Combining Factor Models and External Instruments to Identify Uncertainty Shocks

Martin Bruns[*]

October 2018 - Job Market Paper
Most recent version 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 which are commonly used to explore the real and nominal effects of uncertainty shocks. I propose a Bayesian Proxy Factor-Augmented VAR (BP-FAVAR) to jointly address both issues. I find that real economic activity drops and rebounds following an identified uncertainty shock. The price reaction, while negative in the short run, is indistinguishable from zero after six months. Informational insufficiency issues, while detected by a statistical test, do not qualitatively alter these results.

1 Introduction

Following the seminal paper by Bloom (2009) a fast growing literature analyses the macroeconomic impact of exogenous increases in uncertainty using structural VAR models. An increase in uncertainty is broadly defined as increased difficulties of economic agents to make accurate forecasts. Within this literature, there is a consensus that exogenous increases in uncertainty lead to adverse real effects. These include falling production, hours, and employment. However, there is an ongoing debate about whether these reactions are dominated by supply or demand channels, i.e. whether they are accompanied by a rise or fall in inflation. These nominal reactions are crucial for policy makers: if, for example, central banks know that a rise in uncertainty leads to a decrease in inflation, they can, theoretically, move both real and nominal variables back to their desired targets by employing an expansionary policy. If, on the other hand, uncertainty shocks do not affect prices or are inflationary, central banks are faced with a trade-off between allowing for more inflation or a decline in real activity.

In this paper, I propose a Bayesian Proxy Factor-augmented VAR (BP-FAVAR) model to analyse the real and nominal effects following an uncertainty shock. This novel model offers a unified framework to combine a large information set with a non-recursive identification strategy. It addresses two shortcomings in commonly used small-scale, recursively identified, VAR models: (i) informational insufficiency and (ii) non-credible identification. I find that inflation responds negatively to a positive uncertainty shock in the short run and is indistinguishable from zero after six months. This is comforting news for policy makers given that they can address both real and nominal effects using standard instruments. The dynamic effects depend strongly on the identification scheme. Biases resulting from a recursive scheme cannot be alleviated by augmenting the information set of the model.

The structural VAR literature is inconclusive about the inflationary effects of uncertainty shocks. Leduc and Liu (2016), using a four-variable, recursively identified VAR model, find that uncertainty shocks are deflationary, even in the medium term. Piffer and Podstawski (forthcoming), identifying the work-horse model by Bloom

2

(2009) via an external instrument, find a short-lived drop and fast rebound in prices. Caggiano, Castelnuovo and Nodari (2017), extending a small-scale VAR model to a non-linear setting, find that uncertainty shocks are deflationary only in recessions and have no effect on prices in expansions. Caggiano, Castelnuovo and Pellegrino (2017), employing an interacted VAR model, find that the price reaction is indistinguishable from zero over the whole impulse horizon.

At best, the theoretical literature provides limited guidance for the inflationary effects of uncertainty shocks. Fernández-Villaverde et al. (2015) and Born and Pfeifer (2014) put forward two opposing channels to explain the potential price reaction following an exogenous increase in policy uncertainty: On the one hand, if consumers face difficulties predicting the next period, they will postpone consumption decisions, which will lead to a fall in both economic activity and prices. Therefore, an uncertainty shock would resemble an aggregate demand shock. Using a model with labour market frictions, Leduc and Liu (2016) also reach this conclusion. On the other hand, if firms face difficulties predicting the next period, they will bias their price decision upwards. The reason is that their profit function is concave in prices, making it more costly to set prices too low rather than too high. If this "price-bias-channel" dominates, the reactions of real and nominal variables will resemble a short-lived aggregate supply shock. The aggregate reaction of prices from these two opposing channels is ambiguous.

The responses to an uncertainty shock may depend heavily on the information set of the model, as pointed out by Caggiano, Castelnuovo and Nodari (2017) and Angelini et al. (2018). In particular, omitted variables, such as consumer sentiment (Sims, 2012), total factor productivity (Bachmann and Bayer, 2013), and measures of anticipated risk (Christiano et al., 2014) may bias the impulse responses. In order to avoid having to add potentially omitted variables one by one, I augment the work-horse VAR model by Bloom (2009) with latent factors. These summarise the information contained in a large set of variables and, thus, should alleviate omitted variable biases. Second, as pointed out by Stock and Watson (2012), a recursive identification scheme may be invalid given the contemporaneous interrelation between the real economy and uncertainty as well as the fast moving nature of financial markets.

Therefore, a recent strand of the literature addresses identification issues by departing from recursive schemes and employing external instruments. I follow Piffer and Podstawski (forthcoming) in identifying an uncertainty shock using a proxy based on the price of gold. This proxy captures movements in the price of gold around selected economic and political events. These events are associated with movements in uncertainty. Given that gold can be considered a safe haven asset, these movements should capture exogenous variations in uncertainty.

Estimation of the model is subject to two challenges. The first regards the so-called "curse of dimensionality". Even after shrinking the variable space using latent factors, the model still contains a large number of parameters. The baseline model consists of over 800 parameters, while the effective sample length is only roughly 400. Therefore, estimation is infeasible in a frequentist setting. A second challenge arises from the need to effectively summarise the estimation uncertainty in both the model parameters and the latent factors. This is difficult using bootstrap techniques (see for example Yamamoto, forthcoming). In addition, there are no asymptotic results justifying the use of such techniques, as pointed out by Kilian and Lütkepohl (2017). To jointly address these two challenges I employ a Bayesian approach. It allows for overcoming dimensionality problems by shrinking the parameter space and summarises the estimation uncertainty in a joint posterior distribution. The BP-FAVAR can be considered a combination of the Bayesian FAVAR estimation proposed by Belviso and Milani (2006) and the Bayesian Proxy VAR by Caldara and Herbst (forthcoming). I re-parametrize their model to impose structure on the impact effects of shocks.

The main results are the following: Uncertainty shocks are deflationary in the short run. The price reaction is indistinguishable from zero after about six months. Real variables and the stock market drop and rebound. This suggests that policy makers can employ an expansionary policy to alleviate the adverse effects of an exogenous increase of uncertainty on both prices and real activity. I show evidence that the workhorse model by Bloom (2009) is informationally deficient. When computing impulse responses, I find that alleviating informational deficiency problems has only marginal quantitative effects. This finding is in line with Sims (2012), who points

4

out that informational deficiency is not an either/or but a quantitative issue. In the present case the biases it causes are negligible, which is comforting news.

The remainder of the paper is organized as follows: Section 2 introduces the model setup and explains the identification of uncertainty shocks. Section 3 presents the data and discusses the results. The last section concludes.

2   The Bayesian Proxy FAVAR

In this section, I introduce the Bayesian Proxy FAVAR model. I start by describing the different parts of the model. Then, I discuss how partial identification is achieved. Last, I show how the model can be decomposed into three blocks to facilitate inference.

2.1   Model Description

The Bayesian Proxy Factor-augmented VAR model admits a state-space form, which consists of an observation equation, a transition equation and a proxy equation. First, consider the observation equation, which shows how latent and observable factors map into informational series:

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

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

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 $\boldsymbol{z}_t$. $\boldsymbol{\Lambda}^f$ is a $N \times R$ matrix of factor loadings for latent factors and $\boldsymbol{\Lambda}^z$ 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. $Var(\boldsymbol{\xi}_t) = \boldsymbol{\Omega}$ is assumed to be diagonal.

Next, consider the transition equation which shows the dynamic evolution of the

factors. It writes as a VAR(P) of the following form:

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

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

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 $\boldsymbol{\Pi} = [\boldsymbol{\Pi}_1, ..., \boldsymbol{\Pi}_P]$ of dimension $R + K \times (P(R + K) + 1)$ contains the autoregressive parameters of the VAR. $\boldsymbol{w}_t = [\boldsymbol{1}_{R+K\times1}; \boldsymbol{y}_{t-1}, ..., \boldsymbol{y}_{t-P}]$ 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. $Cov(\boldsymbol{u}_t, \boldsymbol{u}_{t-p}) = 0 \quad \forall t = 1, ..., T, \forall p = 1, ..., \infty$. Also, $\boldsymbol{u}_t$ 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:

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

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

where $\boldsymbol{B}$ 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: $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}'$

As is well known, further restrictions beyond those implied by the covariance matrix are needed to identify $\boldsymbol{B}$. The reason is that the data cannot discriminate between observationally equivalent representations: All $\boldsymbol{B}$ such that $\boldsymbol{B}\boldsymbol{B}' = \boldsymbol{\Sigma}$ yield the same likelihood. Therefore, without imposing further structure, the econometrician cannot distinguish between $\boldsymbol{B}$ and $\tilde{\boldsymbol{B}} = \boldsymbol{B}\boldsymbol{Q}$, where $\boldsymbol{Q}$ is an orthogonal matrix such that $\boldsymbol{Q}\boldsymbol{Q}' = \boldsymbol{I}$.

In order to identify the first column of $\boldsymbol{B}$, which I denote $\boldsymbol{b}$, I augment the model by a "Proxy Equation", as in Caldara and Herbst (forthcoming). It spells out the

relation between structural shock and instrument and is given as[1]:

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

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

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, $\sigma_\nu^2 s$, the less information the instrument contains about the shock of interest.

The full model can be written in compact matrix notation as:

$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{z}_t \\ m_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}^f & \boldsymbol{\Lambda}^z & \boldsymbol{0}_{N \times 1} \\ \boldsymbol{0}_{K \times R} & \boldsymbol{I}_K & \boldsymbol{0}_{K \times 1} \\ \boldsymbol{0}_{1 \times R} & \boldsymbol{0}_{1 \times K} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{f}_t \\ \boldsymbol{z}_t \\ m_t \end{bmatrix} + \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{0}_{K \times 1} \\ 0 \end{bmatrix} \tag{9}$$

$$Var(\begin{bmatrix} \boldsymbol{\xi}_t \end{bmatrix}) = \boldsymbol{\Omega} \tag{10}$$

$$\begin{bmatrix} \boldsymbol{y}_t \\ m_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Pi}(L) \\ \boldsymbol{0}_{1 \times P(K+R)} \end{bmatrix} \boldsymbol{w}_t + \mathcal{B} \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \nu_t \end{bmatrix} \tag{11}$$

$$Var(\begin{bmatrix} \boldsymbol{\epsilon}_t \\ \nu_t \end{bmatrix}) = \begin{bmatrix} \boldsymbol{I}_{R+K} & \boldsymbol{0}_{R+K \times 1} \\ \boldsymbol{0}_{1 \times R+K} & 1 \end{bmatrix}, \tag{12}$$

---

[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 $\boldsymbol{B}$, rather than on $\boldsymbol{B}^{-1}$. Caldara and Herbst (forthcoming) 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 context of uncertainty shocks, however, it is more common to inform the on-impact effects of shocks (see e.g. Bloom, 2009 or Caggiano, Castelnuovo and Nodari, 2017). Therefore, I propose to use a so-called B-model, which imposes structure on the on-impact effects.

where $\mathcal{B} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{\beta} \\ \boldsymbol{\beta}' & \sigma_\nu \end{bmatrix}$, and $\boldsymbol{\beta} = \begin{bmatrix} \beta \\ \boldsymbol{0}_{R+K-1\times 1} \end{bmatrix}$.

2.2 Identification

Shock identification in the BP-FAVAR model is achieved by weighting draws from the posterior of structural parameters. In particular, more weight is given to posterior draws which lead to a close relation between instrument and the shock of interest. To be more precise, consider the joint likelihood of $\boldsymbol{x}_t$, $\boldsymbol{y}_t$ and $\boldsymbol{m}_t$ given a draw of the latent factors[2]:

$$
p(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{m}|\boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}^f, \boldsymbol{\Lambda}^z, \boldsymbol{\Omega}, \beta, \sigma_\nu, \boldsymbol{b}) \tag{13}
$$
$$
=p(\boldsymbol{Y}|\boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}^f, \boldsymbol{\Lambda}^z, \boldsymbol{\Omega})
$$
$$
\cdot (\boldsymbol{m}|\boldsymbol{Y}, \boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}^f, \boldsymbol{\Lambda}^z, \boldsymbol{\Omega}, \beta, \sigma_\nu, \boldsymbol{b})
$$
$$
\cdot (\boldsymbol{X}|\boldsymbol{m}, \boldsymbol{Y}, \boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}^f, \boldsymbol{\Lambda}^z, \boldsymbol{\Omega})
$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, ..., \boldsymbol{x}_T]$, $\boldsymbol{Y} = [\boldsymbol{y}_1, ..., \boldsymbol{y}_T]$ and $\boldsymbol{m} = [m_1, ..., m_T]$ stack the observational series, the factors, and the instrument horizontally. Note that, while the marginal likelihood of $\boldsymbol{Y}$ and the conditional likelihood of $\boldsymbol{X}|\boldsymbol{m}, \boldsymbol{Y}$ depend only on reduced form parameters, the conditional likelihood of $\boldsymbol{m}|\boldsymbol{Y}$ depends, in addition, on structural parameters, β, σ_ν and $\boldsymbol{b}$.

