Using time-varying volatility for identification in Vector Autoregressions: An application to endogenous uncertainty

Published Version:
Carriero A., Clark T.E., Marcellino M. (2021). Using time-varying volatility for identification in Vector Autoregressions: An application to endogenous uncertainty. JOURNAL OF ECONOMETRICS, 225(1), 47-73 [10.1016/j.jeconom.2021.07.001].

Published:
DOI: http://doi.org/10.1016/j.jeconom.2021.07.001

(Article begins on next page)

Using Time-Varying Volatility for Identification in Vector Autoregressions: An Application to Endogenous Uncertainty

Andrea Carriero* Todd E. Clark† Massimiliano Marcellino‡

This draft: July 2021

Abstract

We develop a structural vector autoregression with stochastic volatility in which one of the variables can impact both the mean and the variance of the other variables. We provide conditional posterior distributions for this model, develop an MCMC algorithm for estimation, and show how stochastic volatility can be used to provide useful restrictions for the identification of structural shocks. We then use the model with US data to show that some variables have a significant contemporaneous feedback effect on macroeconomic uncertainty, and overlooking this channel can lead to distortions in the estimated effects of uncertainty on the economy.

J.E.L. Classification: C11, C32, D81, E32

Keywords: Endogeneity, Causality, Stochastic Volatility, Bayesian Methods

*Corresponding author, Queen Mary University of London and University of Bologna, E14NS London UK, a.carriero@qmul.ac.uk.

†Federal Reserve Bank of Cleveland, P.O. Box 6387, Cleveland, OH 44101. todd.clark@researchfed.org

‡Bocconi University, IGIER and CEPR, via Röntgen 1, 20136 Milan. massimiliano.marcellino@unibocconi.it

1 Introduction

Starting from the seminal work of Bloom (2009), the business cycle relationship between uncertainty and output growth and the transmission mechanism from one to the other have received substantial attention in the literature; see Bloom (2014) for an exhaustive survey. Various measures of uncertainty have been put forward, and several efforts have been made to study the macroeconomic effects and broader importance of uncertainty shocks. A non-exhaustive list of studies in this area includes Bachmann et al. (2013), Baker et al. (2016), Basu and Bundick (2017), Bloom (2009), Bloom et al. (2018), Caldara et al. (2016), Caggiano et al. (2014), Carriero et al. (2018, CCM), Cesa-Bianchi, et al. (2020), Jurado et al. (2015, JLN), Rossi and Sekhposyan (2015), and Shin and Zhong (2020).

While the definitions and measurements of uncertainty differ in all these contributions, the common denominator in this line of research is the way in which the effects of uncertainty shocks are identified and assessed. Specifically, most econometric studies typically estimate the effects of uncertainty on economic variables by using structural vector autoregressions with some recursive identification scheme, which all inevitably assume some type of causal direction between uncertainty and economic variables. The assumption typically made is that uncertainty is exogenous; i.e., it does not react contemporaneously to economic variables, while economic variables react contemporaneously to uncertainty.[1]

As is well known, recursive schemes have the advantage of simplicity of implementation and interpretation, but in some cases, they can be hard to defend as a credible identification strategy. This is particularly true when economists have very little a priori, generally accepted, and theoretically grounded reasons to believe that a specific recursive ordering is valid. The study of uncertainty shocks is such a case, since the existing evidence and economic wisdom make us unable to take a stand on the direction of the causality between uncertainty and economic variables such as GDP growth.

The existing literature has shown that both directions of causality are plausible. For example, the case has been made that uncertainty has effects on the economy through firms' behavior. Firms' behavior can be influenced by uncertainty for several reasons, e.g., because of the real option value of waiting before taking investment decisions (e.g., Bernanke 1983, McDonald and Siegel 1986); because of the postponement of hiring and capital investment decisions (e.g., Bloom 2009, Bloom et al. 2018, and Leduc and Liu 2012); and because of the interaction with financial frictions constraining firms' decisions (e.g., Arellano et al. 2019, Gilchrist et al. 2014). From the consumers' side, the effects of uncertainty on the macroeconomy are possible via precautionary savings (e.g., Basu and Bundick 2017

[1]Exogenous as used here and in the rest of the paper is not meant to mean strict exogeneity. Rather, we use it as shorthand for uncertainty being predetermined within the period.

and Fernandez-Villaverde et al. 2011). Equivalently, it is reasonable to conjecture that lower growth, typically associated with higher unemployment, tighter credit conditions, and larger volatility in financial markets, in turn can increase uncertainty. One of the first papers to stress the possible endogeneity of uncertainty is Bachmann et al. 2013). Using an identification strategy in which uncertainty shocks have no long-run effects on aggregate economic activity, they find that the uncertainty shocks then also have no effects in the short run. Instead, various measures of uncertainty substantially increase after a negative shock to aggregate economic activity (see, e.g., Bachmann and Moscarini 2011 and Fajgelbaum et al. 2017).

The concern for bias arising from endogeneity of uncertainty was pointed out by Ludvigson et al. (2019, LMN), who developed an alternative identification strategy based on external information, using event constraints and correlation constraints. Their estimates indicate that macro uncertainty is mostly endogenous; i.e., it mainly reacts to growth conditions rather than being an exogenous source of business cycle fluctuations, whereas financial uncertainty is mostly exogenous.

In this paper, to allow business cycle endogeneity and assess the macroeconomic effects of uncertainty, we develop a structural vector autoregression (VAR) with stochastic volatility and its Bayesian posterior to identify and estimate responses to a shock to a particular endogenous variable of interest, which in our case is uncertainty. As we show, identification of the effects of uncertainty shocks in our model follows from the general results in Lewis (2021) on identification in VARs with time-varying volatility. Our contributions are in the formulation of the specific model, the derivation of the Bayesian posterior for estimation, and the application to the effects of uncertainty shocks. In our implementation, identification is obtained via a heteroskedasticity structure in which the time-varying conditional variances of the variables are driven by an uncertainty measure plus a stochastic idiosyncratic component, or just a stochastic idiosyncratic component. [While we focus on uncertainty shocks, the identification procedure developed in this paper can be applied in any VAR featuring time variation in the volatilities.] Our identification strategy rests on an empirical feature of macroeconomic data — time-varying volatility — that has been well established, starting with the seminal work of Cogley and Sargent (2005) and Primiceri (2005) on VARs with stochastic volatility and Justiniano and Primiceri (2008) on DSGE models with stochastic volatility. More recent studies providing corroborating evidence include, among others, Carriero et al. (2016), Chan and Eisenstat (2018), Clark (2011), Clark and Ravazzolo (2015), D'Agostino et al. (2013), and Diebold et al. (2017).

Methodologically, the model developed in this paper is a structural VAR with time-varying volatility in which one of the variables (the uncertainty measure) can impact both the mean and the variance of the other variables. Differently from most existing approaches in the uncertainty literature based on recursive schemes, our identification strategy permits both a causal channel going from uncertainty to the macroeconomy and the opposite causal channel going from the macroeconomy to uncertainty (which we will refer to as the "feedback channel") to be potentially relevant and quantifiable. We derive and provide conditional posterior distributions for this model, which is a substantial extension of the leverage model of Jacquier et al. (2004), a widely used model in the finance literature. These distributions are nontrivial because our model entails additional layers of complication insofar as it includes VAR dynamics with contemporaneous feedback effects and the stochastic volatility factor enters the conditional mean of the process. The correctness, efficiency, and reliability of the algorithm are established in Monte Carlo experiments with simulated data and with Geweke's (2004) test, prior to use with monthly and quarterly US datasets.

With the model, we show that some variables — specifically consumption, industrial production, and the federal funds rate — have a significant contemporaneous feedback effect on macroeconomic uncertainty. Our empirical application is based on both monthly and quarterly US data. When looking at the dynamic responses, we find that in a model of quarterly data the feedback effect has a significant impact in shaping the response of the variables to an uncertainty shock. Instead, in a model of monthly data we find that this effect is rather small, pointing to a limited effect of the feedback channel. With monthly data, Angelini et al. (2019) reach a conclusion similar to ours, using discrete variance breaks to achieve identification.

To put our approach and results in the broader context of the literature, our identification method belongs to the heteroskedasticity-based identification tradition (see, e.g., Rigobon 2003, Sentana and Fiorentini 2001, and the review in Kilian and Lütkepohl 2017 (chapter 14)). Our methodology differs by modeling the conditional variances via stochastic volatility. Lewis (2020) provides another application of identification via stochastic volatility. The difference relative to earlier applications based on heteroskedasticity is nontrivial, because stochastic volatility allows much more flexibility in the evolution of the conditional variances than regime switching or GARCH specifications, since the time-varying volatilities have their own shocks that are independent from the shocks hitting the level of the variables. Bertsche and Braun (2020) recently developed a related approach to identification in VARs with stochastic volatility (also drawing on the general results of Lewis 2021), with frequentist estimation and inference. They focus on allowing the A matrix of the standard VAR with stochastic volatility to take a form more general than the usual recursive ordering-based format, with the aim of testing the over-identifying restrictions involved in traditional triangular schemes. However, as emphasized by Lanne et al. (2017), with such an approach, additional information is required to be able to attach economic meaning to the structural shocks. We instead condition on the triangularity of the macroeconomic block of variables and focus on modeling the simultaneity between the macro block and the measure of uncertainty.

It is also worth mentioning that stochastic volatility makes the errors of our model non-Gaussian, so that another way to interpret our identification procedure is that it exploits the information in higher-order moments, rather than only in the second moments as in traditional Gaussian VARs. In this sense, our approach also belongs to the literature on identification in non-Gaussian models; see, for example, Lanne et al. (2017).

The paper is structured as follows. Section 2 presents the model and Section 3 discusses identification. Section 4 develops the estimation algorithm, discusses its efficiency and convergence properties, and introduces the prior distributions on the model parameters. Section 5 provides an illustrative application and Monte Carlo experiments based on a small-scale version of the model. Section 6 presents the main empirical results. Section 7 summarizes our main findings and concludes. The paper's appendix contains derivations.[2]

2 A model of endogenous uncertainty

2.1 Model specification

Our interest is in modeling the relationship between a set of economic variables, which we collect in the n-dimensional vector process y_t, and an observable scalar process, which we label m_t. We specify the following model:

$$y_t = \Pi_y(L)y_{t-1} + \Pi_m(L)\ln m_{t-1} + \phi \ln m_t + A^{-1}\Sigma_{y,t}^{0.5}\epsilon_t^* \tag{1}$$

$$\ln m_t = \delta_y(L)y_{t-1} + \delta_m(L)\ln m_{t-1} + \psi\Sigma_{y,t}^{0.5}\epsilon_t^* + \tilde{u}_t, \tag{2}$$

where $\Pi_y(L)$ is an $n \times n$ matrix polynomial, $\Pi_m(L)$ is an $n \times 1$ vector polynomial, ϕ is a $n \times 1$ vector, $\delta_y(L)$ is a $1 \times n$ vector polynomial, $\delta_m(L)$ is a scalar polynomial, and ψ is a $1 \times n$ vector. The $n \times n$ matrix A^{-1}, which describes the contemporaneous relationships among the economic variables, is lower triangular with ones on the main diagonal. $\Sigma_{y,t}$ is an $n \times n$ diagonal matrix of state variables, with σ_{jt}^2 denoting the j-th element on the diagonal. The shocks $\epsilon_t^* \sim iid\ N(0, I_n)$ and $\tilde{u}_t \sim iid\ N(0, \sigma_{\tilde{u}}^2)$ are mutually independent. In the model above we have omitted intercepts for notational simplicity. Intercepts and any other exogenous variables can be added to equations (1) and (2) without any substantial change in the identification and estimation analysis presented below.

[2]Replication files and a supplementary appendix with additional results and robustness checks can be found at `https://didattica.unibocconi.eu/mypage/index.php?IdUte=49257&idr=8345&lingua=eng`.

The model (1)-(2) is a structural VAR for the $(n+1)$-dimensional vector $(y_t' \ \ln m_t)'$, in which the shocks

$$\tilde{\epsilon}_t = \Sigma_{y,t}^{0.5} \epsilon_t^* \tag{3}$$

are serially uncorrelated and conditionally heteroskedastic structural (i.e. mutually uncorrelated) shocks. In our application $\ln m_t$ will be an observable scalar measure of uncertainty.[3] Clearly in a more general context $\ln m_t$ could be any other variable which the researcher is interested in modeling as potentially endogenous, for example a policy variable.

There are two major features that differentiate this model from the VARs typically used in the uncertainty literature (e.g., in studies such as Bloom 2009 and JLN). First, the model allows for bilateral simultaneity between economic variables and uncertainty. Specifically, the model allows for both i) the contemporaneous effects of a shock to uncertainty on the economic variables, as measured by $\partial y_t / \partial \tilde{u}_t = \phi$, and ii) the contemporaneous effect of a shock to economic variables on uncertainty, as measured by $\partial \ln m_t / \partial \tilde{\epsilon}_t = \psi$ (we will refer to this as the "feedback effect"). This bilateral simultaneity is typically not present in the traditional implementations of uncertainty VARs, and is in general not achievable within the class of Gaussian models, since the number of reduced-form coefficients available in such models is insufficient to pin down all of the contemporaneous relations across variables.

The second major feature of the model proposed here is that the disturbance term to the first block of equations (i.e., $A^{-1}\tilde{\epsilon}_t$) is heteroskedastic. The assumption of heteroskedasticity in a VAR of macroeconomic variables has overwhelming support in the recent literature (see, e.g., Chan and Eisenstat 2018 and the other studies cited in the introduction), and many uncertainty measures are constructed on the basis of some variant of a time-varying volatility model (e.g., the measures put forward by JLN and LMN). Lewis (2021) also provides evidence of persistent time variation in the volatility of a wide range of macroeconomic indicators. However, the large majority of uncertainty VARs does not exploit this feature of the data.[4] In this paper we show that — besides providing a better description of the data — the assumption of heteroskedasticity allows us to simultaneously identify the coefficient vectors ϕ and ψ, drawing on Lewis' (2021) general results on identification in VARs with time-varying volatility. Indeed this assumption implies that the VAR is unconditionally not Gaussian, which provides additional identifying information in the form of additional reduced-form moments.

The model in (1)-(2) nests some other models that previously appeared in the literature.

[3]The uncertainty measure could also be treated as unobservable. In this case, an additional step in the MCMC sampler would be needed in order to draw from its conditional posterior distribution. Carriero et al. (2018) consider such an approach, but without allowing for endogenous uncertainty.

[4]Exceptions are Carriero et al. (2018), Creal and Wu (2017), and Shin and Zhong (2020); however, all of these papers do not allow for a contemporaneous feedback.

Setting $n = 1$, $\Sigma_{y,t} = m_t$, $\phi = 0$ and dropping the VAR dynamics as well as the $\Sigma_{y,t}^{0.5}$ in equation (2) provides the model of Jacquier et al. (2004). Setting $\psi = 0$ provides the model of Carriero et al. (2018). Finally, setting $\psi = 0$ and shutting down time variation in volatilities ($\Sigma_{y,t} = \Sigma$) provides the homoskedastic VAR specification of Jurado et al. (2015). All these contributions set either ϕ or ψ to a vector of zeros. As we shall see, this is equivalent to achieving identification by means of a triangular recursive structure in which uncertainty is ordered first ($\psi = 0$) or last ($\phi = 0$) in a VAR.

The model in (1)-(2) is obviously a re-parameterization of a reduced-form model in which both structural shocks $\tilde{\epsilon}_t$ and $\tilde{u}_t$ appear in both the equation for y_t and the equation for $\ln m_t$.[5] This representation can be seen as a semi-structural form of the model, with equation (1) in the structural form (it depends only on the structural shocks $\tilde{\epsilon}_t$) and equation (2) in the reduced form (it depends on a linear combination of $\tilde{\epsilon}_t$ and $\tilde{u}_t$). The semi-structural form is helpful for deriving the model's posterior and building the efficient MCMC algorithm for estimation detailed below.

2.2 Law of motion of the volatilities

The structural VAR in (1)-(2) and the general results on identification in Lewis (2021) admit a wide array of possibilities for modeling the latent state variables in the matrix $\Sigma_{y,t}$. For our approach, a general specification for the unobserved volatility process σ_{jt} for each variable j is given by:

$$\sigma_{jt}^2 = \Pi_{i=1}^k z_{it}^{\beta_{ji}} h_{jt}(\omega_j, \tilde{\eta}_{jt}), \tag{4}$$

where z_{it}, $i = 1, \ldots, k$, are k log-normally distributed observable variables and β_{ji} are loadings measuring the effect of the i-th variable z_{it} on the j-th volatility σ_{jt}^2. The h_{jt} are log-normal unobservable states featuring the Markov property and ω_j are coefficients modeling their evolution. The shocks $\tilde{\eta}_{jt}$ are i.i.d. across time (but they can be mutually contemporaneously correlated) and are assumed to be independent from the shocks to equations (1)-(2). Clearly, equation (4) (for $j = 1, \ldots, n$) represents the transition equation of a state-space system in which the observation equations are given by (1)-(2).

Based on substantial empirical support in Carriero et al. (2016, 2018, 2019) and JLN for some commonality in macroeconomic volatility, we use the specification:

$$\sigma_{jt}^2 = m_t^{\beta_j} h_{jt}, \; j = 1, \ldots, n, \tag{5}$$

where m_t is the measure of uncertainty (and is common to all the volatilities for which

[5] This can be easily seen by inserting the expression for $\ln m_t$ shown in (2) into the right-hand side of (1), which provides a reduced-form representation in which both equations contain both shocks $\tilde{\epsilon}_t$ and $\tilde{u}_t$.

$\beta_j \neq 0$), and h_{jt} are idiosyncratic volatility states with transition equation

$$\ln h_{jt} = \alpha_j + \delta_j \ln h_{jt-1} + \tilde{\eta}_{jt}, \ j = 1, \ldots, n, \tag{6}$$

with $\tilde{\eta}_{jt} \sim iid \ N(0, \sigma^2_{\tilde{\eta}_j})$ and independent from $\tilde{\eta}_{it}$, $i \neq j$, ϵ^*_t, $\tilde{u}_t$. Equation (6) is specified in logarithms, as is common in stochastic volatility models to implicitly implement non-negativity constraints. Note that in this specification σ^2_{jt} is an entirely unobservable state variable, with a log-normal distribution. Taking the logarithms of both sides of (5) shows that the (log) time-varying conditional variance of each variable in y_t is decomposed into a component common to all variables and given by the observable uncertainty measure m_t, plus a variable-specific, unobservable stochastic component h_{jt}:

$$\ln \sigma^2_{jt} = \beta_j \ln m_t + \ln h_{jt}, \ j = 1, \ldots, n. \tag{7}$$

In the expression above, the loading β_j measures the elasticity of the volatility of variable j to the common volatility factor m_t. Moreover, note that specification (5) implies that the measure of uncertainty is allowed to impact not only the conditional mean of the economic variables (which happens through the term $\phi \ln m_t$ in (1)) but also the conditional variance. Even though σ^2_{jt} depends on an observable variable m_t, the term h_{jt} implies that the overall σ^2_{jt} is still an unobservable variable. The next section discusses the features of the volatility process needed for identification of the effects of a shock to uncertainty.

A notable special case of (5) can be obtained by setting $\beta_j = 0$, $j = 1, \ldots, n$: This yields a specification of time-varying volatility analogous to that originally introduced by Cogley and Sargent (2005) and Primiceri (2005), eventually becoming a standard specification for macroeconomic VARs with time variation in volatility.[6] We will use this simpler specification for an illustrative univariate ($n = 1$) example and in the Monte Carlo exercise.

Of course, more general specifications for σ^2_{jt} are possible, as shown in (4). Also, consistent with (4), $\ln h_{jt}$ does not need to be an autoregression, but it could be modeled as a discrete state variable describing a limited number of regimes, as in Markov switching or threshold models. The choice of a good specification is key to obtain efficiency, but, for a range of specifications, the choice does not impact identification. Lewis (2021) shows that heteroskedasticity per se can be sufficient to provide identification under fairly general conditions, regardless of whether the heteroskedasticity is well specified in the model.

