DISCUSSION PAPER SERIES

No. 10610

MONETARY, FISCAL AND OIL SHOCKS: EVIDENCE BASED ON MIXED FREQUENCY STRUCTURAL FAVARS

Massimiliano Marcellino and Vasja Sivec

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

MONETARY, FISCAL AND OIL SHOCKS: EVIDENCE BASED ON MIXED FREQUENCY STRUCTURAL FAVARS

Massimiliano Marcellino and Vasja Sivec

Discussion Paper No. 10610
May 2015
Submitted 13 May 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

MONETARY, FISCAL AND OIL SHOCKS: EVIDENCE BASED ON MIXED FREQUENCY STRUCTURAL FAVARS[†]

Abstract

Large scale factor models have been often adopted both for forecasting and to identify structural shocks and their transmission mechanism. Mixed frequency factor models have been also used in a reduced form context, but not for structural applications, and in this paper we close this gap. First, we adapt a simple technique developed in a small scale mixed frequency VAR and factor context to the large scale case, and compare the resulting model with existing alternatives. Second, using Monte Carlo experiments, we show that the finite sample properties of the mixed frequency factor model estimation procedure are quite good. Finally, to illustrate the method we present three empirical examples dealing with the effects of, respectively, monetary, oil, and fiscal shocks.

JEL Classification: C32, C43 and E32
Keywords: estimation, identification, impulse response function, mixed frequency data, structural FAVAR and temporal aggregation

Massimiliano Marcellino massimiliano.marcellino@unibocconi.it
Bocconi University, IGIER and CEPR

Vasja Sivec vasja.sivec@eui.eu
European University Institute

[†] We would like to thank two anonymous Referees, Juan Dolado, Eric Ghysels, Domenico Giannone, Peter Hansen and seminar participants at the EUI for useful comments on a previous draft. All remaining errors are our own responsibility.

1 Introduction

Since the pioneering work of Sims [1980], vector auto-regressions (hereafter VARs) became a dominant device to identify structural shocks and investigate their propagation mechanism. But VARs are not without flaws. To prevent the curse of dimensionality, they are estimated on a small set of macroeconomic variables. In contrast, economic agents and decision makers generally consider a large set of variables when making their decisions. This discrepancy in information sets can generate statistically biased shock responses and economically counterintuitive results. For example, a typical monetary policy VAR suffers from a price puzzle, namely, after a negative monetary policy shock (unexpected rise in the policy rate) prices initially increase.

Recently, Bernanke, Boivin, and Eliasz [2005] introduced a way to overcome the curse of dimensionality in a structural VAR, see also Marcellino, Favero, and Neglia [2005], Forni, Giannone, Lippi, and Reichlin [2009] and Andreou, Gagliardini, and Ghysels [2013]. The relevant large set of economic variables are assumed to be generated by a factor model, where few common factors explain the bulk of the variation in all the variables and therefore provide an exhaustive summary of the relevant information. Factors, generally estimated by (static or dynamic) principal components, do not have a clear economic interpretation. However, they can be modeled with a VAR, possibly augmented with a few observable variables, and the VAR used to identify structural shocks. In a second step, Bernanke et al. [2005] estimate how the factors load on the macroeconomic variables, and can therefore investigate how the structural shocks affect each of the large set of variables under analysis. The combination of the factor model for the variables and the VAR for the factors is known as a factor augmented VAR (hereafter FAVAR).

Typically a FAVAR is estimated on a dataset with variables sampled at the same frequency. For example, Bernanke et al. [2005] use monthly variables only, but a monthly FAVAR can leave out potentially important indicators that are observed at other than monthly frequency. For example, it leaves out real GDP, which is accepted as the most accurate measure of economic activity but is only available at quarterly frequency. One could aggregate monthly variables to a quarterly level and estimate the model on a quarterly frequency. But then the quarterly model is subject to aggregation bias, meaning that important information gets lost in the aggregation process. This suggests to estimate FAVARs combining data at different frequencies, and various techniques are now available, see e.g., Giannone, Reichlin, and Small [2008] Jungbacker, Koopman, and van der Wel [2011], Bańbura and Modugno [2014], Bańbura and Rünstler [2011], Mariano and Murasawa [2010] and Frale, Marcellino, Mazzi, and Proietti [2010], Frale, Marcellino, Mazzi, and Proietti [2011], Marcellino and Schumacher [2010]. All these studies on mixed frequency (MF) factor models are not of structural nature but focus on reduced form analyses, such as nowcasting and forecasting quarterly GDP growth using monthly or higher frequency indicators.

In this paper we introduce an alternative method to estimate a large MF factor model. We start from the Doz, Giannone, and Reichlin [2011] procedure for estimating plain factor models and extend it to allow for the presence of mixed frequency data, where mixed frequencies are handled along the lines of Mariano and Murasawa [2003] and Mariano and Murasawa [2010]. We then assess the finite sample performance of our procedure in a set of Monte Carlo experiments, comparing it with that of alternative estimators for MF factor or FAVAR models. It turns out the procedure performs quite well even in small samples not only in terms of factor estimation but also to recover the impulse response functions to structural shocks. Finally, it can be easily modified to allow for observable factors, in high or low frequency.

Our second contribution, as anticipated, is to show how to conduct structural economic analyses using MF FAVAR models. We do not provide a new technique to identify structural shocks, but rather show how to apply existing techniques in a mixed frequency FAVAR. So far there are few examples of structural analyses based on mixed frequency data, see e.g. Giannone, Monti, and Reichlin [2010], Chiu, Eraker, Foerster, Kim, and Seoane [2011], Ghysels [2015], Foroni and Marcellino [2014] and Foroni and Marcellino [2015]. However, all these papers are based on VAR or DSGE models. We present three empirical examples.

First, we add quarterly variables to the monthly dataset of McCracken and Ng [2014], in particular GDP, and investigate how monetary policy shocks identified at monthly level affect GDP and other key macroeconomic variables. Second, again using a mixed frequency dataset, we study how monthly oil price shocks propagate to quarterly GDP. Finally, in our third application, we impose quarterly government expenditure as an observable factor, governed by the sum of three latent monthly expenditure growth rates. Using this specification, we can evaluate how monthly government expenditure shocks affect the economy. In all cases we find reasonable results in economic terms from the MF FAVAR, sometimes with interesting differences with respect to the standard FAVARs and VARs.

The remainder of the paper is structured as follows. In Section 2 we introduce the MF FAVAR. We first present dynamic factor models, next explain how we suggest to estimate them in the presence of mixed frequency data, then compare our proposal with other methods suggested in the literature, and finally we discuss estimation in the presence of some observable factors. In Section 3 we use Monte Carlo experiments to analyze the performance of our estimation method when varying the cross-sectional (n), temporal (T) dimensions, the amount of missing observations (generated by the presence of the mixed frequency data) and the frequency of the factors. In Section 4 we present the three empirical applications, studying the effects of, respectively, monetary, oil, and fiscal shocks. Finally, in Section 5 we summarize our main results and conclude.

2 Mixed Frequency FAVAR

2.1 The single frequency FAVAR model

We assume that an n dimensional zero mean stationary vector of variables y_t can be represented as a sum of two components, *a common component* (Λf_t) and *an idiosyncratic component* (e_t):

$$y_t = \Lambda f_t + e_t. \tag{1}$$

f_t is a $k \times 1$ dimensional vector of factors that are common to all the variables in y_t, with the number of factors being (much) smaller than the number of variables ($k < n$). Factors capture the majority of comovements in the evolution of the individual variables. Λ is an $n \times k$ matrix of factor loadings. The loadings determine how the factors affect the dependent variables. Λf_t is called *common component* of the factor model because it represents that part of the variability of y_t that originates from the k factors that are common to all the n variables. On the other hand, e_t is an $n \times 1$ zero mean vector that represents the *idiosyncratic component* of the factor model. This source of variability in y_t can not be captured by the k common factors and is variable specific.

Equation (1) represents a classic factor model. If the $n \times n$ covariance matrix of the idiosyncratic components ($E(e_t e_t') = \Psi$) is a diagonal matrix, then the model in (1) becomes an *exact factor model*. A diagonal covariance matrix can be too restrictive for macroeconomic applications, so we let e_t have some limited cross correlation. Such model is called an *approximate factor model*.

Specifically, following Doz, Giannone, and Reichlin [2012], we impose two conditions:

A1) $0 < \underline{\lambda} < lim\, inf_{n \to \inf} \frac{1}{n} \lambda_{min}(\Lambda' \Lambda) \leq limsup_{n \to \inf} \lambda_{max} \frac{1}{n}(\Lambda' \Lambda) < \overline{\lambda} < \inf$.

A2) $0 < \underline{\psi} < lim\, inf_{n \to \inf} \lambda_{min}(\Psi) \leq lim\, sup_{n \to \inf} \lambda_{max}(\Psi) < \overline{\psi} < \inf$.

λ_{min} and λ_{max} indicate the smallest and the largest eigenvalues of a matrix. Condition A1 ensures that the factors are pervasive, that is, that they affect most dependent variables. Condition A2 ensures that the variance of the idiosyncratic components is greater than zero, but limits the extent of the cross-correlation. Again as in Doz et al. [2012], e_t can be also serially correlated, see their Assumption (A3).

The common factors and the idiosyncratic components are assumed to be uncorrelated at all leads and lags, $E(f_{jt}e_{is}) = 0$ for all $j = 1, ..., k$, $i = 1, ..., n$ and $t, s \in \mathbb{Z}$.

In equation (1), Λ and f_t are unobserved and need to be estimated. This poses an identification problem because there are different combinations of Λ and f_t that deliver the same common component. To identify the factors we assume that the first $k \times k$ entries of the loadings matrix form an identity matrix: $\Lambda = \begin{bmatrix} I_k \\ \Lambda^* \end{bmatrix}$, where Λ^* is an $(n-k) \times k$ matrix of unrestricted loadings. This method of identification is also used in Bernanke et al. [2005].

Finally, we assume that the factors follow a *p-th* order VAR:

$$f_t = A_1 f_{t-1} + ... + A_p f_{t-p} + u_t, \tag{2}$$

where p is finite and u_t is a k dimensional Gaussian white noise process with covariance matrix Σ.

The model presented in equations (1) and (2) represents a *static form* of a dynamic factor model. It is called the static form because factors enter equaion (1) without lags. As shown in Stock and Watson [2005] the static form of a dynamic factor model nests the *dynamic representation*. Suppose that static factors f_t are composed of dynamic factors q_t $(r \times 1)$ and their lags: $f_t = [q_t, q_{t-1}, ..., q_{t-g}]'$, such that $k = r \times (g+1)$. We can then rewrite the model into the *dynamic form*:

$$y_t = \tilde{\lambda}(L) q_t + e_t, \tag{3}$$

$$q_t = \tilde{a}(L) q_{t-1} + v_t, \tag{4}$$

where $\tilde{\lambda}(L)$ and $\tilde{a}(L)$ represent lag polynomials in the dynamic representation and v_t the fundamental shocks that govern the dynamic factors. The number of fundamental shocks (v_t) can be smaller than the number of static shocks (u_t): $u_t = G \times v_t$, where G is of dimension $k \times r$ and $r \leq k$. In the static representation of the factor model we need to choose k and p high enough to capture all the effects that the dynamic factors and their lags $(\lambda(L)q_t)$ exert on the dependent variables. We can then use the static factor model to uncover the effects that the fundamental shocks have on the economy.

Equations (1) and (2) represent the (single frequency) FAVAR model.