The conditional likelihood of $\boldsymbol{m}|\boldsymbol{Y}$ can be written as (see Appendix for deriva-

---

[2]The factors are identified only up to an invertible rotation, i.e. the representations $\boldsymbol{x}_t = \boldsymbol{\Lambda}\boldsymbol{y}_t + \boldsymbol{\xi}_t$ and $\boldsymbol{x}_t = \boldsymbol{\Lambda}\boldsymbol{P}\boldsymbol{P}^{-1}\boldsymbol{y}_t + \boldsymbol{\xi}_t$ yield the same likelihood. Therefore, in order to achieve identification, one has to impose further restrictions. I follow Bernanke et al. (2005) and set the upper $R \times K$ block of $\boldsymbol{\Lambda}^z$ equal to a zero matrix. Furthermore, $\boldsymbol{\Sigma}^f = Cov(\boldsymbol{f}_t)$ is diagonal and $\frac{\boldsymbol{\Lambda}^{f'}\boldsymbol{\Lambda}^f}{N} = \boldsymbol{I}_R$. This normalisation, although not necessary, is sufficient to pin down the factor rotation, as pointed out by Kilian and Lütkepohl (2017).

tion):

$$\boldsymbol{m}|\boldsymbol{Y} \sim N(\boldsymbol{\mu}_{m|Y}, \boldsymbol{V}_{m|Y}) \tag{14}$$

$$\boldsymbol{\mu}_{m|Y} = \beta\boldsymbol{\epsilon}_1 \tag{15}$$

$$\boldsymbol{V}_{m|Y} = \sigma_\nu^2 \boldsymbol{I}_T, \tag{16}$$

where $\boldsymbol{\epsilon}_1 = [\epsilon_{1,1}, ..., \epsilon_{1,T}]$ stacks the structural shocks of interest in a vector.

As seen in equation (15), the conditional likelihood of $\boldsymbol{m}$ is higher the closer are $\boldsymbol{m}$ and $\beta\boldsymbol{\epsilon_1}$. In the posterior sampler, draws are weighted by the conditional likelihood of $\boldsymbol{m}$ (see Appendix C). Therefore, the econometrician will give more weight to posterior draws, which result in structural errors that look like a scaled version of the proxy.

As is apparent from equation (15), since $\boldsymbol{\epsilon}_1$ is obtained from reduced form errors, identification depends heavily on 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 helps alleviate this problem without taking a stand on which of a potentially large set of observational series need to be included.

Compared to recursively identified VARs, the BP-FAVAR has the advantage that when using a Proxy VAR, the researcher is not forced to employ potentially non-credible short run exclusion restrictions. For example, the recursive workhorse model on the empirical identification of uncertainty shocks by Bloom (2009) relies on the assumption that financial markets do not price an exogenous increase in uncertainty within a month. This might be too strong an assumption given the fast-moving nature of financial markets. In the context of Proxy VAR models such short run exclusion restrictions are substituted by external information contained in the proxy variable, m_t.

9

2.3 Inference

The Bayesian approach treats all model parameters and latent factors as random variables whose posterior needs to be sampled from. In order to outline the sampling procedure, first define the parameter space as

$$\boldsymbol{\theta} = (\boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}^f, \boldsymbol{\Lambda}^z, \boldsymbol{\Omega}, \beta, \sigma_\nu, \boldsymbol{b}) \tag{17}$$

The joint posterior of parameters and latent factors is:

$$p(\boldsymbol{\theta}, \boldsymbol{F} | \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{m}), \tag{18}$$

where $\boldsymbol{F} = [\boldsymbol{f}_1, ..., \boldsymbol{f}_T]$, $\boldsymbol{Z} = [\boldsymbol{z}_1, ..., \boldsymbol{z}_T]$. The challenge consists in approximating the marginal posterior distributions of the latent factors,

$$p(\boldsymbol{F} | \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{m}) = \int_{\boldsymbol{\theta}} p(\boldsymbol{\theta}, \boldsymbol{F} | \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{m}) d\boldsymbol{\theta} \tag{19}$$

and the model parameters,

$$p(\boldsymbol{\theta} | \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{m}) = \int_{\boldsymbol{F}} p(\boldsymbol{\theta}, \boldsymbol{F} | \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{m}) d\boldsymbol{F} \tag{20}$$

It is shown by Geman and Geman (1984) that these integrals can be approximated using a multi-move Gibbs sampler, which alternately draws from two distributions: First, draw the latent factors given all model parameters and the data, i.e.

$$p(\boldsymbol{F} | \boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{m}). \tag{21}$$

This draw is generated using filtering techniques. Second, draw the model parameters conditioning on this draw of factors and the data, i.e.[3]

$$p(\boldsymbol{\theta}|\boldsymbol{Z}, \boldsymbol{X}, \boldsymbol{m}). \tag{22}$$

This draw is generated using a Metropolis-within-Gibbs algorithm as in Caldara and Herbst (forthcoming).

Compared to the common approach of first extracting factors via Principal Components and then feeding them into a VAR (see Stock and Watson, 2016 for a review), this Bayesian approach has the advantage that it allows for Bayesian shrinkage of the parameter space. This might seem unnecessary given that the factors already reduce the dimensionality of the estimation problem. However, if, as in the present case, the number of observable factors, K, or the lag length, P, is large, dimensionality issues still arise and can be alleviated using Bayesian shrinkage.[4] Furthermore, as shown by Yamamoto (forthcoming), bootstrap inference in frequentist factor models is far from trivial. In particular, it remains an open issue how to account for the estimation uncertainty in the factors. 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.

Conditional posterior densities of latent factors $\boldsymbol{F}$ given $\boldsymbol{\theta}$: The procedure to generate posterior draws of latent factors, $\boldsymbol{F}$, differs from generating draws of parameters, $\boldsymbol{\theta}$, 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

---

[3]I follow Caldara and Herbst (forthcoming) and generate draws from $p(\boldsymbol{F}|\boldsymbol{\theta}, \boldsymbol{X})$ and $p(\boldsymbol{\theta}|\boldsymbol{Z}, \boldsymbol{X})$ first and account for the additional conditioning on $\boldsymbol{m}$ using an independence Metropolis-Hastings step. See Appendix C for details.

[4]The number of parameters to be estimated in the transition equation is

$$\begin{aligned}
&(R + K)(1 + (R + K)P) + (R + K)^2 \\
=&(4 + 8)(1 + (4 + 8)5) + (4 + 8)(4 + 8 + 1)/2 \\
=&810.
\end{aligned}$$

The sample length is $T = 438$.

described in equation ([3](#)) as follows:

$$p(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}) = p(\boldsymbol{y}_t|\boldsymbol{X},\boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\boldsymbol{y}_t|\boldsymbol{y}_{t+1},\boldsymbol{X},\boldsymbol{\theta}). \tag{23}$$

First note that ([23](#)) describes the posterior of $\boldsymbol{Y}$, 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.[5] 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) (see Appendix E for details). This approach, while straightforward to implement, increases the computational burden slightly compared to Principal Components Analysis. However, it allows incorporating the estimation uncertainty in the latent factors in a consistent way.

Conditional posterior densities of the parameters $\boldsymbol{\theta}$ given latent factors $\boldsymbol{F}$: In order to draw the model parameters $\boldsymbol{\theta}$ given the data and a draw of the factors, $\boldsymbol{Y}$, I form three blocks of parameters: Block 1 refers to parameters of the observation equation ([1](#)), block 2 refers to the parameters of the transition equation ([3](#)), and block 3 refers to the parameters of the proxy equation ([7](#)). Conditional on a draw of the factors and the data the first two blocks can be sampled independently of each other while the last block is sampled conditional on the second block.

Block 1: Observation Equation The idiosyncratic errors $\boldsymbol{\xi}_t$ are assumed to be mutually uncorrelated and normally distributed. Group the factor loading matrices as $\boldsymbol{\Lambda} = [\boldsymbol{\Lambda}^f \quad \boldsymbol{\Lambda}^z]$. Then, conditional on a draw of the factors, we can specify conjugate normal-inverse Gamma priors and draw the posterior for $\boldsymbol{\Lambda}$ and $\boldsymbol{\Omega}$ equation-by-equation using well-known results on Bayesian linear regression models (see e.g.

---

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

Koop, 2003). For each equation i, specify the priors as:

$$\boldsymbol{\omega}_{ii} \sim IG(sc^*, sh^*) \tag{24}$$

$$\boldsymbol{\lambda}_i | \boldsymbol{\omega}_{ii} \sim N(\boldsymbol{\mu}_{\lambda,i}^*, \boldsymbol{\omega}_{ii} \boldsymbol{M}_i^{*-1}), \tag{25}$$

where $\boldsymbol{\lambda}_i$ is the i-th row of $\boldsymbol{\Lambda}$ and $\boldsymbol{\omega}_{ii}$ is the i-th diagonal element of $\boldsymbol{\Omega}$. These priors translate into posterior distributions of the following form:

$$\boldsymbol{\omega}_{ii} | \boldsymbol{X}, \boldsymbol{Y} \sim IG(\bar{sc}_i, \bar{sh}_i) \tag{26}$$

$$\lambda_i | \boldsymbol{\omega}_{ii}, \boldsymbol{X}, \boldsymbol{Y} \sim N(\bar{\mu}_{\lambda,i}, \boldsymbol{\omega}_{ii} \bar{\boldsymbol{M}}_i^{-1}) \tag{27}$$

with

$$\bar{sh}_i = sh^* + T \tag{28}$$

$$\bar{sc}_i = sc^* + \hat{\boldsymbol{\xi}}_i \hat{\boldsymbol{\xi}}_i{}' + (\hat{\boldsymbol{\lambda}}_i^{OLS} - \mu_{\lambda,i}^*)'(\bar{\boldsymbol{M}}_i^{-1} + (\boldsymbol{Y}\boldsymbol{Y}')^{-1})^{-1}(\hat{\boldsymbol{\lambda}}_i^{OLS} - \mu_{\lambda,i}^*) \tag{29}$$

$$\hat{\boldsymbol{\lambda}}_i^{OLS} = \boldsymbol{x}_i \boldsymbol{Y}'(\boldsymbol{Y}\boldsymbol{Y}')^{-1} \tag{30}$$

$$\hat{\boldsymbol{\xi}}_i = \boldsymbol{x}_i - \hat{\boldsymbol{\lambda}}_i^{OLS} \boldsymbol{Y} \tag{31}$$

$$\bar{\boldsymbol{M}}_i = \boldsymbol{M}_i^* + \boldsymbol{Y}\boldsymbol{Y}' \tag{32}$$

$$\bar{\mu}_{\lambda,i} = \bar{\boldsymbol{M}}_i(\boldsymbol{M}_i^{*-1}\mu_{\lambda,i}^* + \boldsymbol{Y}\boldsymbol{Y}'\hat{\boldsymbol{\lambda}}_i^{OLS}) \tag{33}$$

I follow Bernanke et al. (2005) in specifying $sc^* = 3$, $sh^* = 10^{-3}$, $\boldsymbol{M}_i^* = \boldsymbol{I}_{R+K}$ and $\mu_{\lambda,i}^* = \boldsymbol{0}_{(R+K)\times 1}$.

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

$$vec(\boldsymbol{\Pi}) \sim N(\boldsymbol{\mu}_{\boldsymbol{\Pi}}^*, \boldsymbol{V}_{\boldsymbol{\Pi}}^*) \tag{34}$$

$$\boldsymbol{\Sigma} \sim IW(\boldsymbol{S}^*, \tau^*), \tag{35}$$

where $vec(\cdot)$ is the vectorisation operator that stacks the column of a matrix one un-

13

derneath the other into a vector. These priors translate into the following conditional posterior for $vec(\mathbf{\Pi})$:

$$vec(\mathbf{\Pi})|\mathbf{\Sigma}, \mathbf{Y} \sim N(\bar{\boldsymbol{\mu}}_{\mathbf{\Pi}}, \bar{\boldsymbol{V}}_{\mathbf{\Pi}}) \tag{36}$$

$$\bar{\boldsymbol{V}}_{\mathbf{\Pi}} = (\boldsymbol{V}_{\mathbf{\Pi}}^{*-1} + (\boldsymbol{W}\boldsymbol{W}' \otimes \mathbf{\Sigma}^{-1}))^{-1} \tag{37}$$

$$\bar{\boldsymbol{\mu}}_{\mathbf{\Pi}} = \bar{\boldsymbol{V}}_{\mathbf{\Pi}}(\boldsymbol{V}_{\mathbf{\Pi}}^{*-1}\boldsymbol{\mu}_{\mathbf{\Pi}}^* + (\boldsymbol{W} \otimes \mathbf{\Sigma}^{-1})vec(\boldsymbol{Y})) \tag{38}$$

where $\boldsymbol{W} = [\boldsymbol{w}_1, ..., \boldsymbol{w}_T]$, $\boldsymbol{w}_t = [\mathbf{1} \quad \boldsymbol{y}_{t-1} \quad ... \quad \boldsymbol{y}_{t-p}]'$ stacks a vector of 1s and P lags of $\boldsymbol{y}_t$. The conditional posterior for $\mathbf{\Sigma}$ is given as:

$$\mathbf{\Sigma}|\mathbf{\Pi}, \mathbf{Y} \sim IW(\bar{\boldsymbol{S}}, \bar{\tau}) \tag{39}$$

$$\bar{\boldsymbol{S}} = \boldsymbol{S}^* + \boldsymbol{U}\boldsymbol{U}' \tag{40}$$

$$\bar{\tau} = \tau^* + T, \tag{41}$$

where $\boldsymbol{U} = [\boldsymbol{u}_1, ..., \boldsymbol{u}_T]$ stacks the reduced form errors given the current draw of factors. I set $\boldsymbol{\mu}_{\mathbf{\Pi}}^*$ to a zero vector given that all series are transformed to be stationary. $\boldsymbol{V}_{\mathbf{\Pi}}^*$ is a diagonal matrix containing the prior variances of the parameters contained in $\mathbf{\Pi}$. These are set in accordance with standard Minnesota values and given as:

$$v_{\mathbf{\Pi},i,j}^* = \begin{cases} (\lambda/l)^2 & \text{if} \quad i = j \\ (\lambda\sigma_i/l\sigma_j)^2 & \text{if} \quad i \neq j, \end{cases} \tag{42}$$

where σ_i is obtained from univariate AR(1) regressions and $\lambda = 0.2$. $\boldsymbol{S}^*$ is set to $\boldsymbol{I}_{R+K}$, while τ^* is set to $R + K$.

