ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2015)
178, *Part* 3, *pp.* 493–533

Classical time varying factor-augmented vector auto-regressive models—estimation, forecasting and structural analysis

Sandra Eickmeier,

Deutsche Bundesbank, Frankfurt am Main, Germany, and Centre for Applied Macroeconomic Analysis, Canberra, Australia

Wolfgang Lemke

European Central Bank, Frankfurt am Main, Germany

and Massimiliano Marcellino

Bocconi University, Milan, Innocenzo Gasparini Institute for Economic Research, Milan, Italy, and Centre for Economic Policy Research, London, UK

[Received November 2012. Final revision March 2014]

Summary. We propose a classical approach to estimate factor-augmented vector auto-regressive (FAVAR) models with time variation in the parameters. When the time varying FAVAR model is estimated by using a large quarterly data set of US variables from 1972 to 2012, the results indicate some changes in the factor dynamics, and more marked variation in the factors' shock volatility and their loading parameters. Forecasts from the time varying FAVAR model are more accurate, in particular over the global financial crisis period, than forecasts from other benchmark models. Finally, we use the time varying FAVAR model to assess how monetary transmission to the economy has changed.

Keywords: Factor-augmented vector auto-regressive models; Forecasting; Monetary transmission; Time varying parameters

1. Introduction

The recent macroeconometric literature has seen an increasing interest in the application of factor-augmented vector auto-regressive (FAVAR) models for forecasting and structural analysis. They provide means to exploit a large information set and handle the omitted variable problem that is often encountered in standard vector auto-regressive (VAR) models. FAVAR models were originally suggested by Bernanke *et al*. (2005), who modelled a large number of variables as the sum of a common component and an idiosyncratic component. The common component of a variable is the product of a few common factors and variable-specific factor loadings. The factors, which are the driving forces underlying most economic variables, are assumed to follow a VAR process.

Another recent strand of literature has focused on small models with time varying parameters, including evolving variances, to take into consideration the changing sources and sizes of shocks

Address for correspondence: Sandra Eickmeier, Economic Research Centre, Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, D-60431 Frankfurt am Main, Germany.
E-mail: sandra.eickmeier@bundesbank.de

 0964–1998/15/178493

explicitly, and their transmission to the economy; see for example Cogley and Sargent (2005) and Sims and Zha (2006).

A few references have attempted to combine the FAVAR and the time varying parameter approaches, introducing FAVAR models with time varying parameters, and hence combining the benefits of using many variables and allowing for a time varying model structure. Examples include Baumeister *et al.* (2013) and Korobilis (2013), whose applications concern the transmission mechanism of monetary policy in the USA, as well as Del Negro and Otrok (2008), Liu *et al.* (2011) and Mumtaz and Surico (2012), who fitted time varying FAVAR (TVFAVAR) models to study international business cycle and inflation comovements. A common feature of all these contributions is the use of Bayesian procedures. Instead, in this paper we propose a fully classical approach to estimate a FAVAR model with time varying parameters. Our time varying version is fairly flexible, as it can accommodate smooth changes in the factor loadings, in the auto-regressive (AR) coefficients of the factor VAR process, in the contemporaneous relationships between the factors and in the volatility of the common shocks.

We propose to estimate the TVFAVAR process in two stages. The first stage involves estimating the factors as principal components (PCs). As argued by Stock and Watson (2002a, 2008), Banerjee *et al.* (2008) and Bates *et al.* (2013), the PC estimator is consistent for the factors even if the loadings mildly vary over time. The second stage involves estimating the time varying loading coefficients and the AR matrices of the factor VAR process as well as the time varying variances and correlations. Treating the estimated factors as given, the relationships between the observable variables and the factors are represented as a set of univariate regression models with time varying parameters, which evolve as independent random walks. As such, the model is estimated equationwise by converting each equation into state space form, estimating the hyperparameters by maximum likelihood and applying the Kalman filter to back out the time varying parameter paths; see for example Nyblom (1989). Regarding the time varying factor VAR process we employ a representation with a lower triangular matrix of contemporaneous relationships, which renders the VAR equations conditionally independent. This again enables us to estimate the model equationwise, applying standard methods for univariate regression models with time varying parameters. Concerning the volatility specification, we deviate from the common assumption in the literature that volatility is driven by an additional latent factor. We rather specify it as an (exponentially affine) function of lagged factors, which makes our VAR equations conditionally linear. The resulting estimated pattern of volatility is similar to that returned by models where time varying volatility is captured by additional latent variables. Moreover, we think that linking the evolution of volatility to the underlying economic forces, namely the factors, is a sensible modelling choice. Monte Carlo (MC) experiments indicate that our estimation procedure performs well, in particular with a large cross-sectional dimension (that improves factor estimation) and temporal dimension (that reduces the extent of the 'pile-up' problem that was identified by Stock and Watson (1998)).

As an empirical example, we fit our TVFAVAR model to a new large quarterly US data set with more than 300 macroeconomic and financial variables, observed between 1972 and 2012. Our estimation results imply substantial time variation in the variance of the shocks but also in their transmission mechanism, as represented by the factor loadings and factor dynamics. However, time variation is 'sparse' in the sense that changes in only a few parameters govern the time variation of the system, whereas most parameters turn out to be essentially constant over time.

We then use the model to produce out-of-sample forecasts of various macroeconomic and financial variables. The use of a time varying model for a large data set in a forecasting context is another original feature of our paper. In general, it turns out that for several variables and forecast horizons the TVFAVAR model is more accurate than a constant parameter FAVAR model. The latter performs well since recursive estimation introduces anyway a form of parameter time variation. However, the TVFAVAR model still dominates for many monetary and financial variables and performs particularly well during the latest crisis period (which we identify with the period from 2007, quarter 3, to 2009, quarter 4). Overall, these results suggest that allowing for changes in the cross-variable relationships and in the shock volatility can be quite helpful for forecasting macroeconomic and financial indicators, in particular during crisis periods.

Finally, we contribute to the growing literature on time variation in the monetary transmission mechanism by identifying monetary policy shocks and assessing their transmission to the US economy over time between 1972 and 2007.

Boivin *et al.* (2010) have comprehensively overviewed the existing literature and have shown that a consensus on how the monetary transmission mechanism in the USA has evolved is still lacking. The results highlight interesting patterns of time variation. In particular, the volatility of the monetary shocks is substantially smaller after the early 1980s. The negative effect of a same-sized shock on most activity and price measures has declined over time. The effects on activity variables do not appear to be different during recessionary phases compared with expansions. Finally, the negative effect of monetary policy shocks on inflation expectations and long-term interest rates has weakened over time. This could be due to changes in the conduct of monetary policy or to globalization and may have contributed to the decline in the effect on activity and prices.

The paper is organized as follows. In Section 2 we present the model and estimation methodology and compare our approach with related TVFAVAR models. In Section 3 we provide MC evidence on the finite sample performance of our estimation procedure. In Section 4 we present the data. In Section 5 we fit the TVFAVAR model to the data and present evidence on time variation in the parameters. In Section 6 we evaluate the forecasting performance of the TVFAVAR model. In Section 7 we assess changes in the monetary transmission mechanism in the USA over time. Finally, in Section 8 we summarize the main results and conclude.

2. The time varying factor-augmented vector auto-regressive model: representation and estimation

In this section we introduce the TVFAVAR model, discuss its estimation and compare it with related approaches.

2.1. The time varying factor-augmented vector auto-regressive model

Our starting specification is the FAVAR model as proposed by Bernanke *et al.* (2005). Let $X_t' = (x_{1,t}, \ldots, x_{N,t})$ denote a large vector of N zero-mean stationary variables, for $t = 1, \ldots, T$, where both N and T can go to ∞. In the standard dynamic factor model, each element of X_t is assumed to be the sum of a linear combination of G common factors $F_t' = (f_{1,t}, \ldots, f_{G,t})$ and an idiosyncratic component $e_{i,t}$. Hence,

$$x_{i,t} = \Lambda_i' F_t + e_{i,t}, \qquad i = 1, \ldots, N, \tag{2.1}$$

where $e_t' = (e_{1,t}, \ldots, e_{N,t})$. We assume that the factors are orthonormal and uncorrelated with the idiosyncratic errors, and $E(e_t) = 0$ and $E(e_t e_t') = R$, where R is a diagonal matrix. These assumptions identify the model and are common in the FAVAR literature. They can be partly relaxed when the goal of the analysis is purely factor estimation by means of non-parametric methods; see for example Stock and Watson (2002a, b). Some correlation across the idiosyncratic errors could be allowed also in a parametric context by proper specification of the matrix R. However, in a time series context, it is perhaps more interesting to allow for temporal correlation in $e_{i,t}$, and we consider this case later.

The dynamics of the factors are then modelled as a VAR(p) process:

$$F_t = B_1 F_{t-1} + \ldots + B_p F_{t-p} + w_t, \qquad E(w_t) = 0, \quad E(w_t w_t') = W. \tag{2.2}$$

Since each $x_{i,t}$ is assumed to be a zero-mean process (and the respective data are demeaned), equations (2.1) and (2.2) do not contain intercepts.

The VAR equation (2.2) can be interpreted as a reduced form representation of a system of the form

$$PF_t = \mathcal{K}_1 F_{t-1} + \ldots + \mathcal{K}_p F_{t-p} + u_t, \qquad E(u_t) = 0, \quad E(u_t u_t') = S, \tag{2.3}$$

where P is lower triangular with 1s on the main diagonal, and S is a diagonal matrix. The relationship to the reduced form parameters in expression (2.2) is $B_i = P^{-1}\mathcal{K}_i$ and $W = P^{-1}SP^{-1\prime}$. This system of equations may in other contexts be referred to as a 'structural VAR' representation. Although we actually use a triangular contemporaneous relationship in our structural analysis in Section 7, we emphasize that the chosen representation (2.3) is particularly useful for estimating the time varying version that is outlined below, but after estimation of the system matrices other forms of shock identification besides the specific triangular one may be applied.

Having introduced the standard FAVAR model with a constant parameter structure, we now relax the assumption of parameter constancy in four dimensions. Specifically, we allow for time variation in

(a) the AR dynamics of the factors ($\mathcal{K}_1, \ldots, \mathcal{K}_p$),
(b) the contemporaneous relationships captured by the matrix P,
(c) the variances of factor innovations, i.e. the elements of S in expression (2.3), and
(d) the factor loadings in expression (2.1).

Thus, we consider the following time varying version of expressions (2.1) and (2.3):

$$x_{i,t} = \Lambda_{i,t}' F_t + e_{i,t}, \qquad i = 1, \ldots, N, \tag{2.4}$$

and

$$P_t F_t = \mathcal{K}_{1,t} F_{t-1} + \ldots + \mathcal{K}_{p,t} F_{t-p} + u_t, \qquad E(u_t) = 0, \quad E(u_t u_t') = S_t, \tag{2.5}$$

where again the factors are orthonormal and uncorrelated with the idiosyncratic errors, P_t is lower triangular with 1s on the main diagonal and S_t is diagonal. In addition, we specify the idiosyncratic components in expression (2.4) to follow a first-order AR process:

$$e_{i,t} = \rho_i e_{i,t-1} + \xi_{i,t}, \qquad E(\xi_{i,t}) = 0, \quad E(\xi_{i,t}^2) = \sigma_i^2, \quad i = 1, \ldots, N. \tag{2.6}$$

Again, the elements of $\xi_t \equiv (\xi_{1,t}, \ldots, \xi_{N,t})'$ are assumed to be contemporaneously uncorrelated.

Let the time varying parameters $\{P_t, \mathcal{K}_{1,t}, \ldots, \mathcal{K}_{p,t}, \Lambda_{1,t}, \ldots, \Lambda_{N,t}\}$ be collected in a vector α_t. Note that the dimension of this vector is $G(G-1) \times 0.5 + pG^2 + NG$, which can be fairly large. As is common in time-varying parameter regression models (see for example Nyblom (1989)), we assume that the parameters vary slowly over time, as independent random walks

$$\alpha_t = \alpha_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, Q), \tag{2.7}$$

where Q is a diagonal matrix—again, a standard assumption in the literature on time varying parameter models. More specifically it is

$$Q = \begin{pmatrix} Q_P & 0 & 0 \\ 0 & Q_K & 0 \\ 0 & 0 & Q_\Lambda \end{pmatrix},$$

and the elements of Q_P and Q_K are given by q^2_{Pi} and q^2_{Kj}, $i = 1, \ldots, G(G-1) \times 0.5$ and $j = 1, \ldots, pG^2$, whereas those of Q_Λ are $q^2_{\Lambda i}$, $i = 1, \ldots, NG$.

Finally, all elements of $(\xi_t, u_t, \varepsilon_t)$ are assumed to be uncorrelated contemporaneously and over time.

Our TVFAVAR specification is fairly parsimonious, in the sense that the number of parameters governing the innovation variances of the time varying parameters equals the number of parameters in constant parameter FAVAR models. Moreover, our time varying model nests the standard constant parameter FAVAR model, since when all the elements of the Q-matrix are equal to 0 the former reduces to the latter.

We estimate the VAR model and the factor loading relationships equation by equation. As we discuss in Section 2.2, this is possible as each of these equations with time varying parameters can be cast into a linear Gaussian state space model. The crucial point is how to model time variation in factor innovation volatility: if it were assumed to be governed by another latent process, say q_t, such that $S_{t,gg} = \exp(q_t)$ and $q_t = a_i + \phi_i q_{t-1} + \zeta_{i,t}$, this would make the model non-linear in the state vector, preventing estimation based on linear Gaussian state space models, and requiring linear approximation approaches or simulation-based methods. In addition, as the factors F_t are assumed to represent the main driving forces of the economy, they may be considered a natural choice for the drivers of volatility as well.

Owing to these considerations, we assume volatility to be a function of lagged factors F_{t-1}. This guarantees that each single VAR equation with time varying parameters and such-specified time varying innovation volatility can be represented by a linear (conditionally) Gaussian state space model. To be specific, for each of the VAR equations we write innovation volatility as an exponential affine function of the last period's factors:

$$S_{gg,t} = \exp(c_g + b'_g F_{t-1}), \qquad g = 1, \ldots, G. \tag{2.8}$$

Obviously, if $b_g = 0$ we are back to the homoscedastic case. When only the gth element of b_g differs from 0, innovation volatility for factor g depends solely on the lagged levels of this factor.

The approach can be modified by allowing exogenous variables to be determinants of volatility; for an application, see Eickmeier *et al.* (2011a). Moreover, instead of the exponential affine specification, volatility may be modelled as a function of squared past changes in variables, or other functional forms can be chosen. Finally, generalized AR conditional heteroscedasticity disturbances may be accommodated by using the approximation approach by Harvey *et al.* (1992).

We shall see that empirically this approach produces volatility estimates that are in line with those generated by models with additional latent variables capturing the time variation in volatility.

2.2. *Estimating the time varying factor-augmented vector auto-regressive model*

The elements of F_t are estimated as the first G PCs of X_t, with the identifying assumption $E(F_t F'_t) = I$. We then treat the factors as observable, which is justified when N grows faster than $T^{0.5}$ (see Bai and Ng (2006)) and estimate the time varying parameter factor VAR and the loading equations. Note that, as argued by Stock and Watson (2002a, 2008), Banerjee *et al.* (2008) and Bates *et al.* (2013), the factors are still estimated consistently even if there is some time variation in the loading parameters. The intuition underlying this result is that factor estimates at time t are weighted averages of the N x_i-variables at time t only. We come back to this issue in Section 5.1, when presenting the empirical results.

Regarding the cross-sectional relationships, we put each of the N equations (2.4) into state space form. For the ith equation the state vector is $\tilde{\alpha}_t^{(i)} = (\Lambda'_{it}, e_{it})'$. Since the idiosyncratic component in expression (2.4) follows an AR(1) process, rather than being white noise, it becomes part of the state vector besides the time varying loading parameters. The transition equation is given by

$$\tilde{\alpha}_t^{(i)} = \Phi_i \tilde{\alpha}_{t-1}^{(i)} + \tilde{\varepsilon}_t^{(i)},$$

where $\Phi_i = \text{diag}\{(\mathbf{1}_G, \rho_i)\}$, $\tilde{\varepsilon}_t^{(i)} = (\varepsilon_t^{(i)}, \xi_{it})'$, where $\varepsilon_t^{(i)}$ are the respective elements of ε_t in expression (2.7), and hence, $E(\tilde{\varepsilon}_t^{(i)}) = 0$, and $E(\tilde{\varepsilon}_t^{(i)} \tilde{\varepsilon}_t^{(i)'}) = \text{diag}\{(q^{(i)}, \sigma_i^2)\}$, i.e. $q^{(i)}$ contains the random-walk innovation variances of the time varying Λ-parameters (i.e. the respective elements of Q in expression (2.7)) and σ_i^2 is the innovation variance of the idiosyncratic component process. The measurement equation is

$$x_{i,t} = Z_t \tilde{\alpha}_t^{(i)} \tag{2.9}$$

where $Z_t = (F'_t, 1)$. We estimate the $G+2$ hyperparameters $(\rho_i, q^{(i)}, \sigma_i)$ of the ith loading equation by maximum likelihood. We then back out the path of time varying loading parameters by using the Kalman smoother.

Since expression (2.7) preserves the independence between the parameters of the G equations of expression (2.5), we can likewise estimate the time varying parameters that are contained in the P_t- and $\mathcal{K}_{i,t}$-matrices equationwise. For the gth equation in state space form, the state vector containing the time varying parameters is given by

$$\alpha_t^{g\prime} = (-P_{g,1,t}, \ldots, -P_{g,g-1,t}, \mathcal{K}_{g,1,1,t}, \ldots, \mathcal{K}_{g,G,1,t}, \mathcal{K}_{g,1,2,t}, \ldots, \mathcal{K}_{g,G,2,t}, \ldots, \mathcal{K}_{g,1,p,t}, \ldots, \mathcal{K}_{g,G,p,t}),$$

where, for $g = 1$, there are no P-parameters showing up. Note that, owing to the different number of elements coming from the triangular P-matrix, the dimensions of the state vectors are different for each of the G equations.