[6] The restricted specification with $\beta_j = 0$ is analogous to the models of Cogley and Sargent (2005) and Primiceri (2005) in that this paper's restricted specification for volatility takes an AR(1) form, whereas these earlier studies used random walks. In addition, the Primiceri (2005) formulation is slightly different as it allows for correlation across the shocks $\tilde{\eta}_{it}$. This could be also allowed for in our approach with no consequences on identification. Koop and Potter (2007) show that stochastic volatility can be interpreted as a more continuous version of change-points.

3 Identification

To show that the coefficients of the model are identified we rely on the recent results of Lewis (2021). In Lewis' notation, equations (1)-(2) can be written as:

$$A(L)Y_t = H\varepsilon_t, \tag{8}$$

where $Y_t = (y_t' \; \ln m_t)'$, $\varepsilon_t = (\tilde{\epsilon}_t' \; \tilde{u}_t)'$, and

$$A(L) = \begin{bmatrix} I_n - \Pi_y(L)L - \phi\delta_y(L)L & -\Pi_m(L)L - \phi\delta_m(L)L \\ -\delta_y(L)L & 1 - \delta_m(L)L \end{bmatrix}, \tag{9}$$

$$H = \begin{bmatrix} A^{-1} + \phi\psi & \phi \\ \psi & 1 \end{bmatrix}. \tag{10}$$

The impact matrix H is time invariant and full rank (with determinant $|A^{-1} + \phi\psi - \phi\psi||1| = 1$), which satisfies Assumption (B) of Lewis (2021). The $n+1$-dimensional vector ε_t satisfies Assumptions (A) and (C) of Lewis (2021). In particular, it has conditional mean $E[\varepsilon_t|\sigma_t^2] = 0$ (Assumption A.1) and diagonal conditional variance (Assumption A.2)

$$Var[\varepsilon_t|\sigma_t^2] = \begin{bmatrix} \Sigma_{y,t} & 0 \\ 0 & \sigma_{\tilde{u}}^2 \end{bmatrix} \equiv \Sigma_t = diag(\sigma_t^2),$$

where σ_t^2 is a $n+1$ vector comprised of σ_{jt}^2, $j = 1, \ldots, n$, and the coefficient $\sigma_{\tilde{u}}^2$. Moreover, since h_{jt}, $j = 1, \ldots, n$, are mutually independent log-normal processes featuring finite first and second moments for each $t = 1, \ldots, T$, the moments $E[\sigma_t^2]$ and $E[vec(\varepsilon_t\varepsilon_t')vec(\varepsilon_t\varepsilon_t')']$ exist and are finite for each $t = 1, \ldots, T$ (Assumption A.3 and Assumption C). As noted in Lewis (2021) the latter are merely requirements of existence of higher moments.[7] They ensure that all of the expectations used are well defined for an object at a particular point in time, even if the distribution might be different at another point in time. Stationarity has not been assumed and is not required. In this representation, the elements of the lag polynomial $A(L)$ of the reduced-form representation are functions of the conditional mean parameters of the model as represented in equations (1)-(2), the elements of the impact matrix H are functions of A^{-1}, ϕ, and ψ, and Σ_t contains the elements of the conditional variance.

Define $\eta_t = H\varepsilon_t$ and $\zeta_t = vech(\eta_t\eta_t')$, where $vech(\cdot)$ is the vector half operator (vectorizing the lower triangular half of a matrix). The autocovariance matrix of ζ_t is:

$$Cov[\zeta_t\zeta_s'] = L(H \otimes H)GM(H \otimes H)'L', \tag{11}$$

[7] In order to further clarify this point, Lewis (2021) employs a particular notation in which $E_t[z_t]$ and $Cov_{t,s}[z_t, z_s]$ denote unconditional moments of the generic variable z_t at time t. We do not follow this convention to avoid confusion with the more common notation in which the t subscript denotes a conditional expectation.

where L is a matrix selecting the lower triangular elements of $(H \otimes H)vec(\Sigma_t)$, G is a selection matrix such that $G\sigma_t^2 = vec(\Sigma_t)$, and

$$M = Cov[\sigma_t^2, \sigma_s^2]G' + E[\sigma_t^2 vec(\varepsilon_s \varepsilon_s' - \Sigma_s)']. \tag{12}$$

In his Theorem 1, Lewis (2021) shows that under the two conditions (i) M is at least of rank 2, and (ii) M has no proportional rows, the impact matrix H is uniquely identified up to sign and column permutations. Note that condition (ii) does not rule out linearly dependent rows; it only rules out one row being a scalar multiple of another.

As we discuss below, for these conditions to hold it is sufficient that the volatilities in the vector σ_t^2 have idiosyncratic variation. Moreover, it is preferable for the idiosyncratic variation to be sufficiently persistent to ensure there are no issues of weak identification, which appears to be the case in macroeconomic data. For example, Lewis (2021) provides evidence of persistent time variation in the volatility of a wide range of macroeconomic indicators, and we find the same results in a reduced-form analysis based on our dataset, as documented in Section 6.2 below.

3.1 Illustrative example

To shed light on identification and in particular the role of the parameters of the conditional volatility process, consider the special case of the model (1)-(2), with 1 lag and with $n = 1$:

$$\begin{aligned} y_t &= \Pi_y y_{t-1} + \Pi_m \ln m_{t-1} + \phi \ln m_t + \tilde{\epsilon}_t, & (13) \\ \ln m_t &= \delta_y y_{t-1} + \delta_m \ln m_{t-1} + \psi \tilde{\epsilon}_t + \tilde{u}_t, & (14) \end{aligned}$$

where $\tilde{\epsilon}_t = \sqrt{\sigma_{y,t}^2}\epsilon_t^*$, $\tilde{u}_t = \sqrt{\sigma_{\tilde{u}}^2}u_t^*$ are mutually independent (structural) shocks with $u_t^* \sim iid\ N(0,1)$, $\epsilon_t^* \sim iid\ N(0,1)$. Here, y_t is a scalar economic variable of interest, for example, GDP growth, while m_t denotes an observable uncertainty measure. The conditional mean of y_t depends on contemporaneous (log) uncertainty through the term $\phi \ln m_t$. Uncertainty is endogenous, as it depends on the contemporaneous value of y_t through the term $\psi\tilde{\epsilon}_t$. The conditional volatility process for y is $\sigma_{y,t}^2 = m_t^\beta h_t$, with $\ln h_t = \alpha + \delta \ln h_{t-1} + \tilde{\eta}_t$.

In Lewis' matrix notation (8), this smaller model takes the form:

$$\underbrace{\begin{bmatrix} 1-(\Pi_y+\phi\delta_y)L & -(\Pi_m+\phi\delta_m)L \\ -\delta_y L & 1-\delta_m L \end{bmatrix}}_{A(L)} \begin{bmatrix} y_t \\ \ln m_t \end{bmatrix} = \underbrace{\begin{bmatrix} 1+\phi\psi & \phi \\ \psi & 1 \end{bmatrix}}_{H} \begin{bmatrix} \tilde{\epsilon}_t \\ \tilde{u}_t \end{bmatrix}.$$

As we discussed in the previous subsection on the general model, the impact matrix H and the vector $\varepsilon_t = (\tilde{\epsilon}_t' \ \tilde{u}_t')'$ satisfy Assumptions (A), (B), and (C) of Lewis (2021). We will now verify that the conditions required by Lewis' (2021) Theorem 1 are satisfied in this simple example. The matrix G' appearing in (12) is:

$$G' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Adding the two terms[8] of the M matrix appearing in (12) we have:

$$M = \begin{bmatrix} \begin{matrix} Cov_{t,s}[\sigma^2_{y,t},\sigma^2_{y,s}] \\ +E[\sigma^2_{y,t}\sigma^2_{y,s}(\epsilon^{*2}_s-1)] \end{matrix} & \sqrt{\sigma^2_{\tilde{u}}}E[\sigma^2_{y,t}\sqrt{\sigma^2_{y,s}}\epsilon^*_s u^*_s] & \sqrt{\sigma^2_{\tilde{u}}}E[\sigma^2_{y,t}\sqrt{\sigma^2_{y,s}}\epsilon^*_s u^*_s] & \sigma^2_{\tilde{u}}E[\sigma^2_{y,t}(u^{*2}_s-1)] \\ \sigma^2_{\tilde{u}}E[\sigma^2_{y,s}(\epsilon^{*2}_s-1)] & (\sigma^2_{\tilde{u}})^{3/2}E[\sqrt{\sigma^2_{y,s}}\epsilon^*_s u^*_s] & (\sigma^2_{\tilde{u}})^{3/2}E[\sqrt{\sigma^2_{y,s}}\epsilon^*_s u^*_s] & 0 \end{bmatrix}. \tag{15}$$

Consider the determinant of the submatrix composed of the first and last column, which is $-\sigma^4_{\tilde{u}}E[\sigma^2_{y,s}(\epsilon^{*2}_s-1)]E[\sigma^2_{y,t}(u^{*2}_s-1)]$. Recall from (5) that $\sigma^2_{y,t}$ depends on m_t, which in turn depends on the shocks ϵ^*_s, u^*_s via the simultaneous model (1)-(2). Therefore both $E[\sigma^2_{y,s}(\epsilon^{*2}_s-1)]$ and $E[\sigma^2_{y,t}(u^{*2}_s-1)]$ are in general different from 0. It follows that (15) has rank 2, which satisfies both requirements that (i) M is at least of rank 2 and (ii) M has no proportional rows.

The argument above needs to be slightly changed if one wants to consider a simpler specification for the volatilities, in which $\beta = 0$, so that (5) simplifies to $\sigma^2_{y,t} = h_t$. Since h_t is independent from ϵ^*_s, u^*_s, the moment matrix in (15) simplifies to:

$$M = \begin{bmatrix} Cov[h_t,h_s] & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

which has reduced rank 1. However, in this case one can invoke Lewis' Theorem 2, which states that one can also consider the alternative moment matrix

$$\begin{bmatrix} M & E[\sigma^2_s] \end{bmatrix} = \begin{bmatrix} Cov[h_t,h_s] & 0 & 0 & 0 & E[h^2_s] \\ 0 & 0 & 0 & 0 & \sigma^2_{\tilde{u}} \end{bmatrix}, \tag{16}$$

and use the same conditions required for the matrix M alone. The moment matrix in (16) has rank 2 since the submatrix involving the first and last column has determinant

[8] See Appendix A for a derivation of each of these two terms.

$Cov[h_t, h_s]\sigma^2_{\tilde{u}}$, which is nonzero and finite as long as h_t has some amount of persistence, which is ensured by having $\delta \neq 0$ in the AR process for h_t.

The identifying information comes from the time-variation in the volatility term $\sigma^2_{y,t}$. To clarify this point, suppose that $\sigma^2_{\tilde{\eta}} \longrightarrow 0$. In this case $\sigma^2_{y,t}$ will no longer be a random state variable, but rather will converge to a fixed σ^2_y. When this happens, the matrix M in (15) approaches 0:

$$M = \left[\begin{array}{cccc} 0 + \sigma_y^{2\,2} \underbrace{E[(\epsilon_s^{*2} - 1)]}_{=0} & \sqrt{\sigma^2_{\tilde{u}}}\sigma^2_y \sqrt{\sigma^2_y} \underbrace{E[\epsilon^*_s u^*_s]}_{=0} & \sqrt{\sigma^2_{\tilde{u}}}\sigma^2_y \sqrt{\sigma^2_y} \underbrace{E[\epsilon^*_s u^*_s]}_{=0} & \sigma^2_{\tilde{u}}\sigma^2_y \underbrace{E[(u_s^{*2} - 1)]}_{=0} \\ \sigma^2_{\tilde{u}}\sigma^2_y \underbrace{E[(\epsilon_s^{*2} - 1)]}_{=0} & (\sigma^2_{\tilde{u}})^{3/2} \sqrt{\sigma^2_y} \underbrace{E[\epsilon^*_s u^*_s]}_{=0} & (\sigma^2_{\tilde{u}})^{3/2} \sqrt{\sigma^2_y} \underbrace{E[\epsilon^*_s u^*_s]}_{=0} & 0 \end{array} \right],$$

and therefore the required conditions for identification are no longer satisfied. The same applies in the simpler $\beta = 0$ case.

As this example indicates, based on the results in Lewis (2021), the identification of uncertainty shocks in our model with time-varying volatility rests on the presence of random variation in the volatility (with some persistence when the volatility process does not contain a common factor). As noted above, the assumption of heteroskedasticity in a VAR of macroeconomic variables has strong empirical support (see, e.g., Chan and Eisenstat 2018 and the other studies cited in the introduction), and Carriero et al. (2018) provide evidence of idiosyncratic components in stochastic volatility. Lewis (2021) also provides evidence of persistent time variation in the volatility of a wide range of macroeconomic indicators.

The intuition behind the general identification of the model can be explained in two ways: by looking at either the conditional or the unconditional moments of the shocks. Starting with the intuition based on conditional moments, when $\sigma^2_{\tilde{\eta}} > 0$, more moments become available from the reduced form, because time variation in h_t means that we have more reduced-form error-variance matrices (each corresponding to a different point in time). The simplest way to think about this is within the textbook example $u_t = B^{-1}e_t$, where e_t is the structural shock (with an identity variance matrix) and u_t the reduced-form shock, with variance Σ. B^{-1} is a full matrix describing the contemporaneous relationships among the variables. Since Σ has only $n(n+1)/2$ free parameters, while B has n^2 free parameters, there is incomplete identification. Now, assume we have two regimes: one with $\Sigma = \Sigma(h_1)$ and the second with $\Sigma = \Sigma(h_2)$. This doubles the number of parameters from the reduced form, which becomes $n(n + 1)$, and therefore, the order condition for identification is satisfied.[9] This approach to identification makes the implicit assumption that the contemporaneous relationships among the variables (those described by the matrix B^{-1}) are constant over

[9] This line of argument was first put forward by Rigobon (2003), even though it is important to stress that in our approach $\sigma^2_{y,t}$ is a state variable and not a vector of parameters.

time.

The intuition based on unconditional moments is related to Gaussianity. A Gaussian random variable is entirely defined by its first two moments (which are sufficient statistics). In the case of shocks, the first moment is 0 so we are left with the second moments (variances) only. The problem of identification arises precisely because the variance-covariance matrix of the Gaussian shocks has only $n(n+1)/2$ free coefficients, which often are not enough to identify all the contemporaneous relations we would like to (which are typically n^2). However, in the case of a non-Gaussian random variable, higher-order moments can provide additional information for identification. In the case at hand, if $\sigma^2_{\tilde{\eta}} = 0$, then the shocks are Gaussian, and therefore, we have the identification problem. Instead, if $\sigma^2_{\tilde{\eta}} > 0$, the shocks are not Gaussian (they are a mixture of Gaussians with mixture weights $\sqrt{\sigma^2_{y,t}}$), and therefore, we have identification.

3.2 Identification conditions in the general case

In the general case with $n > 1$, the moment matrix M has dimension $(n+1) \times (n+1)^2$. The identification conditions are that (i) M is at least of rank 2 and (ii) M has no proportional rows. It is sufficient to find a $(n+1) \times (n+1)$ submatrix of M which satisfies these conditions. Consider the following submatrix of M collecting the columns in positions $1,\ (n+1)+2,\ 2(n+1)+3,\ 3(n+1)+4,\ \ldots,\ n(n+1)+(n+1)$:

$$\left[\begin{array}{cccc} \left[\begin{array}{c} cov(\sigma^2_{1t}, \sigma^2_{1s}) + \\ E[\sigma^2_{1t}\sigma^2_{1s}(\epsilon^{*2}_{1s} - 1)] \\ cov(\sigma^2_{2t}, \sigma^2_{1s}) + \\ E[\sigma^2_{2t}\sigma^2_{1s}(\epsilon^{*2}_{1s} - 1)] \\ \vdots \\ cov(\sigma^2_{nt}\sigma^2_{1s}) + \\ E[\sigma^2_{nt}\sigma^2_{1s}(\epsilon^{*2}_{1s} - 1)] \\ \sigma^2_{\tilde{u}}E[\sigma^2_{1s}(\epsilon^{*2}_{1s} - 1)] \end{array}\right] & \cdots & \left[\begin{array}{c} cov(\sigma^2_{1t}\sigma^2_{ns}) + \\ E[\sigma^2_{1t}\sigma^2_{ns}(\epsilon^{*2}_{ns} - 1)] \\ cov(\sigma^2_{2t}\sigma^2_{ns}) + \\ E[\sigma^2_{2t}\sigma^2_{ns}(\epsilon^{*2}_{ns} - 1)] \\ \vdots \\ cov(\sigma^2_{nt}, \sigma^2_{ns}) + \\ E[\sigma^2_{nt}\sigma^2_{ns}(\epsilon^{*2}_{ns} - 1)] \\ \sigma^2_{\tilde{u}}E[\sigma^2_{ns}(\epsilon^{*2}_{ns} - 1)] \end{array}\right] & \left[\begin{array}{c} \sigma^2_{\tilde{u}}E[\sigma^2_{1t}(u^{*2}_s - 1)] \\ \sigma^2_{\tilde{u}}E[\sigma^2_{2t}(u^{*2}_s - 1)] \\ \vdots \\ \sigma^2_{\tilde{u}}E[\sigma^2_{nt}(u^{*2}_s - 1)] \\ 0 \end{array}\right] \end{array}\right].$$

Writing this matrix compactly we have:

$$\left[\begin{array}{cc} cov(\sigma^2_t, \sigma^2_s) + E[\sigma^2_t\sigma^2_s(\epsilon^{*2}_s - \mathbf{1})] & \sigma^2_{\tilde{u}}E[\sigma^2_t(u^{*2}_s - 1)] \\ \sigma^2_{\tilde{u}}E[\sigma^2_s \odot (\epsilon^{*2}_s - \mathbf{1})]' & 0 \end{array}\right], \tag{17}$$

where $\odot$ denotes element-by-element multiplication and $\mathbf{1}$ denotes a column vector of ones. Since $\sigma^2_{jt} = m_t^{\beta_j} h_{jt}$ and m_t depends on ϵ^{*2}_t and u^{*2}_t, the expectations $E[\sigma^2_s \odot (\epsilon^{*2}_s - \mathbf{1})]'$ and $E[\sigma^2_t(u^{*2}_s - 1)]$ appearing in the last row and column of (17) are both non-zero. This is sufficient to ensure that the entire matrix in (17) has rank at least 2. Indeed, consider the sub-matrix formed by the last 2 elements of the last 2 rows and columns:

$$\begin{bmatrix} cov(\sigma^2_{nt}, \sigma^2_{ns}) + E[\sigma^2_{nt}\sigma^2_{ns}(\epsilon^{*2}_{ns} - 1)] & \sigma^2_{\tilde{u}} E[\sigma^2_{nt}(u^{*2}_s - 1)] \\ \sigma^2_{\tilde{u}} E[\sigma^2_{ns}(\epsilon^{*2}_{ns} - 1)] & 0 \end{bmatrix}.$$

This matrix has full rank, regardless of the value of the north-west element. The same applies to any other 2×2 sub-matrix formed using the last row and last column of (17), and it is a consequence of the fact that $\tilde{u}_t$ is homoskedastic (hence its volatility has zero autocovariance), while σ^2_t is correlated with ϵ^{*2}_t and u^{*2}_t. The second condition for identification is that there are no two proportional rows (columns) in (17). This condition is satisfied because the states $h_{1t}, \ldots, h_{nt}$ are mutually independent, thereby introducing idiosyncratic variation in the volatility states σ^2_t. Note that persistence in the idiosyncratic states $h_{1t}, \ldots, h_{nt}$ is not necessary, but it is desirable to avoid weak identification issues.