Block 3: Proxy Equation The parameters of the proxy equation are sampled conditional on the parameters of the transition equation. For β and σ_ν the priors are

$$\beta \sim N(\mu_\beta^*, \sigma_\beta^*) \tag{43}$$

$$\sigma_\nu \sim IG(sc_\nu^*, sh_\nu^*). \tag{44}$$

14

$\boldsymbol{b}$ is computed as $\boldsymbol{b} = chol(\boldsymbol{\Sigma})\boldsymbol{Q}_{\cdot,1}$ where $\boldsymbol{Q}_{\cdot,1}$ is the first column of a draw from the uniform Haar distribution (see Rubio-Ramirez et al., 2010 for a discussion).

These priors translate into the following posteriors for β and σ_ν:

$$\beta | \boldsymbol{Y}, \boldsymbol{m}, \boldsymbol{\Pi}, \boldsymbol{\Sigma}, \sigma_\nu, \boldsymbol{b} \sim N(\bar{\mu}_\beta, \bar{\sigma}_\beta) \tag{45}$$

$$\bar{\mu}_\beta = \boldsymbol{m}\boldsymbol{\epsilon_1}(\boldsymbol{\epsilon_1}\boldsymbol{\epsilon_1}')^{-1} \tag{46}$$

$$\bar{\sigma}_\beta = \sigma_\nu(\boldsymbol{\epsilon_1}\boldsymbol{\epsilon_1}')^{-1} \tag{47}$$

$$\sigma_\nu | \boldsymbol{Y}, \boldsymbol{m}, \boldsymbol{\Pi}, \boldsymbol{\Sigma}, \beta, \boldsymbol{b} \sim IG(\bar{sc}_\nu, \bar{sh}_\nu) \tag{48}$$

$$\bar{sc}_\nu = sc_\nu^* + (\boldsymbol{m} - \beta\boldsymbol{\epsilon_1})(\boldsymbol{m} - \beta\boldsymbol{\epsilon_1})' \tag{49}$$

$$\bar{sh}_\nu = sh_\nu^* + T \tag{50}$$

For $\boldsymbol{b}$ the conditional posterior has an unknown form. Therefore, $\boldsymbol{b}$ is sampled using a Metropolis-Hastings step. In particular, at iteration j, a draw $\boldsymbol{b}^{cand}$, will be accepted with probability (see Appendix C for details)

$$\alpha = min(\frac{p(\boldsymbol{m}|\boldsymbol{Y}, \boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{b}^{cand})}{p(\boldsymbol{m}|\boldsymbol{Y}, \boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{b}^{j-1})}, 1) \tag{51}$$

I follow Caldara and Herbst (forthcoming) in specifying the priors as $\mu_\beta^* = 0$, $\sigma_\beta^* = 1$, $sh_\nu^* = 2$, $sc_\nu^* = 0.2$ in order to allow the data to dominate the posterior. In particular, this prior specification implies a zero mean prior correlation between instrument and shock, ρ.

Posterior Sampler The sampler can be summarized as follows (see the Appendix C for a detailed step-by-step procedure):

1. Set starting values

2. Draw $\boldsymbol{y}_t$ via the Carter-Kohn algorithm

3. Draw $\boldsymbol{\Lambda}$ and $\boldsymbol{\Omega}$

4. Draw $\boldsymbol{\Sigma}$ and $\boldsymbol{\Pi}$

5. Draw **b** using a Metropolis-Hastings step

6. Draw β and σ_ν

These steps are repeated a sufficient number of times for the algorithm to converge (see Appendix E.3 for a discussion of the convergence properties of the sampler).

3   Data, Estimation and Results

This section first describes the data and the proxy used for identification. It then shows how the number of factors is determined with the goal of alleviating informational insufficiency issues. Lastly, it discusses how instrument relevance is assessed in the Bayesian context and presents impulse responses from the baseline model as well as from three benchmark models.

3.1   Data and Transformations

The baseline data contained in z_t are monthly US data from Piffer and Podstawski (forthcoming), who updated the data in Bloom (2009). The vector of observational series, x_t, contains 126 of the monthly FRED dataset by McCracken and Ng (2016) that are not already included in z_t (see Appendix F for a detailed description).[6] The length of the estimation sample is constrained by the instrument[7] and lasts from 1979M1 through 2015M7.[8] I follow Piffer and Podstawski (forthcoming) in setting the lag length to $P = 5$ as a baseline (see Appendix A for a robustness exercise setting $P = 9$).

One concern is measurement of the monetary policy stance. Traditionally, the effective 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

---

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

[7]Instrument available at https://sites.google.com/site/michelepiffereconomics/home/research-1

[8]An alternative to shortening the sample is put forward by Braun et al. (2017). They suggest generating synthetic observations for the instrument within their posterior sampler. This procedure implicitly requires parameter stability for time periods when the instrument is unavailable and it assumes that missing values occur at random.

to 2015M11, this variable cannot serve as an indicator of the policy stance. This is why, for this period, I replace the effective federal funds rate by the shadow rate as computed by Wu and Xia (2016). It is based on a term structure model and often considered a better reflection of the policy stance during the zero lower bound period than the federal funds rate. Figure 16 in Appendix F shows the shadow rate.

All informational series contained in $\boldsymbol{x}_t$ are transformed to have zero mean. In addition, they are transformed to induce stationarity as proposed by McCracken and Ng (2016). Missing values are replaced with zeros, which is the unconditional mean of the standardized series. These missing values occur mostly at the beginning of the dataset and amount to less than one percent of total observations. Therefore, the joint dynamics are unlikely to be overly affected by this imputation.

Piffer and Podstawski (forthcoming) argue that a proxy for the uncertainty shock could be based on the price of gold. The intuition behind this idea is that gold is considered a safe haven asset that investors choose in times of heightened uncertainty. This generates movements in the price of gold. The challenge consists in finding price variations that are not correlated with structural shocks other than the uncertainty shock (exogeneity condition). In order to achieve this, the authors collect a series of 38 events, which are considered to be associated with movements in uncertainty (e.g. the fall of the Berlin Wall or the 9/11 terrorist attacks). They then compute the variation of the gold price in narrow windows around these events and argue that these variations are driven exclusively by movements in uncertainty. In addition, they show that this proxy has a low correlation with other structural shocks as computed by Stock and Watson (2012), which is further evidence for exogeneity to their system.

Figure 1 shows the proxy. It peaks during well-known events such as the 9/11 terrorist attacks in 2001 or the bankruptcy of Lehman brothers in 2008.

3.2   Determining the Number of Factors

Choosing the number of latent factors, R, has important consequences for the amount of additional information the BP-FAVAR is based on, compared to the small-scale
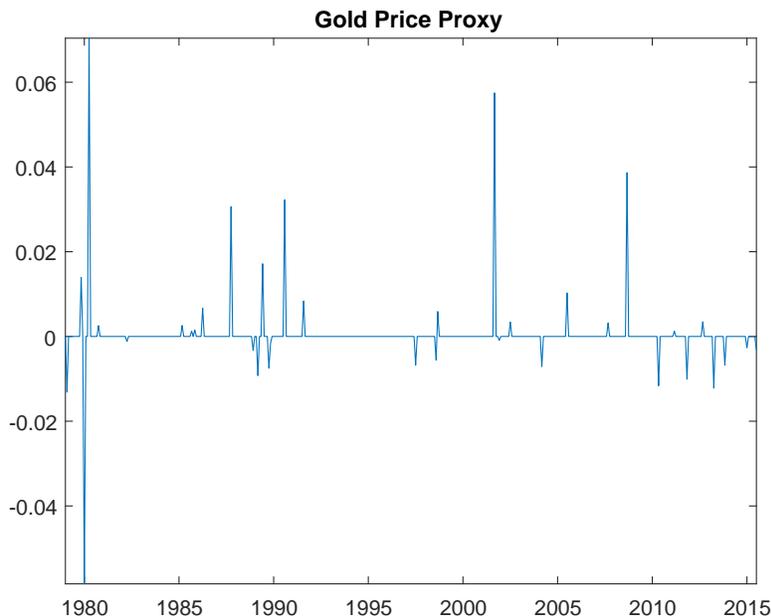
Figure 1: **Uncertainty Proxy.** *Gold price variation around selected events. The sample period is 1979M1 to 2015M7.*

VAR model employed in Bloom (2009) and Piffer and Podstawski (forthcoming). Stock and Watson (2016) stress the importance of using different statistical estimators to determine R. I employ two complementary tools: A scree plot and a sequential testing procedure, as proposed by Forni and Gambetti (2014). The choice of these two criteria is driven by the main purpose of the factors, which is to align the econometrician's and the economic agent's information sets.

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. Figure 2 plots this marginal contribution against the number of factors. It shows that the first factor explains about 60% of variance in $\boldsymbol{x}_t$, while the first four factors explain over 95% of the variance.

The sequential testing procedure chooses the number of factors R so as to ensure that the resulting BP-FAVAR model aligns the econometrician's and the agent's information set. This would not be the case in the presence of omitted variables.
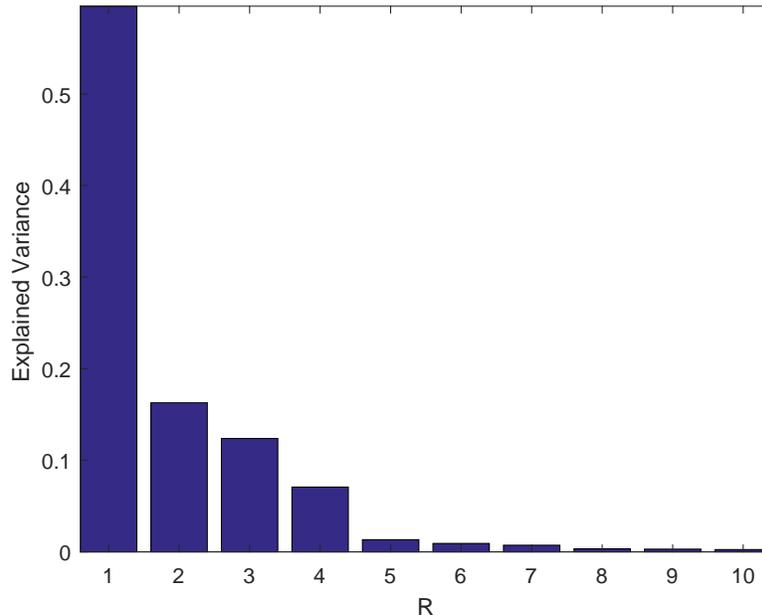
18

Figure 2: **Scree Plot.** *Explained share of variance in $\boldsymbol{x}_t$ as a function of the number of latent factors (R) included in $\boldsymbol{f}_t$*

This procedure is particularly suited when the list of potentially omitted variables is large and the researcher therefore would like to avoid taking a stand on which additional variables to include. In the context of uncertainty shocks, variables which are typically omitted from small-scale VAR models but could potentially be relevant, include consumer sentiment (Bachmann and Sims, 2012), total factor productivity (Bachmann and Bayer, 2013), and measures of anticipated risk (Christiano et al., 2014). Other forward-looking variables could also potentially cause omitted variable biases.

Instead of including one variable after the other in the VAR the sequential testing procedure augments the small-scale VAR model by latent factors extracted from a large set of informational series until the model is informationally sufficient. The model is informationally sufficient if none of the observable variables is Granger-caused by factors. The basis is a multivariate out-of-sample Granger-causality test (see Appendix D for details). The intuition behind this test is the following: If the

19

economy is accurately represented by a factor model, as is assumed here, then the factors contain all relevant information that agents base their decision making on. If these factors do not help predict a vector of variables in the VAR, then the variables in the VAR contain the same information as the factors. Thus, they are sufficient to align the econometrician's and the agent's information set. If, on the other hand, the factors do help predict the VAR variables, then they should be subsequently added to the VAR as additional variables until informational sufficiency cannot be rejected any longer. The sequential testing procedure is as follows:

In a first step, I test

$$H_0 : \boldsymbol{f}_t^{PC} \text{ do not Granger-cause } \boldsymbol{z}_t \tag{52}$$

$$H_1 : \boldsymbol{f}_t^{PC} \text{ Granger-cause } \boldsymbol{z}_t, \tag{53}$$

where $\boldsymbol{f}_t^{PC}$ is a vector containing the first six Principal Components extracted from $\boldsymbol{x}_t$. I then test, for $j = 1, ..., 5$

$$H_0 : \boldsymbol{f}_{t,-(1:j)}^{PC} \text{ do not Granger-cause } \{\boldsymbol{z}_t, \boldsymbol{f}_{t,1:j}\} \tag{54}$$

$$H_1 : \boldsymbol{f}_{t,-(1:j)}^{PC} \text{ Granger-cause } \{\boldsymbol{z}_t, \boldsymbol{f}_{t,1:j}\}. \tag{55}$$

where $\boldsymbol{f}_{t,1:j}^{PC}$ are the first j Principal Components of $\boldsymbol{x}_t$ and $\boldsymbol{f}_{t,-(1:j)}^{PC}$ is a vector containing all but the first j Principal Components.