The state equation is the random walk for α_t^g:

$$\alpha_t^g = \alpha_{t-1}^g + \varepsilon_t^g, \qquad \varepsilon_t^g \sim N(0, Q_g), \quad Q_g = \text{diag}(q_g). \tag{2.10}$$

The measurement equation is given by

$$f_{g,t} = f_t^{g\prime} \alpha_t^g + u_{g,t}, \qquad u_{g,t} \sim N(0, S_{gg,t}), \tag{2.11}$$

where

$$f_t^{g\prime} = (f_{1,t}, \ldots, f_{g-1,t}, f_{1,t-1}, \ldots, f_{G,t-1}, f_{1,t-2}, \ldots, f_{G,t-2}, \ldots, f_{1,t-p}, \ldots, f_{G,t-p}),$$

and $S_{gg,t}$ is given by expression (2.8).

In the first step, we estimate for each equation the hyperparameters (q_g, c_g, b_g) by maximum likelihood. In the second step, we filter out the time varying parameters of each equation by the Kalman filter. For starting the filter, we adopt the frequently used strategy of initializing the time varying parameters with their ordinary least squares estimates. Alternatively, initialization could be based on a diffuse prior approach (as we specify random-walk dynamics for parameters).

However, when taking the filtered states $a_{t|t}^1, \ldots, a_{t|t}^G$ from each equation and reconstructing the respective VAR matrices $P_t, \mathcal{K}_{1,t|t}, \ldots, \mathcal{K}_{p,t|t}$, the resulting local VAR dynamics at time t may imply explosive behaviour. To avoid this we ensure that, at each point in time, all eigenvalues of the auto-regressive matrix corresponding to the reduced form VAR representation in companion form are inside the unit circle. To achieve this, we run the following restricted filtering algorithm, instead of G independent and unrestricted Kalman filters. In essence, the algorithm runs the G Kalman filters and performs an updating step only if the VAR structure that is implied by the filtered states jointly satisfies the stationarity condition.

Let Γ denote the mapping from the family of estimated state vectors $\{a^1_{t|t}, \ldots, a^G_{t|t}\} =: \mathcal{A}_{t|t}$ into the respective VAR matrices $P_{t|t}, \mathcal{K}_{1,t|t}, \ldots, \mathcal{K}_{p,t|t}$. The algorithm ($G$ Kalman filters with joint non-linear restrictions on filtered states) runs as follows.

Step 1: maximize the likelihood that is associated with each of the G state space models (2.10) and (2.11), and obtain the estimates $(\hat{q}_g, \hat{c}_g, \hat{b}_g)$ of (q^g, c_g, b_g), $g = 1, \ldots, G$.
Step 2: given the hyperparameters, initialize the G state space models by some $\mathcal{A}_0$ such that $\{P_0, \mathcal{K}_{1,0}, \ldots, \mathcal{K}_{p,0}\} = \Gamma(\mathcal{A}_0)$ implies a VAR structure without explosive eigenvalues. Set the set of corresponding variance–covariance matrices of initial states $\{\Sigma^1_0, \ldots, \Sigma^G_0\} =: \mathcal{S}_0$. Set $t-1=0$, $\mathcal{A}_{t-1|t-1} = \mathcal{A}_0$ and $\mathcal{S}_{t-1|t-1} = \mathcal{S}_0$.
Step 3: for each of the G state space models do a Kalman filter prediction step, i.e. compute

$$\begin{aligned} a^g_{t|t-1} &= a^g_{t-1|t-1}, \\ \Sigma^g_{t|t-1} &= \Sigma^g_{t-1|t-1} + \hat{Q}^g, \\ \hat{f}_{g,t|t-1} &= f^{g\prime}_t a^g_{t|t-1}, \\ D^g_t &= f^{g\prime}_t \Sigma^g_{t|t-1} f^g_t + \hat{S}_{gg,t} \end{aligned}$$

for $g = 1, \ldots, G$.
Step 4: for each of the G state space models, do a Kalman updating step, i.e.

$$\begin{aligned} K^g_t &= \Sigma^g_{t|t-1} f^g_t D^{g^{-1}}_t, \\ a^g_{t|t} &= a^g_{t|t-1} + K^g_t (f_{g,t} - \hat{f}_{g,t|t-1}), \\ \Sigma^g_{t|t} &= \Sigma^g_{t|t-1} - K^g_t f^{g\prime}_t \Sigma^g_{t|t-1} \end{aligned}$$

for $g = 1, \ldots, G$.
Step 5: compute the corresponding VAR matrices $\{P_t, \mathcal{K}_{1,t|t}, \ldots, \mathcal{K}_{p,t|t}\} = \Gamma(\mathcal{A}_{t|t})$. If the VAR structure satisfies the non-explosiveness condition, set $t := t+1$ and go to step 3. If not, set $\mathcal{A}_{t|t} := \mathcal{A}_{t-1|t-1}, \mathcal{S}_{t|t} := \mathcal{S}_{t-1|t-1}$ and $t := t+1$ and go to step 3.

If an updating step is not performed owing to failure of the non-explosiveness condition, this does *not* mean that the respective states (parameters) will be stuck at their $(t-1)$-magnitudes henceforth. Rather, as new observations on the $f_{g,t}$ come in, an updating step may be feasible in the next or one of the following periods. For the initialization of the filter, we choose the ordinary least squares estimates taken over the whole sample and their respective variance–covariance matrices. They turn out to give rise to a VAR structure that satisfies the stationarity conditions. To obtain smoothed estimates of the time varying parameters we apply the standard Kalman (fixed interval) smoothing algorithm but based on the filtered estimates that have been obtained by the restricted filter in the first step. Although it is not guaranteed *per se* that the thus-constructed smoothed estimates satisfy the non-explosiveness conditions (even if the restricted *filtered* estimates satisfy them by construction), they do so in our empirical application.

2.3. Comparison with related approaches

Unlike the bulk of the existing literature on time varying FAVAR models, which employs Bayesian approaches, we estimate our model by classical (i.e. maximum likelihood) methods.

The likelihood-based approach (using the Kalman filter) is feasible and straightforward in our context, as we use a model representation that allows equation-by-equation estimation, where each equation with time varying parameters is represented as a linear state space model. It is important to note that the model could be likewise estimated by Bayesian methods. Conversely, many of the other time varying FAVAR models in the literature may be estimated by classical approaches, but these would require simulation-based techniques (just like their Bayesian counterparts) or linearizations. Hence, using a frequentist rather than a Bayesian approach here is not a consequence of the model structure *per se* but a convenient choice, as it allows for analytic rather than simulation-based estimation and does not require the specification of prior distributions on the model (hyper)parameters.

In addition, owing to the two-stage approach that was described above, our model is relatively flexible in the sense that it allows for various sources of parameter time variation. In previously employed models either only the factor loadings (Del Negro and Otrok, 2008; Liu *et al.*, 2011) or only the AR parameters of the VAR process on the factors (Baumeister *et al.*, 2013; Mumtaz and Surico, 2012) are allowed to vary over time, but not both as in our approach. An exception is Korobilis (2013), who also adopted a two-stage approach that was similar to ours where the first step involves estimating the factors as PCs and the second stage involves estimating the parameters with Bayesian methods. The two-step approach enables us to circumvent the problem of simultaneously identifying factors and loadings.

All the references cited above allowed for time varying volatility in the factors, and Baumeister *et al.* (2013), Liu *et al.* (2011), Mumtaz and Surico (2012) and Korobilis (2013) also allowed for time variation in the contemporaneous relationships across the factors. As described above, we also feature both sorts of variation, but changes in volatility are modelled differently and are explained by the evolution of the underlying economic forces rather than left unspecified.

Of the references listed above, Del Negro and Otrok (2008), Liu *et al.* (2011) and Mumtaz and Surico (2012) allowed for serial correlation in the idiosyncratic components, which we also do. In addition, Mumtaz and Surico (2012), Del Negro and Otrok (2008) and Korobilis (2013) allowed for time varying volatility in the idiosyncratic components, which our model does not allow for.

3. Finite sample properties of the estimation approach: results from a Monte Carlo simulation

We conduct a small simulation study to explore the properties of our estimation approach. A complete detailed analysis would require much more space, so the analysis here is necessarily constrained to selective aspects. In particular, we focus on the properties of our approach regarding estimation of

(a) the factors,
(b) the paths of time varying AR and loading parameters,
(c) the (time constant) idiosyncratic part of the loading equations and
(d) time varying volatility—for those data-generating processes (DGPs) where this applies.

We fix the cross-sectional dimension as $N=300$, focus on a FAVAR model with two factors only, i.e. $G=2$, and fix the lag length to $p=1$. Moreover, in the VAR process for the factors, we fix the matrix P_t (which was defined in Section 2 and governing contemporaneous relationships between the two factors) to be constant over time and equal to the identity matrix, i.e. $P_t \equiv I_2$. We vary the DGP underlying the MC study along three dimensions: first, constant *versus* time varying volatility; second, moderate *versus* high intensity of parameter variation; third, a small *versus* large number of time series observations. Overall, this makes eight different DGPs. For each DGP, we generate $M=1000$ realizations of 'observed' data $X_t' = (x_{1,t},\ldots,x_{N,t}),\, t=1,\ldots,T$, and then apply our methodology.

3.1. Data-generating processes

Specifically, the DGP of all our MC variants is given as follows. For the factor VAR process,

$$\begin{pmatrix} F_{1,t} \\ F_{2,t} \end{pmatrix} = \begin{pmatrix} \mathcal{K}_{11,t} & \mathcal{K}_{12,t} \\ \mathcal{K}_{21,t} & \mathcal{K}_{22,t} \end{pmatrix} \begin{pmatrix} F_{1,t-1} \\ F_{2,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}, \tag{3.1}$$

$$\begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \sim N\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} S_{1,t} & 0 \\ 0 & S_{2,t} \end{pmatrix} \right\},$$

where

$$S_{i,t} = \exp(c_{i,t} + b_{i,1} F_{1,t-1} + b_{i,2} F_{2,t-1}) \tag{3.2}$$

and

$$\mathcal{K}_{ij,t} = \mathcal{K}_{ij,t-1} + \varepsilon^{\mathcal{K}}_{ij,t}, \qquad \varepsilon^{\mathcal{K}}_{ij,t} \sim N(0, q^{\mathcal{K}}_{ij}), \quad i,j=1,2. \tag{3.3}$$

For the loading equations,

$$x_{i,t} = \Lambda_{i1,t} F_{1,t} + \Lambda_{i2,t} F_{2,t} + e_{i,t}, \tag{3.4}$$

where for the idiosyncratic components

$$e_{i,t} = \rho_i e_{i,t-1} + \xi_{i,t}, \qquad \xi_{i,t} \sim N(0, \sigma_i^2), \quad i=1,\ldots,N, \tag{3.5}$$

and for the time varying loadings

$$\Lambda_{ij,t} = \Lambda_{ij,t-1} + \varepsilon^{\Lambda}_{ij,t}, \qquad \varepsilon^{\Lambda}_{ij,t} \sim N(0, q^{\Lambda}_{ij}), \quad i=1,\ldots,N, \quad j=1,2. \tag{3.6}$$

For all eight variants of the DGP, we have $\rho_i = 0.3$ and $\sigma_i = 1.5$ for all i. For the constant volatility specifications, we have in equation (3.2)

$$c_i = -1.2, \qquad b_{ij} = 0, \qquad i,j=1,2, \tag{3.7}$$

whereas for the time varying volatility specifications

$$c_1 = -1.2, \qquad b_{11} = 0.15, \qquad b_{12} = 0.6, \qquad c_2 = -1.2, \qquad b_{21} = 0.3, \qquad b_{22} = 0.3. \tag{3.8}$$

For the specifications with moderate parameter innovation variability,

$$\begin{aligned} &q^{\mathcal{K}}_{11} = q^{\mathcal{K}}_{22} = 0.04^2, \qquad q^{\mathcal{K}}_{12} = q^{\mathcal{K}}_{21} = 0.01^2, \\ &q^{\Lambda}_{ij} = 0.1^2, \qquad i=1,\ldots,N, \quad j=1,2. \end{aligned} \tag{3.9}$$

For the specifications with high parameter innovation variability,

$$\begin{aligned} &q^{\mathcal{K}}_{11} = q^{\mathcal{K}}_{22} = 0.05^2, \qquad q^{\mathcal{K}}_{12} = q^{\mathcal{K}}_{21} = 0.02^2, \\ &q^{\Lambda}_{ij} = 0.3^2, \qquad i=1,\ldots,N, \quad j=1,2. \end{aligned} \tag{3.10}$$

Table 1 summarizes the eight DGPs used for the MC simulations.

For all MC runs and all DGPs, the parameter processes are initialized with fixed initial values,

Table 1. The eight DGPs for the simulations

Specification	*Volatility*	*Parameter variation intensity*	*T*
I	Constant	Moderate	150
II	Constant	Moderate	600
III	Constant	High	150
IV	Constant	High	600
V	Time varying	Moderate	150
VI	Time varying	Moderate	600
VII	Time varying	High	150
VIII	Time varying	High	600

$$\mathcal{K}_0 = \begin{pmatrix} 0.6 & 0.0 \\ 0.0 & 0.6 \end{pmatrix}, \qquad \Lambda_{ij,0} = 2, \qquad \text{for all } i, j, \tag{3.11}$$

and the factor process is initialized with $f_0 = (0, 0)$. To avoid dependence on initial conditions, there is always a 'burn-in' period of 200 observations that will be discarded for estimation, i.e. only the subsequent 150 or 600 observations are presented to the estimation approach.

3.2. Results

Overall, the results of the MC experiments suggest that our proposed estimation method for time varying parameter FAVAR models is reliable also in finite samples though, as expected, the performance improves with large T and a substantial amount of parameter time variation.

3.2.1. *Factors*

For each MC run, we z-transform the 'observed' data $X'_t = (x_{1,t}, \ldots, x_{N,t})$, $t = 1, \ldots, T$, generated by the respective DGP, and apply PC analysis to obtain a first-run estimate of the factor process. Of course, the factor estimates that are obtained from the PC analysis, say $\tilde{F}$, are not identified. One way to proceed would be to go ahead with $\tilde{F}$, to estimate the TVFAVAR model and then to compare the estimation-implied properties of Y with those of the simulated data. However, this would preclude us from comparing directly the estimated time varying parameters and hyperparameters corresponding to $\tilde{F}$ with those of the original DGP. Thus, instead of estimating the full TVFAVAR model based on $\tilde{F}$, we first rotate and rescale the estimated $\tilde{F}$ so that it resembles the true unobserved $F = (F_1, F_2)$ from the particular realization of the DGP. More precisely, let A be an orthonormal rotation matrix and $\check{F}(A)$ the corresponding rotation of $\tilde{F}$. We grid-search for that A^* which maximizes $\text{corr}\{F_1, \check{F}_1(A)\} + \text{corr}\{F_2, \check{F}_2(A)\}$ and obtain $\check{F}^* \equiv \check{F}(A^*)$. The final estimate $\hat{F}$ is obtained by rescaling $\check{F}^*$ so that the standard deviation corresponds to that of the simulated F:

$$\hat{F}_i = \frac{\sigma(F_i)}{\sigma(\check{F}_i^*)} \check{F}_i^*, \qquad i = 1, 2.$$

Table 2 documents the correlation of the estimated factors $\hat{F}$ with the simulated factors F: for each of the eight DGPs, Table 2 shows the median as well as the 0.05- and 0.95-quantiles corresponding to the distribution of the empirical correlation across the M MC simulation runs.

Table 2. MC simulation results for factors estimated by PCs†

Correlation	*Results for the following volatilities and parameter variation intensities:*							
	Constant				*Time varying*			
	Low		*High*		*Low*		*High*	
T	150	600	150	600	150	600	150	600
DGP	I	II	III	IV	V	VI	VII	VIII
Factor 1								
Median	0.99	0.98	0.99	0.97	0.99	0.98	0.99	0.97
5%-quantile	0.97	0.97	0.94	0.95	0.98	0.97	0.95	0.95
95%-quantile	1.00	0.99	1.00	0.98	1.00	0.99	1.00	0.99
Factor 2								
Median	0.99	0.98	0.99	0.97	0.99	0.98	0.99	0.97
5%-quantile	0.97	0.97	0.94	0.95	0.98	0.97	0.95	0.95
95%-quantile	1.00	0.99	1.00	0.99	1.00	0.99	1.00	0.99

†'Median' denotes the median (across the M DGP realizations) of the correlation between the simulated factor path and the estimated factor path (using PC analysis, rotation and rescaling as described in the main text); '5% (or 95%)-quantile' denotes the 5% (or 95%) quantile (across DGP realizations) of that correlation.

Summing up, the correlation between simulated (unobserved) factors and their estimated counterparts is very high. Median correlations across the eight DGPs are all at least 0.97, and also the lower 5% quantile is never below 0.94. This suggests that even if the true DGP is not governed by a constant parameter or volatility structure, but rather by a time varying structure, the simple PC estimator recovers the underlying factor process in a very reasonable fashion. Heuristically, the cross-section information is obviously very strong for inferring the position of the unobserved common driving forces of the data. Overall, these results confirm the theoretical results of Stock and Watson (1998) and Bates *et al.* (2013), and they are in line with the MC-based findings of Banerjee *et al.* (2008).

3.2.2. *Time varying parameters*

For the time varying elements of the matrix $\mathcal{K}$, governing the AR dynamics of the factors, Table 3 summarizes the estimation results.

First, the median (across DGP realizations) of the average (across time for a given DGP realization) differences between true and estimated parameters is small (almost always below 0.01) for all variants of the DGP and all elements of $\mathcal{K}$.