In the case with no factor structure ($\beta = 0$), we have that $\sigma^2_{jt} = h_{jt}$ and is uncorrelated with ϵ^{*2}_t and u^{*2}_t; hence, the elements $E[\sigma^2_t \sigma^2_s(\epsilon^{*2}_s - \mathbf{1})]$, $E[\sigma^2_s \odot (\epsilon^{*2}_s - \mathbf{1})]'$, and $E[\sigma^2_t(u^{*2}_s - 1)]$ are all zero and the matrix in (17) simplifies to

$$\begin{bmatrix} cov(h^2_t, h^2_s) & 0 \\ 0 & 0 \end{bmatrix},$$

where h_t is the vector of idiosyncratic volatility states $h_{1t}, \ldots, h_{nt}$. Since these states are mutually independent, the matrix $cov(h^2_t, h^2_s)$ is diagonal and has determinant $\Pi^N_{j=1} Cov(h^2_{jt})$. This determinant will be non-zero if there is sufficient persistence, i.e. if $\delta_j \neq 0$, $j = 1, \ldots, n$, in (6), in which case $cov(h^2_t, h^2_s)$ will have full rank n and both identification conditions on the matrix (17) will be satisfied. Note that if just one $\delta_j = 0$, then the matrix (17) contains two rows of zeros (the first being the j-th row and the second being the last row) and therefore the second condition for identification would not be satisfied. Hence in this case the idiosyncratic variation must have some degree of persistence.

3.3 Structure of the impact matrix and identification of macroeconomic shocks

Lewis' result states that H is identified up to a column permutation.[10] This means that the model

$$A(L)Y_t = HP\varepsilon_t = H^*\varepsilon_t,$$

[10] Lewis (2021) also assumes H to have a unit diagonal, whereas the upper left block of the H matrix implied by our model does not, due to the $\phi\psi$ term. However, as Lewis notes, the unit diagonal is just a convenient normalization in his setting. Departing from this normalization with our model and its different normalization does not impact the line of reasoning regarding identification.

where P is a permutation matrix, is observationally equivalent to (8). Therefore, while the theorem ensures statistical identification, economic identification must be achieved by labeling the various shocks of ε_t. In our model this labeling is automatically achieved through the special structure (10) imposed on the impact matrix H. Under this structure, the $n \times n$ north-west block is $\left[A^{-1} + \phi\psi\right]$, which is a uni-triangular matrix A^{-1} plus the products of the elements in the last column ϕ and last row ψ. There are no $H^* = HP$ matrices satisfying such a structure except the one resulting from the trivial permutation in which P is the identity matrix.

The restrictions on A^{-1} could be relaxed while still keeping the system statistically identified, but at the cost of losing economic identification. A lower triangular A^{-1} pins down the macroeconomic shocks ε_t^* and allows us to interpret economically the coefficients in ψ and ϕ. Consider again the impact matrix:

$$H = \begin{bmatrix} \begin{bmatrix} 1+\phi_1\psi_1 & \phi_1\psi_2 & \cdots & \cdots & \phi_1\psi_n \\ a_{21}+\phi_2\psi_1 & 1+\phi_2\psi_2 & & & \vdots \\ a_{31}+\phi_3\psi_1 & a_{32}+\phi_3\psi_2 & \ddots & & \vdots \\ \vdots & \vdots & & 1+\phi_{n-1}\psi_{n-1} & \phi_n\psi_n \\ a_{n1}+\phi_n\psi_1 & a_{n2}+\phi_n\psi_2 & \cdots & a_{nn-1}+\phi_n\psi_{n-1} & 1+\phi_n\psi_n \end{bmatrix} & \begin{bmatrix} \phi_1 \\ \vdots \\ \vdots \\ \phi_{n-1} \\ \phi_n \end{bmatrix} \\ \begin{bmatrix} \psi_1 & \psi_2 & \cdots & \psi_{n-1} & \psi_n \end{bmatrix} & 1 \end{bmatrix}.$$

The structure above tells us a lot about the simultaneity between the macro block and the uncertainty block.

Consider first the case in which $\phi = 0$. In this case, the matrix H becomes a lower uni-triangular matrix. This corresponds to a Cholesky ordering in which the uncertainty measure is ordered last. The shocks ε_t^* are economically identified by a recursive scheme in which the j-th shock impacts contemporaneously only the first j variables. Then consider the case in which $\psi = 0$. In this case, H is not lower uni-triangular, but it still corresponds to a Cholesky ordering in which the shocks ε_t^* are economically identified by a recursive scheme.[11] In this case the uncertainty measure is ordered first, so that it only contemporaneously impacts itself.

When both ϕ and ψ are simultaneously non-zero, uncertainty is ordered neither first nor last; it rather has the simultaneity and feedback loops we want to capture. This simultaneity also means that the shocks to the macro variables ε_t^* have contemporaneous effects on all of

[11] One could see that this is indeed a lower triangular Cholesky scheme with uncertainty ordered first by re-ordering both the equations (i.e., the variables in Y_t) and the columns of H. Specifically, one should move the last equation in the first position, getting $Y_t = (\ln m_t \; y_t')'$, and move the last row and column of H to make them the first row and column.

the other macro variables, through the contemporaneous increase in uncertainty. Still, note that the shocks ε_t^* are economically identified because, after removing the impact effects driven by uncertainty ($\phi\psi$), they follow a recursive Cholesky scheme.

In summary, the requirement for A^{-1} to be lower uni-triangular is desirable to economically identify the macroeconomic shocks ε_t^* and therefore give economic meaning to the coefficients in ψ and ϕ. It has the added advantage that a model with lower triangular A^{-1} has become a workhorse of macro-econometrics since the seminal papers of Cogley and Sargent (2005) and Primiceri (2005).[12] But of course, this also means that different permutations of the variables within the macro block might deliver different shocks ε_t^*, and this can translate into different coefficients ψ and ϕ. In practice, we have experimented with several alternative orderings in both the monthly and the quarterly model, and the effect of having different orderings within the macro block has a rather limited impact on the estimated ψ coefficients. Section 3 in the supplementary appendix provides these results.

Of course, one could use alternative strategies to label the shocks ε_t^*, and avoid imposing a recursive structure in A^{-1}. Our framework and algorithm allow for a non-triangular A^{-1} matrix, and even for a full A^{-1} matrix. In both cases, there would be full statistical identification of the shocks. However if one wanted to have economic identification, some restrictions on A^{-1} would be required in order to be able to label the shocks. In terms of estimation, a full A^{-1} matrix could present more problems simply because it contains more parameters than a triangular one, but the same approach to estimation proposed in Section 4 (i.e., a Random Walk Metropolis step) could be used

3.4 Relation to the literature

The previous discussion clarifies that our method belongs to the family of methods for heteroskedasticity-based identification, considered in papers such as Rigobon (2003) and Lanne and Lütkepohl (2008).[13] Another related strand of research considers identification in non-Gaussian models; see, for example, Lanne et al. (2017).[14] Kilian and Lütkepohl (2017) provide an excellent survey.

There is, however, a key difference between our approach and the approach of these studies: in all of the contributions listed above, volatilities are either deterministic parameters

[12] Bognanni (2018) develops an alternative formulation and estimation algorithm that addresses the ordering dependency which exists with the standard VAR with time-varying parameters and stochastic volatility.

[13] Angelini et al. (2019) extend this approach to identify the effects of uncertainty shocks allowing for endogeneity, finding that — in a monthly dataset — the uncertainty shocks can be treated as exogenous.

[14] The Lanne et al. (2017) approach nests a number of other identification procedures based on conditional heteroskedasticity, including Normandin and Phaneuf (2004), Lanne et al. (2010), and Lütkepohl and Netšunajev (2017).

or are driven by (functions of) the same shocks driving the observable variables. Instead, in our approach, volatilities are unobservable state variables driven by their own shocks. It is precisely this feature that complicates identification, because the unobservable volatility states cannot be recovered even asymptotically.

This problem is particularly evident by comparing our approach with Sentana and Fiorentini (2001), which also considers models with time-varying conditional volatility, but based on a GARCH specification. The GARCH specification is unique in that it allows the reduced-form time-varying covariances to be deterministically recovered from the observations. This does not happen in stochastic volatility models, since in these models the time-varying covariances have their own innovations and therefore are not recoverable from the observable variables. Bertsche and Braun (2020) consider frequentist estimation and inference in VARs with stochastic volatility, and they also must rely on the nonparametric arguments in Lewis (2021) to prove identification.

Finally, our identification strategy differs from the procedure introduced by LMN to identify the effects of uncertainty shocks. The latter achieves identification by imposing "event constraints," in the terminology of LMN, requiring the identified shocks to be coherent with economic reasoning when some extraordinary events happen, and "external variable constraints," which restrict the identified uncertainty shocks based on correlations with stock returns or the price of gold. There is also a conceptual difference in the reported impulse responses insofar as the model we present is point-identified, while LMN's model is set-identified.

4 Model estimation

In this section we describe the Markov Chain Monte Carlo (MCMC) algorithm for the estimation of the model. The model is a structural VAR with time-varying volatility in which one of the regressors (the uncertainty measure) possibly impacts both the mean and the variance of the others. This model nests as a special case the leverage model of Jacquier et al. (2004).

One of the paper's contributions is the derivation of the posterior distribution of the model's parameters and states. The conditional posterior distributions are nontrivial because, with respect to the standard stochastic volatility model, our model entails additional layers of complication insofar as it includes VAR dynamics with contemporaneous feedback effects and the stochastic volatility factor enters the conditional mean of the process. The algorithm proposed here is different from the one in Carriero et al. (2018). In that paper the main complication was the filtering of the unobservable common uncertainty measure, in the absence of simultaneity, while in this algorithm the main challenge is the sampling of the model coefficients, because of the simultaneity.

Section 4.2 discusses the efficiency and convergence of the algorithm. Section 4.3 and Section 4.4 discuss, respectively, the priors used in the empirical application and the computation of impulse response functions.

4.1 MCMC algorithm

We collect the model coefficients in three sets. First, θ_1 groups all the coefficients of the y_t equation, plus the loadings: ϕ, A, $\Pi_y(L)$, $\Pi_m(L)$, β_j, $j = 1, \ldots, n$. Second, θ_2 groups all the coefficients of the uncertainty equation: $\delta_y(L)$, $\delta_m(L)$, ψ, $\sigma^2_{\tilde{u}}$. Third, θ_3 groups all the coefficients of the latent volatility processes: α_j, δ_j, $\sigma^2_{\tilde{\eta}_j}$, $j = 1, \ldots, n$. Both y_t and m_t are observable, while h_t is a vector of state variables. The vector θ contains θ_1, θ_2, and θ_3. We collect the h_{jt}'s in the $n \times 1$ vector process h_t, and we define the following matrices:

$$H_t = \begin{bmatrix} h_{1t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & h_{nt} \end{bmatrix}; \; M_t^{(\beta)} = \begin{bmatrix} m_t^{\beta_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & m_t^{\beta_n} \end{bmatrix},$$

which implies (5) can be written as $\Sigma_{y,t} = M_t^{(\beta)} H_t$, where the superscript β indicates dependence on the vector composed of β_j for $j = 1, \ldots, n$.[15]

Let $\mathbf{h}_{1:T}$ denote the entire time series of the states h_{jt} for $j = 1, \ldots, n$ and $t = 1, \ldots, T$, and $\mathbf{Y}_{1:T}$ denote the entire time series of the data $Y_t = (y_t', m_t)'$. The joint density of data and states is given by:

$$p(\mathbf{h}_{1:T}, \mathbf{Y}_{1:T}|\theta) = \prod_{t=1}^{T} m_t^{-1} \Pi_{j=1}^{n} m_t^{-0.5\beta_j} h_{jt}^{-1.5} p_G(\epsilon_t^*) \times p_G(\tilde{u}_t) \times p_G(\tilde{\eta}_t), \tag{18}$$

where the shocks ϵ_t^*, $\tilde{u}_t$, and $\tilde{\eta}_t$ are those in equations (1), (2), and (6). The joint density in (18) is a non-standard distribution which reflects the novel characteristics of the model proposed here: a multivariate model with stochastic volatility in mean, where the volatility presents a factor structure and simultaneity with respect to the shocks to the dependent variable. The derivation of (18) can be found in the paper's appendix. Combining (18) with a prior $p(\theta)$ yields the posterior density of coefficients and states

$$p(\theta, \mathbf{h}_{1:T}|\mathbf{Y}_{1:T}) \propto p(\mathbf{h}_{1:T}, \mathbf{Y}_{1:T}|\theta) p(\theta), \tag{19}$$

which is not a known distribution but can be simulated via the following Gibbs sampler:

[15] The matrix of the idiosyncratic volatility states H_t should not be confused with the impact matrix H appearing in equation (8), and the matrix of the common volatility states should not be confused with the moment matrix in equation (12).

1. $\mathbf{h}_{1:T}|\theta, \mathbf{Y}_{1:T}$

2. $\theta|\mathbf{h}_{1:T}, \mathbf{Y}_{1:T}$, which is accomplished in three blocks:

   2a $\theta_1|\theta_2, \theta_3, \mathbf{h}_{1:T}, \mathbf{Y}_{1:T}$

   2b $\theta_2|\theta_1, \theta_3, \mathbf{h}_{1:T}, \mathbf{Y}_{1:T}$

   2c $\theta_3|\theta_1, \theta_2, \mathbf{h}_{1:T}, \mathbf{Y}_{1:T}$.

While steps (2b) and (2c) are straightforward, step (2a) is nontrivial due to the simultaneity of our model. We now proceed to analyze these steps in more detail.

4.1.1 Drawing $\mathbf{h}_{1:T}|\theta, \mathbf{Y}_{1:T}$

Under the assumption that the processes h_{jt} are idiosyncratic volatilities with independent shocks $\tilde{\eta}_{it}$, the density $\mathbf{h}_{1:T}|\theta, \mathbf{Y}_{1:T}$ can be decomposed into the product $\Pi_{j=1}^{n} p(\mathbf{h}_{j1:T}|\theta, \mathbf{Y}_{1:T})$ and drawn in blocks $j = 1, \ldots, n$. Since the h_{jt} follow a Markov process, a draw from $p(\mathbf{h}_{j1:T}|\theta, \mathbf{Y}_{1:T})$ can be obtained via a sequence of draws from:

$$p(h_{jt}|h_{jt-1}, h_{jt+1}, \theta, \mathbf{Y}_{1:T}) \propto h_{jt}^{-0.5} \exp\left\{-\frac{e_{jt}^2}{2h_{jt}}\right\} \times h_{jt}^{-1} \exp\left\{\frac{-(\ln h_{jt}^2 - \mu_{jt})^2}{2s_{\tilde{\eta}_j}^2}\right\}, \tag{20}$$

where e_{jt} is the j-th element of the vector $e_t = (M_t^{(\beta)})^{-0.5} A(y_t - \Pi_y(L) y_{t-1} - \Pi_m(L) \ln m_{t-1} - \phi \ln m_t)$, $\mu_{jt} = (\delta_j(\ln h_{jt-1} + \ln h_{jt+1}) + \alpha_j(1-\delta_j))/(\delta_j^2+1)$, and $s_{\tilde{\eta}_j}^2 = \sigma_{\tilde{\eta}_j}^2/(\delta_j^2+1)$. The conditional moments μ_{jt} and $s_{\tilde{\eta}_j}^2$ are slightly different in the first and last period of the sample. Note that e_{jt} is observable under the conditioning set of (20). The derivation of the conditional posterior in (20) is detailed in the paper's appendix. Using an independence chain Metropolis step with the transition equation as proposal, i.e., $q \propto h_{jt}^{-1} \exp\{-(\ln h_{jt}^2 - \mu_{jt})^2/2s_{\tilde{\eta}_j}^2\}$, we can accept/reject with acceptance probability $a = \min(1, \iota)$, where

$$\iota = \sqrt{h_{jt}^{pres}/h_{jt}^{cand}} \exp\left\{\frac{{e_{jt}}^2}{2h_{jt}^{pres}} - \frac{{e_{jt}}^2}{2h_{jt}^{cand}}\right\}.$$

Finally, we initialize the states by drawing a set of initial conditions h_{j0}, $j = 1, \ldots, n$. We assume a log-normal prior $\ln h_{j0} \sim N(\mu_{h_{j0}}, \sigma_{h_{j0}}^2)$, yielding a log-normal posterior $(\ln h_{j0}|\ln h_{jt+1}, \ldots) \sim N(\bar{\mu}_{h_j}, \bar{\sigma}_{h_j}^2)$, where $\bar{\mu}_{h_j} = \bar{\sigma}_{h_j}^2 \left(\frac{\mu_{h_{j0}}}{\sigma_{h_{j0}}^2} + \frac{(\ln h_{j1} - \alpha_j)/\delta_j}{\sigma_{\tilde{\eta}_j}^2/\delta_j^2}\right)$ and $\bar{\sigma}_{h_j}^2 = \frac{\sigma_{h_{j0}}^2 \sigma_{\tilde{\eta}_j}^2/\delta_j^2}{\sigma_{h_{j0}}^2 + \sigma_{\tilde{\eta}_j}^2/\delta_j^2}$.

4.1.2 Drawing $\theta|\mathbf{h}_{1:T},\mathbf{Y}_{1:T}$

Consider again the likelihood and the posterior density given in (18) and (19). These densities depend on ϵ_t^*, $\tilde{u}_t$, and $\tilde{\eta}_{jt}$, where:

$$\begin{aligned}
\epsilon_t^* &= (M_t^{(\beta)})^{-0.5}H_t^{-0.5}A(y_t-\Pi_y(L)y_{t-1}-\Pi_m(L)\ln m_{t-1}-\phi\ln m_t)\\
\tilde{u}_t &= \ln m_t-\delta_y(L)y_{t-1}-\delta_m(L)\ln m_{t-1}\\
&\quad -\psi A(y_t-\Pi_y(L)y_{t-1}-\Pi_m(L)\ln m_{t-1}-\phi\ln m_t)\\
\tilde{\eta}_{jt} &= \ln h_{jt}-\alpha_j-\delta_j\ln h_{jt-1},\quad j=1,\ldots,n.
\end{aligned}$$

Note that the coefficients of the second equation, θ_2, only appear in the equation for $\tilde{u}_t$, and the coefficients of the third equation, θ_3, only appear in the equation for $\tilde{\eta}_{jt}$. However, the coefficients of the first equation θ_1 appear in both equations for ϵ_t^* and $\tilde{u}_t$. Substituting these expressions into the posterior density (19) yields:

$$p(\theta,\mathbf{h}_{1:T}|\mathbf{Y}_{1:T})\propto$$

$$\Pi_{j=1}^n m_t^{-0.5\beta_j}h_{jt}^{-1.5}\exp\left\{-\frac{1}{2}\left(\begin{array}{c}(y_t-\Pi X_t-\phi\ln m_t)'A'(M_t^{(\beta)})^{-1}\\ \times H_t^{-1}A(y_t-\Pi X_t-\phi\ln m_t)\end{array}\right)\right\}p(\theta_1) \qquad (21a)$$

$$\times\frac{1}{\sqrt{\sigma_{\tilde{u}}^2}}\exp\left\{-\frac{1}{2\sigma_{\tilde{u}}^2}\left(\begin{array}{c}\ln m_t-\delta_y(L)y_{t-1}-\delta_m(L)\ln m_{t-1}\\ -\psi A(y_t-\Pi X_t-\phi\ln m_t)\end{array}\right)^2\right\}p(\theta_2) \qquad (21b)$$

$$\times\Pi_{j=1}^n\frac{1}{\sqrt{\sigma_{\tilde{\eta}_i}^2}}\exp\left\{-\frac{(\ln h_{jt}-\alpha_j-\delta_j\ln h_{jt-1})^2}{2\sigma_{\tilde{\eta}_i}^2}\right\}p(\theta_3), \qquad (21c)$$

where we defined

$$\Pi X_t=\Pi_y(L)y_{t-1}+\Pi_m(L)\ln m_{t-1},$$

and where we subsumed the term m_t^{-1} into the proportionality.

