.e. the control variable of the Central Bank. The coefficients b_i, $i = 1, \ldots, p$ on each lagged value of y_t are of dimensions $(n \times n)$, and $v_t \sim N(0, \Omega)$ with Ω a $(n \times n)$ covariance matrix. A new model is introduced in this paper, which builds on the Factor-Augmented VAR (FAVAR) which is used to describe the decomposition of the n-dimensional vector of observables x_t into a lower dimensional vector of k (which is much smaller than n, i.e. $k \ll n$) unobserved factors, f_t. Using this reduced form decomposition we are able consider as many series as we need in order to capture most of the structure underlying the economy. In standard macroeconomic applications n is in the order of some hundreds of variables. The novel element used here is that all the parameters of the FAVAR are stochastic. The time-varying parameters factor-augmented VAR (TVP-FAVAR) takes the form

$$y_t = b_{1,t} y_{t-1} + \ldots + b_{p,t} y_{t-p} + v_t \tag{2}$$

where now $y_t^{'} = [f_t^{'}, r_t]$, with f_t a $(k \times 1)$ vector of latent factors, r_t is again the monetary policy instrument of dimension (1×1), $b_{i,t}$ are $(k \times k)$ coefficient matrices for $i = 1, \ldots, p$ and $t = 1, \ldots, T$, and $v_t \sim N(0, \Omega_t)$ with Ω a $(k \times k)$ full covariance matrix for each $t = 1, \ldots, T$.

The original observed series x_t are linked to the factors and the monetary policy tool through a factor regression (as in Bernanke et al. (2005)), but with drifting parameters and subsequently takes the form

$$x_t = \lambda_t^f f_t + \lambda_t^r r_t + u_t \tag{3}$$

where λ_t^f is $(n \times k)$ and λ_t^r is $(n \times 1)$, and $u_t \sim N(0, H_t)$ with $H_t = diag(\exp(h_{1,t}), \ldots, \exp(h_{n,t}))$ of dimensions $(n \times n)$, for each $t = 1, \ldots, T$. The errors u_t are assumed to be uncorrelated with the factors at all leads and lags and mutually uncorrelated at all leads and lags, namely $E(u_{i,t} f_t) = 0$ and $E(u_{i,t} u_{j,s}) = 0$ for all $i, j = 1, \ldots, n$ and $t, s = 1, \ldots, T$, $i \neq j$ and $t \neq s$. The main TVP-FAVAR model consists of Equations (2) and (3) and for simplicity I will refer to them as the 'FAVAR' and 'factor model' equations, respectively. In order to complete the model specification, it is necessary to characterize all model parameters and their dynamics.

The diagonality assumption of the covariance matrix[2] has the implication that the parameters in Eq. (3) can be estimated equation-by-equation, using

[2]There are several small issues regarding factor models, like why the error covariance matrix is diagonal? It turns out that in some contexts this assumption can be relaxed. Space limitations though do not allow extensive comments on all the features of this

the following univariate regressions, for $i = 1, \ldots, n$

$$x_{i,t} = \lambda_{i,t}^{f} f_t + \lambda_{i,t}^{r} r_t + u_{i,t} \tag{4}$$

where $u_{i,t} \sim N(0, \exp(h_{i,t}))$. Since the factors are already known, the model need not be estimated equation-by-equation. However this approach is preferred for reasons explained in the last paragraph of this subsection.

Equation (2) is a VAR system on the factors and r_t, and consequently the mean equation coefficients and covariance matrix need special treatment. Based on the recent literature on efficiently parametrizing large covariance matrices (c.f. Pourahmadi (1999)), Primiceri (2005) and Cogley and Sargent (2005) use a triangular reduction of the state (factor) error covariance being

$$A_t \Omega_t A_t^{'} = \Sigma_t \Sigma_t^{'} \tag{5}$$

or equivalently

$$\Omega_t = A_t^{-1} \Sigma_t \Sigma_t^{'} (A_t^{'-1}) \tag{6}$$

where $\Sigma_t = diag(\sigma_{1,t}, \ldots, \sigma_{k+1,t})$ and A_t is a unit lower triangular matrix with ones on the main diagonal

$$A_t = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ a_{21,t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{(k+1)1,t} & \ldots & a_{(k+1)k,t} & 1 \end{bmatrix} \tag{7}$$

Stacking all the parameters of Equation (2) in the vectors $B_t = (b_{1,t}^{'}, \ldots, b_{p,t}^{'})$, $\log \sigma_t = (\log \sigma_{1,t}^{'}, \ldots, \log \sigma_{p,t}^{'})$ and $\alpha_t = (a_{j1,t}^{'}, \ldots, a_{j(j-1),t}^{'})$ for $j = 1, \ldots, k+1$, I follow the standard convention and assume that set of drifting parameters, $\lambda_{i,t}, h_{i,t}, B_t, \alpha_t$ and $\log \sigma_t$ follow random walks[3], but augmented with the mixture innovation specification of Giordani and Kohn (2008). This implies that, for each time period, the innovation of the random walk evolution of the parameters is defined as a mixture of two normal components (see also Koop

complicated model. The interested reader may want to consult the excellent papers by Stock and Watson (2005, 2006) which cover forecasting & structural analysis using dynamic factor models.

[3]A random walk model is nonstationary and may lead to parameters that are explosive and tend to infinity. In practice, this shortcoming is not a problem since the data are finite and the parameters evolve only for a short period. This also, partly explains the choice of quarterly data in the empirical section, with not more than 200 time series observations.

et al. (2009))

$$\begin{aligned}\lambda_{i,t} &= \lambda_{i,t-1} + J^{\lambda}_{i,t}\eta^{\lambda}_t \\ h_{i,t} &= h_{i,t-1} + J^{h}_{i,t}\eta^{h}_t \\ B_t &= B_{t-1} + J^{B}_{i,t}\eta^{B}_t \\ \alpha_t &= \alpha_{t-1} + J^{\alpha}_{i,t}\eta^{\alpha}_t \\ \log\sigma_t &= \log\sigma_{t-1} + J^{\sigma}_{i,t}\eta^{\sigma}_t\end{aligned} \tag{8}$$

where $\eta^{\theta}_t \sim N(0, Q_{\theta})$ are innovation vectors independent with each other, as well as u_t and v_t, while Q_{θ} are innovation covariance matrices associated with each of the parameter vectors $\lambda_{i,t}, h_{i,t}, B_t, \alpha_t, \log\sigma_t$, where for brevity we define $\theta \in \{\lambda_i, h_i, B, \alpha, \log\sigma\}$. Some correlation can be allowed between the disturbance terms appearing in (8), which could permit modeling more complex dynamics. However, this flexibility comes at the cost of the proliferation of the parameters that need to be estimated, and the assumption made here is that all error components appearing in equations (2) - (8)are uncorrelated with each other.

The random variables J^{θ}_t, are 0/1 random variables and control structural breaks ('jumps') in the respective innovation error of each of the time-varying parameters. This specification is flexible as it allows the data to determine either one of the two extreme specifications of constant parameters (iff $J^{\theta}_t = 0$ $\forall\ t = 1, \ldots, T$) and of time-varying parameters (iff $J^{\theta}_t = 1\ \forall\ t = 1, \ldots, T$). In between those two extremes lies a specification with few breaks when $J^{\theta}_t = 1$ for only some t. Following Koop et al. (2009) it is easy to show that this framework is appropriate to implement Bayesian testing of constancy of model parameters against time-variation in some or all time-periods, using marginal likelihoods.

Two modeling issues must be clarified at this point. First, estimating equation (4) independently for each variable $x_{i,t}$, $i = 1, \ldots, n$ means that we can define a break indicator J for each row of λ_t. That is, we can have $J^{\lambda}_{i,t} \neq J^{\lambda}_{j,t}$, $i, j = 1, \ldots, n$, $i \neq j$. Subsequently, different dynamic patterns for $\lambda_{i,t}$ and $\lambda_{j,t}$ can be modeled which allows more flexibility than if an index variable J^{λ}_t - pertaining to the whole matrix λ, not just a certain row - was introduced.

Second, following Primiceri (2005), efficient estimation of the $\frac{(k-1)\times k}{2}$ elements of the vector α_t using state-space methods requires the additional assumption that the state covariance Q_{α} is block diagonal, where each block corresponds to parameters belonging to separate equations. In particular, each block consists from the parameters $a_{ij,t}$ which are in the same row of A_t. Subsequently we have the k blocks $\alpha^{block\ 1}_t = \{a_{21,t}\}$, $\alpha^{block\ 2}_t = \{a_{31,t}, a_{32,t}\}$, ..., $\alpha^{block\ k}_t = \{a_{(k+1)1,t}, ..., a_{(k+1)k,t}\}$, so that each block on the diagonal of the covariance matrix Q_{α} is of respective dimensions.

2.2 Estimation

The latent factors have to be treated as latent parameters whose posterior has to be estimated from the data. This is computationally computationally plausible, if we treat the factors as a state variable and use the Kalman filter to derive an estimate conditial on the rest of the model parameters. This approach is avoided because of the difficulty of correctly identifying the factors. Treating the factors as unknown, like the rest of the model parameters, means that strong but arbitrary identifying restrictions have to be imposed in the model, since economic theory cannot provide us with theoretical relationships when we replace observables with statistical factors. If we were to use such restrictions, there is nothing to guarantee that the estimated factors will have sound economic interpretation and be suitable for structural analysis. For example, in the constant parameters dynamic factor model setting, Bernanke et al. (2005) use a triangular identification restriction in the upper $k \times k$ block of the loadings matrix[4], and argue that the Bayesian (likelihood-based in general) estimation produces factors that do not capture information about real-activity and prices. In the time-varying setting, the identification problem is even more accented and will inevitably lead to impulse responses that are hardly in accordance with economic theory[5].

A conceptually and computationally simpler method is used here, and this is to approximate the factors using standard principal components. Empirical studies (see Stock and Watson (1999)) have shown that the first three to seven principal components capture most of the variance in the series x_t, while at the same time there is economic meaning in them (for example the

[4]This identification restriction is similar to the one that is met in cointegration analysis, i.e. the upper block is the identity matrix. This has the implication that the first series in the dataset loads exclusively on the first factor with coefficient 1, the second series loads exclusively on the second factor with coefficient 1 and so on. Hence the ordering of the variables in x_t plays a significant role as it alters the likelihood function, a serious problem that has been noted in the cointegration literature (Strachan (2003)). Unfortunately, when using factor models, Bayesian statisticians and econometricians rely heavily on such identification restrictions and, to my knowledge, there is no formal examination of their implications (apart from a quick reference to this problem in the review paper of Lopes and West (2004)).

[5]Del Negro and Otrock (2008) is one study which uses the Kalman filter (i.e. likelihood-based methods) to estimate the latent factors in a TVP setting. Their application though provides by default enough identification restrictions on the loadings matrix. This matrix has a block structure since data for one country load only to the country's specific factor and with zeros on the rest of the country factors. Even so, due to the fact that their model is heavily parameterized as well, they avoid to account for drifting autoregressive coefficients ($B = B_t$), an assumption that "...would raise additional identification issues" (Del Negro and Otrock, 2008, section 2.1).

first principal component proxies real activity measures). The principal components are computed using either the singular value decomposition (SVD) or the spectral (eigenvalue) decomposition of the data. These estimates will probably not approximate well the true factors (if they exist) implied by the TVP-DFM model. After all, the principal components only approximate the static factors of Equation (3) and do not account for their autoregressive dynamics, as those are described by Equation (2). Only likelihood-based methods can provide estimates of the dynamic factors by means of the updating scheme of the Kalman filter algorithm. Nevertheless, the SVD decomposition gives a meaningful reduced representation of the variables of interest, x_t, while at the same time the flexible modeling approach used here allows to specify endogenously the extent that the parameters vary over time.