Second, estimation results are subject to the 'pile-up problem' that was expounded by Stock and Watson (1998): for moderate degrees of time variation in the parameters and for small samples, the maximum likelihood estimator tends to underestimate the true degree of parameter variation. More precisely, the variance of parameter changes (i.e. the variance of the innovation of the supposed random walk of parameters) tends to be estimated as 0. In our set-up, this problem is most distinct for the off-diagonal parameters of $\mathcal{K}$ for the DGPs with low parameter variation intensity and with small T. Although these parameters are (mildly) time varying, the respective estimated (smoothed) parameter path is essentially a straight line for around 80% of the cases. However, the problem alleviates when moving from a DGP with moderate parameter

Table 3. MC simulation results for time varying $\mathcal{K}$ in expression (3.1)†

	Results for the following volatilities and parameter variation intensities:							
	Constant				*Time varying*			
	Low		*High*		*Low*		*High*	
T	150	600	150	600	150	600	150	600
DGP	I	II	III	IV	V	VI	VII	VIII
$\mathcal{K}_{11}$								
av.diff.	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
corr	0.64	0.85	0.67	0.86	0.64	0.86	0.68	0.87
corr(est.tv)	0.82	0.86	0.81	0.86	0.80	0.86	0.80	0.87
RMSE(est.non-tv)	0.16	0.19	0.18	0.23	0.15	0.18	0.18	0.21
RMSE(est.tv)	0.16	0.16	0.18	0.18	0.15	0.15	0.18	0.17
fraction est.tv	0.44	0.95	0.51	0.98	0.47	0.97	0.60	0.99
$\mathcal{K}_{12}$								
av.diff.	0.00	−0.01	0.00	0.00	0.00	0.00	0.00	0.00
corr	0.28	0.67	0.51	0.79	0.29	0.66	0.51	0.79
corr(est.tv)	0.53	0.78	0.80	0.82	0.56	0.74	0.75	0.81
RMSE(est.non-tv)	0.08	0.09	0.11	0.13	0.08	0.09	0.12	0.12
RMSE(est.tv)	0.10	0.09	0.12	0.13	0.10	0.09	0.12	0.12
fraction est.tv	0.16	0.64	0.27	0.88	0.23	0.71	0.36	0.92
$\mathcal{K}_{21}$								
av.diff.	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
corr	0.25	0.68	0.52	0.80	0.27	0.67	0.52	0.79
corr(est.tv)	0.69	0.76	0.77	0.81	0.66	0.77	0.79	0.81
RMSE(est.non-tv)	0.07	0.09	0.11	0.13	0.07	0.08	0.11	0.12
RMSE(est.tv)	0.10	0.09	0.13	0.12	0.10	0.09	0.12	0.12
fraction est.tv	0.17	0.60	0.29	0.86	0.20	0.64	0.34	0.88
$\mathcal{K}_{22}$								
av.diff.	0.00	0.00	0.02	0.00	0.00	0.00	0.01	0.00
corr	0.65	0.86	0.69	0.81	0.67	0.86	0.69	0.87
corr(est.tv)	0.82	0.86	0.82	0.87	0.82	0.86	0.82	0.87
RMSE(est.non-tv)	0.17	0.20	0.19	0.22	0.16	0.18	0.19	0.23
RMSE(est.tv)	0.16	0.16	0.18	0.18	0.16	0.15	0.18	0.18
fraction est.tv	0.44	0.96	0.53	0.98	0.46	0.97	0.55	0.99

†'av.diff' denotes the median (across DGP realizations) of the average (across time in a particular DGP realization) differences between the simulated parameter path and the estimated (smoothed) path; 'corr' is the median correlation between simulated and estimated parameter paths; 'corr(est.tv)' is the same as 'corr', but the median is taken with respect to those DGPs for which the estimated parameter path has a non-zero standard deviation; 'RMSE(est.non-tv)' is the median of root-mean-squared errors between actual and estimated parameter paths, but computed only for those DGPs that gave rise to an estimated parameter path with zero standard deviation (i.e. parameters were estimated as a constant sequence); 'RMSE(est.tv)' is the same, but now for those DGPs where the estimated parameter path had non-zero standard deviation; 'fraction est.tv' denotes the fraction of DGPs for which the estimated parameter path showed actual time variation, i.e. its standard deviation is non-zero.

variation to a DGP with high variation, or when moving from small to large T. In particular, for high parameter variation intensity and large T, the diagonal elements of $\mathcal{K}$ are identified as time varying in almost 100% of MC runs.

Moreover, for those parameters that are falsely estimated as constant when the true DGP actually exhibits 'mild' time variation, the average size of the parameter can still be estimated to satisfying precision. Table 3 shows that the root-mean-squared error RMSE for those parameters that are estimated as time varying hardly differs from those where this is not so.

Third, the sample correlation between estimated and actual parameters (given that the estimated parameter sequence has a non-zero standard deviation, i.e. the parameter is not estimated as constant) is high: the median correlation for the diagonal elements of $\mathcal{K}$ exceeds 0.8 for all DGP versions, whereas for the off-diagonal elements the median correlation ranges around 0.75 across DGPs. As the time varying parameter paths are generated as random walks, the correlation between them and their respective estimates is, strictly speaking, not a meaningful object. However, for a given finite sample, the correlation can serve as a useful additional descriptive statistic of the relative positioning of the realized and estimated parameter path.

Table 4 shows the corresponding results for the time varying elements of the factor loading matrices Λ. Recall that within each MC set-up we used the same specification for the time

Table 4. MC simulation results for time varying Λ in expression (3.1)†

	Results for the following volatilities and parameter variation intensities:							
	Constant				*Time varying*			
	Low		*High*		*Low*		*High*	
T	150	600	150	600	150	600	150	600
DGP	I	II	III	IV	V	VI	VII	VIII
$\Lambda_{1,1}$								
av.diff.	0.01	−0.02	0.01	−0.14	0.00	0.01	0.01	−0.11
corr	0.67	0.86	0.84	0.89	0.67	0.87	0.84	0.90
corr(est.tv)	0.83	0.86	0.85	0.89	0.83	0.87	0.85	0.90
RMSE(est.non-tv)	0.45	0.56	1.03	—	0.44	0.58	0.99	—
RMSE(est.tv)	0.47	0.54	0.90	1.63	0.45	0.52	0.88	1.61
fraction est.tv	0.52	0.97	0.90	1.00	0.53	0.97	0.91	1.00
$\Lambda_{1,2}$								
av.diff.	0.00	−0.01	−0.01	−0.16	0.00	−0.01	−0.02	−0.14
corr	0.68	0.85	0.85	0.90	0.67	0.87	0.85	0.90
corr(est.tv)	0.82	0.86	0.86	0.90	0.82	0.88	0.86	0.90
RMSE(est.non-tv)	0.46	0.60	0.99	—	0.45	0.54	1.01	—
RMSE(est.tv)	0.45	0.54	0.89	1.66	0.44	0.52	0.87	1.65
fraction est.tv	0.51	0.97	0.91	1.00	0.53	0.97	0.92	1.00
$\Lambda_{1,3}$								
av.diff.	0.00	0.01	−0.02	−0.08	−0.01	0.00	0.00	−0.09
corr	0.63	0.86	0.83	0.90	0.63	0.87	0.83	0.90
corr(est.tv)	0.82	0.86	0.85	0.90	0.83	0.87	0.85	0.90
RMSE(est.non-tv)	0.47	0.58	1.03	—	0.45	0.57	1.03	—
RMSE(est.tv)	0.46	0.53	0.91	1.61	0.45	0.51	0.89	1.59
fraction est.tv	0.48	0.98	0.90	1.00	0.48	0.97	0.90	1.00
$\Lambda_{1,4}$								
av.diff.	−0.01	−0.01	−0.01	−0.09	−0.01	−0.02	0.00	−0.09
corr	0.63	0.87	0.83	0.90	0.67	0.87	0.84	0.90
corr(est.tv)	0.82	0.87	0.85	0.90	0.81	0.87	0.85	0.90
RMSE(est.non-tv)	0.45	0.54	0.99	—	0.44	0.59	1.01	—
RMSE(est.tv)	0.45	0.53	0.90	1.59	0.44	0.52	0.90	1.60
fraction est.tv	0.49	0.98	0.91	1.00	0.51	0.97	0.91	1.00

†$\Lambda_{i,j}$ is the (i, j)-element of Λ. Thus, the table shows results for the loadings for the first factor for the first four elements of y. As described in the main text, the results for the remaining loadings are completely symmetric. Table entries are analogous to those in Table 3.

variation of all $2N$ elements of Λ, and also for the idiosyncratic components of all N cross-sectional equations. Hence, for each of the eight MC variants, the properties of all $2N$ estimated parameter paths for the Λ-elements should be the same. Accordingly, we pick only the first four of the N loadings corresponding to the first factor to study the estimation results, as they are representative for all others.

The results are qualitatively similar to those just presented for $\mathcal{K}$. Overall, the 'pile-up' problem of falsely estimating time varying parameters as constant does not show up that prominently here, mainly because the variation intensity of the Λ-parameters is overall larger than that of the $\mathcal{K}$-parameters. In fact, only for the two MC variants with low parameter variation intensity and smaller T (i.e. I and V) is the fraction of parameter paths identified as time varying just over 50%. For the other MC variants, this fraction ranges between 90% and 100%. Moreover, for cases I and V, where almost half of the parameter paths are erroneously estimated to be constant, the RMSEs for the MC runs identifying time variation hardly differ from those where parameters are estimated as constant. Across all eight MC variants, the correlation between estimated and simulated parameter paths (given that the estimation indeed identifies time variation) is between 0.8 and 0.9.

3.2.3. *Idiosyncratic components*

Turning to the idiosyncratic components (3.5) of the N cross-sectional equations, the estimation results are provided in Table 5. Again, the first four of the N equations have been picked for

Table 5. MC simulation results for idiosyncratic components in expression (3.5)†

	Results for the following volatilities and parameter variation intensities:							
	Constant				*Time varying*			
	Low		*High*		*Low*		*High*	
T	150	600	150	600	150	600	150	600
DGP	I	II	III	IV	V	VI	VII	VIII
σ_1 Median	1.52	1.51	1.55	1.53	1.52	1.51	1.55	1.53
$Q_{95}-Q_{05}$	0.34	0.17	0.42	0.26	0.34	0.16	0.42	0.26
σ_2 Median	1.51	1.51	1.54	1.53	1.51	1.51	1.54	1.53
$Q_{95}-Q_{05}$	0.36	0.18	0.42	0.23	0.36	0.17	0.41	0.23
σ_3 Median	1.52	1.51	1.55	1.52	1.52	1.51	1.55	1.53
$Q_{95}-Q_{05}$	0.35	0.17	0.46	0.25	0.34	0.16	0.43	0.24
σ_4 Median	1.52	1.51	1.55	1.53	1.52	1.51	1.55	1.52
$Q_{95}-Q_{05}$	0.33	0.16	0.43	0.24	0.33	0.17	0.42	0.23
ρ_1 Median	0.33	0.31	0.36	0.33	0.32	0.31	0.36	0.33
$Q_{95}-Q_{05}$	0.36	0.16	0.58	0.34	0.34	0.16	0.54	0.34
ρ_2 Median	0.32	0.31	0.34	0.33	0.32	0.31	0.34	0.33
$Q_{95}-Q_{05}$	0.38	0.18	0.54	0.32	0.37	0.16	0.51	0.34
ρ_3 Median	0.32	0.31	0.35	0.33	0.32	0.31	0.35	0.33
$Q_{95}-Q_{05}$	0.40	0.16	0.58	0.31	0.37	0.15	0.52	0.31
ρ_4 Median	0.32	0.31	0.35	0.33	0.32	0.31	0.34	0.33
$Q_{95}-Q_{05}$	0.37	0.16	0.56	0.29	0.34	0.16	0.53	0.28

†σ_i and ρ_i are the innovation volatility and auto-correlation parameter of the idiosyncratic component of the ith cross-sectional equation. The true parameters are $\sigma_i = 1.5$ and $\rho_i = 0.3$ for $i = 1, \ldots, N$ and all MC variants. Results are shown for the first four cross-sectional equations. Again, results for the remaining $N-4$ equations are similar by design. 'Median' denotes the median of the point estimates of the respective parameter across the M DGP realizations, and $Q_{95}-Q_{05}$ denotes the difference between the 95% and 5% percentile of the distribution of these point estimates.

illustration as the results of the remaining equations are similar by construction. Table 5 provides the median (across MC runs) of the point estimate of σ_i and ρ_i in expression (3.5).

Both parameters are estimated with only a slight positive bias, which is highest for those MC variants that feature high parameter variation intensity and small T. It is also for these two constellations (III and VII) that the estimation uncertainty, measured by the 0.95–0.05-interpercentile range, is highest. This is intuitive, as in these cases relatively more of the total variation in x_i stems from the time varying parameters (and factor variation), compared with the idiosyncratic component.

3.2.4. Time varying volatility

As described above, MC specifications V–VIII feature time varying volatility of factor innovations, whereas variants I–IV do not. For the MC experiment it is assumed that the econometrician knows whether volatility is time varying or not. The parameters c_1 and c_2 in expression (3.2) govern the size of the constant innovation volatility in specifications I–IV, whereas they constitute the 'intercepts' in the assumed exponential affine specification of time varying volatility for specifications V–VIII. Table 6 displays the median (across the M MC runs) of these parameters. Point estimates of c_1 and c_2 appear unbiased for all MC variants. For the variants with time varying volatility, the parameters b_{ij} govern the effect of lagged factors on current innovation volatility. Estimates of b_{12} appear to have a small downward bias, whereas the other b_{ij}-parameters have only a slight such bias, if at all. What is more important, though, is that the estimated volatility paths—computed by plugging the estimated c_i and b_{ij} and estimated factors into expression (3.2)—resemble the simulated parameters very well: first, the median (across MC runs) of the average (across time per MC run) difference between the simulated and estimated volatility paths is essentially 0 for all four MC variants with time varying volatility; second, the median of the correlations between the simulated and estimated volatility paths is very high as it ranges between 0.95 and 0.98 for all specifications.

Table 6. MC simulation results for (time varying) volatility in expression (3.2)†

	Results for the following volatilities and parameter variation intensities:							
	Constant				*Time varying*			
	Low		*High*		*Low*		*High*	
T	150	600	150	600	150	600	150	600
DGP	I	II	III	IV	V	VI	VII	VIII
Factor 1								
c_1	−1.18	−1.19	−1.16	−1.20	−1.20	−1.19	−1.19	−1.21
b_{11}	—	—	—	—	0.14	0.15	0.14	0.15
b_{12}	—	—	—	—	0.58	0.57	0.55	0.56
av.diff.	—	—	—	—	0.00	0.00	0.00	0.00
corr	—	—	—	—	0.97	0.98	0.97	0.97

†The true parameters are $c_1 = c_2 = -1.2$ for all DGPs. For versions V–VIII, $b_{11} = 0.15$, $b_{12} = 0.6$, $b_{21} = 0.3$ and $b_{22} = 0.3$, and they are 0 for the constant volatility variants I–IV. The rows labelled c_1, b_{11}, b_{12}, c_2, b_{21} and b_{22} contain the median of the point estimates of the respective parameters. 'av.diff.' denotes the median (across DGP realizations) of the average (over time) difference between simulated and estimated volatility trajectories, and 'corr' denotes the median of the correlations between the simulated and estimated volatility.

4. A large data set for the USA

After showing the good performance of our estimation procedure with simulated data, we now want to assess it with actual data. Therefore, we have constructed a large balanced data set containing 316 quarterly US time series observed from 1972, quarter 1, to 2012, quarter 4. 114 of them are measures of real economic activity (e.g. gross domestic product (GDP) and components, industrial production, employment measures, capacity utilization and retail sales), 124 series are price measures (e.g. deflators of GDP and components as well as of personal consumption expenditures, consumer and producer prices, and commodity prices), 66 series represent monetary and financial variables (e.g. interest rates, stock prices, house prices, money and credit aggregates, and exchange rates) and 12 series capture (inflation and activity) expectations. Unit labour costs, asset prices, credit and monetary aggregates are divided by the GDP deflator and enter in real terms.

The variables are transformed as usual in dynamic factor analysis. Specifically, series that were not already available in seasonally adjusted form are seasonally adjusted by using the census X12 method (US Bureau of the Census). Variables showing non-stationary behaviour are made stationary through differencing. Most series enter in differences of their logarithms except for interest rates, ratios and expectations, which enter in levels. Following Stock and Watson (2005), outliers are defined as observations of each (stationary) variable with absolute median deviations larger than six times the interquartile range. They are replaced by the median value of the preceding five observations. Finally, the series are demeaned and standardized to have a unit variance. The on-line data appendix (Table A.1) contains details on the data, the transformations and the sources.

5. Estimation results

We estimate the TVFAVAR model along the lines described in Section 2. We rely on five latent factors. Those explain 68% of the variation in the data set. As a check of robustness, we also carried out the structural analysis with six latent factors, which account for 71% of the data set. Our main results remain broadly unaffected, and we prefer the more parsimonious parameterization. (More details on the selection of the number of factors are provided in Eickmeier *et al.* (2011b).)

We use a VAR(2) model for the factor dynamics. The choice of the lag length is suggested both by the need to reduce the number of parameters, and by the consideration that allowing for parameter time variation probably reduces the need for longer lags. We document the estimated factors and provide evidence that the two-step procedure (estimate factors as PCs; then estimate time varying parameters given factors) is adequate. We then summarize the extent of time variation in the FAVAR system and finally provide some diagnostic checking.

5.1. Estimated factors

Fig. 1 shows the estimated factor paths. To assess whether the PC approach is adequate for estimating factors in the presence of time variation in the factor loadings, we derive an alternative factor estimate from a cross-sectional regression of the N variables $x_{i,t}$ on the estimated time varying loadings $\hat{\Lambda}_{i,t|T}$, for each period t. The estimated factors resulting from this exercise (which are displayed in Fig. 1 in dark grey) show a strong similarity to those estimated from the PC analysis; the respective correlation coefficients all exceed 0.98.