2.2 Estimation of the single frequency dynamic factor model

Doz et al. [2012] propose to estimate the model in (1) and (2) using a Quasi Maximum Likelihood (QML) approach where the maximum likelihood estimates of the model are obtained using the *expectation-maximization algorithm* (hereafter EM). The EM algorithm iterates between two steps. In the first (*maximization*) step, it calculates the maximum likelihood estimates of the factor model parameters ($\hat{\theta} = \left\{\hat{\lambda}, \hat{A}, \hat{\Psi}, \hat{\Sigma}\right\}$) conditional on the estimates of the factors. In the second (*expectation*) step, conditional on the parameter estimates, it uses the Kalman filter-smoother to get the factor estimates ($\hat{f}_t$) and the likelihood function of the model. The estimated factors are then used to produce another set of parameter estimates, then another set of estimated factors, and so on until convergence.

When calculating the maximum likelihood estimators, we assume that the idiosyncratic shocks e_t are not auto-correlated and cross correlated, although they often are. Hence, more properly, we obtain *QML* estimates, following White [1982]. Doz et al. [2012] show that this QML approach is valid for estimating the model parameters and the factors, even when the approximating model is mis-specified and the shocks exhibit weak cross correlation and auto-correlation. They show that the QML estimators are consistent for the true factor space, with a consistency rate equal to $min\left\{\sqrt{T}, \frac{n}{log(n)}\right\}$.

The initialization of the EM algorithm requires (consistent) estimates of the factors. For this, we can use principal components (PCA), since consistency of PCA estimates of the factors results from Stock and Watson [2002a], Bai and Ng [2002], Bai [2003].

2.3 Estimation in the presence of mixed frequency data

We now extend the Doz et al. [2012] estimation procedure summarized in the previous subsection in order to handle mixed frequency data. Next, we compare our proposal with two alternative methods.

We closely follow the notation used in Mariano and Murasawa [2010], to whom we refer for additional details. We assume that we have two types of variables, low frequency and high frequency (e.g. quarterly and monthly data). Let $y_{t,1}$ represent an n_1 variate low frequency vector of variables, that are observed only every third period (e.g. only in periods $t = 3, 6, 9, ...$). Let $y_{t,2}$ represent an n_2 vector of high frequency variables that are observed in every period. As before, the total number of variables is n (where $n = n_1 + n_2$) and $y_t = [y'_{1,t}, y'_{2,t}]'$ is an $n \times 1$ dimensional vector. We assume that underlying $y_{1,t}$ there is a process $y^*_{1,t}$. Most of the time $y^*_{1,t}$ is unobservable, except for every third period when it has the same value as the observable $y_{1,t}$. This implies that the quarterly variable is point in time sampled. We retain this assumption for exposition and later relax it. We adopt the same model as in the previous section, but this time we assume that the common factors load on the (sometimes unobservable) process y_t^* instead of on (observable) y_t directly:

$$f_t = A_1 f_{t-1} + ... + A_p f_{t-p} + u_t, \tag{5}$$

$$y_t^* = \Lambda f_t + e_t, \tag{6}$$

where $y_t^* = [y_{1,t}^{'*}, y_{2,t}^{'}]'$. Since y_t^* is unobservable, we need to link it to the observables y_t. This is done with the following equation:

$$y_t^+ = C_t y_t^* + D_t v_t \tag{7}$$

where:

$$y_t^+ = \begin{bmatrix} y_{1,t}^+ \\ y_{2,t} \end{bmatrix} \text{ and } y_{1,t}^+ = \begin{cases} y_{1,t} & \textit{when } y_{1,t} \textit{ is observed} \\ v_{1,t} & \textit{when } y_{1,t} \textit{ is not observed} \end{cases}$$

$$v_t = \begin{bmatrix} v_{1,t}^+ \\ 0 \end{bmatrix} \text{ and } v_{1,t}^+ = \begin{cases} 0 & \textit{when } y_{1,t} \textit{ is observed} \\ v_{1,t} & \textit{when } y_{1,t} \textit{ is not observed} \end{cases}$$

$$C_t = \begin{bmatrix} C_{1,t} : 0_{n_2} \\ 0_{n_1} : I_{n_2} \end{bmatrix} \text{ and } C_{1,t} = \begin{cases} I_{n_1} & \textit{when } y_{1,t} \textit{ is observed} \\ 0_{n_1} & \textit{when } y_{1,t} \textit{ is not observed} \end{cases}$$

$$D_t = \begin{bmatrix} D_{1,t} \\ 0_{n_2} \end{bmatrix} \text{ and } D_{1,t} = \begin{cases} 0_{n_1} & \textit{when } y_{1,t} \textit{ is observed} \\ I_{n_1} & \textit{when } y_{1,t} \textit{ is not observed} \end{cases}$$

I_{n_1}indicates an identity matrix of size $n_1 \times n_1$ and O_{n_1} a matrix of zeros of size $n_1 \times n_1$. We assume that $v_{1,t}$ is a normally distributed random vector of size n_1, $v_{1,t} \sim N(0, I_{n_1})$. But, although v_t is a random vector, it is assumed that all the realizations of the vector v_t are simply zero. Hence, the measurement equation (7) is rewritten as if it consists of the observable variables only, and when the data is missing the missing data are replaced by a $N(0, I_{n_1})$ random vector $v_{1,t}$ whose realizations are zero.[1] Mariano and Murasawa [2010] propose this approach since it implies that equations (5) and (7) form a state space model where, from the point of view of the Kalman filter-smoother, all the variables are observed. Because the loadings of the missing data points are set to zero for the missing variables, the Kalman gain has zeros in the columns that correspond to the missing variables, so that when forecasting a new value of the state vector, the errors corresponding to the missing observation do not contribute to the new value of the state

[1] The value of realizations makes no difference for the method to work.

and to the new value of the state variance. The Kalman filter-smoother simply skips the influence of the missing observations when it estimates the factors and the likelihood.

It is convenient to take one last step and transform the state equation. First we use the companion form for the factor VAR model in (5). That is, we shift the factor lags into the state vector s_t, so that the resulting model becomes a *VAR(1)* model in s_t. Second, we insert equation (6) into equation (7). The resulting model has the familiar state space form:

$$s_t = As_{t-1} + Bu_t \tag{8}$$

$$y_t^+ = H_t s_t + \tilde{v}_t \tag{9}$$

where

$$s_t = \begin{bmatrix} f_{t\,(k\times1)} \\ f_{t-1\,(k\times1)} \\ \vdots \\ f_{t-p+1\,(k\times1)} \end{bmatrix}, \; A = \begin{bmatrix} A_{1\,(k\times k)} & \cdots & A_{p\,(k\times k)} \\ I_{(k\times k)} & \cdots & 0_{(k\times k)} \\ \vdots & \ddots & \vdots \\ 0_{(k\times k)} & \cdots\; I_k & 0_{(k\times k)} \end{bmatrix}, \; B = \begin{bmatrix} \Sigma_{uu\,(k\times k)}^{\frac{1}{2}} \\ 0_{\,([k(p-1)]\times k)} \end{bmatrix}.$$

The model in equations (8)-(9) is a general state space model. We assumed that quarterly variables are point in time sampled, but depending on how we define the matrix H_t we can accommodate other sampling schemes. In particular, we consider: a) point in time sampling (i.e.: $y_t^+ = y_t^*$), b) the sum of three consecutive growth rates (i.e.: $y_t^+ = y_t^* + y_{t-1}^* + y_{t-2}^*$), and c) geometric growth (i.e.: $y_t^+ = \frac{1}{3}y_t^* + \frac{2}{3}y_{t-1}^* + y_{t-2}^* + \frac{2}{3}y_{t-3}^* + \frac{1}{3}y_{t-4}^*)$[2]. For these three aggregation schemes H_t is defined as follows.

a) Point in time sampling

$$H_t = [\underbrace{C_t\Lambda}_{n\times k} \quad \underbrace{0_{(n\times k)} \cdots 0_{(n\times k)}}_{n\times[(k\times(p-1)]}\;] \tag{10}$$

b) Sum of 3 latent growths rates

$$H_t = [\underbrace{C_t^1\Lambda}_{n\times k} \quad \underbrace{C_t^2\Lambda}_{n\times k} \quad \underbrace{C_t^2\Lambda}_{n\times k} \quad \underbrace{0_{:(n\times k)} \cdots 0_{:(n\times k)}}_{n\times[(p-3)\times k]}\;], \tag{11}$$

[2]See Mariano and Murasawa [2010] for further details

where

$$C_t^1 = \begin{bmatrix} C_{1,t} : 0_{n_2} \\ 0_{n_1} : I_{n_2} \end{bmatrix} \quad and,\ C_t^2 = \begin{bmatrix} C_{1,t} : 0_{n_2} \\ 0_{n_1} : 0_{n_2} \end{bmatrix} \tag{12}$$

c) Geometric growth

$$H_t = [\underbrace{\frac{1}{3} \times C_t^1 \Lambda}_{n \times k} \quad \underbrace{\frac{2}{3} \times C_t^2 \Lambda}_{n \times k} \quad \underbrace{\times C_t^2 \Lambda}_{n \times k} \quad \underbrace{\frac{2}{3} \times C_t^2 \Lambda}_{n \times k} \quad \underbrace{\frac{1}{3} \times C_t^2 \Lambda}_{n \times k} \quad \underbrace{0_{(n \times k)} \ldots 0_{(n \times k)}}_{n \times [(p-5) \times k]}\ ] \tag{13}$$

Notice also that if the lag length of the VAR model for the factors (p) is smaller or equal than a) one for point in time sampling, b) three for sum of three latent growth rates or c) five for geometric sampling, then the zero matrices in H_t simply drop out.

The model in equations (8)-(9) can also be extended by adding a moving average component for the idiosyncratic shocks, to explicitly account for auto-correlation. This can be done by adding lags of the idiosyncratic shocks to the state vector and adjusting the matrices accordingly.

The state space representation in (8)-(9) differs slightly from that used by Mariano and Murasawa [2010] while it is in line with Giannone et al. [2008]. Specifically, Mariano and Murasawa [2010] plug the idiosyncratic error term into the state vector, increasing substantially its dimension, which is not desirable in applications with large datasets since estimation becomes very slow, or in practice infeasible. To assess whether the form of the state space matters we have run Monte Carlo experiments (available upon request) that indicate that the differences are generally very small.

Mariano and Murasawa [2010] estimate the model using a quasi-Newton method while we adopt a simpler approach based on Doz et al. [2012] QML estimator. The starting estimate for the factors can be obtained by the EM algorithm of Stock and Watson [2002b], which is a PCA approach applied to an unbalanced dataset. The Stock and Watson EM approach does not assume any time dynamics in the factors. Both Doz et al. [2011] two step approach and our EM algorithm (based, as explained, on Doz et al. [2012]) assume that factors evolve in a VAR model. When factors exhibit time dynamics that can be well approximated by a VAR, which is common in practice, factor estimates based on our EM algorithm are more efficient than those resulting from the Stock and Watson [2002b] EM algorithm. We then use the EM algorithm as described in the previous section to get the maximum likelihood estimates of the parameters and the factors. Our procedure is more efficient than just using Stock and Watson [2002b], and can more easily handle a variety of aggregation schemes.

Finally, Doz et al. [2012] prove the consistency of the QML estimator. Since our specification is nested in their model, the estimation procedure remains consistent. To asses its finite sample performance when using mixed frequency data and observable factors, we will use a set of Monte Carlo experiments, see Section 3.

2.4 Comparison with other estimation methods for MF factor models

Harvey and Pierse [1984] first handled missing data in the Kalman filter context, by modifying the updating and backdating equations of the filter. This approach can become cumbersome when handling systematically missing data, as in the mixed frequency case.