Each time varying parameter is sampled sequentially using the Gibbs sampler. It is easy to see that conditional on the rest of the parameters and the principal component estimates of the factors, each time-varying parameter can be sampled from a conditionally Normal density using a standard state-space filter and smoother (Carter and Kohn (1994), Durbin and Koopman (2002)). Furthermore, conditional on each state variable θ, the covariances of the states, Q_θ, can be sampled using standard formulas. In fact these formulas are essentially the same as in the previous TVP-VAR works of Cogley and Sargent (2005), Primiceri (2005) and Koop et al. (2009), and details are provided in the technical appendix. The indicators J_t^θ are sampled using the algorithm of Gerlach et al. (2000). This is an efficient approach to modelling dynamic mixtures given that J_t^θ can be generated without conditioning on the states θ_t. Again, more implementation details can be found in the technical appendix.

2.3 VAR representation and impulse response functions

It is easy to show that the time-varying FAVAR model admits a VAR representation with drifting parameters. First note that Equations (2) - (3) can be rewritten as

$$g_t = \lambda_t y_t + W_t \epsilon_t^g \tag{9}$$

$$y_t = b_{1,t} y_{t-1} + \ldots + b_{p,t} y_{t-p} + A_t^{-1} \Sigma_t \epsilon_t^y \tag{10}$$

where $g_t^{'} = [x_t^{'}, r_t]$, $y_t^{'} = [f_t^{'}, r_t]$, $W_t = diag(\exp(h_{1,t})/2, \ldots, \exp(h_{n,t})/2, 0)$ such that $W_t W_t^{'} = [H_t, 0]$, the vectors A_t, Σ_t, $b_{1,t}, \ldots, b_{p,t}$ are parameters defined in section 2.1, $(\epsilon_t^g, \epsilon_t^y)$ are *iid* structural disturbances coming from a Normal distribution with zero mean and unit variance, and $\lambda_t = \begin{bmatrix} \lambda_t^f & \lambda_t^r \\ 0_{1\times k} & 1 \end{bmatrix}$.

Inserting (10) into (9) we get the final VAR form which is

$$g_t = \lambda_t b_{1,t} y_{t-1} + \ldots + \lambda_t b_{p,t} y_{t-p} + \zeta_t \tag{11a}$$

$$\zeta_t = \lambda_t (A_t^{-1} \Sigma_t) \epsilon_t^y + W_t \epsilon_t^g \tag{11b}$$

I follow Bernanke and Blinder (1992) and others in that the federal funds rate is assumed to be the monetary policy instrument. The federal funds rate is sorted last in the FAVAR equation (10), and monetary policy is identified in a recursive manner. First, the reduced form model (10) is estimated and then a lower-triangular identification restriction (Cholesky factorization) has to be imposed. This procedure is equivalent to estimating a recursive model (see Lütkepohl (2005)), and implies that the factors respond to monetary policy with one lag (i.e. after one quarter). However, as Bernanke et al. (2005) note, there is no need to impose the same assumption to the idiosyncratic components of the information variables. In particular identification of the monetary policy shocks is implemented with a set of lower-triangular exclusion restrictions using three blocks of variables. The first block includes all the slow-moving variables (like real activity measures), the second block consists only of the monetary policy tool (the federal funds rate) and, finally, in the third block fast-moving variables (like asset prices) are included. The assumption made is that the slow-moving variables are not allowed to respond contemporaneously to monetary policy shock, which is similar to the identification assumption for the factors. However, there is the last block, of fast-moving financial variables, which responds instantly to monetary policy shocks since financial markets are more sensitive to 'news'. The interested reader should consult Bernanke et al. (2005) for exact econometric details underlying this approach.

Methods developed for direct estimation of the structural TVP-FAVAR model are not readily available, but these would be a direct generalization (i.e.conditional on each t) of the constant parameters linear VARs, see Waggoner and Zha (2003). However in this nonlinear setting the computational cost of such an attempt is large as impulse responses must be estimated using simulation methods (Koop et al. (1996)). For these reasons I follow the standard convention in the literature and apply a sequential estimation procedure, where first parameters are estimated from the reduced form model and then the structural shocks are recovered.

3 Preliminary Analysis

3.1 Data

The data-set consists of quarterly observations on 157 U.S. macroeconomic time series spanning the period from 1959:Q1 to 2006:Q3. The series were downloaded from the St. Louis Fed FRED database and a complete description is given in the appendix. One of the series, the Federal Funds rate is used to identify monetary policy. The remaining 156 series are the variables in x_t , which are used to extract factors. These include series like personal income and outlays, GDP and components, assets and liabilities of commercial banks in the United States, productivity and costs measures, and selected interest rates among others. All series are seasonally adjusted, where this is applicable, and transformed to be approximately stationary.

3.2 Priors

The choice of prior distributions is determined on the basis of conjugate priors, which are specified to keep computation of the high dimensional posteriors tractable. Due to the conditionally Gaussian structure of the state equations (8), a reasonable choice for the initial state for all time varying parameters - i.e. the value of the parameter at time $t_0 = 0$) is the Normal density. The choice $\theta_0 \sim N(\mathbf{0}, 4I)$, where θ is a vector summarizing all drifting parameters $\lambda_i, h_i, B, \alpha, \log\sigma$, $\mathbf{0}$ is a vector of 0's, I is the identity matrix, and the dimensions of $\mathbf{0}, I$ correspond to the dimensions of each respective parameter. Similarly the priors on the covariance matrices $Q_{\theta_{/-h}}$ follow the inverse Wishart density, for $\theta_{/-h} \in \{\lambda_i, B, \alpha, \sigma\}$, and on the variances Q_{h_i} the inverse Gamma density, which are the standard conjugate choices (see Koop (2003)). As is the case with Bayesian analysis in general, the challenging task at this point is the choice of prior hyper-parameters, i.e. to give reasonable values for the prior means-variances. Conditional on the factors which are replaced by principal components and the jump indicators J_t^θ which are discussed later in detail, the rest of the TVP-FAVAR model is quite similar to the one used in Primiceri (2005). Subsequently, the hyperparameters are set following this authors' suggestions and further details can be found in the technical appendix.

The only 'nonstandard' parameters in this model are the ones related to the mixture innovation extension. The 0/1 variables J_t^θ are assumed to be random draws from a Bernoulli distribution, $p(J_t^\theta = 1) = \pi_\theta$, $\theta \in \{\lambda_i, h_i, B, \alpha, \log\sigma\}$. The probabilities π_θ control the transition of the index J_t^θ between the two possible states (1:break - 0:no break), and an extra hierarchical layer is introduced in order to update them from the information in the data. A Beta prior of the form $\pi_\theta \sim Beta(\tau_0, \tau_1)$ is placed on this hyper-parameter, which controls the prior belief about the number of breaks through the choice of τ_0 and τ_1. Two choices are applied in this paper, which either reflect ignorance about the number of breaks ($(\tau_0, \tau_1) = (1/2, 1/2)$ which implies $E(\pi_\theta) = 0.5$ and $std(\pi_\theta) = 0.3535$), or the expectation that only few breaks occurred during the sample period($(\tau_0, \tau_1) = (0.01, 10)$ which implies $E(\pi_\theta) = 0.001$ and $std(\pi_\theta) = 0.0095$). For more discussion about the nature of these prior choices the reader is referred to Koop et al. (2009). Note that for simplicity, and in the absence of prior information, τ_0 and τ_1 are the same for all drifting parameters defined in Eq. (8).

A challenging task evident in dynamic factor models is to select the number of static and dynamic factors. A standard strategy is to use available statistical criteria to select the number of factors. However as Bernanke et al. (2005) state, the suggested number of factors from a statistic or a criterion function may not be the actual number of factors used in the model. In that respect, the sensitivity of the results across different number of factors is considered. Ideally we would want to examine and compare all models with 3 to 9 principal components, according to the findings of Stock and Watson (2005). Notice though, that for the sake of brevity only specifications with up to $k = 4$ factors are considered. That way only a maximum of $k + 1 = 5$ series appear in the FAVAR equation, as a means of restraining the number of time-varying parameters to expand without bound. This doesn't necessarily means that there is possible misspecification, since 3 and 4 factors perform really well in many empirical applications.

3.3 Testing parameter evolution

Before discussing macroeconomic issues regarding the time-varying FAVAR, it is interesting to examine what type of time variation is supported by the data, and specifically by the principal components estimates of the factors. Apart from that, different restricted versions of the TVP-FAVAR can be considered where we can begin from the FAVAR with constant parameters and allow several (combinations of) parameters to drift. Estimating and testing all possible model combinations with marginal likelihoods is a necessary task, albeit computationally demanding. The mixture innovation extension makes this process much easier by providing posterior probabilities on the time varying nature of each parameter. That way, the mixture innovation specification can be thought of as a special form of the model selection mixture priors used in Bayesian statistics (see for example George and McCulloch (1997)). Roughly speaking, in this latter literature an indicator variable γ is used to select which regression parameter is zero or not, while here the indicator variable J^{θ} determines which parameter θ is time-varying or constant.

Table 1 presents the posterior probabilities of a break, $p(\pi_{\theta}|Data)$, for each parameter of interest θ using the informative and uninformative choices respectively. It should be noted that there is evident time evolution for all the parameters in the FAVAR equation, using the uninformative prior. The strongest evidence is for the parameters of the FAVAR equation (10), while the ones in the factor equation (9) vary moderately. Even in the case where variation in the parameters is suppressed a priori using the few breaks prior there is strong evidence of time-variation in the FAVAR equation. Koop et al. (2009) report similar evidence on their mixture innovation TVP-VAR. On the other hand, the probabilities on the factor model equation are very close to zero, using the informative prior. This evidence suggests that from now on we should focus on the results from two different models, instead of two different priors on π_{θ}. The first model is the base model which is described by the equations (5) - (10) using the uninformative prior on π_{θ}, and for short it will be denoted as the Benchmark TVP-FAVAR(k,p). The other model will be the TVP-FAVAR where we impose $J_{i,t}^{\lambda} = J_{i,t}^{h} = 0$ and $J_t^B = J_t^A = J_t^{\sigma} = 1$, for all t.

Note that we can get probabilities of a break at each point in time. These can be obtained as the average of the posterior draws of J_t^{θ}. That is, if we have a sequence of S draws from the posterior density $p(J_t^{\theta}|Data)$, then we can easily get the quantity

$$E(J_t^{\theta}|Data) = \frac{1}{S}\sum_{l=1}^{S}\left(J_t^{\theta}\right)_{(l)} \tag{12}$$

which is a time-varying proportion of models visited that had $J_t^{\theta} = 1$, where $J_{t\ (l)}^{\theta}$ is the l-th draw of J_t^{θ}. Presenting all posterior probabilities of jumps for the parameters $\lambda_{i,t}$ and $h_{i,t}$, for each $i = 1, \ldots, n$ is not possible. The same restriction applies to the quantity $E(J_t^{\theta}|Data)$, which would inform us about the evolution of the jump variable in each parameter θ. However, Figure 1 provides a visual assessment of how the median of the posterior loadings in the GDP equation vary over time under the uninformative prior. In this graph, the loadings of GDP on the four factors ($\lambda_{GDP,t}^{f}$) are denoted as ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) while the loading on the Federal funds rate ($\lambda_{GDP,t}^{r}$) is denoted as λ_5. There seems to be no obvious pattern in the variation of the loadings that could be connected with theory or previous experience.

3.4 Why time-varying loadings?

A crucial question to answer is what is the intuition behind time varying loadings when standard principal components are extracted? It is well known that asymptotically the principal components estimate the factors only when the loadings matrix is constant. These estimates though will not fit the data 100%, unless $k = n$ principal components are extracted. Hence, given the principal components estimates, having a nonlinear specification for the loadings matrix has the potential to increase the proportion of variability in x_t explained by the common component $\lambda_t f_t$. Nevertheless, nothing can guarantee that a full time-varying specification will fit better (or fit well at all) than the constant alternative. This explains why the flexible mixture innovation specification is necessary in order to overcome this problem. An extreme scenario would be all the indices $J_{\lambda,t}$ to be zero for each t, and end up with a model equivalent to the constant loadings factor model.