As we noted above, the coefficients of the second equation, θ_2, only appear in (21b), and the coefficients of the third equation, θ_3, only appear in (21c). This means that — conditionally on θ_1 (and on the states and data) — (21b) is the posterior kernel for θ_2 and (21c) is the posterior kernel for θ_3.

In particular, since (21b) is a Normal-Inverse Gamma kernel, $\theta_2|y_t,m_t,h_t,\theta_1,\theta_3$ can be drawn via a Gibbs step based on using equation (2) as a linear regression model. Also, note that $p(\theta_2|y_t,m_t,h_t,\theta_1,\theta_3)\propto p(\theta_2|y_t,m_t,h_t,\theta_1)$. Similarly, since (21c) is a (product of) Normal-Inverse Gamma kernel(s), it follows that $\theta_3|y_t,m_t,h_t,\theta_1,\theta_2$ can be drawn via a Gibbs step based on using equation (6) as a linear regression model. Also note that $p(\theta_3|y_t,m_t,h_t,\theta_1,\theta_2)\propto p(\theta_3|h_t)$.

We are now left with the coefficients θ_1. These coefficients are challenging because — when $\psi \neq 0$ — they appear in both (21a) and (21b). The posterior density $p(\theta_1|y_t, m_t, h_t, \theta_1, \theta_2)$ is proportional to the product of the posterior kernels (21a) and (21b):

$$
\begin{aligned}
& p(\theta_1|y_t, m_t, h_t, \theta_2, \theta_3) \propto \\
& \Pi_{j=1}^{n} m_t^{-0.5\beta_j} \exp\left\{ -\frac{1}{2} \left(\begin{array}{c} (y_t - \Pi X_t - \phi \ln m_t)' A' (M_t^{(\beta)})^{-1} \\ \times H_t^{-1} A (y_t - \Pi X_t - \phi \ln m_t) \end{array} \right) \right\} p(\theta_1) \\
& \times \exp\left\{ -\frac{1}{2\sigma_{\tilde{u}}^2} \left(\begin{array}{c} \ln m_t - \delta_y(L) y_{t-1} - \delta_m(L) \ln m_{t-1} \\ -\psi A (y_t - \Pi X_t - \phi \ln m_t) \end{array} \right)^2 \right\},
\end{aligned}
$$

which is not a known distribution. Therefore, this calls for a Random Walk Metropolis step with acceptance probability

$$
a = \min\left(1, \frac{p(\theta_1^{cand}|\mathbf{y}_{1:T}, \mathbf{m}_{1:T}, \mathbf{h}_{1:T}, \theta_2, \theta_3)}{p(\theta_1^{pres}|\mathbf{y}_{1:T}, \mathbf{m}_{1:T}, \mathbf{h}_{1:T}, \theta_2, \theta_3)}\right).
$$

In order to improve the algorithm's mixing, this step is blocked in several sub-steps involving the coefficients in θ_1. Specifically, we draw separately ϕ, A, $\Pi_y(L)$, $\Pi_m(L)$, $\beta_j, j = 1, \ldots, n$.

4.2 Efficiency and convergence of the algorithm

We have verified the correctness, convergence, and mixing properties of the estimation algorithm in our empirical applications. Detailed results are reported in section 1 of the supplementary appendix. Specifically, Figure 1 of the supplementary appendix reports results from Geweke's (2004) test of correctness of the posterior sampler. Figure 2 of the supplementary appendix reports a summary set of diagnostics that all support convergence and good mixing of the MCMC algorithm. The Monte Carlo experiments described in Section 5 further test the algorithm using simulated data.

4.3 Priors

In the empirical application, we demean the variables and omit intercepts, to reduce the dimensionality of the parameter space. The priors on ϕ, A, $\Pi_y(L)$, $\Pi_m(L)$ are flat. For the loadings β_j, $j = 1, \ldots, n$, we elicit a Gaussian prior $\beta_j \sim N(1, 10^2)$, which is very diffuse (but still proper). For the equation for m_t, we elicit an independent Gaussian prior for each coefficient in the polynomials $\delta_y(L)$ and $\delta_m(L)$, with standard deviation 1 and mean 0, with the only exception being the parameter associated with $\ln m_{t-1}$, whose prior mean is set at 0.5. Following Jacquier et al. (2004), we assume a conjugate prior for ψ and $\sigma_{\tilde{u}}^2$, with $\psi|\sigma_{\tilde{u}}^2 \sim N(0, \sigma_{\tilde{u}}^2 I_n)$ and $\sigma_{\tilde{u}}^2 \sim IG\left(\frac{v_0}{2} = 2, \frac{S_0}{2} = 0.05\right)$. For the equations for h_t, we elicit an independent Gaussian–Inverse Gamma prior, with $\delta_j \sim N(0.99, 0.01)$,

$\alpha_j \sim N(0, 0.01)$, and $\sigma^2_{\tilde{\eta}_j} \sim IG\left(\frac{v_0}{2} = 3, \frac{S_0}{2} = 0.05\right)$, $j = 1, \dots, n$. To ensure stationarity of the idiosyncratic volatility process, we also impose prior (and posterior) truncation for the parameter δ_j using the rejection sampling approach of Cogley and Sargent (2005). For the initial conditions of the states we use a Gaussian prior $\ln h_{j0} \sim N(\mu_{h_0}, \sigma^2_{h_0})$ with $\mu_{h_0} = \ln 0.20$ and $\sigma^2_{h_0} = 10$. These priors are the same across all the presented empirical applications. In the Monte Carlo exercise, to get an accurate read on bias in the coefficient estimates, we use a slightly different, less informative prior for some coefficients, namely $\delta_j \sim N(0.99, 0.1)$, $\alpha_j \sim N(0, 0.1)$, and $\sigma^2_{\tilde{u}} \sim IG\left(\frac{v_0}{2} = 2, \frac{S_0}{2} = 0.0025\right)$.

4.4 Impulse responses

We will use the model to compute the impulse response functions (IRFs) to an uncertainty shock of size $\sqrt{\sigma^2_{\tilde{u}}}$. Following Hamilton (1994, page 10) we obtain responses by simulating the model in the window $t+1, \dots, t+H$ under a baseline and a shocked scenario. The baseline scenario is constructed as follows. Let $i = 1, \dots, I$ be the index of the posterior draws from the MCMC algorithm. We generate I paths of the system by starting from the initial condition $y_t = \ln m_t = 0$. In the baseline scenario, we set all of the shocks to 0 in all periods (this can be thought of as the steady state). In the alternative scenario, we set $\tilde{u}_{t+1} = \sqrt{\sigma^2_{\tilde{u}}}$, i.e., we give a shock of dimension $\sqrt{\sigma^2_{\tilde{u}}}$ to the system in the first period of the window $t+1, \dots, t+H$. The IRFs for each draw are then computed as the difference between the shocked and the baseline scenario: $\left\{ y^i_{t+h,shocked} - y^i_{t+h,baseline} \right\}_{i=1}^{I}$, $h = 1, \dots, H$.

5 Illustrative example and Monte Carlo evaluation

Results based on the general model can be found in Section 6. This section contains an illustrative application and a Monte Carlo evaluation based on a simpler bivariate version of the model, with 1 lag and 1 variable:

$$
\begin{aligned}
y_t &= \Pi_y y_{t-1} + \Pi_m \ln m_{t-1} + \phi \ln m_t + \tilde{\epsilon}_t && \text{(23a)} \\
\ln m_t &= \delta_y y_{t-1} + \delta_m \ln m_{t-1} + \psi \tilde{\epsilon}_t + \tilde{u}_t && \text{(23b)} \\
\ln h_t &= \alpha + \delta \ln h_{t-1} + \tilde{\eta}_t. && \text{(23c)}
\end{aligned}
$$

Note that — since in this illustrative example there is only one variable in the macro block and therefore only one volatility — we removed the assumption of factor structure in the volatility. We did so by setting $\beta = 0$ in (5), which in turns implies $\sigma^2_{y,t} = h_t$ and simplifies (7) into (23c). This simpler specification of the model also facilitates Monte Carlo experiments, whereas the more general specification with $\beta \neq 0$ would greatly complicate the simulation of the model. Indeed, the more general model with $\beta \neq 0$ also involves simultaneity in the conditional variance of the process, and this fact in turn implies that a closed form solution of the model to be used for simulation is not available.

In order to aid interpretation, it is helpful to write down the reduced form of the simultaneous equations (23a)-(23b) above:

$$\begin{bmatrix} y_t \\ \ln m_t \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \ln m_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ u_t \end{bmatrix},$$

where $C_{11}, C_{12}, C_{21}, C_{22}$ are autoregressive coefficients and where the reduced-form errors $(\epsilon_t' \; u_t')'$ have variance $\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}$. The relation between the coefficients of the two representations is:

$$C_{11} = \Pi_y + \phi\delta_y \tag{24a}$$

$$C_{12} = \Pi_m + \phi\delta_m \tag{24b}$$

$$C_{21} = \delta_y \tag{24c}$$

$$C_{22} = \delta_m \tag{24d}$$

$$\Sigma_{11} = \phi^2(\psi^2 h_t + \sigma_{\tilde{u}}^2) + h_t + 2h_t\phi\psi \tag{24e}$$

$$\Sigma_{12} = \phi(\psi^2 h_t + \sigma_{\tilde{u}}^2) + \psi\sigma_\epsilon^2 \tag{24f}$$

$$\Sigma_{22} = \psi^2 h_t + \sigma_{\tilde{u}}^2. \tag{24g}$$

In what follows, we will use the relations above to illustrate the issues possibly arising from shutting down the feedback channel.

5.1 Estimation of the illustrative empirical example

We start with evaluating empirically the relationship between GDP growth and uncertainty in the US. We define y_t as the quarter-on-quarter GDP growth rate and $\ln m_t$ as the JLN measure of macro uncertainty, with quarterly data ranging from 1960Q3 to 2017Q2. Details on the data will be provided in Section 6.1. The posterior means of the parameters (with standard deviations in square brackets) are:

$$y_t = \underbrace{\underset{[0.0705]}{0.2454}}_{\Pi_y} y_{t-1} + \underbrace{\underset{[2.4566]}{2.1583}}_{\Pi_m} \ln m_{t-1} \underbrace{\underset{[2.6035]}{-4.7837}}_{\phi} \ln m_t + \sqrt{h_t}\epsilon_t^*, \tag{25a}$$

$$\ln m_t = \underbrace{\underset{[0.0032]}{-0.0022}}_{\delta_y} y_{t-1} + \underbrace{\underset{[0.0314]}{0.9326}}_{\delta_m} \ln m_{t-1} \underbrace{\underset{[0.0052]}{-0.0088}}_{\psi} \sqrt{h_t}\epsilon_t^* + \tilde{u}_t; \; \sigma_{\tilde{u}}^2 = \underset{[0.0001]}{0.0012} \tag{25b}$$

$$\ln h_t = \underbrace{\underset{[0.0283]}{-0.0416}}_{\alpha} + \underbrace{\underset{[0.0276]}{0.9576}}_{\delta} \ln h_{t-1} + \tilde{\eta}_t; \; \sigma_{\tilde{\eta}}^2 = \underset{[0.0221]}{0.0346} \; ; \; \ln h_0 = \underset{[0.5490]}{0.6620} \; . \tag{25c}$$

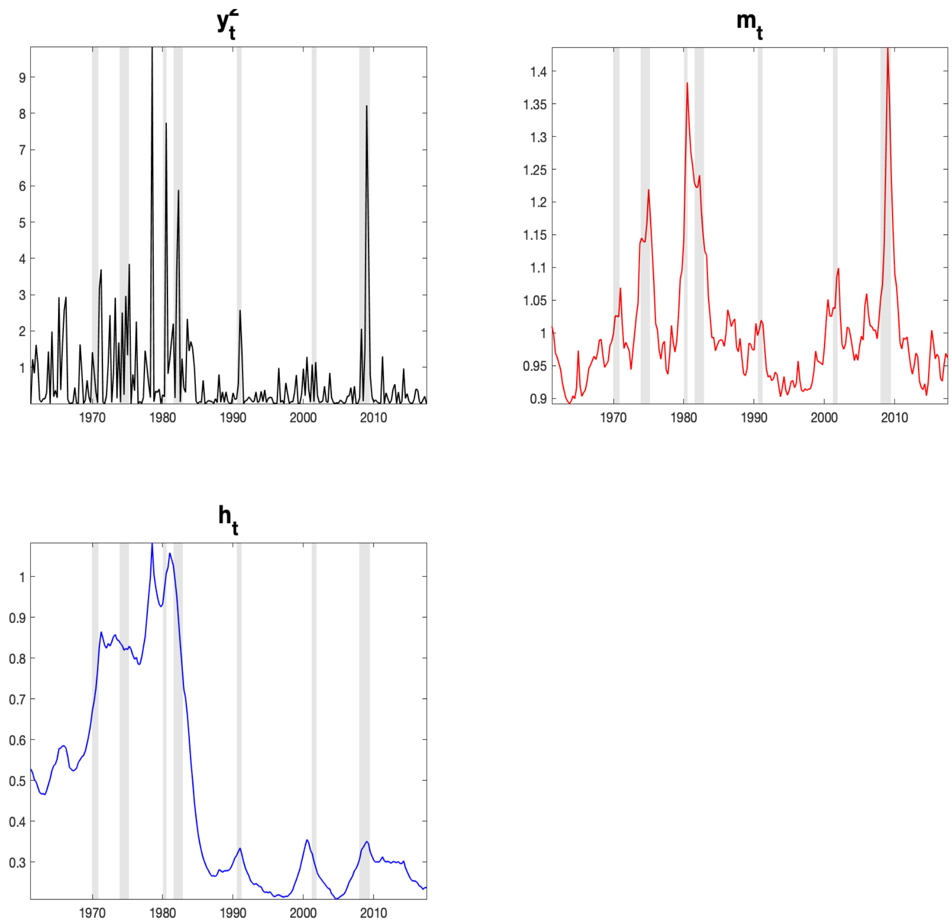

Figure 1: Squared GDP growth rates, uncertainty measure, and reduced-form volatilities. Shaded areas are NBER-dated recessions.

The impact effect of uncertainty on growth as measured by ϕ is negative, and also the effect of growth on uncertainty as measured by ψ is negative, but the value of zero is comfortably included in the 5-95% percentiles of the posterior of ψ.

Figure 1 reports the square of GDP growth, the JLN uncertainty measure, m_t, and the posterior mean of the latent state h_t. The figure highlights the importance of the h_t term for capturing increased volatility during the recessionary periods of the ’70s and early ’80s, and to capture the Great Moderation episode.

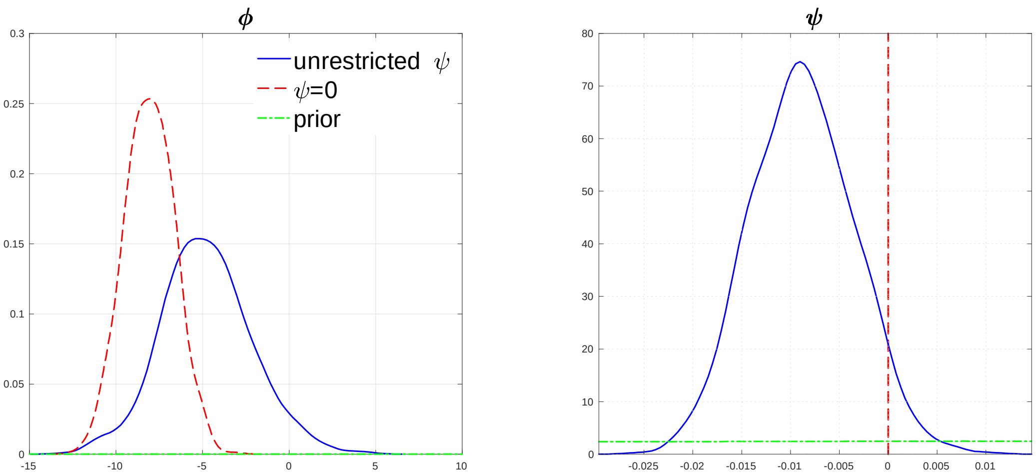

Figure 2: Posterior distributions of ϕ and ψ in (25a) and (25b) when ψ is unrestricted (blue solid line) and when ψ is restricted to 0 (red dashed line).

We now focus on what would happen if we assumed exogeneity of uncertainty, in the sense of setting $\psi = 0$. The results become:

$$
\begin{aligned}
y_t &= \underbrace{\underset{[0.0667]}{0.2365}}_{\Pi_y} y_{t-1} + \underbrace{\underset{[1.4062]}{5.0988}}_{\Pi_m} \ln m_{t-1} \underbrace{\underset{[1.4381]}{-8.0868}}_{\phi} \ln m_t + \sqrt{h_t}\epsilon_t^*, \\
\ln m_t &= \underbrace{\underset{[0.0033]}{-0.0020}}_{\delta_y} y_{t-1} + \underbrace{\underset{[0.0307]}{0.9538}}_{\delta_m} \ln m_{t-1} + \underbrace{0}_{\psi} \sqrt{h_t}\epsilon_t^* + \tilde{u}_t;\ \sigma_{\tilde{u}}^2 = \underset{[0.0001]}{0.0013} \\
\ln h_t &= \underbrace{\underset{[0.0286]}{-0.0426}}_{\alpha} + \underbrace{\underset{[0.0287]}{0.9555}}_{\delta} \ln h_{t-1} + \tilde{\eta}_t;\ \sigma_{\tilde{\eta}}^2 = \underset{[0.0181]}{0.0299}\ ; \ln h_0 = \underset{[0.5217]}{0.6393}.
\end{aligned}
$$

Compared to the case with unrestricted ψ, the posterior of ϕ shifts toward more negative values: the posterior mean goes from -4.7837 to -8.0868 (a decrease of 3.3031). Figure 2 shows the entire posterior distributions of ϕ when ψ is either unrestricted or set to zero. Setting $\psi = 0$ leads to an over-estimate of the negative impact of uncertainty on growth. This happens because, in the more general model with $\psi \neq 0$, following an uncertainty shock there is a decrease in growth, and this in turn increases uncertainty, which further decreases growth, increases uncertainty, and so on. If we shut down the feedback effect of growth on uncertainty by setting $\psi = 0$, the overall impact effect of uncertainty on growth, as measured by ϕ, must increase (in absolute value).

Equation (24b) in this example offers additional insights. Indeed, since the reduced-form coefficient C_{21} cannot change, and $\ln m_t$ and $\ln m_{t-1}$ are highly correlated (δ_m is about 0.93), to compensate for the distortion in ϕ, the parameter Π_m increases by roughly the same amount of the decrease in ϕ. Indeed, the mean of the posterior of Π_m becomes much larger, increasing from 2.1583 to 5.0988, an increase of 2.9405 which almost entirely offsets the 3.3031 decrease in ϕ. This shows that the sum $\phi + \Pi_m$ does not change much, and that under the restriction $\psi = 0$ the model confounds the contemporaneous (ϕ) and the lagged effect (Π_m) of uncertainty on growth.

Moreover, consider equation (24a). Since the product $\phi\delta_y$ is very small, as δ_y is very low relative to ϕ, the coefficient Π_y is virtually unaffected by the exclusion of the feedback channel. Finally, there are virtually no effects on the parameters of the equation for $\ln m_t$ when setting $\psi = 0$. This is not attributable to the insignificance of ψ but rather is a consequence of the fact that ϵ_t^* is uncorrelated with the other regressors, so that its omission does not introduce a distortion.

While Figure 2 makes clear that there is an effect to setting $\psi = 0$, it also shows that the posterior distribution of this parameter comfortably contains the value 0 within the bulk of its probability mass. We can formally compare alternative specifications by using Bayes factors. The unrestricted (i.e., $\psi \neq 0$) and the restricted ($\psi = 0$) specifications have (log) marginal likelihoods of -106.7167 and -118.5822, respectively. In logs, this gives a Bayes factor of 11.86, in strong support of the unrestricted model.

