Bruns, Martin

Working Paper

Proxy VAR models in a data-rich environment

DIW Discussion Papers, No. 1831

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Bruns, Martin (2019) : Proxy VAR models in a data-rich environment, DIW Discussion Papers, No. 1831, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
https://hdl.handle.net/10419/206948

DIW BERLIN

1831

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2019

Proxy VAR Models in a Data-Rich Environment

Martin Bruns

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Proxy VAR models in a data-rich environment

Martin Bruns*

November 22, 2019
Most recent version available at this link

Abstract

Structural VAR models require two ingredients: (i) Informational sufficiency, and (ii) a valid identification strategy. These conditions are unlikely to be met by small-scale recursively identified VAR models. I propose a Bayesian Proxy Factor-Augmented VAR (BP-FAVAR) to combine a large information set with an identification scheme based on an external instrument. In an application to monetary policy shocks I find that augmenting a standard small-scale Proxy VAR by factors from a large set of financial variables changes the model dynamics and delivers price responses which are more in line with economic theory. A second application shows that an exogenous increase in uncertainty affects disaggregated investment series more negatively than consumption series.

JEL classification: C38, E60

Keywords: Dynamic factor models, external instruments, monetary policy, uncertainty shocks

*University of East Anglia. Email: martin.bruns@uea.ac.uk. I am grateful to Helmut Lütkepohl for excellent supervision. Lutz Kilian, Haroon Mumtaz, Michele Piffer, and Barbara Rossi provided helpful suggestions. Participants at the T2M conference, the Royal Economic Society Annual Meeting, the European Economic Association Annual Meeting, the Econometric Society Winter Meeting and the ASSA meeting provided helpful comments. I thank the German Academic Scholarship Foundation for financial support. This project was conducted in large parts at the German Institute for Economic Research (DIW).

1 Introduction

Structural vector-autoregressive models have become the workhorse model of modern macroeconomics. They require two main ingredients: (i) informational sufficiency and (ii) a credible identification scheme. An informationally insufficient model implies that the reduced form innovations do not contain enough information to recover the structural shocks of interest. An invalid identification scheme renders the mapping from reduced form innovations to structural shocks hard to defend, precluding causal inference.

The issue of informational insufficiency has been addressed by augmenting small-scale VAR models by latent factors (see Bernanke et al., 2005). These latent factors are extracted from a large set of informational series and serve to alleviate informational deficiency issues. The issue of identification has been addressed via the use of Proxy VAR models (see Stock and Watson, 2012, Mertens and Ravn, 2013). These models have the advantage that, contrary to the widely used recursively identified models, they do not rely on short-run exclusion restrictions which tend to be hard to defend. Instead, Proxy VAR models employ external instruments which, if they are appropriately chosen, lead to a more credible identification scheme. However, until now the informational content of Bayesian Proxy VAR models has received little attention.

The first contribution of this paper is to propose a Bayesian Proxy Factor-augmented VAR (BP-FAVAR) model. This novel model offers a unified framework to combine a large information set with an identification strategy based on external instruments. I extend the framework proposed by Caldara and Herbst (2019) to allow for latent factors and I account for their estimation uncertainty in a consistent Bayesian framework.

The second contribution of this paper is to apply the BP-FAVAR in two settings which originally use small or medium scale VAR models and a recursive identification scheme. First, I re-visit the classical three-variate monetary policy VAR model by Christiano et al. (1999). I employ the high-frequency proxy identification scheme by Gertler and Karadi (2015) and, motivated by Caldara and Herbst (2019), I augment this small-scale model by latent factors from financial market variables to account for the important feedback effects between financial variables and the real economy. I confirm the findings of Caldara and Herbst (2019) that corporate credit spreads play an important role in the monetary transmission mechanism and help avoid the price puzzle. The use of latent factors allows for a detailed investigation of the reaction of the yield curve and, in line with economic theory, suggests that government bond spreads with a shorter maturity react more strongly to an exogenous monetary contraction than spreads with a longer

maturity. In a second application, I augment the by now classical 8-variate model by Bloom (2009) with latent factors and identify the uncertainty shock using a proxy by Stock and Watson (2012). I find that the latent factors do not strongly affect the dynamics of the system, confirming that the medium-sized Bloom (2009) model captures the most important information in the economy. Second, investigating disaggregated series, I find that unexpected exogenous increases in uncertainty affect investment more severely than consumption.

This study relates firstly to Bayesian FAVAR models such as Bernanke et al. (2005) who introduce the FAVAR model, Belviso and Milani (2006) who provide a structural interpretation of the latent factors, and Amir-Ahmadi and Uhlig (2015) who employ sign restrictions in FAVAR models. Secondly, this study relates to the Bayesian Proxy VAR literature, most directly to the small-scale Bayesian Proxy VAR model by Caldara and Herbst (2019) whose framework I extend to allow for latent factors. Other studies employing external instruments in the Bayesian paradigm are Drautzburg (2016) who estimates a narrative DSGE-VAR model, Bahaj (forthcoming) who applies high-frequency identification in a multi-country framework, and Arias et al. (2018) who propose a Proxy VAR framework amenable to an importance sampler. Lastly, Kerssenfischer (2019) investigates the relation between informational insufficiency and identification in a frequentist setting.

One challenge using factor-augmented VAR models is how to account for the estimation uncertainty of latent factors. This is difficult using bootstrap techniques in the popular approach of estimating factors via principal components (PC) (see for example Yamamoto, 2019). In addition, there are no asymptotic results justifying the use of such techniques, as pointed out by Kilian and Lütkepohl (2017). To address this issue, I exploit the state-space representation of the model and employ the algorithm by Carter and Kohn (1994), treating the latent factors as random variables and sample from their posterior distribution. This procedure is included as an additional Gibbs step in the Metropolis-within-Gibbs sampler of Caldara and Herbst (2019).

The remainder of this paper is organised as follows: Section 2 introduces the model, section 3 discusses the algorithm, priors and starting values, section 4.1 shows the monetary policy application and section 4.2 discusses the uncertainty application. The last section concludes.

2 The Bayesian Proxy FAVAR

In this section I introduce the Bayesian Proxy FAVAR model. I first describe the observation equation, the transition equation and the proxy equation. In a second step, I show how identification is achieved.

First, consider the observation equation, which shows how latent and observable factors map into informational series:

$$\boldsymbol{x}_t = \Lambda^f \boldsymbol{f}_t + \Lambda^z \boldsymbol{z}_t + \boldsymbol{\xi}_t \tag{1}$$

$$\boldsymbol{\xi}_t \sim N(\mathbf{0}, \Omega) \tag{2}$$

where $\boldsymbol{x}_t$ is a $N \times 1$ vector of observable series, $\boldsymbol{f}_t$ is a $R \times 1$ vector of latent factors, and $\boldsymbol{z}_t$ is a $K \times 1$ vector of observable factors. Importantly, $\boldsymbol{x}_t$ does not contain any of the observable factors in $\boldsymbol{z}_t$. Λ^f is a $N \times R$ matrix of factor loadings for latent factors and Λ^z is a $N \times K$ matrix of coefficients for the observable factors. $\boldsymbol{\xi}_t$ is a $N \times 1$ vector of idiosyncratic errors. In general, $\boldsymbol{\xi}_t$ can be serially correlated, i.e. $Cov(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}) \neq 0$, but they are uncorrelated across series, i.e. $Var(\boldsymbol{\xi}_t) = \Omega$ is assumed to be diagonal.

The latent factors $\boldsymbol{f_t}$ and the factor loadings λ^f and λ^z require additional normalisations. For the latent factors, I follow Stock and Watson (2016) in imposing the normalisation that $\lambda^{f\prime}\lambda^f = I$ and $Var(\boldsymbol{f_t})$ is diagonal with decreasing elements. For λ^z I follow Bernanke et al. (2005) and impose that the upper $R \times R$ block is an identity matrix. Note that these normalisations do not affect the space spanned jointly by the latent factors which is the object of interest for this study.

Next, consider the transition equation which shows the dynamic evolution of the factors. It writes as a VAR(P) of the following form:

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_t + \boldsymbol{u}_t \tag{3}$$

$$\boldsymbol{u}_t \sim N(\mathbf{0}, \Sigma), \tag{4}$$

where $\boldsymbol{y}_t = \begin{bmatrix} \boldsymbol{f}_t \\ \boldsymbol{z}_t \end{bmatrix}$ stacks latent and observable factors in a vector. The coefficient matrix $\Pi = [\Pi_1, ..., \Pi_P]$ of dimension $(R+K) \times (P(R+K)+1)$ contains the autoregressive parameters of the VAR. $\boldsymbol{w}_t = [\mathbf{1}_{R+K\times 1}; \boldsymbol{y}_{t-1}, ..., \boldsymbol{y}_{t-P}]$ stacks a constant and P lags of $\boldsymbol{y}_t$. The $(R+K) \times 1$ vector of reduced form errors, $\boldsymbol{u}_t$, is serially uncorrelated, i.e. $Cov(\boldsymbol{u}_t, \boldsymbol{u}_{t-p}) = 0 \quad \forall t = 1, ..., T, \forall p = 1, ..., \infty$. $\boldsymbol{u}_t$ are uncorrelated with all leads and lags of the idiosyncratic errors, $\boldsymbol{\xi}_t$, i.e. $Cov(\boldsymbol{u}_t \boldsymbol{\xi}_{t-j}) = 0 \ \forall j, \forall t = 1, ..., T$.

I impose structure on the on-impact effects of structural shocks by assuming that the reduced form errors map into structural shocks as:

$$\boldsymbol{u}_t = B\boldsymbol{\epsilon}_t \tag{5}$$

$$\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, I_{R+K}), \tag{6}$$

where B is a $(R+K)\times(R+K)$ matrix containing the on-impact effects of the structural shocks. Their variance is normalised to one and they are contemporaneously uncorrelated. This implies the following relation between the reduced form covariance matrix and the matrix of on-impact effects: $\Sigma = BB'$

As is well known, further restrictions beyond those implied by the covariance matrix are needed to identify B. The reason is that the data cannot discriminate between observationally equivalent representations: All B such that $BB' = \Sigma$ yield the same likelihood.

In order to identify the first column of B, which I denote by $\boldsymbol{b}$, I augment the model by a "Proxy Equation", as in Caldara and Herbst (2019).[1] It spells out the relation between structural shock and instrument and is given as:

$$m_t = \beta\epsilon_{1,t} + \sigma_\nu \nu_t \tag{7}$$

$$\nu_t \sim N(0,1), \tag{8}$$

where m_t is a scalar instrument correlated with the shock of interest, $\epsilon_{1,t}$. The shock of interest is ordered first, without loss of generality. Furthermore, m_t is orthogonal to all other shocks, $\boldsymbol{\epsilon}_{-1,t}$, i.e. $E(m_t\boldsymbol{\epsilon}_{-1,t}) = 0\ \forall t$, where $\boldsymbol{\epsilon}_{-1,t}$ stands for a vector containing all but the first shock. In other words, the instrument needs to be both relevant and exogenous in order to be appropriate for identification. β captures the structural relationship between instrument and shock, while ν_t captures any noise contained in the instrument. The higher its variance, σ_ν^2, the less information the instrument contains about the shock of interest.

Since $\boldsymbol{\epsilon}_1$ is obtained from reduced form errors, identification depends heavily on

[1] Unlike their case, however, identification focuses on the on-impact effects of the shocks rather than on the contemporaneous relations of the variables included in the model. Put differently, the model imposes structure on B, rather than on B^{-1}. Caldara and Herbst (2019) estimate a so-called A-model (see Kilian and Lütkepohl, 2017 for a discussion). The A-model specification is appropriate given their aim of identifying a monetary policy equation. In the present context, however, interest lies on the on-impact effects of structural shocks.

the model specification. Therefore, one should pay close attention to which variables are included in the model since an omitted variable bias translates into biases in the identified structural shocks. Augmenting the model with latent factors can help alleviate this problem without taking a stand on which of a potentially large set of observational series need to be included.

3 Inference

The parameters of the BP-FAVAR model are estimated via a Metropolis-within-Gibbs sampler while the latent factors are drawn using the algorithm by Carter and Kohn (1994). In this section, I discuss the choice of priors and sketch an algorithm to generate draws from the posterior distribution.