Also note that presumably we could use a two step estimator for the factors in the spirit of Giannone et al. (2004). These authors - in a dynamic factor model setting with constant parameters - replace the factors at a first step using principal components and estimate the values of all unknown model coefficient using simple likelihood-based methods (as it is done in this paper, but with time-varying parameters). Then at a second step, an updated estimate of the factors is obtained using the Kalman filter and keeping the values of the coefficients fixed to the estimates from the previous step. Subsequently there is a potential to do this here: save the samples from the posteriors of $\{\lambda_i, h_i, B, \alpha, \log\sigma\}$ and use Kalman filtering to get new estimates of the factors. However, unlike the DFM with constant parameters, there is no theoretical justification for doing this. The validity of this approach could be evaluated empirically though, but it is avoided since the principal components already give satisfactory performance, with smaller computational costs.

To assess visually the fit of the two models, the first column of Figure 1 plots the actual (demeaned and standardized) time-series of inflation[6] and a short term interest rate, and their projections on the first 4 static factors implied by the constant and time-varying/mixture-innovation model. That is the lines plotted are:

$$\begin{aligned} &x_{i,t}^{\circ}, \\ &x_{i,t}^{PC} = \hat{\lambda}_i^f \hat{f}_t, \\ &x_{i,t}^{TVL-PC} = \hat{\lambda}_{i,t}^f \hat{f}_t \end{aligned} \tag{13}$$

[6]The data-set contains several measures of inflation, like CPI, PCE deflator and the GDP deflator. In this example, "inflation" is defined as the annual change in the GDP deflator series.

where i=inflation, interest rate, $\hat{f}_t$ is the principal components estimate of the 4 static factors, $\hat{\lambda}^f_{i,t}$ is the mean of the posterior of the time-varying loadings, $\lambda^f_{i,t}$, and $\hat{\lambda}^f_i$ is the OLS estimate from the factor model with constant parameters. Subsequently, $x^{\circ}_{i,t}$, $x^{PC}_{i,t}$ and $x^{TVL-PC}_{i,t}$ are defined as the observed values, the projection from the PC estimates using constant loadings, and the projection from the PC estimates using time-varying loadings, respectively. The second column plots the differences of the mean absolute errors occurring from the two different projections, i.e. $\Delta error = |x^{\circ}_{i,t} - \hat{\lambda}_{i,t}\hat{f}_t| - |x^{\circ}_{i,t} - \hat{\lambda}_i\hat{f}_t|$ so that when $\Delta error < 0$, the time-varying/mixture-innovation model provides a better fit. Note that the principal components estimate of the factors, should normally be associated with a sampling error, nevertheless this error is common to the constant and time varying models and cancels out in the comparison. Additionally, posterior medians of λ_i and $\lambda_{i,t}$ have been used to get point estimates of the projections ($x^{PC}_{i,t}, x^{TVL-PC}_{i,t}$).

Although only two representative series - out of a set of 156 belonging in the vector x_t - are plotted in Figure 1, a similar picture is obtained for all of the series. The reduced standard errors from the estimation with time-varying loadings, suggest that the mixture innovation specification produces projections much closer to the original series, providing this new approach to principal component analysis with a sound empirical justification. One alternative approach for time-varying modeling of the loadings matrix is given in Stock and Watson (2008). They prove that the principal component estimates span the true space of the static factors also in the case of moderate parameter drifts in the factor loadings. They also provide examples by assuming one known break occurring in 1984 and subsequently splitting the sample in two in order to obtain principal components estimates in each subsample. It is obvious that this novel strategy has the disadvantage of assuming a known break date.

4 Empirical Results

4.1 Monetary policy mistakes and the Great Moderation

In principle, it is wise to first examine the nonsystematic policy, i.e movements in the Fed's funds rate that are attributed to exogenous shocks and not to changes in the structure of the economy. In order to achieve that, Figure 1 presents the median posterior estimates of the standard errors in the factors and the Federal Funds rate from the Benchmark model with 4 factors and 2 lags. These are the square root of the main diagonal of the matrices Ω_t, for all t. High variance of monetary policy shocks is connected with higher policy mistakes. It is obvious from Figure 1 (e) that during 1979-1984 the volatility of the shocks in the federal funds rate is quite high relative to the rest of the sample. In this period there was a shift of focus from interest rates (prices) to reserves available to banks (quantities) leading the interest rate to rise at the most rapid rate in the history of U.S.

The standard deviation of the first three factors reveals a very interesting pattern known as the Great Moderation. The variation in the errors gets much lower after approximately 1984 compared to the pre-1984 era. The same is not true for the fourth factor, whose volatility reaches peaks during 1973, 1987, 1990 and 2003 while it explodes in the last 2 years of the sample. The Great Moderation is obvious using the factors, while this is not true when the standard three-variable VAR is used. Additionally, the information contained in the factors has the implication that the standard errors in the Fed's funds rate equation are quite low and smooth (i.e. without many small peaks). The reader is adviced to make comparisons with the standard errors in the time-varying VAR's of Koop et al. (2009) and Primiceri (2005). The observation that three out of the four factors' standard errors have a structural break around 1984, is consistent with the fact that the decline in volatility has occurred broadly across the economy, affecting employment, prices and wages, and consumption.

Notice that it is straightforward to recover the time varying conditional variances of each specific variable in our data-set. These are defined as

$$var(x_{i,t}|\lambda_{i,t}, H_{ii,t}, \Omega_t) = \lambda_{i,t}\Omega_t\lambda_{i,t}^{'} + H_{ii,t} \tag{14}$$

for $i = 1, \ldots, n$. These are the variance decompositions implied by the factor model with time-varying parameters which, in the spirit of Justiniano and Primiceri (2008), allow us to examine which part of the Great Moderation is explained by using a large model with so many variables. For the real GDP series in particular, graphs are plotted for the part of the conditional variance which remains unexplained by the model, $H_{ii,t} = \exp(h_{i,t})$, and the part that is explained by the factors, $\lambda_{i,t}\Omega_t\lambda_{i,t}^{'}$. These are respectively in parts (a) and (b) of Figure 3. The shocks on the factors fully capture the structural break in 1984:Q1, while the idiosyncratic errors on real GDP show a constant downward trend. The same pattern is true for other variables: the factor decomposition $\lambda_{i,t}\Omega_t\lambda_{i,t}^{'}$ clearly explains a possibly large proportion of the Great Moderation (a break between 1982:Q2 - 1984:Q1, depending on the series $i = 1, \ldots, n$), while the idiosyncratic errors either decline or rise slowly, but definitely not in a fashion that could possibly suggest any form of unexplained structural instability.

Two arguments of great empirical value may be derived from the above observation. First, the explanation of the Great Moderation seems to lie heavily in the dynamics of one or more of the variables in this data-set. The TVP-FAVAR model suggests that there is no exogenous power that may have driven the Great Moderation in the US Economy. That is, we should seek the causes of the reduction in the mean and variance of GDP and price inflation, to the evolution in one of the macroeconomic series used in this dataset. For that reason, a more structural framework is needed, that could potentially impose more structure to the relationship between economic fundamentals than the latent factors can. Within the factor model, one solution would be to estimate the factors derived from "blocks of releases", i.e. one factor extracted from proce indexes, another one from exchange rates and so on, with respective structural interpretation (c.f. Belviso and Milani (2006)). This approach needs lots of experimentation in order to achieve empirically the perfect mix between number of factors and interpretability, and it is left for future research. A different route would be to use large scale DSGE models with time-varying volatility, as in Justiniano and Primiceri (2008).

A second remark we can make here is that the Great Moderation was not the result of an ongoing trend, like in Blanchard and Simon (2001). Hence the time-varying model is consistent with the observation of Kim and Nelson (1999) and McConnell and Perez-Quiros (2000), who document an abrupt change in the mid-80's. Notice also that for some series, including inflation, the large drop seems to occur in 1982:Q4, which is also consistent with Stock and Watson (2002) estimates using 168 series. Even so, from a purely statistical point of view, the mixture innovation seems to prefer the random walk evolution of the parameters and not abrupt structural breaks in mid-1980's.

4.2 Impulse responses of main economic indicators

At this point, it is interesting to examine and compare the impulse responses of different time periods, in a data rich environment. The first column of Figure 5 plots medians of the posterior distributions of the impulse responses of inflation, measured by the GDP deflator, and the unemployment rate, for three different representative dates. The right panel of Figure 5 plots the differences of the impulse responses between these dates. The dates are the ones used in Primiceri (2005), i.e 1971:Q1, 1983:Q3 and 1996:Q1, and are chosen arbitrarily to represent the chairmanships of Burns, Volcker and Greenspan. Responses for 2006, which would correspond to the inclusion of a "Bernanke regime" in the analysis, are not included for two reasons. First, there does not seem to be differences between responses in 1996 and any of the three quarters of 2006 in the sample. Second, there are not enough observations for the Bernanke chairmanship, while these few representative observations are at the end of the sample and may be prone to the measurement error associated with using data which, most probably, are going to be revised again in the future. All the results come from the Benchmark FAVAR with 4 factors and 2 lags.

In the top panel of Figure 5, we can see that the response of inflation in 1975 is positive until 8 quarters after the shock. This price puzzle is due to the fact that during the '70s, both inflation and interest rates were high (stagflation). Primiceri (2005) predicts responses of inflation that demonstrate a more accented price puzzle and that are also almost indistinguishable between the three periods. On the other hand, the shapes of the impulse responses of inflation from the TVP-VAR of Koop et al. (2009) are almost identical to the ones presented here. It is expected, thought, that the TVP-FAVAR will give more reliable results, since it can utilise information from much more variables than the previous models. Subsequently we can easily observe that the impulse responses of inflation have less accented a prize puzzle, compared to traditional VARs or TVP-VARs; see also the discussion in Stock and Watson (2005). The responses of unemployment, presented in the bottom panel of Figure 5, also show that following a contractionary monetary policy the job market was affected more intensely in 1975. In contrast to what (Primiceri, 2005, Figure 3) reports using a 3-variable TVP-VAR, there seems to be substantial differences in the responses of unemployment between the three periods.

Figures 6 & 7 present the posterior medians of impulses for 12 variables, coming from the Benchmark TVP-FAVAR(4,2) and the TVP-FAVAR(3,2) with constant λ and H, respectively. Note that in the second model, the TVP-FAVAR with constant loadings, I choose to use 3 lags only for the purpose of parsimony. That is, since the Benchmark model has all the parameters time-varying and 4 factors, it is interesting use a more parsimonious competing model in order to assess how large is the impulse response estimation error. The responses have the expected sign and magnitude: The real economy (GDP, Housing Starts) declines; monetary aggregates, investments, loans and interest rates decline; imports and exports fall; the dollar appreciates. The issue arising in these graphs again is the one of the wide difference between the impulse responses for the three representative time periods. Both models agree to the fact that the responses of GDP, M2, Exchange Rate, Investments at commercial banks, C&I loans, Imports & Exports, and Housing Starts were possibly quite different between these periods.

5 Conclusions

There is a large literature that examines the evolution of monetary policy over the past years. Over these years lots of changes occurred in the economy, like the moderation of GDP and inflation volatility dated circa 1984, and the anchoring of inflation expectations which is dated at the same time as well. Lots of papers try to explain the Great Moderation using small data-sets; see Giannone et al. (2007) for a survey [7]. One of the main contributions of this paper is the support for the fact that by using large data-sets we are able to better understand the nature of correlations and comovements between macroeconomic variables by using factors. This paper examines time-varying comovements and decompositions of a large number of variables.

A second contributions of this paper is to show that all the merits of the constant parameters Dynamic Factor Model (no omitted variable bias with the minimum number of parameters) can be used in a time-varying setting successfully. Using Bayesian methods in order to preserve parsimony in the time-varying parameters, and standard principal components in order to avoid identification issues arising when estimating latent factors, we can end up with sensible time-varying impulse response functions, comparable to the ones used in the time-varying VAR literature.