In addition, we can also run a full filtering exercise, treating our estimated parameter paths as fixed, now treating the factors as unobservable states and then using the Kalman smoother

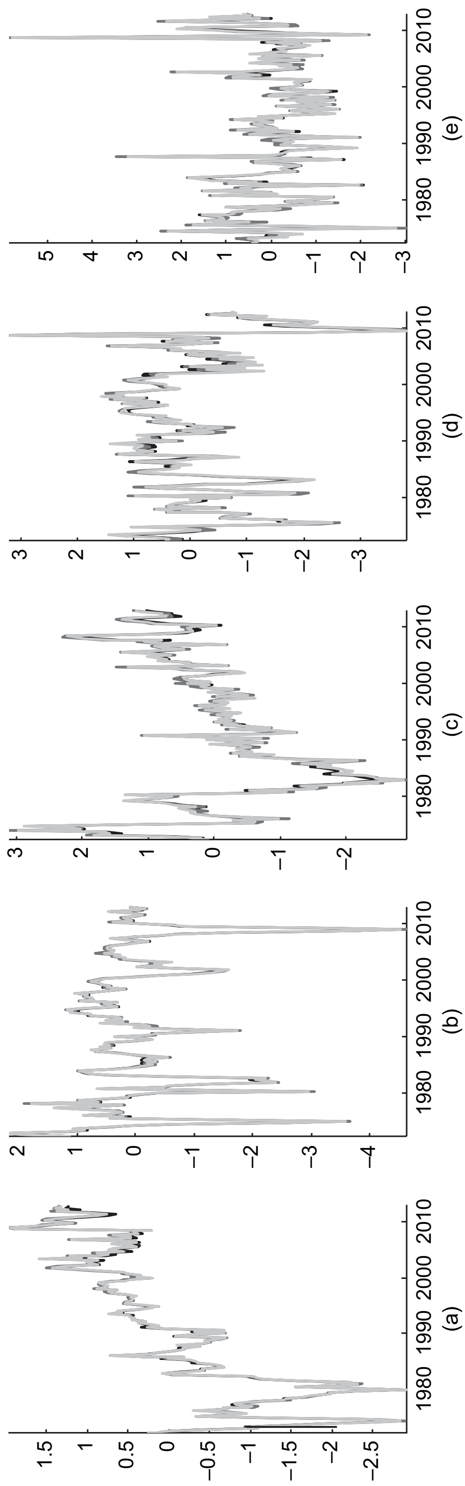

Fig. 1. Factor estimates (——, PCs; ——, re-estimated model (regression); ——, re-estimated model (state space)): (a) factor 1; (b) factor 2; (c) factor 3; (d) factor 4; (e) factor 5

to re-estimate them. For this exercise, the transition equation of the resulting state space model is

$$\begin{pmatrix} F_t \\ F_{t-1} \end{pmatrix} = \begin{pmatrix} \hat{P}_{t|T}^{-1}\hat{\mathcal{K}}_{1,t|T} & \hat{P}_{t|T}^{-1}\hat{\mathcal{K}}_{2,t|T} \\ I_G & \mathbf{0}_G \end{pmatrix} + \begin{pmatrix} F_{t-1} \\ F_{t-2} \end{pmatrix} + \begin{pmatrix} I_G \\ \mathbf{0}_G \end{pmatrix} \tilde{u}_t,$$

$$E(\tilde{u}_t) = 0, \qquad E(\tilde{u}_t \tilde{u}_t') = \hat{P}_{t|T}^{-1} \hat{S}_t \hat{P}_{t|T}^{-1\prime} \tag{5.1}$$

and the measurement equation is

$$X_t = (\hat{\Lambda}_{t|T}, \mathbf{0}_{N\times G}) \begin{pmatrix} F_t \\ F_{t-1} \end{pmatrix} + \tilde{e}_t, \tag{5.2}$$

where objects with circumflexes and subscript $t|T$ denote the parameter paths estimated in the first step, in which the factors had been kept fixed at their PC estimates. Strictly speaking, the 'dual' state space representation (5.1)–(5.2) of a time varying FAVAR model is valid only if the idiosyncratic components in expression (2.6) are serially uncorrelated, i.e. $\rho_i = 0$ for all i. In the relevant case with auto-correlated idiosyncratic errors the idiosyncratic components would enter the state vector which would be of dimension $2G + N$ instead of $2G$ as in equation (5.2). We abstain from conducting the exercise with this large (316-)dimensional state vector, but instead use the misspecified state space representation (5.1)–(5.2), where we ignore the AR structure of the measurement error in equation (5.2). Running the Kalman smoother on this 'approximate' dual state space model (5.1)–(5.2) delivers factor estimates that are likewise very close to the PC estimates, and accordingly also close to the factors that were obtained from the cross-section regression.

Overall, this exercise provides (heuristic) support for our assumption to keep PC-based factor estimates fixed when estimating the time varying parameters.

5.2. Time variation in parameters and volatility

One may wonder whether a constant parameter specification would suffice or whether time variation in the parameters is really needed and, if so, which sources of parameter variation are the most important. One way to quantify the overall degree of time variation in the AR matrix $\mathcal{K}_t$, the contemporaneous relations matrix P_t and the loadings Λ_{it} is to count the number of occasions when the standard deviation of the innovations of the time varying parameters—the respective elements of diag(Q) in expression (2.7)—are significant. However, conducting such a multitude of individual significance tests in the usual fashion may lead to a biased assessment of the overall degree of time variation. Indeed, if these tests are conducted with an effective size of, say, 5%, then even in the extreme case of no time variation at all one would expect to reject the null hypothesis of no time variation 5% of the time. Moreover, a further complication arises as, under the null hypothesis of no variation, the parameter lies on the boundary of the allowable parameter space. Accordingly, we resort to a more direct approach of gauging the overall degree of time variation in the system: we count the number of parameters for which the time evolution estimated by the Kalman smoother is 'a straight line', i.e. for which the standard deviation of the smoothed parameter series is essentially 0.

It turns out that there is actual time variation (i.e. no 'straight line' parameter paths) for five out of the 50 parameters of the $\mathcal{K}$ AR matrix (containing the dynamics of the VAR(2) process for the five factors), one out of the 10 $(= 0.5 \times 5 \times 4)$ parameters of the P-matrix of contemporaneous relationships of the VAR process and 926 out of the 1580 loadings (since there are five loadings, one for each factor, for each of the 316 variables).

Finally, we have assessed whether there is indeed time variation in the volatilies of the shocks, i.e. whether the elements of b_g in equation (2.8) are significant. The corresponding t-statistics are based on the estimated standard errors, which are obtained from the negative inverse of the Hessian of the likelihood function. We find that 12 out of 25 parameters are indeed significant at the 5% level.

In summary, the results in this section based on our estimated TVFAVAR model indicate that most of the time variation in the behaviour of US macroeconomic and financial variables over 1972–2012 is associated with changes in the effect of the factors on the variables under analysis and with changes in the volatility of the shocks (which is linked to lagged factors in our model). The degree of variation in the contemporaneous or dynamic relationships across factors is more subdued.

5.3. Diagnostic checking

We first want to check the adequacy of the chosen VAR lag length. If longer lags were needed, the estimated residuals would be correlated over time. Hence, in Fig. 2 we report the estimated auto-correlation function (ACF) for the standardized VAR residuals, together with asymptotic 95% confidence bands. Overall, there is no major evidence against the assumption of no correlation of the VAR(2) errors. In Fig. 3 we report the cumulative sum statistics for the scaled factor residuals, which also do not indicate any major remaining problems of parameter instability.

Similarly, one may wonder whether our assumption of AR(1) idiosyncratic components, although standard in the literature, is sufficient to clean from temporal correlation. Formal statistical testing is complicated since the joint null hypothesis has a large number of components. To provide at least some indication of the existence of possible problems, we have computed the percentage of the 316 idiosyncratic residuals (one for each of the variables under analysis) for which a given lag of the ACF is outside the asymptotic bands. It turns out that only 6% of the residuals have the first lag of the ACF outside the bands, with only slightly larger values of 10% and 11% for the second and third lag of the ACF. Hence, overall this informal diagnostic check does not provide evidence against our assumption of AR(1) idiosyncratic components.

6. Forecasting with the time varying factor-augmented vector auto-regressive model

In this section we evaluate the forecasting performance of our proposed TVFAVAR approach for a set of key variables. To the best of our knowledge, this is the first application of time varying techniques for forecasting in a data rich environment. We predict variables representing real activity (including growth of GDP, consumption, investment, industrial production and employment as well as the unemployment rate), inflation (changes in the GDP deflator, the consumer price index CPI, the personal consumption deflator, the producer price index PPI and unit labour costs), and a number of financial and monetary variables, for a total of 19 indicators under evaluation.

The factors are estimated as the first $G=5$ PCs of our data set, and they are then modelled together with each target variable as a time varying VAR process whose parameters evolve as independent random walks. The TVFAVAR forecasting model thus includes overall six endogenous variables or factors, and its lag length is, again, set to 2. Hence, for each variable of interest $x_{i,t}$, we have $y_{i,t} := (F_t, x_{i,t})$, with

$$y_{i,t} = A_{1,i,t} y_{i,t-1} + A_{2,i,t} y_{i,t-2} + v_{i,t}, \tag{6.1}$$

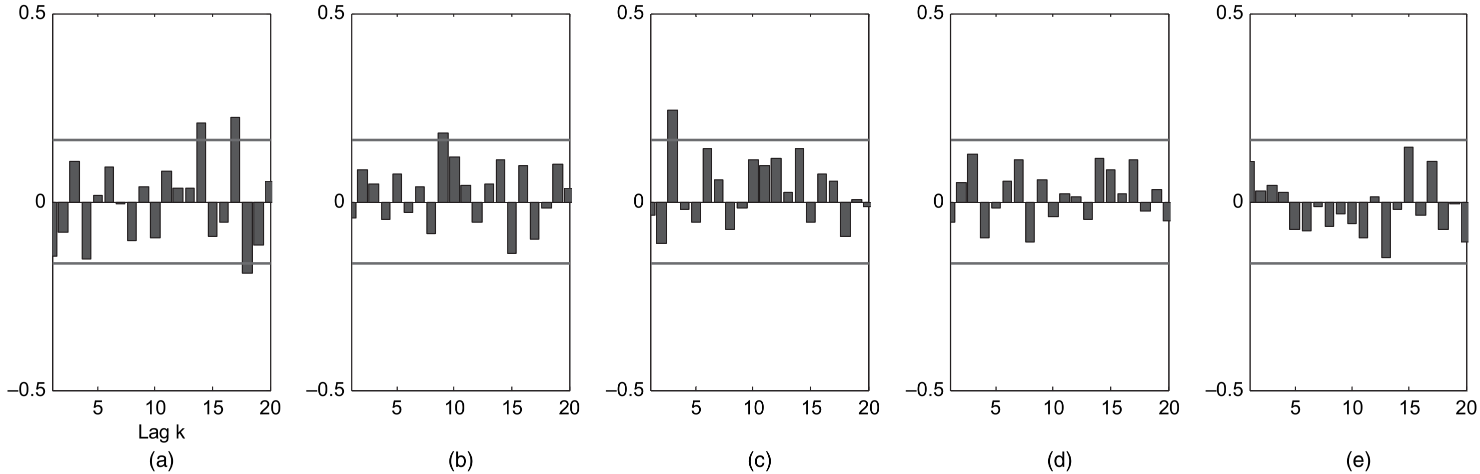

Fig. 2. Tests on the standardized VAR residuals—auto-correlations (■, auto-correlations at different lags; ▬▬, approximate 2-standard-error bounds computed as $2/\sqrt{T}$): (a) factor 1; (b) factor 2; (c) factor 3; (d) factor 4; (e) factor 5

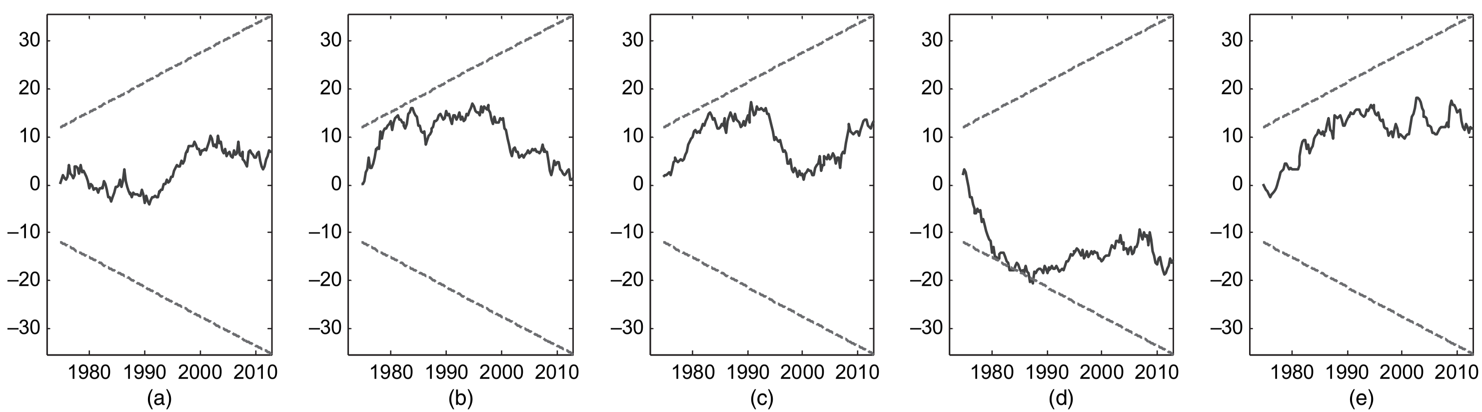

Fig. 3. Tests on the standardized VAR residuals—cumulative sum statistics (——) and approximate 95% confidence bounds (– – –): (a) factor 1; (b) factor 2; (c) factor 3; (d) factor 4; (e) factor 5

where each element of $A_{1,i,t}$ and $A_{2,i,t}$ evolves as an independent random walk and the volatility of $v_{i,t}$ is modelled as in expression (2.8). We take the five lagged latent factors as volatility regressors in the first five equations. The last equation's volatility features these factors as well, but in addition the lagged variable of interest. Note that, with respect to the TVFAVAR specification in Section 2, the forecasting model allows for feedback from the target variable to the factors, and for a direct effect of past values of the target variable on its current evolution. Both features are fairly standard in forecasting models and represent a direct extension of the TVFAVAR model from Section 2.

We take a FAVAR model with constant parameters as the benchmark and compare its root-mean-squared forecast error RMSFE with RMSFEs resulting from an AR process without or with time varying (model TVAR) (random-walk) parameters, the TVFAVAR model assuming constant volatility and the full TVFAVAR model. This exercise allows us to assess whether there are gains not only from using a large information set as summarized by the estimated factors, but also from moving from a constant to a time varying parameter set-up, and from explicitly modelling volatility. For comparability, we set the lag length of the benchmark AR model, the TVAR model, the constant parameter FAVAR model and the TVFAVAR model with constant volatility also set to 2.

To assess whether the RMSFE differences are statistically significant, we compute the Diebold and Mariano (1995) test with the Harvey and Newbold (1997) adjustment of the variance that enters the test statistic, which improves its finite sample performance. Our use of the Diebold–Mariano test with forecasts that are, in many cases, nested is a deliberate choice. MC evidence in Clark and McCracken (2011a, b) indicates that, with nested models, the Diebold–Mariano test compared against normal critical values can be viewed as a somewhat conservative (in the sense of tending to have size modestly below nominal size) test for equal accuracy in finite samples.

The estimation window is expanded quarter by quarter and the models are re-estimated each time. The first estimation window reaches from 1972, quarter 1, until 1994, quarter 1. For each estimation window, the respective estimated model is used to compute forecasts one, two and four quarters ahead ($h=1,2,4$). Hence, the evaluation period ranges from 1995, quarter 1, to 2012, quarter 4. Unlike for the other analyses conducted in this paper, for the forecasting exercise the initial state of the time varying parameter model is based on ordinary least squares estimates by using the first 50 observations (rather than using the respective full estimation window).

For forecast evaluation, we separate the financial crisis period since the very particular behaviour of many variables could distort the outcome of the forecast evaluation exercise. Hence, we report results for the samples from 1995, quarter 1, to 2007, quarter 2, plus from 2010, quarter 1, to 2012, quarter 4, and for the crisis period that we identify with from 2007, quarter 3, to 2009, quarter 4 (the end of the latest recession as dated by the National Bureau of Economic Research).

Table 7 reports the results for the real activity, inflation and interest rates as well as monetary, credit and asset price variables. Each panel contains five groups of results. The first group reports the RMSFEs resulting from the benchmark constant parameter FAVAR model. The second to fifth groups contain relative RMSFEs of the constant parameter and time varying AR models, and of the TVFAVAR models without and with changing volatility *vis-à-vis* the benchmark. Each group has two columns, referring to the full sample without the crisis and to the crisis period only (from 2007, quarter 3, to 2009, quarter 4).

Starting with the full sample without the crisis, adding parameter time variation to the constant parameter FAVAR model is helpful in 26 out of 57 cases (19 variables times three forecast horizons), increasing to 32 out of 57 cases when volatility is also changing. The gains are larger for the nominal variables but are also present for some real and financial indicators.