Besides the PCA based EM algorithm of Stock and Watson [2002b] mentioned above, the two computationally feasible and closest approaches to ours are those by Giannone et al. [2008], Doz et al. [2011] and Bańbura and Modugno [2014].

Doz et al. [2011] use a two step estimator that is based on Kalman filtering. Their approach can be though of as the first step in our EM algorithm. They handle missing data as in Giannone et al. [2008]. Giannone et al. [2008] exploit the fact that the value for a missing data point is irrelevant if its variance is infinite. The Kalman filter puts zero weight on such points and the missing value does not affect the estimates. Bańbura and Modugno [2014] use a selection matrix that modifies the Kalman filter smoother formulae so that only the available data are used in the estimation.

In practice, Giannone et al. [2008], Bańbura and Modugno [2014] and our procedure induce the Kalman filter to skip the missing observations. Hence, not surprisingly, they produce numerically equal results. We believe that our procedure is easier to understand and more closely related to the approach to handle mixed frequencies in other types of models, such as VARs. Moreover, as we will see in the next section, it can be easily modified to allow for some observable factors, which is relevant for economic applications.[3]

Finally, in terms of practical implementation, Frale et al. [2010] and Frale et al. [2011] suggest to use the univariate treatment of the multivariate filter, as discussed by Durbin and Koopman [2012]. However, in our context where the model is larger, the square root filter (see again Durbin and Koopman [2012]) turns out to be more stable

[3]Jungbacker et al. [2011] introduce a more complex procedure based on two different state space representations. A normal representation for when all data are available and a modified representation for when there are missing data. In the modified representation they add the missing data into the state vector, so that missing values are estimated together with the factors, we refer to them for additional details. Jungbacker et al. [2011] report that there are substantial computational gains with their the method. We achieve similar gains because instead of putting the error term into the state vector, as in Mariano and Murasawa [2010], we form a compound error term, leaving the size of the state vector unaltered.

and computationally faster.[4]

2.5 Estimation in the presence of observed factors

Bernanke et al. [2005] assume that, in a FAVAR model for a large set of same frequency variables, one of the factors is observable. It coincides with a short term interest rate as economic theory suggests that monetary policy should affect most variables in the economy, at least in the short term, and therefore it is pervasive. This helps both structural shock identification and the interpretation of the impulse response functions. Hence, we now consider how observable factors can be treated in a mixed frequency context.

To ease the exposition, let us assume that we have only two factors, one latent and one observable and that the observable factor is point in time sampled. The model is the same as in equations (5)-(7). Assume also for convenience that quarterly variables are point in time sampled. The model then becomes:

$$f_t = A_1 f_{t-1} + ... + A_p f_{t-p} + u_t \tag{14}$$

$$y_t^* = \Lambda f_t + e_t \tag{15}$$

$$y_t^+ = C_t y_t^* + D_t v_t \tag{16}$$

where f_t is now $[f_{1,t},\, i_t]'$. $f_{1,t}$ is a latent unobservable factor as before and i_t is an observable factor, in our example the interest rate. Let $y^+_{N-1,t}$ represent all the variables in y_t^+, except the interest rate. Further, assume that the interest rate is ordered last, in the n^{th} place, in the mixed frequency vector of the dependent variables. Then the vector of dependent variables is $y_t^+ = [y^{+'}_{N-1,t},\, i_t]'$.

For simplicity, let us now focus on the last row of equation (15), the interest rate equation, since nothing changes for other parts of the FAVAR model. The last equation is:

$$i_t = \Lambda_N f_t + e_{N,t}, \tag{17}$$

and, since i_t coincides with the observable factor, it must be $\Lambda_N = [0,\, 1]$, $e_{N,t} = 0$ for all time periods. The corresponding variance and covariances of the error term of the interest rate equation are also zero ($\Psi_{N,i} = \Psi_{i,N} = 0$, where $i = 1, ..., N$).

The model can be then estimated using the EM procedure introduced in Section 2.3. Note that a similar procedure can be used when the observable factor is a low frequency

[4] The results of Monte Carlo experiments comparing the two approaches are available upon request.

variable.

3 A Monte Carlo Evaluation of the MF Estimation Procedure

We are interested in assessing how well our proposed estimation method uncovers the factors and impulse responses (hereafter IR) of the dependent variables to shocks in the observable factors, under different assumptions on the sample size T, the size of the cross section n, the number of low frequency variables and the frequency of the factors. We focus on shocks to the observable factor since this case was not considered in the previous literature and it is related to the empirical applications that we will present in the next section, but similar results apply for shocks to unobservable factors (results available upon request).

We use a modified version of the data generating process (hereafter DGP) commonly used in a same frequency setting, see among others Stock and Watson [2002a],Doz et al. [2012] and Doz et al. [2011]. Following Doz et al. [2012], we write the DGP as:

$$\begin{aligned} y_t &= \Lambda f_t + e_t, \\ f_t &= A_1 f_{t-1} + ... + A_p f_{t-p} + u_t, \\ e_t &= d_1 e_{t-1} + \cdots + d_q e_{t-q} + v_t, \end{aligned}$$

Let Λ_{ij} represent the ij^{th} element of Λ, where $i = 1, ..., n$ and $j = 1, ..., k$. We assume $\Lambda_{ij} \sim i.i.d.N(0, 1)$. Let $a_{ij}(l)$ represent the ij^{th} element of A_l, where $l = 1, ..., p$. We assume:

$$a_{ij}(l) = \begin{cases} 1-\rho & if\ i = j \\ 0 & if\ i \neq j \end{cases}, \ i, j = 1, ..., k,$$

and $u_t \sim i.i.d.N(0, (1-\rho^2)I_k)$. Let $d_{ij}(l)$ represent ij^{th} element of d_l, where $l = 1, ..., q$. We assume:

$$d_{ij}(l) = \begin{cases} 1-d & if : i = j \\ 0 & if : i \neq j \end{cases}, \ i, j = 1, ..., n,$$

and $v_t \sim i.i.d.N(0, \mathcal{T})$. And, finally, we assume that the elements of $\mathcal{T}$ satisfy:

$$\begin{aligned} \tau_{ij} &= \sqrt{\alpha_i \alpha_j} \tau^{|i-j|}(1-d^2), \ i, j = 1, ..., n, \\ \alpha_i &= \frac{\beta_i}{1-\beta_i} \Sigma_{j=1}^{k} \Lambda_{ij}^2, \ \beta_i \sim i.i.d.U([u, 1-u]). \end{aligned}$$

In this DGP, the idiosyncratic shocks are allowed to be auto-correlated and also weakly cross correlated, with cross correlation governed by the parameter τ. $\mathcal{T}$ is a Toeplitz matrix. When τ is zero the model becomes an exact factor model and $\mathcal{T}$ is a diagonal matrix. The parameter β_i controls for the ratio between the common component ($\Lambda_i f_t$) variances and the idiosyncratic component (e_{it}) variances (where Λ_i is the i^{th} row of Λ and e_{it} the i^{th} element of e_t). Following Doz et al. [2012], we set it to 50%. u is a parameter that controls the cross sectional heteroscedasticity. We set it to 0.5, which implies cross correlation with the closest two adjacent time series equal to 0.5 (on average) and decays below 0.1 (on average) after the forth closest series. Therefore, cross correlations between the idiosyncratic shocks are clustered.

We estimate the model using the approximating specification that assumes no cross correlation. Hence, the true data generating process is an approximate factor model and we model it using an exact factor model.

We deviate from previous Monte Carlo analyses in two ways. First, we assume that some of the variables are not observed all the time. This is done simply by first simulating the model data and then deleting some of the observations in the data set. In particular, assuming that t is measured in months, some variables are only observed at the end of the quarter (so that all observations corresponding to the first two months of each quarter are deleted). These are the low frequency variables in our simulation study. Second, to align the Monte Carlo study with the MF structural FAVAR used in practice, we assume that two factors generate the data and that one factor is observable. The resulting simulated mixed frequency data set and the observable factor are then used to estimate the space spanned by factors and to produce IRs of the dependent variables to a shock in the observable factor.

We compare five different estimators for the factors. First, the PCA estimator on a data set without imposing the observable factor and without missing observations. This in practice is not feasible, but we use it as a benchmark to asses the effects of the missing observations. Second, the PCA estimator computed after dropping the series with missing observations from the data set. Third, the Stock and Watson [2002b] EM algorithm based approach to estimating factors from unbalanced datasets. Fourth, the Doz et al. [2012] two step estimator where mixed frequencies are handled as in Mariano and Murasawa [2010].[5] In the first step, factors are initialized by PCA and model parameters estimated, in the second step the factors are re-estimated using the Kalman filter. Finally, our QML estimator introduced in the previous Section, which basically keeps iterating the two-step

[5]Doz et al. [2012] estimates latent factors. We modify their method so that it can handle observable factors as proposed in subsection 2.5.

estimator discussed above.

3.1 Recovering the space spanned by the factors

In this section we investigate how well the alternative estimators uncover the space spanned by the factors. We base the evaluation on the trace statistic, a multivariate version of the R^2 measure, also used in Stock and Watson [2002a], Giannone et al. [2008] and Bańbura and Modugno [2014]. It measures how close the estimated factors are to the true factors which generated the data, and is defined as:

$$\frac{Trace(F'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'F)}{Trace(F'F)}, \tag{18}$$

where $F = [F^1, F^2]$ are the true factors and $\hat{F} = [\hat{F}^1, \hat{F}^2]$ their estimated version. The trace statistic lies between one and zero, being equal to one when the factor space is perfectly estimated.

We assume that one of the true factors is observable (i.e. we impose $\hat{F}^2 = F^2$), therefore we only estimate the latent factor and the model parameters[6]. In Table 1 we report the average trace statistic computed over 1000 replications for different values of n and T ($n = 50, 100, 200$, $T = 50, 100, 200$), and a fixed number of low frequency series ($d = 20$). Four main findings emerge. First, for all methods the values increase with n and/or T. Second, the values are already rather large for $n = T = 50$, suggesting that the procedures work well also in finite samples, notwithstanding the presence of missing observations. Third, PCA on full sample, even though based on a larger information set than the other methods, performs generally worse because the observable factor is not imposed. Finally, the DGR and our MF estimators perform comparably and slightly better than principal components, with a slight advantage in all cases for our MF estimator.

Table 2 presents results for $n = 200, T = 200$ and a varying number of series with missing observations: $20, 100, 180$. While the results naturally deteriorate when the number of missing observations increases, the average trace statistics remain quite good also when 180 series are only observable on a quarterly basis, with values in the range $0.98 - 0.99$[7]. The rationale is that the factor structure is quite strong so that few

[6] For the first estimator, the unattainable PCA estimator on the full sample, we do not impose that the second factor is observable. We do this to show the merit of introducing an observable factor, even in a case when one compares it to an estimator where all the factors are latent but one observes all the data.

[7] The trace statistic for our method performs even better than the competitors, if we compute the trace statistic for the latent factor only.

variables already contain substantial information on the factors, as indicated by the still good performance of PCA applied on the reduced sample of monthly only observations. However, adding the quarterly variables improves the trace statistic, the most so when using our MF estimator.

Table 3 presents results for $n = 200, T = 200$ when the observable factor is of quarterly frequency and varying number of series with missing observations: $20, 100, 180$. Naturally the results are slightly worse, compared to the case when the observable factor is of monthly frequency, but the estimators still preform quite well. Our estimator does slightly better than the other estimators, the more so when the number of quarterly series is high.