3.1 Priors and Starting Values

Block 1: Observation Equation For the parameters of the observation equation, $\Lambda = [\Lambda^f \quad \Lambda^z]$ and Ω, I employ equation-by-equation normal-inverse Gamma priors of the form

$$\boldsymbol{\omega}_{ii} \sim IG(sc^*, sh^*) \tag{9}$$

$$\boldsymbol{\lambda}_i | \boldsymbol{\omega}_{ii} \sim N(\boldsymbol{\mu}^*_{\lambda,i}, \boldsymbol{\omega}_{ii} M_i^{*-1}), \tag{10}$$

where $\boldsymbol{\lambda}_i$ is the i-th row of Λ and $\boldsymbol{\omega}_{ii}$ is the i-th diagonal element of Ω. The prior parameters $sc^*, sh^*, \boldsymbol{\mu}^*_{\lambda,i}$, and M_i^{*-1} are chosen to imply diffuse priors (see Appendix A.2 for details). The algorithm starts from the Principal Components estimates for Λ and $F = [\boldsymbol{f}_1, ..., \boldsymbol{f}_T]$ and OLS estimates for Ω.

Block 2: Transition Equation Given a draw of factors, $\boldsymbol{y}_t$ follows a standard $VAR(P)$ model. Therefore, one can employ a version of the Minnesota/ Litterman prior and specify independent normal-inverse Wishart priors:

$$vec(\Pi) \sim N(\boldsymbol{\mu}^*_\Pi, V^*_\Pi) \tag{11}$$

$$\Sigma \sim IW(S^*, \tau^*), \tag{12}$$

where $vec(\cdot)$ is the vectorisation operator that stacks the columns of a matrix one underneath the other into a vector. I set the prior mean for the autoregressive coefficients,

$\boldsymbol{\mu}_{\Pi}^{*} = \mathbf{0}$, implying shrinkage towards white noise processes. This choice is motivated by the stationarity transformations of both $\boldsymbol{z}_t$ and $\boldsymbol{x}_t$. I set the prior variance of $vec(\Pi)$, V_{Π}^{*}, in line with the Minnesota/Litterman prior (see Litterman, 1986). Lastly, I set the improper priors $S^{*} = 0.0001\boldsymbol{I_{K+R}}$ and $\tau^{*} = K + R$, implying a flat prior in the dimension of Σ (see Appendix A.3).

Block 3: Proxy Equation The parameters of the proxy equation are sampled conditional on the parameters of the transition equation and follow Caldara and Herbst (2019) closely. For β I employ a normal distribution

$$\beta \sim N(\mu_{\beta}^{*}, \sigma_{\beta}^{*2}) \tag{13}$$

with $\mu_{\beta}^{*} = 0$ and $\sigma_{\beta}^{*} = 1$.

For σ_{ν} I consider two types of priors: First, I set an uninformative inverse Gamma distribution using

$$\sigma_{\nu} \sim IG(sc_{\nu}^{*}, sh_{\nu}^{*}) \tag{14}$$

with $sc_{\nu}^{*} = 2$ and $sh_{\nu}^{*} = 0.02$. Second, to the extent that the econometrician is confident in the relevance of the proxy, she can employ what Caldara and Herbst (2019) refer to as the "high-relevance" prior and set

$$\sigma_{\nu} = 0.5 std(m_t), \tag{15}$$

implying the dogmatic view that only half the variation of the proxy can be explained by measurement error (see Appendix A.5).

The prior for $\boldsymbol{b}$ is implicit in the above priors given that $\boldsymbol{b}$ is computed as $\boldsymbol{b} = chol(\Sigma)Q_{\cdot,1}$ where $Q_{\cdot,1}$ is the first column of a draw from the uniform Haar distribution (see Rubio-Ramirez et al., 2010 for a discussion).

3.2 Sketch of the Algorithm

The challenge is to generate draws from the joint posterior distribution of latent factors and parameters given the informational series, the observable factors and the instrument:

$$p(\Pi, \Sigma, \Lambda^{f}, \Lambda^{z}, \Omega, \beta, \sigma_{\nu}, \boldsymbol{b}, F | X, Z, \boldsymbol{m}). \tag{16}$$

Instead of attempting to draw directly from this potentially highly non-linear and irregularly shaped distribution, draws are generated from blocks of conditional distributions

in the following steps:

Step 1 **Draw from the conditional posterior of F**

In order to generate draws from

$$p(F|\Pi, \Sigma, \Lambda^f, \Lambda^z, \Omega, \beta, \sigma_\nu, \boldsymbol{b}; X, Z, \boldsymbol{m}), \tag{17}$$

I employ the Carter-Kohn algorithm described in Carter and Kohn (1994) and Frühwirth-Schnatter (1994) (see Appendix A.1).

Step 2 **Draw from the conditional posterior of $\{\Lambda^f, \Lambda^z, \Omega\}$**

To generate draws from

$$p(\Lambda^f, \Lambda^z, \Omega|F, \Pi, \Sigma, \beta, \sigma_\nu, \boldsymbol{b}; X, Z, \boldsymbol{m}), \tag{18}$$

note that, given a draw of the latent factors, F, the observation equation (1) is a system of independent linear equations. Therefore, the parameters can be estimated using well-established results on single-equation models (see Appendix A.2).

Step 3 **Draw from the conditional posterior of Π, Σ**

To generate draws from

$$p(\Pi, \Sigma|\Lambda^f, \Lambda^z, \Omega, F, \beta, \sigma_\nu, \boldsymbol{b}, X, Z, \boldsymbol{m}), \tag{19}$$

note that, conditional on a draw of the factors, the model can be considered a variant of Caldara and Herbst (2019). The conditional posterior in (19) differs from standard reduced form Bayesian VAR models given that $\boldsymbol{m}$ is part of the conditioning set. To address this challenge, I proceed in two steps: first, I generate proposal draws from $p(\Pi, \Sigma|\Lambda^f, \Lambda^z, \Omega, F, \beta, \sigma_\nu, \boldsymbol{b}, X, Z)$ (note that $\boldsymbol{m}$ is not part of the conditioning set) using well-established results on independent normal-inverse Wishart distributions. In a second step, these proposal draws are mapped into the target distribution (19) via a Metropolis-Hastings step (see Appendix A.4). In practice, the proposal distribution is highly similar to the target, leading to an acceptance rate close to 1. This indicates that the instrument does not contain much information about the reduced form parameters, a result also found in Caldara and Herbst (2019).

Step 4 **Draw from the conditional posterior of $\boldsymbol{b}$**

To generate draws from

$$p(\boldsymbol{b}|\Lambda^f, \Lambda^z, \Omega, F, \Pi, \Sigma, \beta, \sigma_\nu; X, Z, \boldsymbol{m}), \tag{20}$$

I adapt the procedure by Caldara and Herbst (2019) to my setting. First, I generate proposal draws b^{cand} as the first column of $chol(\Sigma)Q^{cand}$, where Q^{cand} is an orthogonal matrix. I then map this candidate draw into a draw from the target distribution (20) using a Metropolis-Hastings step (see Appendix A.4).

Step 4 **Draw from the conditional posterior of β, σ_ν**

Lastly, to generate draws from

$$p(\beta, \sigma_\nu|\boldsymbol{b}, \Lambda^f, \Lambda^z, \Omega, F, \Pi, \Sigma; X, Z, \boldsymbol{m}) \tag{21}$$

I employ well-established results on linear regression models for the proxy equation (7). Note that the structural shock can be generated as $\epsilon_{1,t} = chol(\Sigma)Q\boldsymbol{u}_t$.

Compared to the framework in Caldara and Herbst (2019), there are three main differences: First, the above approach includes latent factors and accounts for their estimation uncertainty in a consistent Bayesian setup, thereby allowing the factors to affect the dynamics of Z. Second, the BP-FAVAR allows for inference on a large number of potentially disaggregated series, X, through the mapping from factors to informational series. Third, this approach employs the proxy in order to identify the on-impact effects of structural shocks, $\boldsymbol{b}$, rather than the contemporaneous relations among variables, as in Caldara and Herbst (2019).

Compared to the common approach of replacing F by their Principal Components estimates (see Stock and Watson, 2016 for a review), this approach has two advantages: First, it captures the sampling uncertainty of factors within a coherent Bayesian framework, rather than treating them as observed data. Bootstrap inference could be an alternative. However, as shown by Yamamoto (2019), bootstrap inference in frequentist factor models is far from trivial. A Bayesian approach, on the other hand, offers a unified way of summarising the uncertainty of the model, as pointed out by Huber and Fischer (2018). The joint posterior summarises estimation uncertainty in both the parameters and the latent factors.

In addition, this approach allows for Bayesian shrinkage of the parameter space of

the transition equation. This might seem unnecessary given that the factors already reduce the dimensionality of the estimation problem. However, if, as is often the case in empirical applications, the number of observable factors, K, or the lag length, P, is large, dimensionality issues still arise and can be alleviated using Bayesian shrinkage.

The next section illustrates the use of this algorithm in two applications.

4 Two applications to real data

Having presented the model setup and the algorithm in a generic framework, this section applies the BP-FAVAR to two classical questions: (i) What are the effects of monetary policy shocks? (ii) What are the effects of uncertainty shocks? Both questions have originally been addressed using small/medium scale VAR models with a recursive identification scheme (see Christiano et al., 1999 and Bloom, 2009). I show how a combination of a large information set with recently introduced proxies can change the model dynamics and allow inference on selected disaggregated series.

4.1 Monetary policy shocks

Uncovering the effects of monetary policy shocks within VAR models and their effects on the real economy is by now a classical endeavour. Within this context, two challenges arise: Finding a valid identification scheme and avoiding biases stemming from informational insufficiency (see e.g. Kerssenfischer, 2019 for a discussion).

With regards to the information set, Bernanke et al. (2005) uncover potential biases stemming from a misalignment between the central bank's and economic agent's information sets and recommend augmenting standard small scale VAR models by additional information. More recently, Caldara and Herbst (2019) stress the importance of financial variables in the monetary transmission and augment a small-scale Proxy VAR by a measure of corporate credit spread.

Concerning the identification scheme, the high-frequency identification approach by Gertler and Karadi (2015) has become standard. It relies on variations in financial spreads around Federal Open Market Committee meetings. Proxies based on these variations have proven to be strong external instruments (see Ramey, 2016).

In this section, I re-visit the issue of informational deficiency in the classical small-scale monetary VAR model by Christiano et al. (1999) using the high-frequency proxy proposed by Gertler and Karadi (2015). Instead of adding a single credit spread measure

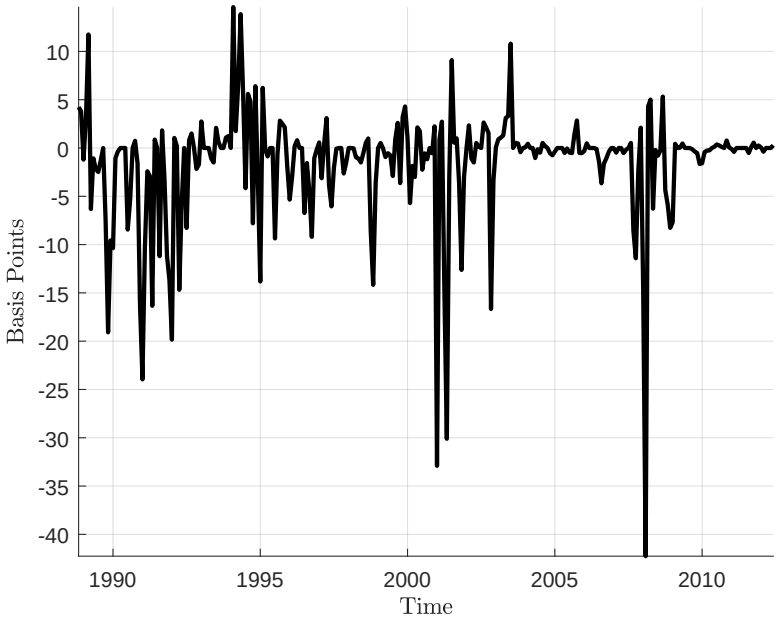

Figure 1: ***Proxy for monetary policy shocks.*** *Surprise in the current federal funds futures rate. See Gertler and Karadi (2015) for details.*

as in Caldara and Herbst (2019), I augment the small-scale VAR model by latent factors from a number of financial series to account for the reaction of central banks to broad financial conditions.