Table 7. Forecast evaluation—RMSFE (third and fourth columns) and relative RMFSE (other columns)†

	Horizon	Results for the following models:									
		Constant FAVAR, absolute		*Constant AR versus constant FAVAR*		*TVAR versus constant FAVAR*		*TVFAVAR, constant volatility, versus constant FAVAR*		*TVFAVAR, time varying volatility, versus constant FAVAR*	
		Normal	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*
ΔGDP	1	0.68	1.15	0.83‡	1.07	0.83‡	1.09	0.99	0.96	0.96	0.95
	2	0.61	1.5	0.89§	0.97	0.89§	1.04	0.98	1.04	0.98	0.92
	4	0.73	1.51	0.78§§	1.05‡	0.8‡	1.09‡	0.98	1.07	0.89	1.02
ΔConsumption	1	0.57	1.29	0.92	0.85	0.92	0.89	1.03	0.93	0.96	0.92
	2	0.67	1.32	0.8§§	0.99	0.81§§	1.02§	0.88§	0.97	0.92	0.99
	4	0.74	1.52	0.77§§	1.02	0.79§§	1.08§§	0.93	1.03	1	0.97
ΔInvestment	1	0.67	0.6	0.87	2.33‡	0.87	2.47‡	0.99	0.94	0.99	0.88
	2	0.63	1.49	0.94	1.09	0.94	1.29§	1.1	0.9	1.06	0.8§
	4	0.72	1.52	0.83§	1.08‡	1.02	1.08‡	1	1.09	0.94	0.98
ΔIndustrial production	1	0.57	1.22	0.78§§	1.14	0.78§§	1.32	1.03	0.7‡	0.93	0.71‡
	2	0.63	2.02	0.86§§	0.97	0.86§§	1.42	1.21	1.09	0.99	0.88
	4	0.72	2.07	0.88	1.02§§	1.97	1.3	1.47	1.02	1.09	0.89
Unemployment rate	1	0.12	0.26	1.13	0.98	1.13	0.99	1.08§	1.11	1.25§§	0.75‡
	2	0.21	0.58	1.13	1.01	1.13	1.01	1.41	0.95	1.31§§	0.67‡
	4	0.34	1.27	1.16	1	1.16	1	1.67	0.94	1.1	0.81
ΔEmployment	1	0.27	0.89	1.06	0.81	1.08	0.91	1.08§§	1.03	1.06	0.9
	2	0.37	1.41	0.99	0.93	1.01	1.23	1.16	1	1.06	0.77§§
	4	0.6	1.67	0.83§	1.12‡	1.93	1.15‡	1.22	0.92	0.91	0.78§
ΔGDP deflator	1	0.39	0.54	0.86	0.92	0.87	1	0.97	1.32	0.92	1.67
	2	0.4	0.76	0.79‡	0.86	0.83§	1.01	0.89§	1.34	0.85§	1.68
	4	0.63	0.83	0.54§§	0.89	0.68§§	1.04§§	0.77§	1.1	0.68‡	1.19
ΔPersonal constant deflator	1	0.45	1.24	0.88‡	1.01	0.91	0.99	0.89§	1.16	0.94	1.15
	2	0.49	1.55	0.87	0.98	0.94	0.97	0.88	0.85	0.92	0.87
	4	0.68	1.49	0.69§§	0.97	0.79‡	0.97	0.79‡	0.87§	0.76‡	0.86§
ΔCPI	1	0.5	1.47	0.98	1.06	1.01	1.14	0.88§	1.36	0.86‡	1.19
	2	0.55	1.85	0.96	0.94	1.06	0.96	0.86	0.88	0.86§	0.88
	4	0.73	1.63	0.72§§	1	0.92	0.94§	0.76§§	0.91‡	0.71§§	0.9‡

(*continued*)

Table 7 (*continued*)

	Horizon	*Results for the following models:*									
		Constant FAVAR, absolute		*Constant AR versus constant FAVAR*		*TVAR versus constant FAVAR*		*TVFAVAR, constant volatility, versus constant FAVAR*		*TVFAVAR, time varying volatility, versus constant FAVAR*	
		Normal	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*
ΔPPI	1	0.69	1.92	0.89§	1.02	0.88‡	1.08	1.02	1.42	0.97	1.06
	2	0.73	2.14	0.93	1.01	0.9	1.01	0.96	1.09§	0.97	0.96
	4	0.87	2.01	0.8‡	1.02	0.78‡	0.99	1.21	0.99	0.82‡	0.98
ΔUnit labour cost manufacturing	1	1.21	1.39	0.95	1.12	0.95	1.18§	1.1	0.92	1.05	1.01
	2	1.2	1.57	0.95§	0.99	0.95§	1.05	1.07	1.13	1.05	0.96
	4	1.14	1.5	0.99	1.03	0.99	1.19	1.08	1.15§	1.03	1.13
Federal Funds rate	1	0.13	0.39	0.73§	0.51	0.82	0.6	1.46§	1.02	1.35	0.74
	2	0.19	0.69	0.97	0.55	1.28	0.8	1.7§	0.83	1.5§	0.65
	4	0.41	0.71	0.85	0.98	1.41	2.72‡	1.21	0.99	1.17	0.9§
10-year government bond yield	1	0.14	0.17	0.97	0.74	1	0.74	1.04	0.94	1.04	0.76
	2	0.23	0.3	0.99	0.68	1.04	0.68	1.12§§	0.87	1.14	0.6
	4	0.32	0.31	1.03	0.85	1.13	0.85	1.19§§	0.89‡	1.17	0.74
ΔM2	1	0.71	1.01	0.99	1.05	0.99	1.05	1.04	0.95	0.95	1
	2	0.88	1.49	0.89§	0.84	0.89§	0.84	0.98	0.89‡	0.91	0.87
	4	1.04	1.39	0.77‡	0.84§§	0.77‡	0.84§§	1.03	0.82	0.89	0.76‡
ΔConsumer loans	1	0.69	1.34	0.98	0.84	1.02	0.81	0.99	0.88	0.96§	0.89
	2	0.85	1.66	0.85	0.88	0.91	0.85§	0.99	0.75	0.92	0.9
	4	1.01	1.45	0.74	0.87§§	0.78	0.86‡	1	0.78§	0.9	0.78
ΔCommercial and industrial loans	1	0.47	1.25	1.01	1.03	1.01	1.03	0.99	1.42	1.05	0.88
	2	0.57	1.21	1.09§	1.43	1.1§	1.44	0.97	1.17‡	1.02	1
	4	0.75	1.68	1.15‡	1.25	1.18‡	1.26	0.97	1.21§	1.02	1.17
ΔReal estate loans	1	0.89	1.52	0.97	0.96	1	0.95	0.96	1.13	1.03	1.15
	2	1	1.17	0.96	1.12	1.03	1.1	1.03	1.28	1.09	1.46
	4	1.15	1.06	0.9	1.1	0.95	1.15§	1.05	1.07	1.04	0.84§

(*continued overleaf*)

Table 7 (*continued*)

	Horizon	*Results for the following models:*									
		Constant FAVAR, absolute		*Constant AR versus constant FAVAR*		*TVAR versus constant FAVAR*		*TVFAVAR, constant volatility, versus constant FAVAR*		*TVFAVAR, time varying volatility, versus constant FAVAR*	
		Normal	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*	*Normal*	*Crisis*
ΔStandard and Poors 500	1	0.84	1.77	0.9‡	1	0.9‡	1	0.96	1.18	0.97	0.97
	2	0.92	1.86	0.88§§	1.09	0.88‡	1.09	0.97	1.36	0.94	1.14
	4	0.9	2.01	0.91§	0.99	0.91§	0.99	1.1	1.1‡	0.99	1.11§
Δ House price	1	0.88	2.25	0.95	0.89	0.92	0.99	0.99	1.02	0.98	1.06
	2	0.85	2.04	0.96	0.95§§	0.97	0.95§	1.04	1.02	1.06	1.13
	4	0.89	1.93	0.95	0.99	0.95	0.98	1.09	0.95	1.06	1.12§§

†The crisis period is defined as from 2007, quarter 3, to 2009, quarter 4. The models compared are the FAVAR model with constant parameters as benchmark (constant FAVAR), the AR model with constant parameters (constant AR), the AR model with time varying parameters (TVAR), the FAVAR model with time varying parameters but constant volatility (TVFAVAR, constant volatility) and the FAVAR model with time varying parameters and time varying volatility (TVFAVAR, time varying volatility). The relative RMSFEs are based on the Diebold and Mariano (1995) test. For details on the forecast design, see Section 6.
‡Different from 1 at the 5% significance level.
§Different from 1 at the 10% significance level.
§§Different from 1 at the 1% significance level.

We should also mention that recursive estimation of the constant parameter FAVAR model helps since it introduces by itself a form of parameter time variation, which is, however, at odds with the underlying assumption of parameter stability, making the resulting estimators biased and inconsistent, though more useful for forecasting.

Interestingly, adding parameter time variation to the AR model is instead virtually useless. This suggests that changes in the cross-variable relationships (as captured by the TVFAVAR model) are more relevant than changes in the AR dynamics (as captured by the time varying AR model), even though we should again remember that the constant parameter AR model is recursively estimated, which allows for some variation in the estimated parameters. Moreover, for several estimation windows, no time variation in the AR process is found so the AR and time varying AR forecasts are identical.

Moving to the crisis subsample, there are three main observations. First, the relative performance of the TVFAVAR model with changing volatility further improves; it now beats the constant parameter FAVAR model in 38 out of 57 cases (even though, naturally, the absolute performance deteriorates for most variables). The differences in RMSFEs are often statistically significant when the relative RMSFEs are either lower than 0.9 or higher than 1.1, even though the crisis sample is very short. Second, adding time variation to the AR model remains virtually useless. Finally, the TVFAVAR model (with or without changing volatility) becomes the best forecasting model for the majority of cases (34 out of 57), with changes in volatility particularly useful and especially for the real activity variables. These results should be interpreted with care, though, since the evaluation sample contains only 10 quarters. However, they suggest that several cross-variable relationships changed during the latest crisis period, as well as the volatility of the shocks driving the variables.

7. Structural analysis

In this section we examine how the transmission of monetary policy in the USA has changed over time. We focus on the sample from 1972, quarter 1, to 2007, quarter 2, to end before the global financial crisis when monetary policy approached the zero lower bound and unconventional monetary policy measures weakened the usefulness of the Federal Funds rate to identify monetary policy shocks. (As a check of robustness, we have also conducted the analysis for the sample period ending just before the Federal Reserve undertook unconventional measures (in 2008, quarter 2). Our main results are unaffected.) We first discuss why changes in the transmission mechanism of monetary policy may have occurred over the sample period and provide an overview of the existing empirical evidence. We then present new evidence based on our TVFAVAR model. We explain how we estimate the latent factors in the structural setting, how we identify monetary policy shocks and how we compute impulse response functions and standard errors around them. As an illustration of our proposed time varying modelling approach, we provide evidence on the time variation in the volatility of monetary policy shocks and assess the evolution in their transmission to a large set of economic variables.

7.1. Existing empirical evidence and possible reasons for changes in the monetary transmission mechanism

The monetary transmission mechanism in the USA may have changed over the period under investigation (1972–2007) as a consequence of several structural changes which comprise three major aspects. First, there was some variation in the conduct and strategy of monetary policy in the late 1970s–early 1980s with a greater emphasis on price stability and, hence, a better anchoring of long-run inflation expectations; see Boivin and Giannoni (2002) and Galvao and

Marcellino (2010) for evidence. Second, liberalization and innovation in financial markets are certainly relevant, which also mostly occurred in the late 1970s–early 1980s (e.g. Boivin *et al.* (2010)). Third, globalization, i.e. greater trade and financial openness, may have resulted in capital market interest rates being increasingly determined by global developments (see for example Boivin and Giannoni (2009)) rather than by domestic forces such as monetary policy.

Despite numerous studies on this topic, the empirical literature is still lacking a consensus on how the transmission of monetary policy shocks in the USA has changed over time. Table A.2 in the on-line appendix overviews recent time series work on monetary transmission on inflation and activity. The evidence is based on a variety of methods which differ in the way that time variation in the parameters is modelled (split-sample *versus* smooth parameter changes), in the way that monetary policy shocks are identified (recursive identification *versus* sign restrictions) and in the amount of information exploited (small-scale VAR models which use a handful of variables *versus* FAVAR models which exploit hundreds of variables). VAR-based references generally focus on the effect of monetary policy on a single measure of real activity and a single inflation measure whereas FAVAR-based analyses assess a wider spectrum of activity, inflation but also financial measures. Table A.2 shows that the evidence on how the transmission of monetary policy shocks on output and inflation has changed is inconclusive, ranging for example for inflation from a decline in the transmission over time, e.g. Boivin *et al.* (2009), over no change, e.g. Primiceri (2005), to an increase (e.g. Baumeister *et al.* (2013)).

Despite inconclusive results regarding the transmission of monetary policy shocks there is, however, a broad consensus that monetary policy shocks have been large in the early 1980s during the Volcker disinflation and have become smaller since then, e.g. Boivin and Giannoni (2002), Eickmeier and Hofmann (2013), Primiceri (2005) and Canova and Gambetti (2009).

In addition to the above-mentioned structural changes that have occurred either relatively quickly (in the case of institutional changes or changes in the conduct of policy) or gradually (in the case of globalization) and probably have permanent effects on the monetary transmission mechanism, economic frictions may lead to asymmetric responses of the economy to monetary policy shocks over the business cycle. Peersman and Smets (2002), for instance, showed for the euro area that monetary policy shocks have a stronger effect on output and prices in recessions than in booms. Results for the USA are missing to our knowledge.

Although it would certainly be very interesting to shed light on all these possible changes, we need to restrict ourselves in this application of our TVFAVAR model. We focus on changes in the transmission to activity, prices, inflation expectations and long-term interest rates, thus tackling the first and third types of permanent structural change, and we leave changes related to financial markets to future research.

7.2. Monetary policy shock identification

For the structural analysis, it is now assumed that X_t is driven by a $((G+1)\times 1)$-vector consisting of G latent factors F_t^* and the Federal Funds rate i_t as the $(G+1)$th observable factor as in Bernanke *et al.* (2005). We shall use $G=5$ factors (but results are very similar with $G=6$ factors). We estimate the space that is spanned by the factors using the first $G+1$ PCs of the data X_t. To remove the observable factor from the space spanned by all $G+1$ factors we split our data set into slow moving variables, i.e. variables that are expected to move with delay after an interest rate shock, and fast moving variables, i.e. variables that move instantaneously in response to an interest rate shock. The slow moving variables comprise, for example, real activity measures, consumer and producer prices, deflators of GDP and its components and wages, whereas the fast moving variables are financial variables such as asset prices, interest rates or commodity prices (for details see the data table in the on-line appendix). We estimate the first G PCs from the set of slow moving variables, denoted by $\hat{F}_t^{\text{slow}}$. We then carry out a multiple regression of F_t on $\hat{F}_t^{\text{slow}}$ and on i_t, i.e.

$$F_t = a\hat{F}_t^{\text{slow}} + bi_t + \nu_t.$$

An estimate of F_t^* is then given by $\hat{a}\hat{F}_t^{\text{slow}}$. In the joint factor vector $F_t \equiv (\hat{F}_t^*, i_t)$, the Federal Funds rate i_t is ordered last. Given this ordering, the VAR representation of our (time varying) VAR model with lower triangular contemporaneous relationship matrix P_t directly identifies the monetary policy shock as the last element of the innovation vector u_t in expression (2.5). Hence, the shock identification works via a Cholesky decomposition, which is here readily given by the lower triangular $P_{t|T}^{-1}$.

The methodology also allows for other identification approaches, such as sign restrictions which need to be satisfied at each point in time. We have checked that, on the basis of our Cholesky identification scheme, non-borrowed reserves and monetary aggregates decline after an unexpected tightening of monetary policy at all points in time. Hence, our results are consistent with the sign restrictions imposed, e.g. in Uhlig (2005) and Benati and Mumtaz (2007), and also with the 1979–1982 period when the Federal Reserve temporarily targeted non-borrowed reserves as opposed to the Federal Funds rate.

7.3. *Computing time varying impulse responses*

The impulse responses are based on the assumption that the system (shock propagation) remains at its time t estimate from time t henceforth. This is common practice and is consistent with our assumption of random-walk parameter evolution, i.e., at time t, we compute impulse responses in the usual fashion from the estimated VAR model:

$$F_t = P_{t|T}^{-1}\mathcal{K}_{1,t|T}F_{t-1} + \ldots + P_{t|T}^{-1}\mathcal{K}_{p,t|T}F_{t-p} + w_t,$$
$$E(w_t w_t') = P_{t|T}^{-1}\hat{S}_t P_{t|T}^{-1'},$$

in conjunction with the estimated loading equations

$$x_{i,t} = \Lambda_{i,t|T}' F_t + \tilde{e}_{i,t}.$$

Confidence bands for the impulse response functions at time t are computed as follows. Recall that we have obtained from the Kalman smoother the estimates of the states $a_{t|T}^g$ (containing the respective elements of the rows of P and $\mathcal{K}$), and the corresponding variance–covariance matrices $\Sigma_{t|T}^g$ for each VAR equation $g = 1, \ldots, G+1$. Moreover, we have for the loading equations the smoothed $\hat{\Lambda}_{i,t|T}$ with the corresponding variance–covariance matrices $V_{i,t|T}$. We generate draws of α^g, $g = 1, \ldots, G+1$, from $N(a_{t|T}^g, \Sigma_{t|T}^g)$. If the VAR matrices that are implied by the set of draws satisfy the non-explosiveness condition, we keep the draw; otherwise we discard it and repeat the previous step. We draw until we have gathered $K = 1000$ successful draws. We then draw K times Λ_i from $N(\hat{\Lambda}_{i,t|T}, V_{i,t|T})$. For a given time t, variable i and horizon h, the desired quantiles of the impulse response function are then obtained from the K draws. A *caveat* of this approach is that we ignore the uncertainty that is associated with the estimation of the hyperparameters.

7.4. *Monetary policy shocks and transmission in our time varying factor-augmented vector auto-regressive model*

We have reported in Section 4 to what extent the volatility estimates of the VAR innovations to unidentified factors are varying over time. Fig. 4 now shows the estimated volatility of the

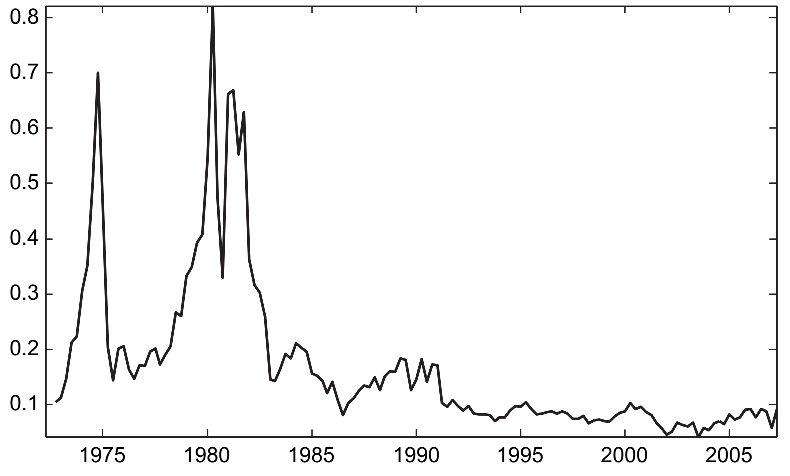

Fig. 4. Time varying volatility of the monetary policy shock

monetary policy shock. Consistent with the literature, the volatility peaks in the early 1980s, which is generally labelled the 'Volcker disinflation', and declines thereafter. We also observe a peak around 1974. One explanation might be that, possibly because of overestimation of the negative effects on activity of the oil embargo in October 1973, the output gap was substantially underestimated and, hence, the Federal Funds rate was much lower than that implied by a simple Taylor rule; see Orphanides (2003). We find indeed a large sequence of expansionary monetary policy shocks around 1974 (not shown) and heightened volatility of the shocks which might reflect this misperception.

On the basis of the TVFAVAR model and the identification scheme described we now assess the evolution of selected impulse response functions to a monetary policy shock over time. We focus on three questions.

(a) Has the transmission to key macroeconomic variables changed over time and, if so, how?
(b) Can we detect asymmetries in the monetary transmission and, more specifically, are monetary policy shocks transmitted to economic activity more strongly during recessions than during booms?
(c) Has the transmission to inflation expectations and long-term interest rates changed over time and, if so, how?