3.2 Recovering the impulse responses

In the preceding section we have seen that our MF method recovers quite well the space spanned by the factors, slightly better than the competing methods. In this section we investigate how well it uncovers the impulse responses to a shock in the observable factor. We run two experiments. In the first experiment we compare the Stock and Watson [2002b] EM algorithm to handle factor estimation in unbalanced datasets with our procedure. In the second experiment we investigate if the mixed frequency data reduces the aggregation bias that is present when one instead aggregates all the variables to a quarterly frequency.

In the first experiment we fix n and T to 200 and the number of quarterly variables to 100. We draw the factor loadings Λ at the first iteration and then retain the same Λ for the remaining replications for comparability.[8] We report the average estimates over 1000 replications. We explore the results along two dimensions. First, we compare the IRs of the low frequency variables (with missing data) and of the high frequency variables. Second, we investigate how the number of low frequency variables affects the IRs.

The general result from the first experiment is that the impulse responses of the dependent variables to a shock in the observable monthly factor track the true impulse responses very closely in both cases, when using Stock and Watson [2002b] or our EM algorithm. In fact, impulse responses of both monthly and quarterly variables, are often indistinguishable from the true impulse response and the true impulse responses never cross 95% confidence bands of the estimated impulse responses[9]. We conclude that when the observable factor is monthly both methods deliver similar results. We do however find differences when the observable factor is quarterly. This case is presented next.

[8] We repeated the experiment several times to make sure that a specific draw of Λ did not affect the results.

[9] Figures from the comparison are not presented, but are available upon request.

Figure 1 plots the IRs of the first 9 quarterly variables from an experiment where $n = 200$, $T = 200$, there are $d = 100$ missing variables and the observable factor is of quarterly frequency. In each figure the black lines with dots represent the true IRs, the black solid lines the IRs estimated with the monthly dataset using our ML procedure (dashed black lines are the $+/-$ 2 std. dev.) and the red lines the IRs estimated using the Stock and Watson [2002b] approach.

There are discrepancies between the true and estimated responses, but our MF FAVAR estimation method generally outperforms the use of the Stock and Watson [2002b] factors. We observe that the IRs obtained with the Stock and Watson [2002b] algorithm sometimes fall outside the 95% confidence bands. This is likely due to the fact that our procedure explicitly takes into account the model generating the quarterly factor whereas the Stock and Watson [2002b] approach does not.

In the second experiment we investigate if the use of mixed frequency data reduces the aggregation bias. In this experiment we fix n and T to 200 and assume that the observable factor is of monthly frequency. We then estimate three models. In the first model we use quarterly data, in the second we use mixed frequency data and in the last (empirically unfeasible) model we use monthly data. We set the number of quarterly series in the mixed frequency dataset to 100. We compare the IRs estimated on the monthly, mixed frequency, and quarterly datasets.[10] In addition, to make sure that the differences between the estimated IRs only result from the different types of datasets, we initialize the three models with the parameter values of the true DGP.

Figure 2 plots the IRs of variables that are quarterly in the mixed frequency dataset (first 9 variables[11]). In each figure the black lines with dots represent the true IRs, the black solid lines the IRs estimated with the monthly dataset (dashed black lines are the $+/-$ 2 std. dev.), solid gray lines are the IRs estimated on the mixed frequency data, and the red lines the IRs estimated with the quarterly dataset.

The figure shows that the IRs estimated on a monthly dataset are in general very close to those estimated on a mixed frequency dataset. In fact, the IRs estimated on the monthly dataset are often not visible because they overlap with the IRs estimated on the mixed frequency dataset almost perfectly (the solid gray line overlaps the solid black line). Both estimated IRs track the true IRs (solid black line with dots) very closely. This is natural because the sample size is quite large, the shocks belong to an always observable factor and we used the true DGP to initialize the EM algorithms. The results

[10] To facilitate comparison of monthly IRs with the quarterly IRs we "skip sample" the monthly IRs to a quarterly frequency, namely, we record the monthly IRs only at times $t = 1, 4, 7, \ldots$.

[11] The variables are representative for the other variables. Note also that the loadings matrix is sampled randomly, therefore one can consider the selected IRs as being chosen randomly.

for monthly variables are similar (not shown).

Figure 3 plots the IRs of the first 9 monthly variables in the mixed frequency dataset (black solid lines) and the IRs obtained with quarterly dataset (red dashed lines). We add 95% confidence bands for both IRs and comapre them (dashed lines). We observe that the confidence bands obtained using the mixed frequency data are narrower than the confidence bands obtained with quarterly data. In addition, while the IRs estimated on the mixed frequency data track the true IRs very closely, the IRs estimated on a quarterly dataset sometimes depart from the true IRs, in particular in the short run. Two sources drive this result. The shock variances of the model estimated on the quarterly dataset are consistently overestimated, and the factor VAR parameters are consistently underestimated, with the first type of bias dominating the latter. Hence, the aggregation bias can be substantial, and the use of mixed frequency data can reduce it. Foroni and Marcellino [2014] and Foroni and Marcellino [2015] obtain similar results for, respectively, DSGE and structural VAR models.

In summary, this section shows that the MF-S-FAVAR, estimated using our method, performs quite well in recovering the space spanned by the factors and the true IRs, even in small samples. It also performs well when one of the observed factors is at quarterly frequency and it reduces the aggregation bias. To further motivate the usefulness of our method and illustrate its practical implementation, we next present three empirical applications.

4 Empirical applications

4.1 Bernanke et al. [2005] Monetary Policy Shocks

In this first application we assess the effects of monetary policy measured at the monthly level on quarterly GDP growth. We start with the original monthly FAVAR model put forward by Bernanke et al. [2005]. We utilize the large data set of McCracken and Ng [2014], X_t, which consists of 135 monthly series from January 1959 to January 2015[12]. The variables summarize all the major developments in the economy and include measures of real output, income and price indicators, interest rates, employment indices, consumption variables, housing prices, etc..

To facilitate comparability[13], we first estimate the same monthly model as in Bernanke

[12]The data set is not yet publicly available. The data set was kindly provided to us by the authors.

[13]We also estimated the model using the original Bernanke et al. [2005] data set, consisting of 120 monthly variables from February 1959 to August 2001. The results we're similar to the results presented in this paper.

et al. [2005], using 3 latent factors, one observable factor (the federal funds rate) and 7 lags for the factor VAR. Bernanke et al. [2005] also use 3 latent factors and indicate that a larger number does not change the results. Next, we add quarterly GDP growth to create a mixed frequency factor model, where quarterly GDP is modeled as a sum of three consecutive unobserved monthly growth rates.[14] [15].

Before discussing the results, it is important to consider as issue not addressed by Bernanke et al. [2005], namely, the number of dynamic factors driving X_t, given the assumed number of static factors (four in our case). We use the Stock and Watson [2005] approach to determine their number.[16] They suggest regressing each variable on own lags and the lags of the static factors, recover the residuals ($\tilde{\epsilon}_{it}$, for $i = 1...n$), and test how many factors drive them. The number of factors driving the estimated residuals is equal to the number of dynamic factors driving the variables X_t. Table 4 displays the values of the Bai and Ng [2002] information criteria associated with different assumptions regarding the numbers of static factors driving $\tilde{\epsilon}_{it}$'s ($\hat{q}$, with $\hat{q} \leq 4$ where 4 is the number of static factors we use for X_t). All criteria favor 4 static factors for $\tilde{\epsilon}_{Xit}$'s and hence 4 dynamic factors for X_t. This implies that static factors for X_t are equivalent to dynamic factors, and we can proceed with our structural analysis by identifying structural shocks directly on the static factor VAR residuals.

Using the monthly FAVAR model, Figure 4 reports the impulse response functions of selected variables to a monetary policy shock identified as in Bernanke et al. [2005] (together with the 90% confidence bands). In the same figure, the dashed black lines represent the IRF obtained using our estimation method for the MF-FAVAR model. Overall, the IRFs are quite similar, and those obtained with our method are most of the time statistically indistinguishable from the IRFs estimated using the Bernanke et al. [2005] approach. This result is not so surprising as in this case there is only one quarterly variable added to the monthly dataset. However, there are few but important signifi-

[14]Namely, $y_t = y_t^* + y_{t-1}^* + y_{t-2}^*$, where y_t is the quarterly GDP growth observed only every 3rd period and y_t^*represents the latent monthly GDP growth. The results are almost identical when we assume that quarterly GDP growth is point in time sampled from monthly GDP growth. Small differences arise when we assume that quarterly GDP is modeled as a geometric mean of unobserved monthly GDP, as in Mariano and Murasawa [2010]. In this version of the paper we only present the results from when quarterly GDP is modeled as a sum of 3 latent growths and refer to other results in the text.

[15]It is not likely that adding a small number of quarterly series would affect the number of static factors needed to model the economy. More so because the quarterly variables can be explained with corresponding monthly variables (i.e. quarterly GDP can be explained with the factor closely related to the monthly variables that represent economic activity and the GDP deflator to the factor that predominantly explains monthly prices.). In addition the IRs of monthly variables were not affected by adding quarterly series.

[16]Bai and Ng [2007] test is not appropriate in our applications because the identifying assumption underlying their test ($\Lambda\Lambda' = I$) is violated.

cant differences. Specifically, while the policy rate and other interest rates show similar increases after a monetary policy shock, the price puzzle in our model is smaller than that obtained with the Bernanke et al. [2005] model[17]. This might be due to a stronger estimated decrease in housing starts and slightly stronger contraction of monetary base indicators.

In Figure 5, we report the response of the monthly (unobservable) GDP growth rate to the monetary policy shock. For comparison, we add in the same graph the response of monthly IP (with the 90% confidence bands) calculated with the Bernanke et al. [2005] method. The response of GDP has similar shape as the response of IP. It is slightly more pronounced between 5-15 months after the shock and otherwise less pronounced. The results obtained with geometric sampling of GDP growth are similar.

4.2 Bernanke, Gertler, Watson, Sims, and Friedman [1997] Oil Price Shocks

In the second application we reconsider the analysis of the effects of oil price shocks by Bernanke et al. [1997]. They set up a small scale VAR for (in this order): 1) the log of real GDP, 2) the log of the GDP deflator, 3) the log of an index of spot commodity prices, 4) an indicator of the state of the oil market and 5) the level of the federal funds rate. As alternative indicators of the state of the oil market, they assess: the log of the nominal PPI for the crude oil products, Hoover-Perez's oil prices, Mork's oil prices and Hamilton's measure of oil price changes (we refer to Bernanke et al. [1997] for additional details on these measures). They estimate the model on monthly data for the period from 1965 to 1995, using interpolated data for real GDP and the GDP deflator based on a cubic spline.

In Bernanke et al. [1997] an oil price shock is followed by a rise in output for the first year and by a slight short-run decline of prices, when using the log level of oil prices. The other three measures produce better results, although immediately after the oil shock one can still observe a slight increase in output. Eventually, Bernanke et al. [1997] prefer the Hamilton's measure for oil prices since it induces positive price response to an oil shock.

We now repeat their exercise but instead of estimating a VAR we estimate a MF FAVAR, using the updated dataset of McCracken and Ng [2014] and classifying the variables into slow and fast moving indicators as in Bernanke et al. [2005]. We also add to the set of slow moving monthly variables the quarterly GDP growth and GDP deflator, both modeled as a sum of three latent monthly growth rates. We then es-

[17]The price puzzle almost disappears for both methods using a shorter dataset, but then other differences in the responses emerge..

timate a MF-FAVAR with two unobservable factors[18] and three observable factors, $F_t = [f_t^1, f_t^2, P_t^{comm}, P_t^{oil}, i_t]$. The first estimated factor turns out to be highly correlated with real measures of economic activity, and the second one with measures of prices. Hence, the VAR for the factors is similar to that by Bernanke et al. [1997], except that we use estimated factors from mixed frequency data as proxies for real variables and price movements. Using the Stock and Watson [2005] test, discussed in the previous application, we estimate that the number of dynamic factors is 5. The results are presented in Table 5. Hence, we can proceed using the static factors for the structural identification.