4.1.1 Data, specification and results

I estimate the BP-FAVAR using for $\boldsymbol{z_t}$ the three variables employed by Christiano et al. (1999): (i) The federal funds rate in levels as a policy variable, (ii) the consumer price index in log-differences, and (iii) industrial production in log-differences. I extract latent factors from $\boldsymbol{x_t}$. $\boldsymbol{x_t}$ contains 16 interest rate and spread series from the monthly FRED dataset by McCracken and Ng (2016) (see Table 6 in Appendix C for a detailed description).[2] As a proxy I use the surprises in the current federal funds future rate, as computed by Gertler and Karadi (2015). Figure 1 plots the proxy. The sample length is constrained by this proxy and runs from 1988M11 to 2012M6. I follow Gertler and Karadi (2015) in setting a lag length of 12 months.

One concern is the measurement of the monetary policy stance. Traditionally, the federal funds rate is considered to be the policy tool of the central bank. However, given that it was constrained by the zero lower bound in the period 2009M1 to 2015M11, this variable cannot serve as an indicator of the policy stance. This is why, starting from 2009M1, I replace the federal funds rate by the shadow rate as computed by Wu and Xia (2016) which is based on a term structure model and often considered a better reflection

[2]Data set available at https://research.stlouisfed.org/econ/mccracken/fred-databases/

of the policy stance during the zero lower bound period than the federal funds rate. Figure 12 in Appendix C shows the shadow rate.

The choice of the number of factors is based on two criteria: The criterion proposed by Bai and Ng (2002) and a scree plot. These two criteria suggest setting $R = 2$ (see Appendix C). I extract factors using a Kalman filtering technique and compare these to the Principal Components factors given the OLS estimates of $\{\Lambda, \Omega, \Pi, \Sigma\}$. Table 1 shows that these two estimation strategies yield highly correlated results, suggesting that the Kalman filter works well. Compared to the PC estimate, the Kalman filter, however, has the advantage of accounting for the estimation uncertainty in the factors, when the remaining parameters are drawn from their posterior distributions rather than replaced by OLS estimates. Figure 2 shows the posterior median and 90% bands of the latent factors estimated using a Kalman filter together with the PC estimate. The posterior distribution of the latent factors follows the PC estimates closely.

	K1	K1	PC1	PC2
K1	1.00			
K2	-0.05	1.00		
PC1	**0.90**	-0.26	1.00	
PC2	0.20	**0.93**	0.00	1.00

Table 1: *Correlation between factors estimated via Kalman filter (top) and Principal Components (bottom) conditional on ordinary least squares estimates of* $\{\Lambda, \Omega, \Pi, \Sigma\}$.

Figure 3 shows the updating of the signal-to-noise ratio β/σ_ν. It measures how much information the instrument contains about the shock of interest compared to its noise, as pointed out by Caldara and Herbst (2019). I employ their high-relevance prior, setting $\sigma_\nu = 0.5 std(m_t)$. Figure 3 suggests that the instrument is informative about the shock of interest given that virtually all of the posterior probability mass of the signal-to-noise ratio is to one side of 0.

Figure 4 shows the updating of $\boldsymbol{b}$, the impact of effect of a one standard deviation monetary policy shock. The prior distribution is not available in closed form but is implicit in the prior distributions of Σ, β and σ_ν. Prior draws are generated imposing the prior mean for β, i.e. setting $\beta = 0$, so that all rotation vectors, $Q_{\cdot,1}$, are accepted with equal probability. A draw from the prior of $\boldsymbol{b}$ is computed as follows:

- Draw Σ^{prior} from its prior inverse Wishart distribution
- Draw $Q_{\cdot,1}^{prior}$ as the first column of a draw from the uniform Haar distribution

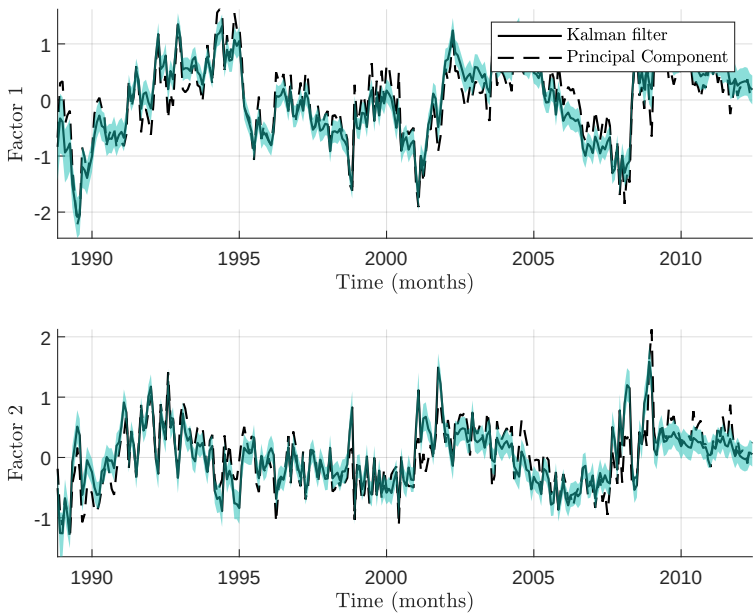

Figure 2: ***Posterior of latent factors.*** *Median posterior draw of latent factors estimated via Kalman filter (solid line). 90% bands (dotted) and Principal Components estimate (dashed).*

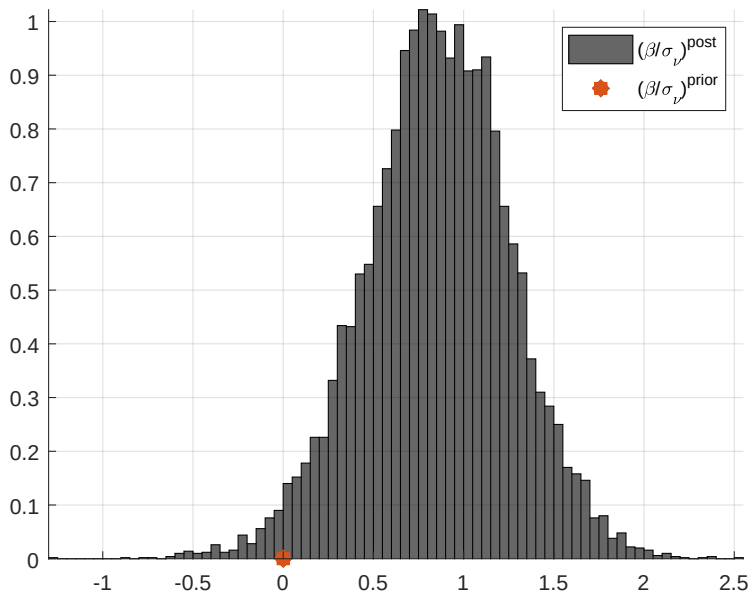

Figure 3: ***Signal-to-noise ratio.*** *Posterior distribution of β/σ_ν (histogram) together with the prior median (star).*

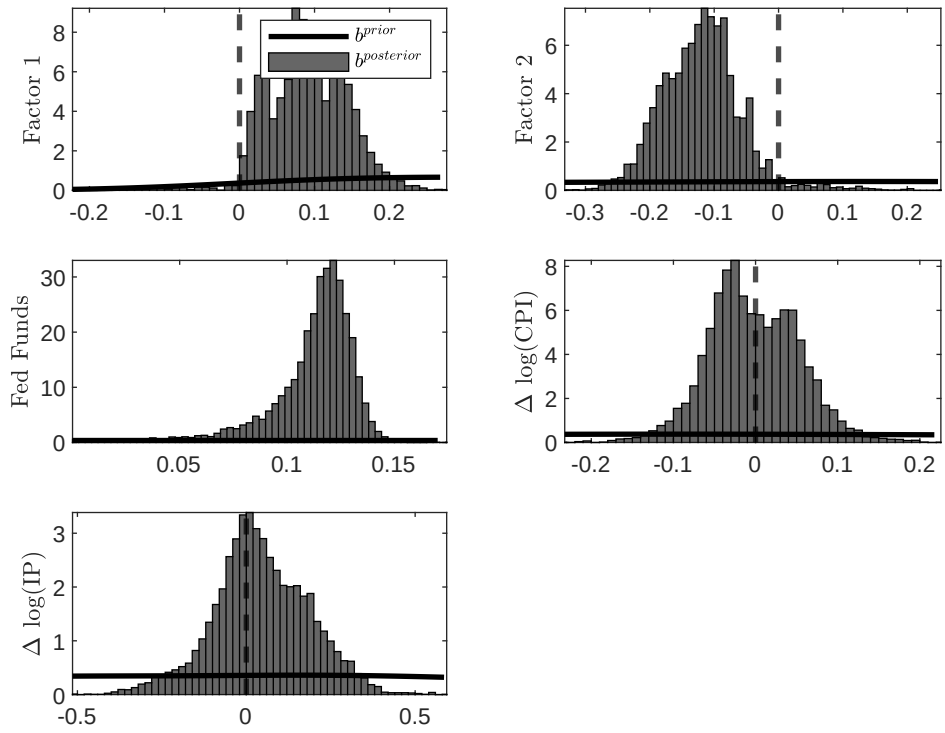

Figure 4: ***Updating of b.*** *Priors (solid line) and posterior (histogram) of* $\boldsymbol{b}$*. Prior draws are computed from the distribution implicit in the priors for* Σ*,* β *and* σ_ν*.*

- Compute $\boldsymbol{b}^{prior} = chol(\Sigma)Q_{\cdot,1}$.

As pointed out by Baumeister and Hamilton (2015), a uniform prior on $Q_{\cdot,1}$ does not necessarily imply a uniform distribution over the structural parameters of interest, which in this case are the elements of $\boldsymbol{b}$. In the present case, the prior of $\boldsymbol{b}$ implicit in the specification is flat in the relevant parameter space suggesting that the posterior is driven primarily by information from the data. Both latent factors react on impact to the monetary policy shock. A one standard deviation monetary policy shock moves the federal funds rate by roughly 10 basis point, an amount slightly smaller than in models excluding the zero lower bound period. For inflation and industrial production growth the impact effect is ambiguous suggesting that these quantities do not react strongly within the first month.

Figure 5 shows the impulse responses to a monetary policy shock which increases the federal funds rate by 25 basis points. In the small-scale model excluding latent factors, there is evidence for the price-puzzle, i.e. an increase in inflation following a monetary contraction. This puzzle is often considered evidence for an omitted variable bias (see e.g. Bernanke et al., 2005). Adding two latent factors from financial series resolves the price puzzle.

Employing latent factors allows to make inference about a wide range of financial

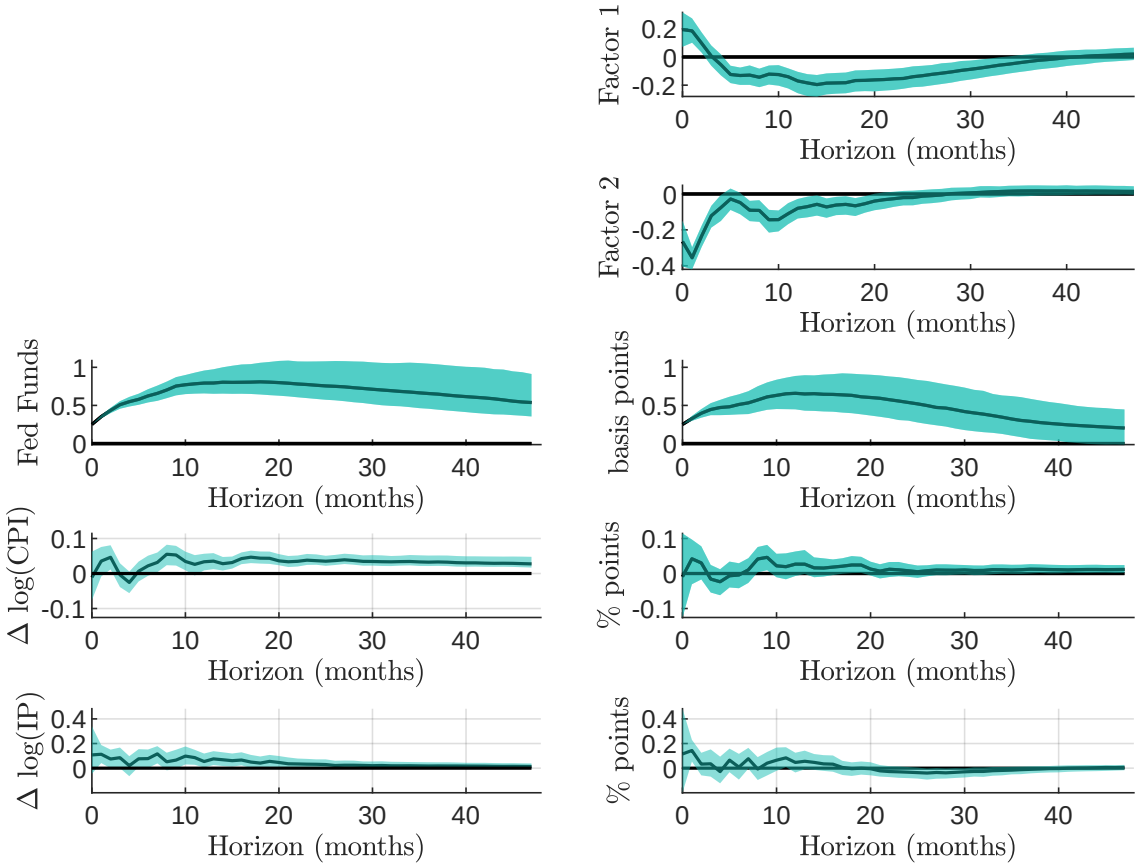

Figure 5: ***Impulse Response Functions of observables and latent factors.*** *Point-wise (non-cumulated) median impulse responses (solid line) with 68% bands. Small-scale VAR (left column) excludes latent factors. BP-FAVAR (right column) includes two latent factors from financial series.*