Figs 5, 6 and 7 show impulse response functions of three key macroeconomic variables (the Federal Funds rate, GDP and the GDP deflator). Figs 8, 9 and 10 show impulse response functions of additional activity and price variables (consumption, investment, industrial production, employment, GDP deflator, PPI finished goods, the personal consumption expenditure (PCE) deflator and unit labour costs) and Figs 11, 12 and 13 show impulse response functions of two inflation expectation measures (taken from the Survey of Professional Forecasters and the survey conducted by the University of Michigan) and the 10-year government bond rate. To focus on transmission only, we show estimates of impulse response functions to a monetary policy shock which raises the Federal Funds rate on impact by 1 percentage point. Figs 5, 8 and 11 show averages of point estimates of impulse responses over the entire sample 1972–2007 (the dotted curves) and, for comparison, impulse responses derived from a constant parameter FAVAR model (full curves). In Figs 6, 9 and 12 we present impulse responses obtained from the TVFAVAR model for each point in time and horizons 0–20 quarters, and, for better vis-

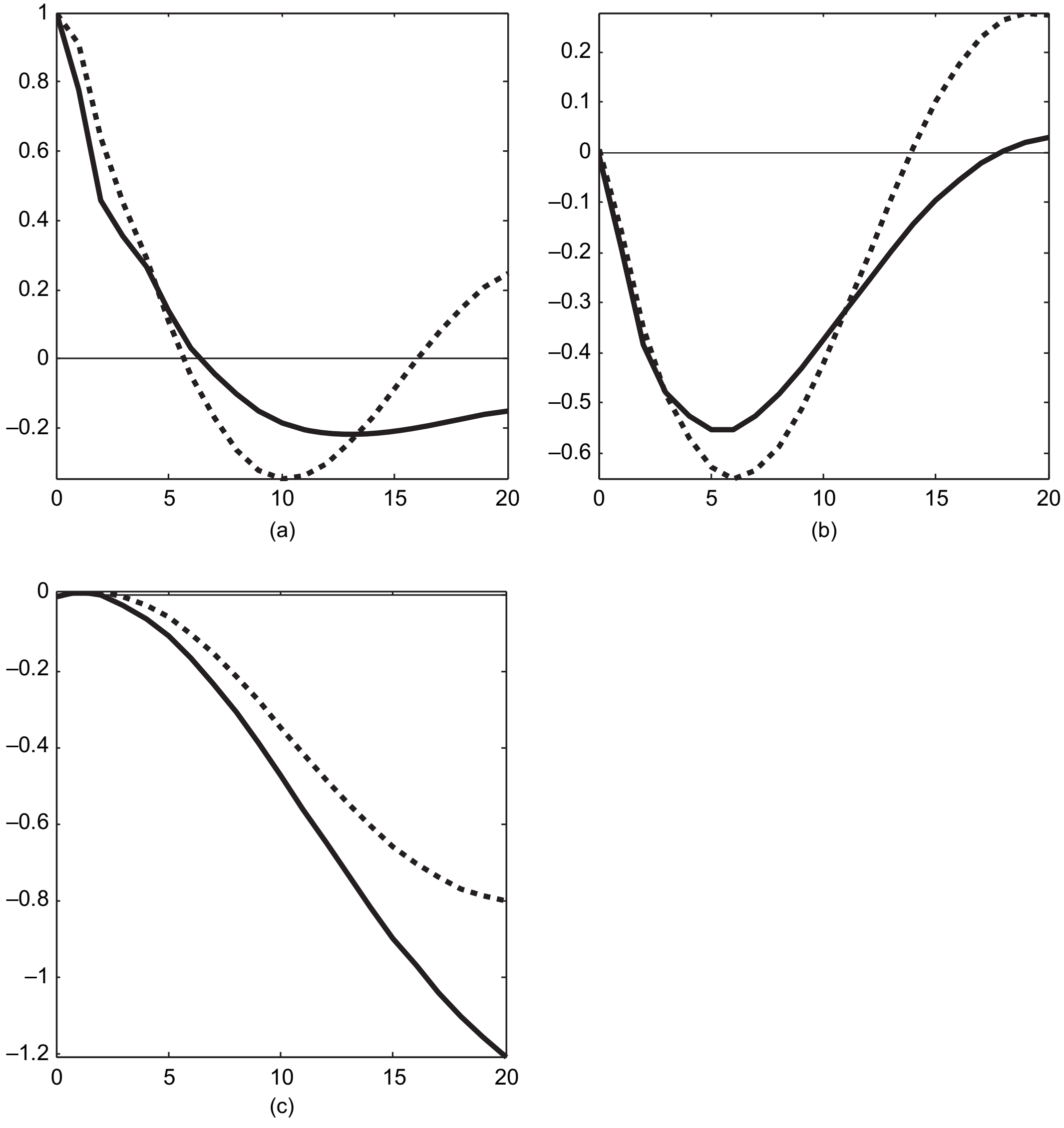

Fig. 5. Impulse response functions of key variables to an unexpected increase in the monetary policy rate by 1 percentage point from a constant parameter FAVAR model (———) and the TVFAVAR model (averages over all periods) (· · · · · ·): (a) Federal Funds rate; (b) GDP (%); (c) GDP deflator (%)

ibility of time variation, we present in Figs 7, 10 and 13 point estimates and 90% confidence bands of impulse responses for each point in time and selected horizons (one, four and eight quarters).

Focusing first on Figs 5, 8 and 11, the constant parameter impulse responses have the expected shape. After an unexpected increase in the Federal Funds rate, GDP and other activity variables decline temporarily and in a hump-shaped manner. The impulse responses then turn to zero after 3–5 years, depending on the activity measure, consistent with real long-run neutrality of monetary policy. The GDP deflator declines persistently. There is no 'price puzzle', i.e. a

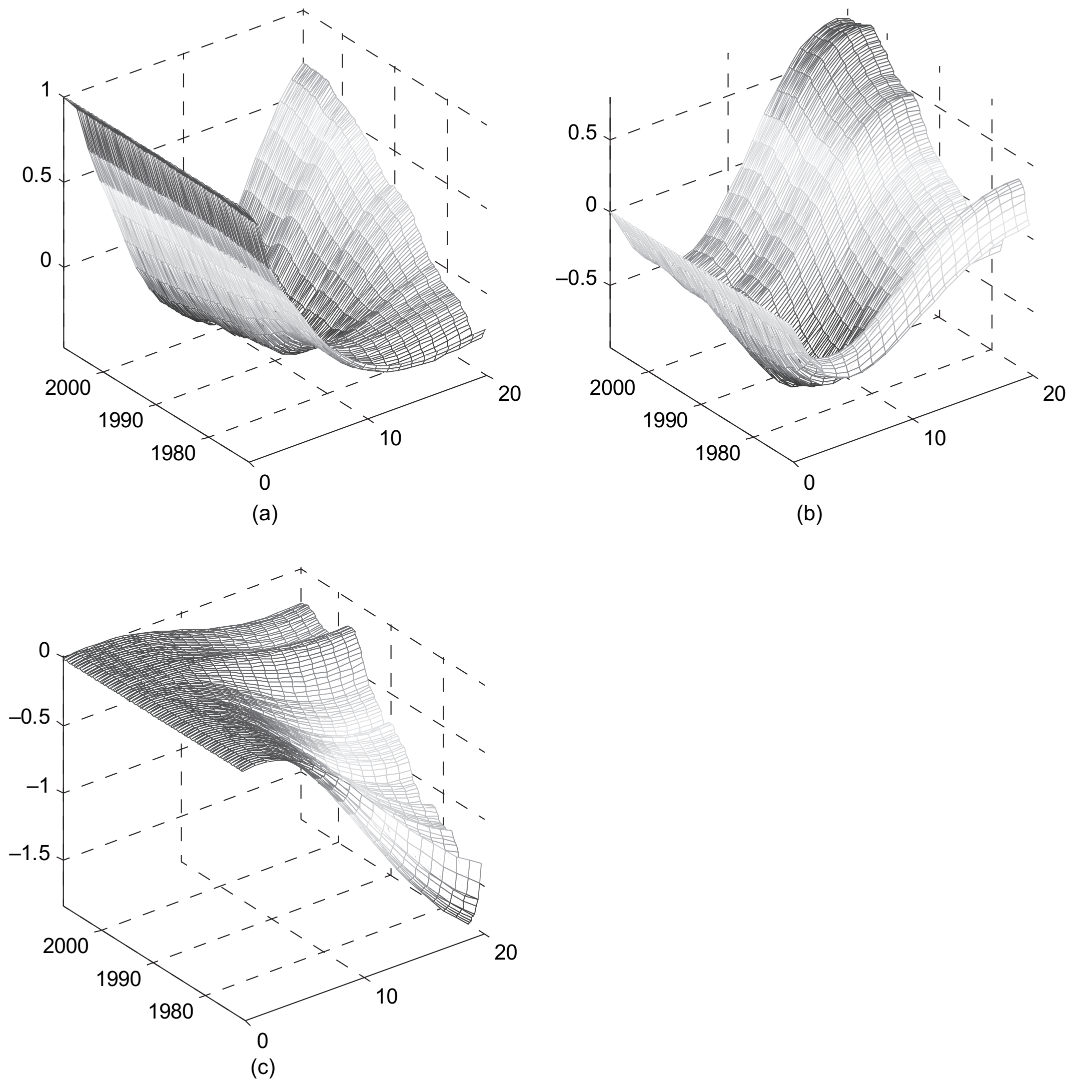

Fig. 6. Impulse response functions of key variables to an unexpected increase in the monetary policy rate by 1 percentage point from the TVFAVAR model (all horizons and points in time): (a) Federal Funds rate; (b) GDP(%); (c) GDP deflator (%)

significantly positive response of prices after a monetary policy tightening, unlike what is found in many empirical monetary studies which use small dimensional models; see Bernanke *et al.* (2010) for a discussion. The graphs for CPI, PPI and unit labour costs display a similar pattern.

Inflation expectations also decline after the shock, although the SPF measure first temporarily increases, which is a pattern that was also found by Boivin *et al.* (2010). Long-term interest rates, reflecting expected future short-term rates and possibly term premia, increase by less than the Federal Funds rate. Figs 5, 8 and 11 also reveal that averages of the time varying impulse responses are similar to their constant parameter counterparts.

Let us now answer the questions related to time variation that were raised at the beginning of this section.

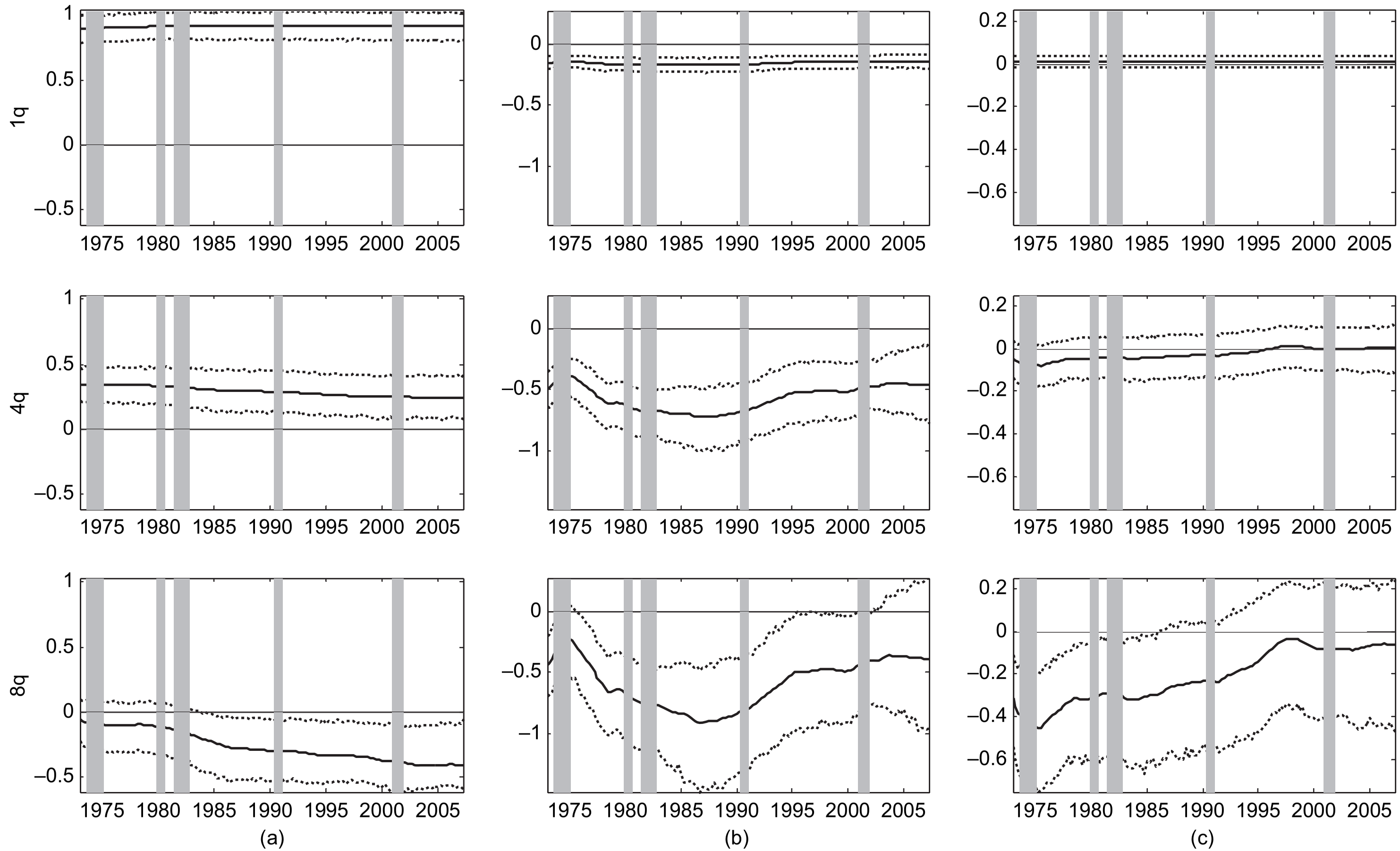

Fig. 7. Impulse response functions of key variables to an unexpected increase in the monetary policy rate by 1 percentage point from the TVFAVAR model (selected horizons) (· · · · · ·, 90% confidence bands; |, National Bureau of Economic Research recessions): (a) Federal Funds rate; (b) GDP (%); (c) GDP deflator (%)

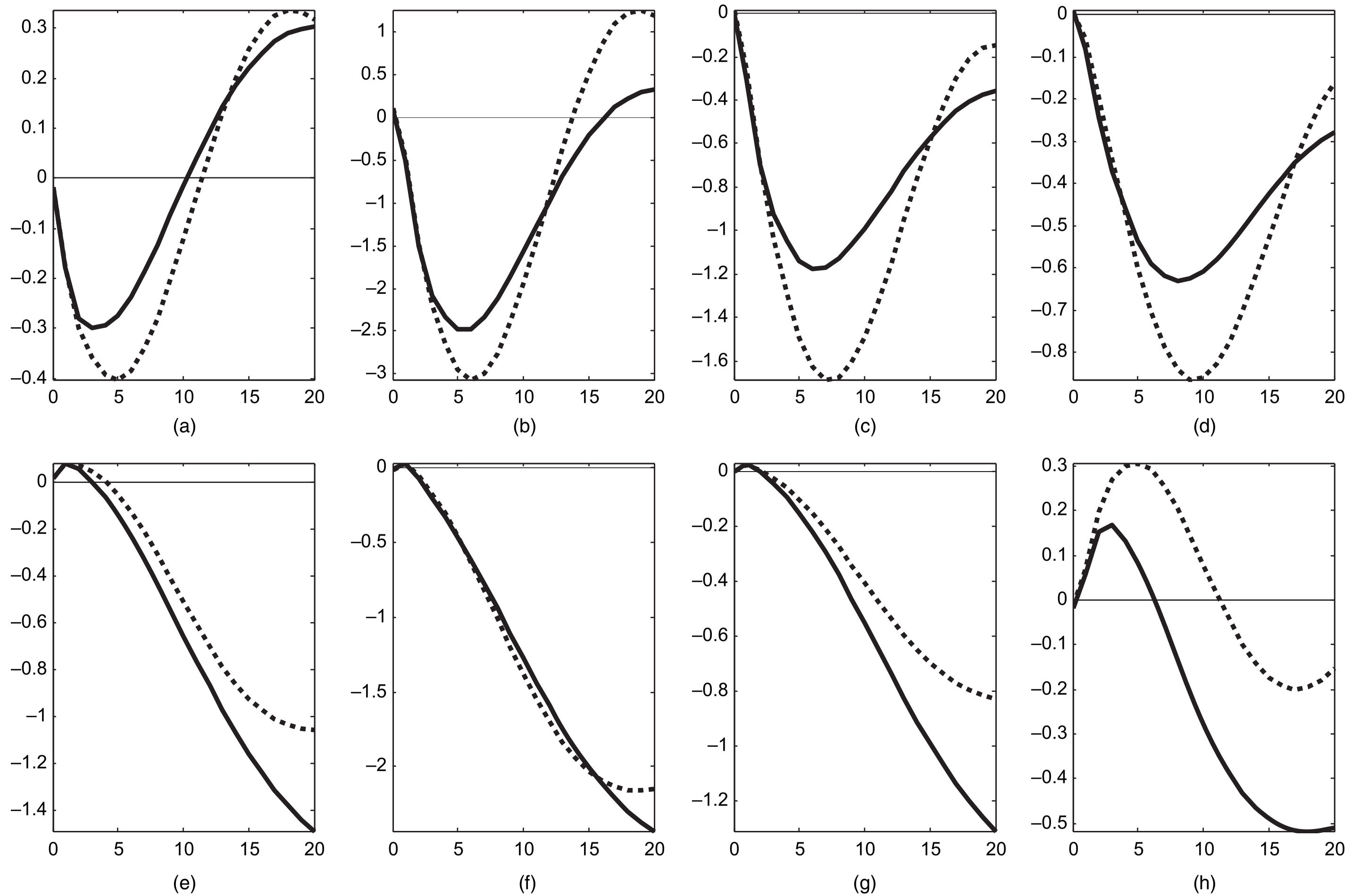

Fig. 8. Impulse response functions of additional activity and price variables from a constant parameter FAVAR model (———) and the TVFAVAR model (averages over all periods) (· · · · · ·): (a) consumption (%); (b) investment (%); (c) industrial production (%); (d) employment (%); (e) CPI (%); (f) PPI (%); (g) PCE deflator (%); (h) unit labour cost manufacturing (%)