Next, we compute the IRFs to oil price shocks using a Cholesky identification, as in Bernanke et al. [1997]. The upper panel of Figure 6 reports the IRFs and the lower panel the cumulated responses (with the 90% confidence bands[19]). After an oil shock, real GDP declines and the GDP deflator rises, after a few periods. This is in line with economic intuition. The responses are statistically significant for GDP growth but not for the GDP deflator. The IRs are often not significant in the application of Bernanke et al. [1997]. After about four months the monetary policy reacts by raising the interest rate, causing the prices to decline but also further depressing the economy. In our application the response of GDP to an oil shock is never positive, whereas it can be positive in Bernanke et al. [1997]. To make sure that the difference is not due to a different time span we also estimated our model for the same period as in Bernanke et al. [1997]. The results were very similar to those obtained for the longer sample, so that we can conclude that the differences are due to different methodology.

An advantage of using a large dataset is that we can also consider the reaction of other variables. For example, in the lower panel of Figure 6, we report the responses of the CPI, IP, employment and hourly earnings. All the reactions are in line with economic theory, since IP decreases, CPI increases, and employment and earnings decrease. This provides additional support for the adopted identification scheme.[20]

[18] We also estimated the model using 3 and 4 latent factors. This did not affect the results significantly. The responses to oil shocks were qualitatively similar to the responses in the model presented above. We present the results from the two latent factors model to facilitate comparability with the VAR model used in Bernanke et al. [1997].

[19] Confidence bands were estimated using sampling with replacement in 500 bootstrap replications. The responses of IR are not cumulated as this variable is in levels.

[20] Kilian and Lewis [2011] criticize the work of Bernanke et al. [1997] on the basis that their results are driven by a specific period and a specific type of shock. Namely, they note that monetary policy response to an oil shock stems from the 1979 oil crisis period and they show that oil price shocks have little impact on interest rate and real output if one instead uses the sample from 1988 onward. In addition Kilian [2009] notes that not all oil price shocks are alike. The response of the economy depends on whether the oil shock is an oil supply shock, demand shock or oil production shock. The IRs that we obtain in our application are qualitatively similar to the IRs that Kilian [2009] obtains for an oil supply shock.

4.3 Ramey [2011] Government Expenditure Shocks

In the last application we investigate how monthly government expenditure shocks (derived from a MF FAVAR with quarterly government expenditure and a set of monthly indicators) affect several macroeconomic variables on a monthly level. Due to its novelty, this application requires a more detailed description. Hence, we describe, in turns, the related literature, the model we implement, and the results.

4.3.1 Related literature

There is no consensus in the literature on the effects of government expenditure shocks. Most researchers agree that GDP and total hours worked increase (though the extent of their reaction is debated), while there is less consensus on the reaction of consumption and real wages. Among others, Fatas and Mihov [2001], Blanchard and Perotti [2002] and Pappa [2005] find that spending shocks raise consumption and real wage. This response is consistent with the new Keynesian models of Rotemberg and Woodford [1992], Devereux, Head, and Lapham [1996] and Gali, Lopez-Salido, and Valles [2007].

Ramey [2011] argues that the positive response of consumption and real wages could be due to timing issues. These arise because government spending changes are announced and are therefore known in advance, before they are implemented. Hence, forward looking agents react to changes in government spending before the changes really occur. If one does not explicitly account for this timing issue in an empirical model, consumption and wages could spuriously increase in response to a government expenditure shock. For this reason, Ramey [2011] uses other variables (instead of government spending) in her study. Specifically, she uses Ramey-Shapiro war dates and shocks to government spending forecasts. Government spending forecasts are forward looking variables, therefore their sudden changes are truly unanticipated. Ramey-Shapiro war dates are instead constructed using a narrative approach. They are characterized as episodes when newspapers suddenly began to forecast large rises in government spending due to prospects of a war. Changes in these variables are less likely to be anticipated. Once controlling for expectations, Ramey [2011] finds that consumption and real wages fall as a response to a spending shock. This result is consistent with the analysis done in Ramey and Shapiro [1998], Edelberg, Eichenbaum, and Fisher [1999] and Burnside, Eichenbaum, and Fisher [2004], and with the response in neoclassical theoretical models (e.g., Aiyagari, Christiano, and Eichenbaum [1992]).

Further investigation should address the issues raised by Kilian and Lewis [2011] and Kilian [2009]. Since this paper is primarily concerned with showing how MF-S-FAVAR can be applied to a large variety of models, we do not further pursue this issue here.

We now try to shed additional light on the effects of government expenditure shocks based on our MF-S-FAVAR framework.

4.3.2 Our MF-FAVAR Model

To investigate the effects of expenditure shocks we estimate a MF-FAVAR model where the majority of the dependent variables are sampled at monthly frequency but one of the observable factors, government expenditures, is quarterly. This enables us to reduce the aggregation bias that is inherent in quarterly models and to avoid the loss of information of low dimensional VARs.

Because our observable factor is the growth rate in quarterly government expenditure (GX), we align it with the dynamics of monthly variables by assuming that the growth rate of quarterly government expenditures is the sum of three consecutive monthly growth rates. These are unobservable but can be estimated by the Kalman filter - smoother, as detailed below, and jointly modeled with the other factors summarizing the dynamics of the economy.

We use the same dataset as in our previous applications. As discussed in the first empirical application, the data consists of various nominal, financial and real indicators (such as consumer prices, producer prices, stocks, commodities and exchange rates, consumption expenditure, production indicators, interest rates and spreads, etc.), from which we extract the factors that describe the dynamics of the economy.

We extract three latent factors and impose two observable factors, the federal funds rate and the real government expenditure growth rate, so that it is $F_t = [GX_t, f_t^1, f_t^2, f_t^3, i_t]$. Monthly real government expenditure is ordered first as it takes longer than a month for a government to implement a change in spending decision or for the automatic stabilizers to respond. This assumption is often used in models estimated on a quarterly frequency. We believe that it is even more plausible in a monthly model. GX_t is modeled as the sum of three consecutive latent monthly growth rates, GX_t^*.

We also restrict the VAR dynamics for government expenditure (only). Mariano and Murasawa [2010] note that when they use higher order VARs to construct the monthly GDP series, the monthly GDP series becomes too volatile. This is due to too many parameters in the model, for a variable with many missing observations. For this reason, they use only one lag of the regressors in the GDP equation. We encountered a similar issue, the resulting estimated monthly GX_t^* variable was too volatile and the impulse responses exhibited a volatile pattern, and we used a similar remedy.[21]

[21] The rest of the factors evolve in a $VAR(p)$ model.

4.3.3 Empirical Results

In this section we describe the impulse responses of the economy to a government expenditure shock on a monthly frequency. Figure 7 plots the response of the core variables (with the 95% confidence bands[22]). The response of prices is not statistically significant on impact but in the long term prices rise although the increase is close to not being significant. Consumer expectations improve and the number of housing starts increases. Labor market indicators improve as total employment immediately increases and unemployment and unemployment duration decline in a delayed fashion. Capacity utilization also improves. The combination of higher prices, better economic conditions and an increase in consumer credit triggers a (still delayed) increase in the federal funds rate. These results are aligned with basic findings of economic theory.

Hourly earnings and average hours worked slightly increase on impact and return to zero after a few months. Real personal consumption increases on impact and consumption of services increases in a delayed fashion. Therefore the response of earnings and consumption in our model is closer to the results in Fatas and Mihov [2001], Blanchard and Perotti [2002] and Pappa [2005], than to the ones obtained by Ramey [2011].

In summary, this application further shows that using a MF-S-FAVAR can shed interesting light on relevant economic issues.

5 Conclusions

In this paper we suggest to extend the FAVAR model to the mixed frequency case (MF-FAVAR) and use it for structural analyses, in order to better exploit all the available information, improve shock identification, and avoid temporal aggregation and variable omission biases.

We illustrate how the MF-FAVAR can be estimated using Kalman filter based techniques and show, by means of Monte Carlo experiments, that the resulting parameter and impulse response estimators work reasonably well also in finite samples.

We then use the MF-FAVAR to evaluate the effects of monetary, oil, and fiscal shocks, comparing the results with those in existing studies. Overall, we obtain reasonable responses in economic terms, sometimes with interesting differences with respect to earlier studies based on same frequency data.

The structural MF-FAVAR model can be applied in a variety of other contexts, and therefore we believe that it is an important item to be added to the standard toolbox of economists.

[22] Confidence bands were estimated using sampling with replacement in 500 bootstrap replications.

References

S.R. Aiyagari, L.J. Christiano, and M. Eichenbaum. The output, employment, and interest rate effects of government consumption. *Journal of Monetary Economics*, 30 (1):73–86, 1992.

E. Andreou, P. Gagliardini, and E. Ghysels. Mixed frequency large scale factor models. *University of North Carolina - Chapel Hill, mimeo*, 2013.

J. Bai. Inferential theory for factor models of large dimensions. *Econometrica*, 71(1): 135–171, 2003.

J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002.

J. Bai and S. Ng. Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics*, 25(1), 2007.

M. Bańbura and M. Modugno. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, 29 (1):133–160, 2014.

M. Bańbura and G. Rünstler. A look into the factor model black box: publication lags and the role of hard and soft data in forecasting gdp. *International Journal of Forecasting*, 27(2):333–346, 2011.

B.S. Bernanke, M. Gertler, M. Watson, C.A Sims, and B.M. Friedman. Systematic monetary policy and the effects of oil price shocks. *Brookings papers on economic activity*, 1997(1):91–157, 1997.

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

O. Blanchard and R. Perotti. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of economics*, 117(4):1329–1368, 2002.

C. Burnside, M. Eichenbaum, and J.D.M. Fisher. Fiscal shocks and their consequences. *Journal of Economic theory*, 115(1):89–117, 2004.

C.W. Chiu, B. Eraker, A.T. Foerster, T.B. Kim, and H.D. Seoane. Estimating vars aampled at mixed or irregular spaced frequencies: a bayesian approach. *Federal Reserve Bank of Kansas City, working paper, 11-11*, 2011.

M.B. Devereux, A.C. Head, and B.J. Lapham. Monopolistic competition, increasing returns, and the effects of government spending. *Journal of Money, Credit and Banking*, 28(2):233–254, 1996.

C. Doz, D. Giannone, and L. Reichlin. A two-step estimator for large approximate dynamic factor models based on kalman filtering. *Journal of Econometrics*, 164(1): 188–205, 2011.

C. Doz, D. Giannone, and L. Reichlin. A quasi–maximum likelihood approach for large, approximate dynamic factor models. *Review of Economics and Statistics*, 94(4):1014–1024, 2012.

J. Durbin and S.J. Koopman. Time series analysis by state space methods. *Oxford University Press*, 2012.

W. Edelberg, M. Eichenbaum, and J.D.M. Fisher. Understanding the effects of a shock to government purchases. *Review of Economic Dynamics*, 2(1):166–206, 1999.

A. Fatas and I. Mihov. The effects of fiscal policy on consumption and employment: theory and evidence. *CEPR discussion paper, No. 2760*, 2001.

M. Forni, D. Giannone, M. Lippi, and L. Reichlin. Opening the black box: Structural factor models with large cross-sections. *Econometric Theory*, 25(05):1319–1347, 2009.