Figure 3 shows the distribution of the test statistic across 1000 samples obtained via a standard residual bootstrap procedure together with the test statistic computed using the actual sample data. If the actual test statistic lies outside the bootstrap distribution test statistic, then this indicates that the Null of no Granger causality can be rejected.

Table 1 shows the corresponding p-values. It suggests that for $R = 0$, informational sufficiency can be rejected. This can be considered evidence that the workhorse model of Bloom (2009) is, indeed, informationally deficient and needs to be extended in order to alleviate potential omitted variables biases. The Null Hypothesis of informational sufficiency cannot be rejected once the model is augmented by at least
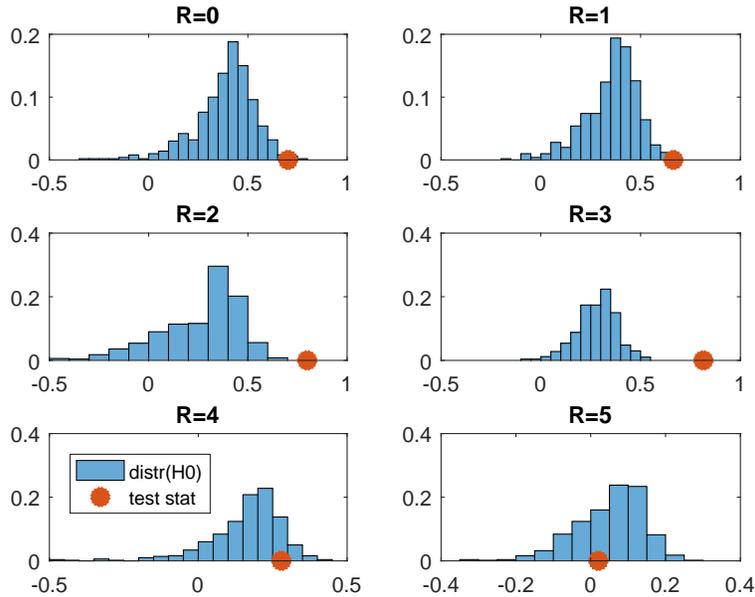
20

Figure 3: **Test for Informational Sufficiency.** *The histogram shows the bootstrap test statistic under the Null of no Granger-causality. It is based on a multivariate one-step-ahead out-of-sample Granger-causality test with lag length 4. The bootstrap test statistic is based on 1000 replications. The sample is split as $T = T_1 + T_2$, $T_1 = T_2 = 0.5T$. A larger test statistic indicates Granger-causality. Rejection of the Null of no Granger-causality indicates informational deficiency.*

four factors. Therefore, the multiple testing procedure suggests using $R = 4$.

Combining the information from the scree plot and the sequential testing procedure, I choose to set $R = 4$.

Figure 4 shows the last accepted draw of estimated factors from the Bayesian estimation, together with the Principal Components estimation used as a starting value. The two estimation procedures yield similar results suggesting that Principal Components can be used as an approximation to determine R.

3.3   Relevance of the Instrument

The instrument is appropriate to identify the uncertainty shock to the extent that it contains enough information about the shock, i.e. it is relevant. The relevance of the

R	p-value
0	0.0000
1	0.0020
2	0.0000
3	0.0000
4	0.1240
5	0.6680

Table 1: **Test for Informational Sufficiency.** *p-values are obtained as the fraction of bootstrap test statistics under the Null of no Granger causality exceeding the actual test statistic.*
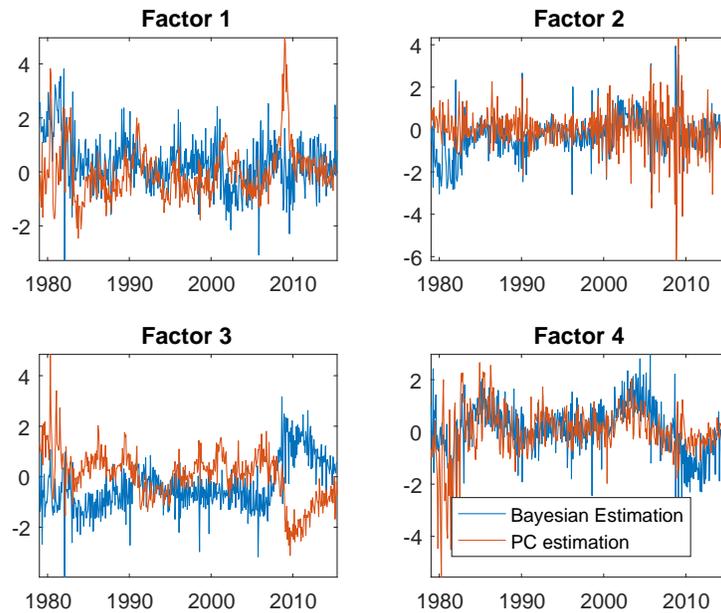

Figure 4: **Factors** *Last accepted draw from posterior sampler and Principal Components estimation of factors*
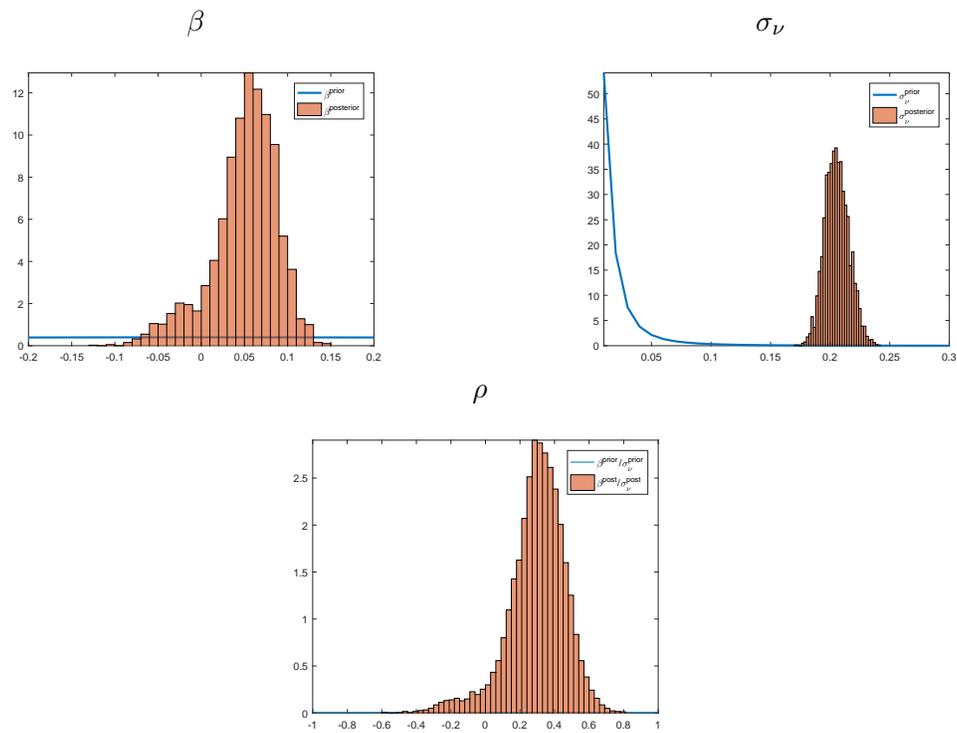
Figure 5: **Instrument Relevance** *Updating of β (top left), σ_ν (top right) and β/σ_ν (bottom panel)*

instrument in the Bayesian context is assessed by analysing the posterior updating of the proxy equation. A high posterior correlation between instrument and shock suggests relevance.

Figure 5 shows the updating of the relevant quantities β, σ_ν and their ratio β/σ_ν. This ration is the signal-to-noise ratio and measure how much information the instrument contains about the shock of interest.

The top left panel shows that while using a prior for β that is flat over the relevant parameter space, the posterior is centred around 0.1 suggesting that the data support a structural relationship between structural error $\epsilon_{1,t}$ and instrument m_t. The top right panel shows the updating of σ_ν. The prior is chosen to have mean 0.02 and infinite variance. The posterior suggests a standard deviation of this noise measurement centred around 0.2. The bottom panel shows what this implies for the signal-to-noise ratio. While the prior is centred around zero and flat over the whole parameter space, the posterior is centred around 0.15 and has little probability mass near zero. This strongly suggests that the instrument contains relevant information about the structural shock.

3.4   Updating of b

Given that both the recursive identification scheme and the proxy identification scheme impose structure on the impact effect of shocks, differences between these two approaches will be most apparent in the identification of b. It contains the impact effects of the uncertainty shocks on the latent and observable factors. The prior distribution is not available in closed form but is implicit in the prior distributions of Σ, $Q_{\cdot,1}$, 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 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
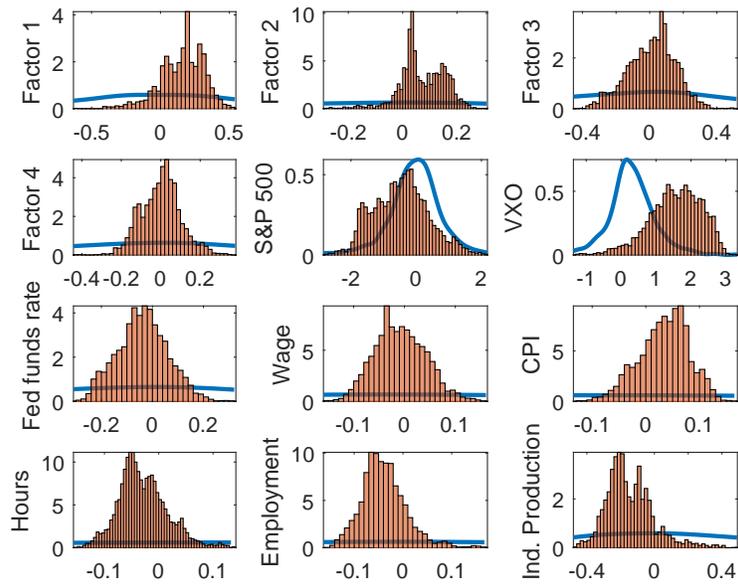
- Compute $b^{prior} = chol(\Sigma)Q_{\cdot,1}$.

Figure 6: **Update of b:** *Updating of the first column of* **B** *(not normalised). The blue line shows the prior distribution of* **b** *computed as the distribution implicit in the priors on* $\boldsymbol{\Sigma}$ *and* $\boldsymbol{Q}_{\cdot,1}$ *and plotted using a Kernel smoother. The bars show the posterior.*

As pointed out by Baumeister and Hamilton (2015), a uniform prior on $\boldsymbol{Q}_{.,1}$ does not necessarily imply a uniform distribution over the structural parameters of interest, which in this case are the elements of $\boldsymbol{b}$. Figure 6 shows that, indeed, the implicit prior on $\boldsymbol{b}$ has some curvature on the impact effect of the S&P 500 index and the VXO. However, it has good coverage of the relevant parameter space. More importantly, it is flat in the relevant parameter regions for the variables of prime interest, namely CPI, hours, employment, and industrial production as well as the latent factors. Therefore, the implicit prior on $\boldsymbol{b}$ should not overly affect the posterior of $\boldsymbol{b}$.

3.5 Impulse Responses

The BP-FAVAR extends the workhorse model by Bloom (2009) in two ways: First, instead of imposing exclusion restrictions on $\boldsymbol{b}$, the BP-FAVAR achieves identification via a proxy. Second, the BP-FAVAR addresses informational deficiency of the Bloom (2009) model, which could potentially bias the responses of variables to the uncertainty shock. The workhorse model contains the following variables: $\Delta log(S\&P500)$, VXO, federal funds rate, $\Delta log(wages)$, $\Delta log(CPI)$, hours, $\Delta log(employment)$, and $\Delta log(IP)$. In order to isolate the effects of the identification scheme and the information set, I include three benchmark models to compare to the baseline BP-FAVAR:

BP-VAR. This model can be considered a Bayesian adaptation of the model by Piffer and Podstawski (forthcoming). I use their model specification, i.e.

$$\boldsymbol{y}_t = \boldsymbol{z}_t, \tag{56}$$

and identification is achieved via the gold price proxy. Differences between the BP-FAVAR and the BP-VAR will be driven primarily by informational issues.