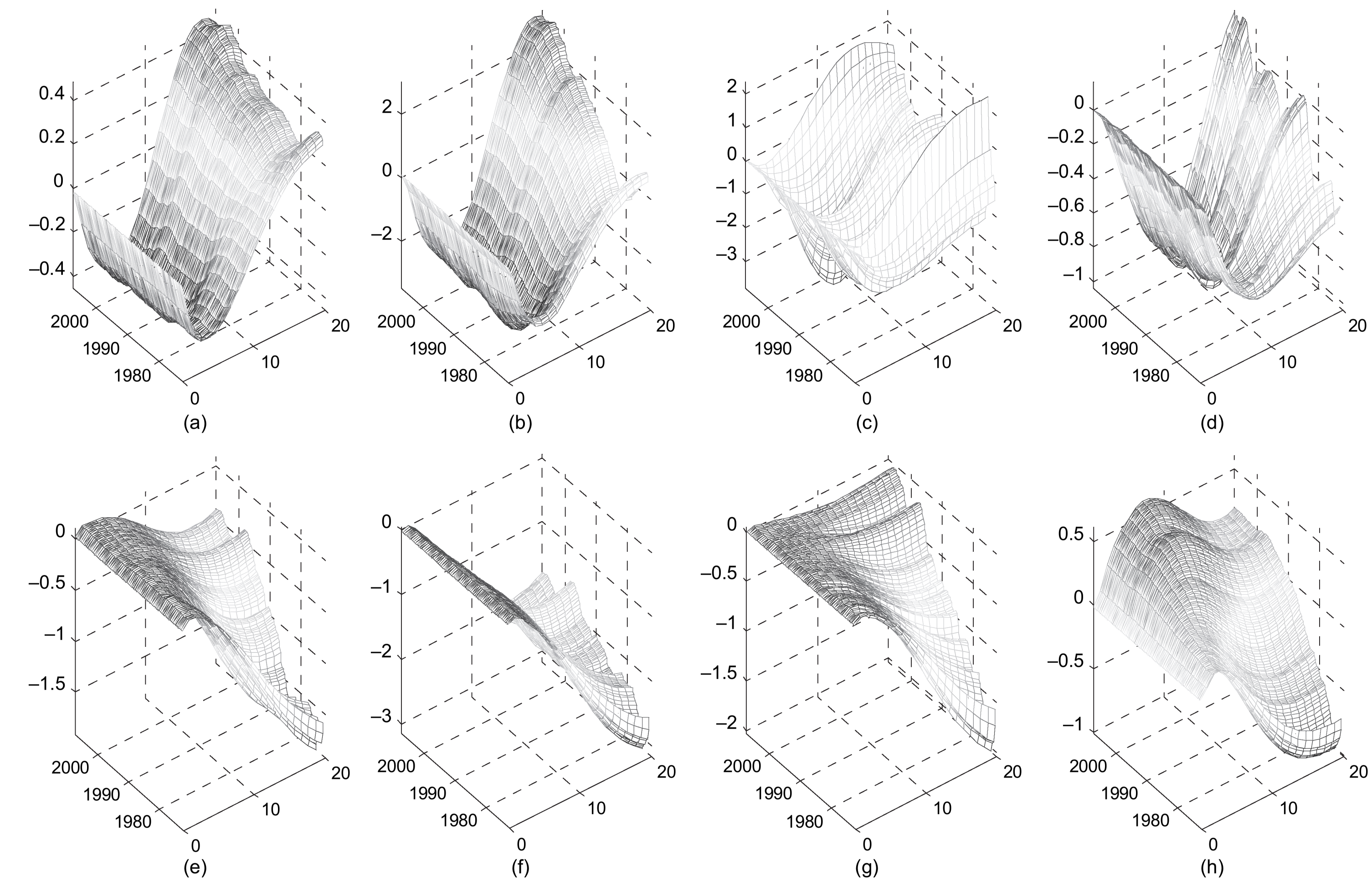

Fig. 9. Impulse response functions of additional activity and price variables from the TVFAVAR model (all horizons and points in time): (a) consumption (%); (b) investment (%); (c) industrial production (%); (d) employment (%); (e) CPI (%); (f) PPI (%); (g) PCE deflator (%); (h) unit labour cost manufacturing (%)

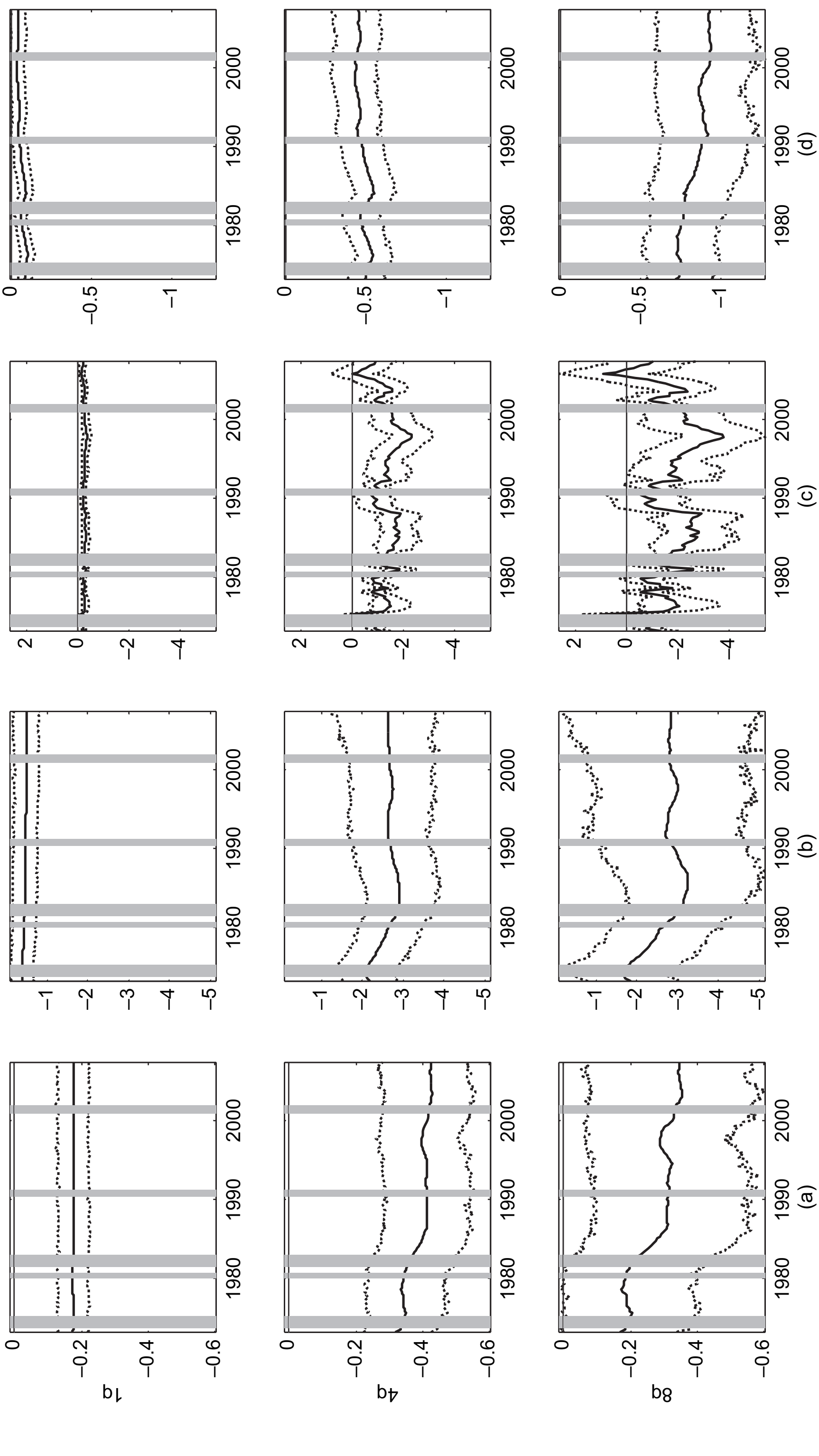

Fig. 10. Impulse response functions of additional activity and price variables from the TVFAVAR model (selected horizons) (· · · · · ·, 90% confidence bands; |, National Bureau of Economic Research recessions): (a) consumption (%); (b) investment (%); (c) industrial production (%); (d) employment (%); (e) CPI (%); (f) PPI (%); (g) PCE deflator (%); (h) unit labour cost manufacturing (%)

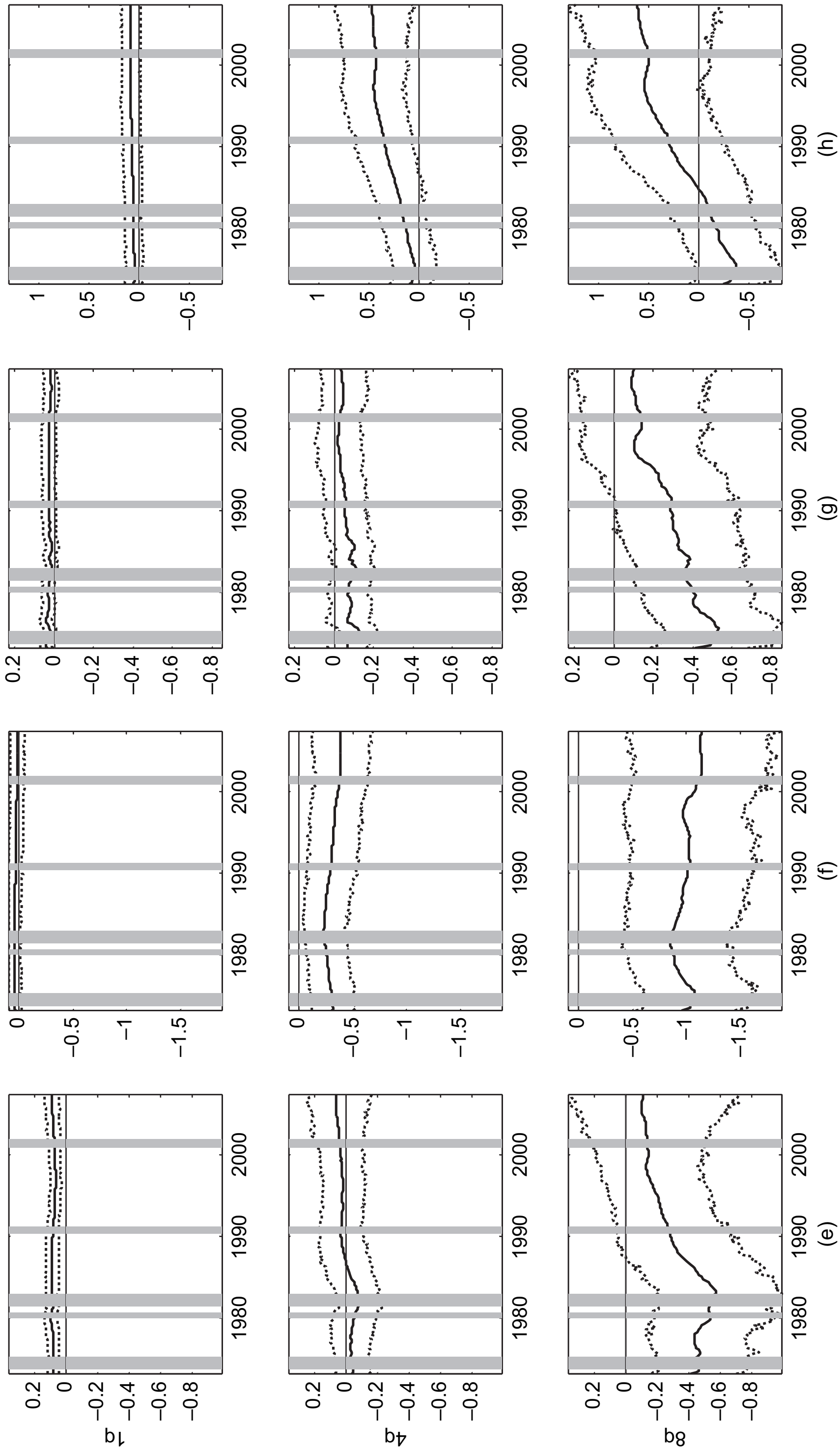

Fig. 10 (*continued*)

(a) Figs 5 and 6 reveal that, whereas the effect of monetary policy shocks on the Federal Funds rate itself has not changed much, there are notable changes in the impulse responses of GDP and the GDP deflator over time. Whereas the effects on GDP and the GDP deflator after one quarter have barely changed, the effects at longer horizons are estimated to have considerably weakened since the 1980s, in line with Boivin and Giannoni (2009) and Eickmeier and Hofmann (2013). The pattern observed for GDP carries over to industrial production and employment, but not to the other real activity variables. The pattern that is observed for the GDP deflator is also apparent in the graphs for CPI, the PCE deflator and unit labour costs, but not for PPI (Figs 8–10).

(b) Inspection of the time varying impulse responses of the activity variables (Figs 7 and 10) does not point to sizably different effects of monetary policy shocks during recessions *versus* expansions. Hence, unlike Peersman and Smets (2002) for the euro area, we do not find evidence of asymmetry in the monetary transmission for the USA. One possible explanation for this discrepancy between the findings for the two regions is that there are less frictions in the USA than in the euro area economy. Another explanation might be that Peersman and Smets (2002) modelled parameter variation differently, allowing parameters to take only two values, one for recessions and one for booms, whereas we also allow for gradual changes and trending parameters over time. Moreover, it is apparent from Peersman and Smets (2002) that the asymmetry that was found might be due to a stronger effect also on the policy rate whereas we find no asymmetries in the effect of policy shocks on the Federal Funds rate.

(c) Figs 11–13 finally show that the negative effect on inflation expectations has become smaller over time, in line with Boivin *et al.* (2010). The decline starts in the 1970s for both inflation expectation measures. The changes for the Survey of Professional Forecasters measure is mostly apparent for longer horizons. The timing of the decline is roughly consistent with a change in the conduct of monetary policy towards more aggressive reactions to output and inflation and, consequently, a better anchoring of long-term inflation expectations. A smaller response of inflation expectations may have also contributed to a decline in the effect on the term premium and, hence, long-term interest rates which is, however, only apparent for short horizons. A further interesting point is that this decline started in the mid-1980s, and—at least the timing—is consistent with the initial years of globalization; see Kosel *et al.* (2006). A smaller effect on long-term rates and inflation expectations may also have contributed to the weakening of the negative responses of output and price measures.

Summing up, our results confirm previous findings in the literature that the size of monetary policy shocks has decreased since the early 1980s. We find weaker effects on activity and prices, which could be partly due to a better conduct of monetary policy and, consequently, to a better anchoring of inflation expectations and, possibly, globalization. Finally, we do not find evidence for different reactions of activity variables to monetary policy shocks in recessions *versus* non-recession periods.

8. Conclusions

In this paper we have proposed a FAVAR specification that is suited to model large data sets allowing for general patterns of time variation in the factor loadings, the factor dynamics and their innovation variance–covariance structure. Contrarily to previous literature, which is mostly Bayesian, we propose a fully classical (i.e. maximum-likelihood-based) approach for estimation, inference, forecasting and structural analysis.

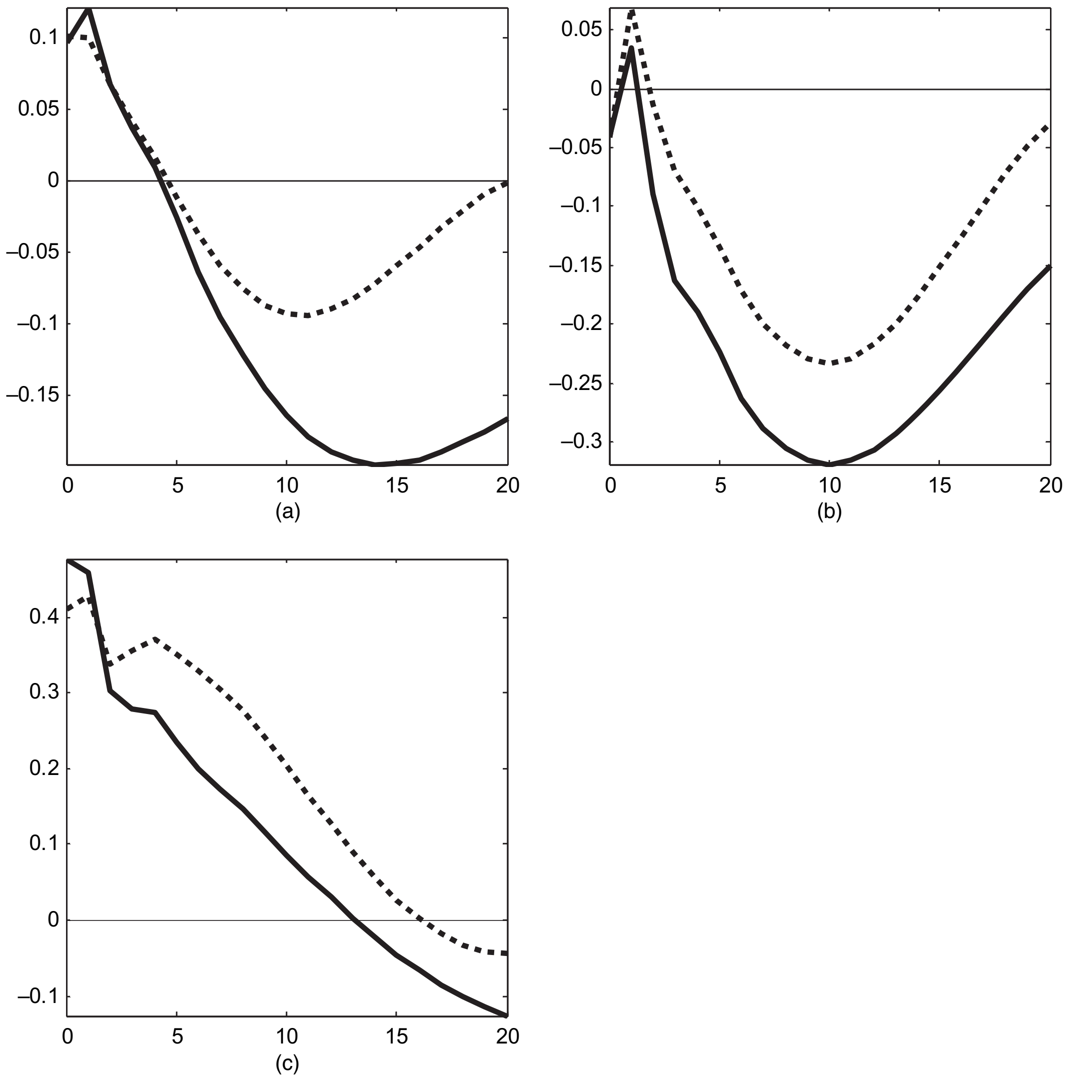

Fig. 11. Impulse response functions (percentage points) of inflation expectations and long-term government bond yields from the constant parameter FAVAR model (———) and the TVFAVAR model (averages over all periods) (· · · · · ·): (a) Survey of Professional Forecasters inflation expectation; (b) University of Michigan survey inflation expectation; (c) 10-year government bond yield

The three main technical features underlying our approach are, first, the use of PC-based factor estimates (which is justified by the theoretical results in Stock and Watson (2002a, b, 2008)), second, a representation of the factor dynamics as a VAR process with triangular contemporaneous structure, which renders equation-by-equation estimation feasible, and, third, a specification of volatility as a function of past factors. MC evidence suggests that our estimation procedure works well, in particular when the numbers of variables and observations are sufficiently large.