5.2 A Monte Carlo evaluation

This section describes a range of Monte Carlo (MC) experiments, with the aim of illustrating the main features of the model, checking the estimation algorithm, and assessing the effects of some sources of misspecification. We base the data generating process (DGP) on the estimated model (25a)-(25c). All the coefficients are set as in the empirical estimates of equation (25). With this DGP, we simulate 500 time series for y_t, m_t, and h_t with $t = 1, \dots, 250$. We then estimate the model using 60,000 draws for each set of time series, storing at each replication the posterior mean of each parameter, reported in charts below. This provides the empirical distribution of the point estimates across the MC replications. The prior distributions are overlaid on the results in all of the charts provided.

We have examined a wide range of Monte Carlo designs, considering the following issues:

1. Designs 1-3: What are the effects of setting either $\psi = 0$ or $\phi = 0$ in the model?

2. Design 4: What is the effect of the heteroskedasticity of the system going to 0?

3. Design 5: What is the role of the Gaussianity assumption for the shocks ε_t^*?

The complete set of results is available in section 7 of the supplementary appendix. In what follows, we only present a selection of the results focusing only on three key coefficients: Π_m, ψ, ϕ. Results are in Figure 3. In the figure, the rows of panels correspond to alternative Monte Carlo designs, and the columns of panels provide the results for each coefficient. Specifically, the results consist of the empirical distribution of the point estimates across the MC replications, under two alternative scenarios (identified by a blue solid line and a red dashed line).

Omitted and redundant feedback effects We start with Design 1, in which the DGP features $\psi \neq 0$ and the researcher either does or does not impose the restriction $\psi = 0$ in estimation. The results for this MC design are shown in the first line of panels in Figure 3. When the researcher chooses to leave the coefficient ψ unrestricted, the estimated model does not suffer from misspecification, and the resulting posteriors are entirely in line with the true DGP. However, if the researcher imposes the restriction $\psi = 0$, the resulting posteriors are markedly distorted and fail to recover the true values of the coefficients in the DGP. It is interesting to note that such distortions resemble the pattern we have found in the empirical application. Specifically, the posterior mean of ϕ moves to the left, while that of Π_m shifts to the right, so that the posterior mean of the sum $\phi + \Pi_m$ does not move much. As we have already emphasized, this is because the reduced-form parameter $C_{21} = \Pi_m + \phi\delta_m$ cannot change, and — since δ_m is equal to 0.93 — the parameter Π_m must increase by roughly the same amount of the decrease in ϕ.

Next, we consider Design 2, in which we use the coefficient values of the estimated model (25a)-(25c) as DGP, with the relevant exception of setting $\psi = 0$. This means that there is no contemporaneous feedback effect of macro variables on uncertainty; that is, uncertainty is exogenous. As before, the researcher can either impose the restriction or not in the estimated model. This case resembles the case of redundant variables in the estimated model, which typically leads to inefficient — but yet unbiased — estimators. This is precisely what happens, as illustrated in the second row of Figure 3. In this case, both models recover the correct values for the coefficients, but the model imposing the correct restriction $\psi = 0$ attains more precise estimates.

Design 3 considers the effects of erroneously imposing the restriction $\phi = 0$, i.e., no contemporaneous effects of uncertainty on y_t, while in the DGP such a restriction does not hold. We use the coefficient values of the estimated model (25a)-(25c) for the DGP, without any modification being required, since ϕ is already high and significant in these estimates. Results from this design are presented in the third row of Figure 3. As expected, imposing the restriction $\phi = 0$ distorts the results. In particular, the contemporaneous effect

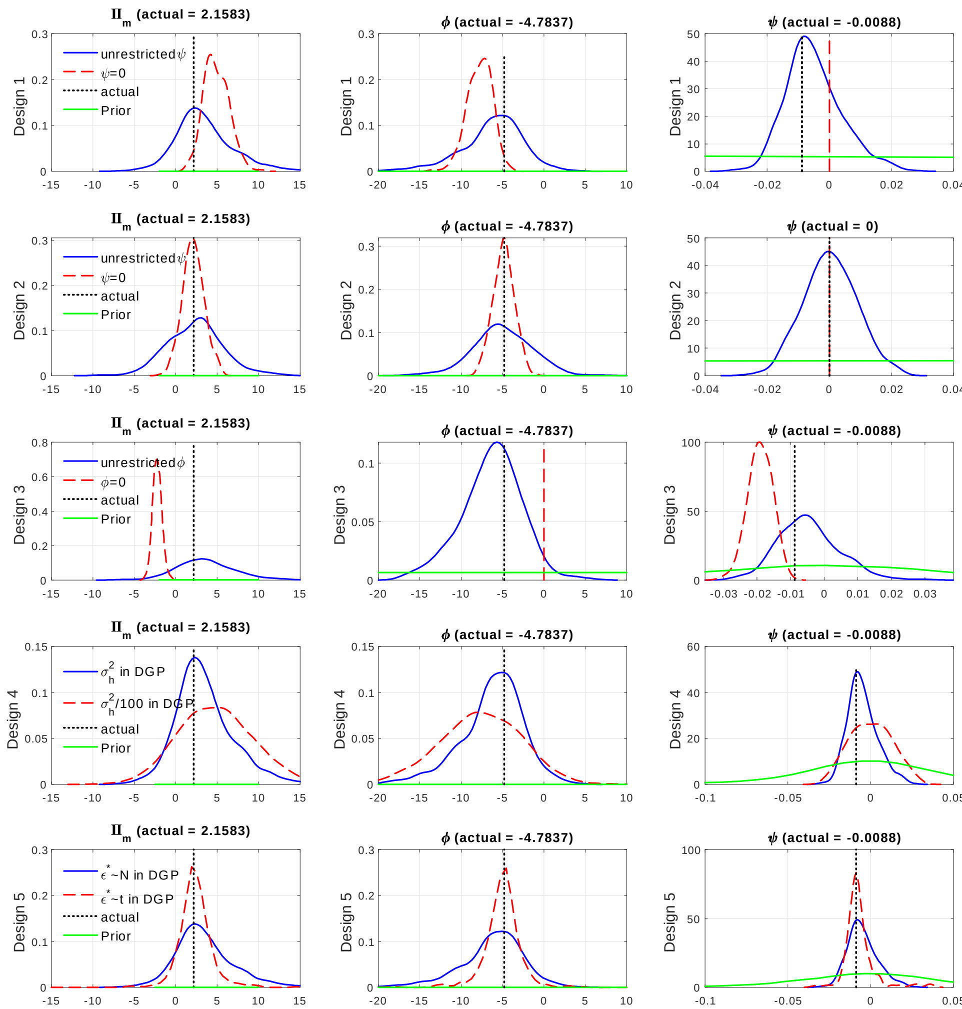

Figure 3: Excerpts from Monte Carlo experiments, designs 1-5. Empirical distributions of posterior means across Monte Carlo replications.

of uncertainty on output (as measured by ϕ) is underestimated (set at $\phi = 0$), and since the reduced-form parameter $C_{21} = \Pi_m + \phi\delta_m$ cannot change, the lagged effect Π_m gets correspondingly over-estimated. Moreover, the distortion in ϕ implies a distortion in ϵ_t^* (it no longer features zero mean), and therefore, ψ gets overestimated (in absolute value).

In light of the three designs described above, the identification problem faced by the standard approach can be rephrased as follows. In the reduced form we have the parameter $C_{21} = \Pi_m + \phi\delta_m$, which is uniquely identified in the likelihood. Recursive identification schemes impose either $\phi = 0$ or $\psi = 0$, which both amount to a potential omitted variable problem. We have that:

- If $\psi = 0$ (uncertainty ordered first), then the estimated ϕ will be more negative than it is in the DGP, and we have an overestimate of Π_m. In this case the identification scheme attributes too much of the impact variation to uncertainty and too little to lagged uncertainty.

- If $\phi = 0$ (uncertainty ordered last), then the estimated ϕ will be less negative than it is in the DGP, and we have an underestimate of Π_m. In this case the identification scheme attributes too little of impact variation to uncertainty and too much to lagged uncertainty.

The approach we propose in this paper solves these issues, since it allows estimating a model that does not put any restrictions on either ψ or ϕ. The results of the Monte Carlo evaluation highlight the importance of properly modeling the endogeneity of uncertainty, and support the interpretation of the empirical findings about the relationship between GDP growth and macro uncertainty. In addition, there is no tendency to spuriously estimate a significant contemporaneous dependence of uncertainty on macro conditions when none exists in the DGP.

Effects of shutting down the heteroskedasticity As we have discussed, identification in this model comes from the heteroskedasticity, and hence it is natural to ask what happens as heteroskedasticity vanishes. Design 4 answers this question, by letting the variance $\sigma^2_{\tilde{\eta}}$ be close to 0. In this case $\sigma^2_{y,t}$ will no longer be a random state variable, but rather will converge to a fixed σ^2_y, and this will make the model converge to a standard homoskedastic VAR. We simulate the model using either $\sigma^2_{\tilde{\eta}} = 0.0346$ or $\sigma^2_{\tilde{\eta}} = 0.000346$ (the remaining coefficients being as in Design 1). Then, we estimate the model and compare the results. This is illustrated in the fourth row of Figure 3. In this case, as the idiosyncratic volatility's variance becomes very small, the model estimates generally differ widely from the true values. This highlights that identification in our model rests on the time-varying volatility.[16] In the empirical application, we will provide overwhelming evidence of its existence.

Effects of mis-specifying the distributional assumption on ε^*_t Design 5 looks at what happens if the distribution of ε^*_t is misspecified. To do so, we simulate artificial data using the estimated model (25), but we modify the shock to (25a) so that it has a t-distribution with 3 degrees of freedom, as opposed to a Gaussian distribution. Then, we estimate the model and compare the results. This is illustrated in the last row of Figure 3. The figure shows that this type of misspecification does have an impact on the

[16] Lewis (2020) provides a framework for inference in models identified via heteroskedasticity in the presence of weak identification.

bias of the estimates. This effect is limited in the case at hand (which is calibrated on our empirical application) but in general it would lead to inconsistency in the posterior in large samples. Sensitivity to this assumption is a particular fragility of the stochastic volatility framework, as the same would not happen in either standard SVARs identified by short or run-long restrictions, or under long-span regime switching specifications of time-varying volatility. The most natural way to address this issue would be to eliminate the possible misspecification by modeling the shocks as a t-student, which is a feasible but computationally demanding extension of our framework.[17] For our application presented below, normality tests performed on the posterior means of the components of ϵ_t^* comfortably support normality for both our specifications and for both datasets. Results can be found in the supplementary appendix's section 6.

6 The economic effects of (endogenous) uncertainty

We now study the relationship between macroeconomic uncertainty and economic variables. We do so using both quarterly and monthly data for the US. The data are described in Section 6.1. Section 6.2 discusses the choice of the specification. Section 6.3 provides the results.

6.1 Data

6.1.1 Uncertainty measures

Assessing the relationship between uncertainty and economic variables requires choosing a concept and measure of uncertainty. The uncertainty literature features a range of both concepts and measures, reflecting the broad definition in the opening paragraph of Bloom (2014, p.153): "Uncertainty is an amorphous concept. It reflects uncertainty in the minds of consumers, managers, and policymakers about possible futures. It is also a broad concept, including uncertainty over the path of macro phenomena like GDP growth, micro phenomena like the growth rate of firms, and non-economic events like war and climate change."

Bloom (2014) goes on to indicate that uncertainty cannot be perfectly measured but can be proxied by a range of measures. One very common proxy is the volatility of stock prices. As noted in Bloom (2009, p.627), stock market volatility "...is strongly linked to other measures of productivity and demand uncertainty." Sources such as Bloom (2014)

[17] A few studies — e.g., Chiu et al. 2017, Clark and Ravazzolo 2015, and Curdia et al. 2015 — have considered models with both stochastic volatility and fat-tails in the conditional innovations. The evidence from this literature seems mixed, with some but not all finding fat tails to be helpful to model fit and forecasting.

discuss some reasons why; as a matter of timeliness, asset prices such as stock prices react quickly to news on the economy, including news relating to economic uncertainty. Other examples of studies using stock volatility-based measures include Basu and Bundick (2017) and Caggiano et al. (2014). Baker et al. (2016) develop an alternative measure of economic policy uncertainty, based on newspaper coverage.

Jurado et al. (2015, p.1177) take a more specific stand on the concept and measure of uncertainty: "At a general level, uncertainty is typically defined as the conditional volatility of a disturbance that is unforecastable from the perspective of economic agents." JLN argue that, for various reasons, common proxies such as stock market volatility need not be tightly linked to such a concept.[18] JLN go on to develop measures of macroeconomic and financial uncertainty based on forecast error variances of large sets of macro indicators and asset returns. Conceptually similar measures have been developed in studies such as Carriero et al. (2018) and Jo and Sekkel (2019).

In this paper we focus on the measure of macroeconomic uncertainty put forward by JLN. In estimates omitted in the interest of brevity, we have verified the robustness of our VAR-based results to instead using the macroeconomic uncertainty measure of Carriero et al. (2018). At the monthly frequency, the correlation of the JLN measure of macro uncertainty with the Carriero et al. estimate is about 0.8.[19]

6.1.2 Other data

For the macroeconomic block of the system, given by equation (1), we considered VARs at both quarterly and monthly frequency. The variables included in the quarterly model are (with acronym and transformation in parenthesis): GDP (GDP, $100^*\Delta \ln$), consumption (CONS, $100^*\Delta \ln$), private investment (INVES, $100^*\Delta \ln$), hours (HOURS, $100^*\Delta \ln$), compensation of employees (COMPE, $100^*\Delta \ln$), GDP deflator (PRICE, $100^*\Delta \ln$), and the federal funds rate (FFR, Δ). The variables included in the monthly model are (with acronym and transformation in parenthesis): Total nonfarm payroll employment (PAYEM, $100^*\Delta \ln$), industrial production (IP, $100^*\Delta \ln$), weekly hours: goods-producing (HOURS,

[18] Berger et al. (2020) present evidence that it is realized stock market volatility, as opposed to uncertainty about the future, that leads to economic fluctuations, and develop a structural model to account for the finding.

[19] In unreported results, we also estimated VARs with a news index-based measure of policy uncertainty (using data back to 1960, with the uncertainty measure provided by Professor Bloom from earlier versions of Baker et al. 2016). In these estimates, shocks to uncertainty have economic effects similar to those reported for the paper's baseline measure. News-based policy uncertainty appears to be endogenous, responding contemporaneously to the business cycle. We interpret that pattern as consistent with the newspaper story basis of the uncertainty measure as being a fast-moving variable, like financial indicators.

100*Δ ln), real consumer spending (SPEND, 100*Δ ln), orders (ORDER, index/100), earnings (EARNI, 100*Δ ln), PCE price index (PCEPI, 100*Δ ln), the federal funds rate (FFR, Δ), and the S&P 500 (S&P, Δ ln).

This specification of monthly variables is very similar to those considered in JLN and Bloom (2009), and contains many of the same variables in the VAR of Caldara et al. (2016).[20] We use four lags (these are sufficient to provide white noise residuals in both the monthly and quarterly applications) and estimate over the sample 1961m7 to 2016m11, for a total of $T = 659$ observations. The set of quarterly variables aligns with those used in structural models such as those of Smets and Wouters (2007). The sample period covers 1960Q3 to 2017Q2. All variables are demeaned prior to estimation to reduce the computational burden. We obtained the macroeconomic data of the monthly model from the FRED-MD database developed in McCracken and Ng (2016) and made available on the website of the Federal Reserve Bank of St. Louis. We took the macro data of the quarterly model from the FAME database of the Federal Reserve Board of Governors.

Although the model is estimated with data transformed as indicated above, for comparability to previous studies, the impulse responses are cumulated and transformed back to the units typical in the literature. Accordingly, the units of the reported impulse responses are percentage point changes (based on 100 times log levels for variables in logs or rates for variables not in log terms). The fact that the model is estimated using some variables differenced for stationarity (e.g., GDP, consumption, and investment) implies that, for some of these variables, the long-run effects of transitory shocks do not die out.

6.2 Volatilities

As a preliminary step, we have estimated some unrestricted time-varying volatilities for the variables of the monthly and quarterly data sets. Specifically, we have used the Bayesian VAR with stochastic volatility described in Carriero et al. (2019). The data show overwhelming evidence of variation in the conditional volatilities. Figure 4 shows the posterior distributions of the time-varying volatilities of these variables.

Moreover, as was the case for the larger data set used in Carriero et al. (2019), there is strong evidence of commonality in the volatilities of our monthly and quarterly variable sets. In particular, the first principal component computed over the resulting estimated volatilities explains 61.47% of the total variance in the monthly dataset.[21] This first principal component is highly correlated with the macroeconomic uncertainty measure of JLN, with

[20] We obtained very similar results with a model augmented to include a credit spread.

[21] In quarterly data, the first principal component computed over the estimated volatilities explains 81.62% of the total variance, and its correlation with the macroeconomic uncertainty measure of JLN is 66.16%.

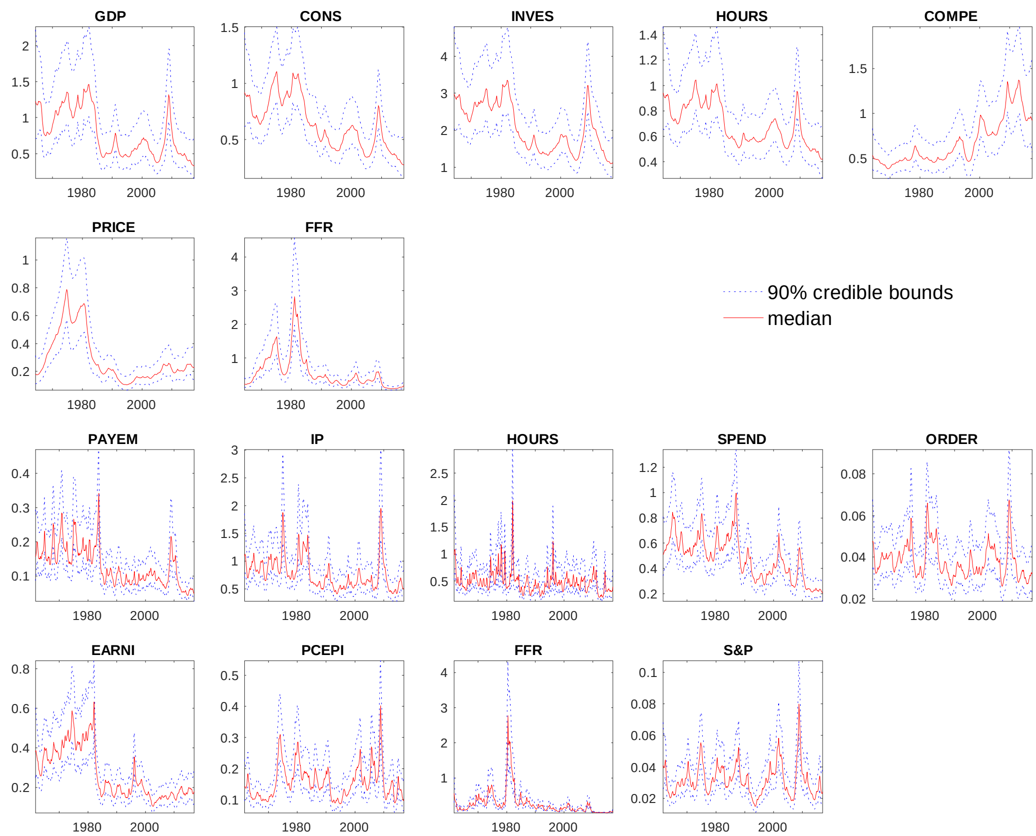

Figure 4: Reduced-form volatilities from an unrestricted Bayesian VAR. Top rows are quarterly results, bottom rows are monthly results.

a correlation of 71.75% in the monthly dataset.

In light of this evidence of commonality in volatilities, in the remainder of the paper we will report results for the volatility specification of equation (5), which features a factor structure with non-zero loadings β. The loadings $\beta_j,\ j = 1, \ldots, n$, measure the proportion with which the common factor impacts the volatility of variable j. The posterior distributions of the loadings for both datasets are displayed in Figure 5. The figure shows that these coefficients generally vary a lot across different variables. They are also generally different from 1.

