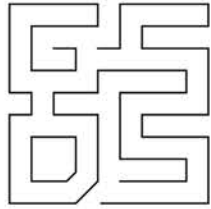
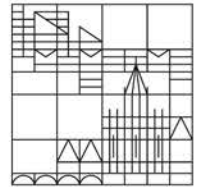


GRADUATE SCHOOL  
OF DECISION SCIENCES



Universität  
Konstanz



GSDS  
Working Paper  
No. 2019-02

# A regularized structural factor-augmented vector autoregressive model

**Maurizio Daniele  
Julie Schnaitmann**

July 2019

## **Graduate School of Decision Sciences**

All processes within our society are based on decisions – whether they are individual or collective decisions. Understanding how these decisions are made will provide the tools with which we can address the root causes of social science issues.

The GSDS offers an open and communicative academic environment for doctoral researchers who deal with issues of decision making and their application to important social science problems. It combines the perspectives of the various social science disciplines for a comprehensive understanding of human decision behavior and its economic and political consequences.

The GSDS primarily focuses on economics, political science and psychology, but also encompasses the complementary disciplines computer science, sociology and statistics. The GSDS is structured around four interdisciplinary research areas: (A) Behavioural Decision Making, (B) Intertemporal Choice and Markets, (C) Political Decisions and Institutions and (D) Information Processing and Statistical Analysis.

GSDS – Graduate School of Decision Sciences  
University of Konstanz  
Box 146  
78457 Konstanz

**Phone:** +49 (0)7531 88 3633

**Fax:** +49 (0)7531 88 5193

**E-mail:** [gsds.office@uni-konstanz.de](mailto:gsds.office@uni-konstanz.de)

**-gsds.uni-konstanz.de**

**ISSN: 2365-4120**

July 2019

© 2019 by the author(s)

# A regularized structural factor-augmented vector autoregressive model\*

Maurizio Daniele<sup>a</sup>

University of Konstanz

GSDS

Julie Schnaitmann<sup>b</sup>

University of Konstanz

GSDS

first draft: July 20, 2018

this version: July 30, 2019

## Abstract

We propose a regularized factor-augmented vector autoregressive (FAVAR) model that allows for sparsity in the factor loadings. In this framework, factors may only load on a subset of variables which simplifies the factor identification and their economic interpretation. We identify the factors in a data-driven manner without imposing specific relations between the unobserved factors and the underlying time series. Using our approach, the effects of structural shocks can be investigated on economically meaningful factors and on all observed time series included in the FAVAR model. We prove consistency for the estimators of the factor loadings, the covariance matrix of the idiosyncratic component, the factors, as well as the autoregressive parameters in the dynamic model. In an empirical application, we investigate the effects of a monetary policy shock on a broad range of economically relevant variables. We identify this shock using a joint identification of the factor model and the structural innovations in the VAR model. We find impulse response functions which are in line with economic rationale, both on the factor aggregates and observed time series level.

*Keywords:* Factor-augmented VAR model, SVAR,  $L_1$ -regularization, Monetary policy shock

*JEL classification:* C32, C38, C55, E52

---

\* An earlier version of the paper was presented at the (EC)<sup>2</sup> 2018 in Rome, the QFFE 2019 in Marseille and the IAAE 2019 in Nicosia. Financial support by the Graduate School of Decision Sciences (GSDS) and the German Science Foundation (DFG) grant number BR 2941/3-1 is gratefully acknowledged.

<sup>a</sup> Department of Economics, Universitätsstraße 1, D-78457 Konstanz, Germany. Phone: +49-7531-88-2657, email: Maurizio.Daniele@uni-konstanz.de.

<sup>b</sup> Corresponding Author: Department of Economics, Universitätsstraße 1, D-78457 Konstanz, Germany. Phone: +49-7531-88-2941, email: Julie.Schnaitmann@uni-konstanz.de.

# 1 Introduction

In this paper, we propose a regularized factor-augmented vector autoregressive (FAVAR) model to investigate the effects of a structural macroeconomic shock on the economy in the presence of many observed time series. Since the seminal work of [Sims \(1980\)](#), small scale vector autoregressive (VAR) models are conventionally used to analyze the dynamic effects of structural shocks on economic systems. Structural shocks arise as linear combinations of the reduced form innovations which depend on the variables included. Since VAR models typically incorporate only relatively few variables, the information set spanned by those models is rather limited and important characteristics on the underlying economy may be omitted. This problem is known as “non-fundamentalness”, which implies that there is no direct mapping between the reduced form and structural innovations and misspecified structural shocks are obtained.<sup>1</sup> In an empirical application, this informational deficiency may e.g. result in misleading impulse response patterns in the form of price puzzles as described in [Sims \(1992\)](#) and [Ramey \(2016\)](#). Hence, in order to span the entire space of the structural shocks, it is crucial to incorporate all relevant variables. However, as the number of parameters increases with the square of the number of included variables, the extent to which a VAR model can be enlarged is limited by the number of observations.

To circumvent these drawbacks, dimension reduction techniques that enable the use of the informational content of many time series for structural analysis have obtained increasing attention. A frequently used approach is the FAVAR model introduced by [Bernanke, Boivin, and Elias \(2005\)](#). FAVAR models have been used to investigate structural monetary policy, fiscal or oil price shocks. An incomplete list comprises the studies by [Del Negro and Otrok \(2007\)](#), [Boivin and Giannoni \(2007\)](#), [Boivin, Giannoni, and Mihov \(2009\)](#), [Mumtaz and Surico \(2009\)](#), [Kilian and Lewis \(2011\)](#) and [Stock and Watson \(2016\)](#).

A FAVAR model decomposes the co-movement of the observed time series into a common and idiosyncratic component. The common component is composed of latent and observed factors that affect the underlying time series according to the corresponding weights represented by the factor loadings matrix. The observable factors behave similar to the latent factors, but they are measured without estimation error. The idiosyncratic innovations allow for weak

---

<sup>1</sup> An overview of the literature is provided in a review paper by [Alessi, Barigozzi, and Capasso \(2011\)](#).

cross-sectional and serial correlation in the spirit of [Chamberlain and Rothschild \(1983\)](#). As the number of latent and observed factors is much smaller than the number of time series included, the information contained in a large panel of variables is condensed into a small number of factors. Hence, the FAVAR framework can be used as a dimension reduction strategy. To allow for dynamics, the FAVAR model incorporates an autoregressive structure in both types of factors. The vector autoregressive system based on the FAVAR model has a rich information set but involves only a small number of factors that need to be estimated. Due to the large number of time series incorporated in the high-dimensional FAVAR process, the model accommodates enough variables to span many macroeconomic shocks without an omitted variable bias.

There is an identification problem associated with the FAVAR approach. In contrast to the observed factors that can be structurally interpreted as they are not subject to an identification problem, only the common component corresponding to the latent factors can be estimated consistently. To identify the latent factor and factor loadings estimates, restrictions have to be imposed on the FAVAR model. In the context of structural analysis, it is crucial to introduce a scheme that allows an economic interpretation of the factors. The named factors identification scheme is used in the factor literature which deals with the analysis of structural shocks. Examples of named factor identification schemes can be found in [Bernanke et al. \(2005\)](#), [Stock and Watson \(2016\)](#) or [Bai, Li, and Lu \(2016\)](#). This scheme associates each latent factor with a unique observed time series, by which the factor is defined. However, these time series may not represent an entire economic sector appropriately. Furthermore, the naming time series potentially impose implausible relations between the factors and the remaining observed time series that affect the structure of the factor model and its dynamics.

In this paper, we propose a regularized FAVAR model that allows for sparsity in the factor loadings matrices of the latent and observed factors. In this framework, factors may only load on a subset of variables which represent different economic sectors. This allows to identify and interpret the factors economically without imposing restrictive identification restrictions through a named factors scheme. Our estimation procedure relies on a penalized quasi-maximum likelihood approach and is based on  $L_1$ -norm regularization of the factor loadings. The estimation of the factor loadings for the latent and observed factors is conducted in a single step. This allows for a similar degree of shrinkage for both type of factors, which is necessary to retain

their structural interpretation.

In the theoretical part of the paper, we prove the consistency for the estimators of the factor loadings, the factors and the covariance matrix of the idiosyncratic component under the average Frobenius norm. Moreover, the autoregressive parameters in the dynamic model are as well consistently estimated.

Using our approach, the impact of structural shocks can be investigated on economically meaningful factors and observed time series using impulse response functions. Moreover, we propose a joint identification scheme of the factor model and the structural innovations in the VAR model that identifies a shock in the observed factors. In a monetary policy application, the key interest lies in the monetary policy shock which amounts to a shock in the structural innovation of the Federal Funds rate (FFR). As many time series react contemporaneously to changes in the FFR, it is commonly treated as an observed factor, see e.g. [Bernanke et al. \(2005\)](#) or [Boivin et al. \(2009\)](#).

Using monthly US macroeconomic data, we use our framework to investigate the effects of a monetary policy shock on the factors and the underlying time series. We are able to identify the factors in a data-driven manner and extract five latent factors that relate to the labor market, prices, industrial production, the stock market and credit spreads. The impulse response functions are in line with economic rationale both on the factor aggregates and the observed time series level. More specifically, we do not observe a price puzzle and the obtained impulse response patterns are economically plausible. In particular, following a tight monetary policy, industrial production falls, credit conditions deteriorate and the level of employment decreases. These results are similar to those obtained by [Forni and Gambetti \(2010\)](#) in a structural dynamic factor model framework. However, the latter is very sensitive to changes in the model specification in comparison to our regularized FAVAR model.

The remainder of this paper is organized as follows. In [Section 2](#) we describe the regularized FAVAR approach and propose an estimation procedure, where we impose shrinkage on the factor loadings of both the observed and unobserved factors. We further provide an identification scheme that jointly identifies the factor model and the structural innovations in the dynamic equation. In [Section 3](#) the large sample properties of the sparse factor loadings, the latent factor estimator, the covariance matrix of the idiosyncratic innovations, as well as the coefficients of

the dynamic equation are provided. The implementation issues are outlined in Section 4. In Section 5 our empirical application investigates the effects of a monetary policy shock on the US economy. Finally, Section 6 concludes.

The following notation is used throughout the paper:  $\pi_{\max}(A)$  and  $\pi_{\min}(A)$  are the maximum and minimum eigenvalue of a matrix  $A$ . Further,  $\|A\|$  and  $\|A\|_F$  denote the spectral and Frobenius norm, respectively. They are defined as  $\|A\| = \sqrt{\pi_{\max}(A'A)}$  and  $\|A\|_F = \sqrt{\text{tr}(A'A)}$ . For some constant  $c > 0$  and a non-random sequence  $b_N$ , we use the notation  $b_N = \mathcal{O}(N)$ , if  $N^{-1}b_N \rightarrow c$ , for  $N \rightarrow \infty$ . Moreover,  $b_N = o(N)$ , if  $N^{-1}b_N \rightarrow 0$ , for  $N \rightarrow \infty$ . Similarly, for a random sequence  $d_N$ , we say  $d_N = \mathcal{O}_p(N)$ , if  $N^{-1}d_N \xrightarrow{p} c$ , for  $N \rightarrow \infty$  and  $d_N = o_p(N)$ , if  $N^{-1}d_N \xrightarrow{p} 0$ , for  $N \rightarrow \infty$ , where  $\xrightarrow{p}$  denotes convergence in probability.

## 2 Econometric modeling framework

### 2.1 Regularized factor-augmented vector autoregressive model

We define the factor-augmented vector autoregressive (FAVAR) model as

$$x_t = \Lambda^f f_t + \Lambda^g g_t + e_t, \quad (2.1)$$

$$\begin{bmatrix} f_t \\ g_t \end{bmatrix} = \begin{bmatrix} \Phi^{ff}(L) & \Phi^{fg}(L) \\ \Phi^{gf}(L) & \Phi^{gg}(L) \end{bmatrix} \begin{bmatrix} f_{t-1} \\ g_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t^f \\ \eta_t^g \end{bmatrix}, \quad t = 1, \dots, T, \quad (2.2)$$

where  $x_t$  is a  $(N \times 1)$  vector of the observable time series,  $f_t$  is a  $(r_1 \times 1)$  vector of latent factors,  $g_t$  is a  $(r_2 \times 1)$  vector of observed factors,  $\Lambda^f$  is a  $(N \times r_1)$  matrix of factor loadings of the latent factors, and  $\Lambda^g$  is a  $(N \times r_2)$  matrix of factor loadings of the observed factors.  $e_t$  is a  $(N \times 1)$  vector of idiosyncratic innovations which may be cross-sectionally and serially correlated, as in the approximate factor model framework. Their covariance matrix is given by  $\Sigma_e = \mathbb{E}[e_t e_t']$ . Moreover,  $\eta_t^f$  denotes a  $(r_1 \times 1)$  and  $\eta_t^g$  a  $(r_2 \times 1)$  vector of factor innovations associated with the unobserved and observed factors, respectively, and  $\Phi^{ff}(L)$ ,  $\Phi^{fg}(L)$ ,  $\Phi^{gf}(L)$  and  $\Phi^{gg}(L)$  are lag polynomials of order  $p$ . We define the vector of latent and observed factors as  $h_t = [f_t', g_t']'$ .

In the FAVAR specification the latent factors are not identified without further restrictions. Their interpretability hinges on the identification restrictions imposed on the factor model. For an exact identification of the model,  $r_1^2 + r_1 r_2$  restrictions have to be imposed see e.g.

Bernanke et al. (2005) or Bai et al. (2016). Among alternative identification schemes, Bai et al. (2016) analyze the named factor scheme which yields estimated factors that can be interpreted economically. Intuitively, the idea is to re-order the underlying observed time series such that the first  $r_1$  time series define the latent factors. It is assumed that these time series only load onto one factor.<sup>2</sup> Hence, the ordering of the time series is crucial for the structure of the factor model. To circumvent this a priori selection of the factors, we introduce a regularized FAVAR model where we impose sparsity on the factor loadings matrices. This leads to a factor loadings matrix which is economically interpretable and enables us to attribute economic meaning to the estimated factors.

A second issue that we target is the estimation of the regularized factor-augmented vector autoregressive model. The complication in this model arises because it consists of latent and observable factors. This implies that the conventionally used methodologies to extract the latent factors do not work in presence of the observed factors. The effect of the observed factors in explaining the covariance structure of the observed data has to be accounted for. The approaches introduced in the literature have tackled this problem in various ways. Bernanke et al. (2005) propose a two-step principal components approach to estimate the FAVAR model which does not take into account that the observed factors also contribute to the factor space used to estimate the latent factors. Bai et al. (2016) address this issue by introducing a two-step quasi maximum likelihood (QML) approach to estimate the FAVAR model. In a first step the impact of the observed factors is linearly projected out. They extract the latent factors and their factor loadings using QML. In a second step, they estimate the factor loadings of the observed factors by linear projection.

We propose a procedure for the joint estimation of the factors loadings of both types of factors. Hence, we do not need to project out the informational content explained by the observed factors prior to estimating the latent factors. This allows us to jointly impose shrinkage on the factor loadings estimates which leads to a similar degree of shrinkage for both types of factors.

---

<sup>2</sup> The resulting factor loadings matrix is given by  $\Lambda_{NF}^f = [I_{r_1}, \Lambda^{f'}]'$ , where  $I_{r_1}$  is a  $(r_1 \times r_1)$  identity matrix and  $\Lambda^{f'} = \Lambda_2^f (\Lambda_1^f)^{-1}$ , with  $\Lambda^f = [\Lambda_1^{f'}, \Lambda_2^{f'}]'$ .



## 2.2 Estimation of the regularized FAVAR model

The factors and factor loadings in equation (2.1) can be estimated by either principal component analysis (PCA)<sup>3</sup> or quasi maximum likelihood estimation under normality.<sup>4</sup> In the following we pursue estimating the factor model by QML, which allows us to introduce sparsity in the factor loadings by penalizing the likelihood function. Moreover, contrary to PCA all model parameters including the covariance matrix  $\Sigma_e$  can be estimated jointly, while PCA-based second stage estimates of  $\Sigma_e$  require consistent estimation of  $\Lambda$  and  $F$  in the first stage, see e.g. [Bai and Li \(2016\)](#).

We propose a penalized likelihood approach to estimate the regularized FAVAR model. Intuitively, we can estimate the factor loadings of the observed and latent factors,  $\Lambda^f$  and  $\Lambda^g$ , jointly using the covariance matrix of the observed time series,  $x_t$ . In matrix notation, the observation equation of the model is defined as

$$X = \Lambda^f F' + \Lambda^g G' + e = \begin{bmatrix} \Lambda^f & \Lambda^g \end{bmatrix} \begin{bmatrix} F' \\ G' \end{bmatrix} + e = \Lambda H + e, \quad (2.3)$$

where  $X = [x_1, \dots, x_T]$  and  $e = [e_1, \dots, e_T]$  are  $(N \times T)$  matrices,  $F = [f_{1t}, \dots, f_{r_1 t}]$  and  $G = [g_{1t}, \dots, g_{r_2 t}]$  are  $(T \times r_1)$  and  $(T \times r_2)$  matrices, respectively. The covariance matrix of the observed time series  $X$  based on the factor model in (2.3) can be written as  $\Sigma = \Lambda \Sigma_H \Lambda' + \Sigma_e$ , where  $\Sigma_H$  is the composite covariance matrix of the observed and unobserved factors and  $\Sigma_e$  denotes the covariance matrix of the idiosyncratic component.

Using the previous result, we obtain the following expression for the negative quasi log-likelihood function for the covariance matrix of the data in the FAVAR model

$$\mathcal{L}(\Lambda, \Sigma_H, \Sigma_e) = \log \left| \det (\Lambda \Sigma_H \Lambda' + \Sigma_e) \right| + \text{tr} \left[ S_x (\Lambda \Sigma_H \Lambda' + \Sigma_e)^{-1} \right], \quad (2.4)$$

where  $S_x = \frac{1}{T} \sum_{t=1}^T x_t x_t'$  is the sample covariance matrix of the observed data. In the first step of our model estimation, we treat  $\Sigma_e$  as a diagonal matrix and define  $\Phi_e = \text{diag}(\Sigma_e)$  denoting a diagonal matrix that contains only the main diagonal elements of  $\Sigma_e$  to reduce the number

<sup>3</sup> See e.g. [Bai and Ng \(2002\)](#) or [Stock and Watson \(2002b\)](#) for a detailed treatment of the PCA, in approximate factor models.

<sup>4</sup> [Bai and Li \(2012\)](#) deal with the consistent estimation of the strict factor model, whereas [Bai and Li \(2016\)](#) analyze the approximate factor model estimation by QML.

of parameters in the estimation. As the number of parameters in  $\Sigma_e$  is potentially of order  $\mathcal{O}(N^2)$ , the joint estimation with the factor loadings and factors becomes intractable. Thus, our unpenalized objective function reduces to

$$\mathcal{L}(\Lambda, \Sigma_H, \Phi_e) = \log \left| \det (\Lambda \Sigma_H \Lambda' + \Phi_e) \right| + \text{tr} \left[ S_x (\Lambda \Sigma_H \Lambda' + \Phi_e)^{-1} \right]. \quad (2.5)$$

As  $\Sigma_e$  incorporates correlations of general form, equation (2.5) may be seen as a quasi log-likelihood function, since it imposes the innovation term structure of a strict factor model. However, Bai and Li (2016) show in the approximate factor model framework that imposing this restrictions on  $\Sigma_e$  does not affect the consistency of the QML estimator. The diagonality assumption on  $\Sigma_e$  is relaxed in a second step based on the soft-thresholding estimator introduced in Section 2.3.

In order to introduce sparsity in the factor loadings matrix  $\Lambda$ , we shrink each element of  $\Lambda$  towards zero. This is incorporated based on a penalized maximum likelihood estimation (MLE) of the objective function in (2.5) by separate  $L_1$ -norm penalties on the factor loadings associated with the observed and unobserved factors, respectively. More specifically, we focus on the following penalized optimization problem

$$\begin{aligned} \min_{\{\Lambda, \Phi_e\}} & \log \left| \det (\Lambda \Sigma_H \Lambda' + \Phi_e) \right| + \text{tr} \left[ S_x (\Lambda \Sigma_H \Lambda' + \Phi_e)^{-1} \right] \\ & + \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} \left| \lambda_{ik}^f \right| + \mu_2 \sum_{i=1}^N \sum_{l=r_1+1}^r \left| \lambda_{il}^g \right|, \end{aligned} \quad (2.6)$$

where  $\mu_1$  and  $\mu_2$  determine the degree of penalization of the factor loadings corresponding to the unobserved and observed factors, respectively. A clear separation of both sets of factor loadings has the advantage of offering a more flexible treatment of both components.

The latent factors  $f_t$  can be estimated by generalized least squares (GLS)

$$\tilde{f}_t = \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f \right)^{-1} \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} x_t, \quad (2.7)$$

where the estimates  $\tilde{\Lambda}^f$  and  $\tilde{\Phi}_e$  are the ones obtained from the optimization of the objective function in (2.6).

### 2.3 Estimation of the idiosyncratic component covariance matrix $\Sigma_e$

As the diagonality assumption on  $\Sigma_e$  that we impose in the first step of our estimation is restrictive, we introduce a procedure that relaxes the assumption and allows for the estimation of a possibly dense idiosyncratic innovation covariance matrix. More specifically, we re-estimate  $\Sigma_e$  by means of the principal orthogonal complement thresholding (POET) estimator introduced by [Fan, Liao, and Mincheva \(2013\)](#). The POET estimator allows for sparsity in the idiosyncratic error covariance matrix by shrinking the off-diagonal elements of the sample covariance matrix of the residuals obtained from the estimation of our regularized FAVAR model towards zero using the soft-thresholding method. The estimated idiosyncratic error covariance matrix  $\tilde{\Sigma}_e^\tau$  based on the POET method is defined as

$$\tilde{\Sigma}_e^\tau = s_{ij}^\tau, \quad s_{ij}^\tau = \begin{cases} s_{e,ii}, & i = j \\ \mathcal{S}(s_{e,ij}, \tau), & i \neq j \end{cases},$$

where  $s_{e,ij}$  is the  $ij$ -th element of the sample covariance matrix  $S_e = \frac{1}{T} \sum_{t=1}^T (x_t - \tilde{\Lambda} \tilde{h}_t)(x_t - \tilde{\Lambda} \tilde{h}_t)'$  of the estimated factor model residuals,  $\tilde{h}_t = (\tilde{f}_t', g_t')'$  are the estimated factors,  $\tau = \frac{1}{\sqrt{N}} + \sqrt{\frac{\log N}{T}}$  is a threshold and  $\mathcal{S}(\cdot)$  denotes the soft-thresholding operator defined as

$$\mathcal{S}(\sigma_{e,ij}, \tau) = \text{sign}(\sigma_{e,ij})(|\sigma_{e,ij}| - \tau)_+. \quad (2.8)$$

### 2.4 Identification restrictions

The FAVAR model in (2.1) and (2.2) is not identified, as it can be expressed as

$$\begin{aligned} x_t &= \Lambda A^{-1} \cdot A h_t + e_t \\ A h_t &= A \Phi_1 A^{-1} \cdot A h_{t-1} + \dots + A \Phi_p A^{-1} \cdot A h_{t-p} + A \eta_t, \end{aligned}$$

where  $h_t = (f_t', g_t')'$  and  $A$  denotes an invertible  $[(r_1 + r_2) \times (r_1 + r_2)]$  matrix. Moreover,  $\Lambda^* = \Lambda A^{-1}$ ,  $h_{t-j}^* = A h_{t-j}$  for all  $j = 0, \dots, p$ ,  $\Phi_i^* = A \Phi_i A^{-1}$  for all  $i = 1, \dots, p$  and  $\eta_t^* = A \eta_t$  are the uniquely identified quantities given a specific choice of  $A$ . [Bai et al. \(2016\)](#) show that  $r_1^2 + r_1 r_2$  restrictions are necessary to uniquely identify the FAVAR model. Hence, we impose restrictions on the factor model and the dynamics of the factors.

The covariance matrix of the VAR innovations  $\eta_t^* = (\eta_t^{*f'}, \eta_t^{*g'})'$  is given by

$$\Omega^* = \begin{bmatrix} \mathbb{E} \begin{bmatrix} \eta_t^{*f} & \eta_t^{*f'} \end{bmatrix} & \mathbb{E} \begin{bmatrix} \eta_t^{*f} & \eta_t^{*g'} \end{bmatrix} \\ \mathbb{E} \begin{bmatrix} \eta_t^{*g} & \eta_t^{*f'} \end{bmatrix} & \mathbb{E} \begin{bmatrix} \eta_t^{*g} & \eta_t^{*g'} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \Omega_{ff}^* & \Omega_{fg}^* \\ \Omega_{gf}^* & \Omega_{gg}^* \end{bmatrix}. \quad (2.9)$$

We consider the following sets of identification restrictions on the covariance matrix of the factor innovations and the factor model.<sup>5</sup>

$$IRa: \Omega_{fg}^* = 0 \text{ and } \Sigma_F = I_{r_1}.$$

$$IRb: \Omega_{fg}^* = 0 \text{ and } \Lambda_1 = I_{r_1}, \text{ where } \Lambda_1 \text{ is the upper } r_1 \times r_1 \text{ submatrix of } \Lambda.$$

Both identification schemes share the restriction that  $\Omega_{fg}^* = 0$ . This restriction assures that the observed factors  $g_t$  are not rotated. Moreover, as this assumption is imposed on the covariance matrix of the structural factor innovations this implies that the structural shocks associated with the observed factors are contemporaneously uncorrelated with those of the unobserved ones.

The first set of identifying restrictions *IRa* is used for our regularized FAVAR model. As long as some degree of sparsity is imposed on the factor loadings (i.e.  $\mu_1, \mu_2 > 0$ ), the regularized FAVAR model is identified up to a unitary generalized permutation matrix,<sup>6</sup> by the combination of *IRa* and the  $L_1$ -norm penalties on  $\Lambda$ . Hence, the regularized FAVAR model is fully identified, by ordering the columns of the estimated factor loadings matrix  $\tilde{\Lambda}$  according to their sparsity and by fixing the sign of each column. We choose the sign of the estimated factors such that they align to the corresponding observable time series.

However, if there is no sparsity imposed on the factor loadings matrix (i.e.  $\mu_1 = \mu_2 = 0$ ), we need restrictions in addition to *IRa* to identify the FAVAR model. For this case, we further use the normalization  $\frac{1}{N} \Lambda' \Sigma_e^{-1} \Lambda = Q$ , where the diagonal entries of  $Q$  are assumed to be distinct and arranged in a decreasing order. These restrictions are usually imposed in the MLE framework for the approximate factor model (see e.g. [Lawley and Maxwell \(1971\)](#)).<sup>7</sup> The second restriction

<sup>5</sup> Similar identification restrictions are common in the factor model literature, see e.g. [Bai and Ng \(2013\)](#) or [Bai and Wang \(2014\)](#).

<sup>6</sup> A unitary generalized permutation matrix is the only isometry for the  $L_1$ -norm penalty (see [Horn and Johnson \(2012\)](#)).

<sup>7</sup> Alternatively, we could also impose the restriction  $\Omega_{ff}^* = I_{r_1}, \Omega_{fg}^* = 0$  and  $\frac{1}{N} \Lambda' \Sigma_e^{-1} \Lambda = Q$  (see e.g. identification restriction *IRa* in [Bai et al. \(2016\)](#)). This imposes an orthogonality restriction on the covariance matrix of the dynamic factor innovations. However, this version implies that the factor loadings of the unobserved factors are rotated.

*IRb* is conventionally<sup>8</sup> referred to as named factor identification. The first  $r_1$  time series offer a direct identification for the unobserved factors and are assumed to only load on the specific factors, respectively. However, as a consequence the ordering of the observed time series matters crucially for this set of restrictions.