Recursively identified VAR. This model is akin to Bloom (2009). For consistency the variable selection is the same as in the previous model, but identification is achieved by imposing a lower-triangular structure on $\boldsymbol{B}$, i.e.

$$\boldsymbol{B} = chol(\boldsymbol{\Sigma}). \tag{57}$$

26
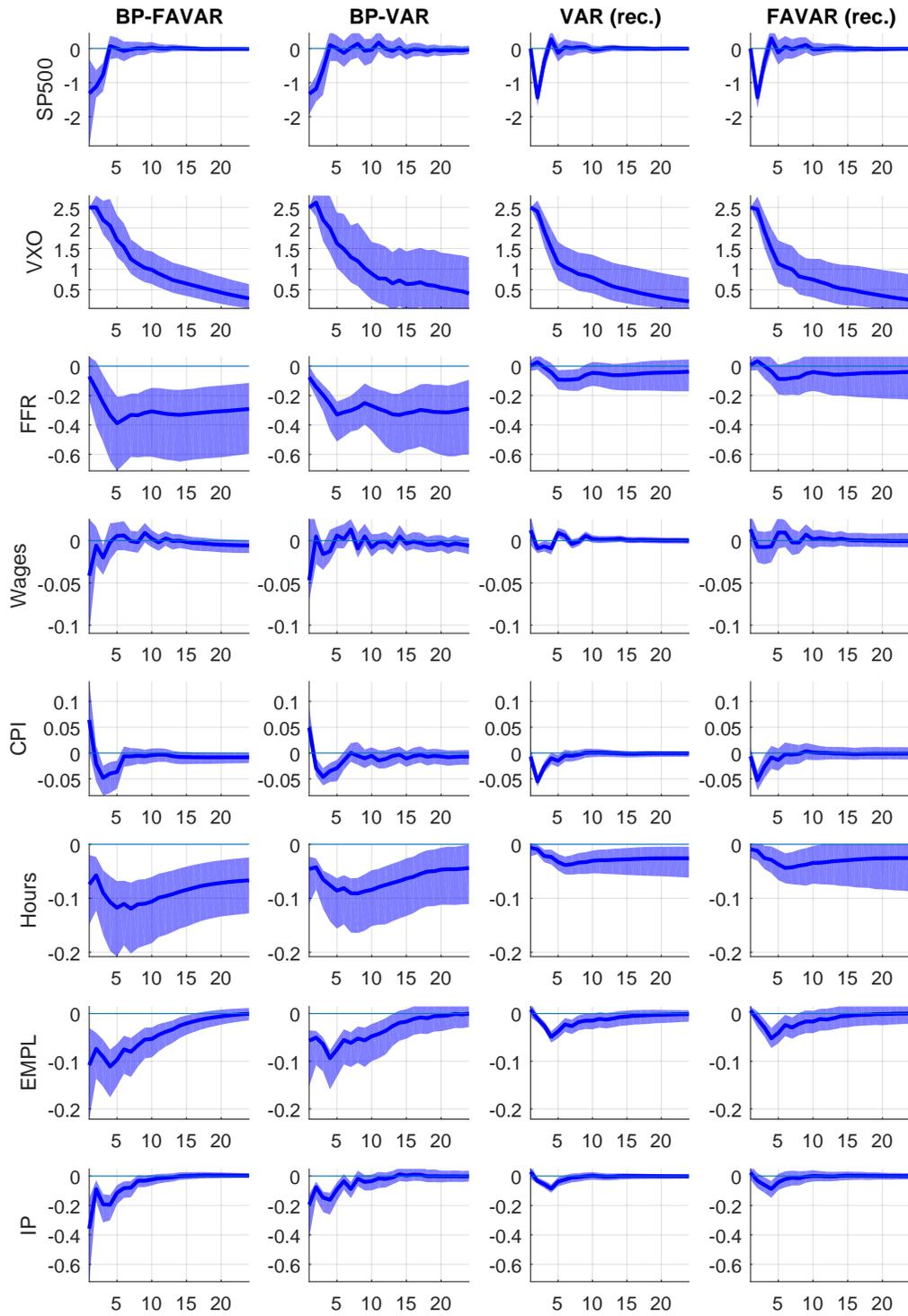
Figure 7: **Impulse Responses:** *The sample is 1979M1 - 2015M7. The first column shows IRFs from the baseline BP-FAVAR. The second column shows IRFs from a small-scale Proxy VAR. The third column shows IRFs from a small-scale recursively identified VAR. The last column shows IRFs from a recursively identified FAVAR. The BP-FAVAR and the recursive FAVAR are based on $R = 4$ latent factors. The bands are computed point-wise 68 % posterior credible bands based on 50000 draws of the posterior sampler discarding the first 40000 as burn-in.*

The uncertainty shock is ordered second, i.e. after the S&P500 index. This assumes that the stock market does not react within a month to an exogenous increase in uncertainty and that financial shocks are the only shocks, apart from the uncertainty shock itself, which influence the uncertainty measure within the month. Differences between the BP-FAVAR and the recursive VAR can be driven by both informational and identification issues.

Recursively identified FAVAR. This scheme adds R latent factor to the small-scale VAR while keeping $\boldsymbol{B}$ lower-triangular. The uncertainty shock is ordered in position $R+2$, i.e. after the latent factors and the S&P500 index. This assumes that neither the latent factors nor the stock market react within a month to an exogenous increase in uncertainty. Differences between the BP-FAVAR and the recursive FAVAR in the impact effects of shocks will be driven primarily by the identification scheme.

Figure 7 shows the impulse responses obtained from the baseline and the three benchmark models: The response variables are those employed by Bloom (2009), Piffer and Podstawski (forthcoming) and other studies. The shock is normalized to generate an increase of 2.5 in the VXO on impact, which is comparable in magnitude to these studies. Estimation is based on 50000 Gibbs draws, discarding the first 40000 draws as a burn-in sample, as in Belviso and Milani (2006) and Ahmadi and Uhlig (2015).

All four models replicate the main findings of Bloom (2009): A rapid drop and subsequent rebound of employment, production and hours worked. It is also in line with Basu and Bundick (2017) who argue that the co-movement among these variables is a key empirical feature that theoretical models should be able to reproduce.

For the stock market, the two recursively identified models (column 3 and 4) exclude an on-impact effect of the uncertainty shock on the stock market index. This results in a biased reaction of the S&P 500 index: The models identified via a proxy (columns 1 and 2) show that the stock market reacts on impact and quickly rebounds. This is in line with the fast moving nature of financial markets, which price increases in uncertainty within the period.

For the real variables, the recursive scheme suggests a moderate negative reaction

of hours, employment and industrial production of less than -0.05%. In the two models identified via proxies this reaction is estimated to be up to -0.2% for industrial production and roughly -0.1% for hours and employment. Given the similarity in the price reaction across models, the recursive models provide a skewed view of the nominal and real interactions following an uncertainty shock.

The inclusion of factors does not qualitatively alter the results. The BP-FAVAR produces results broadly in line with the BP-VAR. The biases resulting from a recursive identification in the small-scale VAR cannot be alleviated by the inclusion of factors. As pointed out by Sims (2012), informational insufficiency is not an either/or concept but can have quantitatively very different effects on impulse responses, depending on the application. In the present case, statistical tests detect informational insufficiency, but alleviating this issue does not have a severe quantitative impact.

All in all, the results are comforting news for policy makers who employ instruments that affect real and nominal variables in the same direction. For example, a Central Bank could use an expansionary policy when faced with an increase in uncertainty and thus move both inflation and real variables closer to their respective targets.

4 Conclusion

This paper aims at recovering the interrelations between real and nominal variables following an identified uncertainty shock by employing a Bayesian Proxy FAVAR.

The first contribution is the empirical finding that uncertainty shocks are deflationary in the short run and indistinguishable from zero after six months. This finding relates to the work by Born and Pfeifer (2014) and Fernández-Villaverde et al. (2015) who show in a theoretical modelling framework that the inflationary effects of uncertainty shocks are driven by two opposing channels: an aggregate demand and a price-bias channel. The aggregate effect of these two channels is unknown ex ante. My results lend empirical support to the dominance of the aggregate demand over the price-bias channel.

My second contribution is methodological. I combine a recent strand of the

Bayesian VAR literature that uses external instruments for identification (Caldara and Herbst, forthcoming) with the Bayesian factor model literature (Belviso and Milani, 2006, 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. I detect informational insufficiency of the small-scale workhorse model, but find that it has limited quantitative impacts on the impulse responses. This relates to the work by Sims (2012) who also finds that informational deficiency is not an either/or but a quantitative issue.

Future empirical research concerned with the identification of uncertainty shocks in a structural VAR context, especially if it is conducted with few variables, should pay close attention to the information set. Informational insufficiency is detected even in the relatively rich workhorse model by Bloom (2009). While it has only limited quantitative impacts in this case, this might be very different when reducing the information set.

From a methodological point of view, the BP-FAVAR model offers potential for a number of extensions: First, identification of two or more shocks is generally possible in this setup. As pointed out by Kilian and Lütkepohl (2017), avoiding non-credible short-run exclusion restrictions is particularly important in factor models. Therefore, the BP-FAVAR could be extended to identify multiple shocks via external instruments. Second, a combination of proxies with sign restrictions in this context is a natural point of departure given the similarity in model setup between the BP-VAR and VARs identified via sign restrictions. The combination of these two approaches is likely to lead to sharper inference.

References

Ahmadi, P. A. and Uhlig, H. (2015), Sign restrictions in bayesian favars with an application to monetary policy shocks, Technical report, National Bureau of Economic Research.

Angelini, G., Bacchiocchi, E., Caggiano, G. and Fanelli, L. (2018), 'Uncertainty across volatility regimes'.

Bachmann, R. and Bayer, C. (2013), 'wait-and-see business cycles?', *Journal of Monetary Economics* **60**(6), 704–719.

Bachmann, R. and Sims, E. R. (2012), 'Confidence and the transmission of government spending shocks', *Journal of Monetary Economics* **59**(3), 235–249.

Baker, S. R., Bloom, N. and Davis, S. J. (2016), 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics* **131**(4), 1593–1636.

Basu, S. and Bundick, B. (2017), 'Uncertainty shocks in a model of effective demand', *Econometrica* **85**(3), 937–958.

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. (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.

Boivin, J., Giannoni, M. P. and Mihov, I. (2009), 'Sticky prices and monetary policy: Evidence from disaggregated us data', *The American Economic Review* **99**(1), 350–384.

31

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

Braun, R., Brüggemann, R. et al. (2017), Identification of svar models by combining sign restrictions with external instruments, Technical report, Department of Economics, University of Konstanz.

Caggiano, G., Castelnuovo, E. and Nodari, G. (2017), 'Uncertainty and monetary policy in good and bad times'.

Caggiano, G., Castelnuovo, E. and Pellegrino, G. (2017), 'Estimating the real effects of uncertainty shocks at the zero lower bound', *European Economic Review* **100**, 257–272.

Caldara, D. and Herbst, E. (forthcoming), 'Monetary policy, real activity, and credit spreads: Evidence from bayesian proxy svars', *American Economic Journal: Macroeconomics* .

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

Christiano, L. J., Motto, R. and Rostagno, M. (2014), 'Risk shocks', *American Economic Review* **104**(1), 27–65.

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.

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.

Forni, M. and Gambetti, L. (2014), 'Sufficient information in structural vars', *Journal of Monetary Economics* **66**, 124–136.

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

Gelper, S. and Croux, C. (2007), 'Multivariate out-of-sample tests for granger causality', *Computational statistics & data analysis* **51**(7), 3319–3329.

Geman, S. and Geman, D. (1984), 'Stochastic relaxation, gibbs distributions, and the bayesian restoration of images', *IEEE Transactions on pattern analysis and machine intelligence* (6), 721–741.

Geweke, J. (1992), Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, *in* 'BAYESIAN STATISTICS', Citeseer.

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.

Jurado, K., Ludvigson, S. C. and Ng, S. (2015), 'Measuring uncertainty', *The American Economic Review* **105**(3), 1177–1216.

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

Koop, G. (2003), *Bayesian Econometrics*, John Wiley & Sons Ltd,.

Koop, G., Korobilis, D. et al. (2010), 'Bayesian multivariate time series methods for empirical macroeconomics', *Foundations and Trends in Econometrics* **3**(4), 267–358.

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

Ludvigson, S. C., Ma, S. and Ng, S. (2018), Uncertainty and business cycles: exogenous impulse or endogenous response?, Technical report, National Bureau of Economic Research.

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

Piffer, M. and Podstawski, M. (forthcoming), 'Identifying uncertainty shocks using the price of gold', *Economic Journal* .

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.

Sims, E. R. (2012), News, non-invertibility, and structural vars, *in* 'DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments', Emerald Group Publishing Limited, pp. 81–135.

Stock, J. H. and Watson, M. W. (2012), Disentangling the channels of the 2007-2009 recession, Technical report, National Bureau of Economic Research.