C. Foroni and M. Marcellino. Mixed frequency structural models: estimation, and policy analysis. *Journal of Applied Econometrics, 29, 1118-1114*, 2014.

C. Foroni and M. Marcellino. Mixed frequency structural VARs. *Journal of the Royal Statistical Society, forthcoming*, 2015.

C. Frale, M. Marcellino, G.L. Mazzi, and T. Proietti. Survey data as coincident or leading indicators. *Journal of Forecasting*, 29(1-2):109–131, 2010.

C. Frale, M. Marcellino, G. L. Mazzi, and T. Proietti. Euromind: a monthly indicator of the euro area economic conditions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(2):439–470, 2011.

J. Gali, J. D. Lopez-Salido, and J. Valles. Understanding the effects of government spending on consumption. *Journal of the European Economic Association*, 5(1):227–270, 2007.

E. Ghysels. Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics, forthcoming*, 2015.

D. Giannone, L. Reichlin, and D. Small. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676, 2008.

D. Giannone, F. Monti, and L. Reichlin. Incorporating conjunctural analysis in structural models. *The science and practice of monetary policy today, Springer E-book (Ed. Wieland V.)*, pages 41–57, 2010.

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

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

L. Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *The American Economic Review*, pages 1053–1069, 2009.

L. Kilian and L. T. Lewis. Does the Fed Respond to Oil Price Shocks? *Economic Journal*, 121(555):1047–1072, 09 2011.

M. Marcellino and C. Schumacher. Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp. *Oxford Bulletin of Economics and Statistics*, 72(4):518–550, 2010.

M. Marcellino, C.A. Favero, and F. Neglia. Principal components at work: the empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics*, 20 (5):603–620, 2005.

R.S. Mariano and Y. Murasawa. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4):427–443, 2003.

R.S. Mariano and Y. Murasawa. A coincident index, common factors, and monthly real gdp. *Oxford Bulletin of Economics and Statistics*, 72(1):27–46, 2010.

M.W. McCracken and S. Ng. Fred-md: A monthly database for macroeconomic research. *FRB St. Louis, mimeo*, 2014.

E. Pappa. New-keynesian or rbc transmission? the effects of fiscal shocks in labour markets. *CEPR discussion paper, No. 5313*, 2005.

V.A. Ramey. Identifying government spending shocks: It's all in the timing*. *The Quarterly Journal of Economics*, 126(1):1–50, 2011.

V.A. Ramey and M.D. Shapiro. Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference Series on Public Policy*, 48:145–194, 1998.

J.J. Rotemberg and M. Woodford. Oligopolistic pricing and the effects of aggregate demand on economic activity. *Journal of Political Economy*, 100(6):1153–1207, 1992.

C.A. Sims. Macroeconomics and Reality. *Econometrica*, 48(1):1–48, January 1980.

J.H. Stock and M.W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179, 2002a.

J.H. Stock and M.W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162, 2002b.

J.H. Stock and M.W. Watson. Implications of dynamic factor models for var analysis. *NBER working paper*, (11467), 2005.

H. White. Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, pages 1–25, 1982.

Tables and figures

Table 1: Trace statistic from MC experiments, varying n and T

n=50 d=20: Estimator\Time	T = 50	T = 100	T = 200
PCA on full sample	0.9284	0.9524	0.9608
PCA on the reduced sample	0.9402	0.9531	0.9640
SW estimator	0.9554	0.9690	0.9740
DGR estimator	0.9541	0.9693	0.9752
MF estimator	0.9678	0.9792	0.9836

This table reports trace statistic - a measure of how well the estimated factors track the true factors (eq.(14) on p.13). We fix the sample size to $n = \{50\}$, the number of quarterly series to $d = \{20\}$ and vary sample length $T = \{50, 100, 200\}$. The DGP is described in Section 3.

n=100 d=20: Estimator\Time	T = 50	T = 100	T = 200
PCA on full sample	0.9655	0.9760	0.9802
PCA on the reduced sample	0.9781	0.9835	0.9859
SW estimator	0.9820	0.9868	0.9891
DGR estimator	0.9800	0.9862	0.9888
MF estimator	0.9838	0.9890	0.9911

This table reports trace statistic - a measure of how well the estimated factors track the true factors (eq.(14) on p.13). We fix the sample size to $n = \{100\}$, the number of quarterly series to $d = \{20\}$ and vary sample length $T = \{50, 100, 200\}$. The DGP is described in Section 3.

n=200 d=20: Estimator\Time	T = 50	T = 100	T = 200
PCA on full sample	0.9829	0.9882	0.9902
PCA on the reduced sample	0.9895	0.9919	0.9932
SW estimator	0.9913	0.9937	0.9948
DGR estimator	0.9901	0.9932	0.9946
MF estimator	0.9913	0.9941	0.9953

This table reports trace statistic - a measure of how well the estimated factors track the true factors (eq.(14) on p.13). We fix the sample size to $n = \{200\}$, the number of quarterly series to $d = \{20\}$ and vary sample length $T = \{50, 100, 200\}$. The DGP is described in Section 3.

Table 2: Trace statistic from MC experiments, varying number of quarterly series

T=200, n=200: Estimator \ Qrt. series	d=20	d=100	d=180
PCA on the full sample	0.9902	0.9902	0.9902
PCA on the reduced sample	0.9932	0.9882	0.9519
SW estimator	0.9948	0.9920	0.9714
DGR estimator	0.9946	0.9905	0.9746
MF estimator	0.9953	0.9932	0.9841

This table reports trace statistic - a measure of how well the estimated factors track the true factors (eq.(14) on p.13). We fix the sample size to $n = \{200\}$, sample lenght to $T = \{200\}$ and vary the number of quarterly series $d = \{20, 100, 180\}$. The DGP is described in Section 3.

Table 3: Trace statistic from MC experiments, quarterly (unobservable) factor

T=200, n=200: Estimator \ Qrt. series	d=20	d=100	d= 180
PCA on full sample	0.9901	0.9901	0.9901
PCA on the reduced sample	0.9909	0.9836	0.9239
SW estimator	0.9910	0.9850	0.9350
DGR estimator	0.9913	0.9851	0.9545
MF estimator	0.9916	0.9869	0.9620

This table reports trace statistic - a measure of how well the estimated factors track the true factors (eq.(14) on p.13). We fix the sample size to $n = \{200\}$, sample lenght to $T = \{200\}$ and vary the number of quarterly series $d = \{20, 100, 180\}$. The observable factor is a quarterly variable. The DGP is described in Section 3.

Table 4: Number of dynamic factors in Application 1, Stock and Watson (2005) test

$\dot{\hat{q}}$ \ Criteria	PC1	PC2	PC3	IC1	IC2	IC3
1	0.9464	0.9476	0.942	-0.0474	-0.0458	-0.0531
2	0.9164	0.9188	0.9077	-0.0756	-0.0723	-0.0869
3	0.8996	0.9033	0.8866	-0.0925	-0.0876	-0.1095
4	0.8892	0.8942	0.872	-0.1046	-0.098	-0.1273

This table reports the values of Bai and Ng (2002) information criteria used in the Stock and Watson (2005) test for selecting the number of dynamic factors ($\dot{\hat{q}}$). Number of dynamic factors is estimated to be the one with the smallest value of the information criteria.

Table 5: Number of dynamic factors in Application 2, Stock and Watson (2005) test

$\dot{\hat{q}}$ \ Criteria	PC1	PC2	PC3	IC1	IC2	IC3
1	0.9116	0.9126	0.9086	-0.0742	-0.0727	-0.079
2	0.8824	0.8845	0.8763	-0.0943	-0.0912	-0.1038
3	0.8614	0.8645	0.8523	-0.1113	-0.1066	-0.1255
4	0.8551	0.8591	0.8428	-0.1142	-0.1079	-0.1332
5	0.8567	0.8618	0.8414	-0.1097	-0.1018	-0.1334

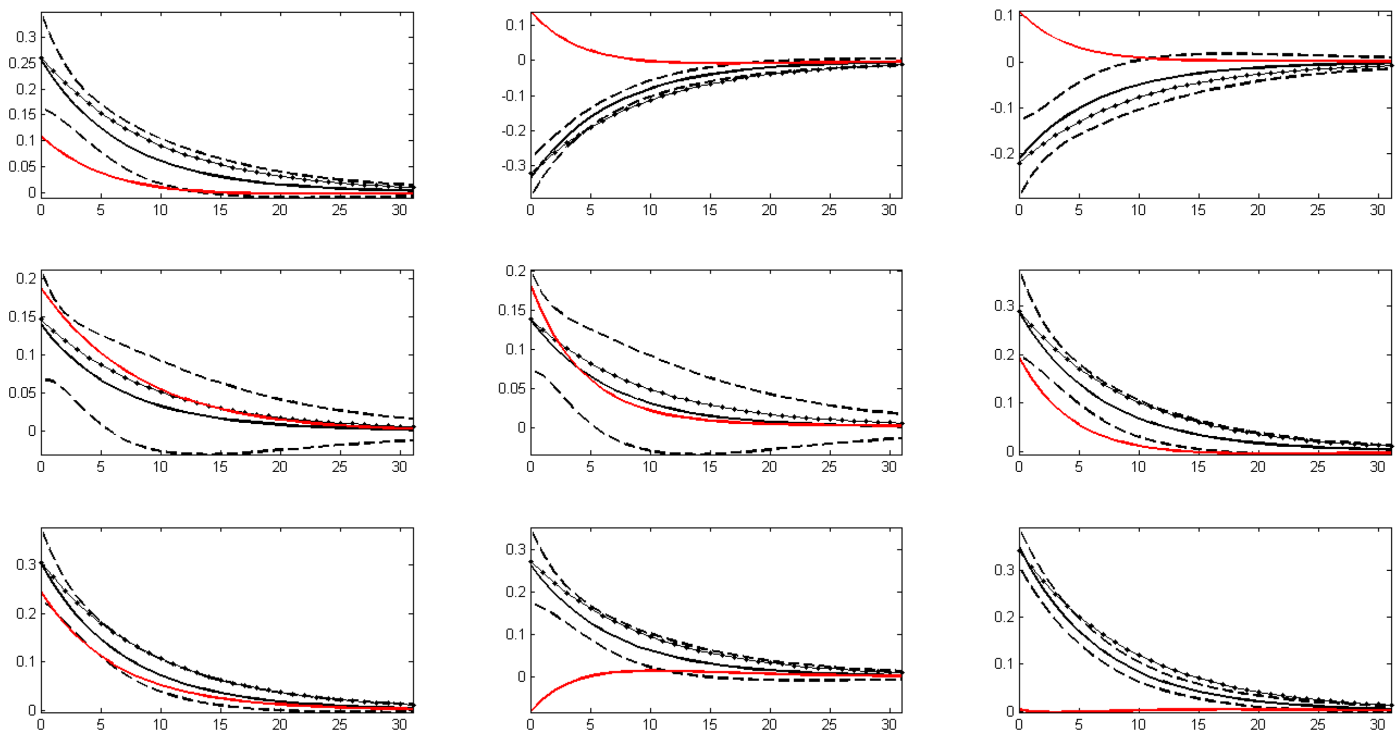
This table reports the values of Bai and Ng (2002) information criteria used in the Stock and Watson (2005) test for selecting the number of dynamic factors ($\dot{\hat{q}}$). Number of dynamic factors is estimated to be the one with the smallest value of the information criteria.