Based on the different sets of identification schemes the resulting estimates change. In scheme *IRa*, the identification restrictions are chosen to have no impact on the rotation of the factor loadings corresponding to the unobserved factors  $\tilde{\Lambda}^f$ . Hence, the interpretation of  $\hat{\Lambda}^f$  is not distorted by any rotation and the factor loadings estimates are unchanged  $\hat{\Lambda}^f = \tilde{\Lambda}^f$ . The factor loadings associated with the observed factors are transformed according to  $\hat{\Lambda}^g = \tilde{\Lambda}^g + \tilde{\Lambda}^f \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1}$ . Moreover, we rotate the estimated factors by  $\hat{F} = \tilde{F} - \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} G$  and the autoregressive parameters by  $\hat{\Phi}_i = \tilde{A} \tilde{\Phi}_i \tilde{A}^{-1}$ , for  $i = 1, \dots, p$ , with the rotation matrix  $\tilde{A} = \begin{bmatrix} I_{r_1} & -\tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} \\ 0 & I_{r_2} \end{bmatrix}$ . In case of *IRb*, let  $\tilde{\Lambda}_1$  be the first  $r_1 \times r_1$  block of  $\tilde{\Lambda}$ . The resulting estimated factor loadings are  $\hat{\Lambda}^f = \tilde{\Lambda} \tilde{\Lambda}_1^{-1}$  and  $\hat{\Lambda}^g = \tilde{\Lambda}^g + \tilde{\Lambda}^f \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1}$ . The estimated factors are given by  $\hat{F} = \left( \tilde{F} - \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} G \right) \tilde{\Lambda}'_1$  and the autoregressive parameters are obtained as  $\hat{\Phi}_i = \tilde{A} \tilde{\Phi}_i \tilde{A}^{-1}$ , for  $i = 1, \dots, p$ , where we use the rotation matrix  $\tilde{A} = \begin{bmatrix} \tilde{\Lambda}_1 & -\tilde{\Lambda}_1 \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} \\ 0 & I_{r_2} \end{bmatrix}$ . Note that quantities with a '˜' are estimates of the identified quantities denoted by a '\*'.

## 2.5 Impulse responses

Starting from the VAR representation of the factors in equation (2.2), we obtain the vector moving average representation of the model and the impulse response functions. By rewriting the dynamic equation of the factors as  $(I_r - \Phi_1 L - \dots - \Phi_p L^p) h_t = \eta_t$  and using the inverted lag polynomial to write the factors  $h_t$  as a function of their innovations  $\eta_t$ , we obtain

$$h_t = (I_r - \Phi_1 L - \dots - \Phi_p L^p)^{-1} \eta_t = \Psi(L) \eta_t,$$

where  $\Psi(L) = I_r + \Psi_1 L + \Psi_2 L^2 + \dots$  and the  $\Psi_i$  are moving average coefficients. Hence,  $\Psi(L)$  is the conventional inverted lag polynomial and the  $\Psi_i$  can be interpreted as the matrices of responses to the innovations  $\eta_t$ . In our setting, we transform the latent and observed factors to stationary time series, if necessary. Furthermore, we accumulate the moving average parameters

<sup>8</sup> Restriction *IRb* is also employed by [Bai et al. \(2016\)](#).

to analyze the effect of a shock in the factor innovations on the level of the factors, when the underlying observed time series  $x_t$  are given in first-differences.

In the factor-augmented VAR model, we can also calculate the responses of the observed time series by employing

$$x_t = \Lambda \Psi(L) \eta_t.$$

In addition, we can investigate the impact of a shock in the factor innovations on all time series included in the factor model. Impulse responses on the individual time series in levels are obtained by pre-multiplying the moving average coefficients  $\Psi_i$  by the factor loadings matrix  $\Lambda$  and accumulating them, when necessary.

Shocks to the factor innovations  $\eta_t$  cannot be interpreted as structural shocks as they are contemporaneously correlated. However, as we identify the factor model and the structural innovations jointly by the schemes described in Section 2.4, we obtain innovations with a block diagonal structure. The structural innovations are given by pre-multiplying the factors and factor innovations by the rotation matrix. The contemporaneous impact matrix is obtained by inverting the rotation matrix and is given by

$$A^{-1} = \begin{bmatrix} I_{r_1} & \Omega_{fg} \Omega_{gg}^{-1} \\ 0 & I_{r_2} \end{bmatrix}, \quad (2.10)$$

where  $\Omega_{fg}$  and  $\Omega_{gg}$  are the corresponding elements of the covariance matrix of the factor innovations  $\eta_t$ , which is denoted as  $\Omega = \begin{bmatrix} \Omega_{ff} & \Omega_{fg} \\ \Omega_{gf} & \Omega_{gg} \end{bmatrix}$ . The covariance matrix of the structural

innovations  $\eta_t^*$  is given by the block diagonal matrix  $\Omega^* = \begin{bmatrix} \Omega_{f.g} & 0 \\ 0 & \Omega_{gg} \end{bmatrix}$ , where  $\Omega_{f.g} = \Omega_{ff} - \Omega_{fg} \Omega_{gg}^{-1} \Omega_{gf}$ . This implies that a structural shock in the innovation of the observed

factors are contemporaneously uncorrelated with those of the latent factors. Shocks to the structural innovation of  $g_t$  can impact all latent factors contemporaneously. We do not impose zero restrictions on the last  $r_2$  columns of the impact matrix  $A^{-1}$  in (2.10). For the named factor identification scheme (*IRb*), we obtain a block diagonal covariance matrix of the following

$$\text{form } \Omega^* = \begin{bmatrix} \Lambda_1 \Omega_{f.g} \Lambda_1' & 0 \\ 0 & \Omega_{gg} \end{bmatrix}.$$

In case of only one observable factor, i.e.  $r_2 = 1$ , the covariance of the structural innovations  $\Omega_{gg}$  is a scalar. Hence, in the FAVAR context the structural innovations of the observed factors are identified by the factor model identification irrespective of the identification scheme used.

For more than one observable factor, i.e.  $r_2 > 1$ , we can generalize our model. The inverse of the contemporaneous impact matrix can be written as  $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$  and the structural innovations of the observed factors are given by  $\eta_{g,t}^* = A_{22}\eta_{g,t}$ . Hence, their covariance matrix is expressed as

$$\Omega_{gg}^* = \mathbb{E} \left[ \eta_{g,t}^* \eta_{g,t}^{*'} \right] = A_{22} \Omega_{gg} A_{22}'. \quad (2.11)$$

In this setting, we impose restrictions on  $A_{22}$  to achieve identification of the structural innovations with respect to the observable factors  $g_t$ . For example, a recursive structure within the observable factors could be used.

Moreover, we derive the impact matrix for the observed data  $x_t$ . It describes the contemporaneous impact of an exogenous structural shock on the observed time series. Starting from the implied structural moving average representation for the observed data in lag operator notation,  $x_t = \Lambda \Psi(L) A^{-1} \eta_t^*$  where  $\Psi(L) = I_r + \Psi_1 L + \Psi_2 L^2 + \dots$ , the model in period  $t$  reads  $x_t = \Lambda A^{-1} \eta_t^* + \Lambda \Psi_1 A^{-1} \eta_{t-1}^* + \Lambda \Psi_2 A^{-1} \eta_{t-2}^* + \dots$ . The impact matrix is denoted as  $B_0 = \Lambda^* = \Lambda A^{-1}$ . Using the estimated quantities, the impact matrix is given by

$$\hat{B}_0 = \begin{bmatrix} \tilde{\Lambda}^f & \tilde{\Lambda}^f \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} + \tilde{\Lambda}^g \end{bmatrix} = \hat{\Lambda}. \quad (2.12)$$

Hence, the sparsity in  $\hat{\Lambda}$  provides structure to the contemporaneous impact matrix. The zeros in the factor loadings yield timing restrictions which are data-driven but can be explained economically. Moreover, the strength and direction of the effect is guided by the regularized factor loadings.

Inference on the impulse response functions of the regularized FAVAR model is conducted using a residual-bootstrap based on the dynamic factor equation (2.2). As shown in the following section, all estimators based on the regularized FAVAR are consistent. Hence, the estimated factors are taken as given in the VAR bootstrap. This approach is in the spirit of the factor augmented regression literature as in Bai and Ng (2006) and Gonçalves and Perron (2014).

### 3 Asymptotic properties

In this section we establish the consistency of the first step estimators of the regularized FAVAR model introduced in equation (2.6). The proofs of the stated theorems are similar to those in Daniele, Pohlmeier, and Zagidullina (2018) and Bai et al. (2016). Subsequently, we state the necessary assumptions.

**Assumption 3.1** (Data generating process).

(i)  $\{e_t, h_t\}_{t \geq 1}$  are strictly stationary, for  $h_t = (f_t', g_t')$ . In addition,  $\frac{1}{T} \sum_{t=1}^T f_t g_t' = 0$  and  $\mathbb{E}[e_{it}] = \mathbb{E}[e_{it} h_{kt}] = 0$ , for all  $i \leq N$ ,  $k \leq r$  and  $t \leq T$ .

(ii) There exists a constant  $c > 0$  such that, for all  $N$ ,

$$c^{-1} < \pi_{\min} \left( \frac{\Lambda' \Lambda}{N^\beta} \right) \leq \pi_{\max} \left( \frac{\Lambda' \Lambda}{N^\beta} \right) < c, \text{ where } 1/2 \leq \beta \leq 1.^9$$

(iii) There exists  $r_1, r_2 > 0$  and  $b_1, b_2 > 0$ , such that for any  $s > 0$ ,  $i \leq N$  and  $k \leq r$ ,

$$\mathbb{P}(|e_{it}| > s) \leq \exp(-(s/b_1)^{r_1}), \quad \mathbb{P}(|h_{kt}| > s) \leq \exp(-(s/b_2)^{r_2})$$

(iv) Define the mixing coefficient:

$$\alpha(T) = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_T^\infty} |\mathbb{P}(A) \mathbb{P}(B) - \mathbb{P}(AB)|,$$

where  $\mathcal{F}_{-\infty}^0$  and  $\mathcal{F}_T^\infty$  denote the  $\sigma$ -algebras generated by  $\{(h_t, e_t) : -\infty \leq t \leq 0\}$  and  $\{(h_t, e_t) : T \leq t \leq \infty\}$

Strong mixing: There exist  $r_3 > 0$  and  $C > 0$  satisfying: for all  $T \in \mathbb{Z}^+$ ,

$$\alpha(T) \leq \exp(-CT^{r_3})$$

(v) There exist constants  $c_1, c_2, c_3 > 0$  such that  $c_2 \leq \pi_{\min}(\Sigma_{e0}) \leq \pi_{\max}(\Sigma_{e0}) \leq c_1$  and

$$\max_{i \leq N} \max_{k \leq r} |\lambda_{ik0}| < c_3.$$

<sup>9</sup> The lower limit 1/2 for  $\beta$  is necessary for a consistent estimation of the factors. See Lemma A.4 in Section A.1 in the Appendix A.



The assumptions in 3.1 impose regularity conditions on the data generating process and are similar to those imposed by Bai and Liao (2016). Condition (i) induces strict stationarity for  $e_t$  and  $h_t$  and requires that both terms are uncorrelated. Furthermore, it implies that the observed and unobserved factors are orthogonal, which simplifies our technicalities. Condition (ii) relaxes the pervasiveness assumption commonly imposed in the approximate factor model literature and allows for weak factors. Condition (iii) requires exponential-type tails, which allows to use large deviation theory for  $\frac{1}{T} \sum_{t=1}^T e_{it}e_{jt} - \sigma_{e,ij}$  and  $\frac{1}{T} \sum_{t=1}^T h_{jt}e_{it}$ . Condition (iv) imposes a strong mixing condition to allow for weak serial dependence. Further, Condition (v) implies bounded eigenvalues of the idiosyncratic error covariance matrix, which is a common identifying assumption in the factor model framework.

To control the sparsity in both  $\Lambda$  and  $\Sigma_e$ , we impose the following sparsity assumptions

**Assumption 3.2** (Sparsity).

$$(i) L_N = \sum_{i=1}^N \mathbb{1} \{ \lambda_{ik} \neq 0 \} = \mathcal{O}(N), \quad \forall k = 1, \dots, r$$

$$(ii) S_N = \max_{i \leq N} \sum_{j=1}^N \mathbb{1} \{ \sigma_{e,ij} \neq 0 \}, S_N^2 d_T = o_p(1) \text{ and } S_N \max(\mu_1, \mu_2) = o(1),$$

where  $\mathbb{1} \{ \cdot \}$  defines an indicator function that is equal to one if the boolean argument in braces is true.

Condition (i) defines the quantity  $L_N$  that represents the number of non-zero elements in the factor loadings matrix  $\Lambda$ . As the number of factors  $r$  are assumed to be constant, (i) upper bounds the number of non-zero elements in each column of  $\Lambda$  by  $N$ . Condition (ii) specifies  $S_N$  that corresponds to the maximum number of non-zero elements in each row of  $\Sigma_e$ , following the definition in Fan et al. (2013).

**Theorem 3.1** (Consistency of the first step estimators of the regularized FAVAR model).

Under Assumptions 3.1 and 3.2 the regularized FAVAR model in (2.6) satisfies for  $T$  and  $N \rightarrow \infty$ , the following properties

$$\begin{aligned} \frac{1}{N} \left\| \tilde{\Lambda} - \Lambda \right\|_F^2 &= \mathcal{O}_p \left( \max(\mu_1, \mu_2) + \sqrt{\frac{N(d_T + \max(\mu_1, \mu_2))}{L_N}} \right), \\ \frac{1}{N} \left\| \tilde{\Phi}_e - \Phi_e \right\|_F^2 &= \mathcal{O}_p \left( \frac{\log N^\beta}{N} + \frac{\log N}{T} \right), \end{aligned}$$

where  $d_T = \frac{\log N^\beta}{N} + \frac{1}{N^\beta} \frac{\log N}{T}$ , for  $1/2 \leq \beta \leq 1$ .

Hence, for  $\log(N) = o(T)$  and the regularization parameters  $\mu_1 = o(1)$ ,  $\mu_2 = o(1)$ , we obtain

$$\frac{1}{N} \left\| \tilde{\Lambda} - \Lambda \right\|_F^2 = o_p(1), \quad \frac{1}{N} \left\| \tilde{\Phi}_e - \Phi_e \right\|_F^2 = o_p(1).$$

Furthermore, for all  $t \leq T$ :

$$\left\| \tilde{f}_t - f_t \right\| = o_p(1)$$

For the second step estimator of the idiosyncratic error covariance matrix, specified in Section 2.3, we get

$$\left\| \tilde{\Sigma}_e^T - \Sigma_e \right\| = \mathcal{O}_p \left( S_N \sqrt{\max(\mu_1, \mu_2)^2 + \frac{N(d_T + \max(\mu_1, \mu_2))}{L_N}} \right).$$

Moreover, for the autoregressive matrices in the dynamic equation (2.2) we have

$$\left\| \tilde{\Phi}_i - \Phi_i \right\|_F = \left( \sum_{t=\bar{p}}^T u_t \psi_t' \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t \psi_t' \right)^{-1} (\iota_i \otimes I_r) + o_p(1),$$

where  $\iota_i$  is the  $i$ -th column of the  $r \times r$  identity matrix.

The proof of Theorem 3.1 is given in Appendix A.1. Under the imposed regularity conditions Theorem 3.1 establishes the average consistency in the Frobenius norm of the first step estimators of the factor loadings matrix and idiosyncratic error covariance matrix based on our regularized FAVAR model. More specifically, we can see that  $\Lambda$  and  $\Phi$  are estimated consistently, even though we impose the strict factor model structure on  $\Sigma_e$  in the first step of our estimation procedure. Consequently, the latent factors  $f_t$  estimated based on GLS are consistent as long as  $1/2 \leq \beta \leq 1$ . Intuitively, the lower limit on  $\beta$  ensures that the factors are not too weak such that there is still a clear distinction between the common and idiosyncratic component. Furthermore, the second step estimator for the idiosyncratic error covariance matrix introduced in Section 2.3 is consistent under the spectral norm as long as  $S_N^2 d_T = o_p(1)$  and  $S_N \max(\mu_1, \mu_2) = o(1)$ . Finally, the autoregressive parameter matrices in the dynamic VAR equation in (2.2) are consistently estimated under the Frobenius norm as well.

The following theorem summarizes the consistency results for the complete set of estimators based on our regularized FAVAR model after rotating them with the matrix  $\tilde{A} = \begin{bmatrix} I_{r_1} & -\tilde{\Omega}_{fg}\tilde{\Omega}_{gg}^{-1} \\ 0 & I_{r_2} \end{bmatrix}$ .

**Theorem 3.2** (Consistency of the rotated estimators based on the regularized FAVAR model). *Under Assumptions 3.1 and 3.2 the regularized FAVAR model in (2.6) satisfies for  $T$  and  $N \rightarrow \infty$ , the following*

$$\begin{aligned} \max_{i \leq N} \left\| \hat{\lambda}_i^g - \lambda_i^{*g} \right\| &= \mathcal{O}_p \left( \sqrt{\mu_1} + \sqrt{\mu_2} + \sqrt{d_T} \right) \\ \left\| \hat{f}_t - f_t^* \right\| &= \mathcal{O}_p \left( N^{-\beta/2} \right) + \mathcal{O}_p \left( \frac{\mu_1 \sqrt{L_N}}{N^\beta} + \sqrt{\frac{N(d_T + \mu_2)}{N^{2\beta}}} \right), \\ \left\| \hat{\Phi}_i - \Phi_i^* \right\|_F &= \left( \sum_{t=\bar{p}}^T u_t \psi_t' \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t \psi_t' \right)^{-1} (\iota_i \otimes I_r) \\ &\quad + \mathcal{O}_p \left( \frac{\mu_1 \sqrt{L_N}}{N^\beta} + \sqrt{\frac{N(d_T + \mu_2)}{N^{2\beta}}} \right), \end{aligned}$$

where  $\iota_i$  is the  $i$ -th column of the  $r \times r$  identity matrix and  $1/2 \leq \beta \leq 1$ .

The proof of Theorem 3.2 is given in Appendix A.1. Theorem 3.2 shows that the estimators of the factor loadings of the observed factors, of the unobserved factors and of the coefficients of the dynamic equation in (2.2) are consistently estimated under the Euclidean and Frobenius norm after rotating them based on the matrix  $\tilde{A}$ . Note that a further analysis of the estimated factor loadings of the unobserved factors  $\tilde{\lambda}_i$  is redundant, as they are unaffected by the rotation.

## 4 Implementation of the regularized FAVAR estimator

### 4.1 Majorize-minimize EM algorithm

For the implementation of the regularized FAVAR model, we employ the majorize-minimize EM algorithm by [Bien and Tibshirani \(2011\)](#). The idea of this algorithm is to replace the optimization of the nonconvex objective function in equation (2.4) by a sequence of convex problems that are numerically easy to solve by algorithms for convex optimization. In that respect, we majorize the log-likelihood function in (2.4) by the tangent plane of the concave part

$\log \left| \det \left( \tilde{\Sigma} \right) \right|$ , which leads to the following expression

$$\mathcal{L}_m^* = \log \left| \det \left( \tilde{\Sigma}_m \right) \right| + \text{tr} \left[ \left( 2\tilde{\Sigma}_{H,m} \tilde{\Lambda}'_m \right) \tilde{\Sigma}_m^{-1} \left( \Lambda - \tilde{\Lambda}_m \right) \right] + \text{tr} \left[ S_x \Sigma^{-1} \right], \quad (4.1)$$

where  $\tilde{\Sigma}_m = \tilde{\Lambda}_m \tilde{\Sigma}_{H,m} \tilde{\Lambda}'_m + \hat{\Phi}_{e,m}$  and  $\Sigma = \Lambda \tilde{\Sigma}_{H,m} \Lambda' + \hat{\Phi}_{e,m}$  with  $\tilde{\Sigma}_{H,m} = \begin{bmatrix} I_{r_1} & 0 \\ 0 & \tilde{\Sigma}_{G,m} \end{bmatrix}$ .  $\tilde{F}$  denotes a initial estimate for the latent factors that can be obtained by unpenalized MLE and the subscript  $m$  is the  $m$ -th step in the iterative procedure. As we use standardized time series, the covariance matrix of the observed factors reduces to an identity matrix and  $\tilde{\Sigma}_{H,m} = I_{r_1+r_2}$ . In the following, we augment the majorized log-likelihood function in (4.1) by  $L_1$ -penalty terms for the factor loadings  $\Lambda^f$  and  $\Lambda^g$ , which leads to the following optimization problem for our regularized FAVAR model:

$$\begin{aligned} \min_{\{\Lambda\}} \log \left| \det \left( \tilde{\Sigma}_m \right) \right| + \text{tr} \left[ \left( 2\tilde{\Lambda}'_m \right) \tilde{\Sigma}_m^{-1} \left( \Lambda - \tilde{\Lambda}_m \right) \right] + \text{tr} \left[ S_x \Sigma^{-1} \right] \\ + \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} \left| \lambda_{ik}^f \right| + \mu_2 \sum_{i=1}^N \sum_{l=r_1+1}^r \left| \lambda_{il}^g \right|. \end{aligned} \quad (4.2)$$

As the optimization problem in equation (4.2) has the appealing property of being entirely convex it can be easily solved by a convex optimizer. In this respect, we rely on an efficient projected gradient descent algorithm proposed by [Bien and Tibshirani \(2011\)](#) and optimize the following minimization problem

$$\min_{\{\Lambda\}} \frac{1}{2c} \sum_{i=1}^N \sum_{j=1}^r \left( \lambda_{ij} - \tilde{\lambda}_{ij,m} + c \cdot \tilde{A}_{ij,m} \right)^2 + \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} \left| \lambda_{ik}^f \right| + \mu_2 \sum_{i=1}^N \sum_{l=r_1+1}^r \left| \lambda_{il}^g \right|, \quad (4.3)$$

where  $c$  determines the depth of the projection in the gradient decent algorithm<sup>10</sup> and

$$\tilde{A}_m = \left[ \tilde{\Sigma}_m^{-1} - \tilde{\Sigma}_m^{-1} S_x \tilde{\Sigma}_m^{-1} \right] \left( 2\tilde{\Lambda}'_m \right),$$

which corresponds to the first-derivative of  $\mathcal{L}_m^*$  with respect to each element of  $\Lambda$ . The minimization of the objective function (4.3) leads to the following first order conditions with respect

<sup>10</sup> In all our applications we set  $c = 0.01$ .

to each element of the quantities  $\Lambda^f$  and  $\Lambda^g$

$$\begin{aligned}\frac{\partial}{\partial \lambda_{ik}^f} &= \sum_{i=1}^N \sum_{k=1}^{r_1} \left( \lambda_{ik} - \tilde{\lambda}_{ik} + c \cdot \tilde{A}_{ik,m} \right) + c \cdot \mu_1 \sum_{k=1}^{r_1} \sum_{i=1}^N \nu_{ik}^f \stackrel{!}{=} 0, \\ \frac{\partial}{\partial \lambda_{il}^g} &= \sum_{i=1}^N \sum_{l=r_1+1}^r \left( \lambda_{il} - \tilde{\lambda}_{il} + c \cdot \tilde{A}_{il,m} \right) + c \cdot \mu_2 \sum_{l=r_1+1}^r \sum_{i=1}^N \nu_{il}^g \stackrel{!}{=} 0,\end{aligned}$$

where  $\nu_{ik}^f$  and  $\nu_{il}^g$  are the subgradients of  $|\lambda_{ik}^f|$  and  $|\lambda_{il}^g|$ , respectively. Hence, solving for a specific  $\tilde{\lambda}_{ik}^f$  or  $\tilde{\lambda}_{il}^g$  leads to the following updating formulas for the factor loadings estimates

$$\tilde{\lambda}_{ik,m+1}^f = \mathcal{S} \left( \tilde{\lambda}_{ik,m}^f - c \cdot \tilde{A}_{ik,m}, c \cdot \mu_1 \right), \quad \text{for } i = 1, \dots, N; k = 1, \dots, r_1 \quad (4.4)$$

$$\tilde{\lambda}_{il,m+1}^g = \mathcal{S} \left( \tilde{\lambda}_{il,m}^g - c \cdot \tilde{A}_{il,m}, c \cdot \mu_2 \right), \quad \text{for } i = 1, \dots, N; l = r_1 + 1, \dots, r, \quad (4.5)$$

where  $\mathcal{S}(\cdot)$  is the soft-thresholding function defined in (2.8).

To obtain an update for the estimate of the covariance matrix of the idiosyncratic error  $\Phi_e$ , we use the formula in the EM algorithm suggested by [Bai and Li \(2012\)](#)

$$\tilde{\Phi}_{e,m+1} = \text{diag} \left[ S_x - \tilde{\Lambda}_{m+1} \tilde{\Lambda}_m' \left( \tilde{\Lambda}_m \tilde{\Lambda}_m' + \tilde{\Phi}_{e,m} \right)^{-1} S_x \right].$$

Our iterative estimation procedure for the regularized FAVAR model is therefore described by the following steps:

Starting from the FAVAR model  $X = \Lambda^f F' + \Lambda^g G' + e$

*Step 1:* Transform the data by linearly projecting the observed factors  $G$  from the observed data. This is achieved by post-multiplying  $M = I_T - G(G'G)^{-1}G'$  and obtain the transformed data  $XM = \Lambda^f F'M + eM$  or  $\tilde{X} = \Lambda^f \tilde{F}' + \tilde{e}$ .

*Step 2:* Set  $m = 1$  and obtain an initial estimate for the factor loading matrix  $\tilde{\Lambda}_m^f$ , factors  $\tilde{F}_m$  and the diagonal idiosyncratic error covariance matrix  $\tilde{\Phi}_{e,m}$ , e.g. by using unpenalized MLE on  $\tilde{X}$ . Get an initial estimate of the factor loadings of the observed factor by  $\tilde{\Lambda}_m^g = \left( X - \tilde{\Lambda}_m^f \tilde{F}_m' \right) G (G'G)^{-1}$ .

*Step 3:* Update  $\tilde{\lambda}_{ik,m}^f$  by  $\tilde{\lambda}_{ik,m+1}^f = \mathcal{S} \left( \tilde{\lambda}_{ik,m}^f - c \cdot \tilde{A}_{ik,m}, c \cdot \mu_1 \right)$ , for  $i = 1, \dots, N; k = 1, \dots, r_1$  and  $\tilde{\lambda}_{il,m}^g$  by  $\tilde{\lambda}_{il,m+1}^g = \mathcal{S} \left( \tilde{\lambda}_{il,m}^g - c \cdot \tilde{A}_{il,m}, c \cdot \mu_2 \right)$ , for  $i = 1, \dots, N; l =$

$r_1 + 1, \dots, r$ .

*Step 4:* Update  $\tilde{\Phi}_e$  using the EM algorithm in [Bai and Li \(2012\)](#), according to

$$\tilde{\Phi}_{e,m+1} = \text{diag} \left[ S_x - \tilde{\Lambda}_{m+1} \tilde{\Lambda}'_m \left( \tilde{\Lambda}_m \tilde{\Lambda}'_m + \tilde{\Phi}_{e,m} \right)^{-1} S_x \right]$$

*Step 5:* If  $\left\| \tilde{\Lambda}_{m+1}^f - \tilde{\Lambda}_m^f \right\|$  and  $\left\| \tilde{\Phi}_{e,m+1} - \tilde{\Phi}_{e,m} \right\|$  are sufficiently small, stop the procedure, otherwise set  $m = m + 1$  and return to *Step 3*.

*Step 6:* Estimate the unobserved factors by  $\tilde{f}_t = \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f \right)^{-1} \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} x_t$ , where  $\tilde{\Lambda}^f$  and  $\tilde{\Phi}_e$  are the parameter estimates after convergence.