In order to answer more and more questions in the future, factor models can play a significant leading role since their advantages are many. At the same time, the fact that dynamic factor models are atheoretic time series models, can be tackled if they are combined with DSGE models. For example, Boivin and Giannoni (2006a) show how factors can be used in a DSGE setting, combining the merits of large data-sets, with those of structural economic models. We can anticipate that using factors in a time-varying DSGE model, would be a future challenge that will extend Justiniano and Primiceri (2008) and may provide even more interesting, new empirical findings.

[7] By the way, this is one of the few papers in this literature that uses an extensive data-set in order to examine the Great Moderation without the pitfalls of omitted variable bias

Acknowledgments

The author is grateful to John Geweke, Gary Koop, John Maheu and Simon Potter for helpful discussions and comments.

References

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*, 135171.

Belviso, F. and F. Milani (2006). Structural factor-augmented VARs (SFAVARs) and the effects of monetary policy. *Topics in Macroeconomics 6*, 1443–1443.

Bernanke, B. S. and A. S. Blinder (1992). The Federal funds rate and the channels of monetary transmission. *American Economic Review 82*, 901–21.

Bernanke, B. S. and J. Boivin (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics 50*, 525–546.

Bernanke, B. S., 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.

Blanchard, O. J. and J. Simon (2001). The long and large decline in U.S. output volatility. *Brookings Papers on Economic Activity 1*, 135–164.

Boivin, J. and M. Giannoni (2006a). DSGE models in a data-rich environment. NBER Technical Working Papers 0332, National Bureau of Economic Research, Inc.

Boivin, J. and M. P. Giannoni (2006b). Has monetary policy become more effective? *The Review of Economics and Statistics 88*, 445–462.

Boivin, J. and S. Ng (2005). Understanding and comparing factor-based forecasts. *International Journal of Central Banking 3*, 117–151.

Carter, C. K. and R. Kohn (1994). On Gibbs sampling for state space models. *Biometrika 81*, 541–553.

Cogley, T. and T. Sargent (2001). Evolving post-World War II inflation dynamics. *NBER Macroeconomic Annual 16*, 331–373.

Cogley, T. and T. Sargent (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics 8*, 262–302.

Del Negro, M. and C. Otrock (2008). Dynamic factor models with time-varying parameters. Technical report, Federal Reserve Bank of New York.

Durbin, J. and S. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–616.

Felices, G. and T. Wieladek (2009). Are sovereign debt crises fundamentals decoupling from global factors? Bank of england working papers, forthcoming, Bank of England.

Gambetti, L., E. Pappa, and F. Canova (2008). The structural dynamics of us output and inflation: what explains the changes? *Journal of Money, Credit and Banking 40*, 369–388.

George, E. I. and R. E. McCulloch (1997). Approaches for Bayesian variable selection. *Statistica Sinica 7*, 339–373.

Gerlach, R., C. Carter, and R. Kohn (2000). Efficient Bayesian inference in dynamic mixture models. *Journal of the American Statistical Association 95*, 819–828.

Giannone, D., M. Lenza, and L. Reichlin (2007). Explaining the great moderation: It is not the shocks. CEPR Discussion Papers 6600, C.E.P.R. Discussion Papers.

Giannone, D., L. Reichlin, and L. Sala (2004). Monetary policy in real time. In M. Gertler and K. Rogoff (Eds.), *NBER Macroeconomics Annual*, pp. 161–200. Cambridge: MIT Press.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*, 665–676.

Giordani, P. and R. Kohn (2008). Efficient Bayesian inference for multiple change-point and mixture innovation models. *Journal of Business and Economic Statistics 26*, 66–77.

Justiniano, A. and G. E. Primiceri (2008). The time-varying volatility of macroeconomic fluctuations. *American Economic Review 98*(3), 604–41.

Kim, C.-J. and C. Nelson (1999). Has the U.S. economy become more stable? a Bayesian approach based on a markov-switching model of the business cycle. *The Review of Economics and Statistics 81*, 608–616.

Koop, G. (2003). *Bayesian Econometrics*. Wiley, Chichester.

Koop, G., R. Leon-Gonzales, and R. Strachan (2009). On the evolution of the monetary policy transmission mechanism. *Journal of Economic Dynamics and Control 33*, 997–1017.

Koop, G., M. Pesaran, and S. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*, 77–118.

Lopes, H. F. and M. West (2004). Bayesian model assessment in factor analysis. *Statistica Sinica 14*, 41–67.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag.

McConnell, M. and G. Perez-Quiros (2000). Output fluctuations in the united states: What has changed since the early 1980s. *American Economic Review 90*, 1464–1476.

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation. *Biometrica 86*, 677–690.

Primiceri, G. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies 72*, 821–852.

Sims, C. and T. Zha (2006). Were there regime switches in macroeconomic policy? *American Economic Review 96*, 54–81.

Stock, J. H. and M. W. Watson (1999). Forecasting inflation. *Journal of Monetary Economics 44*, 293335.

Stock, J. H. and M. W. Watson (2002). Has business cycle changed and why? *NBER Macroeconomics Annual 17*, 159218.

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

Stock, J. H. and M. W. Watson (2006). Forecasting using many predictors. In G. Elliott, C. Granger, and A. Timmerman (Eds.), *Handbook of Economic Forecasting*, Volume 1, pp. 515–554. North Holland.

Stock, J. H. and M. W. Watson (2008). Forecasting in dynamic factor models subject to structural instability. In J. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics, A Festschrift in Honour of Professor David F. Hendry.* Oxford: Oxford University Press.

Strachan, R. (2003). Valid Bayesian estimation of the cointegrating error correction model. *Journal of Business and Economic Statistics 21*, 185–195.

Waggoner, D. and T. Zha (2003). A Gibbs sampler for structural vector autoregressions. *Journal of Economic Dynamics & Control 28*, 349366.

A Data and Transformations

All series were downloaded from St. Louis' FRED database and cover the quarters Q1:1959 to Q3:2006. The series HHSNTN, PMNO, PMDEL, PMNV, MOCMQ, MSONDQ (series numbered 152 - 157 in the following table) were kindly provided by Mark Watson and come from the Global Insights Basic Economics Database. All series were seasonally adjusted: either taken adjusted from FRED or by applying to the non-seasonally adjusted series a quarterly X11 filter based on an AR(4) model (after testing for seasonality). Some series in the database were observed only on a monthly basis and quarterly values were computed by averaging the monthly values over the quarter. Following Bernanke et al. (2005), the fast moving variables are interest rates, stock returns, exchange rates and commodity prices. The rest of the variables in the dataset are the slow moving variables (output, employment/unemployment etc). All variables are transformed to be approximate stationary. In particular, if $z_{i,t}$ is the original untransformed series, the transformation codes are (column Tcode below): 1 - no transformation (levels), $x_{i,t} = z_{i,t}$; 2 - first difference, $x_{i,t} = z_{i,t} - z_{i,t-1}$; 4 - logarithm, $x_{i,t} = \log z_{i,t}$; 5 - first difference of logarithm, $x_{i,t} = \log z_{i,t} - \log z_{i,t-1}$.

#	Mnemonic	Tcode	Description
1	CBI	1	Change in Private Inventories
2	FINSAL	5	Final Sales of Domestic Product
3	FSDP	5	Final Sales to Domestic Purchasers
4	GDP	5	Gross Domestic Product, 1 Decimal
5	GDPC96	5	Real Gross Domestic Product, 3 Decimal
6	FINSLC96	5	Real Final Sales of Domestic Product, 3 Decimal
7	FGCE	5	Federal Consumption Expenditures & Gross Investment
8	FGSL	5	Federal Grants-in-Aid to State & Local Governments
9	DGI	5	Federal National Defense Gross Investment
10	NDGI	5	Federal Nondefense Gross Investment
11	TGDEF	1	Net Government Saving
12	SLINV	5	State & Local Government Gross Investment
13	SLEXPND	5	State & Local Government Current Expenditures
14	EXPGSC96	5	Real Exports of Goods & Services, 3 Decimal
15	IMPGSC96	5	Real Imports of Goods & Services, 3 Decimal
16	CIVA	1	Corporate Inventory Valuation Adjustment
17	CP	5	Corporate Profits After Tax
18	CNCF	5	Corporate Net Cash Flow
19	DIVIDEND	5	Net Corporate Dividends
20	RENTIN	5	Rental Income of Persons with Capital Consumption Adjustment (CCAdj)
21	GDPDEF	5	Gross Domestic Product: Implicit Price Deflator
22	GDPCTPI	5	Gross Domestic Product: Chain-type Price Index
23	FPI	5	Fixed Private Investment
24	GGSAVE	1	Gross Government Saving
25	GSAVE	5	Gross Saving
26	PRFI	5	Private Residential Fixed Investment
27	CMDEBT	5	Household Sector: Liabilites: Household Credit Market Debt Outstanding
28	INDPRO	1	Industrial Production Index
29	NAPM	1	ISM Manufacturing: PMI Composite Index
30	HCOMPBS	5	Business Sector: Compensation Per Hour
31	HOABS	5	Business Sector: Hours of All Persons
32	RCPHBS	5	Business Sector: Real Compensation Per Hour
33	ULCBS	5	Business Sector: Unit Labor Cost
34	COMPNFB	5	Nonfarm Business Sector: Compensation Per Hour

35	HOANBS	5	Nonfarm Business Sector: Hours of All Persons
36	COMPRNFB	5	Nonfarm Business Sector: Real Compensation Per Hour
37	ULCNFB	5	Nonfarm Business Sector: Unit Labor Cost
38	UNRATE	1	Unemployment Rate: All Workers, 16 Years & Over
39	UEMPLT5	5	Civilians Unemployed - Less Than 5 Weeks
40	UEMP5TO14	5	Civilian Unemployed for 5-14 Weeks
41	UEMP15OV	5	Civilians Unemployed - 15 Weeks & Over
42	UEMP15T26	5	Civilians Unemployed for 15-26 Weeks
43	UEMP27OV	5	Civilians Unemployed for 27 Weeks and Over
44	NDMANEMP	5	All Employees: Nondurable Goods Manufacturing
45	MANEMP	5	Employees on Nonfarm Payrolls: Manufacturing
46	SRVPRD	5	All Employees: Service-Providing Industries
47	USTPU	5	All Employees: Trade, Transportation & Utilities
48	USWTRADE	5	All Employees: Wholesale Trade
49	USTRADE	5	All Employees: Retail Trade
50	USFIRE	5	All Employees: Financial Activities
51	USEHS	5	All Employees: Education & Health Services
52	USPBS	5	All Employees: Professional & Business Services
53	USINFO	5	All Employees: Information Services
54	USSERV	5	All Employees: Other Services
55	USPRIV	5	All Employees: Total Private Industries
56	USGOVT	5	All Employees: Government
57	USLAH	5	All Employees: Leisure & Hospitality
58	AHECONS	5	Average Hourly Earnings: Construction
59	AHEMAN	5	Average Hourly Earnings: Manufacturing
60	AHETPI	5	Average Hourly Earnings: Total Private Industries
61	AWOTMAN	1	Average Weekly Hours: Overtime: Manufacturing
62	AWHMAN	1	Average Weekly Hours: Manufacturing
63	HOUST	4	Housing Starts: Total: New Privately Owned Housing Units Started
64	HOUSTNE	4	Housing Starts in Northeast Census Region
65	HOUSTMW	4	Housing Starts in Midwest Census Region
66	HOUSTS	4	Housing Starts in South Census Region
67	HOUSTW	4	Housing Starts in West Census Region
68	HOUST1F	4	Privately Owned Housing Starts: 1-Unit Structures
69	PERMIT	4	New Private Housing Units Authorized by Building Permit