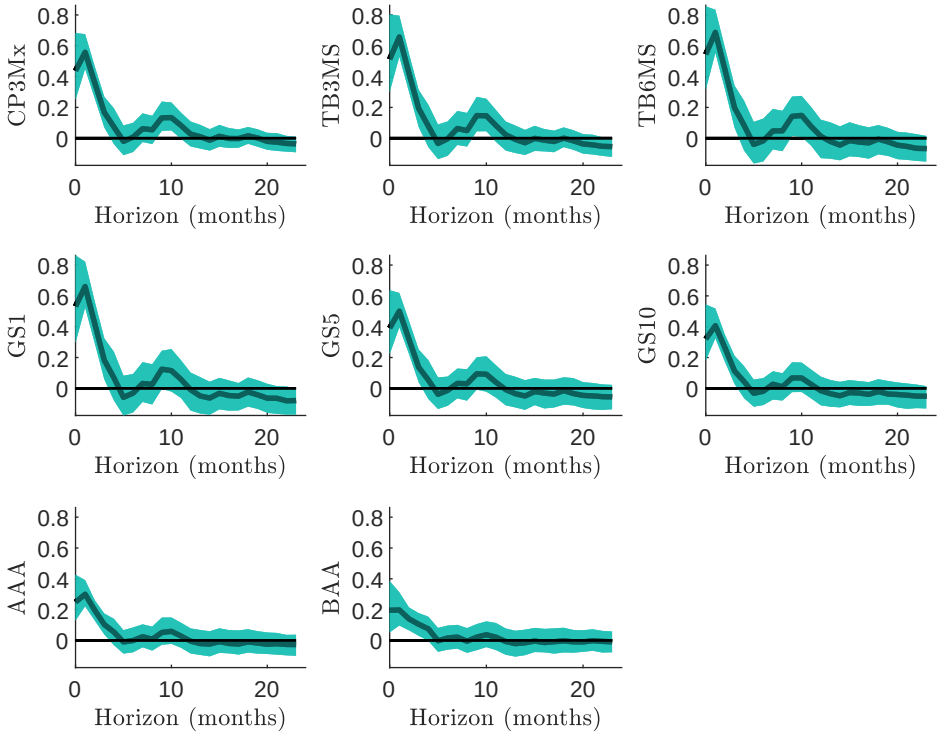

Figure 6: ***Impulse Response Functions of informational series.*** *Point-wise (non-cumulated) median impulse responses (solid line) with 68% bands. See Appendix C table 6 for a detailed variable description.*

series. Figure 6 shows the response of spreads with different maturities to a monetary contraction which increases the federal funds rate by 25 basis points. Firstly, I confirm Caldara and Herbst (2019)'s finding that the Baa spread increases after a monetary contraction. This effect is even stronger for the Aaa spread. Secondly, the factor setup allows me to analyse the reaction of the yield curve in more detail. I find that, in line with economic intuition, government spreads with shorter maturities respond more strongly to a monetary contraction than spreads with longer maturities.

4.2 Uncertainty shocks

A second popular example of the use of Proxy VAR models is the identification of uncertainty shocks in the workhorse model by Bloom (2009). The aim is to trace out the effects of unexpected exogenous variations in uncertainty, measured as realised volatility of stock market futures (VXO), on the economy. Within this context, the main challenge is the two-way causality between the business cycle and uncertainty, which precludes the use of timing restrictions. To address this issue, a number of studies propose external instruments. Examples are, among others, Stock and Watson (2012), Carriero et al. (2015), and Piffer and Podstawski (2017).

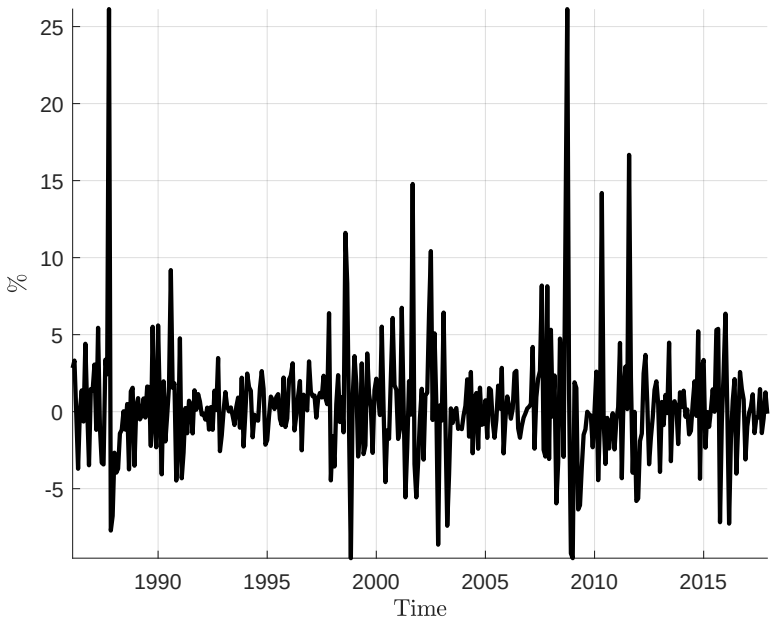

Figure 7: ***Proxy for uncertainty shocks.*** *Residuals of an AR(2) process of the VXO, see Stock and Watson (2012).*

In this section, I re-visit the question of the effects of uncertainty shocks in a data-rich environment. I employ a popular proxy by Stock and Watson (2012) based on the residuals of the VXO and augment the workhorse model by Bloom (2009) by latent factors. This allows me to trace out the effects of uncertainty shocks on disaggregated consumption and investment series.

4.2.1 Data, specification and results

The data contained in $\boldsymbol{z}_t$ are eight monthly US data series from Bloom (2009), using the transformations as in Piffer and Podstawski (2017). The vector of observational series, $\boldsymbol{x}_t$, contains 128 monthly series from the FRED dataset by McCracken and Ng (2016) (see Appendix C for a detailed description). As an instrument m_t I employ the residuals of an AR(2) process of the VXO, as proposed by Stock and Watson (2012).[3] The sample runs from 1961M2 to 2015M6. I follow Piffer and Podstawski (2017) in setting the lag length to $P = 5$.

Figure 7 shows the proxy. It is continuous and peaks in well-known periods of heightened uncertainty such as the 2008 collapse of Lehman brothers or the 1989 fall of the Berlin wall. As a number of factors I set $R = 2$ (see Appendix C for details). I estimate the latent factors using Principal Components due to the higher dimensionality

[3]An important requirement for m_t in this setup is that it is continuous. This precludes, for example, the use of the proxy in Piffer and Podstawski (2017) which is non-zero for only 37 months.

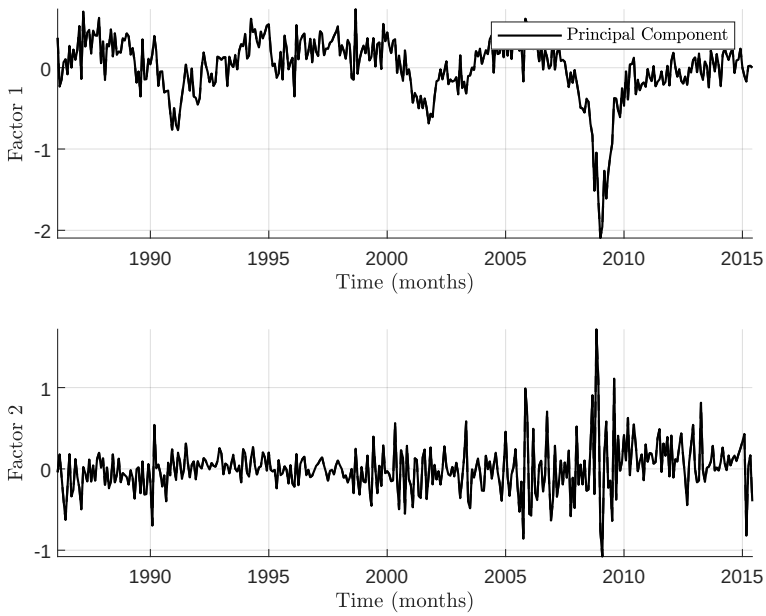

Figure 8: ***Latent factors.*** *Principal Components estimate of latent factors.*

of the system which renders the Kalman filter computationally demanding. As suggested in the previous section, the PC estimates capture the dynamics of the Kalman filter well but could potentially lead to an underestimation of the estimation uncertainty. Figure 8 plots the resulting estimate.

Figure 9 shows the impulse responses of the model containing the Bloom (2009) variables together with the impulse responses from the factor-augmented model. The BP-FAVAR model replicates the main dynamics of the model using only the Bloom (2009) variables. This is also in line with Piffer and Podstawski (2017)'s Proxy VAR suggesting that their instrument and Stock and Watson (2012)'s instrument have similar properties. The similarity of the two models suggests that the 8-variate workhorse model captures the relevant information in the economy well and is unlikely to be affected by omitted variable biases. In particular, both models show a short-lived drop in stock market returns following an exogenous increase in uncertainty. Price inflation drops and rebounds together with the three measures of real activity (hours worked, employment growth and industrial production growth) suggesting a dominance of demand-side behaviour. This confirms findings by Leduc and Liu (2016) that uncertainty shocks have similar effects as aggregate demand shocks.

Figure 10 shows the response of selected informational series. It suggests that consumption-related quantities such as real personal consumption expenditure (top left) and industrial production of consumer goods (top right), react less strongly to an uncertainty shock than investment-related quantities such as industrial production of business equipment (bottom left) or industrial production of durable materials (bottom right).