It is worth noting that all of the results we present below are substantially the same under an alternative specification, closer to the popular VAR specification of Cogley and Sargent (2005) and Primiceri (2005), in which the volatilities σ^2_{jt} do not depend on a common uncertainty factor, but rather they are modeled as idiosyncratic, mutually independent, state variables (this is achieved by setting $\beta_j = 0,\ j = 1, \ldots, n$). This is largely due to the fact that the specification of the conditional variance part of the model mainly enhances the

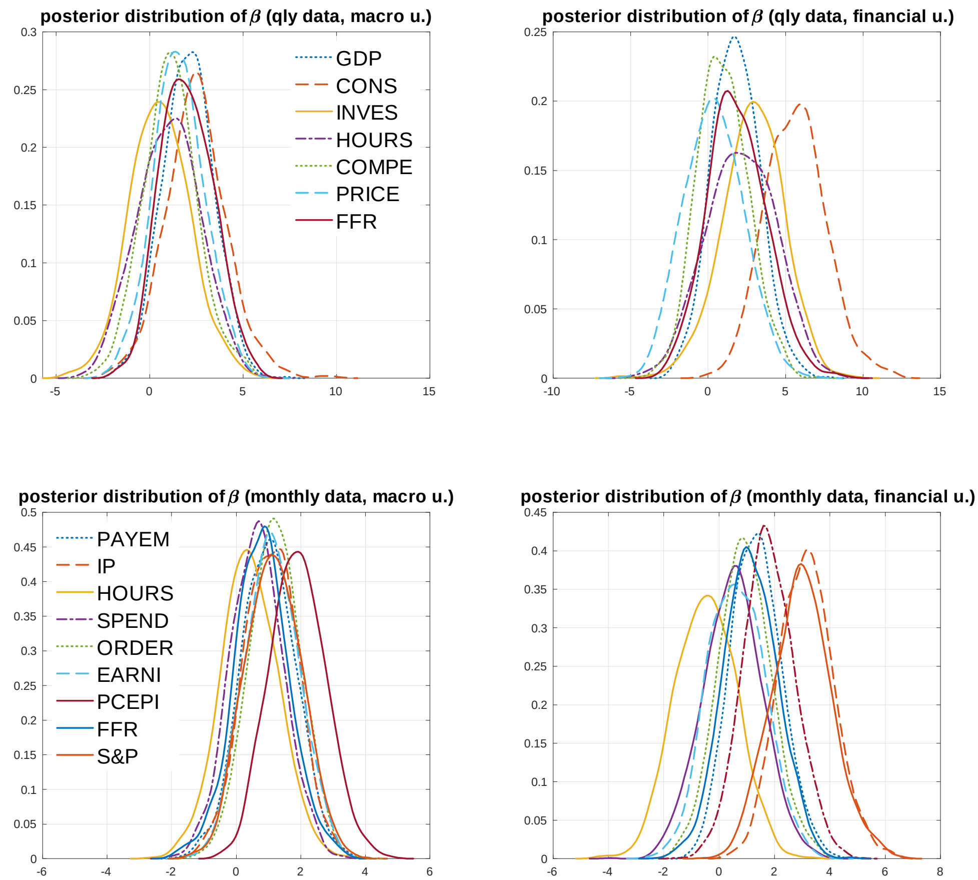

Figure 5: Posterior distributions of the β loadings in (7).

efficiency of the estimates of the conditional mean, but has only a limited effect on the point estimates.

6.3 Macroeconomic uncertainty shocks

Figure 6 shows the posterior distributions of the standard effect coefficients ϕ and the feedback coefficients ψ for the monthly and quarterly VARs.

Focusing first on the standard effect coefficients ϕ, which are on the left-hand side panel of Figure 6, they appear to be largely in line with previous findings about the effects of uncertainty on macroeconomic variables. In particular uncertainty has a large depressive effect on hours, employment, industrial production, consumer spending, and hours. These depressive effects also appear in the quarterly dataset on investment, output, consumption, and hours. A shock to uncertainty also leads to a loosening of the federal funds rate, and an increase in inflation as measured by wages and the PCE price index. These results confirm those of several other studies, such as Baker et al. (2016), Bloom (2009), Carriero et al.

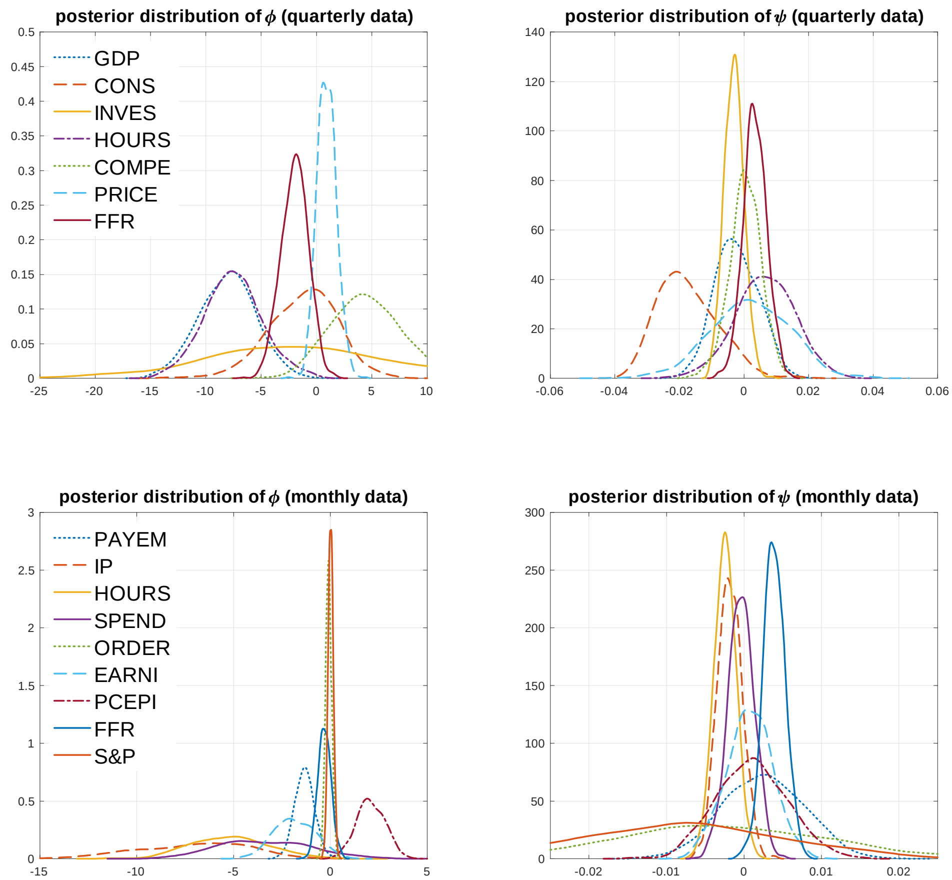

Figure 6: Posterior distributions of the ϕ and ψ coefficients in (1) and (2). VARs with macro (JLN) uncertainty.

(2018), Gilchrist et al. (2014), Jo and Sekkel (2019), and Jurado et al. (2015). They are also in line with the international evidence in Carriero et al. (2020).

Turning our focus to the feedback coefficients ψ, which are on the right-hand side of Figure 6, there are some variables for which these coefficients appear significantly different from 0. Specifically, these variables are consumption, investment, industrial production, hours, and the federal funds rate. The sign of the posterior means is in line with what macroeconomic reasoning would suggest, for example variables such as hours and industrial production tend to reduce uncertainty, whereas the federal funds rate increases macroeconomic uncertainty.

As mentioned in Section 3.3, the ϕ and ψ coefficients depicted in Figure 6 are conditional on the particular ordering of the variables within the macro block. In practice, we have experimented with several alternative orderings in both the monthly and the quarterly model, finding a rather limited impact on the estimated ϕ and ψ coefficients. Section 3 in

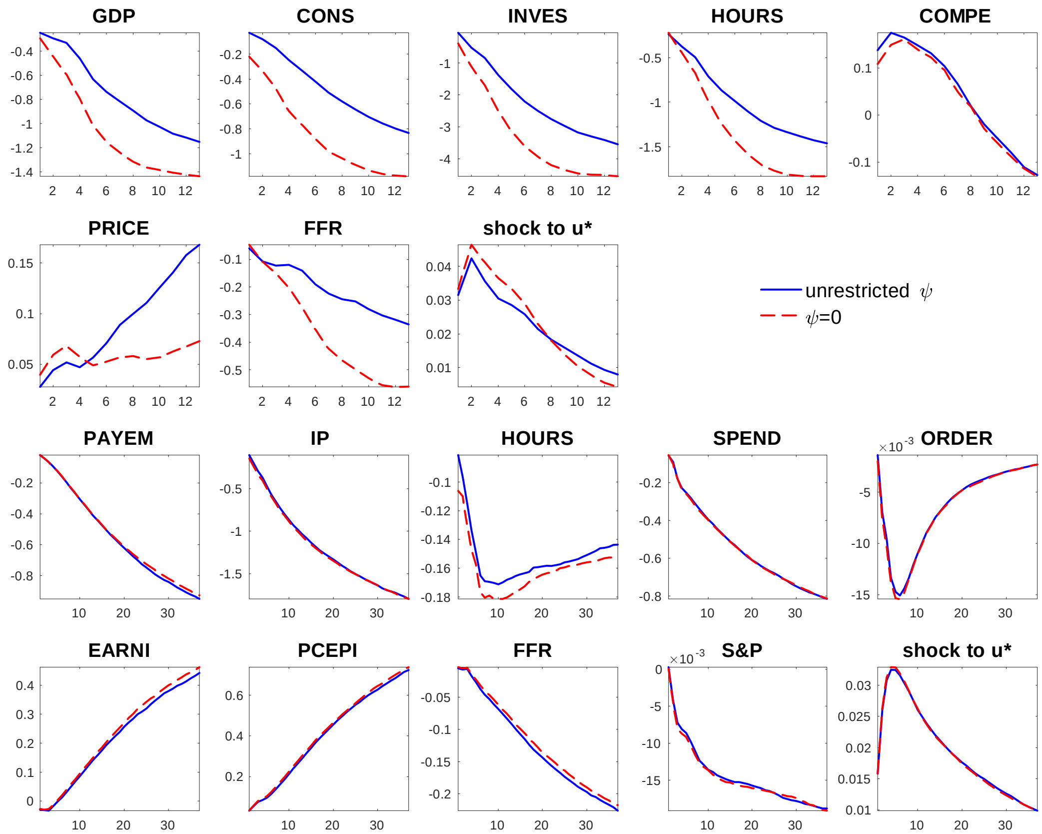

Figure 7: Impulse responses (posterior medians) to a macro uncertainty shock with $\psi = 0$ (red dotted line) and $\psi \neq 0$ (blue solid line). Top rows are quarterly results, bottom rows are monthly results.

the supplementary appendix provides these results.

The coefficients in Figure 6 represent the impact effects on the system. In order to have a complete picture, we now investigate how the differences in the impact effect translate into differences in the dynamics of the system. To do so, we consider the consequences that shutting down the feedback coefficients ψ (which capture the immediate response of uncertainty to economic conditions) has on the coefficients ϕ (which capture the immediate response of economic conditions to uncertainty) and on the impulse responses. This amounts to ordering uncertainty first in a VAR identified through a recursive Cholesky scheme. Figure 7 shows the posterior medians of the impulse responses to a macroeconomic uncertainty shock. In the figure, the unrestricted model is denoted by the solid blue lines, while the model with the feedback effects restricted to zero ($\psi = 0$) is denoted by red dashed lines.

In the top panels of Figure 7, which are based on the quarterly data, it is clear that the differences in the impact effects translate into a different dynamic adjustment after the

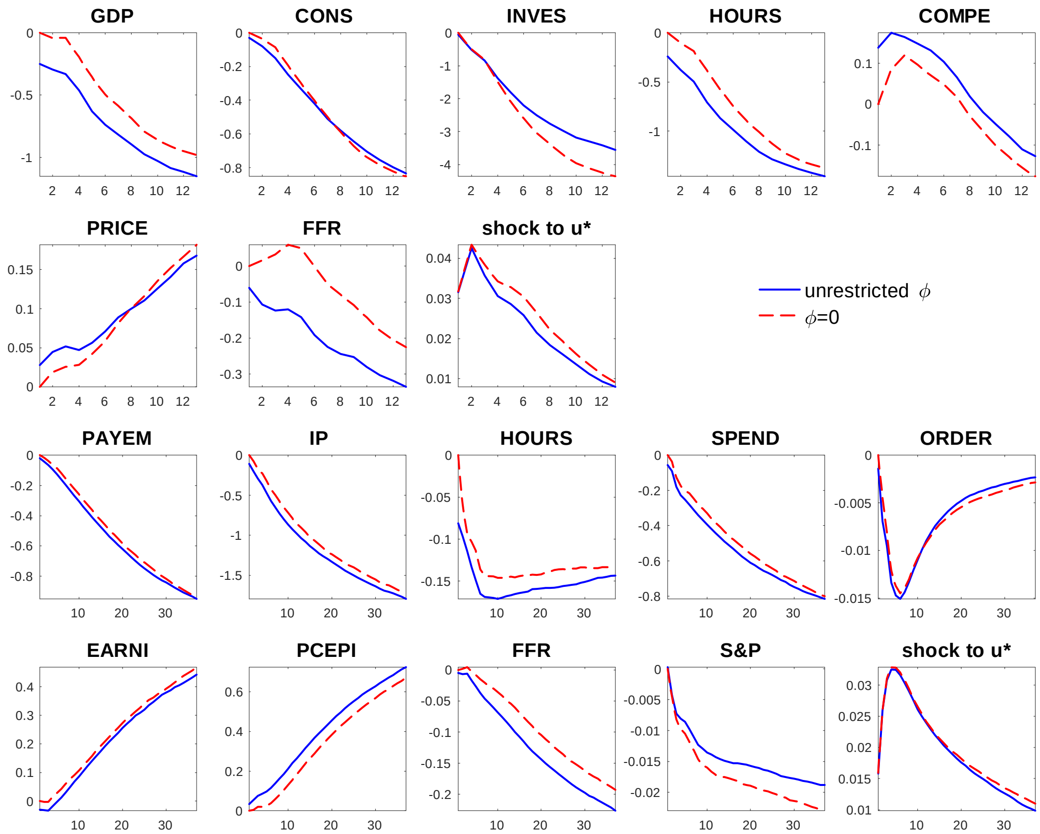

Figure 8: Impulse responses (posterior medians) to a macro uncertainty shock with $\phi = 0$ (red dotted line) and $\phi \neq 0$ (blue solid line). Top rows are quarterly results, bottom rows are monthly results.

shock. Instead, the differences in the bottom panels of Figure 7, which are based on the monthly model, are barely noticeable.[22]

It is interesting to also consider the case of shutting down the standard uncertainty effect (the contemporaneous effect of uncertainty on economic conditions), i.e., to set $\phi = 0$. This amounts to ordering uncertainty last in a VAR identified through a recursive Cholesky scheme. Since — as seen in Figure 6 — the coefficients ϕ are broadly different from zero, and considering also the results we obtained with the MC design in which the researcher

[22]This is not due to a scaling effect, the fact that the charts plot the time series evolution of the median, or that they might conceal differences in higher moments, rather than the posterior median. In order to check that this is not the case, we have examined (see section 4 of the supplementary appendix) the entire posterior distribution of the impulse responses at some selected horizons. These distributions show that, with monthly data, setting $\psi = 0$ has only a slight effect on impact and in the very short run. After that, the differences between the two models quickly die out, and the distributions of the impulse responses become virtually identical.

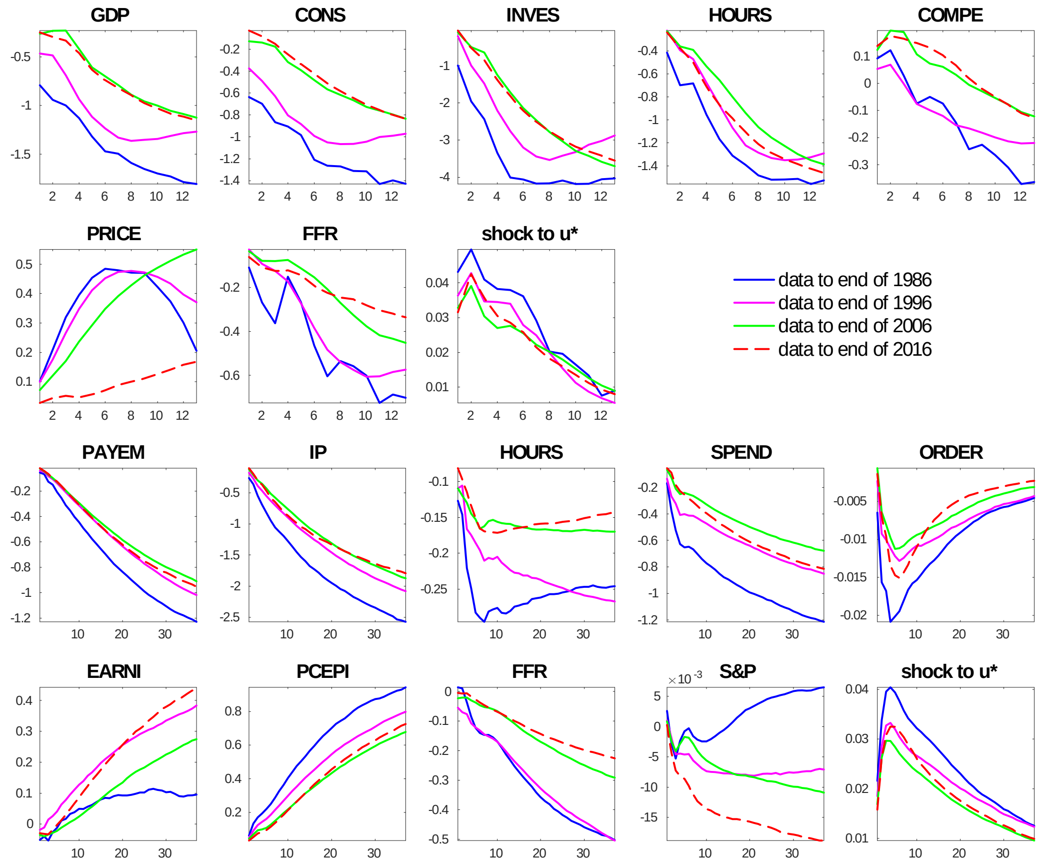

Figure 9: Impulse responses (posterior medians) for different estimation samples.

erroneously imposes $\phi = 0$, we expect shutting down this channel to noticeably affect the impulse responses.

Figure 8 shows the posterior medians of the impulse responses to a macroeconomic uncertainty shock (section 4 of the supplementary appendix displays the entire posterior distribution of the impulse responses at some selected horizons). Clearly, shutting down the standard channel also produces somewhat different impulse responses in the monthly model, and the differences do not completely die out even at the 36-month- (or 12-quarter-) ahead horizons. These results, combined with the Monte Carlo evidence we discussed above (in particular, the design in which the researcher erroneously imposes $\phi = 0$), imply that setting $\phi = 0$ — or, equivalently, ordering macroeconomic uncertainty last in a recursive VAR — could lead to distorted estimation of the effects of macro uncertainty shocks on macroeconomic variables, and a confusion between its contemporaneous and lagged effects.

Finally, in order to assess possible parameter instabilities, we estimated the model over different subsamples and produced impulse responses for each one. As indicated above, the impulse responses reflect most of the model's parameters and thereby provide a simple check of stability. Figure 9 shows the impulse responses computed using the alternative sub-samples ending in 1986, 1996, and 2006, and the full sample estimates based on data up to the end of the sample. As is clear from the Figure, the dynamic response of almost all variables is qualitatively similar in the alternative subsamples. The largest quantitative differences arise when using the sample ending in 1986, i.e. a sample that does not include the interest rate targeting regime and the Great Moderation.