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. (forthcoming), 'Bootstrap inference for impulse response functions in factor-augmented vector autoregressions', *Journal of Applied Econometrics* .
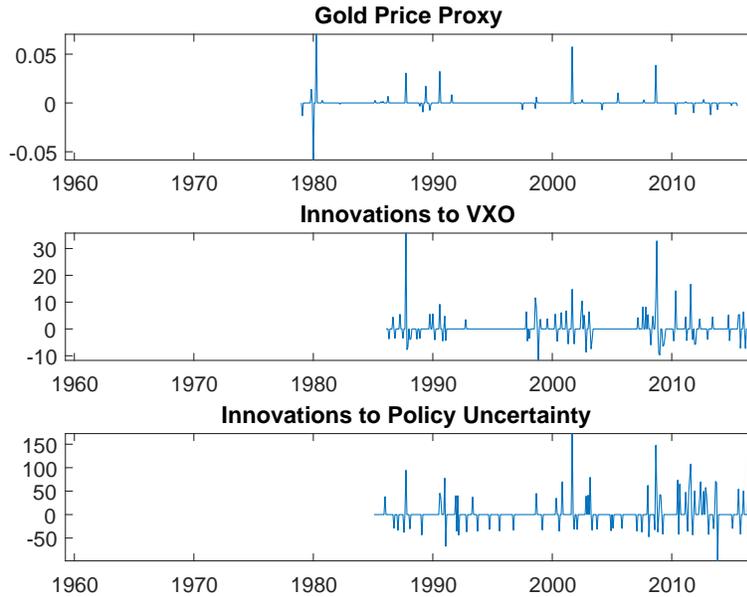
Gold Price Proxy

Innovations to VXO

Innovations to Policy Uncertainty

Figure 8: ***Uncertainty Proxies*** *The top panel shows the baseline gold price proxy for uncertainty by Piffer and Podstawski (forthcoming). The middle panel shows the Stock and Watson (2012) uncertainty proxy computed as residuals of an AR(2) process of the VXO. The bottom panel shows the Stock and Watson (2012) uncertainty proxy computed as residuals of an AR(2) process of the Baker et al. (2016) policy uncertainty proxy. All proxies are winsorized at the 10 % level.*

A Robustness Checks

A.1 Stock and Watson (2012) Instruments

Stock and Watson (2012) were the first to propose two proxies for uncertainty shocks. They employ the innovations in the VXO and in the policy uncertainty index by Baker et al. (2016). These innovations are computed as residuals from an AR(2) process. While innovations to the VXO are likely to be correlated with the uncertainty shocks of interest in this study, this is less likely for the policy uncertainty instrument. The reason is that the concept of uncertainty it is designed to approximate differs from the one I have in mind in this study. Thus, in analysing impulse responses, I concentrate on innovations to the VXO.

I reproduce the instrument as follows: In a first step I estimate an AR(2) process
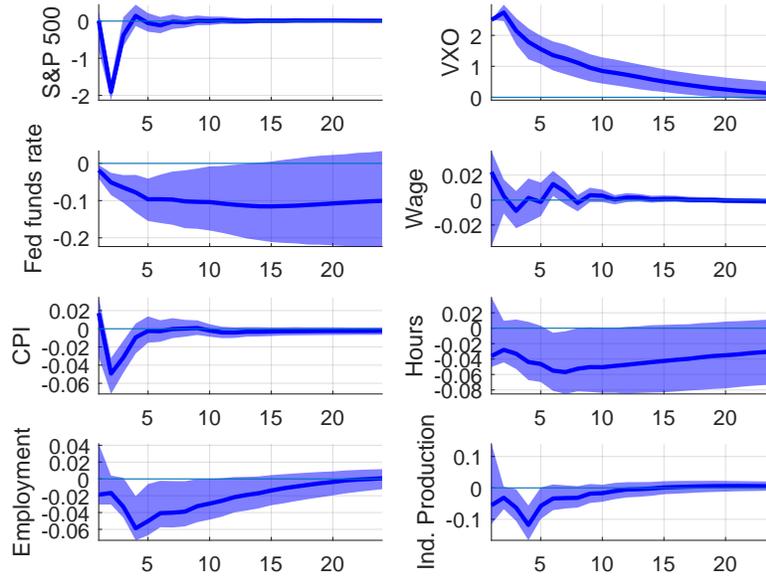
Figure 9: **Stock and Watson (2012) innovations to the VXO** *The model is specified as in the baseline with the gold price proxy replaced by the Stock and Watson (2012) VXO innovations. These are computed as the residuals from an AR(2) process of the VXO.*

including a constant for the VXO index as in Stock and Watson (2012). In a second step, I truncate the proxy at the 10th and 90th percentile and set all observations within these two quantiles to zero. This is to ensure that the proxy only captures quantitatively relevant spikes in uncertainty.

Figure 8 plots the resulting proxies while Table 2 shows the correlations among the proxies. The instruments have some degree of correlation, but they clearly measure distinct variations in uncertainty.

Table 2: Correlation among proxies

Gold proxy	1.00		
VXO innovations	0.38	1.00	
Policy Uncertainty innovations	0.50	0.37	1.00

Figure 9 shows the impulse response functions when employing the innovations

to the VXO as proxies. While the identification of the impact effects is not as sharp as in the baseline, the dynamics are broadly similar.

A.2 Increasing the Lag Length

In order to assess robustness with respect to the lag length, I set $P = 9$. Figure 10 shows that the main results remain unchanged.

A.3 Alternative Uncertainty Measures

Jurado et al. (2015) argue that indices like the VXO used in the baseline do not accurately capture the type of macroeconomic uncertainty that the researcher is interested in. They propose alternative indeces based on large panels of time series. Jurado et al. (2015) provide indeces extracted from macro series and Ludvigson et al. (2018) provide indeces extracted from financial and real series.[9] Figure 11 shows that they are highly correlated over the whole sample period.

Figure 12 replaces the VXO by the macro uncertainty index, Figure 13 replaces the VXO by the financial uncertainty index and Figure 14 replaces it by the real uncertainty index. The results remain qualitatively unchanged.

B Conditional likelihood of m_t

This section re-parametrises Caldara and Herbst (forthcoming) 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 $\boldsymbol{y}_t$.

---

[9]Details can be found here https://www.sydneyludvigson.com/data-and-appendixes/
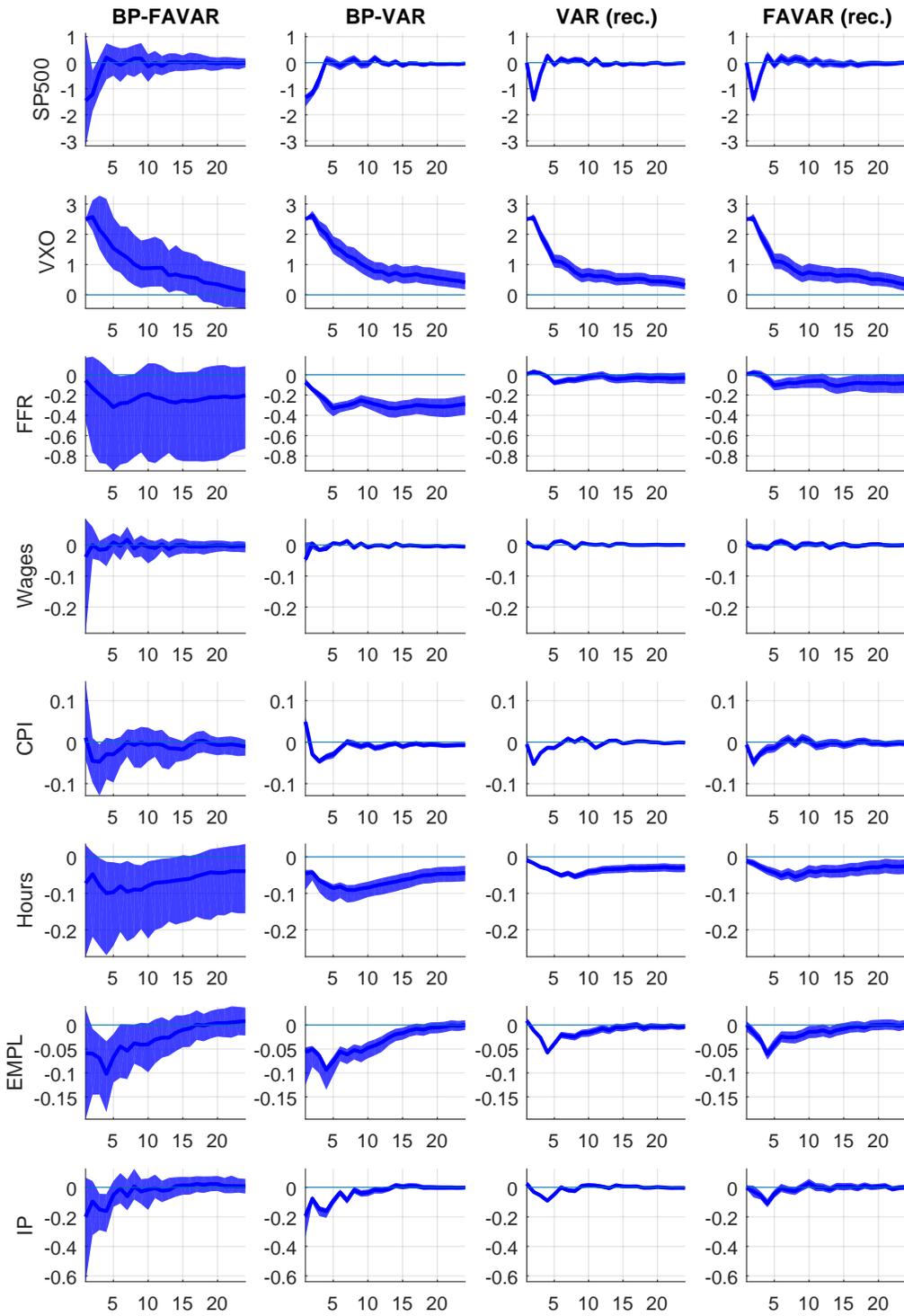
Figure 10: **Impulse Responses (P=9):** *The sample is 1979M1 - 2015M7. The first column shows IRFs from the baseline BP-FAVAR. The second column shows IRFs from a small-scale Proxy VAR. The third column shows IRFs from a small-scale recursively identified VAR. The last column shows IRFs from a recursively identifed FAVAR. The bands are computed point-wise 68 % posterior credible bands based on 10000 draws of the posterior sampler.*
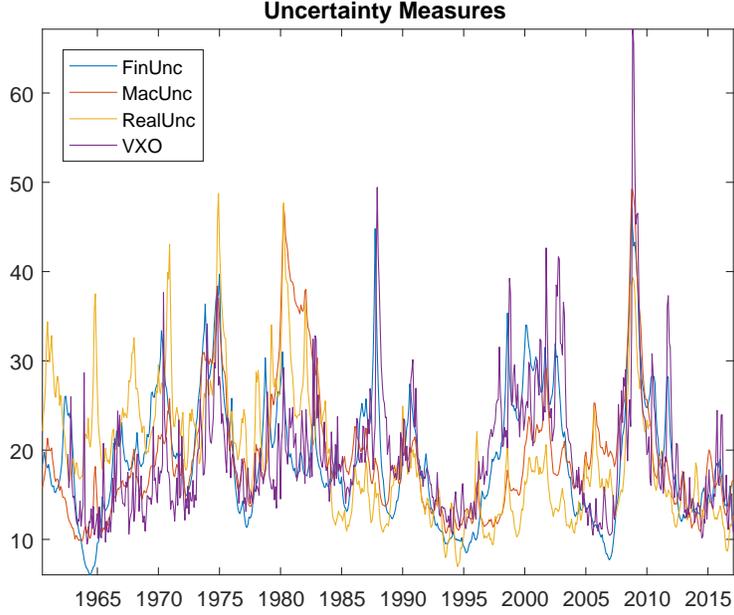
38

Figure 11: ***Uncertainty Measures*** *The uncertainty measures are taken from Ludvigson et al.*
(2018) and rescaled to match the mean of the VXO for comparability.

Restate the model for convenience:

$$\boldsymbol{x}_t = \boldsymbol{\Lambda}\boldsymbol{y}_t + \boldsymbol{\xi}_t \tag{58}$$

$$\begin{bmatrix} \boldsymbol{y}_t \\ m_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Pi} & 0 \\ \boldsymbol{0} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_t \\ m_{t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{B} & \boldsymbol{\beta} \\ \boldsymbol{\beta}' & \sigma_\nu \end{bmatrix} \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \nu_t \end{bmatrix} \tag{59}$$

$$Var\left(\begin{bmatrix} \boldsymbol{y}_t - \boldsymbol{\Pi}\boldsymbol{w}_t \\ m_t \end{bmatrix} \right) = \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{B}\boldsymbol{\beta}' \\ \boldsymbol{\beta}\boldsymbol{B}' & \boldsymbol{\beta}'\boldsymbol{\beta} + \sigma_\nu^2 \end{bmatrix}, \tag{60}$$

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

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

$$Var\left(\begin{bmatrix} \boldsymbol{y}_t - \boldsymbol{\Pi}\boldsymbol{w}_t \\ m_t \end{bmatrix} \right) = \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{B}^c\boldsymbol{Q}\boldsymbol{\beta}' \\ \boldsymbol{\beta}\boldsymbol{Q}'\boldsymbol{B}^{c'} & \boldsymbol{\beta}'\boldsymbol{\beta} + \sigma_\nu^2 \end{bmatrix} \tag{61}$$

Figure 12: **Macroeconomic Uncertainty** *The model is specified as in the baseline with VXO replaced by the macroeconomic uncertainty indicator by* Ludvigson et al. (2018).