Table 6: Number of dynamic factors in Application 3, Stock and Watson (2005) test

$\dot{\hat{q}}$ \ Criteria	PC1	PC2	PC3	IC1	IC2	IC3
1	0.8879	0.8889	0.8842	-0.1078	-0.1062	-0.1135
2	0.8404	0.8425	0.8331	-0.1564	-0.1531	-0.1678
3	0.8168	0.8199	0.8058	-0.181	-0.1761	-0.1981
4	0.7975	0.8017	0.7828	-0.2042	-0.1976	-0.2269
5	0.7807	0.7859	0.7624	-0.228	-0.2199	-0.2564

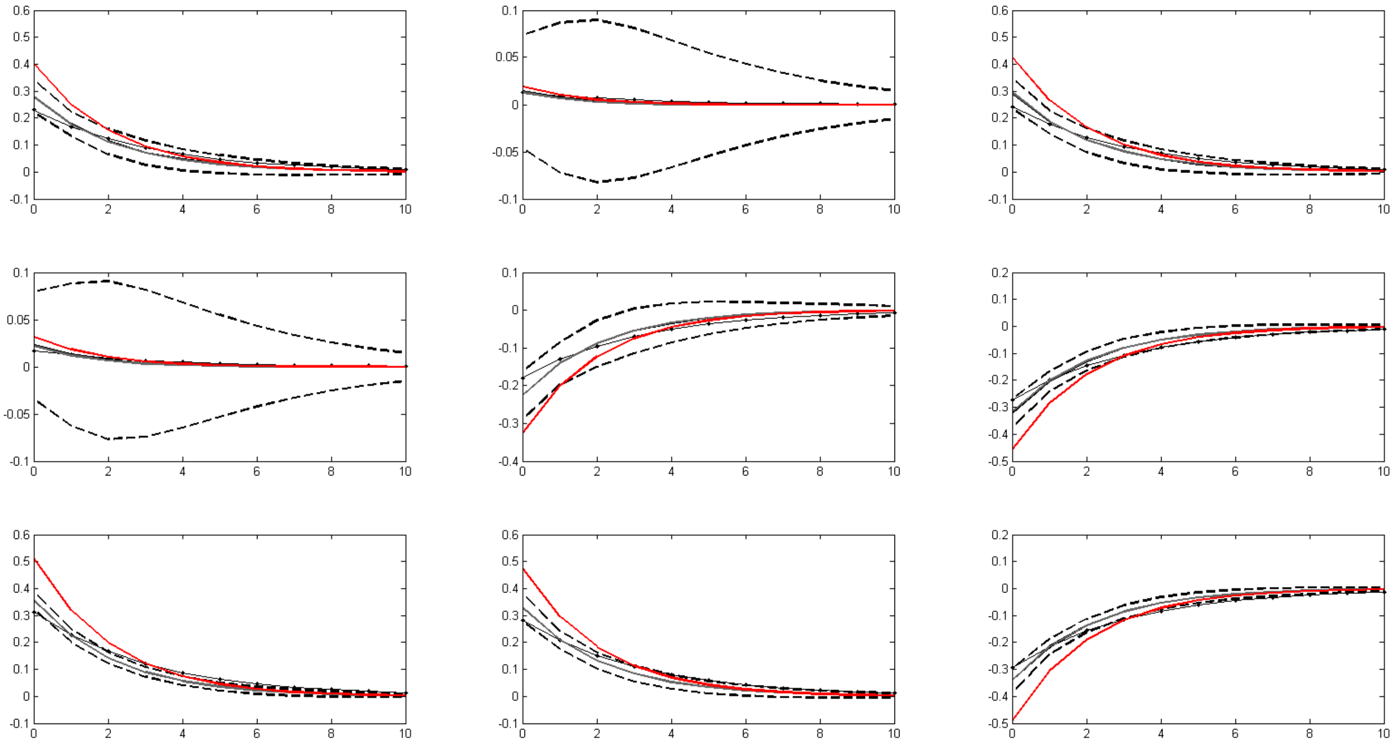
This table reports the values of Bai and Ng (2002) information criteria used in the Stock and Watson (2005) test for selecting the number of dynamic factors ($\dot{\hat{q}}$). Number of dynamic factors is estimated to be the one with the smallest value of the information criteria.

Figure 1: IRs of quarterly variables to a unit shock in the quarterly factor



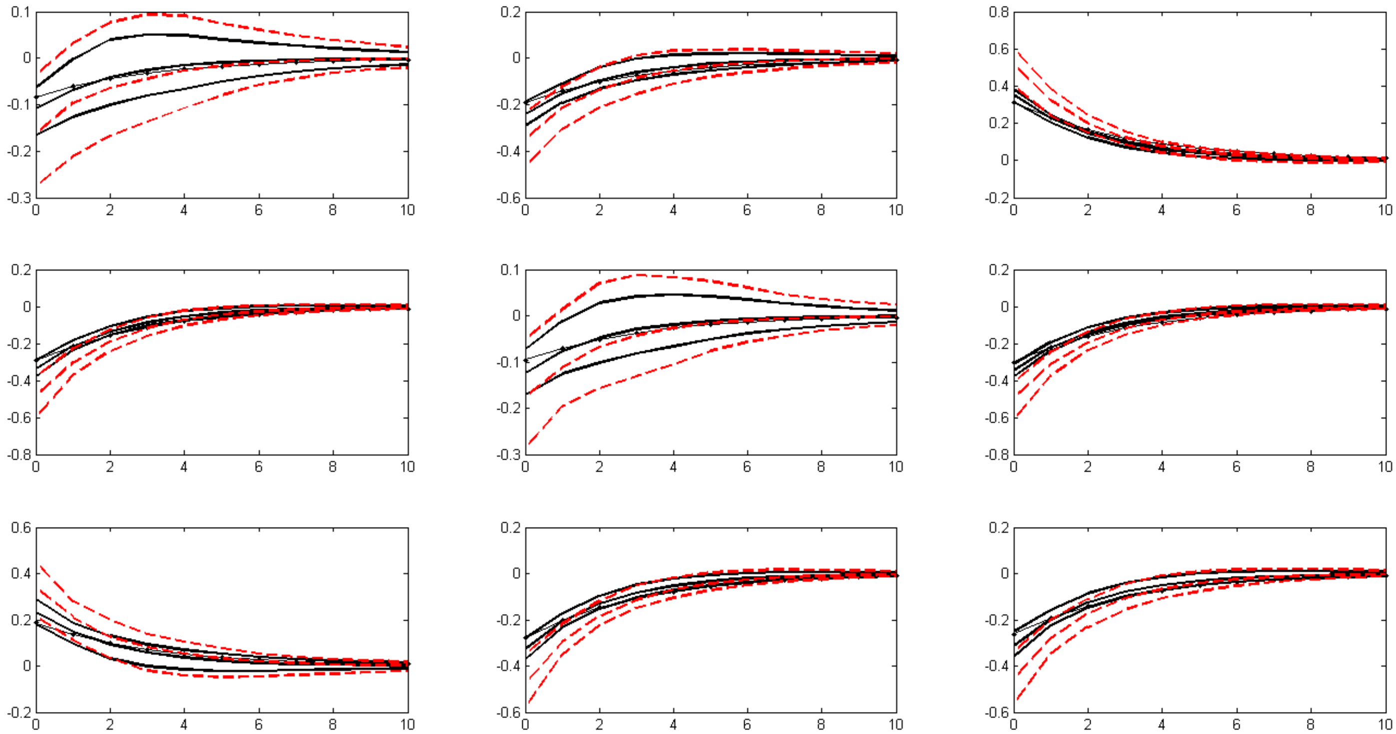
IRs of the first 9 quarterly variables: **true** - solid black with dots, **ML** - solid black, **ML +/- 2 std. dev.** - dashed black, **SW** - solid red. IRs are reported on a monthly frequency. We fix the sample size to $n = \{200\}$, sample length to $T = \{200\}$ and the number of quarterly series to $d = \{100\}$. The DGP is described in Section 3.

Figure 2: (Quarterly) IRs to a unit shock in the monthly observable factor, quarterly variables



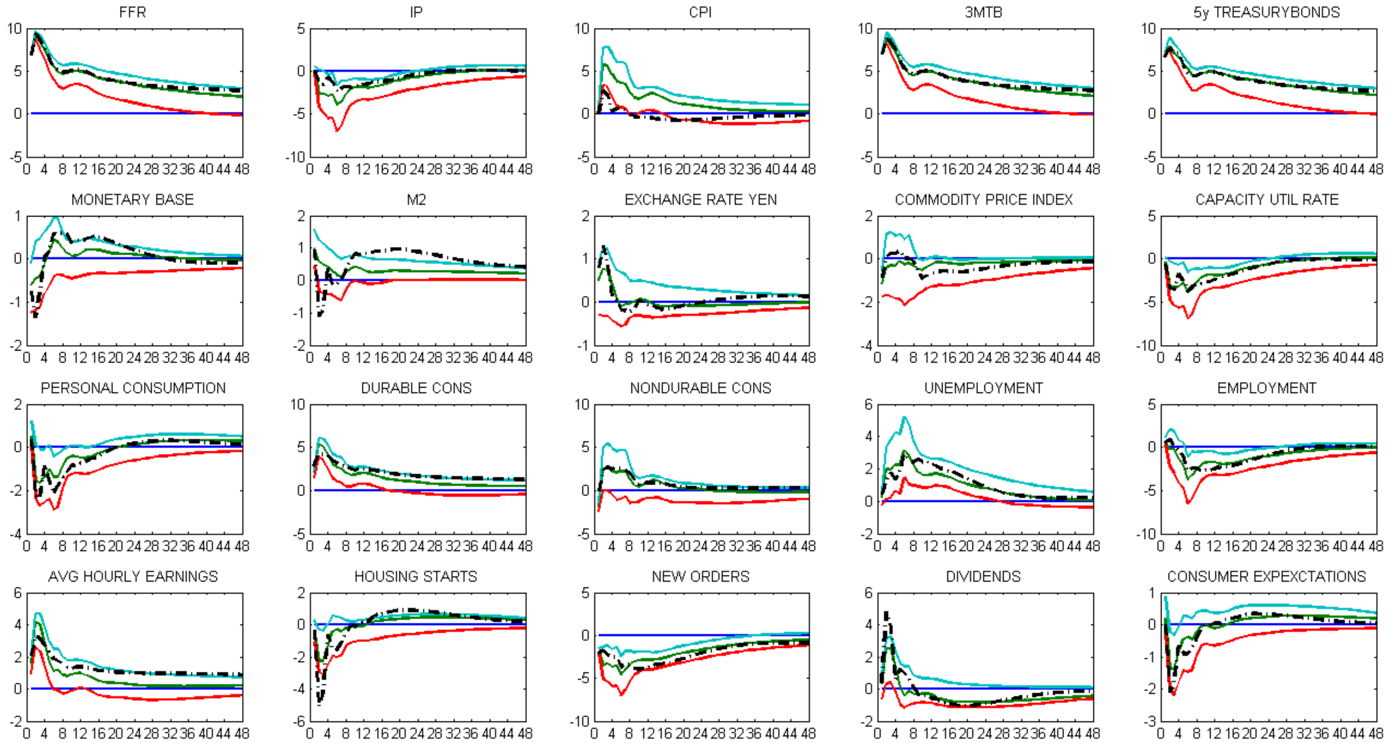
The figure displays the impulse responses of the first 9 quarterly variables to a unit shock in the observable factor, calculated using **monthly dataset** (solid black line, dashed black lines are the +/- 2 std. dev. bands), **mixed frequency dataset** (solid gray line) and **quarterly dataset** (solid red line). The **true IRs** are represented with a solid black line with dots. IRs are reported on a quarterly frequency. Sample size is $n = \{200\}$ and sample length is $T = \{200\}$. The number of quarterly series in a mixed frequency dataset is set to $d = \{100\}$. The DGP is described in Section 3.