7 Conclusions

Uncertainty is a key variable to understanding economic dynamics, attracting growing interest following the seminal work of Bloom (2009) and the Great Recession of 2007-2009. Several theoretical and empirical papers are by now available on the effects of uncertainty. A general finding from the empirical studies is that uncertainty leads to a deterioration in economic conditions. However, this outcome could be at least partly due to an endogeneity problem. If economic conditions have a contemporaneous effect on uncertainty, ruling it out a priori could result in overestimation of the effects of uncertainty.

In this paper we have developed an econometric model where current and past values of uncertainty affect the current levels of economic variables, and uncertainty is in turn affected by them also contemporaneously. Our model includes stochastic volatility, with time-varying conditional variances of the variables that can be driven by an uncertainty measure plus an idiosyncratic component, or just a stochastic idiosyncratic component. As we show, identification of the model follows from the general results in Lewis (2021) on identification in VARs with time-varying volatility. We derive and provide the relevant conditional posteriors for the states and coefficients of the model, which can be used to estimate the model with a Gibbs sampler.

While the focus of this paper is on uncertainty shocks, the model can be used for any situation in which the researcher wishes to model some of the variables in a vector autoregression as endogenous and there is evidence of time-varying volatility.

Our empirical results point to the conclusion that there is some evidence for the endogeneity of uncertainty. Specifically, we found that some ψ coefficients are nonzero for variables such as industrial production, the federal funds rate, and consumption. We also found that the overall impact of the feedback effect on the system dynamics is visible in the quarterly model while it is rather small in the monthly model. These findings imply that, to reliably assess macroeconomic uncertainty and its effect, it is necessary to depart from a simple recursive ordering and use a more sophisticated approach to identification, such as the one we develop.

Our modeling approach does not put any restrictions on either ψ or ϕ. Still, if a researcher wanted to use a recursive VAR, our results provide two important suggestions for identification. First, when using monthly data, ordering macroeconomic uncertainty first is likely to be harmless. Second, ordering uncertainty last is likely to produce misleading results.

Acknowledgments

Previous versions of this paper circulated with the titles "Endogenous Uncertainty" and "The Identifying Information in Vector Autoregressions with Time-Varying Volatilities: An Application to Endogenous Uncertainty." We gratefully acknowledge research assistance from John Zito and comments from the Editor Serena Ng, two anonymous referees, Daniel Lewis, Kurt Lunsford, Helmut Lütkepohl, Ulrich Müller, Davide Pettenuzzo, Mikkel Plagborg-Møller, Giorgio Primiceri, Harald Uhlig, Molin Zhong, and seminar participants at Bologna University, Cleveland State University, the Dolomiti Macro Meeting, University of Essex, Montreal Econometrics Seminar, University of North Carolina, University of Nottingham, Oxford University, Rice University, UCLA, University of California-Riverside, European Central Bank, and Federal Reserve Board. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System. Marcellino thanks MIUR – PRIN Bando 2017 – prot. 2017TA7TYC for financial support.

Appendix

A. Derivation of the moment matrix

The first term of the matrix appearing in (12) is:

$$
\begin{aligned}
& Cov[\sigma_t^2, \sigma_s^2]G' = \{E[\sigma_t^2 \sigma_s^{2\prime}] - E[\sigma_t^2]E[\sigma_s^{2\prime}]\}G' \\
= & \left\{ E\left[\begin{pmatrix} \sigma_{y,t}^2 \\ \sigma_{\tilde{u}}^2 \end{pmatrix} \begin{pmatrix} \sigma_{y,s}^2 & \sigma_{\tilde{u}}^2 \end{pmatrix} \right] - E\begin{bmatrix} \sigma_{y,t}^2 \\ \sigma_{\tilde{u}}^2 \end{bmatrix} E\begin{bmatrix} \sigma_{y,s}^2 & \sigma_{\tilde{u}}^2 \end{bmatrix} \right\} G' \\
= & \begin{bmatrix} E[\sigma_{y,t}^2 \sigma_{y,s}^2] - E[\sigma_{y,t}^2]E[\sigma_{y,s}^2] & E[\sigma_{\tilde{u}}^2 \sigma_{y,t}^2] - E[\sigma_{\tilde{u}}^2]E[\sigma_{y,t}^2] \\ E[\sigma_{\tilde{u}}^2 \sigma_{y,s}^2] - E[\sigma_{\tilde{u}}^2]E[\sigma_{y,s}^2] & E[\sigma_{\tilde{u}}^4] - E[\sigma_{\tilde{u}}^2]^2 \end{bmatrix} G' \\
= & \begin{bmatrix} Cov_{t,s}[\sigma_{y,t}^2, \sigma_{y,s}^2] & 0 \\ 0 & 0 \end{bmatrix} G' \\
= & \begin{bmatrix} Cov_{t,s}[\sigma_{y,t}^2, \sigma_{y,s}^2] & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

The second term of the M matrix appearing in (12) is:

$$
\begin{aligned}
& E[\sigma_t^2 vec(\varepsilon_s\varepsilon_s' - \Sigma_s)'] \\
= \; & E\left[\begin{bmatrix} \sigma_{y,t}^2 \\ \sigma_{\tilde{u}}^2 \end{bmatrix} vec\left(\begin{bmatrix} \sqrt{\sigma_{y,s}^2} & 0 \\ 0 & \sqrt{\sigma_{\tilde{u}}^2} \end{bmatrix}\begin{pmatrix} \epsilon_s^* \\ u_s^* \end{pmatrix}\begin{pmatrix} \epsilon_s^* & u_s^* \end{pmatrix}\begin{bmatrix} \sqrt{\sigma_{y,s}^2} & 0 \\ 0 & \sqrt{\sigma_{\tilde{u}}^2} \end{bmatrix} - \begin{bmatrix} \sigma_{y,s}^2 & 0 \\ 0 & \sigma_{\tilde{u}}^2 \end{bmatrix}\right)'\right] \\
= \; & E\left[\begin{bmatrix} \sigma_{y,t}^2 \\ \sigma_{\tilde{u}}^2 \end{bmatrix}\left(\begin{bmatrix} \sigma_{y,s}^2(\epsilon_s^{*2}-1) & \sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*\sqrt{\sigma_{\tilde{u}}^2} & \sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*\sqrt{\sigma_{\tilde{u}}^2} & \sigma_{\tilde{u}}^2(u_s^{*2}-1) \end{bmatrix}\right)\right] \\
= \; & E\begin{bmatrix} \sigma_{y,t}^2\sigma_{y,s}^2(\epsilon_s^{*2}-1) & \sigma_{y,t}^2\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*\sqrt{\sigma_{\tilde{u}}^2} & \sigma_{y,t}^2\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*\sqrt{\sigma_{\tilde{u}}^2} & \sigma_{y,t}^2\sigma_{\tilde{u}}^2(u_s^{*2}-1) \\ \sigma_{\tilde{u}}^2\sigma_{y,s}^2(\epsilon_s^{*2}-1) & (\sigma_{\tilde{u}}^2)^{3/2}\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^* & (\sigma_{\tilde{u}}^2)^{3/2}\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^* & (\sigma_{\tilde{u}}^2)^2(u_s^{*2}-1) \end{bmatrix} \\
= \; & \begin{bmatrix} E[\sigma_{y,t}^2\sigma_{y,s}^2(\epsilon_s^{*2}-1)] & \sqrt{\sigma_{\tilde{u}}^2}E[\sigma_{y,t}^2\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*] & \sqrt{\sigma_{\tilde{u}}^2}E[\sigma_{y,t}^2\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*] & \sigma_{\tilde{u}}^2E[\sigma_{y,t}^2(u_s^{*2}-1)] \\ \sigma_{\tilde{u}}^2E[\sigma_{y,s}^2(\epsilon_s^{*2}-1)] & (\sigma_{\tilde{u}}^2)^{3/2}E[\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*] & (\sigma_{\tilde{u}}^2)^{3/2}E[\sqrt{\sigma_{y,s}^2}\epsilon_s^* u_s^*] & 0 \end{bmatrix}.
\end{aligned}
$$

B. Derivation of the joint density of data and states

The joint density of the data and the states $p(y_t, m_t, h_t|\theta)$ can be obtained via the change of variable theorem. We start by re-writing the shocks as follows:

$$
\begin{aligned}
\epsilon_t^* &= \Sigma_{y,t}^{-0.5}A(y_t - \Pi_y(L)y_{t-1} - \Pi_m(L)\ln m_{t-1} - \phi\ln m_t), \\
&= (M_t^{(\beta)})^{-0.5}H_t^{-0.5}A(y_t - \Pi_y(L)y_{t-1} - \Pi_m(L)\ln m_{t-1} - \phi\ln m_t), \qquad \text{(B.1)} \\
\tilde{u}_t &= \ln m_t - \delta_y(L)y_{t-1} - \delta_m(L)\ln m_{t-1} - \psi\Sigma_{y,t}^{0.5}\epsilon_t^*, \\
&= \ln m_t - \delta_y(L)y_{t-1} - \delta_m(L)\ln m_{t-1} \\
&\quad -\psi A(y_t - \Pi_y(L)y_{t-1} - \Pi_m(L)\ln m_{t-1} - \phi\ln m_t) \qquad \text{(B.2)} \\
\tilde{\eta}_{jt} &= \ln h_{jt} - \alpha_j - \delta_j\ln h_{jt-1}, \quad j = 1, \ldots, n. \qquad \text{(B.3)}
\end{aligned}
$$

Collect these shocks in the vector:

$$
\mathbf{r}_t = \begin{bmatrix} \epsilon_t^* \\ u_t^* \\ \tilde{\eta}_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_n & 0 & 0 \\ 0 & \sigma_{\tilde{u}}^2 & 0 \\ 0 & 0 & \Sigma_{\tilde{\eta}} \end{bmatrix}\right),
$$

where $\tilde{\eta}_t = (\tilde{\eta}_{1t}, \ldots, \tilde{\eta}_{nt})'$ and $\Sigma_{\tilde{\eta}}$ is a diagonal matrix with elements $\sigma_{\tilde{\eta}_j}^2$, $j = 1, \ldots, n$. The vector $\mathbf{r}_t$ is a vector of independent Gaussian shocks. The Jacobian of the system (B.1)-(B.3) is

$$
JN_t = \begin{bmatrix} \partial\epsilon_t^*/\partial y_t & \partial\epsilon_t^*/\partial m_t & \partial\epsilon_t^*/\partial h_t \\ \partial\tilde{u}_t/\partial y_t & \partial\tilde{u}_t/\partial m_t & \partial\tilde{u}_t/\partial h_t \\ \partial\tilde{\eta}_t/\partial y_t & \partial\tilde{\eta}_t/\partial m_t & \partial\tilde{\eta}_t/\partial h_t \end{bmatrix},
$$

and can be used in a change of variable to get:

$$
p(y_t, m_t, h_t|\theta) = |JN_t| \times p_G(\epsilon_t^*, \tilde{u}_t, \tilde{\eta}_t).
$$

Since $\partial\tilde{\eta}_t/\partial y_t = \partial\tilde{\eta}_t/\partial m_t = 0$ and $\partial\tilde{\eta}_t/\partial h_t = H_t^{-1}$, the determinant $|JN_t|$ simplifies to:

$$
\begin{aligned}
|JN_t| &= |H_t^{-1}|\left|\begin{pmatrix} \partial\epsilon_t^*/\partial y_t & \partial\epsilon_t^*/\partial m_t \\ \partial\tilde{u}_t/\partial y_t & \partial\tilde{u}_t/\partial m_t \end{pmatrix}\right| \\
&= |H_t^{-1}||\partial\epsilon_t^*/\partial y_t||\partial\tilde{u}_t/\partial m_t - \partial\tilde{u}_t/\partial y_t \cdot (\partial\epsilon_t^*/\partial y_t)^{-1} \cdot \partial\epsilon_t^*/\partial m_t|. \qquad \text{(B.4)}
\end{aligned}
$$

The derivatives are:

$$
\begin{aligned}
\partial\epsilon_t^*/\partial y_t &= (M_t^{(\beta)})^{-0.5} H_t^{-0.5} A,\ \partial\epsilon_t^*/\partial m_t = -(M_t^{(\beta)})^{-0.5} H_t^{-0.5} A\phi m_t^{-1}, \\
\partial\tilde{u}_t/\partial y_t &= \psi A,\ \partial\tilde{u}_t/\partial m_t = m_t^{-1} - \psi A\phi m_t^{-1},
\end{aligned}
$$

which implies that

$$
\partial\tilde{u}_t/\partial y_t \cdot (\partial\epsilon_t^*/\partial y_t)^{-1} \cdot \partial\epsilon_t^*/\partial m_t = -\psi A\phi m_t^{-1},
$$

and therefore the third determinant on the right hand side of (B.4) simplifies to $|m_t^{-1}|$. Hence, the determinant (B.4) is:

$$
\begin{aligned}
|JN_t| &= |H_t^{-1}||(M_t^{(\beta)})^{-0.5} H_t^{-0.5} A||m_t^{-1}| \\
&= m_t^{-1}\Pi_{j=1}^n m_t^{-0.5\beta_j} h_{jt}^{-1.5},
\end{aligned}
$$

and the density in (18) is:

$$
p(y_t, m_t, h_t|\theta) = m_t^{-1}\Pi_{j=1}^n m_t^{-0.5\beta_j} h_{jt}^{-1.5} \underbrace{p_G(\epsilon_t^*)}_{\text{eq. (1)}} \times \underbrace{p_G(\tilde{u}_t)}_{\text{eq. (2)}} \times \underbrace{p_G(\tilde{\eta}_t)}_{\text{eq. (6)}},
$$

where the shocks ϵ_t^*, $\tilde{u}_t$, and $\tilde{\eta}_t$ are those in equations (1), (2), and (6).

C. Derivation of the conditional posterior of the states

In this subsection we derive the expression for the conditional posterior of the states. To do so we consider the data density (18) for the generic idiosyncratic volatility of variable j at time t (i.e. h_{jt}) and recognize that i) since mutual independence of the shocks ensures that $p_G(\epsilon_t^*) = \Pi_{j=1}^n p_G(\epsilon_{jt}^*)$ and $p_G(\tilde{\eta}_t) = \Pi_{j=1}^n p_G(\tilde{\eta}_{jt})$, all the terms not involving variable j can be subsumed in the integrating constant; ii) due to the Markov property featured by h_{jt}, all the terms involving time periods beyond $t-1$ or $t+1$ can be ignored. This gives:

$$
\begin{aligned}
p(h_{jt}|h_{jt-1}, h_{jt+1}, \theta, \mathbf{Y}_{1:T}) \quad \propto \quad & h_{jt}^{-1.5} \exp\left(\frac{-\epsilon_{jt}^{*2} - \epsilon_{jt+1}^{*2}}{2}\right) && \text{(C.1a)} \\
& \times \exp\left(\frac{-\tilde{u}_t^2 - \tilde{u}_{t+1}^2}{2\sigma_{\tilde{u}}^2}\right) && \text{(C.1b)} \\
& \times \exp\left(\frac{-\tilde{\eta}_{jt}^2 - \tilde{\eta}_{jt+1}^2}{2\sigma_{\tilde{\eta}}^2}\right), && \text{(C.1c)}
\end{aligned}
$$

where we also subsumed $m_t^{-1}\Pi_{j=1}^n m_t^{-0.5\beta_j}$ and all of the terms $h_{it}^{-1.5}$ with $i \neq j$ in the integrating constant. Define:

$$e_t = H_t^{0.5}\epsilon_t^* = (M_t^{(\beta)})^{-0.5}A(y_t - \Pi_y(L)y_{t-1} - \Pi_m(L)\ln m_{t-1} - \phi \ln m_t),$$

where we used $H_t^{0.5}(M_t^{(\beta)})^{-0.5}H_t^{-0.5} = (M_t^{(\beta)})^{-0.5}$, which holds true since both $H_t^{0.5}$ and $M_t^{(\beta)}$ are diagonal. Note that e_t is observable conditioning on m_t and y_t (and the coefficients θ). Using $\epsilon_{jt}^* = h_{jt}^{-0.5}e_{jt}$, the term $\exp(-\epsilon_{jt}^{*2}/2)$ in (C.1a) can be written as:

$$\exp\left(-\frac{\epsilon_{jt}^{*2}}{2}\right) = \exp\left(-\frac{e_{jt}^2}{2h_{jt}}\right). \tag{C.2}$$

A similar expression holds for the term $\exp(-\epsilon_{jt+1}^{*2}/2)$ in (C.1a) but this term is redundant as it depends on h_{jt+1}, not h_{jt}. The shocks

$$u_t = \ln m_t - \delta_y(L)y_{t-1} - \delta_m(L)\ln m_{t-1}$$

are observable when conditioning on m_t and y_t (and the coefficients θ). Moreover,

$$\tilde{u}_t = u_t - \psi\tilde{\epsilon}_t = u_t - \psi(M_t^{(\beta)})^{0.5}H_t^{0.5}\epsilon_t^* = u_t - \psi(M_t^{(\beta)})^{0.5}e_t$$

are not only observable, but they do not depend on h_{jt}, which implies that the entire term (C.1b) is redundant. Finally, by completing the squares, the term $\exp((-\tilde{\eta}_{jt}^2 - \tilde{\eta}_{jt+1}^2)/2\sigma_{\tilde{\eta}}^2)$ in (C.1c) can be written as:

$$\exp\left(\frac{-\tilde{\eta}_{jt}^2 - \tilde{\eta}_{jt+1}^2}{2\sigma_{\tilde{\eta}}^2}\right) \propto \exp\left(\frac{-(\ln h_{jt}^2 - \mu_{jt})^2}{2s_{\tilde{\eta}_j}^2}\right), \tag{C.3}$$

where $\mu_{jt} = (\delta_j(\ln h_{jt-1} + \ln h_{jt+1}) + \alpha_j(1-\delta_j))/(\delta_j^2+1)$ and $s_{\tilde{\eta}_j}^2 = \sigma_{\tilde{\eta}_j}^2/(\delta_j^2+1)$.[23]

Using (C.2) and (C.3), the density (C.1) can be written as:

$$p(h_{jt}|h_{jt-1}, h_{jt+1}, \theta, \mathbf{Y}_{1:T}) \propto h_{jt}^{-0.5}\exp\left(-\frac{e_{jt}^2}{2h_{jt}}\right) \times h_{jt}^{-1}\exp\left(\frac{-(\ln h_{jt}^2 - \mu_{jt})^2}{2s_{\tilde{\eta}_j}^2}\right).$$

This is the conditional posterior of the states appearing in equation (20).

References

[1] Angelini, G., E. Bacchiocchi, G. Caggiano, and L. Fanelli (2019), "Uncertainty Across Volatility Regimes," *Journal of Applied Econometrics* 34, 437-455. `https://doi.org/10.1002/jae.2672`

[23] These conditional moments are slightly different in the first and last periods of the sample: in $t = 0$ we have $\mu_{j0} = (\ln h_{j0} - \alpha_j)/\delta_j$ and $s_{\tilde{\eta}_j}^2 = \sigma_{\tilde{\eta}_j}^2/\delta_j^2$. In $t = T$ we have $\mu_{jT} = \alpha_j + \delta_j \ln h_{jT-1}$ and $s_{\tilde{\eta}_j}^2 = \sigma_{\tilde{\eta}_j}^2$.

[2] Arellano, C., Y. Bai, and P. Kehoe (2019), "Financial Frictions and Fluctuations in Volatility," *Journal of Political Economy* 127, 2049-2103. https://doi.org/10.1086/701792

[3] Bachmann, R., and G. Moscarini (2011), "Business Cycles and Endogenous Uncertainty," Discussion paper, Society for Economic Dynamics.