*Step 7:* Re-estimate the covariance matrix of the idiosyncratic errors based on the procedure introduced in [Section 2.3](#).

*Step 8:* Based on the estimated and observed factors,  $\tilde{h}_t = \left( \tilde{f}_t', g_t' \right)'$ , we estimate the following VAR regression:  $\tilde{h}_t = \Phi_1 \tilde{h}_{t-1} + \Phi_2 \tilde{h}_{t-2} + \dots + \Phi_p \tilde{h}_{t-p} + \eta_t$ , we obtain the residuals  $\tilde{\eta}_t$  from the previous regression and calculate  $\tilde{\Omega} = \frac{1}{T-p} \sum_{t=p+1}^T \tilde{\eta}_t \tilde{\eta}_t'$ .

*Step 9:* For the identification of our regularized FAVAR model, we use the identification restriction *IRa*. Thus, the factor loadings estimates for the unobserved factors are unchanged  $\hat{\Lambda}^f = \tilde{\Lambda}^f$ . Further, we estimate  $\hat{\Lambda}^g = \tilde{\Lambda}^g + \tilde{\Lambda}^f \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1}$ ,  $\hat{F} = \tilde{F} - G \tilde{\Omega}_{gg}^{-1} \tilde{\Omega}_{gf}$  and the autoregressive parameters by  $\hat{\Phi}_i = A \tilde{\Phi}_i A^{-1}$ , for  $i = 1, \dots, p$ , with the rotation matrix  $A = \begin{bmatrix} I_{r_1} & -\tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} \\ 0 & I_{r_2} \end{bmatrix}$ .

## 4.2 Selection of the regularization parameters

We provide a selection criterion to estimate the regularization parameters  $\mu_1$  and  $\mu_2$ . The information criterion is based on an adaptation of the Bayesian information criterion, comparable to the criteria in [Bai and Ng \(2002\)](#), and takes the following form

$$IC(\mu_1, \mu_2) = \mathcal{L} \left( \tilde{\Lambda}, S_{\tilde{H}}, \tilde{\Sigma}_e^T \right) + \kappa_{\mu_1, \mu_2} \sqrt{\frac{\log(2N)}{N} + \frac{\log N}{NT}}, \quad (4.6)$$

where  $\kappa_{\mu_1, \mu_2}$  denotes the number of non-zero elements in the factor loadings matrix  $\tilde{\Lambda}$  for specific  $\mu_1$  and  $\mu_2$ .  $\mathcal{L} \left( \tilde{\Lambda}, S_{\tilde{H}}, \tilde{\Sigma}_e^T \right)$  is the value of the log-likelihood function in [equation \(2.4\)](#) evaluated at the estimates of the factor loadings  $\tilde{\Lambda}$ , the sample covariance matrix of the factors  $S_{\tilde{H}}$  and the

covariance matrix of the idiosyncratic component  $\tilde{\Sigma}_e^\tau$ . The penalty function in (4.6) relates to the convergence rate of the estimator  $\tilde{\Lambda}$  and vanishes for  $N, T \rightarrow \infty$ . In order to select the optimal penalty parameters, we calculate the information criterion in (4.6) for a grid of different values for  $\mu_1$  and  $\mu_2$  and select the ones that minimize the criterion. The grids for the regularization parameters are set to  $\mu_1 = [0, 0.05, \mu_{1,\max}]$  and  $\mu_2 = [0, 0.1, \mu_{2,\max}]$ , where  $\mu_{1,\max}$  and  $\mu_{2,\max}$  denote the highest values of the penalty parameters such that we do not select an empty loadings matrix and the Assumptions 3.2 in Section 3 are still fulfilled.

## 5 Empirical illustration: The effects of monetary policy shocks

In this section we use our regularized FAVAR model to investigate the effects of a monetary policy shock on economic variables. We use the Federal Funds rate (FFR) as the monetary policy instrument. The monetary policy shock is defined as a shock to the innovation of the FFR. In our FAVAR setting, we treat the FFR as an observable factor.

### 5.1 Data description

The data set  $x_t$  consists of 126 macroeconomic and financial time series on monthly frequency. It spans the period from January 1985 until December 2016 which results in 384 monthly observations. The data set is mainly based on the FRED-MD database by McCracken and Ng (2016). To represent a broad range of economic sectors we augment the data set by manufacturing data, stock market data and various short-term interest rate spreads. The resulting data set is comparable to the one used by Forni and Gambetti (2010), Stock and Watson (2016) and Kersefischer (2019). A detailed overview can be found in Table B.1 in Appendix B. We transform the data such that  $x_t$  contains only stationary time series.<sup>11</sup> Furthermore, without loss of generality we standardize  $x_t$ .

### 5.2 Overview of models

This section provides an overview of the models that are compared in the empirical analysis. The regularized FAVAR (**RFVAR**) leads to a data driven identification of the unobserved factors. We obtain an economic structure by shrinking single elements in the factor loadings

<sup>11</sup> The specific transformations can be found in Table B.1 in Appendix B.

matrix to zero. Hence, the estimated factors load only on a subset of the observed time series that correspond to different sectors of the economy. For the identification of our model we use scheme *IRa* in Section 2.4 that relies on a statistical identification of the FAVAR model that is invariant to the ordering of the data. Hence, we circumvent imposing a restrictive a priori assumption on the structure of the model. We initialize our model with eight unobserved factors according to the  $IC_1$  selection criterion by Bai and Ng (2002) and are left with five unobserved factors after the regularization.

We set the lag order for our regularized FAVAR model to  $p = 12$ . In a dynamic factor model setting, this lag order captures the dynamics in the observed and estimated factors sufficiently. Concerning the choice of the regularization parameters we use the procedure described in Section 4.2.<sup>12</sup>

The commonly used scheme in the literature to obtain economically identified factors is based on the named factor identification. This corresponds to the *IRb* identification scheme as pointed out in Section 2.4, which implies that the first  $r_1$  time series drive the dynamics of the unobserved factors. In our analysis we denote this model as **FAVAR-NF** and the following time series are used as naming variables: civilian unemployment rate, consumer price index, industrial production index, S&P 500 composite index and Baa corporate bond yield (see Table B.2). This specific choice of time series is guided by the selection of our regularized FAVAR model. More specifically, these are the time series with the highest loadings in absolute value on the estimated factors. Moreover, the selection is also economically sensible as relevant sectors of the economy are represented by those time series. We use  $p = 12$  lags in the named factor FAVAR model.

In the following, we introduce two additional models that are restricted versions of the general dynamic factor model proposed by Forni, Hallin, Lippi, and Reichlin (2000). Closely related models are studied by Stock and Watson (2002b, 2002a, 2005). More specifically, the

---

<sup>12</sup> Different robustness checks in our empirical application reveal that the structure of the regularized FAVAR model does not change much for different values of  $\mu_2$ . Hence, to reduce the computational time, one could fix  $\mu_2$  to a value in the interval  $[0.7, 1.3]$  and optimize only for values of  $\mu_1$ .



DFM is given by

$$x_t = \Lambda^f f_t + e_t, \quad (5.1)$$

$$f_t^\dagger = \Theta(L)f_{t-1}^\dagger + u_t, \quad (5.2)$$

where  $f_t$  in equation (5.1) is a  $(q \times 1)$ -vector of latent static factors and  $\Theta(L)$  is a  $p$ -th order lag polynomial. The first model we consider is the pure dynamic factor model (**DFM**) by [Forni and Gambetti \(2010\)](#). In this model  $f_t^\dagger = f_t$ ,  $u_t = R\varepsilon_t$ ,  $R$  is a  $(q \times q_1)$ -dimensional matrix and  $\varepsilon_t$  is  $(q_1 \times 1)$ -vector of primitive shocks.

In the second model, we augment the factors in equation (5.2) by observable variables  $y_t$ . Hence, in this setting  $f_t^\dagger = [f_t', y_t']'$ . We denote this model as **VAR-F** in what follows. In comparison to the FAVAR model introduced in equations (2.1) and (2.2), the model specification in (5.1) only includes latent variables on the right hand side.

For the model specification of the DFM, we follow [Forni and Gambetti \(2010\)](#) by setting  $q = 16$ ,  $q_1 = 4$  and  $p = 12$ .<sup>13</sup> The identification of the structural shocks relies on a recursive Cholesky scheme, which includes industrial production (IP), consumer prices (CPI), the FFR and the excess bond premium (EBP). A tight monetary policy shock increases the FFR and has no contemporaneous effect on IP and CPI. The FFR can only be affected by industrial production, consumer prices and itself contemporaneously, whereas EBP reacts to shocks to the three others. The model settings for the VAR-F are comparable to [Keressenfischer \(2019\)](#) and are set to  $q_1 = 9$  and  $p = 12$ . As observable variables, we include IP, CPI and the FFR. For the identification of the structural shocks corresponding to the observable variables we use a recursive Cholesky scheme. As above, IP and CPI are not affected by a contractionary monetary policy shock on impact, whereas the FFR increases.

### 5.3 Results for the factor model analysis

To get some insights on the data, we start with an unregularized factor model estimated by PCA. The number of included factors is determined by the [Bai and Ng \(2002\)](#)  $IC_1$  criterion, where we set the maximum number of allowed factors ( $r_{\max}$ ) to ten.

Figure C.1 shows  $R^2$  of univariate regressions of the latent factors on the observed time

<sup>13</sup> The selection criterion by [Bai and Ng \(2007\)](#) also suggests 4 dynamic factors for our dataset.

series. Of the eight factors that are selected by the criterion, only five factors have blocks with high explanatory power. It is well documented in the literature that the information criteria of [Bai and Ng \(2002\)](#) tend to overestimate the true number of factors when there is remaining correlations in the idiosyncratic component (see e.g. [Ahn and Horenstein \(2013\)](#) and [Caner and Han \(2014\)](#)). Furthermore, the standard factor model is only identified statistically and the estimated factors may not be economically meaningful. For the structural analysis we are interested in the dynamics of a model that is economically interpretable. Hence, it is crucial to economically identify the unobserved factors in the FAVAR model.

In our regularized FAVAR model we enhance the economic interpretability by shrinking elements of the factor loadings matrix to zero. We illustrate the  $R^2$  results of univariate regressions associated with a sparse factor structure in [Figure C.2](#). The separation of the unobserved factors is notably more distinct compared to the factors estimated by PCA. More specifically, we obtain a block structure in the factor loadings which leads to factors that are linked to different sectors in the economy. Furthermore, the RFAVAR model estimates five latent factors. Hence, it shrinks the factors that do not add additional explanatory power to zero. The economic groups associated to the latent factors are: the labor market, prices, industrial production, the stock market and credit spreads. In our context, the labor market factor is mostly linked to employment time series. The corresponding time series plots are shown in [Figure C.3](#). The obtained factors are closely aligned with the underlying economic time series which is due to the fact that the informational content of various time series is used to construct the latent factor estimates. [Table B.1](#) provides an overview of which factor drives each variable. A comparable number of factors has been found by [Stock and Watson \(2016\)](#).

The same analysis is repeated for the named factor FAVAR (FAVAR-NF) model. The resulting  $R^2$  are plotted in [Figures C.4](#). The choice of the naming variables in the FAVAR-NF model leads to factors that have a comparable allocation to economic sectors as the RFAVAR model. This result is anticipated as the selection of the naming variables is guided by the RFAVAR. However, the obtained factor time series in [Figures C.5](#) are less aligned to the observed time series in comparison to the RFAVAR. This effect is more pronounced for naming variables that do not represent the entire sector. In these cases the resulting factors are distorted by sector unrelated variables, as in the labor market and credit spread factors.

It is important to note that the model identification crucially depends on the ordering of the observed time series. For example, [Stock and Watson \(2016\)](#) use the following naming scheme in an oil price application: Real personal consumption expenditures, industrial production index, civilian employment rate, S&P 500 composite index, trade weighted US dollar index major currencies and producer price index.<sup>14</sup> The resulting  $R^2$  are plotted in [Figures C.6](#). Even though the naming variables are chosen based on economic reasoning the selection is rather arbitrary and leads to very different results. Hence, the selection of the naming variables constitutes a restrictive assumption on the structural model.

We omit this analysis for the DFM and VAR-F models because these frameworks are not concerned with the interpretability of the factors.

#### 5.4 Results from the impulse response analysis

In the following, we analyze the effects of a monetary policy shock on the observed time series based on our model specification (RFAVAR). Additionally, we compare the results to the ones obtained for the FAVAR-NF, DFM and VAR-F. Hence, our focus lies on the dynamic responses to a shock in the monetary policy instrument (FFR). This corresponds to a partial identification of the structural model, where we identify the monetary policy shock. More specifically, the innovations corresponding to the FFR are contemporaneously uncorrelated with the latent factor innovations, whereas the latter can be contemporaneously correlated. Our RFAVAR framework allows for the structural analysis of the dynamics for both the factors and the underlying observed time series as pointed out in [Section 2.5](#).

In a first step, we elaborate on the effect on the estimated factors for the RFAVAR and the FAVAR-NF model. In both settings, the obtained factors are economically identified and serve as proxies of economic aggregates. The contemporaneous structural effect on the factors are depicted in [Table B.3](#), where the last column in both panels is associated with the monetary policy shock.

For our RFAVAR model, the strength of the impact is given in Panel A of [Table B.3](#). The sign of the contemporaneous impact is in line with economic reasoning. More specifically, the price, industrial production and the labor market factors react negatively on impact in response

---

<sup>14</sup> As [Stock and Watson \(2016\)](#) work on quarterly frequency they can use data that is not available on monthly frequency. We deviate by using a different proxy for employment and by omitting government spending.

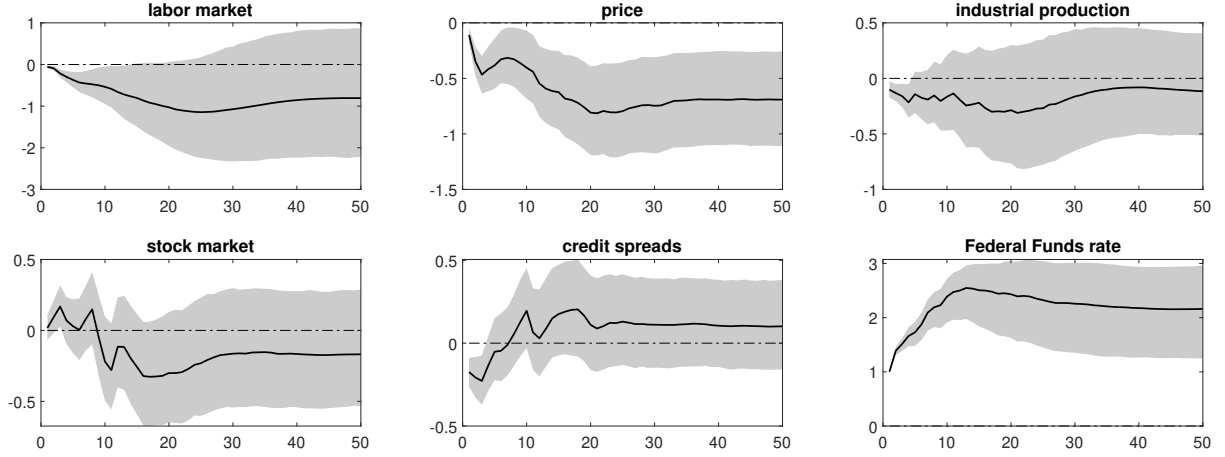


Figure 1: Accumulated impulse responses to a monetary policy shock on the factors for the regularized FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals.

to a tight monetary policy shock.

The impulse responses of the factors to a 100bp shock in the innovation of the Federal Funds rate are illustrated in Figure 1. The responses of all factors are transitory and stabilize to a new level after one or two years. More precisely, the labor market reacts significantly negative on impact and no longer reacts after 12 months. Moreover, as expected by economic rationale, the price level, the level of industrial production and the credit spread factor are impacted negatively by an exogenous increase in the monetary policy rate. The stock market factor does not react significantly to a monetary policy shock for most of the periods. It slightly increases after five months for about one month.

For the FAVAR-NF model, the impact matrix is reported in Panel B in Table B.3. Note that we only provide standard errors for the estimates in the last column of Panel B. This follows from the structure of the rotation matrix  $\tilde{A}$  given at the end of Section 2.4. As the first  $r_1 \times r_1$  block of  $\tilde{A}$  depends on  $\tilde{\Lambda}_1$ , which is kept fixed in the residual-bootstrap based on the dynamic factor equation (2.2), we cannot compute bootstrap-based standard errors for the estimated quantities in the first  $r_1 \times r_1$  block of the impact matrix. However, this is not harmful for the upcoming analysis, as we are interested in the dynamic effects of a monetary policy shock, whose contemporaneous impacts are given by the last column of the impact matrix.

Furthermore, in Figure C.7, we plot the impulse response functions associated with the named factor scheme. Both, the contemporaneous impacts and the resulting impulse responses on the

estimated factors are not always in line with the economic theory. For example, unemployment reacts negatively in response to a tight monetary policy shock, whereas industrial production increases for about ten months. Moreover, we obtain relatively large confidence intervals for the remaining factors leading to statistically insignificant impulse responses. Even though the choice of the naming variables is economically motivated, the implied structural dynamics fail to span the monetary policy shock.

In a second step, we investigate the impulse responses on the observable variables  $x_t$  for all models. For the RFAVAR and FAVAR-NF models the contemporaneous impact matrix on  $x_t$  is given by  $\hat{\Lambda}$  corresponding to the rotated factor loadings estimate, as outlined in Section 2.5. The impact matrix for our regularized FAVAR model is given in Figure 2. The contemporaneous impact of a tightening monetary policy shock is in accordance with economic reasoning. Following an exogenous increase in the FFR, the short term interest rates go up on impact. Further, price and IP as well as employment time series react negatively, whereas unemployment series rise. The stock market plummets and credit conditions deteriorate.

The contemporaneous impact matrices for the FAVAR-NF model is depicted in Figure C.8. The economic implications are implausible, as the sign of the effects is often incorrect. For example, the labor market, industrial production and prices react positively on impact for a contractionary monetary policy shock. In contrast to our RFAVAR model, the factor loadings matrix of the FAVAR-NF model is more dense and an economic sector association is not possible. Hence, even though the  $R^2$  plot of univariate regressions of the named factors on the observed time series in Figure C.6 shows a high explanatory power of the latent factors, the implied rotation scheme yields a factor loadings matrix which we can no longer explain economically.

In the following, we analyze the impulse responses for selected time series, i.e. we concentrate on the effects of a tight monetary policy shock on the consumer price index (CPI), the civilian unemployment rate, the IP index and the three months Treasury bill. Figure 3 shows the level effects for our RFAVAR model on the specific variables. CPI and the IP index series react negatively for around two years, respectively and stabilize to a new level for the remaining periods. The unemployment rate increases for approximately 24 months, whereas the short-term interest rate reacts positively on impact for 12 months.

The impulse responses for the FAVAR-NF are provided in Figure C.9. Overall, the results

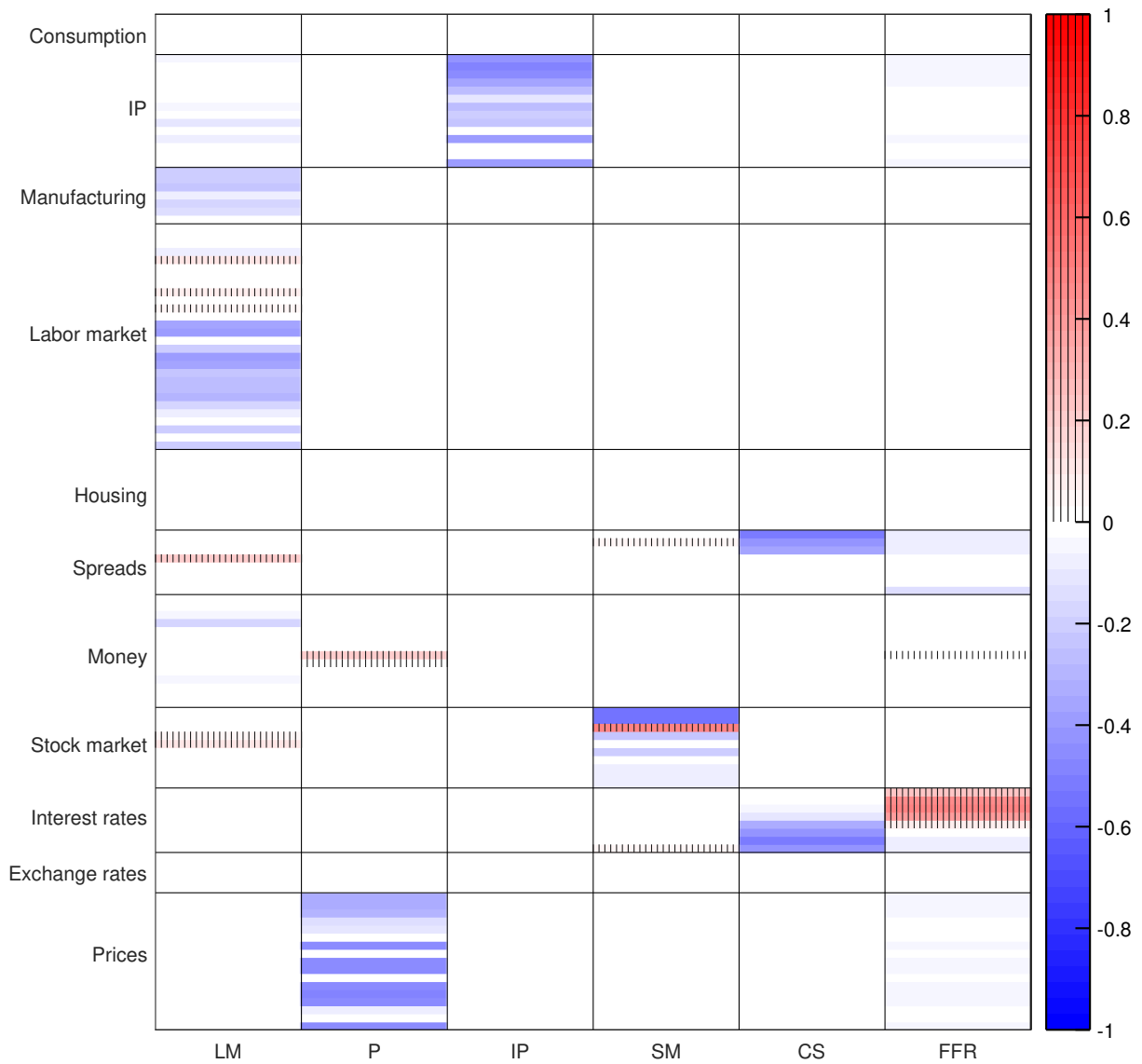


Figure 2: Impact matrix on the observed variables  $x_t$  for the regularized FAVAR  
 This graph shows the contemporaneous impact matrix of a 100bp shock in the factor innovations to the observed time series for the regularized FAVAR model. The factors are abbreviated as follows: 'LM' labor market, 'P' price, 'IP' industrial production, 'SM' stock market, 'CS' credit spread and 'FFR' Federal Funds rate.

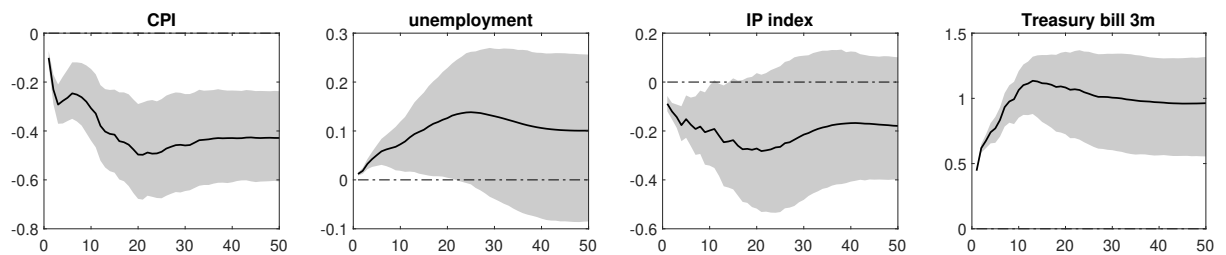


Figure 3: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the regularized FAVAR  
 The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals.

show contradicting effects contemporaneously as well as over time for the unemployment rate and the IP index.

For the DFM, the impulse responses for CPI, unemployment, industrial production and the three month Treasury bill rate are illustrated in Figure C.10. The point estimates for CPI, the IP index and unemployment are economically not reasonable. Moreover, the DFM is sensitive to the number of included latent factors and lags, which leads to a high degree of estimation noise.

The accumulated impulse responses for the VAR-F model are given in Figure C.11. The responses of unemployment and IP are in line with economic rationale. Moreover, the impulse response of CPI is very volatile and statistically insignificant. The three month Treasury bill rate reacts negatively over all horizons. The results are very sensitive to the number of included factors and lags. If we do not include enough factors or lags, the impulse response patterns are hardly interpretable (e.g. a price puzzle is obtained). Similar results are obtained by [Kerssenfischer \(2019\)](#).

## 5.5 Robustness checks

We evaluate the robustness of our findings in various ways: First, we use the shadow rate provided by [Wu and Xia \(2016\)](#) instead of the Federal Funds rate when the policy rate is at the zero lower bound between December 16, 2008 and December 15, 2015.<sup>15</sup> Figure C.12 shows both, the shadow rate and the Federal Funds rate, in one graph. The replacement of the policy rate does not affect the results qualitatively for our regularized FAVAR model. In Figure C.13 we plot the impulse responses of the observed variables to a shock in the innovation of the shadow rate. The impulse response patterns are very similar to those in Figure 3. For the named factor scheme, the positive impact on the IP index is not significant in Figure C.14 and the other three impulse responses are not substantially different. The impulse response patterns of the DFM model in Figure C.15 are similar to those obtained in the previous section. Finally, for the VAR-F model in Figure C.16, the impulse responses of CPI, the unemployment rate and the IP index remain economically plausible when we use the shadow rate instead of the FFR. More precisely, CPI reacts significantly negative for the first 6 months, whereas the IP index decreases significantly for approximately three years. Overall, using the shadow rate in periods where the

---

<sup>15</sup> The shadow rate is retrieved from Jing Cynthia Wu's website <https://sites.google.com/view/jingcynthiawu/shadow-rates>.

FFR is at the zero lower bound improves the results of our competing models slightly but does not affect our RFAVAR model qualitatively.

Second, we augment the maximum number of latent factors used in the initial steps to  $r_{\max} = 20$ . Figure C.17 contains the  $R^2$  of univariate regressions of the principal components factors on the observed time series. Qualitatively, it is similar to the one obtained for  $r^* = 10$  factors depicted in Figure C.1. In both settings, there are five factors which have separate block-wise explanatory power. Once we employ shrinkage onto the factor loadings, we obtain five latent factors for both initial number of factors. Our method is robust against alternative initial number of factors. This can be seen in Figures C.2 and C.18 which are based on different starting points for the shrinkage but yield similar latent factors and factor loadings estimates.

Third, we alter the lag order of the regularized FAVAR model to allow for different dynamics in the impulse responses. The resulting response functions of the factors and observed time series of a RFAVAR(2) model can be found in Figures C.19 and C.20. Moreover, the results for a RFAVAR(3) and RFAVAR(6) model are depicted in Figures C.21, C.22 and C.23, C.24, respectively. The shape of the impulse responses and the evolution over the horizons does not change qualitatively. However, for the first six to twelve months an increasing lag order leads to impulse response functions that react less smoothly to structural shocks.