70	NONREVSL	5	Total Nonrevolving Credit Outstanding, SA, Billions of Dollars
71	USGSEC	5	U.S. Government Securities at All Commercial Banks
72	OTHSEC	5	Other Securities at All Commercial Banks
73	TOTALSL	5	Total Consumer Credit Outstanding
74	BUSLOANS	5	Commercial and Industrial Loans at All Commercial Banks
75	CONSUMER	5	Consumer (Individual) Loans at All Commercial Banks
76	LOANS	5	Total Loans and Leases at Commercial Banks
77	LOANINV	5	Total Loans and Investments at All Commercial Banks
78	INVEST	5	Total Investments at All Commercial Banks
79	REALLN	5	Real Estate Loans at All Commercial Banks
80	BOGAMBSL	5	Board of Governors Monetary Base, Adjusted for Changes in Reserve Req.
81	TRARR	5	Board of Governors Total Reserves, Adjusted for Changes in Reserve Req.
82	BOGNONBR	5	Non-Borrowed Reserves of Depository Institutions
83	REQRESNS	5	Required Reserves, Not Adjusted for Changes in Reserve Requirements
84	RESBALNS	5	Reserve Balances with Fed. Res. Banks, Not Adj. for Changes in Res. Req.
85	BORROW	5	Total Borrowings of Depository Institutions from the Federal Reserve
86	EXCRESNS	5	Excess Reserves of Depository Institutions
87	NFORBRES	1	Net Free or Borrowed Reserves of Depository Institutions
88	M1SL	5	M1 Money Stock
89	CURRSL	5	Currency Component of M1
90	CURRDD	5	Currency Component of M1 Plus Demand Deposits
91	DEMDEPSL	5	Demand Deposits at Commercial Banks
92	TCDSL	5	Total Checkable Deposits
93	TVCKSSL	5	Travelers Checks Outstanding
94	M2SL	5	M2 Money Stock
95	M2OWN	5	M2 Own Rate
96	SVSTCBSL	5	Savings and Small Time Deposits at Commercial Banks

97	SVSTSL	5	Savings and Small Time Deposits - Total
98	SVGCBSL	5	Savings Deposits at Commercial Banks
99	SVGTI	5	Savings Deposits at Thrift Institutions
100	SAVINGSL	5	Savings Deposits - Total
101	STDCBSL	5	Small Time Deposits at Commercial Banks
102	STDTI	5	Small Time Deposits at Thrift Institutions
103	STDSL	5	Small Time Deposits - Total
104	M2MSL	5	M2 Minus Small Time Deposits
105	M2MOWN	5	M2 Minus Own Rate
106	MZMSL	5	MZM Money Stock
107	DDDFCBNS	5	Demand Deposits Due to Foreign Commercial Banks
108	DDDFOINS	5	Demand Deposits Due to Foreign Official Institutions
109	USGVDDNS	5	U.S. Government Demand Deposits and Note Balances - Total
110	USGDCB	5	U.S. Government Demand Deposits at Commercial Banks
111	CURRCIR	5	Currency in Circulation
112	FEDFUNDS	1	Effective Federal Funds Rate
113	TB3MS	1	3-Month Treasury Bill: Secondary Market Rate
114	TB6MS	1	6-Month Treasury Bill: Secondary Market Rate
115	GS1	1	1-Year Treasury Constant Maturity Rate
116	GS3	1	3-Year Treasury Constant Maturity Rate
117	GS5	1	5-Year Treasury Constant Maturity Rate
118	GS10	1	10-Year Treasury Constant Maturity Rate
119	MPRIME	1	Bank Prime Loan Rate
120	AAA	1	Moody's Seasoned AAA Corporate Bond Yield
121	BAA	1	Moody's Seasoned BAA Corporate Bond Yield
122	sTB3MS	1	TB3MS - FEDFUNDS
123	sTB6MS	1	TB6MS - FEDFUNDS
124	sGS1	1	GS1 - FEDFUNDS
125	sGS3	1	GS3 - FEDFUNDS
126	sGS5	1	GS5 - FEDFUNDS
127	sGS10	1	GS10 - FEDFUNDS
128	sMPRIME	1	MPRIME - FEDFUNDS
129	sAAA	1	AAA - FEDFUNDS
130	sBAA	1	BAA - FEDFUNDS
131	EXSZUS	5	Switzerland / U.S. Foreign Exchange Rate
132	EXJPUS	5	Japan / U.S. Foreign Exchange Rate

133	EXUSUK	5	U.S. / U.K Foreign Exchange Rate
134	EXCAUS	5	Canada / U.S. Foreign Exchange Rate
135	PPIACO	5	Producer Price Index: All Commodities
136	PPICRM	5	Producer Price Index: Crude Materials for Further Processing
137	PPIFCF	5	Producer Price Index: Finished Consumer Foods
138	PPIFCG	5	Producer Price Index: Finished Consumer Goods
139	PFCGEF	5	Producer Price Index: Finished Consumer Goods Excluding Foods
140	PPIFGS	5	Producer Price Index: Finished Goods
141	PPICPE	5	Producer Price Index Finished Goods: Capital Equipment
142	PPIENG	5	Producer Price Index: Fuels & Related Products & Power
143	PPIIDC	5	Producer Price Index: Industrial Commodities
144	PPIITM	5	Producer Price Index: Intermediate Materials: Supplies & Components
145	CPIAUCSL	5	Consumer Price Index For All Urban Consumers: All Items
146	CPIUFDSL	5	Consumer Price Index for All Urban Consumers: Food
147	CPIENGSL	5	Consumer Price Index for All Urban Consumers: Energy
148	CPILEGSL	5	Consumer Price Index for All Urban Consumers: All Items Less Energy
149	CPIULFSL	5	Consumer Price Index for All Urban Consumers: All Items Less Food
150	CPILFESL	5	Consumer Price Index for All Urban Cons.: All Items Less Food & Energy
151	OILPRICE	5	Spot Oil Price: West Texas Intermediate
152	HHSNTN	1	U. Of Mich. Index Of Consumer Expectations
153	PMNO	1	NAPM New Orders Index
154	PMDEL	1	NAPM Vendor Deliveries Index
155	PMNV	1	NAPM Inventories Index
156	MOCMQ	5	New Orders (Net) - Consumer Goods & Materials
157	MSONDQ	5	New Orders, Nondefense Capital Goods

B Tables and Figures

Table 1: Average posterior probabilities of $J = 1$ under informative and uninformative priors

	$E(\pi_\theta \| data)$	
Parameter θ	Informative prior (few breaks)	Uninformative prior
$\lambda_{i,t}$	0 - 0.108	0.571 - 0.680
$\log h_{i,t}$	0.032 - 0.091	0.325 - 0.542
B_t	0.3386	0.9181
α_t	0.2138	0.8731
$\log \sigma_t$	0.8254	0.9781

Figure 1: Projections on first 4 principal components. Time series (left column) and difference in absolute errors (right column)

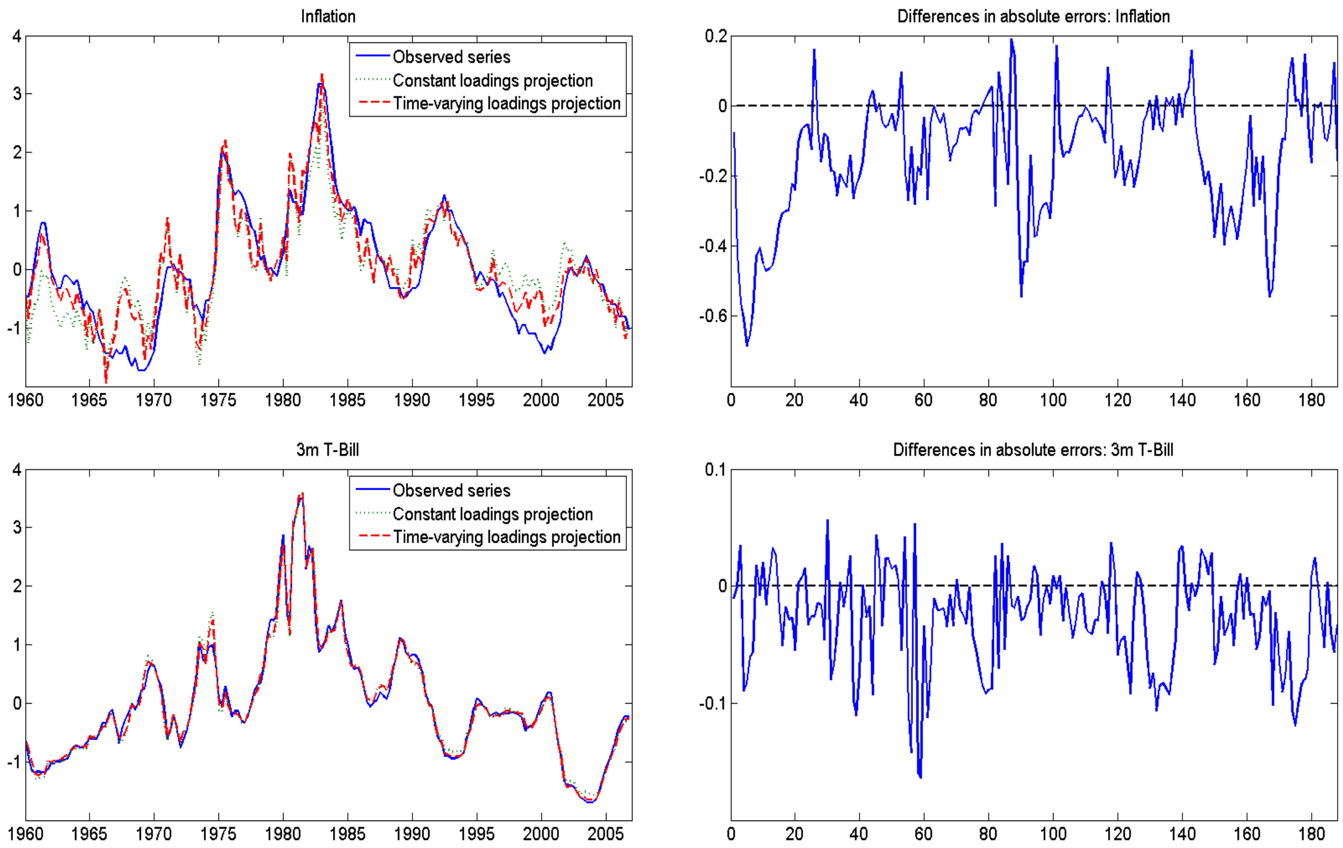


Figure 2: Time-varying standard deviations of errors on (a) 1st factor, (b) 2nd factor, (c) 3rd factor, (d) 4th factor and (e) Federal Funds Rate, from the Benchmark TVP-FAVAR(4,2))