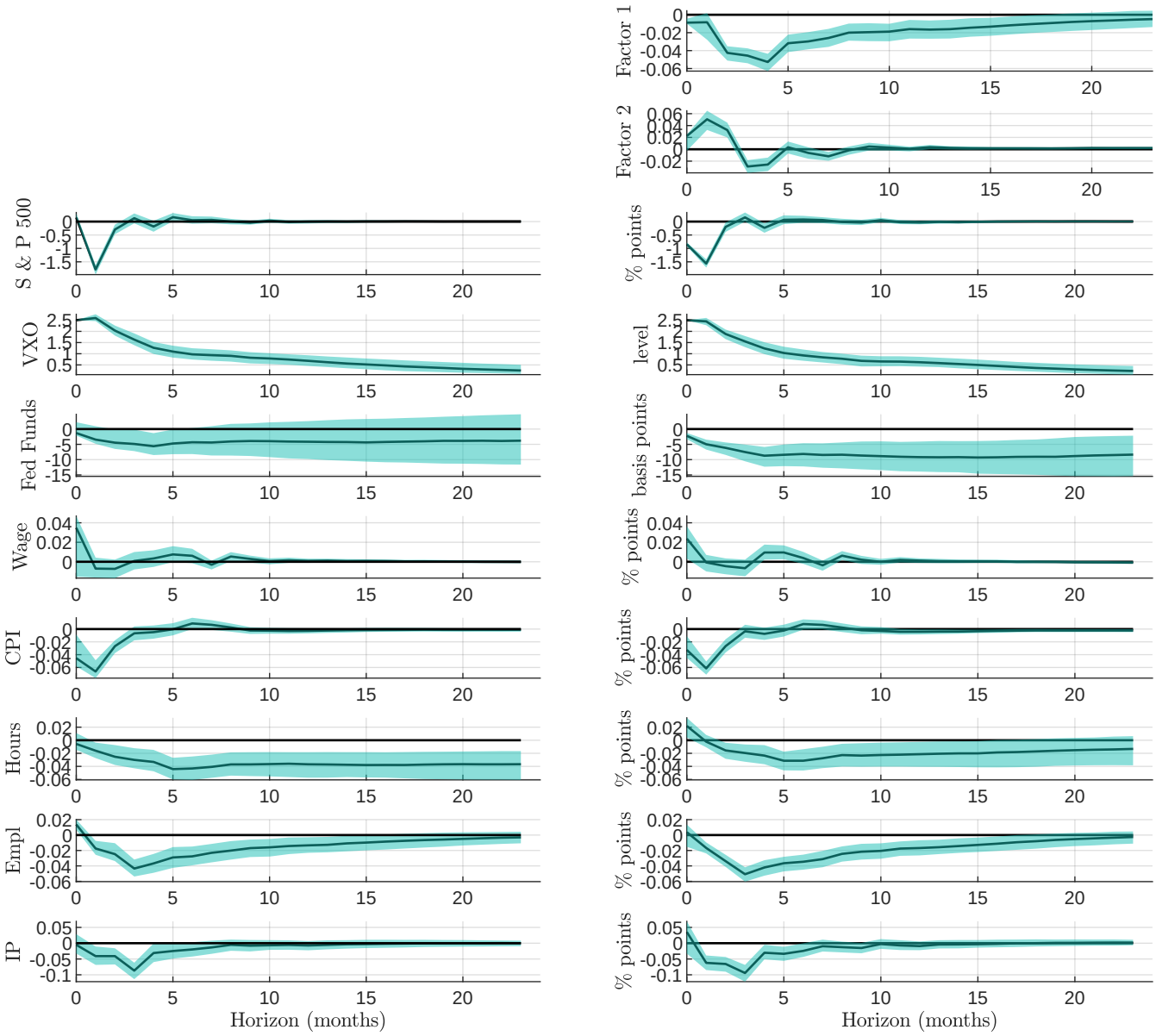

Figure 9: ***Impulse Response Functions of observables and latent factors.*** *Point-wise (non-cumulated) median impulse responses (solid line) with 68% bands. Small-scale VAR (left column) excludes latent factors. BP-FAVAR (right column) includes two latent factors from financial series.*

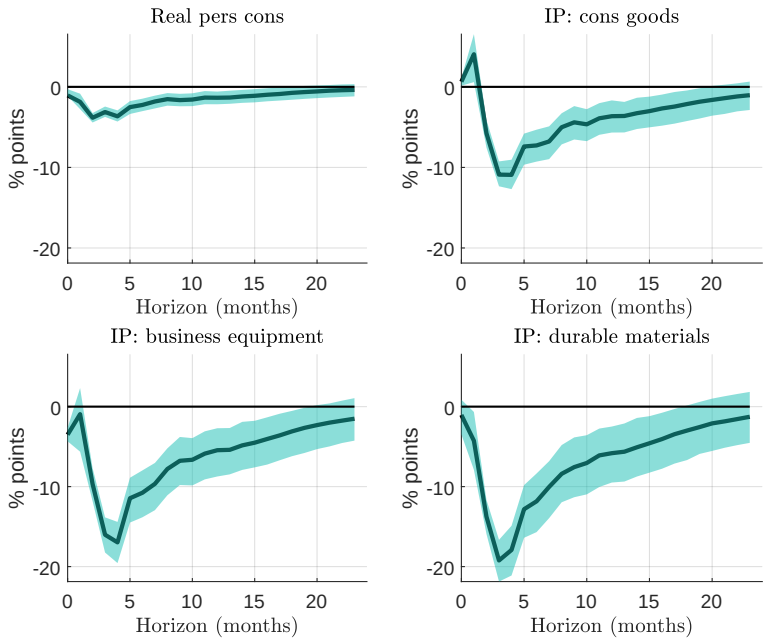

Figure 10: ***Impulse Response Functions of informational series.*** *Point-wise (non-cumulated) median impulse responses (solid line) with 68% bands. See Appendix C table 6 for a detailed variable description.*

These findings lend support to the "wait-and-see" channel of uncertainty shocks (Bloom, 2009) stating that firms postpone their investment decision in light of increased uncertainty while not ruling out demand-related channels, notably the "precautionary savings" motive discussed in Fernández-Villaverde et al. (2015) and Born and Pfeifer (2014) which states that consumers prefer to increase savings and reduce consumption when they perceive it harder to predict the future.

5 Conclusion

This paper proposes a Bayesian Proxy FAVAR to jointly address informational insufficiency and identification issues in structural VAR models. I combine a recent strand of the Bayesian VAR literature that uses external instruments for identification (Caldara and Herbst, 2019) with the Bayesian factor model literature (Belviso and Milani, 2006, Amir-Ahmadi and Uhlig, 2015). I show how a state-space model can be set up to jointly exploit the advantages of both approaches. The resulting Bayesian Proxy factor-augmented VAR model avoids two shortcomings of commonly employed small-scale recursively identified VAR models, namely a non-credible identification scheme and informational insufficiency. In two applications I find that an augmented information set changes the dynamics of small-scale VAR models and that interesting conclusions can be drawn about disaggregated series.

Future empirical research can use the BP-FAVAR to avoid omitted variable problems and investigate a large number of informational series in settings where identification via proxies is appropriate.

References

Amir-Ahmadi, P. and Uhlig, H. (2015), 'Sign restrictions in Bayesian FAVARs with an application to monetary policy shocks', *NBER Working paper 21738* .

Arias, J. E., Rubio-Ramírez, J. F. and Waggoner, D. F. (2018), 'Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications', *Econometrica* **86**(2), 685–720.

Bahaj, S. (forthcoming), 'Sovereign spreads in the euro area: Cross border transmission and macroeconomic implications', *Journal of Monetary Economics* .

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

Baumeister, C. and Hamilton, J. D. (2015), 'Sign restrictions, structural vector autoregressions, and useful prior information', *Econometrica* **83**(5), 1963–1999.

Belviso, F. and Milani, F. (2006), 'Structural factor-augmented VARs (SFAVARs) and the effects of monetary policy', *Topics in Macroeconomics* **6**(3).

Bernanke, B. S., Boivin, J. and Eliasz, P. (2004), Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach, Technical report, National Bureau of Economic Research.

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

Bloom, N. (2009), 'The impact of uncertainty shocks', *Econometrica* **77**(3), 623–685.

Born, B. and Pfeifer, J. (2014), 'Policy risk and the business cycle', *Journal of Monetary Economics* **68**, 68–85.

Caldara, D. and Herbst, E. (2019), 'Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs', *American Economic Journal: Macroeconomics* **11**(1), 157–92.

Carriero, A., Mumtaz, H., Theodoridis, K. and Theophilopoulou, A. (2015), 'The impact of uncertainty shocks under measurement error: A proxy svar approach', *Journal of Money, Credit and Banking* **47**(6), 1223–1238.

Carter, C. K. and Kohn, R. (1994), 'On Gibbs sampling for state space models', *Biometrika* **81**(3), 541–553.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (1999), 'Monetary policy shocks: What have we learned and to what end?', *Handbook of Macroeconomics* **1**, 65–148.

Cowles, M. K. and Carlin, B. P. (1996), 'Markov Chain Monte Carlo convergence diagnostics: a comparative review', *Journal of the American Statistical Association* **91**(434), 883–904.

Drautzburg, T. (2016), A narrative approach to a fiscal dsge model, Technical report, Federal Reserve Bank of Philadelphia.

Fernández-Villaverde, J., Guerrón-Quintana, P., Kuester, K. and Rubio-Ramírez, J. (2015), 'Fiscal volatility shocks and economic activity', *The American Economic Review* **105**(11), 3352–3384.

Frühwirth-Schnatter, S. (1994), 'Data augmentation and dynamic linear models', *Journal of time series analysis* **15**(2), 183–202.

Gertler, M. and Karadi, P. (2015), 'Monetary policy surprises, credit costs, and economic activity', *American Economic Journal: Macroeconomics* **7**(1), 44–76.

Geweke, J. (1992), 'Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments', *Bayesian Statistics* **4**, 169–193.

Huber, F. and Fischer, M. M. (2018), 'A Markov switching factor-augmented VAR model for analyzing US business cycles and monetary policy', *Oxford Bulletin of Economics and Statistics* **80**(3), 575–604.

Kerssenfischer, M. (2019), 'The puzzling effects of monetary policy in vars: Invalid identification or missing information?', *Journal of Applied Econometrics* **34**(1), 18–25.

Kilian, L. and Lütkepohl, H. (2017), *Structural vector autoregressive analysis*, Cambridge University Press.

Leduc, S. and Liu, Z. (2016), 'Uncertainty shocks are aggregate demand shocks', *Journal of Monetary Economics* **82**, 20–35.

Litterman, R. B. (1986), 'Forecasting with Bayesian vector autoregressions: five years of experience', *Journal of Business & Economic Statistics* **4**(1), 25–38.

McCracken, M. W. and Ng, S. (2016), 'Fred-md: A monthly database for macroeconomic research', *Journal of Business & Economic Statistics* **34**(4), 574–589.

Mertens, K. and Ravn, M. O. (2013), 'The dynamic effects of personal and corporate income tax changes in the united states', *The American Economic Review* **103**(4), 1212–1247.

Piffer, M. and Podstawski, M. (2017), 'Identifying uncertainty shocks using the price of gold', *The Economic Journal* **128**(616), 3266–3284.

Ramey, V. A. (2016), Macroeconomic shocks and their propagation, Technical report, National Bureau of Economic Research.

Rubio-Ramirez, J. F., Waggoner, D. F. and Zha, T. (2010), 'Structural vector autoregressions: Theory of identification and algorithms for inference', *The Review of Economic Studies* **77**(2), 665–696.

Stock, J. H. and Watson, M. W. (2012), 'Disentangling the channels of the 2007-2009 recession', *NBER Working paper 18094* .

Stock, J. H. and Watson, M. W. (2016), 'Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics', *Handbook of macroeconomics* **2**, 415–525.

Wu, J. C. and Xia, F. D. (2016), 'Measuring the macroeconomic impact of monetary policy at the zero lower bound', *Journal of Money, Credit and Banking* **48**(2-3), 253–291.

Yamamoto, Y. (2019), 'Bootstrap inference for impulse response functions in factor-augmented vector autoregressions', *Journal of Applied Econometrics* **34**(2), 247–267.

A Posterior Inference

A.1 Conditional posterior of F

The procedure to generate posterior draws of latent factors, F, differs from generating draws of parameters, in that one has to generate the whole dynamic evolution of factors for each $t = 1, ..., T$. For this to be feasible I exploit the Markov property of the system described in equation (3) as follows:

$$p(Y|X, \theta) = p(\boldsymbol{y}_t|X, \theta) \prod_{t=1}^{T-1} p(\boldsymbol{y}_t|\boldsymbol{y}_{t+1}, X, \theta). \tag{22}$$

First note that (22) describes the posterior of Y, which contains both latent and observable factors. The reason for including the observable factors is the dynamic interdependence between latent and observable factors, which needs to be accounted for. Given that the observable factors are non-random, their distribution has a zero variance.[4] Second, note that this is a product of $R + K$-dimensional conditional distributions. Given the assumption of Gaussianity of $\boldsymbol{\xi}_t$ and $\boldsymbol{u}_t$, this representation can be combined with the observation equation (1) and is amenable to the Carter-Kohn algorithm described in Carter and Kohn (1994) and Frühwirth-Schnatter (1994).