Forth, we shorten the time span to the period prior to the global financial crisis. Hence, the data set spans a period for January 1985 to December 2006. We set the initial number of factors to  $r_{\max} = 10$ . Based on the Bai and Ng (2002)  $IC_1$  criterion, we initialize the factor model with  $r^* = 9$  factors. The  $R^2$  of univariate regressions of the PCA based factors onto the observed time series can be found in Figure C.25. For the pre-crisis period, the block-wise dependence is pronounced for four or five factors. When we impose sparsity onto the factor loadings, we retrieve four latent factors: prices, credit spreads, real activity, and the stock market. Figure C.26 shows the  $R^2$  of univariate regressions of the regularized factors on the observed time series and Figure C.27 shows the time series of the latent factors. A block-structure is still present, however, the real activity sector is represented by the industrial production sector. The labor market sector is no longer a separate latent factor in the pre-crisis period because the two real activity proxies covary heavily. For the entire sample, the labor market factor reacts more sluggish in comparison to industrial production which leads to two separate real activity factors



(i.e. Figure C.3). Our regularized factor model yields economically sensible latent factors and factor loadings estimates in the pre-crisis period. Figure C.28 contains the contemporaneous effects of a 100bp structural shock to the FFR onto the observed time series. The direction of the impact is economically plausible for all cases. Lastly, Figure C.29 shows impulse responses of four observed time series: the consumer price index and the industrial production index react negatively in response to a tight monetary policy shock, the level of employment in the manufacturing sector plummets, whereas the three months Treasury bill increases.

## 6 Conclusions

In this paper, we propose a regularized factor-augmented vector autoregressive model that simplifies the factor identification and their economic interpretation. Our estimation procedure relies on a penalized quasi-maximum likelihood approach and is based on a  $L_1$ -norm regularization of the factor loadings matrix. The named factor identification scheme conventionally used in the FAVAR literature to identify the factors, imposes specific relations between the factors and observed time series. More specifically, the ordering of the variables determines the structure of the model. The sparse factor loadings structure in our regularized FAVAR model allows for a direct factor identification. Hence, we are able to identify the estimated factors in a data-driven manner without implicitly assuming their form a priori in our identification scheme. In this framework, the effects of structural shocks can be investigated on economically meaningful estimated factors and on all observed time series included in the model.

We prove the average consistency under the Frobenius norm for the estimators of the factor loadings and the covariance matrix of the idiosyncratic component based on the regularized FAVAR model. The factors estimated based on GLS are shown to be consistent. Moreover, the autoregressive parameters in the dynamic equation are consistently estimated as well.

In an empirical application, we investigate the effects of structural monetary policy shocks on a broad range of economically relevant variables. We choose to identify this shock using a joint identification of the factor model and the structural innovations in the vector autoregressive model. We extract five latent factors that relate to the labor market, prices, industrial production, the stock market and credit spreads. Furthermore, the Federal Funds rate is used as an observed factor. We find impulse response functions which are in line with economic rationale both on the

factor aggregates and the observed time series level. More specifically, we do not observe a price puzzle and the obtained impulse response patterns are economically plausible. In particular, following a tight monetary policy, industrial production falls, credit conditions deteriorate and the level of employment decreases.

## References

- AHN, S. C. AND A. R. HORENSTEIN (2013): “Eigenvalue ratio test for the number of factors,” *Econometrica*, 81, 1203–1227.
- ALESSI, L., M. BARIGOZZI, AND M. CAPASSO (2011): “Non-Fundamentalness in Structural Econometric Models: A Review,” *International Statistical Review*, 79, 16–47.
- BAI, J. AND K. LI (2012): “Statistical analysis of factor models of high dimension,” *The Annals of Statistics*, 40, 436–465.
- (2016): “Maximum likelihood estimation and inference for approximate factor models of high dimension,” *Review of Economics and Statistics*, 98, 298–309.
- BAI, J., K. LI, AND L. LU (2016): “Estimation and inference of FAVAR models,” *Journal of Business & Economic Statistics*, 34, 620–641.
- BAI, J. AND Y. LIAO (2016): “Efficient estimation of approximate factor models via penalized maximum likelihood,” *Journal of Econometrics*, 191, 1–18.
- BAI, J. AND S. NG (2002): “Determining the number of factors in approximate factor models,” *Econometrica*, 70, 191–221.
- (2006): “Confidence intervals for diffusion index forecasts and inference for factor-Augmented regressions,” *Econometrica*, 74, 1133–1150.
- (2007): “Determining the Number of Primitive Shocks in Factor Models,” *Journal of Business & Economic Statistics*, 25, 52–60.
- (2013): “Principal components estimation and identification of static factors,” *Journal of Econometrics*, 176, 18–29.
- BAI, J. AND P. WANG (2014): “Identification theory for high dimensional static and dynamic factor models,” *Journal of Econometrics*, 178, 794–804.
- BERNANKE, B. S., J. BOIVIN, AND P. ELIASZ (2005): “Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach,” *The Quarterly Journal of Economics*, 120, 387–422.

- BIEN, J. AND R. J. TIBSHIRANI (2011): “Sparse estimation of a covariance matrix,” *Biometrika*, 98, 807–820.
- BOIVIN, J. AND M. P. GIANNONI (2007): “Global Forces and Monetary Policy Effectiveness,” in *International Dimensions of Monetary Policy*, University of Chicago Press, 429–478.
- BOIVIN, J., M. P. GIANNONI, AND I. MIHOV (2009): “Sticky prices and monetary policy: Evidence from disaggregated US data,” *American Economic Review*, 99, 350–384.
- CANER, M. AND X. HAN (2014): “Selecting the correct number of factors in approximate factor models: The large panel case with group bridge estimators,” *Journal of Business & Economic Statistics*, 32, 359–374.
- CHAMBERLAIN, G. AND M. ROTHSCILD (1983): “Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets,” *Econometrica*, 51, 1281–1304.
- DANIELE, M., W. POHLMEIER, AND A. ZAGIDULLINA (2018): “Sparse Approximate Factor Estimation for High-Dimensional Covariance Matrices,” Working paper, University of Konstanz.
- DEL NEGRO, M. AND C. OTROK (2007): “99 Luftballons: Monetary policy and the house price boom across US states,” *Journal of Monetary Economics*, 54, 1962–1985.
- FAN, J., Y. LIAO, AND M. MINCHEVA (2011): “High Dimensional Covariance Matrix Estimation in Approximate Factor Models.” *Annals of Statistics*, 39, 3320–3356.
- (2013): “Large covariance estimation by thresholding principal orthogonal complements,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75, 603–680.
- FORNI, M. AND L. GAMBETTI (2010): “The dynamic effects of monetary policy: A structural factor model approach,” *Journal of Monetary Economics*, 57, 203–216.
- FORNI, M., M. HALLIN, M. LIPPI, AND L. REICHLIN (2000): “The generalized dynamic-factor model: Identification and estimation,” *Review of Economics and Statistics*, 82, 540–554.
- GONÇALVES, S. AND B. PERRON (2014): “Bootstrapping factor-augmented regression models,” *Journal of Econometrics*, 182, 156–173.
- HORN, R. A. AND C. R. JOHNSON (2012): *Matrix analysis*, Cambridge university press.

- KERSSENFISCHER, M. (2019): “The puzzling effects of monetary policy in VARs: Invalid identification or missing information?” *Journal of Applied Econometrics*, 34, 18–25.
- KILIAN, L. AND L. T. LEWIS (2011): “Does the Fed respond to oil price shocks?” *The Economic Journal*, 121, 1047–1072.
- LAWLEY, D. AND A. MAXWELL (1971): *Factor Analysis as a Statistical Method*, second ed., Butterworths, London.
- MCCRACKEN, M. W. AND S. NG (2016): “FRED-MD: A monthly database for macroeconomic research,” *Journal of Business & Economic Statistics*, 34, 574–589.
- MUMTAZ, H. AND P. SURICO (2009): “The transmission of international shocks: a factor-augmented VAR approach,” *Journal of Money, Credit and Banking*, 41, 71–100.
- RAMEY, V. A. (2016): “Macroeconomic shocks and their propagation,” in *Handbook of Macroeconomics*, Elsevier, vol. 2, 71–162.
- SIMS, C. A. (1980): “Macroeconomics and reality,” *Econometrica*, 48, 1–48.
- (1992): “Interpreting the macroeconomic time series facts: The effects of monetary policy,” *European Economic Review*, 36, 975–1000.
- STOCK, J. H. AND M. W. WATSON (2002a): “Forecasting using principal components from a large number of predictors,” *Journal of the American Statistical Association*, 97, 1167–1179.
- (2002b): “Macroeconomic forecasting using diffusion indexes,” *Journal of Business & Economic Statistics*, 20, 147–162.
- (2005): “Implications of Dynamic Factor Models for VAR Analysis,” Tech. Rep. w11467, National Bureau of Economic Research.
- (2016): “Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics,” in *Handbook of Macroeconomics*, Elsevier, vol. 2, 415–525.
- WU, J. C. AND F. D. XIA (2016): “Measuring the macroeconomic impact of monetary policy at the zero lower bound,” *Journal of Money, Credit and Banking*, 48, 253–291.

## Appendix

### A Proofs

#### A.1 Consistency of the regularized FAVAR Model Estimator

**Proof.** Theorem 3.1 (Consistency of the first step estimators of the regularized FAVAR model)

To establish the consistency of the regularized FAVAR process, we proceed in a similar fashion as in Daniele et al. (2018). Initially, we define the following penalized log-likelihood

$$\mathcal{L}_p(\Lambda, \Sigma_e) = Q_1(\Lambda, \Sigma_e) + Q_2(\Lambda, \Sigma_e) + Q_3(\Lambda, \Sigma_e), \quad (\text{A.1})$$

where

$$\begin{aligned} Q_1(\Lambda, \Sigma_e) &= \frac{1}{N} \log |\Sigma_e| + \frac{1}{N} \text{tr} \left( S_e \Sigma_e^{-1} \right) - \frac{1}{N} \log |\Sigma_e| - \frac{1}{N} \text{tr} \left( S_e \Sigma_e^{-1} \right) \\ &\quad + \frac{1}{N} \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} \left| \lambda_{ik}^f \right| - \left| \lambda_{ik}^f \right| + \frac{1}{N} \mu_2 \sum_{i=1}^N \sum_{l=r_1+1}^r \left| \lambda_{il}^g \right| - \left| \lambda_{il}^g \right|, \\ Q_2(\Lambda, \Sigma_e) &= \frac{1}{N} \text{tr} \left[ (\Lambda - \Lambda)' \Sigma_e^{-1} (\Lambda - \Lambda) - (\Lambda - \Lambda)' \Sigma_e^{-1} \Lambda \left( \Lambda' \Sigma_e^{-1} \Lambda \right)^{-1} \Lambda' \Sigma_e^{-1} (\Lambda - \Lambda) \right], \\ Q_3(\Lambda, \Sigma_e) &= \frac{1}{N} \log |\Lambda \Lambda' + \Sigma_e| + \frac{1}{N} \text{tr} \left( S_x (\Lambda \Lambda' + \Sigma_e)^{-1} \right) - Q_2(\Lambda, \Sigma_e) \\ &\quad - \frac{1}{N} \log |\Sigma_e| - \frac{1}{N} \text{tr} \left( S_e \Sigma_e^{-1} \right), \end{aligned}$$

and  $\Lambda = \begin{bmatrix} \Lambda^f & \Lambda^g \end{bmatrix}$ . Hence, the penalized log-likelihood in (A.1) can be written as

$$\begin{aligned} \mathcal{L}_p(\Lambda, \Sigma_e) &= \frac{1}{N} \log |\Lambda \Lambda' + \Sigma_e| + \frac{1}{N} \text{tr} \left( S_x (\Lambda \Lambda' + \Sigma_e)^{-1} \right) \\ &\quad - \frac{1}{N} \log |\Sigma_e| - \frac{1}{N} \text{tr} \left( S_e \Sigma_e^{-1} \right) \\ &\quad + \frac{1}{N} \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} \left| \lambda_{ik}^f \right| - \left| \lambda_{ik}^f \right| + \frac{1}{N} \mu_2 \sum_{i=1}^N \sum_{l=r_1+1}^r \left| \lambda_{il}^g \right| - \left| \lambda_{il}^g \right|. \end{aligned} \quad (\text{A.2})$$

Consider the following set,

$$\Psi_\delta = \left\{ (\Lambda, \Sigma_e) : \delta^{-1} < \pi_{\min} \left( \frac{\Lambda' \Lambda}{N^\beta} \right) \leq \pi_{\max} \left( \frac{\Lambda' \Lambda}{N^\beta} \right) < \delta, \right. \\ \left. \delta^{-1} < \pi_{\min}(\Sigma_e) \leq \pi_{\max}(\Sigma_e) < \delta \right\}, \quad \text{for } 1/2 \leq \beta \leq 1.$$

Further, we define  $\Phi_e = \text{diag}(\Sigma_e)$ , which corresponds to a covariance matrix that contains only the diagonal elements of  $\Sigma_e$  on its main diagonal. To control the sparsity in both  $\Lambda$  and  $\Sigma_e$ , we impose the following sparsity assumptions:

$$L_N = \sum_{k=1}^r \sum_{i=1}^N \mathbb{1} \{ \lambda_{ik} \neq 0 \} = \mathcal{O}(N), \\ S_N = \max_{i \leq N} \sum_{j=1}^N \mathbb{1} \{ \sigma_{e,ij} \neq 0 \},$$

where  $\mathbb{1} \{ \cdot \}$  denotes the indicator function that is equal to one if the boolean argument in braces is true. Hence,  $L_N$  is the number of non-zero elements in the factor loadings matrix  $\Lambda$  and  $S_N$  denotes the maximum number of non-zero elements in each row of  $\Sigma_e$ .

The following lemma will be useful for the forthcoming derivations.

**Lemma A.1.**

$$(i) \max_{i,j \leq N} \left| \frac{1}{T} \sum_{t=1}^T u_{it} u_{jt} - \mathbb{E} [u_{it} u_{jt}] \right| = \mathcal{O}_p \left( \sqrt{(\log N)/T} \right), \\ (ii) \max_{i \leq r, j \leq N} \left| \frac{1}{T} \sum_{t=1}^T f_{it} u_{jt} \right| = \mathcal{O}_p \left( \sqrt{(\log N)/T} \right).$$

**Proof.** See *Lemmas A.3 and B.1* in [Fan, Liao, and Mincheva \(2011\)](#). □

Under Assumptions [3.1](#) and [3.2](#) and *Lemma S.1.2.* in [Daniele et al. \(2018\)](#), we have that

$$\sup_{(\Lambda, \Sigma_e) \in \Psi_\delta} |Q_3(\Lambda, \Sigma_e)| = \mathcal{O}_p \left( \frac{\log N^\beta}{N} + \frac{1}{N^\beta} \frac{\log N}{T} \right).$$

Furthermore, by *Lemma S.1.3.* in [Daniele et al. \(2018\)](#), we obtain

$$Q_1(\tilde{\Lambda}, \tilde{\Sigma}_e) + Q_2(\tilde{\Lambda}, \tilde{\Sigma}_e) \leq d_T, \tag{A.3}$$

where  $d_T = \frac{\log N^\beta}{N} + \frac{1}{N^\beta} \frac{\log N}{T}$ .

In the following, we establish the consistency results for the diagonal idiosyncratic error covariance matrix estimator  $\tilde{\Phi}_e$  and the factor loadings estimator  $\tilde{\Lambda}$ .

**Lemma A.2.**

$$\frac{1}{N} \left\| \tilde{\Phi}_e - \Phi_e \right\|_F^2 = \mathcal{O}_p \left( \frac{\log N}{T} + d_T \right) = o_p(1).$$

**Proof.** By the definition of  $Q_1(\tilde{\Lambda}, \tilde{\Sigma}_e)$  and  $Q_2(\tilde{\Lambda}, \tilde{\Sigma}_e)$  and equation (A.3) we define

$$B_1 + B_2 \leq d_T, \tag{A.4}$$

where

$$\begin{aligned} B_1 &= \frac{1}{N} \log |\tilde{\Sigma}_e| + \frac{1}{N} \text{tr} \left( S_e \tilde{\Sigma}_e^{-1} \right) - \frac{1}{N} \log |\Sigma_e| - \frac{1}{N} \text{tr} \left( S_e \Sigma_e^{-1} \right), \\ B_2 &= \frac{1}{N} \text{tr} \left[ \left( \tilde{\Lambda} - \Lambda \right)' \tilde{\Sigma}_e^{-1} \left( \tilde{\Lambda} - \Lambda \right) - \left( \tilde{\Lambda} - \Lambda \right)' \tilde{\Sigma}_e^{-1} \tilde{\Lambda} \left( \tilde{\Lambda}' \tilde{\Sigma}_e^{-1} \tilde{\Lambda} \right)^{-1} \tilde{\Lambda}' \tilde{\Sigma}_e^{-1} \left( \tilde{\Lambda} - \Lambda \right) \right] \\ &\quad + \frac{1}{N} \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} \left| \tilde{\lambda}_{ik}^f \right| - \left| \lambda_{ik}^f \right| + \frac{1}{N} \mu_2 \sum_{i=1}^N \sum_{l=r_1+1}^r \left| \tilde{\lambda}_{il}^g \right| - \left| \lambda_{il}^g \right|. \end{aligned}$$

The result follows by the same argument as in *Lemma S.1.4.* in [Daniele et al. \(2018\)](#). □

To establish the consistency of the factor loadings matrix  $\tilde{\Lambda}$ , we analyze both sets of factor loadings corresponding to the latent and observed factors separately. Initially, we lower bound the first term in  $B_2$  as in *Lemma S.1.5.* in [Daniele et al. \(2018\)](#) and obtain

$$\begin{aligned} &\frac{1}{N} \text{tr} \left[ \left( \tilde{\Lambda} - \Lambda \right)' \tilde{\Sigma}_e^{-1} \left( \tilde{\Lambda} - \Lambda \right) - \left( \tilde{\Lambda} - \Lambda \right)' \tilde{\Sigma}_e^{-1} \tilde{\Lambda} \left( \tilde{\Lambda}' \tilde{\Sigma}_e^{-1} \tilde{\Lambda} \right)^{-1} \tilde{\Lambda}' \tilde{\Sigma}_e^{-1} \left( \tilde{\Lambda} - \Lambda \right) \right] \\ &\geq \mathcal{O}_p \left( \frac{L_N}{N} \right) \max_{i \leq N} \left\| \tilde{\lambda}_i - \lambda_i \right\|^2. \end{aligned} \tag{A.5}$$

Furthermore, the consistency results for  $\tilde{\Lambda}^f$  and  $\tilde{\Lambda}^g$  are summarized in the following lemma.



**Lemma A.3.**

$$\max_{i \leq N} \|\tilde{\lambda}_i - \lambda_i\| = \mathcal{O}_p \left( \max(\mu_1, \mu_2) + \sqrt{\frac{N(d_T + \max(\mu_1, \mu_2))}{L_N}} \right),$$

where  $d_T = \frac{\log N^\beta}{N} + \frac{1}{N^\beta} \frac{\log N}{T}$ .

**Proof.** By the definition of  $B_2$ , and equations (A.4) and (A.5) we have

$$\mathcal{O}_p \left( \frac{L_N}{N} \right) \max_{i \leq N} \|\tilde{\lambda}_i - \lambda_i\|^2 + \frac{1}{N} \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} |\tilde{\lambda}_{ik}^f| - |\lambda_{ik}^f| + \frac{1}{N} \mu_2 \sum_{i=1}^N \sum_{l=r_1+1}^r |\tilde{\lambda}_{il}^g| - |\lambda_{il}^g| \leq d_T. \quad (\text{A.6})$$

We start by analyzing  $\tilde{\Lambda}^f$ . The left hand side of (A.6) can be further lower bounded by

$$\begin{aligned} & \mathcal{O}_p \left( \frac{L_N}{N} \right) \max_{i \leq N} \|\tilde{\lambda}_i^f - \lambda_i^f\|^2 - \frac{1}{N} \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} |\lambda_{ik}^f| - |\tilde{\lambda}_{ik}^f| \leq d_T + \mu_2 r_2 \\ & \mathcal{O}_p \left( \frac{L_N}{N} \right) \max_{i \leq N} \|\tilde{\lambda}_i^f - \lambda_i^f\|^2 - \frac{1}{N} \mu_1 \sum_{i=1}^N \sum_{k=1}^{r_1} |\tilde{\lambda}_{ik}^f - \lambda_{ik}^f| \leq d_T + \mu_2 r_2 \\ & \mathcal{O}_p \left( \frac{L_N}{N} \right) \max_{i \leq N} \|\tilde{\lambda}_i^f - \lambda_i^f\|^2 - \mathcal{O} \left( \frac{L_N}{N} \right) \mu_1 \sqrt{r_1} \sqrt{\max_{i \leq N} \|\tilde{\lambda}_i^f - \lambda_i^f\|^2} \leq d_T + \mu_2 r_2 \end{aligned}$$

Solving for  $\max_{i \leq N} \|\tilde{\lambda}_i^f - \lambda_i^f\|$  yields

$$\begin{aligned} \max_{i \leq N} \|\tilde{\lambda}_i^f - \lambda_i^f\| & \leq \mu_1 + \sqrt{\mu_1^2 + \mathcal{O}_p \left( \frac{N(d_T + \mu_2)}{L_N} \right)} \\ & \leq \mathcal{O}_p \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right). \end{aligned}$$

Equivalently, by the same argument as above for  $\tilde{\Lambda}^g$ , we obtain

$$\max_{i \leq N} \|\tilde{\lambda}_i^g - \lambda_i^g\| = \mathcal{O}_p \left( \mu_2 + \sqrt{\frac{N(d_T + \mu_1)}{L_N}} \right),$$

which completes the proof. □

The consistency of the latent factor estimator is established in the following lemma.

**Lemma A.4.**

$$\frac{1}{T} \sum_{t=1}^T \left\| \tilde{f}_t - f_t \right\|^2 = o_p(1).$$

**Proof.** The latent factor estimator in equation (2.7) yields

$$\tilde{f}_t - f_t = - \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f \right)^{-1} \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \left( \tilde{\Lambda}^f - \Lambda^f \right) f_t + \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f \right)^{-1} \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} e_t. \quad (\text{A.7})$$

As  $L_N = \mathcal{O}(N)$ , the first term on the right-hand side is upper bounded by

$$\begin{aligned} & \mathcal{O}_p \left( N^{-\beta} \right) \sqrt{\sum_{i=1}^N \left\| \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \right)_i \left( \tilde{\lambda}_i^f - \lambda_i^f \right) \right\|^2} \|f_t\| \\ & \leq \mathcal{O}_p \left( N^{-\beta} \right) \sqrt{\mathcal{O}_p \left( \sum_{i=1}^N \left\| \tilde{\lambda}_i^f - \lambda_i^f \right\|^2 \right)} \\ & \leq \mathcal{O}_p \left( N^{-\beta} \right) \sqrt{\mathcal{O}_p \left( L_N \max_{i \leq N} \left\| \tilde{\lambda}_i^f - \lambda_i^f \right\|^2 \right)} \\ & = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \right) \mathcal{O}_p \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) = o_p(1). \end{aligned} \quad (\text{A.8})$$

In the following, we bound the second term on the right-hand side of (A.7). For this we analyze the term  $\tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} e_t$ .

$$\begin{aligned} & \mathcal{O}_p \left( N^{-\beta} \right) \left\| \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} - \Lambda^{f'} \Phi_e^{-1} \right) e_t \right\|_F \leq \\ & \mathcal{O}_p \left( N^{-\beta} \right) \left\| \left( \tilde{\Lambda}^f - \Lambda^f \right)' \tilde{\Phi}_e^{-1} e_t \right\|_F + \mathcal{O}_p \left( N^{-\beta} \right) \left\| \Lambda^{f'} \left( \tilde{\Phi}_e^{-1} - \Phi_e^{-1} \right) e_t \right\|_F. \end{aligned}$$

By Lemma A.3., the first term is bounded by

$$\begin{aligned}
& \mathcal{O}_p\left(N^{-\beta}\right) \sqrt{\sum_{i=1}^N \left\| \left( \tilde{\lambda}_i^f - \lambda_i^f \right) \left( \tilde{\Phi}_e^{-1} e_t \right)_i \right\|^2} \\
& \leq \mathcal{O}_p\left(N^{-\beta}\right) \sqrt{L_N \max_{i \leq N} \left\| \tilde{\lambda}_i^f - \lambda_i^f \right\|^2} \mathcal{O}_p(1) \\
& = \mathcal{O}_p\left(\frac{\sqrt{L_N}}{N^\beta}\right) o_p(1) = o_p(1).
\end{aligned} \tag{A.9}$$

The second term can be bounded using Lemma A.2. according to

$$\begin{aligned}
\mathcal{O}_p\left(N^{-\beta}\right) \left\| \Lambda^{f'} \left( \tilde{\Phi}_e^{-1} - \Phi_e^{-1} \right) e_t \right\|_F & = \mathcal{O}_p\left(N^{-\beta}\right) \sqrt{\sum_{i=1}^N \left\| \left( \Lambda^{f'} \Phi_e^{-1} \right)_i \left( \phi_{ie} - \tilde{\phi}_{ie} \right) \left( \tilde{\Phi}_e^{-1} e_t \right)_i \right\|^2} \\
& = \mathcal{O}_p\left(\frac{\log N}{N^\beta} \left\| \tilde{\Phi}_e - \Phi_e \right\|_F\right) = o_p(1).
\end{aligned} \tag{A.10}$$

Hence, equations (A.8), (A.9) and (A.10) yield

$$\begin{aligned}
\left\| \tilde{f}_t - f_t \right\| & = \mathcal{O}_p\left(N^{-\beta}\right) \sum_{i=1}^N \left\| \left( \Lambda^{f'} \Phi_e^{-1} \right)_i e_{it} \right\| + \mathcal{O}_p\left(\frac{\mu_1 \sqrt{L_N}}{N^\beta} + \sqrt{\frac{N(d_T + \mu_2)}{N^{2\beta}}}\right) \\
& = \mathcal{O}_p\left(N^{-\beta/2}\right) + o_p(1) = o_p(1).
\end{aligned}$$

□

To establish the consistency of the second step estimator of the idiosyncratic error covariance matrix  $\tilde{\Sigma}_e^\tau$ , we first compute the convergence rate of idiosyncratic errors  $e_{it}$ .

**Lemma A.5.**

$$\max_{i \leq N} \frac{1}{T} \sum_{t=1}^T |\tilde{e}_{it} - e_{it}|^2 = \mathcal{O}_p\left(\max(\mu_1, \mu_2)^2 + \frac{N(d_T + \max(\mu_1, \mu_2))}{L_N}\right).$$

**Proof.** As  $\tilde{e}_{it} - e_{it} = (\tilde{\lambda}_i - \lambda_i) \tilde{f}'_t + \lambda_i (\tilde{f}_t - f_t)'$ , we obtain by Lemma A.3. and Lemma A.4.

$$\begin{aligned} \max_{i \leq N} \frac{1}{T} \sum_{t=1}^T |\tilde{e}_{it} - e_{it}|^2 &\leq 2 \max_{i \leq N} \|\tilde{\lambda}_i - \lambda_i\|^2 \frac{1}{T} \sum_{t=1}^T \|\tilde{f}_t\|^2 + 2 \max_{i \leq N} \|\lambda_i\|^2 \frac{1}{T} \sum_{t=1}^T \|\tilde{f}_t - f_t\|^2 \\ &\leq \mathcal{O}_p \left( \max_{i \leq N} \|\tilde{\lambda}_i - \lambda_i\|^2 \right) + \mathcal{O}_p \left( \frac{1}{T} \sum_{t=1}^T \|\tilde{f}_t - f_t\|^2 \right) \\ &= \mathcal{O}_p \left( \max(\mu_1, \mu_2)^2 + \frac{N(d_T + \max(\mu_1, \mu_2))}{L_N} \right). \end{aligned}$$

□

By Lemma A.5 we have shown that  $\max_{i \leq N} \frac{1}{T} \sum_{t=1}^T |\tilde{e}_{it} - e_{it}|^2 = o_p(1)$ . Hence, by a similar argument as in Lemma S.1.9. in Daniele et al. (2018), we have

$$\max_{i, j \leq N} |\tilde{\sigma}_{ij} - \sigma_{ij}| = \mathcal{O}_p \left( \sqrt{\max(\mu_1, \mu_2)^2 + \frac{N(d_T + \max(\mu_1, \mu_2))}{L_N}} \right). \quad (\text{A.11})$$

In what follows, we are going to determine the convergence rate of the idiosyncratic error covariance matrix estimator based on second step soft-thresholding estimator introduced in Section 2.3.