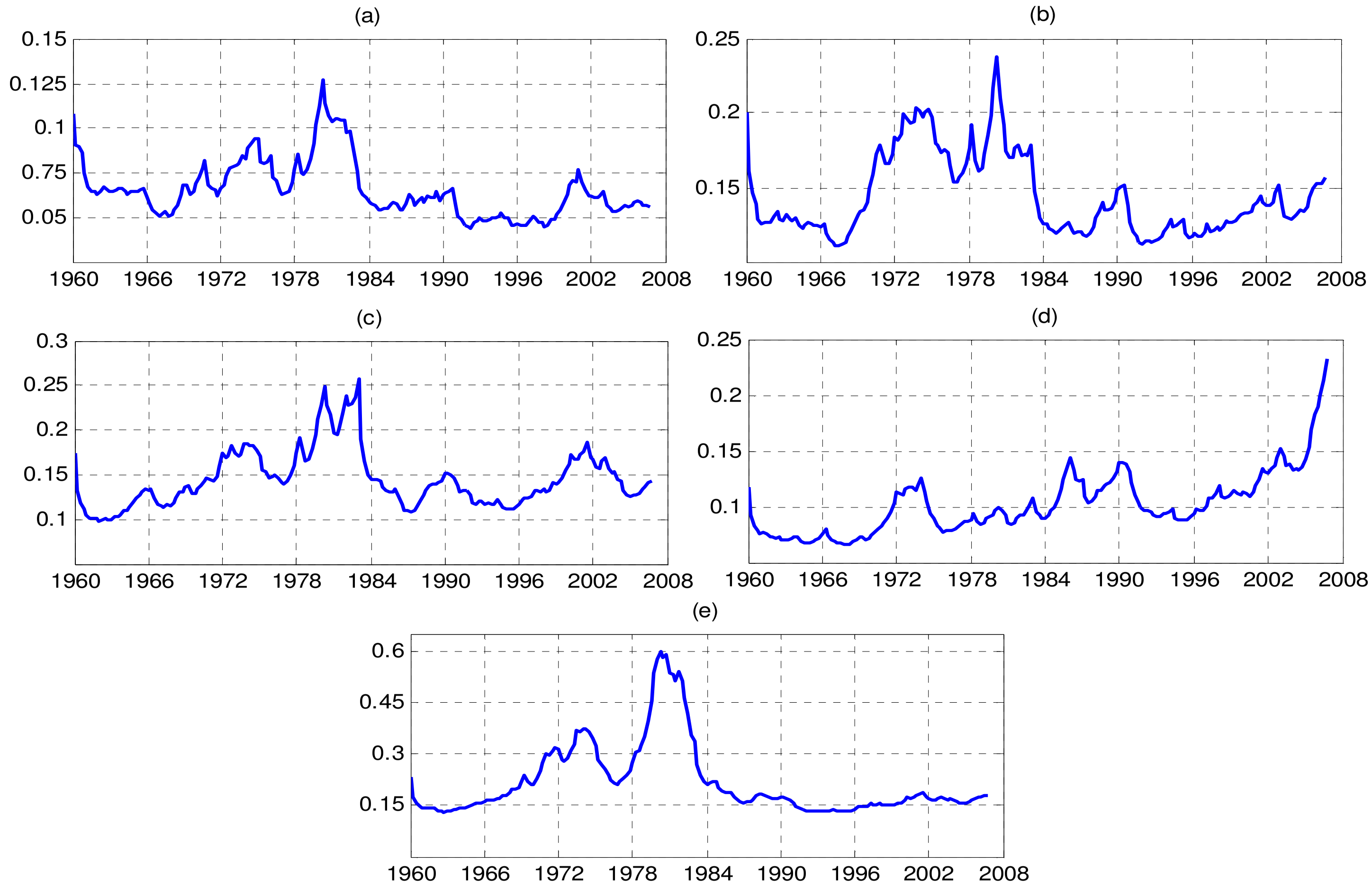


Figure 3: Factor model decomposition of real GDP volatility: (a) time-varying variance of innovation error and, (b) time-varying variance attributable to the 4 factors and the federal funds rate

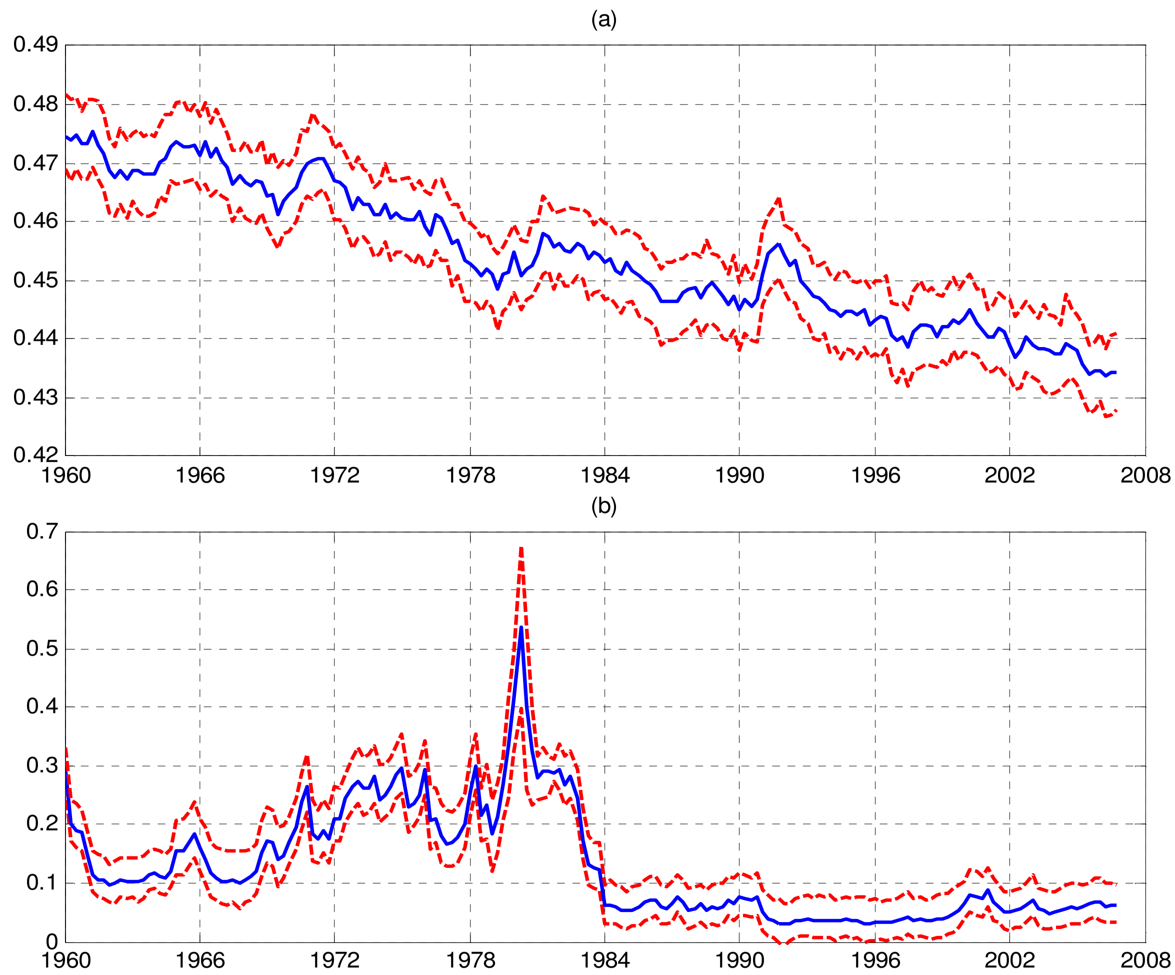


Figure 4: Time varying loadings in the real GDP equation: 4 factors (λ_{1-4}) and the federal funds rate (λ_5

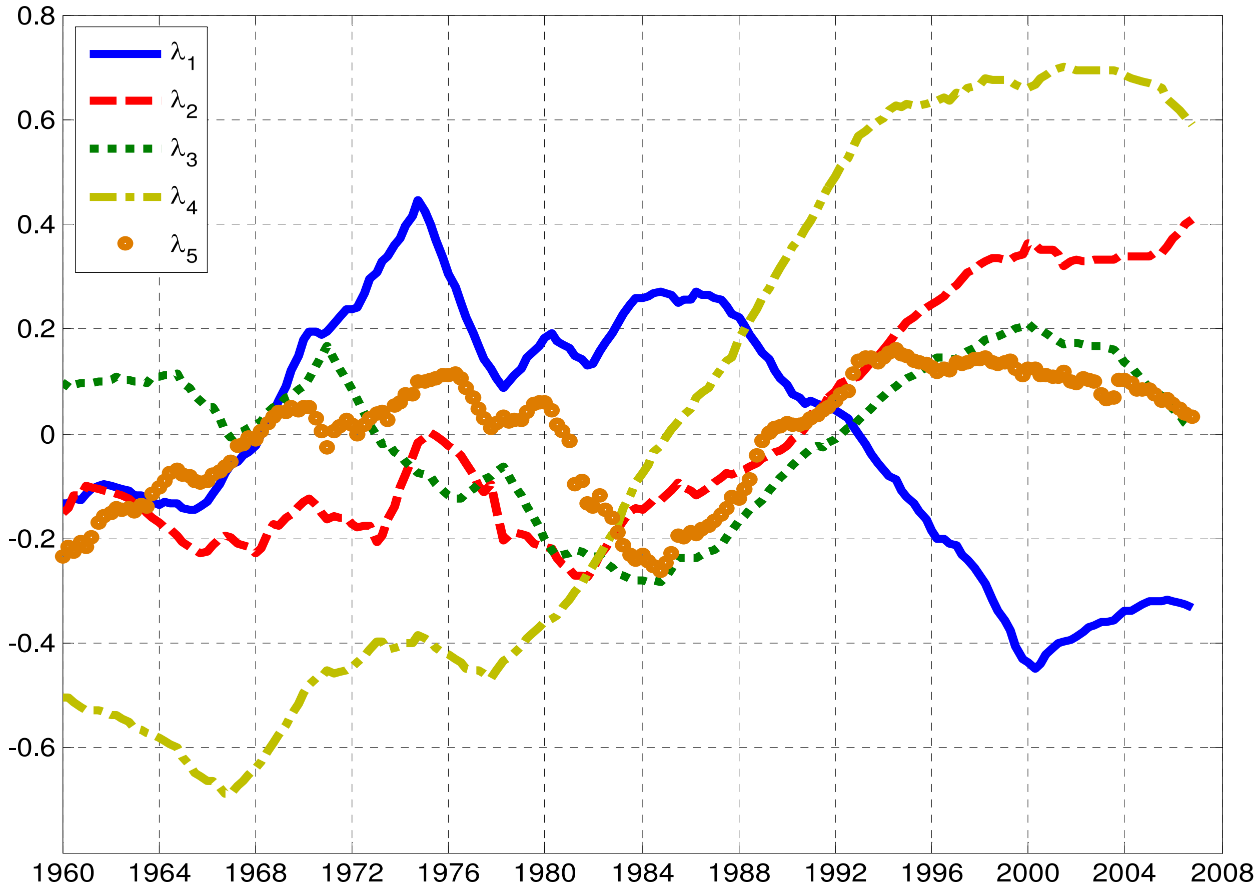


Figure 5: Impulse responses (medians) to a 25 basis point innovation to the Fed Funds Rate, and differences between the responses for three time periods from Benchmark FAVAR(4,2): Inflation (GDP deflator) (top) and Unemployment (bottom