A.1.1 State-space form

Start by rewriting observation equation (1) and transition equation (3) as

$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{z}_t \end{bmatrix} = \mathcal{H}\mathcal{B}_t + \mathcal{W}_t \tag{23}$$

$$\mathcal{B}_t = \mathcal{F}\mathcal{B}_{t-1} + \mathcal{V}_t \tag{24}$$

$$Var(\mathcal{W}_t) = \mathcal{R} \tag{25}$$

$$Var(\mathcal{V}_t) = \mathcal{Q} \tag{26}$$

where

$$\mathcal{H} = \begin{bmatrix} \Lambda^f & \Lambda^z & 0 & ... & 0 \\ \mathbf{0} & I & 0 & ... & 0 \end{bmatrix}; \quad \mathcal{B}_t = \begin{bmatrix} \lfloor_t' & \lfloor_{t-1}' & ... & \lfloor_{t-p}' \end{bmatrix}'; \quad \lfloor_t = \begin{bmatrix} \boldsymbol{f}_t \\ \boldsymbol{z}_t \end{bmatrix}$$

[4] Here, I refer to the variance across draws. The variance across time is, of course, non-zero.

$$\mathcal{W}_t = \begin{bmatrix} \boldsymbol{\xi}_t \\ \mathbf{0} \end{bmatrix}; \quad \mathcal{F} = \begin{bmatrix} & \Pi & \\ I & \mathbf{0} & \mathbf{0} \end{bmatrix}; \quad \mathcal{V}_t = \begin{bmatrix} \boldsymbol{u}_t \\ \mathbf{0} \end{bmatrix}$$

$$\mathcal{R} = \begin{bmatrix} \Omega & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}; \quad \mathcal{Q} = \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Then consider the following factorisation:

$$p(\mathcal{B}_{1:T}|X, \boldsymbol{\theta}) = p(\mathcal{B}_T|\boldsymbol{x}_{1:T}, \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\mathcal{B}_t|\mathcal{B}_{t+1}, X, \theta) \tag{27}$$

Given the linear Gaussian form of the state space model we have that

$$\mathcal{B}_T|\boldsymbol{x}_{1:T}, \theta \sim N(\mathcal{B}_{T|T}, \mathcal{P}_{T|T}) \tag{28}$$

$$\mathcal{B}_{t|T}|\mathcal{B}_{t+1|T}, \boldsymbol{x}_{1:T}, \theta \sim N(\mathcal{B}_{t|t,\mathcal{B}_{t+1|T}}, \mathcal{P}_{t|t,\mathcal{B}_{t+1|T}}) \tag{29}$$

with

$$\mathcal{B}_{T|T} = E(\mathcal{B}_T|\boldsymbol{x}_{1:T}, \boldsymbol{\theta}) \tag{30}$$

$$\mathcal{P}_{T|T} = Cov(\mathcal{B}_T|\boldsymbol{x}_{1:T}, \boldsymbol{\theta}) \tag{31}$$

$$\mathcal{B}_{t|t,\mathcal{B}_{t+1|T}} = E(\mathcal{B}_t|\mathcal{B}_{t|t}, \mathcal{B}_{t+1|t}, \boldsymbol{\theta}) \tag{32}$$

$$\mathcal{P}_{t|t,\mathcal{B}_{t+1|T}} = Cov(\mathcal{B}_t|\mathcal{B}_{t|t}, \mathcal{B}_{t+1|t}, \boldsymbol{\theta}) \tag{33}$$

A.1.2 Carter-Kohn algorithm

In a first step, we can run a Kalman filter to obtain a series of Kalman-filtered draws of the state variable $\mathcal{B}_t$ $\mathcal{B}_{t|t}$ for $t = 1, ...T$. To initialise, we set $\mathcal{B}_{1|0} = 0$ and $\mathcal{P}_{1|0} = \mathbf{I}$. Then, iterate forward as:

$$\mathcal{B}_{t|t} = \mathcal{B}_{t|t-1} + \kappa_{t|t-1}\eta_{t|t-1} \tag{34}$$

where $\eta_{t|t-1} = \mathcal{B}_t - \mathcal{F}\mathcal{B}_{t|t-1}$ denotes the forecast error, $\mathfrak{f}_{t|t-1} = \mathcal{H}\mathcal{P}_{t|t-1}\mathcal{H}' + \mathcal{R}$ its variance and $\kappa_{t|t-1} = \mathcal{P}_{t|t-1}\mathcal{H}\mathfrak{f}_{t|t-1}^{-1}$ the "Kalman-gain"

$$\mathcal{P}_{t|t-1} = \mathcal{F}\mathcal{P}_{t-1|t-1}\mathcal{F}' + \mathcal{Q} \tag{35}$$

Then, conditioning on the last of these Kalman-filtered draws, $\mathcal{B}_{T|T}$ and $\mathcal{P}_{T|T}$, we can run the filter backwards to obtain a series $\mathcal{B}_{t|t+1}$ for $t = 1, ..., T-1$ as follows:

$$\mathcal{B}^*_{t|t,\mathcal{B}_t+1|T} = \mathcal{B}_{t|t} + \mathcal{P}_{t|t}\mathcal{F}^{*'} J^{-1}_{t+1|t}\psi_{t+1|t} \tag{36}$$

$$\mathcal{P}^*_{t|t,\mathcal{B}_t+1|T} = \mathcal{P}_{t|t} - \mathcal{P}_{t|t}\mathcal{F}^{*'} J^{-1}_{t+1|t}\mathcal{F}^*\mathcal{P}_{t|t} \tag{37}$$

where $\psi_{t+1|t} = \mathcal{B}^*_{t+1} - \mathcal{F}^*\mathcal{B}_{t|t}$ and $J_{t+1|t} = \mathcal{F}^* P_{t|t}\mathcal{F}^{*'} + \mathcal{Q}^*$. Note that $\mathcal{Q}^*$ refers to the top $R \times R$ block of $\mathcal{Q}$ and that $\mathcal{F}^*$ and $\mathcal{B}^*$ denote the first R rows of $\mathcal{F}$ and $\mathcal{B}$, respectively. This is required because $\mathcal{Q}$ is singular given the presence of observable factors.

Plugging these draws into (27) results in an unconditional posterior draw of the state variable, $\mathcal{B}_1, ..., \mathcal{B}_T$. Its top $R + K$ block represents an unconditional posterior draw of factors, $\boldsymbol{y}_t$.

A.2 Conditional posterior of Λ, Ω

This section discusses conditional posterior inference on Λ, Ω. It draws partially on Bernanke et al. (2004). Restate the observation equation for convenience

$$\boldsymbol{x}_t = \Lambda^f \boldsymbol{f}_t + \Lambda^z \boldsymbol{z}_t + \boldsymbol{\xi}_t \tag{38}$$

$$\boldsymbol{\xi}_t \sim N(\mathbf{0}, \Omega) \tag{39}$$

with $\Lambda = [\lambda^f \quad \lambda^z]$. Given a draw of the latent and observable factors, $\boldsymbol{y}_t = [\boldsymbol{f}'_t \quad \boldsymbol{z}'_t]'$, and under the assumption that Ω is diagonal, (38) amounts to N independent linear regressions.

Assuming (conjugate) normal-inverse Gamma priors of the form

$$\boldsymbol{\omega}_{ii} \sim IG(sc^*, sh^*) \tag{40}$$

$$\boldsymbol{\lambda}_i|\boldsymbol{\omega}_{ii} \sim N(\boldsymbol{\mu}^*_{\lambda,i}, \boldsymbol{\omega}_{ii}M^{*-1}_i), \tag{41}$$

delivers posterior distributions of the form

$$\boldsymbol{\omega}_{ii} \sim IG(sc, sh) \tag{42}$$

$$\boldsymbol{\lambda}_i|\boldsymbol{\omega}_{ii} \sim N(\boldsymbol{\mu}_{\lambda,i}, \boldsymbol{\omega}_{ii}M^{-1}_i), \tag{43}$$

with

$$
\begin{aligned}
sh &= sc^* + T \\
sc &= sc^* + \hat{\xi}_i \hat{\xi}_i{}' + (\hat{\lambda}_i - \boldsymbol{\mu}^*_{\lambda,i})'(M_i^{*-1} + (Y_i Y_i')^{-1})(\hat{\lambda}_i - \boldsymbol{\mu}^*_{\lambda,i}) \\
M_i &= M_i^* + Y_i Y_i' \\
\boldsymbol{\mu}_{\lambda,i} &= M_i^{-1}(M_i^* \boldsymbol{\mu}^*_{\lambda,i} + Y_i Y_i')\hat{\lambda}_i,
\end{aligned}
$$

where $\hat{\xi}_i$ are the fitted errors from the i-th regression and $\hat{\lambda}_i$ is the OLS estimate of λ_i, and Y_i are the regressors of the i-th equation. Note that employing uninformative priors centred atound zero, i.e. setting $sc^* = 0$, $sh^* = 0$, $\boldsymbol{\mu}^*_{\lambda,i} = \mathbf{0}$, and $M_i^{*-1} = 0$ collapses the posterior towards the OLS estimate:

$$
\begin{aligned}
sh &= T \\
sc &= \hat{\xi}_i \hat{\xi}_i{}' \\
M_i &= Y_i Y_i' \\
\boldsymbol{\mu}_{\lambda,i} &= \hat{\lambda}_i,
\end{aligned}
$$

A.3 Conditional posterior of Σ, Π

This section discusses conditional posterior inference on Σ, Π. It proceeds in two steps: First, generate draws from

$$p(\Pi, \Sigma | \Lambda^f, \Lambda^z, \Omega, F, \beta, \sigma_\nu, \boldsymbol{b}, X, Z)$$

(note that $\boldsymbol{m}$ is not part of the conditioning set). Second, map these into draws from

$$p(\Pi, \Sigma | \Lambda^f, \Lambda^z, \Omega, F, \beta, \sigma_\nu, \boldsymbol{b}, X, Z, \boldsymbol{m})$$

(note that $\boldsymbol{m}$ is now part of the conditioning set).

For the first step, the derivations can be found, for example, in Kilian and Lütkepohl (2017). Restate the transition equation for convenience

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_t + \boldsymbol{u}_t \tag{44}$$

$$\boldsymbol{u}_t \sim N(\mathbf{0}, \Sigma), \tag{45}$$

Given a draw of latent and observable factors, $\boldsymbol{y}_t$, equation (44) is a standard VAR(P)

model. Therefore, employing independent normal-inverse Wishart priors of the form

$$vec(\Pi) \sim N(\boldsymbol{\mu}_{\Pi}^{*}, V_{\Pi}^{*}) \tag{46}$$

$$\Sigma \sim IW(S^{*}, \tau^{*}), \tag{47}$$

delivers posterior distributions of the form

$$vec(\Pi)|\Sigma \sim N(\boldsymbol{\mu}_{\Pi}, V_{\Pi}) \tag{48}$$

$$\Sigma|\Pi \sim IW(S, \tau), \tag{49}$$

where

$$V_{\Pi} = (V_{\Pi}^{*-1} + (WW' \otimes \Sigma^{-1}))^{-1} \tag{50}$$

$$\boldsymbol{\mu}_{\Pi} = V_{\Pi}(V_{\Pi}^{*-1} + (W \otimes \Sigma^{-1})vec(Y)) \tag{51}$$

$$S = S^{*} + (Y - \Pi W)(Y - \Pi W)' \tag{52}$$

$$\tau = \tau^{*} + T \tag{53}$$

In a second step, draws from the above distribution are mapped into draws from the full conditional distribution using a Metropolis-Hastings step (see below for the acceptance probability).

A.4 Conditional posterior of b

This section re-parametrises Caldara and Herbst (2019) to allow for identification of impact effects. For the posterior sampler, we will need to be able to evaluate the conditional likelihood of m_t given $\boldsymbol{y}_t$.