**Lemma A.6.**

$$\|\tilde{\Sigma}_e^T - \Sigma_e\| = \mathcal{O}_p \left( S_N \sqrt{\max(\mu_1, \mu_2)^2 + \frac{N(d_T + \max(\mu_1, \mu_2))}{L_N}} \right).$$

**Proof.** The result follows from equation (A.11) and Theorem A.1. of Fan et al. (2013).

□

In the following, we continue with deriving the convergence rates for the autoregressive matrices  $\Phi_i$ , for  $i = 1, \dots, p$ , in the dynamic equation (2.2). We proceed in the same fashion as in Proposition A.2. in Bai et al. (2016). In order to improve the upcoming technicalities, we define  $\bar{p} = p + 1$  and  $\bar{T} = T - p - 1$ .

**Lemma A.7.** For  $\frac{N \log N}{T} = o(1)$ , we have

$$(a) \frac{1}{\bar{T}} \sum_{t=\bar{p}}^T \tilde{h}_{t-q} \tilde{h}'_{t-s} - \frac{1}{\bar{T}} \sum_{t=\bar{p}}^T h_{t-q} h'_{t-s} = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right), \text{ for } q, s = 0, \dots, p.$$

$$(b) \frac{1}{T} \sum_{t=\bar{p}}^T (\tilde{h}_{t-q} - h_{t-q}) \tilde{h}'_{t-s} = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right), \text{ for } q, s = 0, \dots, p.$$

$$(c) \frac{1}{T} \sum_{t=\bar{p}}^T u_t \tilde{h}'_{t-q} - \frac{1}{T} \sum_{t=\bar{p}}^T u_t h'_{t-q} = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right), \text{ for } q = 0, \dots, p.$$

**Proof.** We start with expression (a). As  $\tilde{h}_t = [\tilde{f}'_t, g'_t]'$  and  $h_t = [f'_t, g'_t]'$ , the left hand side of (a) can be expressed as  $\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & 0 \end{bmatrix}$ , where  $W_{11} = \frac{1}{T} \sum_{t=\bar{p}}^T (\tilde{f}_{t-q} - f_{t-q}) (\tilde{f}_{t-s} - f_{t-s})' + \frac{1}{T} \sum_{t=\bar{p}}^T (\tilde{f}_{t-q} - f_{t-q}) f'_{t-s} + \frac{1}{T} \sum_{t=\bar{p}}^T f_{t-q} (\tilde{f}_{t-s} - f_{t-s})'$ ,  $W_{12} = \frac{1}{T} \sum_{t=\bar{p}}^T (\tilde{f}_{t-q} - f_{t-q}) g'_{t-s}$  and  $W_{21} = \frac{1}{T} \sum_{t=\bar{p}}^T g_{t-q} (\tilde{f}_{t-s} - f_{t-s})'$ . Now, we analyse each of the three quantities separately. Based on Lemma A.4, we can see that the first term on the right hand side of  $W_{11}$  is  $\mathcal{O}_p \left( \frac{\mu_1^2 L_N + N(d_T + \mu_2)}{N^{2\beta}} \right)$ .

Using equation (A.7), we obtain the following expression for the second term in  $W_{11}$

$$\begin{aligned} \frac{1}{T} \sum_{t=\bar{p}}^T (\tilde{f}_{t-q} - f_{t-q}) f'_{t-s} &= - \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f \right)^{-1} \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \left( \tilde{\Lambda}^f - \Lambda^f \right) \frac{1}{T} \sum_{t=\bar{p}}^T f_{t-q} f'_{t-s} \\ &\quad + \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f \right)^{-1} \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \frac{1}{T} \sum_{t=\bar{p}}^T e_{t-q} f'_{t-s}. \end{aligned} \tag{A.12}$$

The first term on the right hand side of equation (A.12) is bounded by the expression  $\mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right)$ , similarly as in Lemma A.4, as  $\frac{1}{T} \sum_{t=\bar{p}}^T f_{t-q} f'_{t-s} = \mathcal{O}_p(1)$ . To bound the second term in equation (A.12), we first analyze the expression  $\tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \frac{1}{T} \sum_{t=\bar{p}}^T e_{t-q} f'_{t-s}$ .

$$\begin{aligned} \mathcal{O}_p \left( N^{-\beta} \right) &\left\| \left( \tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} - \Lambda^{f'} \Phi_e^{-1} \right) \frac{1}{T} \sum_{t=\bar{p}}^T e_{t-q} f'_{t-s} \right\|_F \leq \\ \mathcal{O}_p \left( N^{-\beta} \right) &\left[ \left\| \left( \tilde{\Lambda}^f - \Lambda^f \right)' \tilde{\Phi}_e^{-1} \frac{1}{T} \sum_{t=\bar{p}}^T e_{t-q} f'_{t-s} \right\|_F + \left\| \Lambda^{f'} \left( \tilde{\Phi}_e^{-1} - \Phi_e^{-1} \right) \frac{1}{T} \sum_{t=\bar{p}}^T e_{t-q} f'_{t-s} \right\|_F \right]. \end{aligned} \tag{A.13}$$

Based on Lemma A.1 and Lemma A.3, the first term in (A.13) is upper bounded by

$$\mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \right) \sqrt{\max_{i \leq N} \|\tilde{\lambda}_i^f - \lambda_i^f\|^2} \mathcal{O}_p \left( \sqrt{\frac{N \log N}{T}} \right).$$

Hence, as  $\frac{N \log N}{T} = o(1)$ , the first term of equation (A.13) is  $o_p(1)$ . Using Lemma A.2, the

second term in (A.13) is upper bounded by

$$\begin{aligned} & \mathcal{O}_p\left(N^{-\beta}\right)\left\|\tilde{\Phi}_e-\Phi_e\right\|_F\left\|\frac{1}{T} \sum_{t=\bar{p}}^T e_{t-q} f'_{t-s}\right\|_F \\ & \leq o_p(1) \sqrt{\frac{N \log N}{T}}=o_p(1) . \end{aligned}$$

Finally, the second term on the right hand side of equation (A.12) is bounded by

$$\begin{aligned} & \mathcal{O}_p\left(N^{-\beta}\right) \sqrt{\max _{i \leq N}\left(\Lambda^{f'} \Phi_e^{-1}\right)_i}\left\|\frac{1}{T} \sum_{t=\bar{p}}^T e_{t-q} f'_{t-s}\right\|_F \\ & \leq \mathcal{O}_p\left(N^{-\beta}\right) \sqrt{\frac{N \log N}{T}}=o_p(1) . \end{aligned}$$

Thus, the second term in  $W_{11}$  is  $\mathcal{O}_p\left(\frac{\sqrt{L_N}}{N^{\beta}}\left(\mu_1+\sqrt{\frac{N\left(d_T+\mu_2\right)}{L_N}}\right)\right)$ . It can be shown that the last term in  $W_{11}$  is of the same order. By summarizing these results, we have that  $W_{11}=\mathcal{O}_p\left(\frac{\sqrt{L_N}}{N^{\beta}}\left(\mu_1+\sqrt{\frac{N\left(d_T+\mu_2\right)}{L_N}}\right)\right)$ . By similar arguments as for  $W_{11}$ , it can be shown that  $W_{12}$  and  $W_{21}$  are as well  $\mathcal{O}_p\left(\frac{\sqrt{L_N}}{N^{\beta}}\left(\mu_1+\sqrt{\frac{N\left(d_T+\mu_2\right)}{L_N}}\right)\right)$ . Hence, (a) follows by this.

Similar as in Bai et al. (2016), expression (b) can be written as

$$\begin{bmatrix} \frac{1}{T} \sum_{t=\bar{p}}^T\left(\tilde{f}_{t-q}-f_{t-q}\right) \tilde{f}'_{t-s} & \frac{1}{T} \sum_{t=\bar{p}}^T\left(\tilde{f}_{t-q}-f_{t-q}\right) g'_{t-s} \\ 0 & 0 \end{bmatrix} .$$

Both above terms are of order  $\mathcal{O}_p\left(\frac{\sqrt{L_N}}{N^{\beta}}\left(\mu_1+\sqrt{\frac{N\left(d_T+\mu_2\right)}{L_N}}\right)\right)$ , as shown in (a). (b) follows by this result.

The left hand side of part (c) can be expressed as

$$\begin{bmatrix} \frac{1}{T} \sum_{t=\bar{p}}^T u_t\left(\tilde{f}_{t-q}-f_{t-q}\right)' \\ 0 \end{bmatrix} .$$

By using equation (A.7),  $\frac{1}{T} \sum_{t=\bar{p}}^T u_t (\tilde{f}_{t-q} - f_{t-q})'$  can be written as

$$\begin{aligned} \frac{1}{T} \sum_{t=\bar{p}}^T u_t (\tilde{f}_{t-q} - f_{t-q})' &= -\frac{1}{T} \sum_{t=\bar{p}}^T u_t f'_{t-q} (\tilde{\Lambda}^f - \Lambda^f)' \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f (\tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f)^{-1} \\ &\quad + \frac{1}{T} \sum_{t=\bar{p}}^T u_t e'_{t-q} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f (\tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f)^{-1}, \end{aligned}$$

which is bounded by

$$\begin{aligned} &\left\| \frac{1}{T} \sum_{t=\bar{p}}^T u_t f'_{t-q} (\tilde{\Lambda}^f - \Lambda^f)' \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f (\tilde{\Lambda}^{f'} \tilde{\Phi}_e^{-1} \tilde{\Lambda}^f)^{-1} \right\|_F \\ &\leq \mathcal{O}_p(N^{-\beta}) \sqrt{\sum_{i=1}^N \left\| (\tilde{\Lambda}_i^f - \Lambda_i^f)' \right\|^2} = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right) = o_p(1) \end{aligned}$$

Hence, (c) follows from this result. □

**Lemma A.8.**

$$\left\| \tilde{\Phi}_i - \Phi_i \right\|_F = \left( \sum_{t=\bar{p}}^T u_t \psi_t' \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t \psi_t' \right)^{-1} (\iota_i \otimes I_r) + \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right),$$

where  $\Phi = (\Phi_1, \dots, \Phi_p)$  and  $\psi = (h'_{t-1}, \dots, h'_{t-p})'$ .

**Proof.** We denote  $\tilde{\Phi}$  the estimator of  $\Phi$ , which is obtained by estimating the regression

$$\tilde{h}_t = \Phi_1 \tilde{h}_{t-1} + \dots + \Phi_p \tilde{h}_{t-p} + error,$$

which yields the estimator

$$\tilde{\Phi} = \left( \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{\psi}_t' \right) \left( \sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{\psi}_t' \right)^{-1},$$

with  $\psi_t = (h'_{t-1}, \dots, h'_{t-p})'$ .

By Lemma A.7 and the same arguments as in the proof of Proposition A.2. in Bai et al.

(2016), we have that

$$\left\| \tilde{\Phi} - \Phi \right\|_F = \left( \sum_{t=\bar{p}}^T u_t \psi_t' \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t \psi_t' \right)^{-1} + \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right).$$

The result follows by post-multiplying  $\iota_i \otimes I_r$  on both sides. □

In the following, we will focus on the consistency of the parameter estimates after rotation. As the rotation matrix  $A$  in Section 2.4 depends on the covariance matrix  $\Omega$  of the VAR innovations  $\eta_t$  in equation (2.2), we first concentrate on the consistency of  $\Omega$ . For this, we introduce the following two lemmas that are similarly established by Bai et al. (2016).

**Lemma A.9.** *For any two compatible matrices  $\mathcal{B}$  and  $\mathcal{C}$  and the corresponding estimates  $\tilde{\mathcal{B}}$  and  $\tilde{\mathcal{C}}$ , we have*

$$\tilde{\mathcal{B}}\tilde{\mathcal{C}}^{-1}\tilde{\mathcal{B}}' - \mathcal{B}\mathcal{C}^{-1}\mathcal{B}' = (\tilde{\mathcal{B}} - \mathcal{B})\mathcal{C}^{-1}\mathcal{B}' + \mathcal{B}\mathcal{C}^{-1}(\tilde{\mathcal{B}} - \mathcal{B})' - \mathcal{B}\mathcal{C}^{-1}(\tilde{\mathcal{C}} - \mathcal{C})\mathcal{C}^{-1}\mathcal{B}' + \mathcal{R},$$

where

$$\begin{aligned} \mathcal{R} = & -(\tilde{\mathcal{B}} - \mathcal{B})\tilde{\mathcal{C}}^{-1}(\tilde{\mathcal{C}} - \mathcal{C})\mathcal{C}^{-1}\mathcal{B}' + (\tilde{\mathcal{B}} - \mathcal{B})\tilde{\mathcal{C}}^{-1}(\tilde{\mathcal{B}} - \mathcal{B})' \\ & + \mathcal{B}\tilde{\mathcal{C}}^{-1}(\tilde{\mathcal{C}} - \mathcal{C})\mathcal{C}^{-1}(\tilde{\mathcal{C}} - \mathcal{C})\mathcal{C}^{-1}\tilde{\mathcal{B}}' - \mathcal{B}\tilde{\mathcal{C}}^{-1}(\tilde{\mathcal{C}} - \mathcal{C})\mathcal{C}^{-1}(\tilde{\mathcal{B}} - \mathcal{B})'. \end{aligned}$$

**Proof.** See Lemma B.1. in Bai et al. (2016). □

**Lemma A.10.**

$$\frac{1}{T}\tilde{H}'M_{\tilde{\Phi}}\tilde{H} - \frac{1}{T}H'M_{\Phi}H = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right),$$

where

$$\frac{1}{T}\tilde{H}'M_{\tilde{\Phi}}\tilde{H} = \frac{1}{T}\sum_{t=\bar{p}}^T \tilde{h}_t \tilde{h}_t' - \left( \frac{1}{T}\sum_{t=\bar{p}}^T \tilde{h}_t \tilde{\psi}_t' \right) \left( \frac{1}{T}\sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{\psi}_t' \right)^{-1} \left( \frac{1}{T}\sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{h}_t' \right),$$



and  $\frac{1}{T}H'M_\Phi H$  is defined similarly.

**Proof.** The proof is identically conducted as in Bai et al. (2016). Hereby, we consider the following three expressions that are bounded using Lemma A.7 (a).

$$\begin{aligned}\frac{1}{T} \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{h}'_t - \frac{1}{T} \sum_{t=\bar{p}}^T h_t h'_t &= \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right), \\ \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{\psi}'_t - \frac{1}{T} \sum_{t=\bar{p}}^T h_t \psi'_t &= \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right), \\ \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{\psi}'_t - \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t \psi'_t &= \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right).\end{aligned}$$

The result follows based on the above results and Lemma A.9. □

The covariance matrix estimator of the innovations  $\eta_t$  can be bounded according to the following lemma.

**Lemma A.11.**

$$\left\| \tilde{\Omega} - \Omega \right\|_F = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right).$$

**Proof.** The estimator of  $\tilde{\eta}_t$  is defined as

$$\tilde{\Omega} = \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{\eta}_t \tilde{\eta}'_t,$$

where  $\tilde{\eta}_t$  are the residuals of the regression  $\tilde{h}_t = \Phi_1 \tilde{h}_{t-1} + \dots + \Phi_p \tilde{h}_{t-p} + \text{error}$ . Hence,

$$\tilde{\eta}_t = \tilde{h}_t - \left( \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{\psi}'_t \right) \left( \sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{\psi}'_t \right)^{-1} \tilde{\psi}_t,$$

where  $\tilde{\psi}_t = (\tilde{h}'_{t-1}, \dots, \tilde{h}'_{t-p})'$ . Based on the previous result we have

$$\tilde{\Omega} = \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{h}'_t - \left( \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{\psi}'_t \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{\psi}'_t \right)^{-1} \left( \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{h}'_t \right). \quad (\text{A.14})$$

Hence, given (A.14) we obtain

$$\begin{aligned} \tilde{\Omega} - \Omega &= \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{h}'_t - \left( \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{h}_t \tilde{\psi}'_t \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{\psi}'_t \right)^{-1} \left( \frac{1}{T} \sum_{t=\bar{p}}^T \tilde{\psi}_t \tilde{h}'_t \right) \\ &\quad - \frac{1}{T} \sum_{t=\bar{p}}^T h_t h'_t + \left( \frac{1}{T} \sum_{t=\bar{p}}^T h_t \psi'_t \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t \psi'_t \right)^{-1} \left( \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t h'_t \right). \end{aligned} \quad (\text{A.15})$$

The expression (A.15) is bounded by Lemma A.10. These results yield

$$\left\| \tilde{\Omega} - \Omega \right\|_F = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right).$$

□

The estimated rotation matrix  $\tilde{A} = \begin{bmatrix} I_{r_1} & -\tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} \\ 0 & I_{r_2} \end{bmatrix}$  can be bounded using the following lemma.

**Lemma A.12.**

$$\left\| \tilde{A}_{21} - A_{21} \right\|_F = \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right),$$

where  $\tilde{A}_{21} = -\tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1}$  and  $A_{21} = -\Omega_{fg} \Omega_{gg}^{-1}$ .

**Proof.** Given the definitions of  $\tilde{A}_{21}$  and  $A_{21}$ , we obtain

$$\begin{aligned} \tilde{A}_{21} - A_{21} &= \Omega_{fg} \Omega_{gg}^{-1} - \tilde{\Omega}_{fg} \tilde{\Omega}_{gg}^{-1} \\ &= - \left( \tilde{\Omega}_{fg} - \Omega_{fg} \right) \left( \tilde{\Omega}_{gg}^{-1} - \Omega_{gg}^{-1} \right) - \Omega_{fg} \left( \tilde{\Omega}_{gg}^{-1} - \Omega_{gg}^{-1} \right) - \left( \tilde{\Omega}_{fg} - \Omega_{fg} \right) \Omega_{gg}^{-1}. \end{aligned} \quad (\text{A.16})$$

Based on Lemma A.11, we can upper bound the expression (A.16) by

$$\begin{aligned} \left\| \tilde{A}_{21} - A_{21} \right\|_F &\leq - \left\| \tilde{\Omega}_{fg} - \Omega_{fg} \right\|_F \left\| \tilde{\Omega}_{gg}^{-1} - \Omega_{gg}^{-1} \right\|_F - \left\| \Omega_{fg} \right\| \left\| \tilde{\Omega}_{gg}^{-1} - \Omega_{gg}^{-1} \right\|_F - \left\| \tilde{\Omega}_{fg} - \Omega_{fg} \right\|_F \left\| \Omega_{gg}^{-1} \right\| \\ &= \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right). \end{aligned}$$

□

**Lemma A.13.**

$$\max_{i \leq N} \left\| \hat{\lambda}_i^g - \lambda_i^{*g} \right\| = \mathcal{O}_p \left( \sqrt{\mu_1} + \sqrt{\mu_2} + \sqrt{d_T} \right).$$

**Proof.** The rotated factor loadings of the observed factors are defined as  $\hat{\lambda}_i^g = \tilde{\lambda}_i^g + \tilde{\lambda}_i^f \tilde{A}_{21}$  and  $\lambda_i^{*g} = \lambda_i^g + \lambda_i^f A_{21}$ . Hence, their difference can be expressed as

$$\hat{\lambda}_i^g - \lambda_i^{*g} = \tilde{\lambda}_i^g - \lambda_i^g + \left( \tilde{\lambda}_i^f - \lambda_i^f \right) A_{21} + \lambda_i^f \left( \tilde{A}_{21} - A_{21} \right) + \left( \tilde{\lambda}_i^f - \lambda_i^f \right) \left( \tilde{A}_{21} - A_{21} \right). \quad (\text{A.17})$$

The above expression can be bounded using Lemma A.3 and Lemma A.12 according to

$$\begin{aligned} \max_{i \leq N} \left\| \hat{\lambda}_i^g - \lambda_i^{*g} \right\| &\leq \max_{i \leq N} \left\| \tilde{\lambda}_i^g - \lambda_i^g \right\| + \max_{i \leq N} \left\| \tilde{\lambda}_i^f - \lambda_i^f \right\| \left\| A_{21} \right\|_F \\ &\quad + \max_{i \leq N} \left\| \lambda_i^f \right\| \left\| \tilde{A}_{21} - A_{21} \right\|_F + \max_{i \leq N} \left\| \tilde{\lambda}_i^f - \lambda_i^f \right\| \left\| \tilde{A}_{21} - A_{21} \right\|_F \\ &= \mathcal{O}_p \left( \mu_1 + \mu_2 + \sqrt{\frac{N(d_T + \mu_1)}{L_N}} + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \\ &\leq \mathcal{O}_p \left( \sqrt{\mu_1} + \sqrt{\mu_2} + \sqrt{d_T} \right). \end{aligned}$$

□

**Lemma A.14.**

$$\left\| \hat{f}_t - f_t^* \right\| = \mathcal{O}_p \left( N^{-\beta/2} \right) + \mathcal{O}_p \left( \frac{\mu_1 \sqrt{L_N}}{N^\beta} + \sqrt{\frac{N(d_T + \mu_2)}{N^{2\beta}}} \right).$$

**Proof.** The rotated unobserved factors are defined as  $\hat{f}_t = \tilde{f}_t - \tilde{A}_{21} g_t$  and  $f_t^* = f_t - A_{21} g_t$ .

Thus, the difference of both terms is given by

$$\hat{f}_t - f_t^* = \tilde{f}_t - f_t - \left( \tilde{A}_{21} - A_{21} \right) g_t. \quad (\text{A.18})$$

Using the euclidean norm, we determine the upper bound of (A.18), using Lemma A.4 and Lemma A.12, by

$$\begin{aligned} \left\| \hat{f}_t - f_t^* \right\| &\leq \left\| \tilde{f}_t - f_t \right\| - \left\| \tilde{A}_{21} - A_{21} \right\|_F g_t. \\ &= \mathcal{O}_p \left( N^{-\beta/2} \right) + \mathcal{O}_p \left( \frac{\mu_1 \sqrt{L_N}}{N^\beta} + \sqrt{\frac{N(d_T + \mu_2)}{N^{2\beta}}} \right) = o_p(1). \end{aligned}$$

□

Finally, for the rotated coefficient matrices of the dynamic equation in (2.2) we establish the following lemma.

**Lemma A.15.**

$$\left\| \hat{\Phi}_i - \Phi_i^* \right\|_F = \left( \sum_{t=\bar{p}}^T u_t \psi_t' \right) \left( \frac{1}{T} \sum_{t=\bar{p}}^T \psi_t \psi_t' \right)^{-1} (\iota_i \otimes I_r) + \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right),$$

where  $\iota_i$  is the  $i$ -th column of the  $r \times r$  identity matrix.

**Proof.** Given the definitions  $\hat{\Phi}_i = \tilde{A} \tilde{\Phi}_i \tilde{A}^{-1}$  and  $\Phi_i^* = A \Phi_i A^{-1}$ , we obtain

$$\begin{aligned} \hat{\Phi} - \Phi^* &= \left( \tilde{A} - A \right) \left( \tilde{\Phi} - \Phi \right) \left( \tilde{A}^{-1} - A^{-1} \right) + A \left( \tilde{\Phi} - \Phi \right) \left( \tilde{A}^{-1} - A^{-1} \right) \\ &\quad + \left( \tilde{A} - A \right) \left( \tilde{\Phi} - \Phi \right) A^{-1} + A \left( \tilde{\Phi} - \Phi \right) A^{-1} + \left( \tilde{A} - A \right) \Phi A^{-1} \\ &\quad + \left( \tilde{A} - A \right) \Phi \left( \tilde{A}^{-1} - A^{-1} \right) + A \Phi \left( \tilde{A}^{-1} - A^{-1} \right) \end{aligned} \quad (\text{A.19})$$

Based on Lemma A.8 and Lemma A.12, we get the expression

$$\begin{aligned}
\left\| \hat{\Phi} - \Phi^* \right\|_F &\leq \left\| \tilde{A} - A \right\|_F \left\| \tilde{\Phi} - \Phi \right\|_F \left\| \tilde{A}^{-1} - A^{-1} \right\|_F + A \left\| \tilde{\Phi} - \Phi \right\|_F \left\| \tilde{A}^{-1} - A^{-1} \right\|_F \\
&\quad + \left\| \tilde{A} - A \right\|_F \left\| \tilde{\Phi} - \Phi \right\|_F A^{-1} + A \left\| \tilde{\Phi} - \Phi \right\|_F A^{-1} + \left\| \tilde{A} - A \right\|_F \Phi A^{-1} \\
&\quad + \left\| \tilde{A} - A \right\|_F \Phi \left\| \tilde{A}^{-1} - A^{-1} \right\|_F + A \Phi \left\| \tilde{A}^{-1} - A^{-1} \right\|_F \\
&= \left( \sum_{t=\bar{p}}^T u_t \psi'_t \right) \left( \frac{1}{\bar{T}} \sum_{t=\bar{p}}^T \psi_t \psi'_t \right)^{-1} + \mathcal{O}_p \left( \frac{\sqrt{L_N}}{N^\beta} \left( \mu_1 + \sqrt{\frac{N(d_T + \mu_2)}{L_N}} \right) \right).
\end{aligned}$$

The result follows by post-multiplying  $\iota_i \otimes I_r$  on both sides.