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

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

$$\mu_{m|Y} = \boldsymbol{\beta} \boldsymbol{Q}' \mathbf{B}^{c'} \boldsymbol{\Sigma}^{-1} \boldsymbol{u}_t \tag{63}$$

$$= \beta \boldsymbol{\epsilon}_{1,t} \tag{64}$$

$$V_{m|Y} = \boldsymbol{b}\boldsymbol{b}' + \sigma_\nu^2 - \boldsymbol{b}\boldsymbol{Q}'\mathbf{B}^{c'}\boldsymbol{\Sigma}^{-1}\boldsymbol{B}^c\boldsymbol{Q}\boldsymbol{b}' \tag{65}$$

$$= \sigma_\nu^2 \tag{66}$$

Note that, once we condition on $\boldsymbol{y}_t$, the likelihood of m_t does not depend on $\boldsymbol{x}_t$. In addition, note that the conditional likelihood of m_t does not depend on the full matrix $\boldsymbol{B}$, but only on its first column, $\boldsymbol{b}$ because the model is partially identified.

Figure 13: **Financial Uncertainty** *The model is specified as in the baseline with VXO replaced by the financial uncertainty indicator by* Ludvigson et al. (2018).

Therefore, we can rewrite (62) as

$$m_t | \boldsymbol{y}_t, \boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{b}, \beta, \sigma_\nu \sim N(\mu_{m|Y}, V_{m|Y}), \tag{67}$$

C Metropolis-within-Gibbs sampler for the BP-FAVAR

This section outlines the posterior sampling procedure for the Bayesian Proxy Factor Augmented VAR. It combines posterior samplers for Bayesian FAVARs, e.g. Koop et al. (2010) and Ahmadi and Uhlig (2015) with the algorithm for Bayesian Proxy VARs proposed by Caldara and Herbst (forthcoming). It provides posterior draws of the parameters $[\boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\Lambda}, \beta, \sigma_\nu]$ as well as of the factors $\boldsymbol{y}_t$.
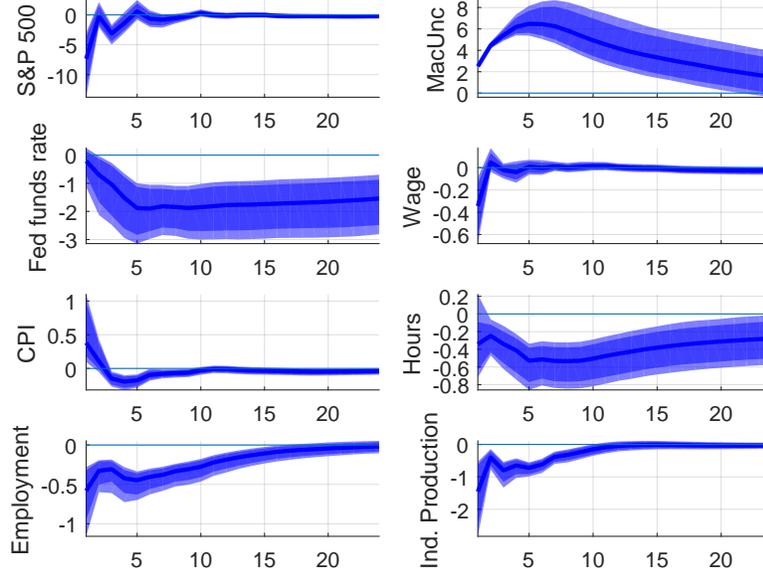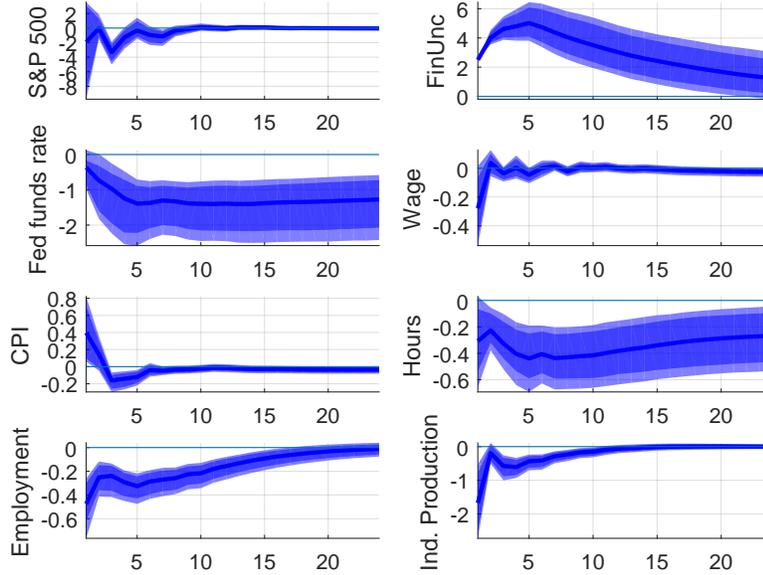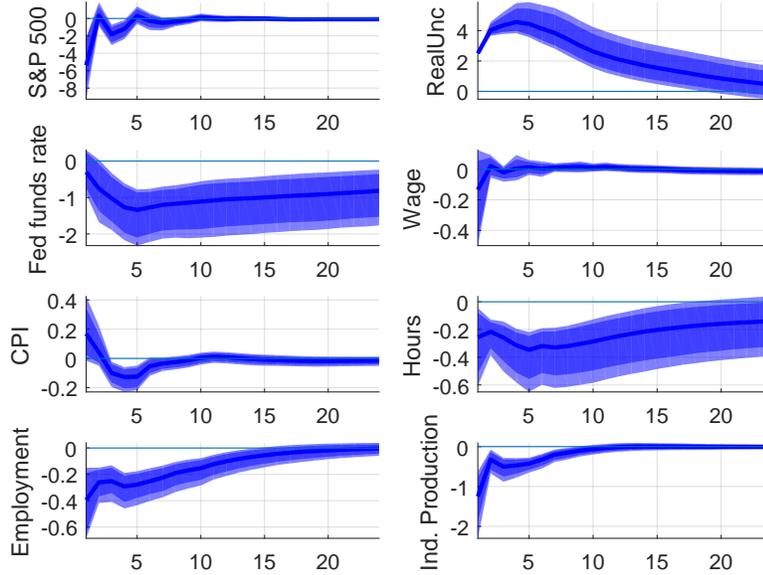
Figure 14: **Real Uncertainty** *The model is specified as in the baseline with VXO replaced by the real uncertainty indicator by* [Ludvigson et al. (2018)](#).

Start by rewriting the posterior distribution as

$$
\begin{aligned}
p(\boldsymbol{\theta}, \boldsymbol{Y} | \boldsymbol{X}, \boldsymbol{m}) &\propto p(\boldsymbol{X}, \boldsymbol{m} | \boldsymbol{\theta}, \boldsymbol{Y}) p(\boldsymbol{\theta}, \boldsymbol{Y}) \\
&= p(\boldsymbol{m} | \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{X}) p(\boldsymbol{X} | \boldsymbol{\theta}, \boldsymbol{Y}) p(\boldsymbol{\theta}, \boldsymbol{Y}) \\
&= p(\boldsymbol{m} | \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{X}) p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{Y}) \\
&= p(\boldsymbol{m} | \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{X}) p(\boldsymbol{\theta}, \boldsymbol{Y} | \boldsymbol{X}) p(\boldsymbol{X}) \\
&\propto p(\boldsymbol{m} | \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{X}) p(\boldsymbol{\theta}, \boldsymbol{Y} | \boldsymbol{X}) \\
&\propto p(\boldsymbol{m} | \boldsymbol{\theta}, \boldsymbol{Y}) p(\boldsymbol{\theta}, \boldsymbol{Y} | \boldsymbol{X})
\end{aligned}
$$

Note that the second term in the last line is the posterior distribution of all model parameters and factors given the data. Standard results (see Ahmadi and Uhlig, 2015) exist to generate draws from this posterior. The first term is the conditional likelihood of $\boldsymbol{m} | \boldsymbol{Y}$. Note that the last transformation is justified by the fact that once conditioned on $\boldsymbol{Y}$, $\boldsymbol{X}$ does not contain further information about $\boldsymbol{m}$. The posterior

sampler weights draws from $p(\boldsymbol{\theta}, \boldsymbol{Y}|\boldsymbol{X})$ with the conditional likelihood of $\boldsymbol{m}|\boldsymbol{Y}$ using an independence Metropolis-Hastings step as in Caldara and Herbst (forthcoming). It is in this sense that $\boldsymbol{m}$ informs the estimation of reduced form parameters.

1. Set starting values

   In order to obtain starting values of the reduced form parameters $[\boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\Lambda}]$, I estimate the model once using the two-step-procedure proposed by Boivin et al. (2009), which takes the restrictions implied by the observation equation into account when extracting factors. I use Principal Component Analysis to obtain R factors $\boldsymbol{f}_t^{PC}$ from $\boldsymbol{x}_t$ and the factor loadings $\boldsymbol{\Lambda}^{PC}$. Then run the regression

$$\boldsymbol{x}_t = const + \boldsymbol{\Lambda}^{\boldsymbol{f}} \boldsymbol{f}_t^{PC} + \boldsymbol{\Lambda}^{\boldsymbol{z}} \boldsymbol{z}_t + \boldsymbol{v}_t$$

   and construct $\tilde{X} = \hat{\boldsymbol{\Lambda}}^{\boldsymbol{f}} \boldsymbol{f}_t^{PC}$, the fitted values orthogonalized with respect to the observable factors. Then extract R factors from $\tilde{\boldsymbol{X}}$ and repeat the procedure 20 times as in Boivin et al. (2009). Save $[\boldsymbol{\Lambda}^0, \boldsymbol{f}_t^0, \boldsymbol{\Omega}^0]$

   Lastly, estimate a reduced-form VAR in $\boldsymbol{y}_t^0 = \begin{bmatrix} \boldsymbol{f}_t^0 \\ \boldsymbol{z}_t \end{bmatrix}$ to obtain $[\boldsymbol{\Sigma}^0, \boldsymbol{\Pi}^0]$.

   For the remaining parameters, I start the algorithm from $\beta^0 = 0$, $\sigma_\nu^0 = 0.5 std(m_t)$ (Caldara and Herbst, forthcoming call this the "high relevance prior", which imposes that half the variance in the proxy can be attributed to measurement error)

   At each stage j proceed with the following steps:

2. Draw $\boldsymbol{f}_t^j$ using the Carter-Kohn backward recursion of the Kalman filter. Set $\boldsymbol{y}_t^j = \begin{bmatrix} \boldsymbol{f}_t^j \\ \boldsymbol{z}_t \end{bmatrix}$

3. Draw $\boldsymbol{\Lambda}^j$ from its conditional normal posterior given in equation 27.

$$\lambda_i | \boldsymbol{\omega}_{ii}^{j-1}, \boldsymbol{X}, \boldsymbol{Y} \sim N(\bar{\boldsymbol{\mu}}_{\lambda,i}, \boldsymbol{\Omega}_{ii^{j-1}} \bar{\boldsymbol{M}}_i^{-1})$$

43

Impose the normalisation on $\boldsymbol{\Lambda}$.

4. Compute $\boldsymbol{\xi}_t^j = \boldsymbol{x}_t - \boldsymbol{\Lambda}^j \boldsymbol{y}_t^j$ and draw the diagonal elements of $\boldsymbol{\Omega}$ from their posterior inverse Gamma distributions

$$\boldsymbol{\omega}_{ii}|\boldsymbol{X}, \boldsymbol{Y} \sim IG(\bar{sc}_i/2, \bar{sh}_i/2)$$

5. Sample $\boldsymbol{\Sigma}^{cand}$ from an inverse Wishart

$$\boldsymbol{\Sigma}^{cand} \sim IW(\bar{\boldsymbol{S}}, \bar{\tau})$$

6. Sample $\boldsymbol{\Pi}^{cand}$ from a multivariate normal using the Minnesota values as priors for the posterior covariance.

$$vec(\boldsymbol{\Pi}^{cand}) \sim N(\bar{\boldsymbol{\mu}}_{\boldsymbol{\Pi}}, \bar{\boldsymbol{V}}_{\boldsymbol{\Pi}})$$

7. With probability α set $\boldsymbol{\Pi}^j = \boldsymbol{\Pi}^{cand}$ and $\boldsymbol{\Sigma}^j = \boldsymbol{\Sigma}^{cand}$, otherwise set $\boldsymbol{\Pi}^j = \boldsymbol{\Pi}^{j-1}$ and $\boldsymbol{\Sigma}^j = \boldsymbol{\Sigma}^{j-1}$, where

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

8. Draw $\boldsymbol{Q}_{.,1}^{cand}$ as the first column of an orthogonal matrix form a uniform Haar distribution using the algorithm by Rubio-Ramirez et al. (2010). Set $\boldsymbol{Q}_{.,1}^j = \boldsymbol{Q}_{.,1}^{cand}$ with probability α and $\boldsymbol{Q}_{.,1}^j = \boldsymbol{Q}_{.,1}^{j-1}$ else.