Write transition and proxy equation in stacked form for convenience and compute the unconditional variance:

$$\begin{bmatrix} \boldsymbol{y}_t \\ m_t \end{bmatrix} = \begin{bmatrix} \Pi & 0 \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_t \\ m_{t-1} \end{bmatrix} + \begin{bmatrix} B & \boldsymbol{\beta} \\ \boldsymbol{\beta}' & \sigma_{\nu} \end{bmatrix} \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \nu_t \end{bmatrix} \tag{54}$$

$$Var(\begin{bmatrix} \boldsymbol{y}_t - \Pi \boldsymbol{w}_t \\ m_t \end{bmatrix}) = \begin{bmatrix} \boldsymbol{\Sigma} & B\boldsymbol{\beta}' \\ \boldsymbol{\beta} B' & \boldsymbol{\beta}'\boldsymbol{\beta} + \sigma_{\nu}^{2} \end{bmatrix}, \tag{55}$$

where $\boldsymbol{\beta} = [\beta \quad \mathbf{0}]'$

The likelihood is invariant to observationally equivalent rotations of B. Therefore we can replace $B = B^c Q$, where B^c is, for example, the lower-triangular Cholesky decomposition of Σ.

$$Var(\begin{bmatrix} \boldsymbol{y}_t - \Pi \boldsymbol{w}_t \\ m_t \end{bmatrix}) = \begin{bmatrix} \Sigma & B^c Q \boldsymbol{\beta}' \\ \boldsymbol{\beta} Q' B^{c'} & \boldsymbol{\beta}'\boldsymbol{\beta} + \sigma_\nu^2 \end{bmatrix} \tag{56}$$

Then, using the rules for the conditional mean of multivariate normal distributions, we obtain the conditional likelihood

$$m_t | \boldsymbol{y}_t, \Pi, \Sigma, \boldsymbol{b}, \beta, \sigma_\nu \sim N(\mu_{m|Y}, V_{m|Y}), \tag{57}$$

$$\mu_{m|Y} = \boldsymbol{\beta} Q' B^{c'} \Sigma^{-1} \boldsymbol{u}_t \tag{58}$$

$$= \beta \boldsymbol{\epsilon}_{1,t} \tag{59}$$

$$V_{m|Y} = \boldsymbol{b}\boldsymbol{b}' + \sigma_\nu^2 - \boldsymbol{b} Q' B^{c'} \Sigma^{-1} B^c Q \boldsymbol{b}' \tag{60}$$

$$= \sigma_\nu^2 \tag{61}$$

Note that the conditional likelihood of m_t does not depend on the full matrix B, but only on its first column, $\boldsymbol{b}$ because the model is partially identified. Therefore, we can rewrite (57) as

$$m_t | \boldsymbol{y}_t, \Pi, \Sigma, \boldsymbol{b}, \beta, \sigma_\nu \sim N(\mu_{m|Y}, V_{m|Y}), \tag{62}$$

Next, we can use the above result in a Metropolis step to generate a draw of $\boldsymbol{b}$: given a draw of Π, Σ, draw $Q_{\cdot,1}^{cand}$ as the first column of an orthogonal matrix form a uniform Haar distribution using the algorithm by Rubio-Ramirez et al. (2010). Set $Q_{\cdot,1} = Q_{\cdot,1}^{cand}$ with probability α and $Q_{\cdot,1}$ equal to the previous draw, $Q_{\cdot,1}^{j-1}$, otherwise.

$$\alpha = min(\frac{p(\boldsymbol{m}|Y, \Pi, \Sigma, Q_{\cdot,1}^{cand})}{p(\boldsymbol{m}|Y, \Pi, \Sigma, Q_{\cdot,1}^{j-1})}, 1)$$

Compute structural errors $\epsilon_{1,t} = (chol(\Sigma) Q_{\cdot,1})^{-1} U$.

A.5 Conditional posterior of β, σ_ν

Given a draw of structural shocks, $\epsilon_{1,t}$, the proxy equation is a linear equation. I consider two types of priors: a flat and a "high-relevance" prior. Restate the proxy equation for

convenience

$$m_t = \beta\epsilon_{1,t} + \sigma_\nu \nu_t \tag{63}$$
$$\nu_t \sim N(0,1), \tag{64}$$

First, for the flat prior, results from appendix A.2 apply and the normal-inverse Gamma priors of the form

$$\beta \sim N(\mu^*_\beta, \sigma^{*2}_\beta) \tag{65}$$
$$\sigma_\nu \sim IG(sc^*_\nu, sh^*_\nu) \tag{66}$$

map into posteriors of the form

$$\beta|\sigma_\nu \sim N(\mu_\beta, \sigma^2_\beta) \tag{67}$$
$$\sigma_\nu|\beta \sim IG(sc_\nu, sh_\nu) \tag{68}$$

with (assuming zero-centred un-informative priors)

$$\mu_\beta = \hat{\beta} \tag{69}$$
$$\sigma^2_\beta = (\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1')^{-1} \tag{70}$$
$$sc_\nu = sc^*_\nu + T \tag{71}$$
$$sh_\nu = \hat{\nu}\hat{\nu}', \tag{72}$$

where $\hat{\nu}$ are the fitted errors from equation (63) and $\hat{\beta}$ is the OLS estimate of β.

Second, for the "high-relevance" prior, the posterior of β is unaffected. σ_ν, however, is not updated and stays at $\sigma_\nu = 0.5std(m_t)$ throughout the sampler.

A.6 Convergence of the Posterior Sampling Algorithm

The convergence properties of the reduced form parameters of a Bayesian FAVAR model are discussed in detail in Amir-Ahmadi and Uhlig (2015). They show that a Gibbs sampling procedure, similar to the one employed for the reduced form parameters here, converges for appropriate lengths of the sampler. The convergence properties of the structural parameters, however, need to be assessed. In particular the first column of B containing the on-impact effects of the shock of interest are of importance. In order to do so, I follow Amir-Ahmadi and Uhlig (2015) and employ the convergence diagnostic

proposed by Geweke (1992). A detailed discussion of this convergence diagnostic can be found, for example, in Cowles and Carlin (1996).

This diagnostic assesses the convergence of each element η_i of parameter vector, $\boldsymbol{\eta}$. The assessment is based on a comparison of means across different parts of this chain. If the means are close to each other, the procedure detects convergence.

In a first step, extract from each (univariate) posterior draw $\{\eta_i\}_{i=1}^{D}$ the following subseries: $\eta_{1i}, ..., \eta_{0.1D,i}$, i.e. the first 10 % of draws for parameter i, and $\eta_{0.6D+1,i}, ..., \eta_{D,i}$, i.e. the last 40% of draws, where D is the length of the MCMC chain. Compute $\hat{\mu}_{first}$ and $\hat{\mu}_{last}$, the mean, as well as $\hat{\sigma}_{first}$ and $\hat{\sigma}_{last}$, the standard deviation, of these subseries. Then the test statistic is

$$CD = \frac{\hat{\mu}_{first} - \hat{\mu}_{last}}{\frac{\hat{\sigma}_{first}}{\sqrt{0.1D}} + \frac{\hat{\sigma}_{last}}{\sqrt{0.4D}}} \tag{73}$$

Under the conditions mentioned in Geweke (1992), CD has an asymptotic standard normal distribution

The final output is a p-value indicating whether or not we can reject the Null hypothesis of convergence, i.e. equality of mean across the chain, at a given significance level.

B Criteria to determine the number of factors R

This section describes two criteria to determine the number of factors: The Bai and Ng (2002) criterion and the scree plot.

Bai and Ng (2002) suggest the following criterion to determine the number of factors:

$$\begin{aligned} BN(R) =& \log \frac{1}{NT} \sum_{t=1}^{T} (X_t - \Lambda^{PC} \boldsymbol{f}_t^{PC})'(X_t - \Lambda^{PC} \boldsymbol{f}_t^{PC}) \\ &+ R \frac{N+T}{NT} log(min(N,T)) \end{aligned} \tag{74}$$

where Λ^{PC} and $\boldsymbol{f}_t^{PC}$ are the principal components estimators of the factor loadings and the factors, respectively. Bai and Ng (2002) suggest to set R^* such that (74) is minimised.

A scree plot summarizes the marginal contribution of the r-th factor to the average explanatory power of N regressions of $\boldsymbol{x}_t$ against the first r factors as computed via Principal Components (see Stock and Watson, 2016 for details).

C Additional Figures and Tables

C.1 Monetary Policy Application

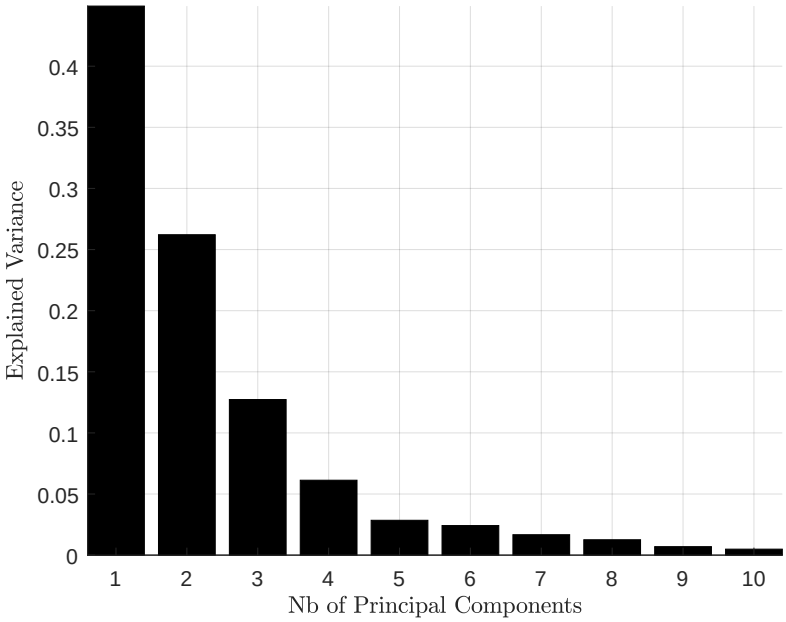

Figure 11: ***Scree Plot (Monetary Policy Application).*** *Explained share of variance in* $\boldsymbol{x}_t$ *as a function of the number of latent factors (R) included in* $\boldsymbol{f}_t$

R	Criterion
1	-2.0439
2	**-2.2853**
3	-2.2226
4	-2.1295
5	-1.9393
6	-1.7756
7	-1.5723
8	-1.3702
9	-1.1378
10	-0.8515

Table 2: ***Bai and Ng (2002) criterion (Monetary Policy Application).*** *See Appendix B for a description.*

Variable	Criterion
1	0.1465
2	0.2257
3	0.1428
4	0.1555
5	0.7312

Table 3: ***Geweke (1992) convergence diagnostic (Monetary Policy Application).*** *Null hypothesis: convergence. See Appendix A.6 for a description.*

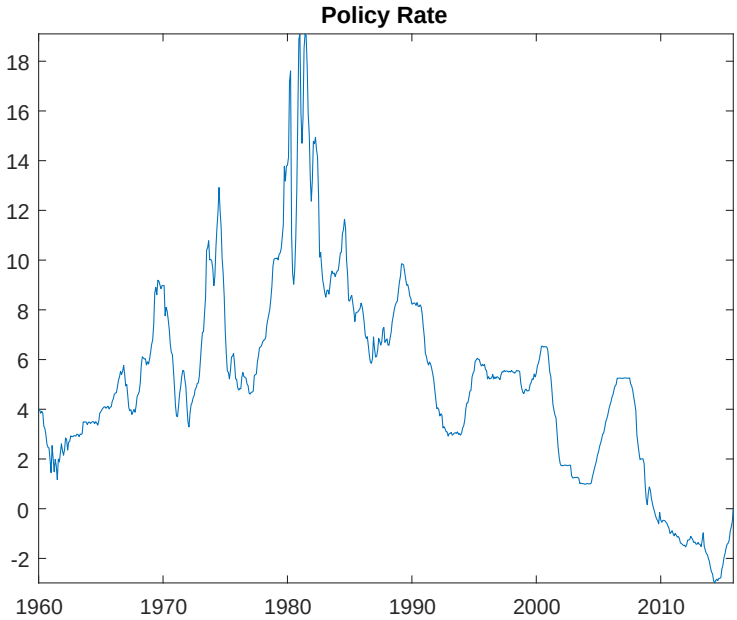

Figure 12: **Shadow Rate:** Wu and Xia (2016) shadow rate available from 1960M1 to 2015M11.