□

## B Tables

Table B.1: Data and Transformations

	Description of the variables	Transformation code	Latent factor	FFR
1	real personal income	5		0
2	real personal income ex transfer receipts	5		0
3	real personal consumption expenditures	5		0
4	real manufacturing and trade industries sales	5	5	-0.0025
5	retail and food services sales	5		0
6	IP index	5	1,5	-0.0444
7	IP final products and nonindustrial supplies	5	1,5	-0.0494
8	IP final products	5	5	-0.0473
9	IP consumer goods	5	5	-0.0362
10	IP durable consumer goods	5	5	-0.0267
11	IP nondurable consumer goods	5	5	-0.0103
12	IP business equipment	5	1,5	-0.0283
13	IP materials	5	1,5	-0.0230
14	IP durable materials	5	1,5	-0.0300
15	IP nondurable materials	5	5	-0.0026
16	IP manufacturing	5	1,5	-0.0445
17	IP residential utilities	5		0
18	IP fuels	5		0
19	capacity utilization	2	1,5	-0.0402
20	US ISM Manufacturers survey: production index	1	1	-0.0110
21	US ISM Manufacturers survey: employment index	1	1	-0.0119
22	US ISM Purchasing Managers Index	1	1	-0.0138
23	US ISM Manufacturers survey: supplier delivery index	1	1	-0.0038
24	US ISM Manufacturers survey: new orders index	1	1	-0.0096
25	US ISM Manufacturers survey: inventories index	1	1	-0.0075
26	US ISM Manufacturers survey: prices paid index	1	1,2	-0.0021
27	help wanted index	2		0
28	ratio of help wanted/number unemployed	2	1	-0.0005
29	civilian labor force	5		0
30	civilian employment	5	1	-0.0041
31	civilian unemployment rate	2	1	0.0060
32	average duration of unemployment	2		0
33	civilians unemployed - less than 5 weeks	5		0
34	civilians unemployed for 5-14 weeks	5		0

*Continued on next page*

Table B.1 – *cont.*

	Description of the variables	Transformation code	Latent factor	FFR
35	civilians unemployed - 15 weeks and over	5	1	0.0050
36	civilians unemployed for 15-26 weeks	5		0
37	civilians unemployed for 27 weeks and over	5	1	0.0030
38	initial claims	5		0
39	All employees: total nonfarm	5	1	-0.0194
40	All employees: goods-producing industries	5	1	-0.0222
41	All employees: mining	5		0
42	All employees: construction	5	1	-0.0120
43	All employees: manufacturing	5	1,5	-0.0217
44	All employees: durable goods	5	1,5	-0.0219
45	All employees: nondurable goods	5	1	-0.0130
46	All employees: service-providing industries	5	1	-0.0139
47	All employees: trade, transportation and utilities	5	1	-0.0152
48	All employees: wholesale trade	5	1	-0.0159
49	All employees: retail trade	5	1	-0.0098
50	All employees: financial activities	5	1	-0.0044
51	All employees: government	5		0
52	Average hourly earnings: goods-producing	1	1	-0.0110
53	Average weekly overtime hours: manufacturing	2		0
54	Average weekly hours: manufacturing	1	1	-0.0121
55	Housing starts: total new privately owed	5		0
56	Housing starts: NE	5		0
57	Housing starts: MW	5		0
58	Housing starts: S	5		0
59	Housing starts: W	5		0
60	New private housing permits	5		0
61	New private housing permits: NE	5		0
62	New private housing permits: MW	5		0
63	New private housing permits: S	5		0
64	New private housing permits: W	5		0
65	Moodys Seasoned Aaa Corporate Bond Yield, Percent	1	3	-0.0912
66	Moodys Seasoned Baa Corporate Bond Yield, Percent	2	3,4	-0.0722
67	30-Year Fixed Rate Mortgage Average in the United States, Percent	2	3	-0.0644
68	Excess Bond Premium	1	1	0.0119
69	Spread UK-US	2		0
70	Spread CAN-US	2		0

*Continued on next page*

Table B.1 – *cont.*

	Description of the variables	Transformation code	Latent factor	FFR
71	Spread SW-US	2		0
72	Spread JPN-US	2		-0.1383
73	New orders for durable goods	5		0
74	New orders for nondefense capital goods	5		0
75	Unfilled orders for durable goods	5	1	-0.0019
76	Business Inventories	5	1	-0.0100
77	Inventories to sales ratio	2		0
78	M1 money stock	5		0
79	M2 money stock	5		0
80	Real M2 money stock	5	2	0.0232
81	St. Louis adjusted monetary base	5	2	0.0023
82	Total reserves of depository institutions	5		0
83	Commercial and industrial loans	5	1	-0.0029
84	Real estate loans at all commercial banks	5		0
85	Total nonrevolving credit	5		0
86	nonrevolving consumer credit to personal income	2		0
87	S&P 500 composite	5	4	0.0092
88	S&P industrials	5	4	0.0091
89	S&P common stock dividend yields	2	4	-0.0083
90	S&P common stock price-earnings ratio	5	1,4	0.0066
91	VXO	1	1,4	0.0063
92	Dow Jones Industrial	5	4	0.0034
93	Dow Jones Utilities	5		0
94	Nasdaq Composite	5	4	0.0015
95	Dow Jones	5	4	0.0015
96	Nasdaq Industrial	5	4	0.0013
97	3 month financial commercial paper rate	2	1	0.2361
98	3 month Treasury bill	2		0.4455
99	6 month Treasury bill	2	1,3	0.4731
100	1 year Treasury rate	2	3	0.4023
101	5 year Treasury rate	2	3	0.0871
102	10 year Treasury rate	2	3	-0.0207
103	Moody AAA corporate bond yield	2	3	-0.0912
104	Moody BAA corporate bond yield	2	3,4	-0.0722
105	Trade weighted US Dollar index major currencies	5		0
106	Switzerland US foreign exchange rate	5		0
107	Japan US foreign exchange rate	5		0
108	US UK foreign exchange rate	5		0

*Continued on next page*



Table B.1 – *cont.*

Description of the variables		Transformation code	Latent factor	FFR
109	Canada US foreign exchange rate	5		0
110	PPI: finished goods	5	2	-0.0349
111	PPI: finished consumer goods	5	2	-0.0353
112	PPI: intermediate materials	5	2	-0.0332
113	PPI: crude materials	5	2	-0.0145
114	crude oil	5	2	-0.0113
115	PPI: metals	5		0
116	CPI: all items	5	2	-0.0505
117	CPI: apparel	5		0
118	CPI: transportation	5	2	-0.0480
119	CPI: commodities	5	2	-0.0508
120	CPI: durables	5		0
121	CPI: all items less food	5	2	-0.0501
122	CPI: all items less shelter	5	2	-0.0516
123	CPI: all items less medical care	5	2	-0.0509
124	Personal consumption expenditure: chain index	6	2	-0.0090
125	Personal consumption expenditure: durable goods	5		0
126	Personal consumption expenditure: nondurable goods	5	2	-0.0491

Note: This table shows the different observed variables on monthly frequency for a sampling period January 1985 until December 2016. The data is retrieved from the [McCracken and Ng \(2016\)](#) FRED-MD data base and datastream. The transformation codes are labeled as follows: 1 = no transformation, 2 =  $\Delta x_t$ , 3 =  $\Delta^2 x_t$ , 4 =  $\log(x_t)$ , 5 =  $\Delta \log(x_t)$ , 6 =  $\Delta^2 \log(x_t)$ . The last two columns are associated with our regularized FAVAR model. The column “Latent factor” shows the latent factors the specific time series load on. The latent factors are denoted as 1 = labor market, 2 = price, 3 = industrial production, 4 = stock market, 5 = credit spread. The column “FFR” denotes the estimated loadings onto the observable factor.

Table B.2: Naming variables scheme

Index	Description of the variables
6	Industrial Production (IP) index
31	Civilian unemployment rate
87	S&P 500 composite
66	Moody’s Seasoned Baa Corporate Bond Yield
116	Consumer Price Index (CPI): all items

Note: This table lists the observed time series that are used as naming variables for the named factor identification case (*IRb*).

Table B.3: Contemporaneous impact matrices

<b>Panel A: regularized FAVAR model - RFAVAR</b>					
1	0	0	0	0	-0.0555 (0.0316)
0	1	0	0	0	-0.1093 (0.0676)
0	0	1	0	0	-0.1014 (0.0721)
0	0	0	1	0	0.0169 (0.0895)
0	0	0	0	1	-0.1742 (0.0836)
0	0	0	0	0	1
<b>Panel B: named factors scheme - FAVAR-NF</b>					
-1.2426	0.3996	0.2934	0.2029	0.0824	-0.4235 (0.0449)
-0.8293	-0.9389	0.0315	-0.2053	0.0276	-0.1201 (0.0738)
-0.6111	-0.0810	-0.2619	-0.1330	-1.1376	0.0337 (0.0870)
1.7036	-0.3395	0.9813	1.1365	0.3018	0.1589 (0.0752)
-1.1194	-0.1157	-0.9908	0.7985	0.5888	-0.1626 (0.0836)
0	0	0	0	0	1

Note: This tables shows the contemporaneous impact matrices associated with the different identification schemes. Panel A shows the impact matrix for our structural regularized FAVAR model (*IRa*). Panel B gives the impact matrix associated with the named factor identification scheme (*IRb*). Standard errors are given in brackets.

## C Figures

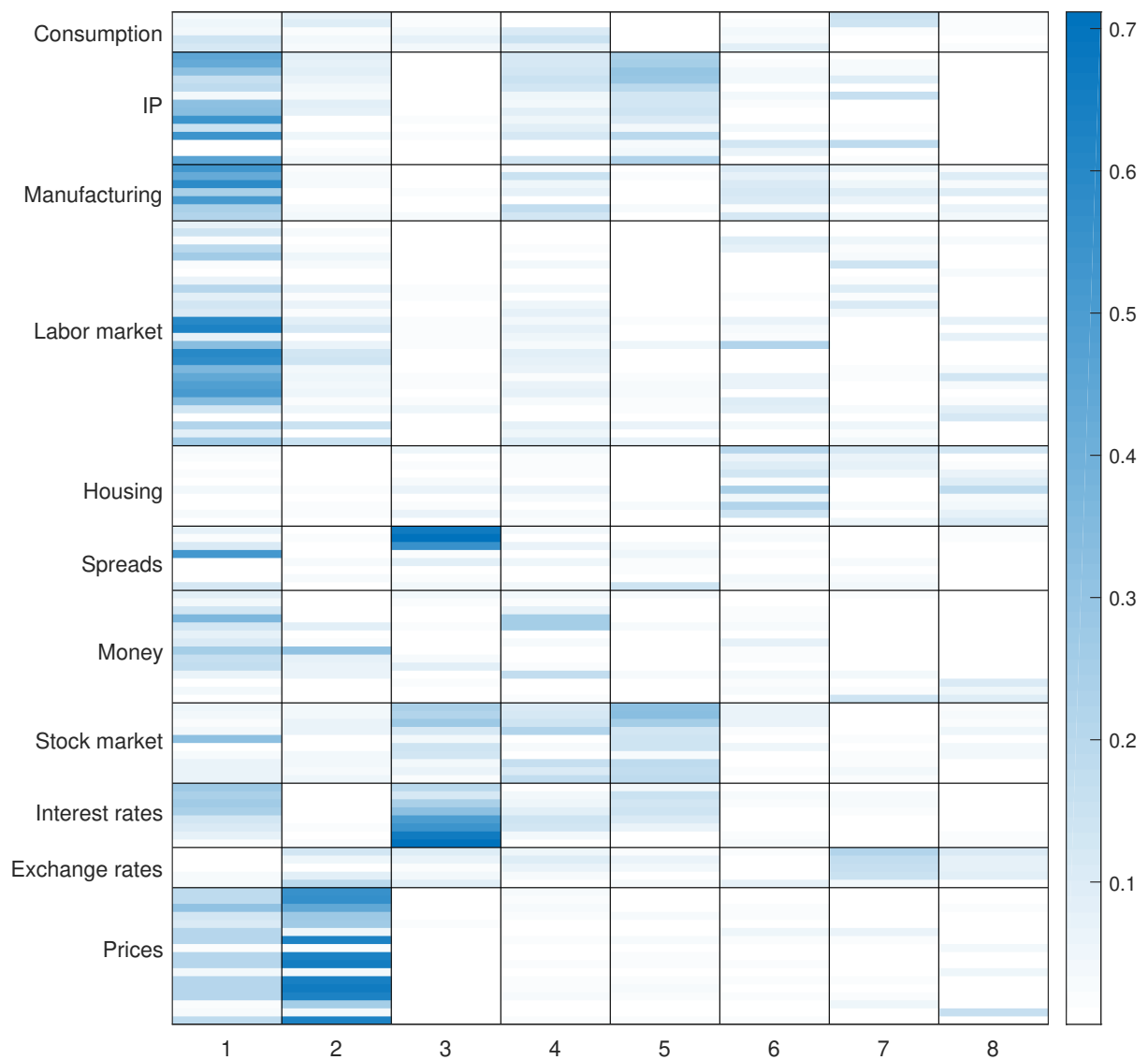


Figure C.1:  $R^2$  plot for a factor model based on principal components analysis

This graph shows  $R^2$  for univariate regressions of the observed variables on each of the eight factors. The number of factors is chosen based on the [Bai and Ng \(2002\)](#)  $IC_1$  criterion.

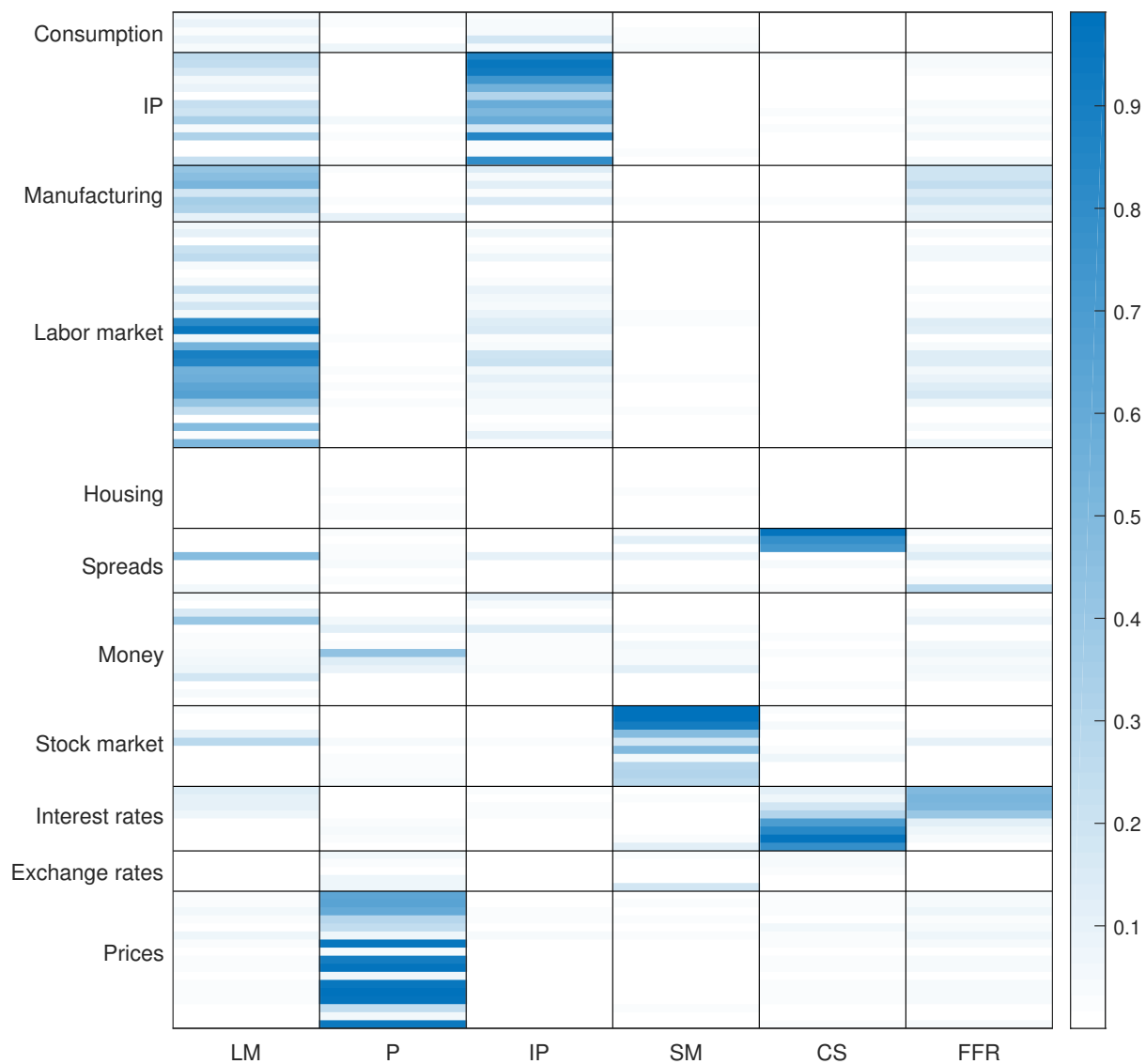
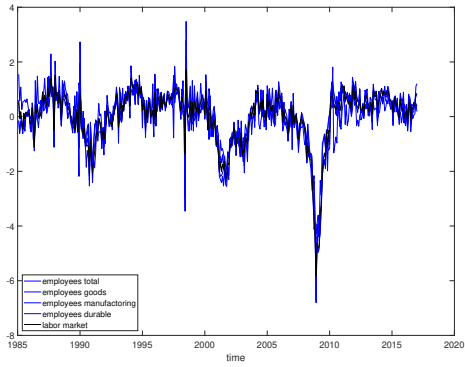
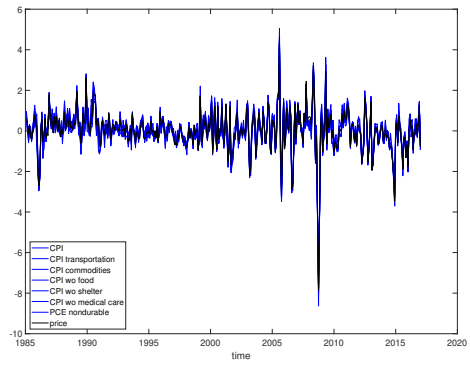


Figure C.2:  $R^2$  plot for the regularized FAVAR

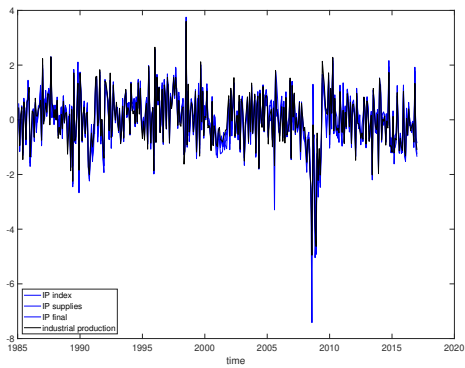
This graph shows  $R^2$  for univariate regressions of the observed variables on each of the factors. The factors are abbreviated as follows: 'LM' labor market, 'P' price, 'IP' industrial production, 'SM' stock market, 'CS' credit spread and 'FFR' Federal Funds rate.



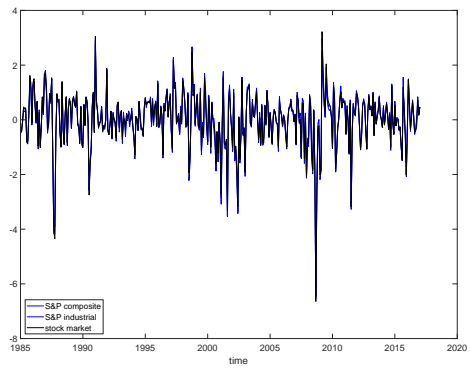
(a) Labor Market



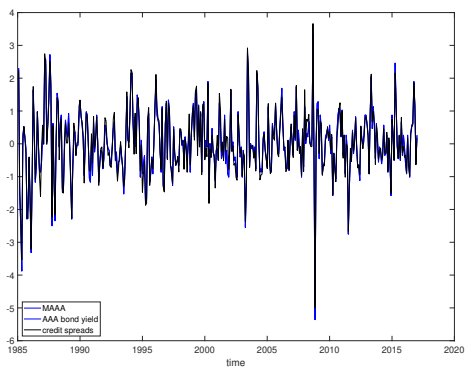
(b) Price



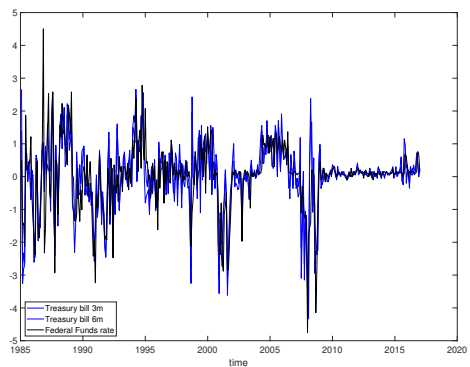
(c) Industrial Production



(d) Stock Market



(e) Credit Spread



(f) Federal Funds rate

Figure C.3: Factor plots for the regularized FAVAR

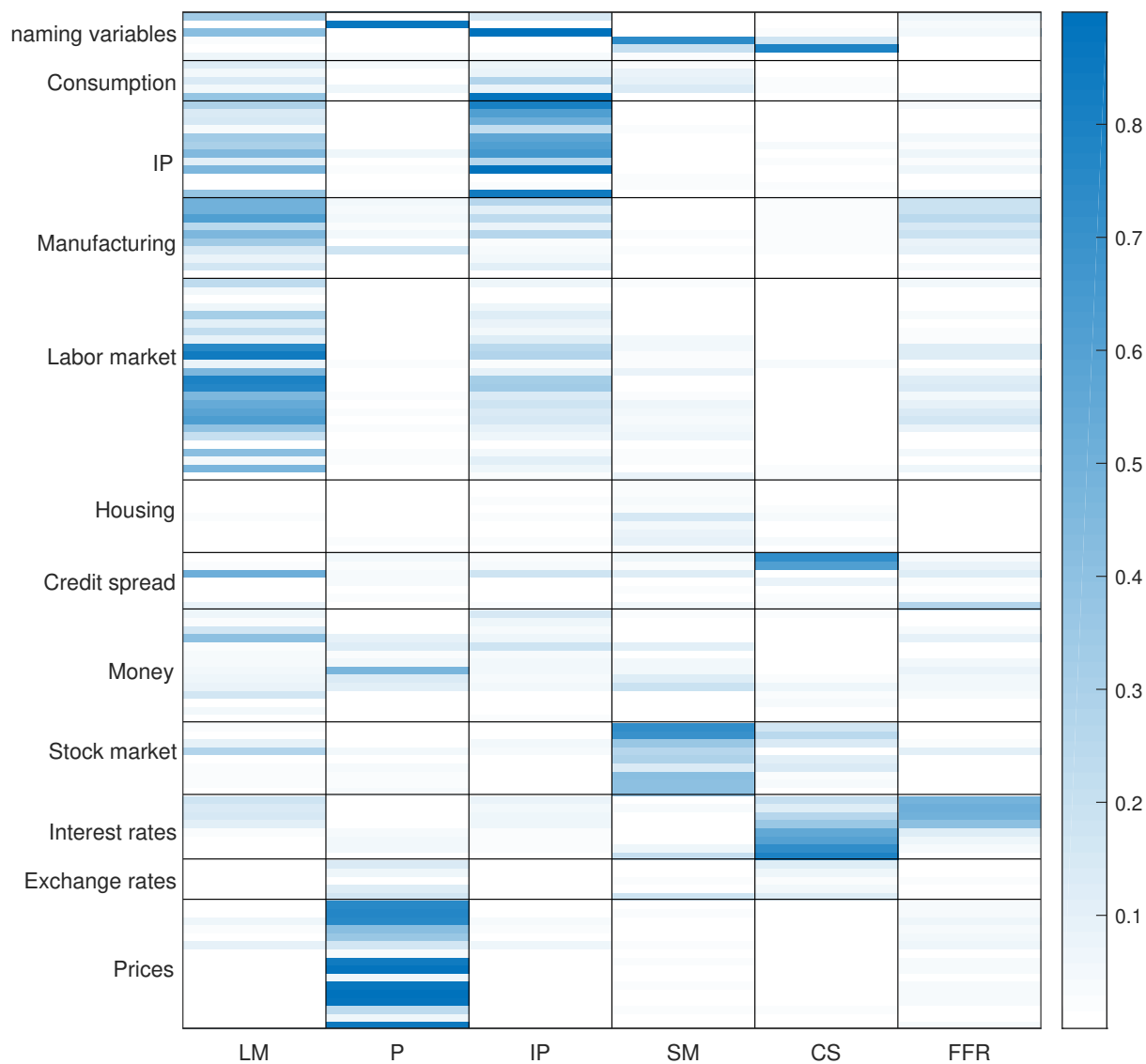
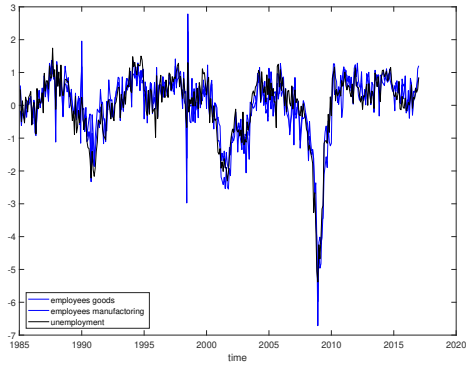
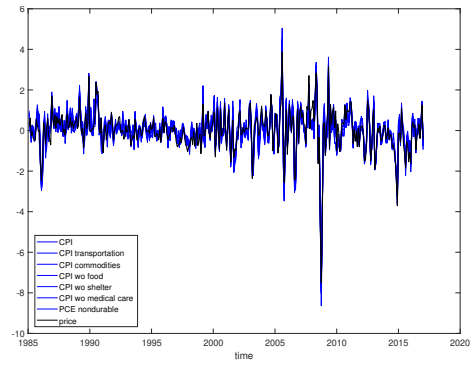


Figure C.4:  $R^2$  plot for the named factor FAVAR

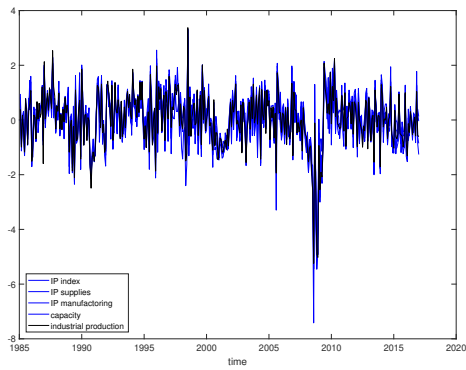
This graph shows  $R^2$  for univariate regressions of the observed variables on each of the factors. The factors are abbreviated as follows: 'LM' labor market, 'P' price, 'IP' industrial production, 'SM' stock market, 'CS' credit spread and 'FFR' Federal Funds rate. The naming variables are given in Table B.2.