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

9. Compute structural errors $\epsilon_t^j = (chol(\boldsymbol{\Sigma}^j)\boldsymbol{Q}_{.,1}^j)^{-1}\boldsymbol{U}^j$. Draw β^j from its posterior

normal distribution

$$\beta^j \sim N(\bar{\mu}_\beta, \bar{\sigma}_\beta)$$

10. Draw σ_ν^j from its posterior inverse Gamma distribution

$$\sigma_\nu^j \sim IG(\bar{sh}, \bar{sc})$$

D Multivariate out-of-sample Granger-causality test

We would like to assess whether a vector $\boldsymbol{f}_t$ Granger-causes a vector $\boldsymbol{z}_t$. The series $\boldsymbol{f}_t$ is said to Granger-cause the series $\boldsymbol{z}_t$ if the past of $\boldsymbol{f}_t$ has additional power for forecasting $\boldsymbol{z}_t$ after controlling for the past of $\boldsymbol{z}_t$. Gelper and Croux (2007) base their test statistic on the comparison of two nested VAR models:

$$\boldsymbol{z}_t = \boldsymbol{\Phi}(L)\boldsymbol{z}_{t-1} + \boldsymbol{v}_t^r \tag{68}$$

$$\boldsymbol{z}_t = \boldsymbol{\Phi}(L)\boldsymbol{z}_{t-1} + \boldsymbol{\Psi}(L)\boldsymbol{f}_{t-1} + \boldsymbol{v}_t^f \tag{69}$$

The restricted model (68) has only past values of $\boldsymbol{z}_t$ as regressors, while the unrestricted model (69) has both the past of $\boldsymbol{z}_t$ and $\boldsymbol{f}_t$ as regressors. $\boldsymbol{v}_t^r$ are the residuals of the restricted, while $\boldsymbol{v}_t^f$ are the residuals of the full model. The test statistic is based on the out-of-sample forecast performance of these two models. Compared to in-sample tests, this approach is less susceptible to overfitting.

The unrestricted model can be written out as

$$\boldsymbol{z}_t = \boldsymbol{\phi_0} + \boldsymbol{\phi_1}\boldsymbol{z}_{t-1} + ... + \boldsymbol{\phi_p}\boldsymbol{z}_{t-p} + \boldsymbol{\psi_1}\boldsymbol{f}_{t-1} + ... + \boldsymbol{\psi_p}\boldsymbol{f}_{t-p} + \boldsymbol{v}_t^f \tag{70}$$

where $\boldsymbol{\phi}_j$, $j = 0, ..., p$ are of dimension $K \times K$ and $\boldsymbol{\psi}_j$ are of dimension $K \times R$. Then the Null hypothesis of no Granger causality can be written as

$$H_0 : \boldsymbol{\psi_1} = \boldsymbol{\psi_2} = \cdots = \boldsymbol{\psi_p} = 0 \tag{71}$$

The out-of-sample test proceeds as follows: First, split the sample in half as $T = T_1 + T_2$, where $T_1 = T_2 = 0.5T$ (assuming T is even) and construct one-step-ahead forecasts $\hat{z}^r_{T_1+1}$ and $\hat{z}^f_{T_1+1}$ based on the restricted and the full model, respectively. Then, expand the estimation sample by one and construct $\hat{z}^r_{T_1+2}$ and $\hat{z}^f_{T_1+2}$. The last forecasts, $\hat{z}^r_T$ and $\hat{z}^f_T$ are based on an estimation sample of size $T - 1$. As a second step, construct the series of one-step-ahead forecast errors $\hat{v}^r_t = \hat{z}^r_t - z_t$ and $\hat{v}^f = \hat{z}^f_t - z_t$ for the two competing models and save them in the vectors $\hat{v}^r$ and $\hat{v}^f$ of size T_2. As a third step, construct the test statistic comparing the forecasting performance of the two models as

$$MSFE = log(\frac{|\hat{v}^{r'}\hat{v}^r|}{|\hat{v}^{f'}\hat{v}^f|}), \tag{72}$$

where MSFE is the mean squared forecast error and $|\cdot|$ stands for the determinant of a matrix. If the full model provides better forecasts, MSFE takes a larger value, indicating Granger-causality.

The asymptotic distribution of the test statistic is unknown. Critical values will therefore be based on a residual bootstrap. It proceeds in the following steps:

1. Estimate the model under the Null, i.e. model (68) and compute the test statistic, denoted here by s_0, as described above

2. Generate $Nb = 1000$ new time series $z^*_1, ..., z^*_t$ according to model (68) using the parameter estimates and resampling the residuals with replacement. For each bootstrap sample, compute the test statistic resulting in $s^*_1, ..., s^*_{Nb}$

3. The percentage of bootstrap test statistics, $s^*_1, ..., s^*_{Nb}$, exceeding s_0 is an approximation of the p-value.

Gelper and Croux (2007) show that the test performs well in a Monte Carlo setting as well as in an application to real data.

E   Carter-Kohn Algorithm

This section lays out the Carter-Kohn algorithm. It is used to sample the factors $\boldsymbol{Y}$ given all model parameters, $\boldsymbol{\theta}$, and the data, $\boldsymbol{X}$, i.e. it generates draws from $p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{X})$.

E.1   State-space form

Start by rewriting observation and transition equation as

$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{z}_t \end{bmatrix} = \mathcal{H}\mathcal{B}_t + \mathcal{W}_t \tag{73}$$

$$\mathcal{B}_t = \mathcal{F}\mathcal{B}_{t-1} + \mathcal{V}_t \tag{74}$$

$$Var(\mathcal{W}_t) = \mathcal{R} \tag{75}$$

$$Var(\mathcal{V}_t) = \mathcal{Q} \tag{76}$$

where

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

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

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

Then consider the following factorisation:

$$p(\mathcal{B}_{1:T}|\mathbf{X}, \boldsymbol{\theta}) = p(\mathcal{B}_T|\boldsymbol{x}_{1:T}, \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\mathcal{B}_t|\mathcal{B}_{t+1}, X, \theta) \tag{77}$$

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

$$\mathcal{B}_T|\boldsymbol{x}_{1:T}, \theta \sim N(\mathcal{B}_{T|T}, \mathcal{P}_{T|T}) \tag{78}$$

$$\mathcal{B}_{t|T}|\mathcal{B}_{t+1|T}, \boldsymbol{x}_{1:T}, \theta \sim N(\mathcal{B}_{t|t, \mathcal{B}_{t+1|T}}, \mathcal{P}_{t|t, \mathcal{B}_{t+1|T}}) \tag{79}$$

with

$$\mathcal{B}_{T|T} = E(\mathcal{B}_T|\boldsymbol{x}_{1:T}, \boldsymbol{\theta}) \tag{80}$$

$$\mathcal{P}_{T|T} = Cov(\mathcal{B}_T|\boldsymbol{x}_{1:T}, \boldsymbol{\theta}) \tag{81}$$

$$\mathcal{B}_{t|t, \mathcal{B}_{t+1|T}} = E(\mathcal{B}_t|\mathcal{B}_{t|t}, \mathcal{B}_{t+1|t}, \boldsymbol{\theta}) \tag{82}$$

$$\mathcal{P}_{t|t, \mathcal{B}_{t+1|T}} = Cov(\mathcal{B}_t|\mathcal{B}_{t|t}, \mathcal{B}_{t+1|t}, \boldsymbol{\theta}) \tag{83}$$

E.2   Kalman-filter

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:

$$\mathcal{B}_{t|t} = \mathcal{B}_{t|t-1} + \kappa_{t|t-1}\eta_{t|t-1} \tag{84}$$

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"

$$\mathcal{P}_{t|t-1} = \mathcal{F}\mathcal{P}_{t-1|t-1}\mathcal{F}' + \mathcal{Q} \tag{85}$$

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:

$$\mathcal{B}_{t|t, \mathcal{B}_{t+1|T}}^* = \mathcal{B}_{t|t} + \mathcal{P}_{t|t}\mathcal{F}^{*'}J_{t+1|t}^{-1}\psi_{t+1|t} \tag{86}$$

$$\mathcal{P}_{t|t, \mathcal{B}_{t+1|T}}^* = \mathcal{P}_{t|t} - \mathcal{P}_{t|t}\mathcal{F}^{*'}J_{t+1|t}^{-1}\mathcal{F}^*\mathcal{P}_{t|t} \tag{87}$$

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 (77) 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, $\boldsymbol{y}_t$.

E.3 Convergence of the Posterior Sampling Algorithm

The convergence properties of the reduced form parameters of a Bayesian FAVAR model are discussed in detail in 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 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

$$CD = \frac{\hat{\mu}_{first} - \hat{\mu}_{last}}{\frac{\hat{\sigma}_{first}}{\sqrt{0.1D}} + \frac{\hat{\sigma}_{last}}{\sqrt{0.4D}}} \tag{88}$$

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.

Table 3 shows that for 10 out of 12 parameters, the Null of convergence cannot be rejected. Note that the first R parameters refer to the on-impact effect on the latent factors, which do not have an economic interpretation.

	p-value
r=1	0.7348
r=2	0.0007
r=3	0.1343
r=4	0.1363
k=1	0.8634
k=2	0.1635
k=3	0.7797
k=4	0.0871
k=5	0.3087
k=6	0.9196
k=7	0.0444
k=8	0.1515

Table 3: **Geweke (1992) test for convergence of the MCMC chain** *The sample split is $T = T_1 + T_2 + T_3$, where $T_1 = 0.1T$, $T_2 = 0.5T$ and $T_3 = 0.4T$. P-values are computed as quantiles of a Chi-squared distribution.*

F   Data Description

Table 4: 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

id	tcode	fred	description
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

id	tcode	fred	description
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

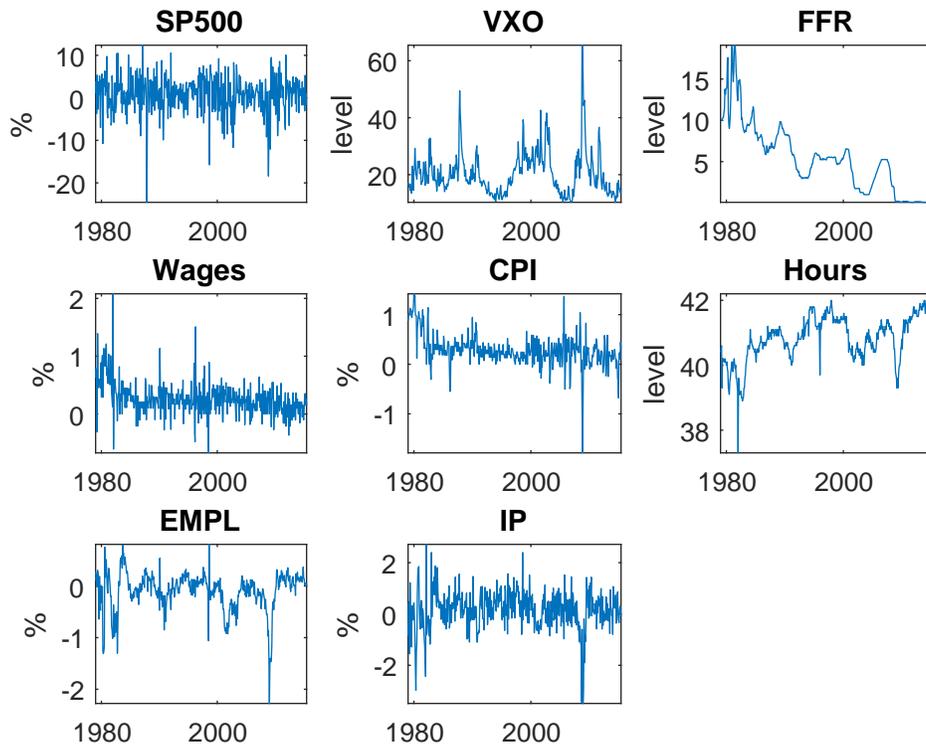| id	tcode	fred	description
80*	5	S&P 500	S&P's Common Stock Price Index: Composite
81*	5	S&P: indust	S&P's Common Stock Price Index: Industrials
82*	2	S&P div yield	S&P's Composite Common Stock: Dividend Yield
83*	5	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio
135*	1	VXOCLSx	VXO

Figure 15: ***Observable Factors*** *The sample length is 1979M1 to 2015M7. The observable factors are the variables included in* Piffer and Podstawski (forthcoming)*.*
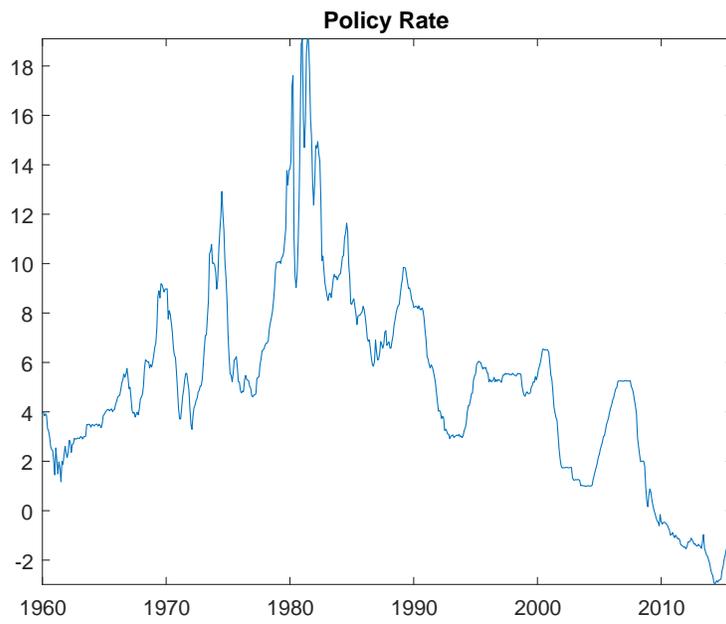
Policy Rate

Figure 16: **Shadow Rate:** Wu and Xia (2016) shadow rate available from 1960M1 to 2015M11.