C.2 Uncertainty Application

R	Criterion
1	1.1692
2	**1.0286**
3	1.0317
4	1.0583
5	1.1060
6	1.1558
7	1.2062
8	1.2570
9	1.3079
10	1.3589

Table 4: ***Bai and Ng (2002) criterion (Uncertainty Application).*** *See Appendix B for a description.*

Variable	Criterion
1	0.7888
2	0.4595
3	0.9055
4	0.0434
5	0.8398
6	0.6726
7	0.4275
8	0.4834
9	0.4275
10	0.4432

Table 5: ***Geweke (1992) convergence diagnostic (Uncertainty Application).*** *Null hypothesis: convergence. See Appendix A.6 for a description.*

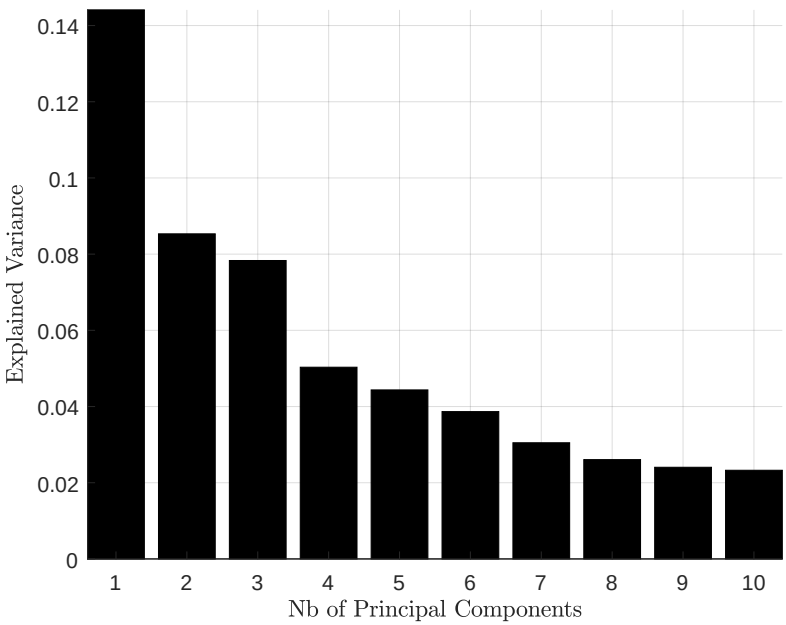

Figure 13: ***Scree Plot (Uncertainty Application).*** *Explained share of variance in* $\boldsymbol{x}_t$ *as a function of the number of latent factors (R) included in* $\boldsymbol{f}_t$

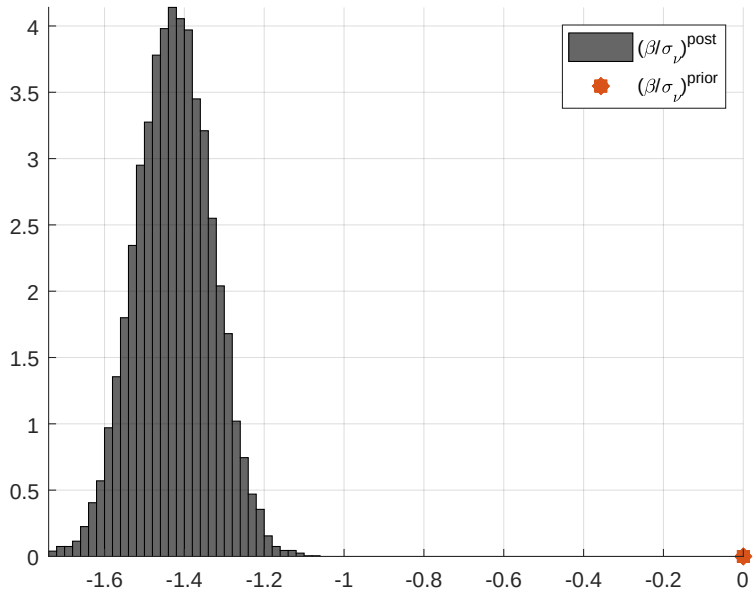

Figure 14: ***Signal-to-noise ratio (Uncertainty Application).*** *Posterior distribution of* β/σ_ν *(histogram) together with the prior median (star).*

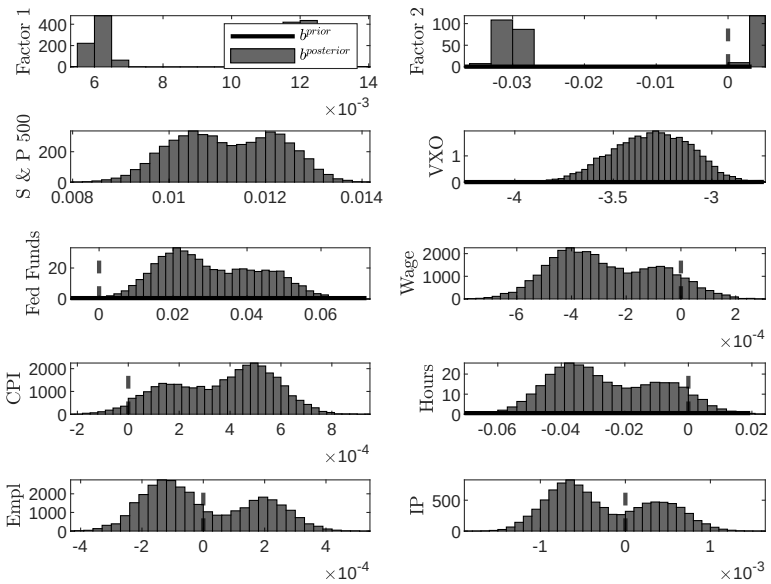

Figure 15: ***Updating of b (Uncertainty Application).*** *Priors (solid line) and posterior (histogram) of* $\boldsymbol{b}$*. Prior draws are computed from the distribution implicit in the priors for* Σ*,* $Q_{\cdot,1}$*,* β *and* σ_ν*.*

Table 6: Data

Output and Income

id	tcode	fred	description
1	5	RPI	Real Personal Income
2	5	W875RX1	Real personal income ex transfer receipts
6	5	INDPRO	IP Index
7	5	IPFPNSS	IP: Final Products and Nonindustrial Supplies
8	5	IPFINAL	IP: Final Products (Market Group)
9	5	IPCONGD	IP: Consumer Goods
10	5	IPDCONGD	IP: Durable Consumer Goods
11	5	IPNCONGD	IP: Nondurable Consumer Goods
12	5	IPBUSEQ	IP: Business Equipment
13	5	IPMAT	IP: Materials
14	5	IPDMAT	IP: Durable Materials
15	5	IPNMAT	IP: Nondurable Materials
16	5	IPMANSICS	IP: Manufacturing (SIC)
17	5	IPB51222s	IP: Residential Utilities
18	5	IPFUELS	IP: Fuels
19	1	NAPMPI	ISM Manufacturing: Production Index
20	2	CUMFNS	Capacity Utilization: Manufacturing

Labor Market			
id	tcode	fred	description
21*	2	HWI	Help-Wanted Index for United States
22*	2	HWIURATIO	Ratio of Help Wanted/No. Unemployed
23	5	CLF16OV	Civilian Labor Force
24	5	CE16OV	Civilian Employment
25	2	UNRATE	Civilian Unemployment Rate
26	2	UEMPMEAN	Average Duration of Unemployment (Weeks)
27	5	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks
28	5	UEMP5TO14	Civilians Unemployed for 5-14 Weeks
29	5	UEMP15OV	Civilians Unemployed - 15 Weeks & Over
30	5	UEMP15T26	Civilians Unemployed for 15-26 Weeks
31	5	UEMP27OV	Civilians Unemployed for 27 Weeks and Over
32*	5	CLAIMSx	Initial Claims
33	5	PAYEMS	All Employees: Total nonfarm
34	5	USGOOD	All Employees: Goods-Producing Industries
35	5	CES1021000001	All Employees: Mining and Logging: Mining
36	5	USCONS	All Employees: Construction
37	5	MANEMP	All Employees: Manufacturing
38	5	DMANEMP	All Employees: Durable goods
39	5	NDMANEMP	All Employees: Nondurable goods
40	5	SRVPRD	All Employees: Service-Providing Industries
41	5	USTPU	All Employees: Trade, Transportation & Utilities
42	5	USWTRADE	All Employees: Wholesale Trade
43	5	USTRADE	All Employees: Retail Trade
44	5	USFIRE	All Employees: Financial Activities
45	5	USGOVT	All Employees: Government
46	1	CES0600000007	Avg Weekly Hours : Goods-Producing
47	2	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing
48	1	AWHMAN	Avg Weekly Hours : Manufacturing
49	1	NAPMEI	ISM Manufacturing: Employment Index
127	6	CES0600000008	Avg Hourly Earnings : Goods-Producing
128	6	CES2000000008	Avg Hourly Earnings : Construction
129	6	CES3000000008	Avg Hourly Earnings : Manufacturing

Housing			
id	tcode	fred	description
50	4	HOUST	Housing Starts: Total New Privately Owned
51	4	HOUSTNE	Housing Starts, Northeast
52	4	HOUSTMW	Housing Starts, Midwest
53	4	HOUSTS	Housing Starts, South
54	4	HOUSTW	Housing Starts, West
55	4	PERMIT	New Private Housing Permits (SAAR)
56	4	PERMITNE	New Private Housing Permits, Northeast (SAAR)
57	4	PERMITMW	New Private Housing Permits, Midwest (SAAR)
58	4	PERMITS	New Private Housing Permits, South (SAAR)
59	4	PERMITW	New Private Housing Permits, West (SAAR)
Consumption, Orders and inventories			
3	5	DPCERA3M086SBEA	Real personal consumption expenditures
4*	5	CMRMTSPLx	Real Manu. and Trade Industries Sales
5*	5	RETAILx	Retail and Food Services Sales
60	1	NAPM	ISM : PMI Composite Index
61	1	NAPMNOI	ISM : New Orders Index
62	1	NAPMSDI	ISM : Supplier Deliveries Index
63	1	NAPMII	ISM : Inventories Index
64	5	ACOGNO	New Orders for Consumer Goods
65*	5	AMDMNOx	New Orders for Durable Goods
66*	5	ANDENOx	New Orders for Nondefense Capital Goods
67*	5	AMDMUOx	Unfilled Orders for Durable Goods
68*	5	BUSINVx	Total Business Inventories
69*	2	ISRATIOx	Total Business: Inventories to Sales Ratio
130*	2	UMCSENTx	Consumer Sentiment Index

Money and Credit

id	tcode	fred	description
70	6	M1SL	M1 Money Stock
71	6	M2SL	M2 Money Stock
72	5	M2REAL	Real M2 Money Stock
73	6	AMBSL	St. Louis Adjusted Monetary Base
74	6	TOTRESNS	Total Reserves of Depository Institutions
75	7	NONBORRES	Reserves Of Depository Institutions
76	6	BUSLOANS	Commercial and Industrial Loans
77	6	REALLN	Real Estate Loans at All Commercial Banks
78	6	NONREVSL	Total Nonrevolving Credit
79*	2	CONSPI	Nonrevolving consumer credit to Personal Income
131	6	MZMSL	MZM Money Stock
132	6	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding
133	6	DTCTHFNM	Total Consumer Loans and Leases Outstanding
134	6	INVEST	Securities in Bank Credit at All Commercial Banks
Interest Rates and Exchange Rates			
84	2	FEDFUNDS	Effective Federal Funds Rate
85*	2	CP3Mx	3-Month AA Financial Commercial Paper Rate
86	2	TB3MS	3-Month Treasury Bill:
87	2	TB6MS	6-Month Treasury Bill:
88	2	GS1	1-Year Treasury Rate
89	2	GS5	5-Year Treasury Rate
90	2	GS10	10-Year Treasury Rate
91	2	AAA	Moody's Seasoned Aaa Corporate Bond Yield
92	2	BAA	Moody's Seasoned Baa Corporate Bond Yield
93*	1	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS
94	1	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS
95	1	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS
96	1	T1YFFM	1-Year Treasury C Minus FEDFUNDS
97	1	T5YFFM	5-Year Treasury C Minus FEDFUNDS
98	1	T10YFFM	10-Year Treasury C Minus FEDFUNDS
99	1	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS
100	1	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS
101	5	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies
102*	5	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate
103*	5	EXJPUSx	Japan / U.S. Foreign Exchange Rate
104*	5	EXUSUKx	U.S. / U.K. Foreign Exchange Rate
105*	5	EXCAUSx	Canada / U.S. Foreign Exchange Rate

Prices			
id	tcode	fred	description
106	6	WPSFD49207	PPI: Finished Goods
107	6	WPSFD49502	PPI: Finished Consumer Goods
108	6	WPSID61	PPI: Intermediate Materials
109	6	WPSID62	PPI: Crude Materials
110*	6	OILPRICEx	Crude Oil, spliced WTI and Cushing
111	6	PPICMM	PPI: Metals and metal products:
112	1	NAPMPRI	ISM Manufacturing: Prices Index
113	6	CPIAUCSL	CPI : All Items
114	6	CPIAPPSL	CPI : Apparel
115	6	CPITRNSL	CPI : Transportation
116	6	CPIMEDSL	CPI : Medical Care
117	6	CUSR0000SAC	CPI : Commodities
118	6	CUUR0000SAD	CPI : Durables
119	6	CUSR0000SAS	CPI : Services
120	6	CPIULFSL	CPI : All Items Less Food
121	6	CUUR0000SA0L2	CPI : All items less shelter
122	6	CUSR0000SA0L5	CPI : All items less medical care
123	6	PCEPI	Personal Cons. Expend.: Chain Index
124	6	DDURRG3M086SBEA	Personal Cons. Exp: Durable goods
125	6	DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods
126	6	DSERRG3M086SBEA	Personal Cons. Exp: Services
Stock Market			
80*	5	S&P 500	S&P's Common Stock Price Index: Composite
81*	5	S&P: indust	S&P's Common Stock Price Index: Industrials
82*	2	S&P div yield	S&P's Composite Common Stock: Dividend Yield
83*	5	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio
135*	1	VXOCLSx	VXO