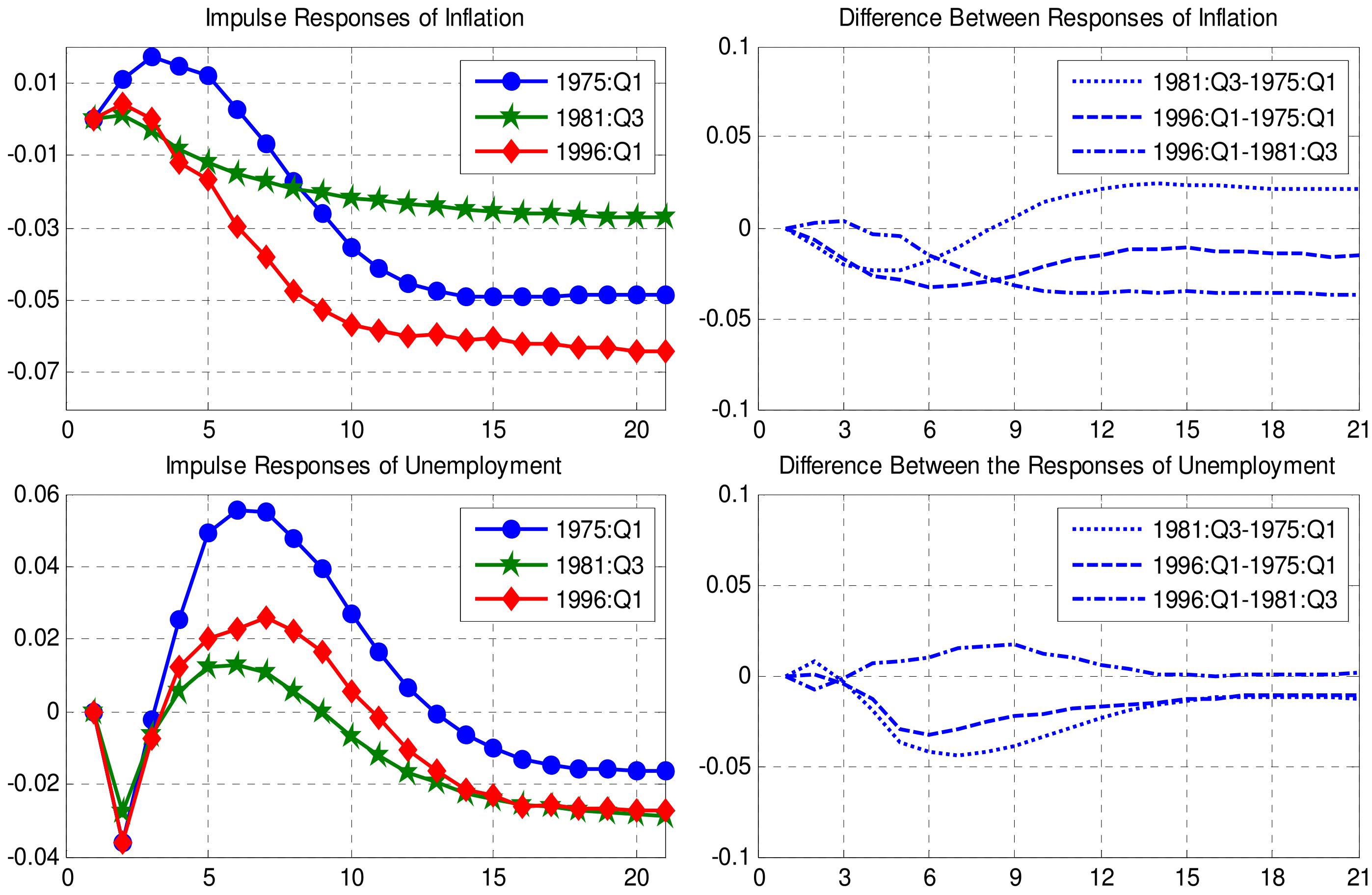


Figure 6: Impulse Responses from Benchmark FAVAR(4,2): i) 1975:Q1 (blue line), ii) 1981:Q3 (green line) and iii) 1996:Q1 (red line)

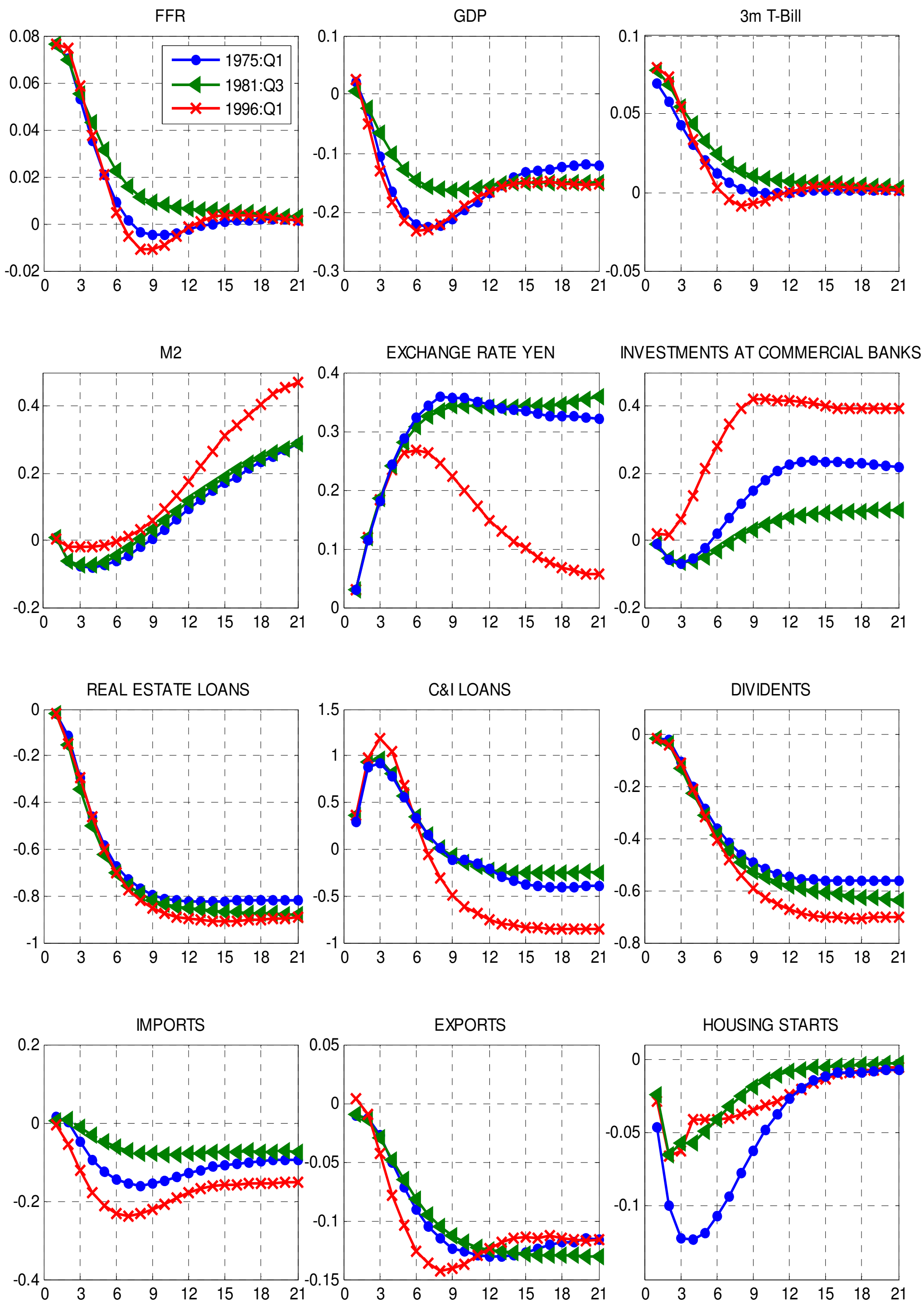
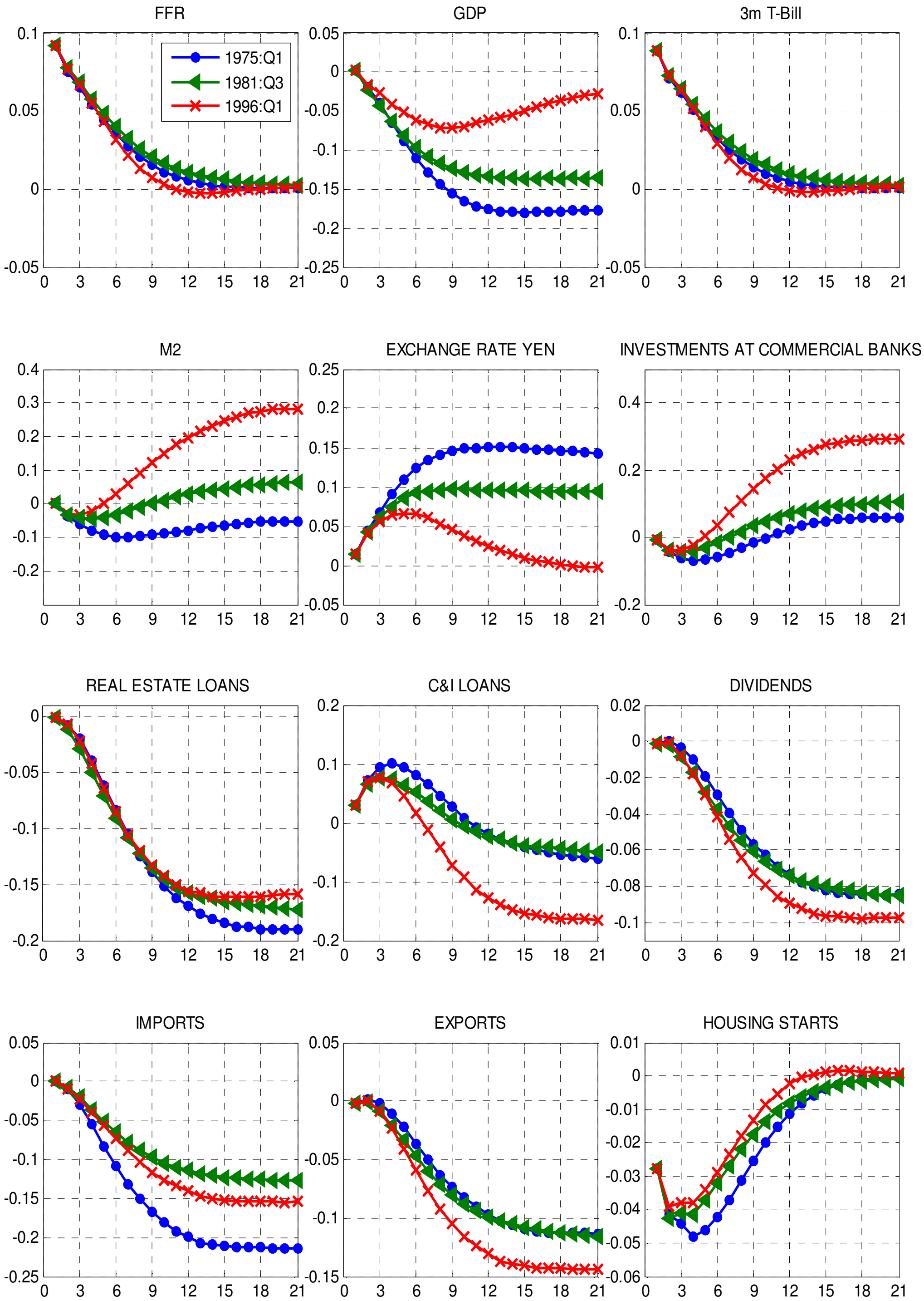


Figure 7: Impulse Responses from TVP-FAVAR (3,2) with constant loadings: i) 1975:Q1 (blue line), ii) 1981:Q3 (green line) and iii) 1996:Q1 (red line)