When our TVFAVAR method is employed to model a large data set of US variables over the period 1972–2012, several interesting results emerge. First, we identify minor changes in the factor dynamics and contemporaneous relationships, but much more marked variation in factor

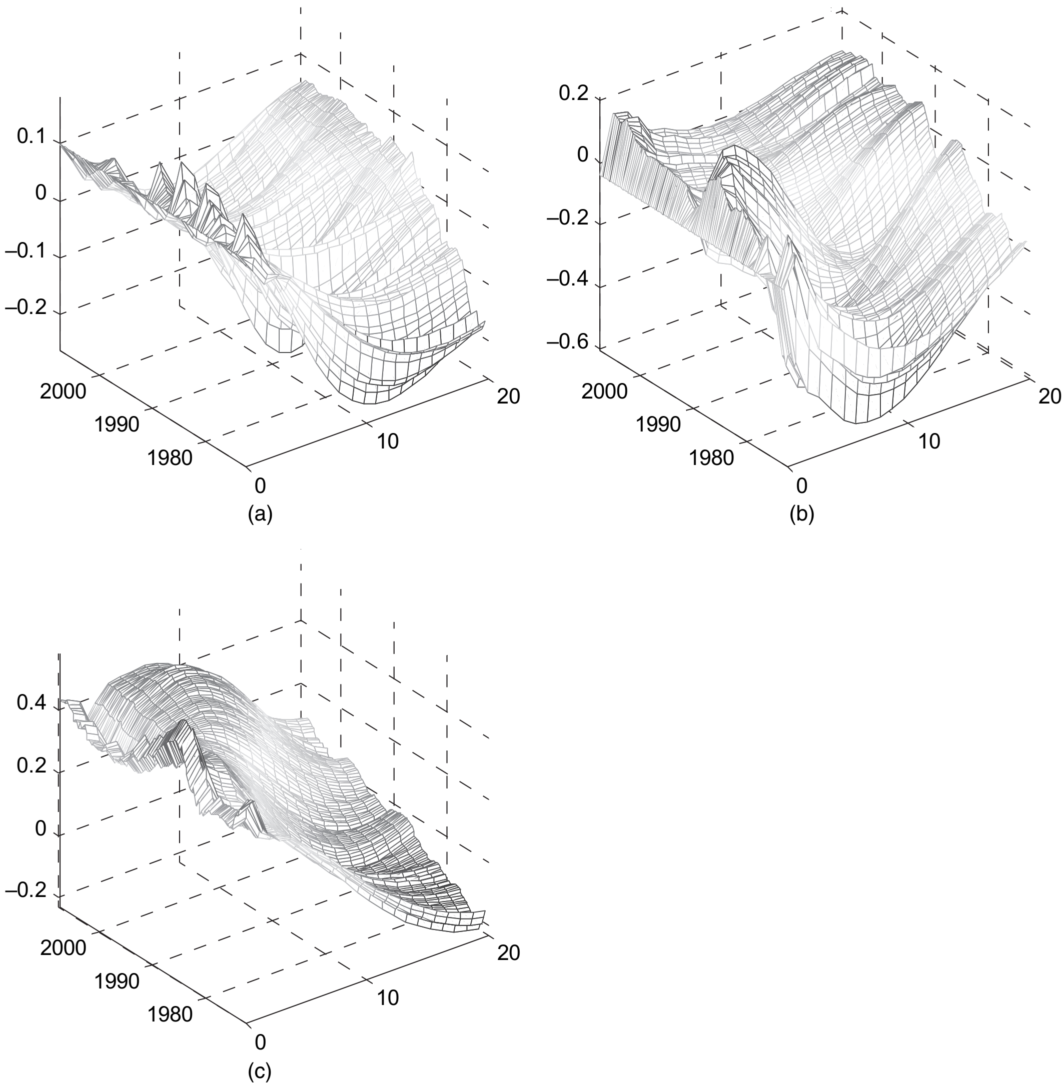

Fig. 12. Impulse response functions (percentage points) of inflation expectations and long-term government bond yields from the TVFAVAR model (all horizons and points in time): (a) Survey of Professional Forecasters inflation expectation; (b) University of Michigan survey inflation expectation; (c) 10-year government bond yield

volatility and their direct effect on key macroeconomic variables. Therefore, according to our model, changes both in the volatility of the shocks and in their transmission to the economy matter.

Second, pseudorealtime forecasts from the TVFAVAR model are more accurate than those from a constant parameter FAVAR model for several variables, in particular during the latest crisis period and when also allowing for changes in the volatility. The constant parameter FAVAR model performs well since recursive estimation introduces anyway a form of parameter time variation, which is, however, at odds with the underlying assumption of stable parameters.

Finally, we illustrate how the TVFAVAR model can be used to identify monetary policy shocks and their transmission to the economy. We find that the volatility of monetary shocks is sub-

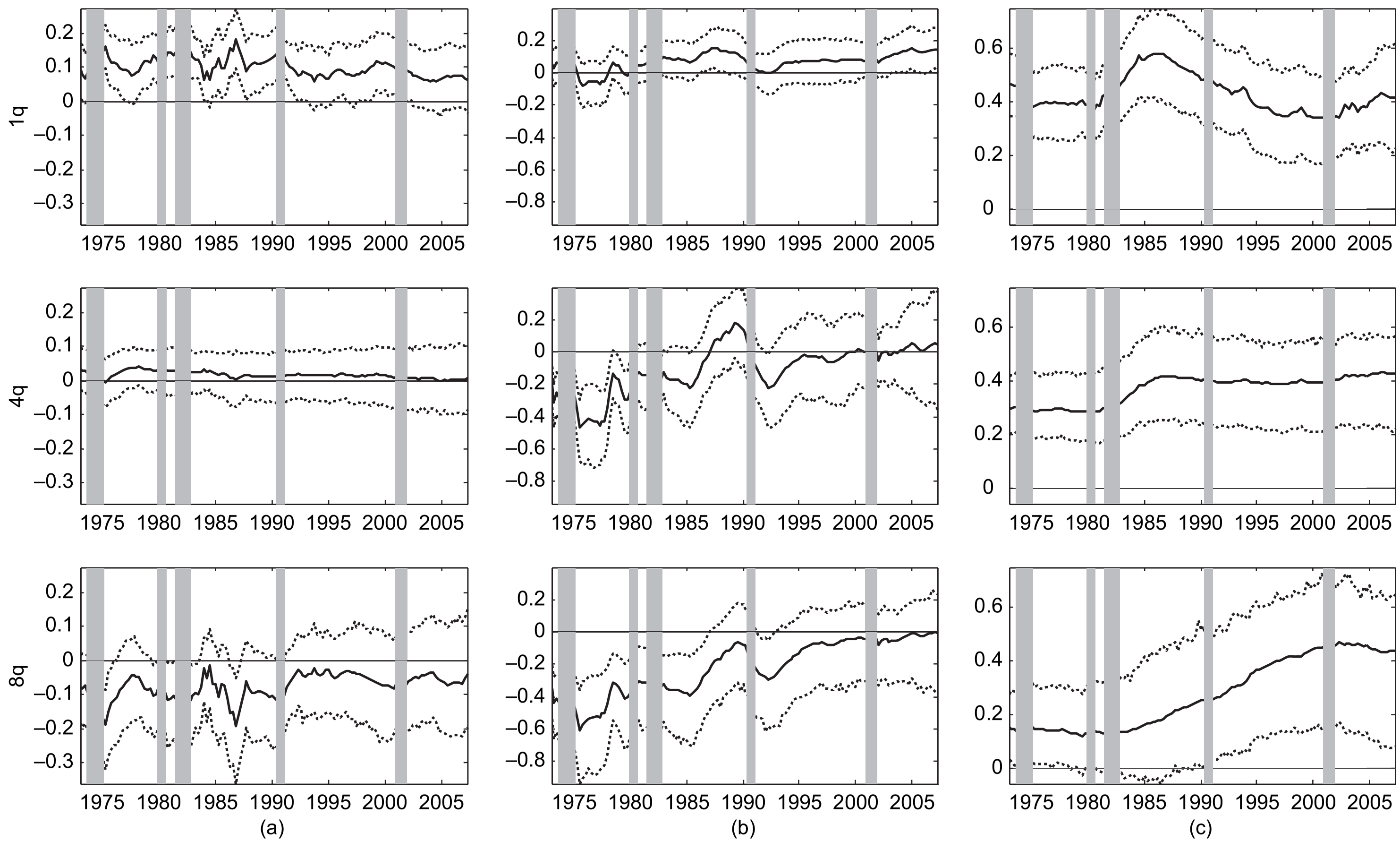

Fig. 13. Impulse response functions (percentage points) of inflation expectations and long-term government bond yields from the TVFAVAR model (selected horizons) (· · · · · ·, 90% confidence bands; ▮, National Bureau of Economic Research recessions): (a) Survey of Professional Forecasters inflation expectation; (b) University of Michigan survey inflation expectation; (c) 10-year government bond yield

stantially smaller after the early 1980s and that a constant size shock appears to have smaller effects on GDP, prices, inflation expectations and long-term interest rates over the more recent period, which are consistent with changes in the conduct of monetary policy and, consequently, a better anchoring of inflation expectations and, possibly, globalization. Moreover, we do not find evidence that the real economy reacts differently to monetary policy shocks in recession periods compared with expansions.

Acknowledgements

The views expressed in this paper are those of the authors and do not necessarily represent the view of the European Central Bank or the Deutsche Bundesbank. Sandra Eickmeier thanks the Monetary Policy Strategy Division of the European Central Bank for its hospitality. We thank two referees, Herman van Dijk, Sylvia Kaufmann, Hashem Pesaran, Lucrezia Reichlin, Jim Stock, Harald Uhlig, Giovanni Urga and Mark Watson, as well as participants at the National Bureau of Economic Research Summer Institute 2010 (Boston) and the conference on 'High-dimensional econometric modelling' at Cass Business School (London), a joint Swiss National Bank–Oesterreichische Nationalbank–Bundesbank workshop and seminar participants at the Deutsche Bundesbank and the European Central Bank for useful comments and discussions. We are also grateful to Christiane Baumeister for useful discussions and for providing us with commodity and PCE price data. Many thanks go to Angela Abbate, Guido Schultefrankenfeld and Michael Richter for their help on the data sets.

References

Bai, J. and Ng, S. (2006) Confidence intervals for diffusion index forecasts with a large number of predictors and inference for factor-augmented regressions. *Econometrica*, **74**, 1133–1150.

Banerjee, A., Marcellino, M. and Masten, I. (2008) Forecasting macroeconomic variables using diffusion indexes in short samples with structural change. In *Forecasting in the Presence of Structural Breaks and Model Uncertainty* (eds M. E. Wohar and D. E. Rapach). New York: Elsevier.

Bates, B. J., Plagborg-Møller, M., Stock, J. H. and Watson, M. W. (2013) Consistent factor estimation in dynamic factor models with structural instability. *J. Econmetr.*, **177**, 289–304.

Baumeister, C., Liu, P. and Mumtaz, H. (2013) Changes in the effects of monetary policy on disaggregate price dynamics. *J. Econ. Dynam. Control*, **37**, 543–560.

Benati, L. and Mumtaz, H. (2007) U.S. evolving macroeconomic dynamics—a structural investigation. *Working Paper 746*. European Central Bank, Frankfurt am Main.

Bernanke, B., Boivin, J. and Eliasz, P. (2005) Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *Q. J. Econ.*, **120**, 387–422.

Boivin, J. and Giannoni, M. P. (2002) Assessing changes in the monetary transmission mechanism. *Fed. Resve Bnk New York Econ. Pol. Rev*., May, 97–111.

Boivin, J. and Giannoni, M. P. (2009) Global forces and monetary policy effectiveness. In *International Dimensions of Monetary Policy* (eds J. Gali and M. Gertler), ch. 8, pp. 429–478. Chicago: University of Chicago Press.

Boivin, J., Giannoni, M. P. and Mihov, I. (2009) Sticky prices and monetary policy: evidence from disaggregated US data. *Am. Econ. Rev.*, **99**, 350–384.

Boivin, J., Kiley, M. T. and Mishkin, F. S. (2010) How has the monetary transmission mechanism evolved over time? In *Handbook of Monetary Economics*, vol. 3 (eds B. M. Friedman and M. Woodford), ch. 8, pp. 369–422. New York: Elsevier.

Canova, F. and Gambetti, L. (2009) Structural changes in the US economy: is there a role for monetary policy? *J. Econ. Dynam. Control*, **33**, 477–490.

Clark, T. and McCracken, M. (2011a) Nested forecast model comparisons: a new approach to testing equal accuracy. *Mimeo*. Federal Reserve Bank of St Louis, St Louis.

Clark, T. and McCracken, M. (2011b) Testing for unconditional predictive ability. In *Oxford Handbook of Economic Forecasting* (eds P. Clements and F. Hendry). Oxford: Oxford University Press.

Cogley, T. and Sargent, T. J. (2005) Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Rev. Econ. Dynam*., **8**, 262–302.

Del Negro, M. and Otrok, C. (2008) Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *Staff Report 326*. Federal Reserve Bank of New York, New York.

Diebold, F. and Mariano, R. (1995) Comparing predictive accuracy. *J. Bus. Econ. Statist.*, **13**, 253–263.
Eickmeier, S. and Hofmann, B. (2013) Monetary policy, housing booms and financial (im)balances. *Macrecon. Dynam.*, **17**, 830–870.
Eickmeier, S., Lemke, W. and Marcellino, M. (2011a) The changing international transmission of financial shocks: evidence from a classical time-varying FAVAR. *Discussion Paper 05/2011*. Deutsche Bundesbank, Frankfurt am Main.
Eickmeier, S., Lemke, W. and Marcellino, M. (2011b) Classical time-varying FAVAR models—estimation, forecasting and structural analysis. *Discussion Paper 04/2011*. Deutsche Bundesbank, Frankfurt am Main.
Galvao, A. and Marcellino, M. (2010) Endogenous monetary policy regimes and the Great Moderation. *Discussion Paper 7827*. Centre for Economic Policy Research, London.
Harvey, S. L. and Newbold, P. (1997) Testing the equality of prediction mean squared errors. *Int. J. Forecast.*, **13**, 281–291.
Harvey, A. C., Ruiz, E. and Sentana, E. (1992) Unobserved component time series models with ARCH disturbances. *J. Econmetr.*, **52**, 129–157.
Korobilis, D. (2013) Assessing the transmission of monetary policy using dynamic factor models. *Oxf. Bull. Econ. Statist.*, **75**, 157–179.
Kose, A., Prasad, E. and Terrones, M. (2006) How do trade and financial integration affect the relationship between growth and volatility? *J. Int. Econ*., **69**, 176–202.
Liu, P., Mumtaz, H. and Theophilopoulou, A. (2011) International transmission of shocks: a time-varying factor-augmented var approach to the open economy. *Working Paper 425*. Bank of England, London.
Mumtaz, H. and Surico, P. (2012) Evolving international inflation dynamics: world and country-specific factors. *J. Eur. Econ. Ass.*, **10**, 716–734.
Nyblom, J. (1989) Testing for the constancy of parameters over time. *J. Am. Statist. Ass.*, **84**, 223–230.
Orphanides, A. (2003) Historical monetary policy analysis and the Taylor Rule. *J. Monet. Econ.*, **50**, 983–1022.
Peersman, G. and Smets, F. (2002) Are the effects of monetary policy greater in recessions than in booms? In *Monetary Transmission in Diverse Economies* (eds L. Mahadeva and P. Sinclair), pp. 36–55. Cambridge: Cambridge University Press.
Primiceri, G. (2005) Time varying structural vector autoregressions and monetary policy. *Rev. Econ. Stud.*, **72**, 821–852.
Sims, C. A. and Zha, T. (2006) Were there regime switches in monetary policy? *Am. Econ. Rev*., **96**, 54–81.
Stock, J. and Watson, M. (1998) Median unbiased estimation of coefficient variance in a time-varying parameter model. *J. Am. Statist. Ass.*, **93**, 349–358.
Stock, J. and Watson, M. (2002a) Forecasting using principal components from a large number of predictors. *J. Am. Statist. Ass.*, **97**, 1167–1179.
Stock, J. and Watson, M. (2002b) Macroeconomic forecasting using diffusion indexes. *J. Bus. Econ. Statist.*, **20**, 147–162.
Stock, J. and Watson, M. (2005) Implications of dynamic factor models for VAR analysis. *Working Paper 11467*. National Bureau of Economic Research, Cambridge.
Stock, J. and Watson, M. (2008) Forecasting in dynamic factor models subject to structural instability. In *The Methodology and Practice of Econometrics, a Festschrift in Honour of Professor David F. Hendry* (eds J. Castle and N. Shephard). Oxford: Oxford University Press.
Uhlig, H. (2005) What are the effects of monetary policy on output?: results from an agnostic identification approach. *J. Monet. Econ.*, **52**, 381–419.

Supporting information
Additional 'supporting information' may be found in the on-line version of this article:

'Appendix Table A.1: Data'.