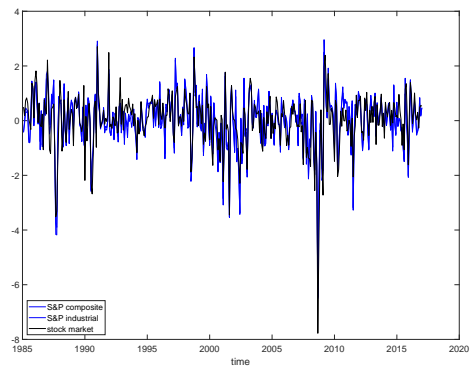
(a) Labor Market



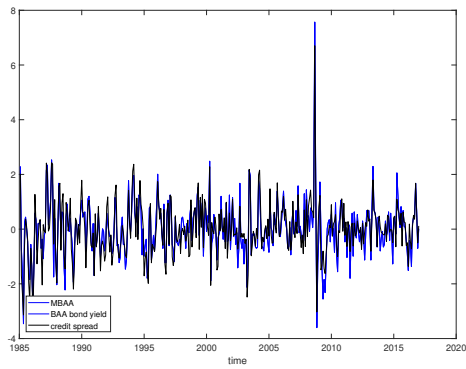
(b) Price



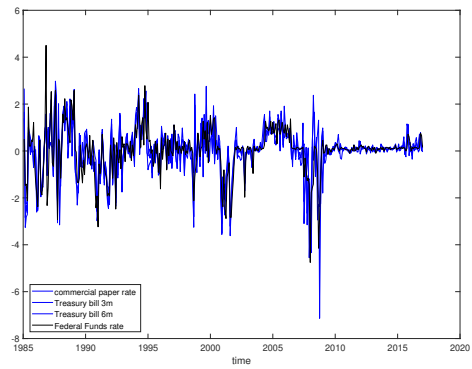
(c) Industrial Production



(d) Stock Market



(e) Credit Spread



(f) Federal Funds rate

Figure C.5: Factor plots for the named factor FAVAR

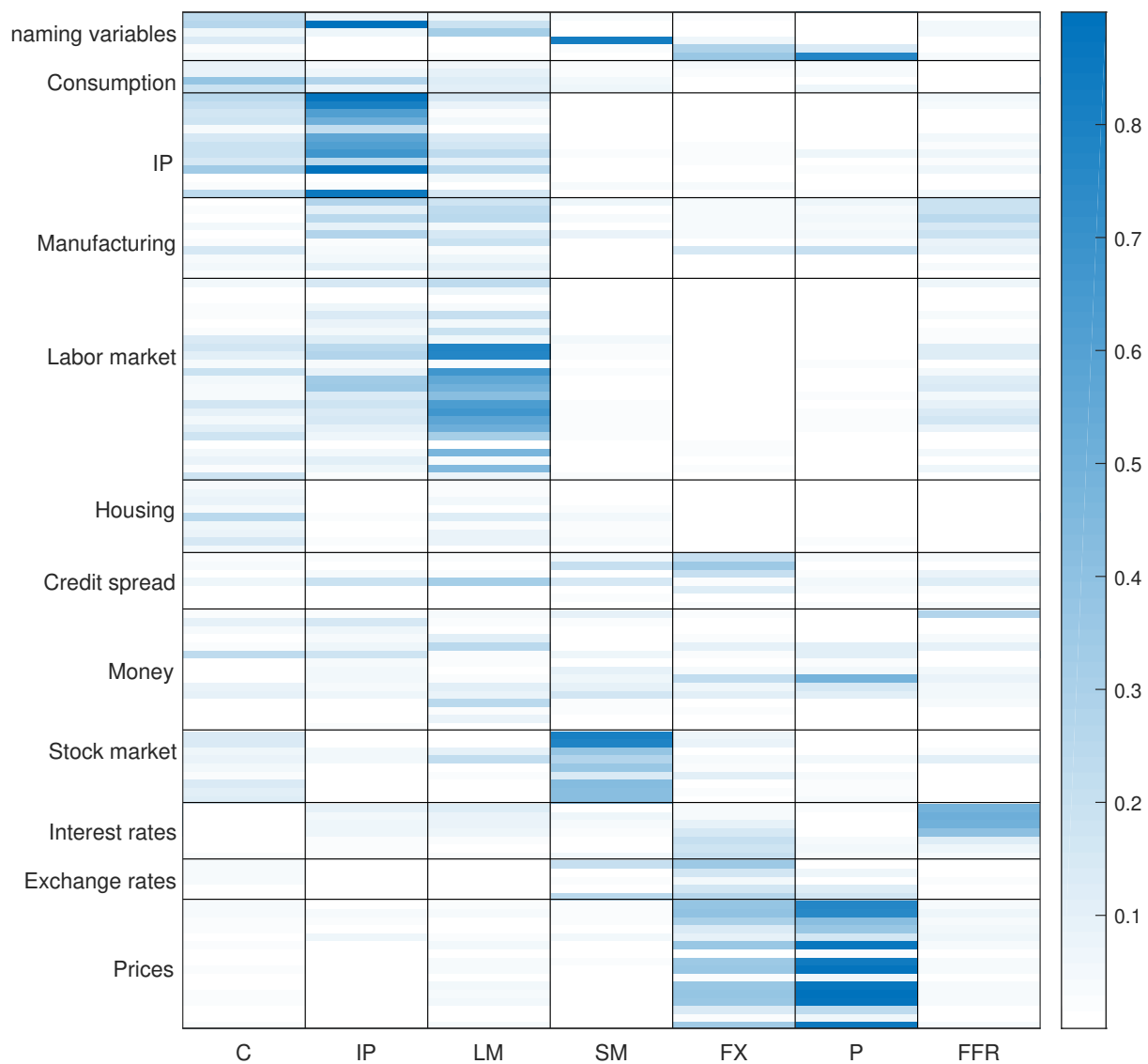


Figure C.6:  $R^2$  plot for the named factor FAVAR: oil price application

This graph shows  $R^2$  for univariate regressions of the observed variables on each of the factors. The factors are abbreviated as follows: 'LM' labor market, 'P' price, 'IP' industrial production, 'SM' stock market, 'C' consumption, 'FX' is a trade weighted currency index and 'FFR' Federal Funds rate. The naming variables are based on [Stock and Watson \(2016\)](#).



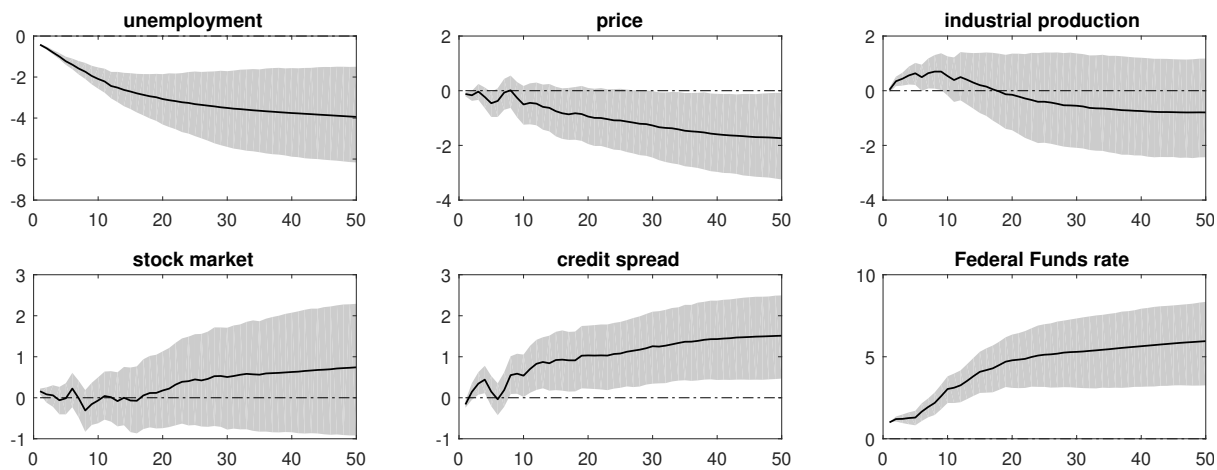


Figure C.7: Accumulated impulse responses to a monetary policy shock on the factors for the named factor FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals.

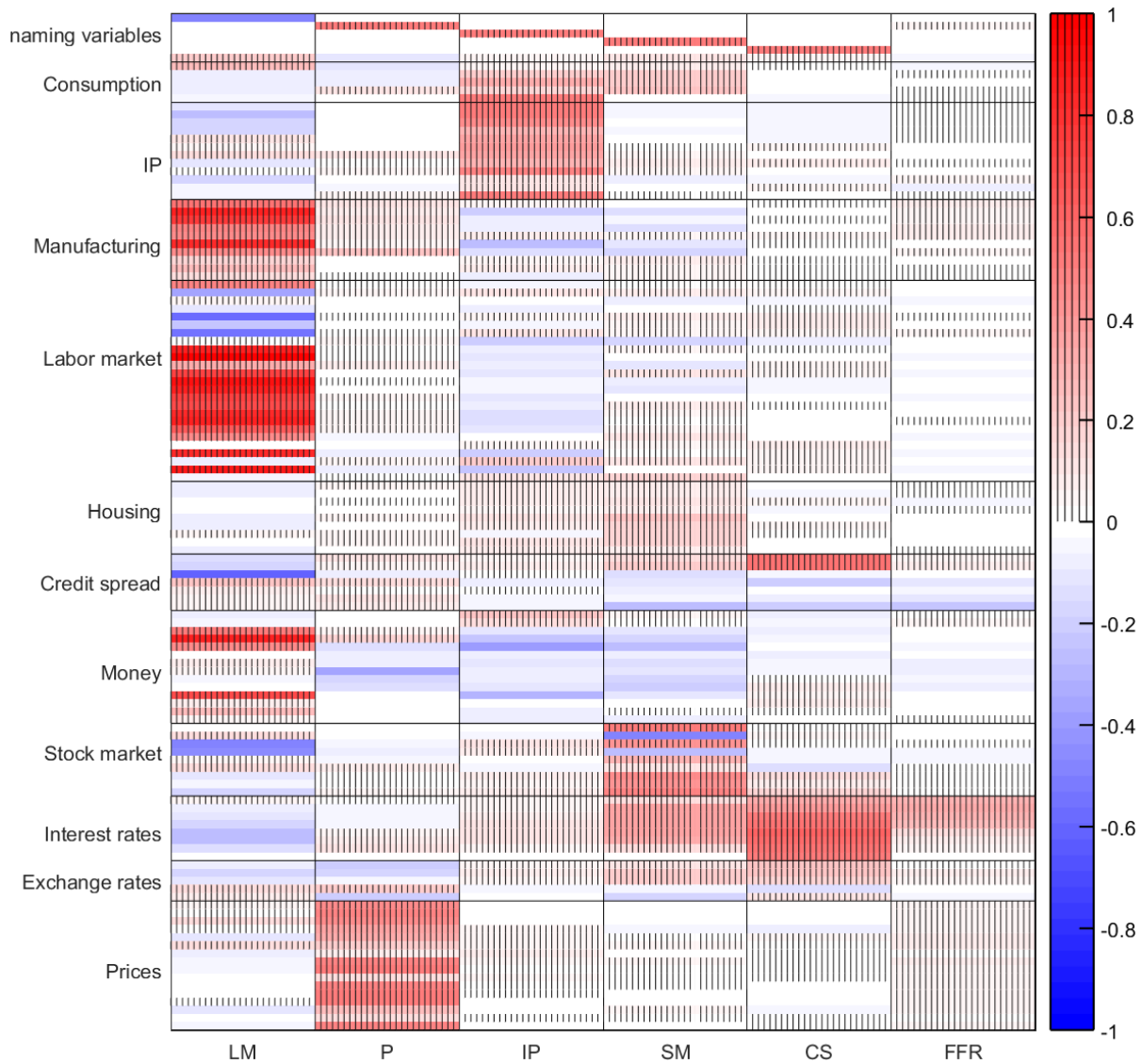


Figure C.8: Impact matrix on the observed variables  $x_t$  for the named factor FAVAR  
 This graph shows the contemporaneous impact matrix of a 100bp shock in the factor innovations to the observed time series for a factor model with named factor identification. The factors are abbreviated as follows: 'LM' labor market, 'P' price, 'IP' industrial production, 'SM' stock market, 'CS' credit spread and 'FFR' Federal Funds rate. The naming variables for the first scheme are given in Table B.2.

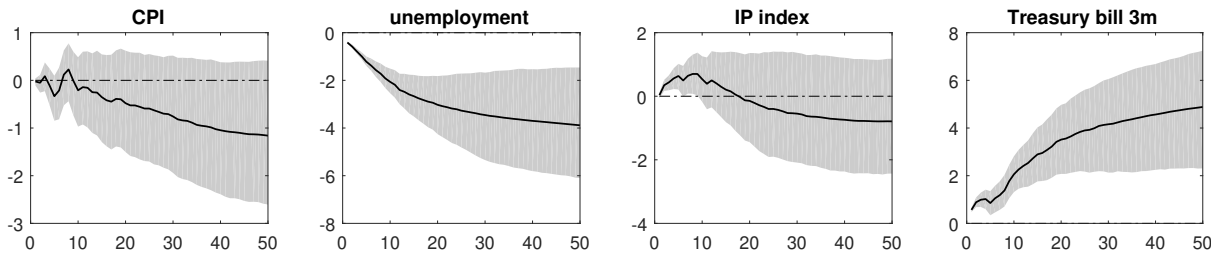


Figure C.9: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the named factor FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals.

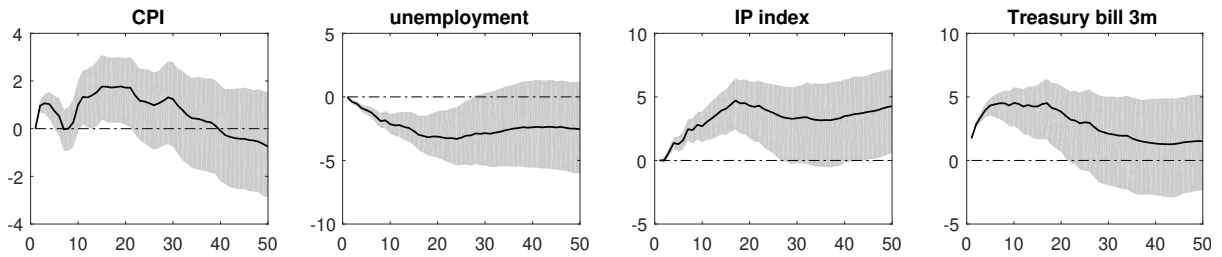


Figure C.10: Accumulated impulse responses on the observed variables  $x_t$  for the dynamic factor model

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals.

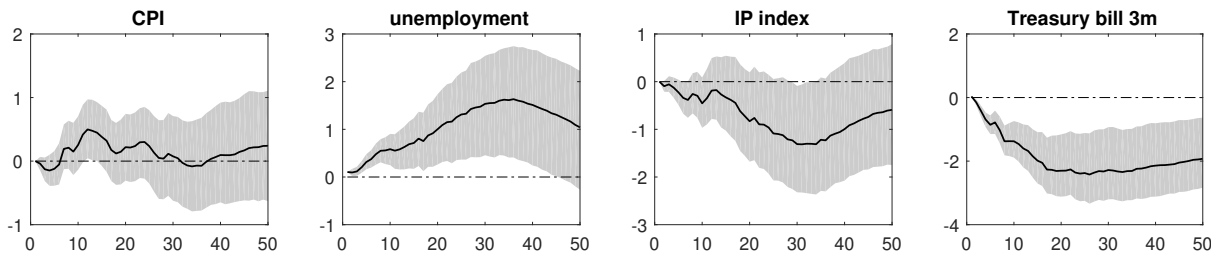


Figure C.11: Accumulated impulse responses on the observed variables  $x_t$  for the VAR-F model  
The graph shows impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals.

# Robustness Checks

## C.1 Shadow rate

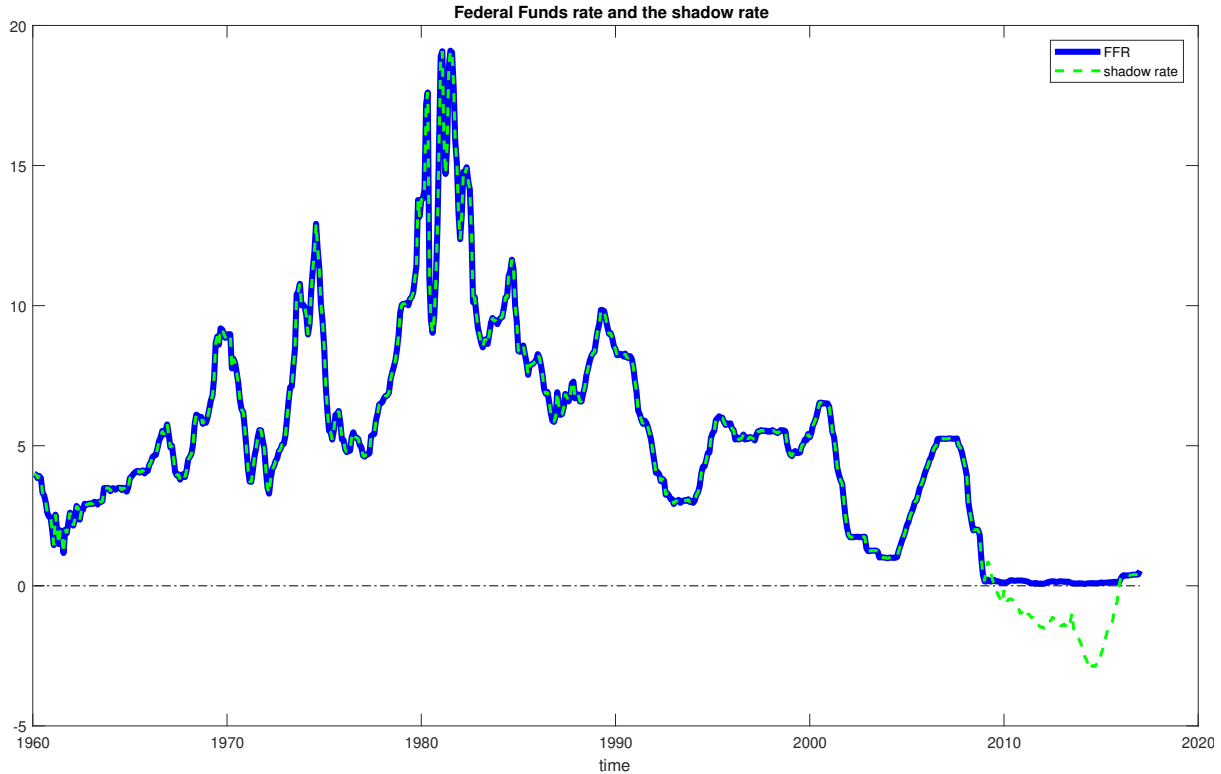


Figure C.12: Shadow and Federal funds rate  
 This graph shows the shadow rate and the Federal Funds rate. The shadow rate takes effect whenever the FFR is at the zero lower bound, i.e. when it is below 25bp.

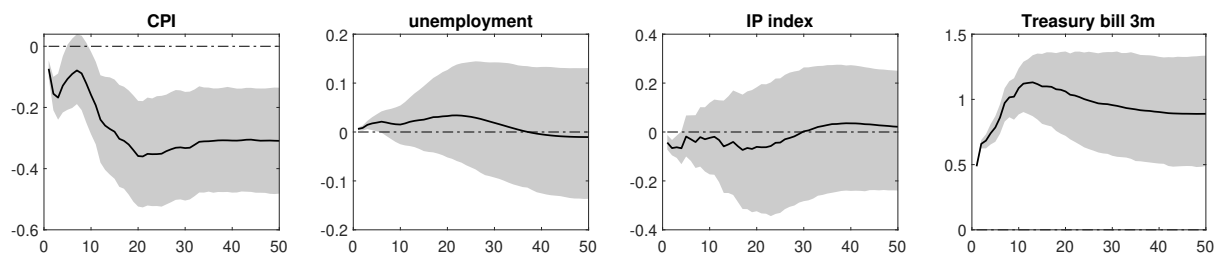


Figure C.13: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the regularized FAVAR  
 The graph shows accumulated impulse responses to a 100bp shock in the innovation of the shadow rate. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(12) model.

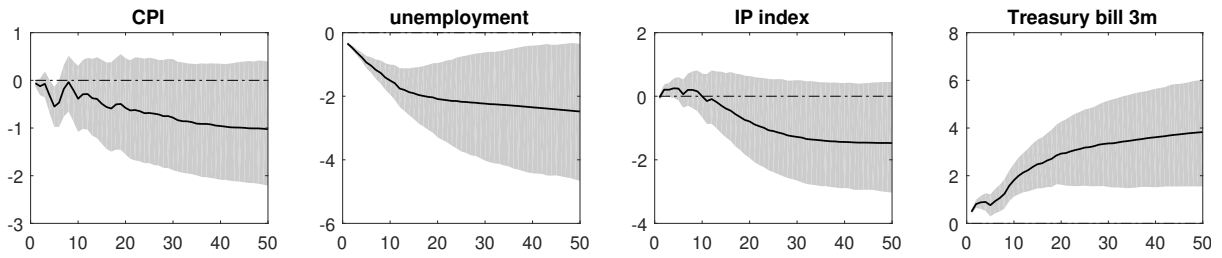


Figure C.14: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the named factor FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the shadow rate. The dashed lines correspond to 68% bootstrap confidence intervals.

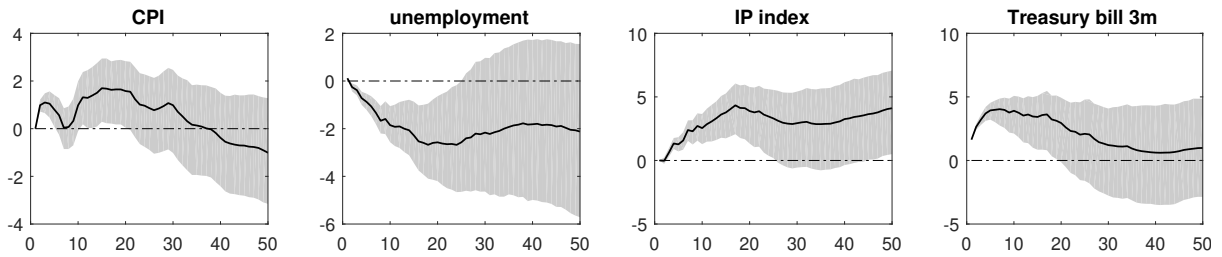


Figure C.15: Accumulated impulse responses on the observed variables  $x_t$  for the dynamic factor model

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the shadow rate. The dashed lines correspond to 68% bootstrap confidence intervals.

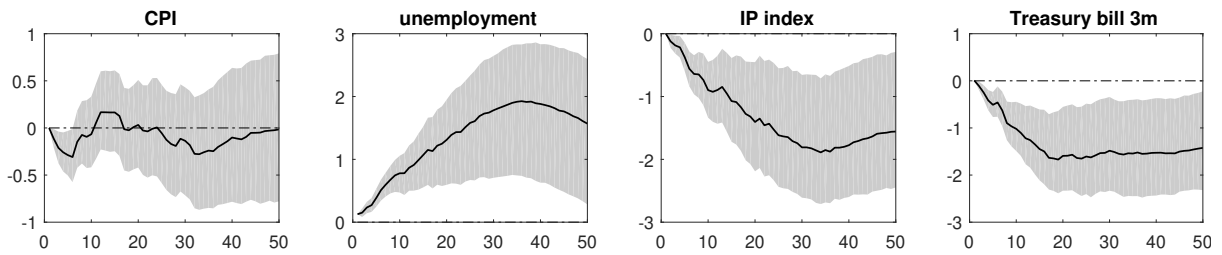


Figure C.16: Accumulated impulse responses on the observed variables  $x_t$  for the VAR-F model

The graph shows impulse responses to a 100bp shock in the innovation of the shadow rate. The dashed lines correspond to 68% bootstrap confidence intervals.

## C.2 Different initial number of factors

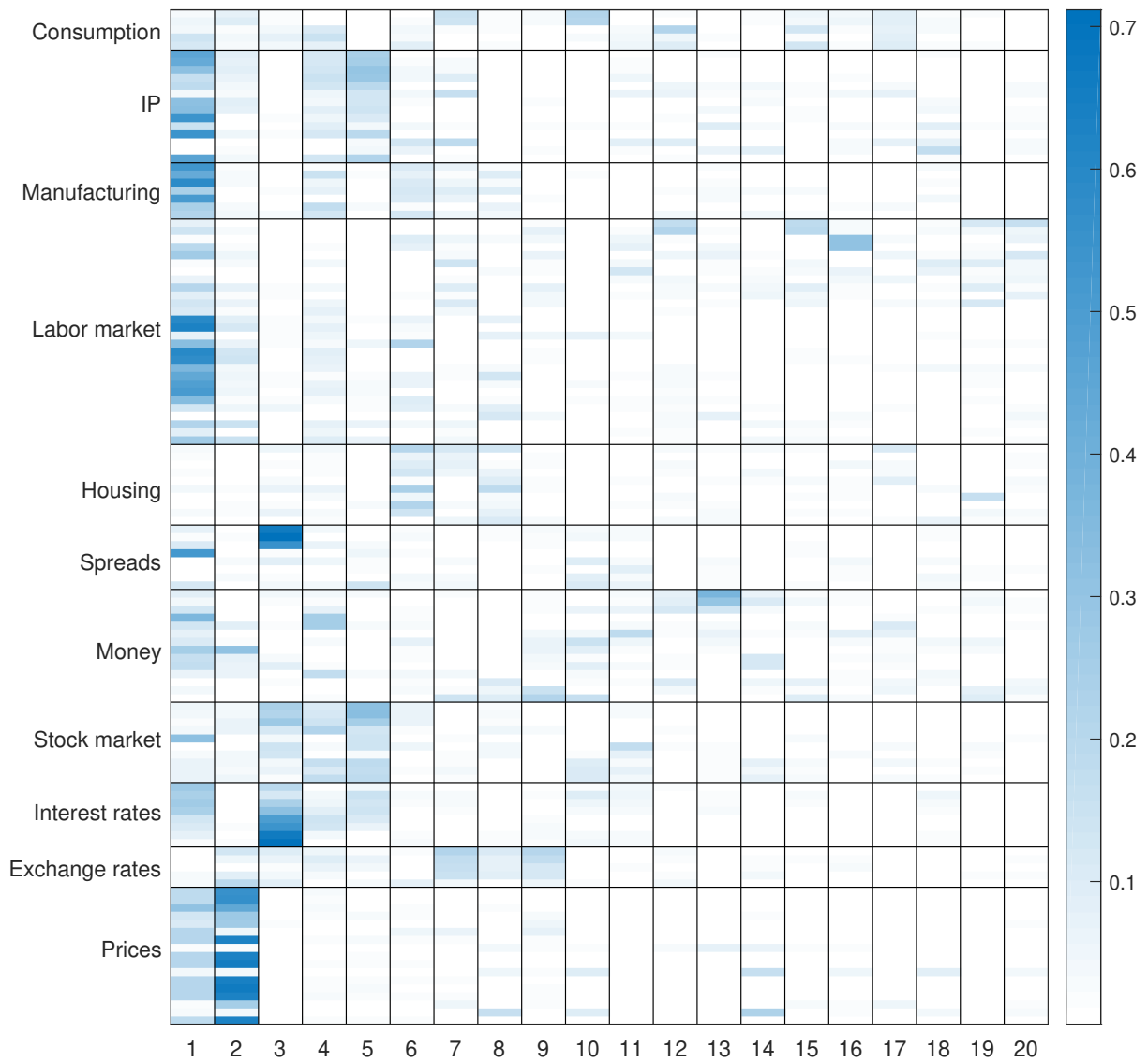


Figure C.17:  $R^2$  plot for a factor model based on principal components analysis  
This graph shows  $R^2$  for univariate regressions of the observed variables on each of the 20 factors.

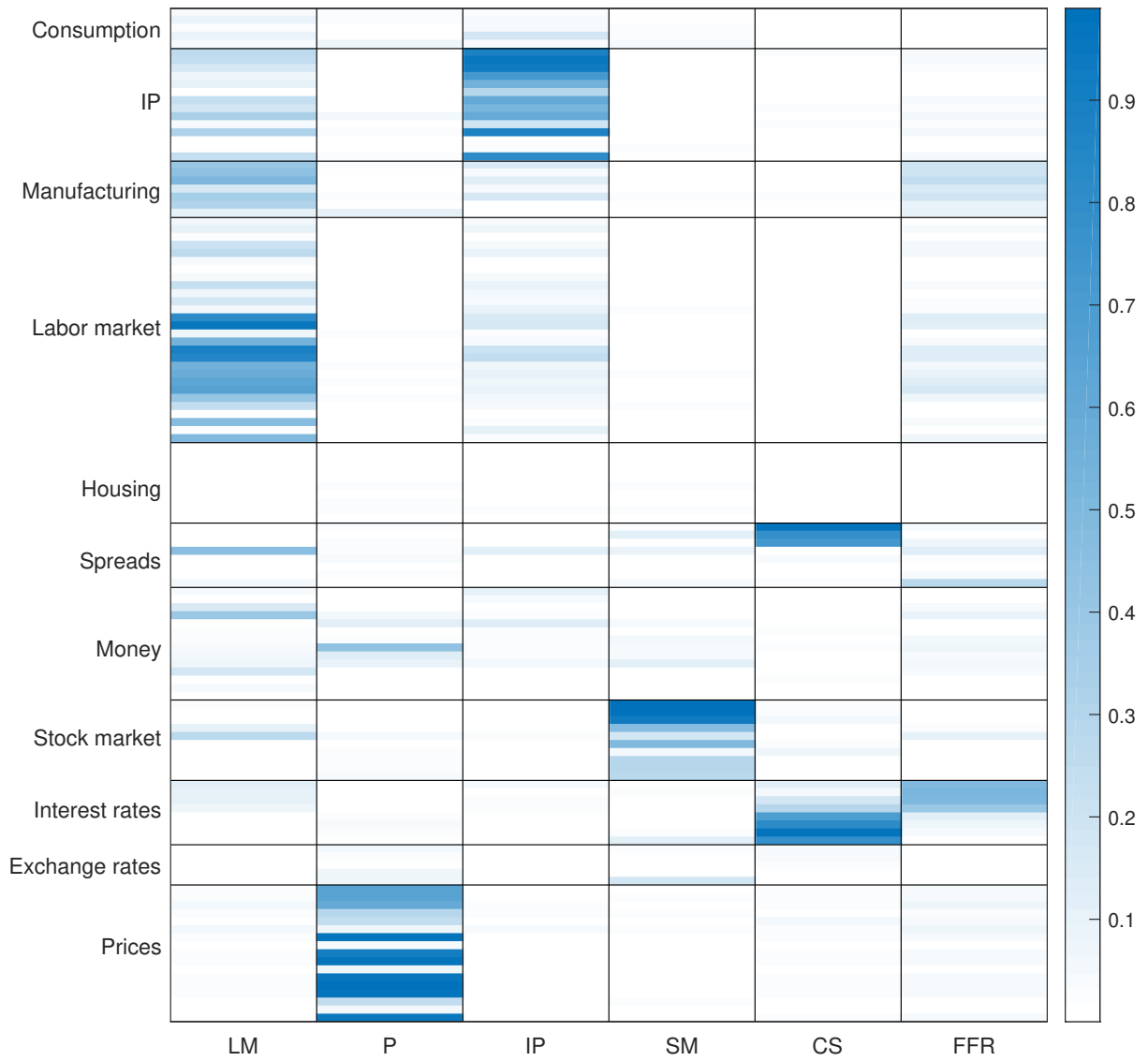


Figure C.18:  $R^2$  plot for the regularized FAVAR

This graph shows  $R^2$  for univariate regressions of the observed variables on each of the factors. The factors are abbreviated as follows: 'LM' labor market, 'P' price, 'IP' industrial production, 'SM' stock market, 'CS' credit spread and 'FFR' Federal Funds rate.