Figure 3: (Quarterly) IRs of monthly variables to a unit shock in the observable factor, confidence bands



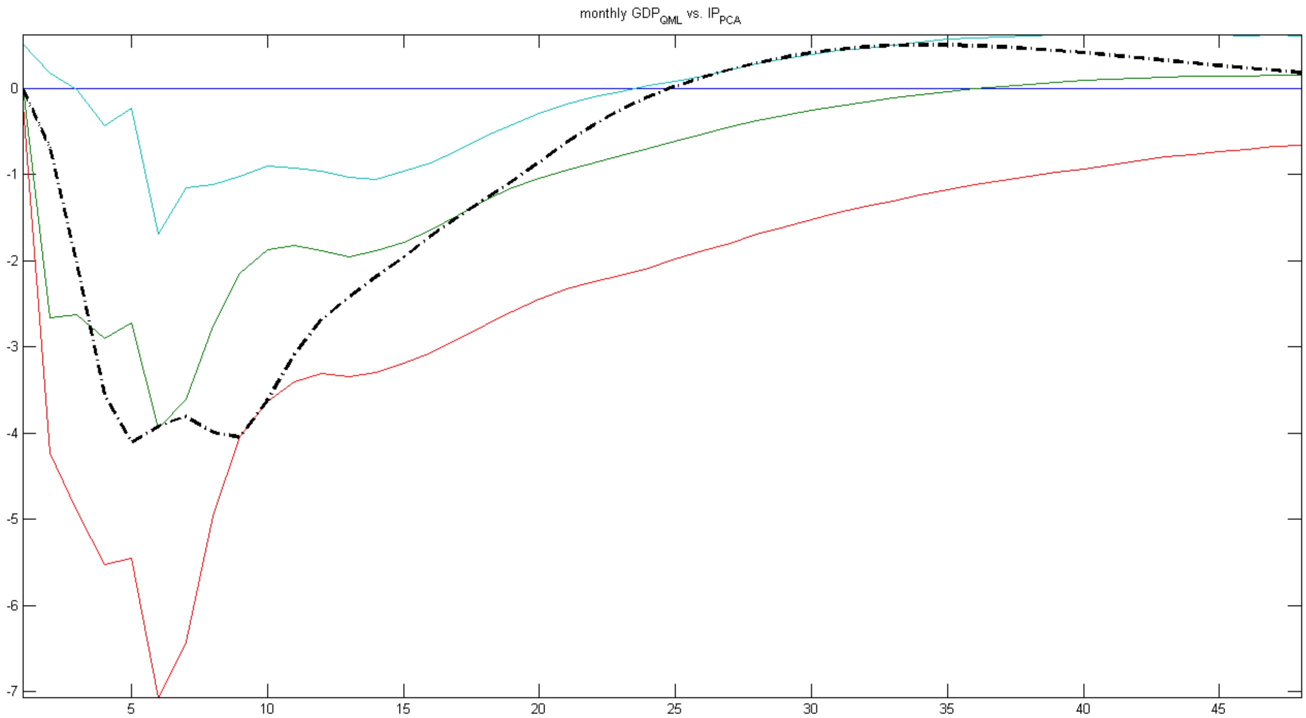
The figure displays the impulse responses of the first 9 monthly variables to a unit shock in the observable factor, calculated using **mixed frequency dataset** (solid black line, dashed black lines are the +/- 2 std. dev. bands) and **quarterly dataset** (dashed red line). The **true IRs** are represented with a solid black line with dots. IRs are reported on a quarterly frequency. Sample size is $n = \{200\}$ and sample length is $T = \{200\}$. The number of quarterly series in a mixed frequency dataset is set to $d = \{100\}$. The DGP is described in Section 3.

Figure 4: IRs of selected variables to a monetary policy shock



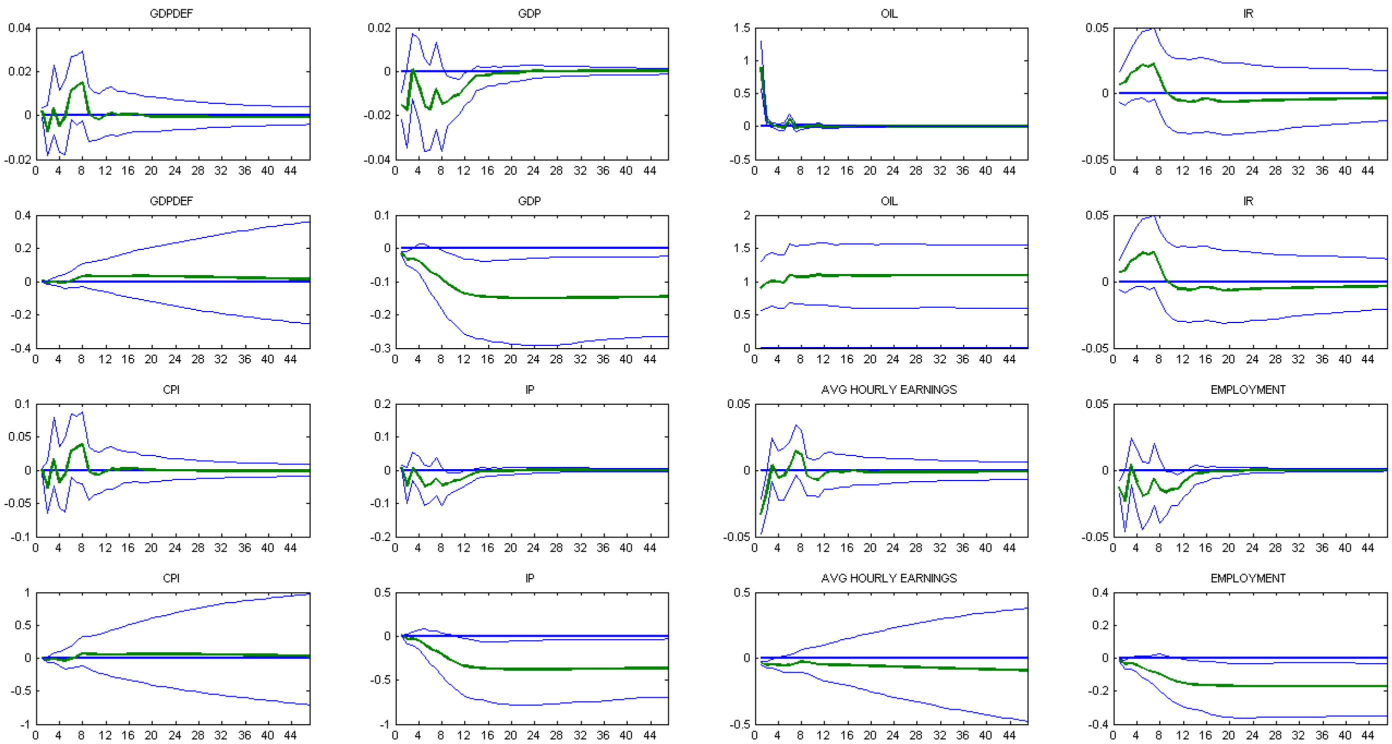
The figure displays the IRs of selected variables to a monetary policy shock, estimated on the McCracken and Ng (2014) dataset. Solid lines represent the IRs (and the confidence bands) calculated using the Bernanke et al. (2005) method and the black dashed lines the IRs estimated with MF-S-FAVAR.

Figure 5: IRs of the industrial production index and monthly GDP to a monetary policy shock



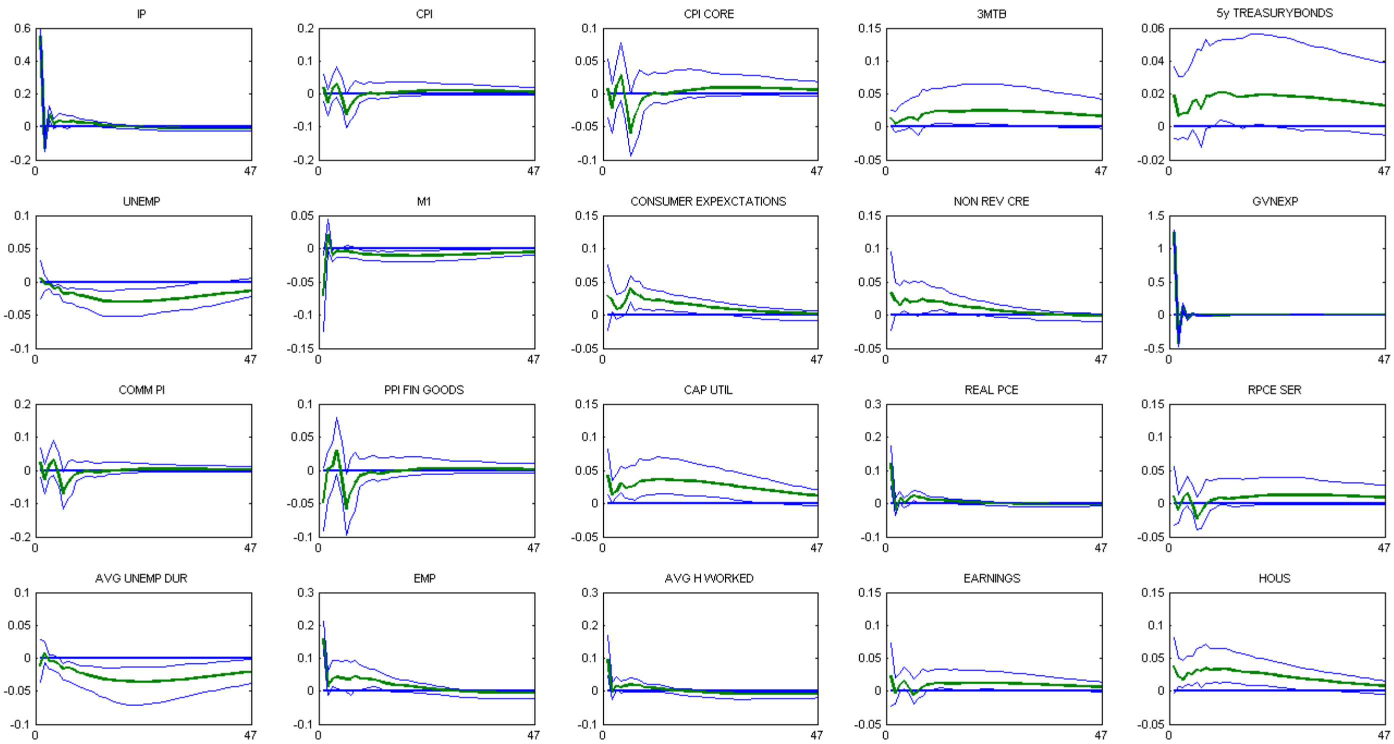
The figure displays the IR of the industrial production index (solid line), calculated using the PCA approach, and of the latent monthly GDP (dashed black line), estimated with MF-S-FAVAR. The IRs were estimated on the McCracken and Ng (2014) dataset

Figure 6: MF-S-FAVAR, the effects of oil price shocks



IRs to an oil price shock estimated on the McCracken and Ng (2014) dataset. For each variable we plot the level (upper figure) and cumulated (lower figure) responses, except for IR that is already in levels. Confidence bands are calculated using 500 bootstrap replications.

Figure 7: MF-FAVAR, IRs to a government expenditure shock - Cholesky identification, government expenditure modelled as a sum of three latent monthly growth rates



IRs of selected variables to a monthly government expenditure shock. IRs are estimated using the McCracken and Ng (2014) dataset. Blue lines are the 95% confidence bands estimated using 500 bootstrap replications.