Technical Appendix

The basic model consists of Equations (2) to (8), but here for the purpose of estimation and inference I will use the formulation in Equations (9)-(10). Writing the vectors $y_t = \begin{bmatrix} f_t \\ r_t \end{bmatrix}$ and $g_t = \begin{bmatrix} x_t \\ r_t \end{bmatrix}$ analytically, and stacking the autoregressive coefficients in one vector, Equations (9) - (10) can be rewritten in the following form

$$\begin{bmatrix} x_t \\ r_t \end{bmatrix} = \lambda_t \begin{bmatrix} f_t \\ r_t \end{bmatrix} + W_t \epsilon_t^g \tag{T.1}$$

$$\begin{bmatrix} f_t \\ r_t \end{bmatrix} = Z_t' B_t + A_t^{-1} \Sigma \epsilon_t^y \tag{T.2}$$

where $\lambda_t = \begin{bmatrix} \lambda_t^f & \lambda_t^r \\ 0_{1\times k} & 1 \end{bmatrix}$ a $(n\times 1)\times(k\times 1)$ matrix, $W_t = diag(\exp(h_{1,t})/2, \ldots, \exp(h_{n,t})/2, 0)$, $B_t = (b_{1,t}', \ldots, b_{p,t}')$, $Z_t' = I_{k+1} \otimes \left[\begin{bmatrix} f_{t-1} \\ r_{t-1} \end{bmatrix}', \ldots, \begin{bmatrix} f_{t-p} \\ r_{t-p} \end{bmatrix}' \right]$ with the symbol $\otimes$ denoting the Kronecker product. Note that (T.1) now has $(n \times 1)$ equations in total (n columns of x_t and one column of r_t) even though the equations of interest are only the first n ones. The last equation has known parameters ($r_t = 0 \cdot f_t + 1 \cdot r_t + 0 \cdot \epsilon_{n+1,t}^g$), and it is used only to help write the model (T.1) - (T.2) as a VAR system (notice that y_t now enters both the R.H.S of (T.1) and the L.H.S. of (T.2), so we can easily plug in the second eqution into the first one, as shown in the main text). The drifting parameters evolve according to (8)

$$\begin{aligned}
\lambda_{i,t} &= \lambda_{i,t-1} + J_{i,t}^{\lambda} \eta_t^{\lambda} \\
h_{i,t} &= h_{i,t-1} + J_{i,t}^{h} \eta_t^{r} \\
B_t &= B_{t-1} + J_t^B \eta_t^B \\
\alpha_t^{block\ 1} &= \alpha_{t-1}^{block\ 1} + J_t^{\alpha^{block\ 1}} \eta_t^{\alpha^{block\ 1}} \\
\alpha_t^{block\ 2} &= \alpha_{t-1}^{block\ 2} + J_t^{\alpha^{block\ 2}} \eta_t^{\alpha^{block\ 2}} \\
&\vdots \\
\alpha_t^{block\ k} &= \alpha_{t-1}^{block\ k} + J_t^{\alpha^{block\ k}} \eta_t^{\alpha^{block\ k}} \\
\log \sigma_t &= \log \sigma_{t-1} + J_t^{\sigma} \eta_t^{\sigma}
\end{aligned} \tag{T.3}$$

with errors $\eta_t^\theta \sim N(0, Q_\theta)$, $\theta = \lambda_i, h_i, B, \alpha_t^{block\ 1}, \alpha_t^{block\ 2}, \ldots, \alpha_t^{block\ k}, \log \sigma$ where for simplicity I drop the subscript t in the notation. Formulation (T.3) is different than equation (8), because the $\frac{(k+1)\times k}{2}$ elements in the parameter vector α_t are sampled in k blocks[1]. Each block has elements $\alpha_t^{block\ 1} = \{a_{21,t}\}$, $\alpha_t^{block\ 2} = \{a_{31,t}, a_{32,t}\}$, ..., $\alpha_t^{block\ \mathrm{k}} = \{a_{(k+1)1,t}, ..., a_{(k+1)k,t}\}$ representing the elements in the k rows of the matrix A_t, as these are defined in equation (7). The implication of the block sampling scheme is that now there is a not a unique break index $J_{A,t}$ for the lower triangular matrix A_t, but k different indexes $J_{\alpha^{block\ \mathrm{j}},t}$ for each block of elements $\alpha_t^{block\ \mathrm{j}}$, $j = 1, ..., k$. Recall also that the parameters in (T.1) are drawn independently from each i univariate equation, $i = 1, ..., n$, due to the diagonality assumption of the covariance matrix H_t. This allows to define an index J_t for each row $\lambda_{i,t}$ of λ_t (which is denoted as $J_{i,t}^{\lambda}$ in (T.3)), and that way model more complex dynamics (see also the main text).

Priors and sensitivity

The initial state of the time varying parameters (at time $t = 0$) is drawn randomly from Normal distributions of the form:

$$\theta_0 \sim N\left(\underline{m}_\theta, \underline{V}_\theta\right)$$

Note that uninformative priors are based on the choice $\underline{m}_\theta = 0$, $\underline{V}_\theta = 4I$ for all $\theta = \lambda_i, h_i, B, \alpha_t^{block\ 1}, \alpha_t^{block\ 2}, \ldots, \alpha_t^{block\ \mathrm{k}}, \log\sigma$, $i = 1, ..., n$. As it is the usually the case in state-space models, there is little sensitivity to the choice of the initial values for the state variables (the time varying parameters θ). This is true especially since the final parameter estimates are the smoothed ones which utilise the information in the whole sample, not just time t. In time-varying paramters models in general, the crucial choice is the prior on the covariance of the state variables. In that respect, I use conjugate priors but the hyperparameters are fine-tuned following Primiceri (2005) for reasons explained in that paper. An inverse-Wishart prior is placed on the covariance matrices, $Q_{\theta_{/-h}}$, where I define $\theta_{/-h} = \lambda_i, B, \alpha_t^{block\ 1}, \alpha_t^{block\ 2}, \ldots, \alpha_t^{block\ \mathrm{k}}, \log\sigma$. Each h_i is scalar, so that the inverse-Gamma prior is defined instead. The prior densities and hyperparameters are defined as

$$\begin{aligned} Q_{\theta_{/-h}} &\sim iW(l_{\theta_{/-h}} \cdot (1 + n_{\theta_{/-h}}) \cdot V_{\theta_{/-h}}^{OLS}, 1 + n_{\theta_{/-h}}) \\ Q_h &\sim iG(l_h \cdot (1 + n_h) \cdot V_h^{OLS}, 1 + n_h) \end{aligned}$$

[1]Since the factors are known (principal component estimates), equations (T.1) - (T.2) are estimated independently. Subsequently (T.2) can be estimated using standard methods developed in the TVP-VAR literature. Here I follow Primiceri (2005) and the reader should consult the appendix of this paper for full estimation details. Also (T.1) is estimated equation by equation, due to the diagonality assumption of the covariance matrix W_t,which essentially reduces estimation to n time-varying parameter regression.

where n_θ denotes the number of elements on each state vector θ, V_θ^{OLS} is the variance of the OLS estimate for θ, l_θ are tunning constants whose values are based on Primiceri (2005, Section 4.4), and $iW(a,b)$ and $iG(a,b)$ denote the inverse-Wishart and inverse-Gamma distributions respectively with scale parameter a and shape parameter b.

The jump variables J_t^θ come from a Bernoulli density. An initial 'guess' has to be made for the value of J_1^θ, i.e. when time $t = 1$, while the states in the subsequent periods are updated using the Gerlach et al. (2000) algorithm. Since there is no prior information to set the initial condition J_1^θ, either the assumption that all the parameters remain constant in the first period ($J_1^\theta = 0$) or that all parameters change in the first period ($J_1^\theta = 1$) may be used. It turns out that these two initial 'guesses' imply observationally equivalent models, since the posterior results are not affected by this choice. The prior densities are formulated as

$$J_t^\theta \sim Bernoulli(\pi_\theta)$$

The only hyperpamater associated with each Bernoulli prior is its respective probability, denoted by π_θ. One extra hierarchical layer is introduced in the model by placing a prior density on this hyperparameter. It is easy to prove that a conjugate prior density for π_θ in this instance is the Beta density, which gives

$$\pi_\theta \sim Beta(\tau_0, \tau_1)$$

where for simplicity it is assumed that all probabilities share the same prior values (τ_0, τ_1), i.e. they are common for all parameters $\theta = \lambda_i, h_i, B, \alpha_t^{block\ 1}$, $\alpha_t^{block\ 2}, \ldots, \alpha_t^{block\ \mathrm{k}}, \log\sigma$. The reference prior for the Beta distribution is $\tau_0 = \tau_1 = \frac{1}{2}$.

Posterior Analysis Using the Gibbs Sampler

Conditional on using the well-defined conjugate priors of the previous section and a Kalman filter/smoother, it is easy to construct a Gibbs sampler which will converge to the true posterior densities of the parameters. Note that in my implementation of the Kalman filter/smoother I used the Carter and Kohn (1994) method, but other efficient methods can be used like the Durbin and Koopman (2002) or DeJong and Shephard (1995). The reader can see that conditional on each time period t, the TVP-FAVAR model collapses to the constant parameters FAVAR and simple regression model arguments apply in this case. For the sake of brevity full formulae are not presented here - the TVP-FAVAR model has many parameters to attempt providing full details. Instead I will give direct references to journal articles for more information. Note that MATLAB code is provided by the author so that the reader can replicate the results and use the TVP- models in his/her own research.

Given the starting values for the parameters, and the final rotated solution for the factors, the Gibbs sampler makes M cycles where we consecutively sample from the following conditional densities:

- Draw θ_t using the CK algorithm. In the special case of the volatilities h, σ we must use the KSC algorithm
- Draw $Q_{\theta_{/-h}} \sim iW(\rho_0^{\theta_{/-h}}, \rho_1^{\theta_{/-h}})$ and $Q_{h_i} \sim iG(\rho_0^{h_i}, \rho_1^{h_i})$, where $\rho_0^{\theta} = \left(\sum_{t=2}^{T} \left(\theta_t' - \theta_{t-1}'\right)' \left(\theta_t' - \theta_{t-1}'\right) + l_\theta \cdot (1 + n_\theta) \cdot V_\theta^{OLS}\right)^{-1}$ and $\rho_1^{\theta} = 1 + n_\theta + \sum_{t=1}^{T} J_t^{\theta}$
- Draw $J^\theta | data$ using the GCK algorithm
- Draw $\pi_\theta | data \sim Beta(\bar{\tau}_0^\theta, \bar{\tau}_1^\theta)$, where $\bar{\tau}_0^\theta = \tau_0 + \sum_{t=1}^{T} J_t^\theta$, and $\bar{\tau}_1^\theta = \tau_1 + T - \sum_{t=1}^{T} J_t^\theta$

Since the covariance in equation (T.2) is decomposed into a lower triangular and a diagonal matrix, further transformations of the model are required in order to ensure that each parameter is in appropriate state-space form. These are given for example in Koop et al. (2009, Section A.1.2)

The Carter and Kohn (1994) (CK) algorithm:

The first step in this algorithm is to use the prediction and update steps of the Kalman filter, usually called the "forward iterations" (for $t = 1$ to T, we update consecutively each θ_t conditional on data at time t). Then smoothing takes place, where we update again our estimates of θ_t using the information in future periods, i.e. we condition on the data observed at time $t+1$. Since for $t = T$ there are no future observations available (i.e. $T+1$), we keep θ_T fixed from the previous step, and start the "bavkward iterations" (i.e. for $t = T-1$ to 1, we update consecutively each θ_t conditional on $t+1$). Exact details are provided in Carter and Kohn (1994). Note only that here the state variance is $J_t^\theta \times Q_\theta$ which means that if for a certain t we have $J_t^\theta = 0$, then θ_t is not updated, but stays constant to its previous value, θ_{t-1}.

The Kim, Shephard and Chib (1998) (KSC) algorithm:

The Carter and Kohn algorithm can be used only in the case of λ_i, B, $\alpha_t^{block\ 1}$, $\alpha_t^{block\ 2}, \ldots, \alpha_t^{block\ k}$. In the case of the volatilities h_i, σ a modification of the above algorithm is needed. That is because the errors in the measurement equations are non-normal. Kim, Shephard and Chib (1998) use a mixture of normal approximation of the measurement errors, so that the state-space becomes conditionally Normal, and subsequently the Carter and Kohn algorithm can be applied. This means that some extra parameters have to be updated from the data, like an index variable $S_t|data$ which indexes each of the 7 mixture components. I avoid though full details and formulas, since exact details can be found in Kim et al. (1998), Koop et al. (2009) and Primiceri (2005).

The Gerlach, Carter and Kohn (2000) (GCK) algorithm:
Gerlach et al. (2000) point out that previous attempts to draw J^{θ} (see Carter and Kohn, 1996 and McCulloch andTsay,1993) can be highly innefective, since there may be high correlation between J^{θ} and the states θ_t. They propose an algorithm that draws from J_t^{θ} for each t, without conditioning on the states θ_t. Mathematically, the density they obtain draws from is $p\left(J_t^{\theta}|data, J_{s\neq t}^{\theta}\right)$ where $s, t = 1, \ldots, T$, and $J_{s\neq t}^{\theta}$ denotes all the elements of J^{θ} apart from J_t^{θ}. This density can be decomposed as in equation (3) of Gerlach et al. (2000) and then estimated using their algorithm, described in Section 3.2 of the same paper.