### C.3 Different lag order

#### C.3.1 RFAVAR(2) model

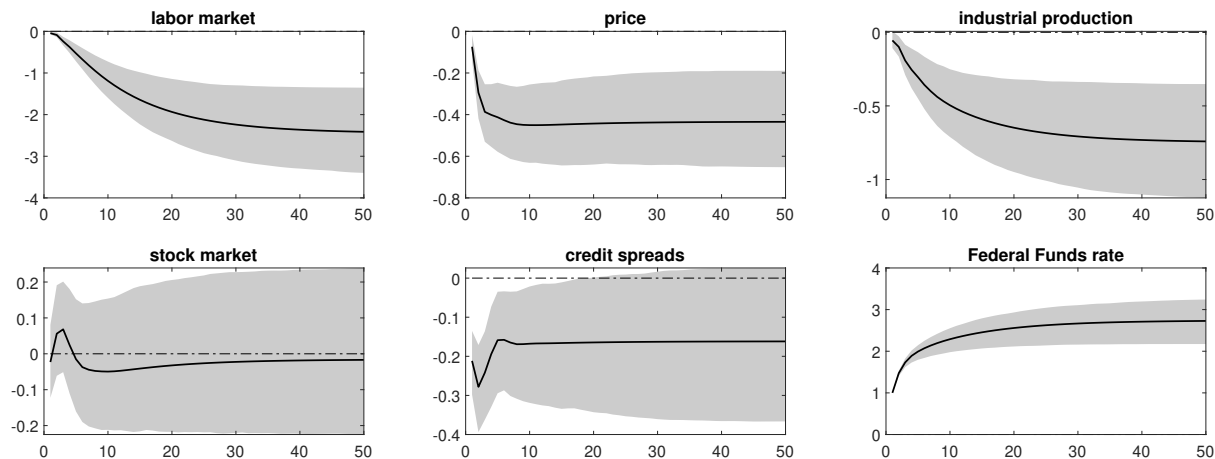


Figure C.19: Accumulated impulse responses to a monetary policy shock on the factors for the regularized FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(2) model.

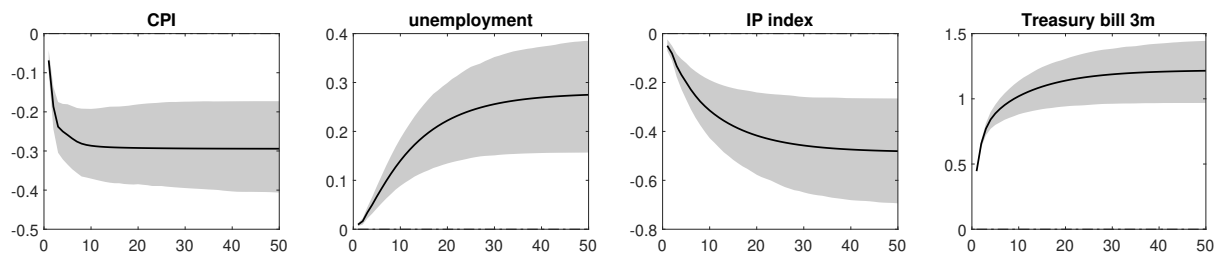


Figure C.20: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the regularized FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(2) model.



### C.3.2 RFAVAR(3) model

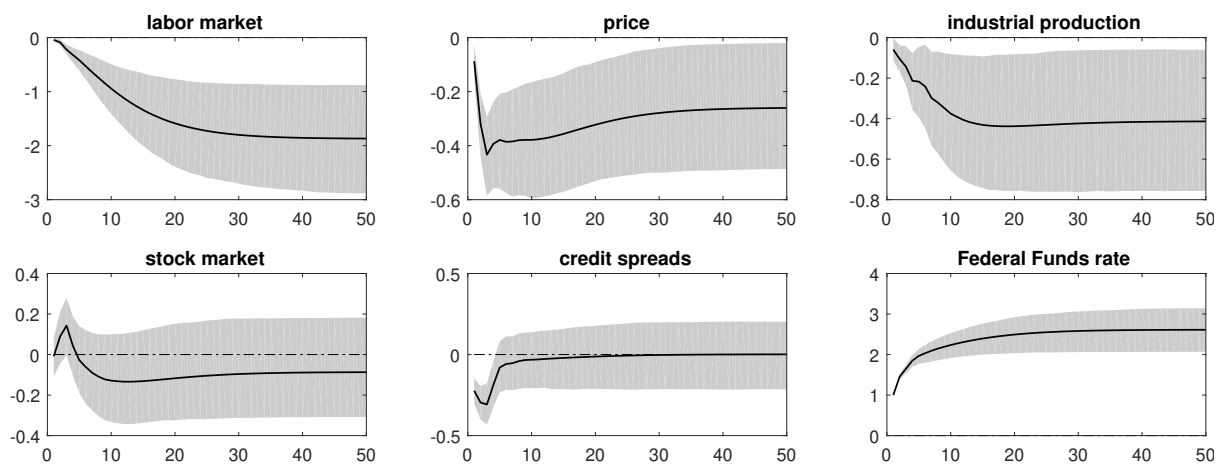


Figure C.21: Accumulated impulse responses to a monetary policy shock on the factors for the regularized FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(3) model.

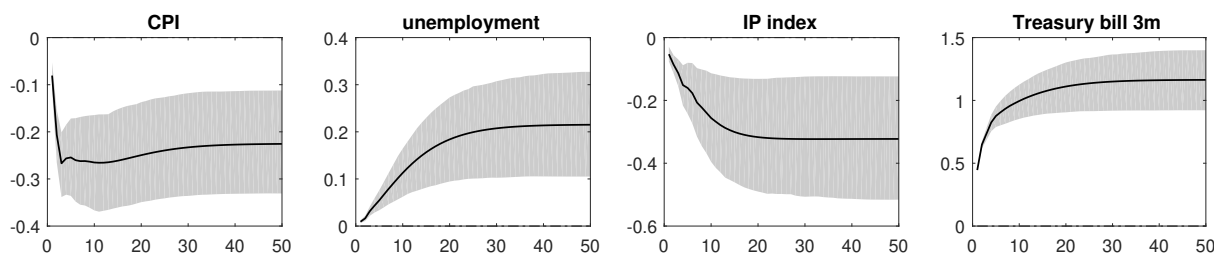


Figure C.22: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the regularized FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(3) model.

### C.3.3 RFAVAR(6) model

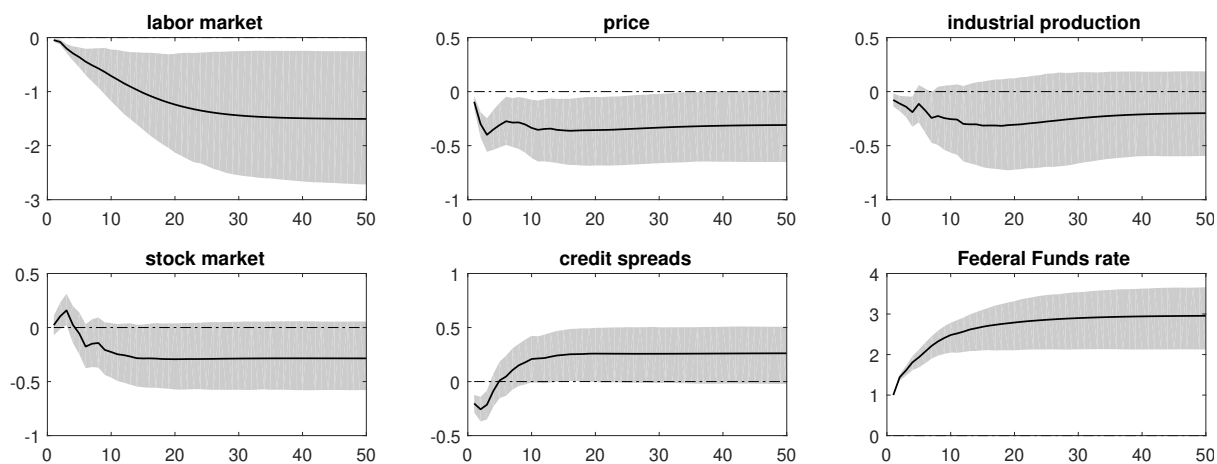


Figure C.23: Accumulated impulse responses to a monetary policy shock on the factors for the regularized FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(6) model.

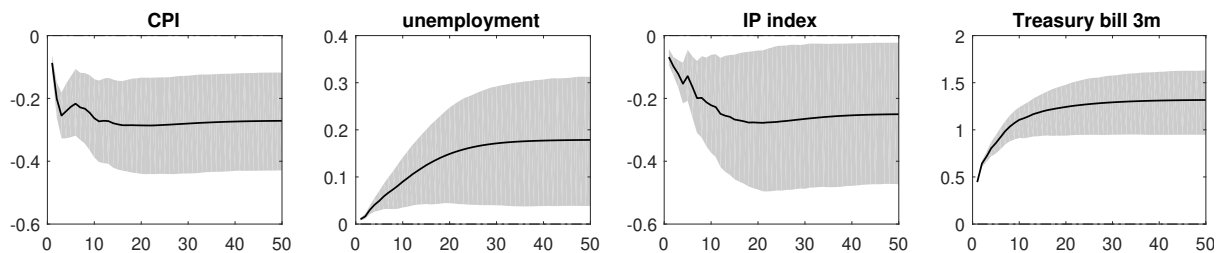


Figure C.24: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the regularized FAVAR

The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(6) model.

## C.4 Pre-crisis time span

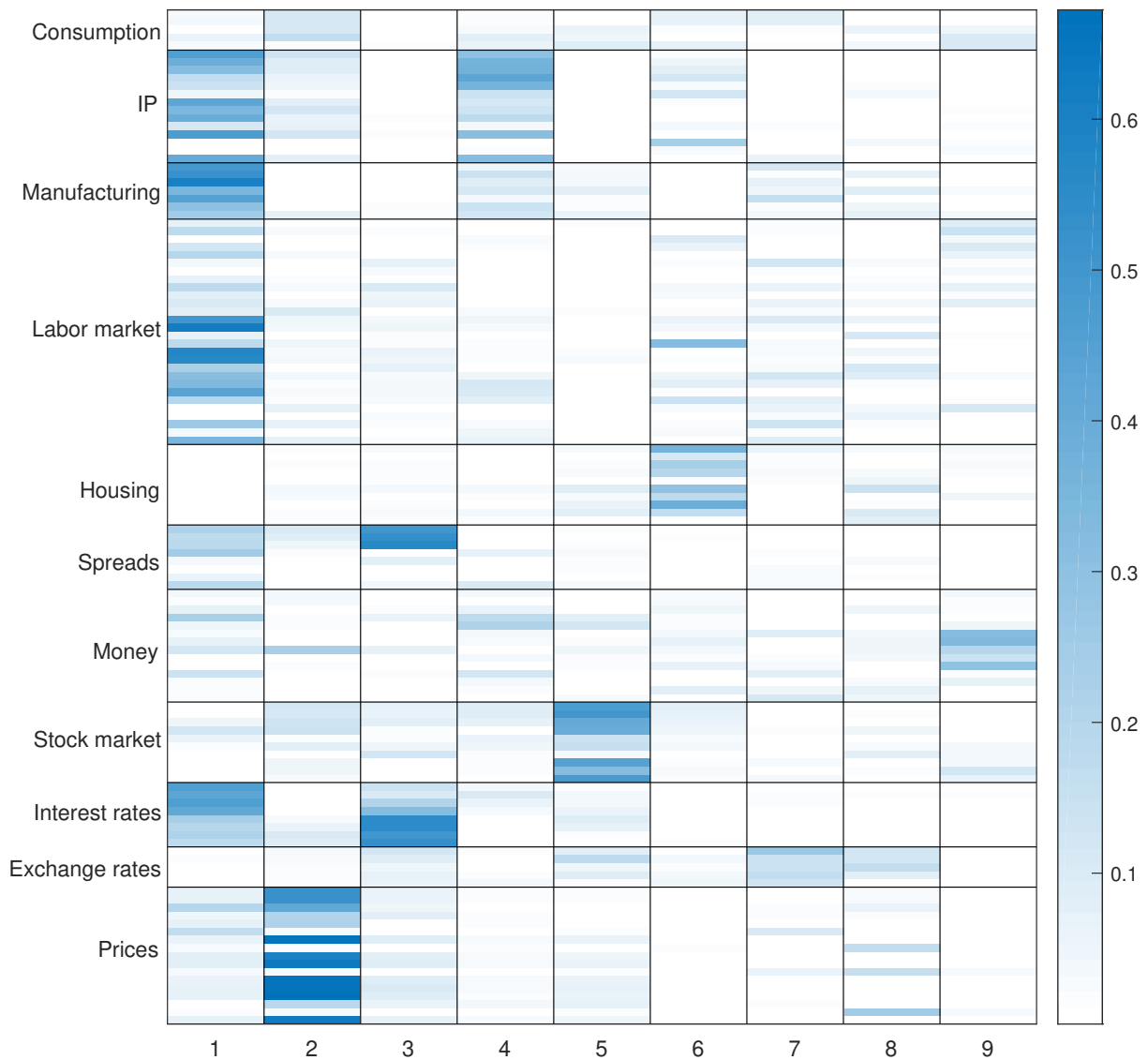


Figure C.25:  $R^2$  plot for the FAVAR based on principal components analysis

This graph shows  $R^2$  for univariate regressions of the observed variables on each of the nine factors. The number of factors is chosen based on the [Bai and Ng \(2002\)](#)  $IC_1$  criterion. The sample spans the pre-crisis period from January 1985 to December 2006.

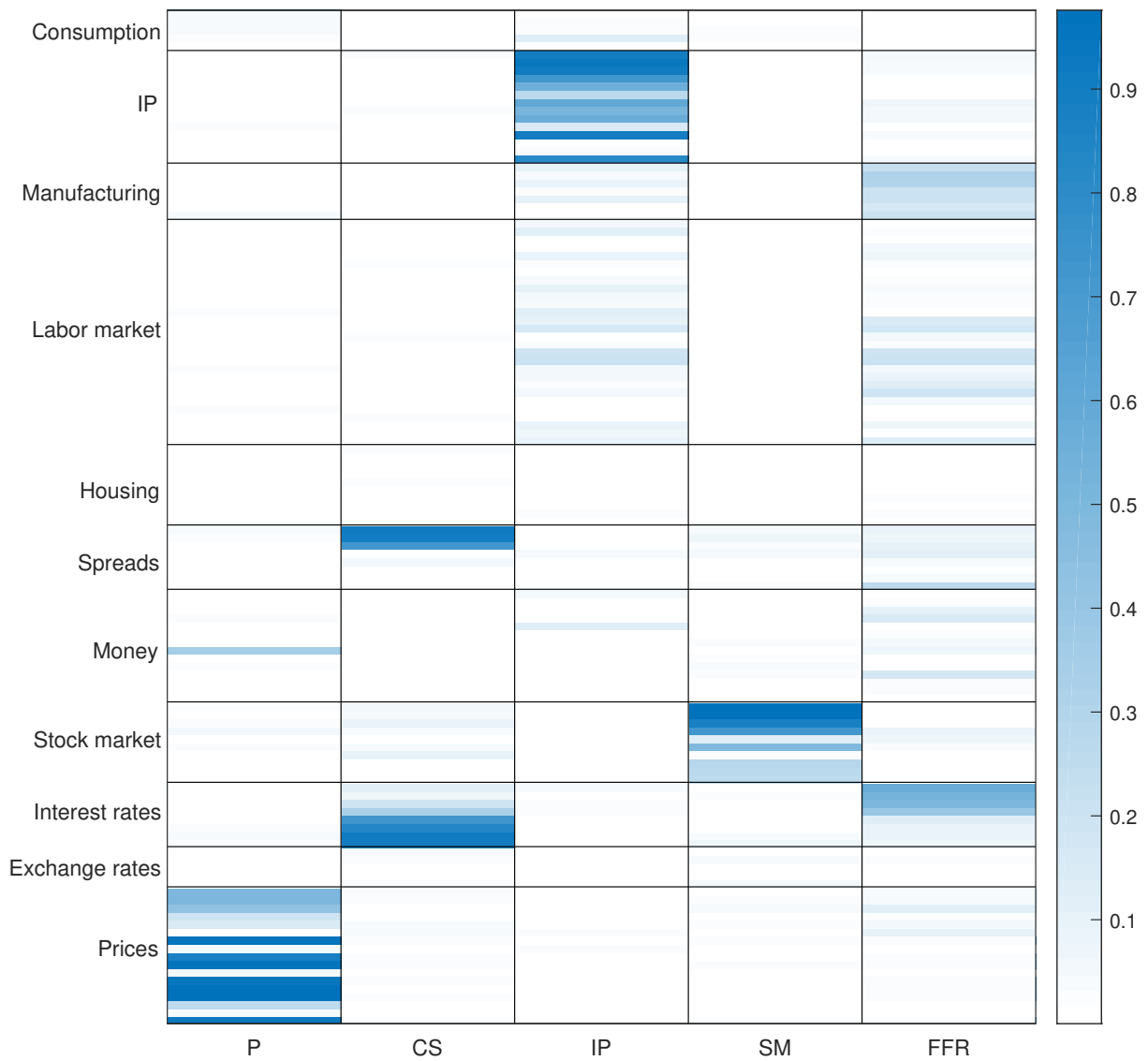
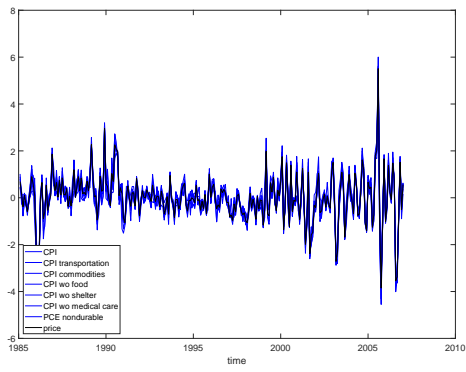
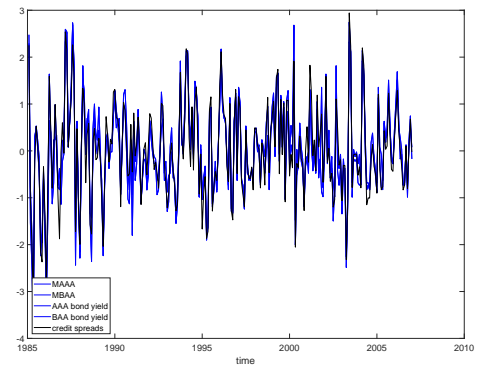


Figure C.26:  $R^2$  plot for the regularized FAVAR

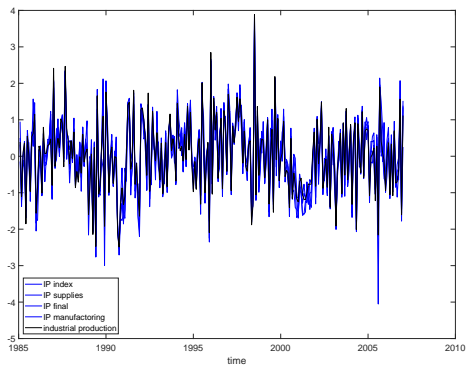
This graph shows  $R^2$  for univariate regressions of the observed variables on each of the factors. The factors are abbreviated as follows: 'P' price, 'CS' credit spreads, 'IP' industrial production, 'SM' stock market and 'FFR' Federal Funds rate. The sample spans the pre-crisis period from January 1985 to December 2006.



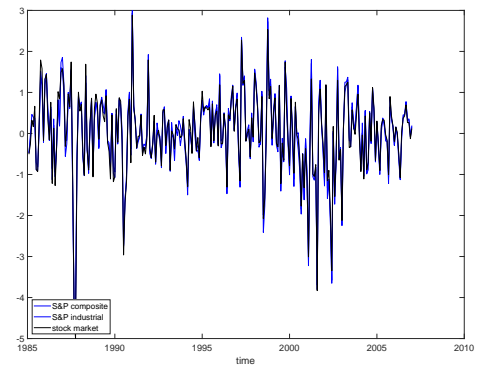
(a) Price



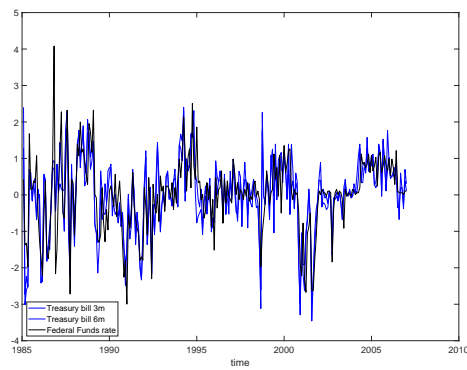
(b) Credit Spread



(c) Industrial Production



(d) Stock Market



(e) Federal Funds Rate

Figure C.27: Factor plots for the regularized FAVAR

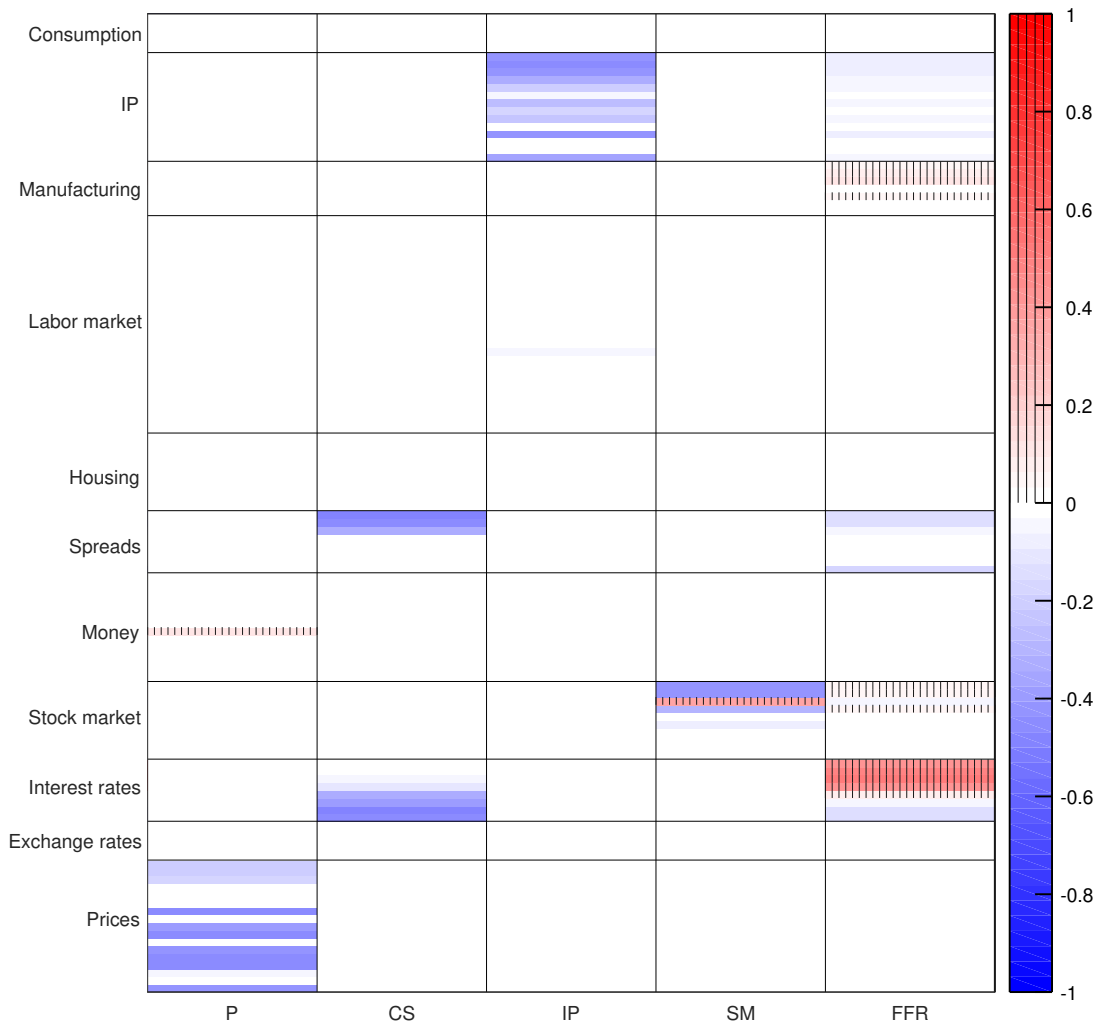


Figure C.28: Impact matrix on the observed variables  $x_t$  for the regularized FAVAR  
 This graph shows the contemporaneous impact matrix of a 100bp shock in the factor innovations to the observed time series for the regularized FAVAR model. The factors are abbreviated as follows: 'P' price, 'CS' credit spreads, 'IP' industrial production, 'SM' stock market and 'FFR' Federal Funds rate. The sample spans the pre-crisis period from January 1985 to December 2006.

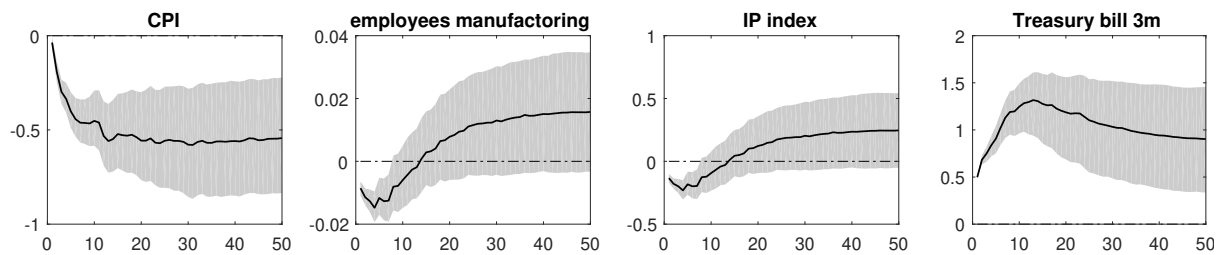


Figure C.29: Accumulated impulse responses to a monetary policy shock on the observed variables  $x_t$  for the regularized FAVAR  
 The graph shows accumulated impulse responses to a 100bp shock in the innovation of the FFR. The dashed lines correspond to 68% bootstrap confidence intervals. The impulse responses are based on a RFAVAR(12) model. The sample spans the pre-crisis period from January 1985 to December 2006.