[4] Bachmann, R., S. Elstner, and E. Sims (2013), "Uncertainty and Economic Activity: Evidence from Business Survey Data," *American Economic Journal: Macroeconomics* 5, 217-249. https://doi.org/10.1257/mac.5.2.217

[5] Baker S., N. Bloom, and S. Davis (2016), "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics* 131, 1593-1636. https://doi.org/10.1093/qje/qjw024

[6] Basu S., and B. Bundick (2017), "Uncertainty Shocks in a Model of Effective Demand," *Econometrica* 85, 937-958. https://doi.org/10.3982/ECTA13960

[7] Berger, David, I. Dew-Becker, and S. Giglio (2020), "Uncertainty Shocks as Second-Moment News Shocks," *Review of Economic Studies* 87, 40-76. https://doi.org/10.1093/restud/rdz010

[8] Bernanke, B. (1983), "Irreversibility, Uncertainty, and Cyclical Investment," *Quarterly Journal of Economics* 98, 85-106. https://doi.org/10.2307/1885568

[9] Bertsche, D., and R. Braun (2020), "Identification of Structural Vector Autoregressions by Stochastic Volatility," *Journal of Business and Economic Statistics*, forthcoming. https://doi.org/10.1080/07350015.2020.1813588

[10] Bloom, N. (2009), "The Impact of Uncertainty Shocks," *Econometrica* 77, 623-685. https://doi.org/10.3982/ECTA6248

[11] Bloom, N. (2014), "Fluctuations in Uncertainty," *Journal of Economic Perspectives* 28, 153-176. https://doi.org/10.1257/jep.28.2.153

[12] Bloom N., M. Floetotto, M. Jaimovich, I. Saporta-Eksten, and S. Terry (2018), "Really Uncertain Business Cycles," *Econometrica* 86, 1031-1065. https://doi.org/10.3982/ECTA10927

[13] Bognanni, M. (2018), "A Class of Time-Varying Parameter Structural VARs for Inference under Exact or Partial Identification," Federal Reserve Bank of Cleveland, Working Paper no. 18-11. https://doi.org/10.26509/frbc-wp-201811

[14] Caggiano, G., E. Castelnuovo, and N. Groshenny (2014), "Uncertainty Shocks and Unemployment Dynamics: An Analysis of Post-WWII US Recessions," *Journal of Monetary Economics* 67, 78-92. https://doi.org/10.1016/j.jmoneco.2014.07.006

[15] Caldara, D., C. Fuentes-Albero, S. Gilchrist, and E. Zakrajsek (2016), "The Macroeconomic Impact of Financial and Uncertainty Shocks," *European Economic Review* 88, 185-207. https://doi.org/10.1016/j.euroecorev.2016.02.020

[16] Carriero, A., T. Clark, and M. Marcellino (2016), "Common Drifting Volatility in Large Bayesian VARs," *Journal of Business and Economic Statistics* 34, 375-390. https://doi.org/10.1080/07350015.2015.1040116

[17] Carriero, A., T. Clark, and M. Marcellino (2018), "Measuring Uncertainty and Its Impact on the Economy," *Review of Economics and Statistics* 100, 799-815. https://doi.org/10.1162/REST_a_00693

[18] Carriero, A., T. Clark, and M. Marcellino (2019), "Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors," *Journal of Econometrics* 212, 137-154. https://doi.org/10.1016/j.jeconom.2019.04.024

[19] Carriero, A., T. Clark, and M. Marcellino (2020), "Assessing International Commonality in Macroeconomic Uncertainty and its Effects," *Journal of Applied Econometrics* 35, 273-293. https://doi.org/10.1002/jae.2750

[20] Cesa-Bianchi, A., M. Pesaran, and A. Rebucci (2020), "Uncertainty and Economic Activity: A Multi-Country Perspective," *Review of Financial Studies* 33, 3393-3445. https://doi.org/10.1093/rfs/hhz098

[21] Chan, J.C.C., and E. Eisenstat (2018), "Bayesian Model Comparison for Time-Varying Parameter VARs with Stochastic Volatility," *Journal of Applied Econometrics* 33, 509-532. https://doi.org/10.1002/jae.2617

[22] Chiu, C., H. Mumtaz, and G. Pinter (2017), "Forecasting with VAR Models: Fat Tails and Stochastic Volatility," *International Journal of Forecasting* 33, 1124-1143. https://doi.org/10.1016/j.ijforecast.2017.03.001

[23] Clark, T. (2011), "Real-Time Density Forecasts from BVARs with Stochastic Volatility," *Journal of Business and Economic Statistics* 29, 327-341. https://doi.org/10.1198/jbes.2010.09248

[24] Clark, T., and F. Ravazzolo (2015), "Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics* 30, 551-575. https://doi.org/10.1002/jae.2379

[25] Curdia, V., M. Del Negro, and D. Greenwald (2015), "Rare Shocks, Great Recessions," *Journal of Applied Econometrics* 29, 1031-1052. https://doi.org/10.1002/jae.2395

[26] D'Agostino, A., L. Gambetti, and D. Giannone (2013), "Macroeconomic Forecasting and Structural Change," *Journal of Applied Econometrics* 28, 82-101. https://doi.org/10.1002/jae.1257

[27] Del Negro, M., and G. Primiceri (2015), "Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum," *Review of Economic Studies* 82, 1342-1345. https://doi.org/10.1093/restud/rdv024

[28] Diebold, F., F. Schorfheide, and M. Shin (2017), "Real-Time Forecast Evaluation of DSGE Models with Stochastic Volatility," *Journal of Econometrics* 201, 322-332. https://doi.org/10.1016/j.jeconom.2017.08.011

[29] Fajgelbaum, P., E. Schaal, and M. Taschereau-Dumouchel (2017), "Uncertainty Traps," *Quarterly Journal of Economics*, 132, 1641-1692. https://doi.org/10.1093/qje/qjx021

[30] Fernandez-Villaverde, J., P. Guerron-Quintana, J. Rubio-Ramirez, and M. Uribe (2011), "Risk Matters: The Real Effects of Volatility Shocks," *American Economic Review* 101, 2530-2561. https://doi.org/10.1257/aer.101.6.2530

[31] Geweke, J. (2004) "Getting It Right: Joint Distribution Tests of Posterior Simulators," *Journal of the American Statistical Association* 99, 799-804. https://doi.org/10.1198/016214504000001132

[32] Gilchrist S., J. Sim, and E. Zakrajsek (2014), "Uncertainty, Financial Frictions, and Investment Dynamics," NBER Working Paper 20038. https://doi.org/10.3386/w20038

[33] Hamilton, J. (1994), *Time Series Analysis*, Princeton, N.J.: Princeton University Press.

[34] Jacquier, E., N. Polson, and P. Rossi (2004), "Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors," *Journal of Econometrics* 122, 185-212. https://doi.org/10.1016/j.jeconom.2003.09.001

[35] Jo, S., and R. Sekkel (2019), "Macroeconomic Uncertainty Through the Lens of Professional Forecasters," *Journal of Business and Economic Statistics* 37, 436-446. https://doi.org/10.1080/07350015.2017.1356729

[36] Jurado, K., S. Ludvigson, and S. Ng (2015), "Measuring Uncertainty," *American Economic Review* 105, 1177-1216. https://doi.org/10.1257/aer.20131193

[37] Justiniano, A., and G. Primiceri (2008), "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review* 98, 604-641. https://doi.org/10.1257/aer.98.3.604

[38] Kilian, L., and H. Lütkepohl (2017), *Structural Vector Autoregressive Analysis*, Cambridge University Press. https://doi.org/10.1017/9781108164818

[39] Koop, G., and S. Potter (2007), "Estimation and Forecasting in Models with Multiple Breaks," *The Review of Economic Studies* 74, 763-789. https://doi.org/10.1111/j.1467-937X.2007.00436.x

[40] Lanne, M., and H. Lütkepohl (2008), "Identifying Monetary Policy Shocks via Changes in Volatility," *Journal of Money, Credit, and Banking* 40, 1131-1149. https://doi.org/10.1111/j.1538-4616.2008.00151.x

[41] Lanne, M., H. Lütkepohl, and K. Maciejowska (2010), "Structural Vector Autoregressions with Markov Switching," *Journal of Economic Dynamics and Control* 34, 121-131. https://doi.org/10.1016/j.jedc.2009.08.002

[42] Lanne, M., M. Meitz, and P. Saikkonen (2017), "Identification and Estimation of non-Gaussian Structural Vector Autoregressions," *Journal of Econometrics* 196, 288-304. https://doi.org/10.1016/j.jeconom.2016.06.002

[43] Leduc S., and Z. Liu (2012), "Uncertainty Shocks Are Aggregate Demand Shocks," *Journal of Monetary Economics* 82, 20-35. https://doi.org/10.1016/j.jmoneco.2016.07.002

[44] Lewis, D. (2021), "Identifying Shocks via Time-Varying Volatility," forthcoming, *Review of Economics Studies*. https://doi.org/10.1093/restud/rdab009

[45] Lewis, D. (2020), "Robust Inference in Models Identified via Heteroskedasticity," forthcoming, *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_00963

[46] Ludvigson, S., S. Ma, and S. Ng (2019), "Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?" *American Economic Journal: Macroeconomics*, forthcoming

[47] Lütkepohl, H., and A. Netšunajev (2017), "Structural Vector Autoregressions with Smooth Transition in Variances," *Journal of Economic Dynamics and Control* 84, 43-57. https://doi.org/10.1016/j.jedc.2017.09.001

[48] McCracken, M., and S. Ng, (2016), "FRED-MD: A Monthly Database for Macroeconomic Research," *Journal of Business and Economic Statistics* 34, 574-589. https://doi.org/10.1080/07350015.2015.1086655

[49] McDonald, R., and D. Siegel (1986), "The Value of Waiting to Invest," *Quarterly Journal of Economics* 101, 707-727. https://doi.org/10.2307/1884175

[50] Normandin, M., and L. Phaneuf (2004), "Monetary Policy Shocks: Testing Identification Conditions Under Time-Varying Conditional Volatility," *Journal of Monetary Economics* 51, 1217-1243. https://doi.org/10.1016/S0304-3932(04)00069-8

[51] Primiceri, G.E. (2005), "Time-Varying Structural vector autoregressions and Monetary Policy," *Review of Economic Studies* 72, 821-852. https://doi.org/10.1111/j.1467-937X.2005.00353.x

[52] Rigobon, R. (2003), "Identification through Heteroskedasticity," *Review of Economics and Statistics* 85, 777-792. https://doi.org/10.1162/003465303772815727

[53] Rossi, B., and T. Sekhposyan (2015), "Macroeconomic Uncertainty Indices Based on Nowcast and Forecast Error Distributions," *American Economic Review* 105, 650-655. https://doi.org/10.1257/aer.p20151124

[54] Sentana, E., and G. Fiorentini (2001), "Identification, Estimation and Testing of Conditionally Heteroskedastic Factor Models," *Journal of Econometrics* 102, 143-164. https://doi.org/10.1016/S0304-4076(01)00051-3

[55] Shin, M., and M. Zhong (2020), "A New Approach to Identifying the Real Effects of Uncertainty Shocks,"*Journal of Business and Economic Statistics* 8, 367-379. https://doi.org/10.1080/07350015.2018.1506342

[56] Smets, F., and R. Wouters (2007), "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review* 97, 586-606. https://doi.org/10.1257/aer.97.3.586

Using Time-Varying Volatility for Identification in Vector Autoregressions: An Application to Endogenous Uncertainty

Andrea Carriero[*] Todd E. Clark[†] Massimiliano Marcellino[‡]

This version: July 2021

Contents

[*]Queen Mary, University of London, a.carriero@qmul.ac.uk
[†]Federal Reserve Bank of Cleveland, todd.clark@clev.frb.org
[‡]Bocconi University, IGIER and CEPR, massimiliano.marcellino@unibocconi.it

This supplementary appendix provides additional results and robustness checks. These results refer to two different specifications of the volatility process for the model. We use "specification 1" to refer to a model restricted to make all factor loadings β_j equal 0 and "specification 2" to refer to the model used in the paper, in which the factor loadings are not restricted to 0.

1 Algorithm diagnostics

In this section we evaluate the correctness, convergence, and mixing properties of the MCMC sampler. Figure 1 reports the quantile-quantile plots of Geweke's (2004) test of correctness of the posterior sampler, computed using the univariate version of the model. In particular, the plots compare the quantiles of simulated draws of a few observations of the data and states (y_t, $\ln m_t$, $\ln h_t$) and the model's parameters, for draws from the marginal-conditional simulator and from the successive-conditional simulator.[1] The quantiles from the two simulators line up closely. Accordingly, the Geweke diagnostic indicates correctness of the sampler.

Figure 2 reports the potential scale reduction factors and inefficiency factors for all of the coefficients in the model for the monthly dataset (results are very similar for all the other specifications). Results are organized in groups. Results for the coefficients in the y_t equations (θ_1) are reported in the plots in the first column on the left-hand side, results for the coefficients in the uncertainty equation (θ_2) are reported in the plots in the central column, and results for the coefficients of the idiosyncratic volatilities processes (θ_3) are reported in the plots in the last column, on the right-hand side.

[1] Our choice of statistics and quantile-quantile patterns is patterned on the implementation of the Geweke (2004) sampler test in the online appendix of Del Negro and Primiceri (2015).

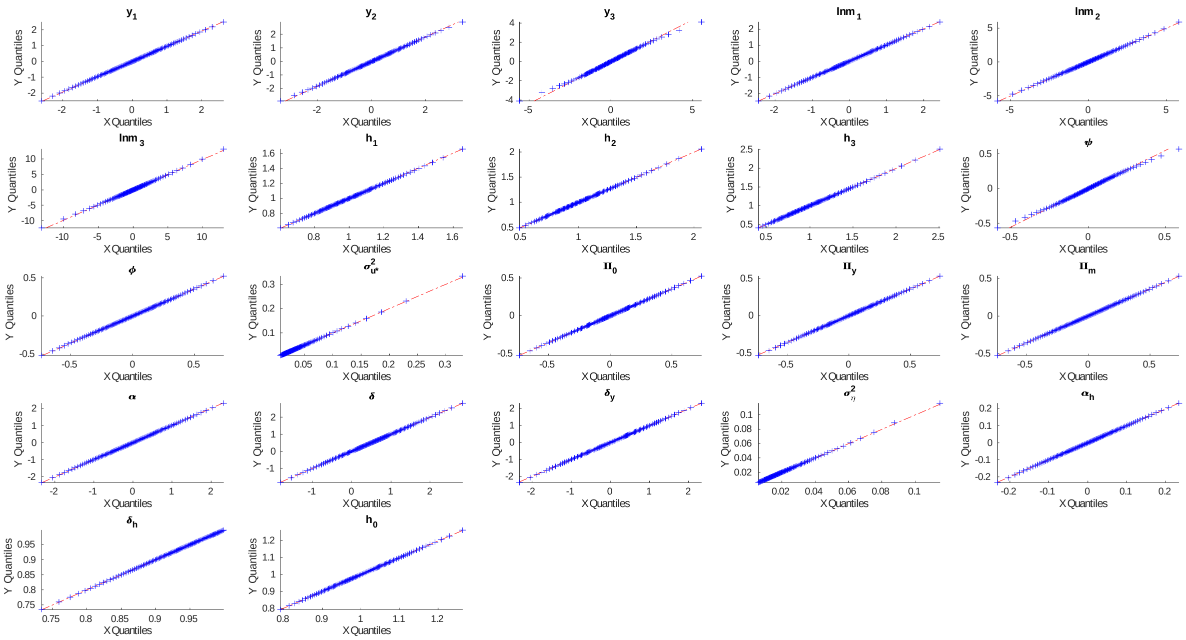

Figure 1: Geweke (2004) test of correctness of the posterior sampler, q-q plots.

2 Results from Specification 1 ($\beta = 0$)

This section contains the entire set of results for Specification1. Results are organised as follows:

- Figure 3: model with $\beta = 0$, posteriors of ϕ and ψ
- Figure 4: IRFs of the model with $\beta = 0$, cases $\psi = 0$ vs $\psi \neq 0$
- Figure 5: IRFs of the model with $\beta = 0$, cases $\phi = 0$ vs $\phi \neq 0$
- Figure 6: Comparison of models with and without loadings.

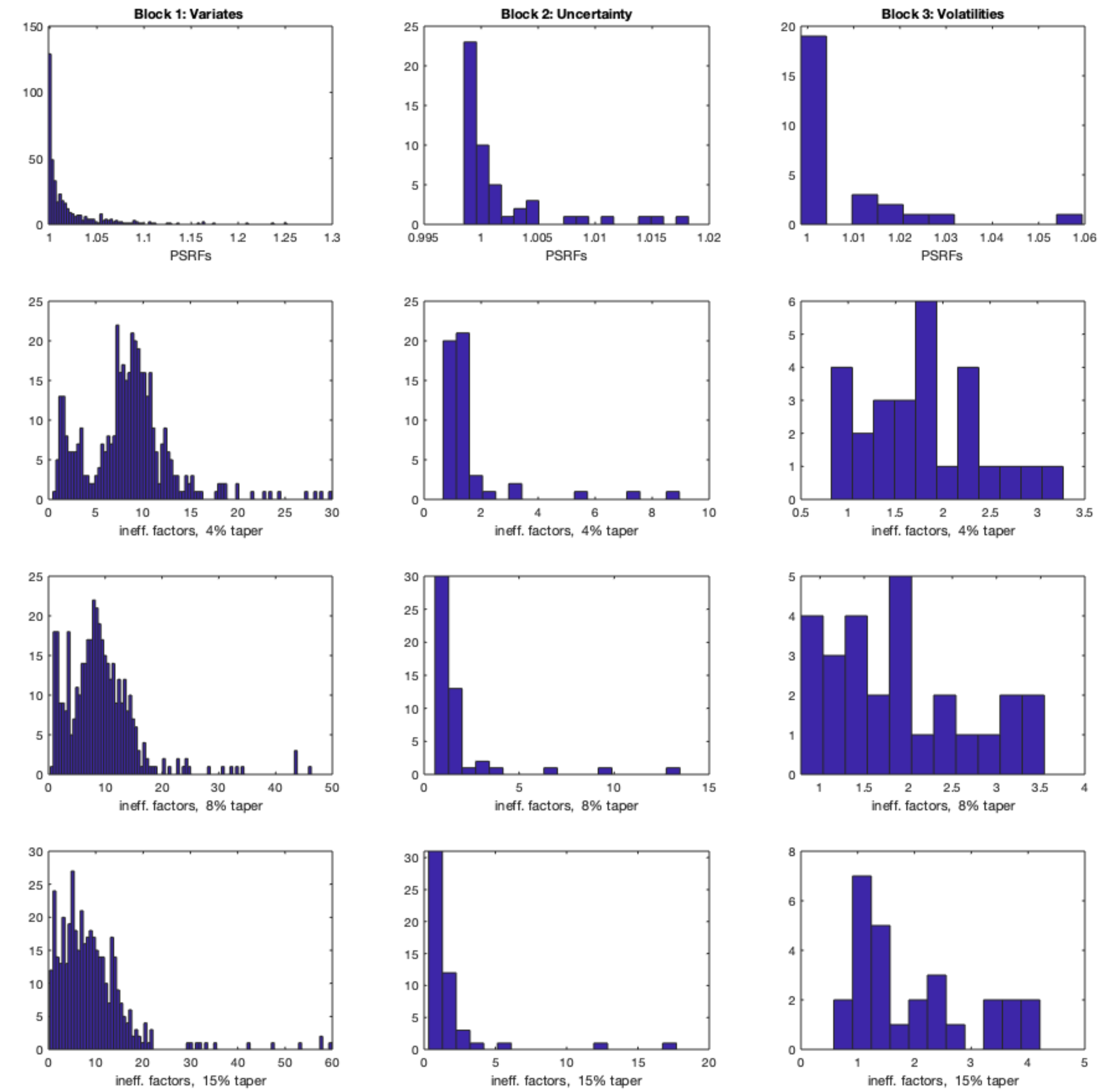

Figure 2: Potential scale reduction factors and inefficiency factors for the simulated draws of the model coefficients.

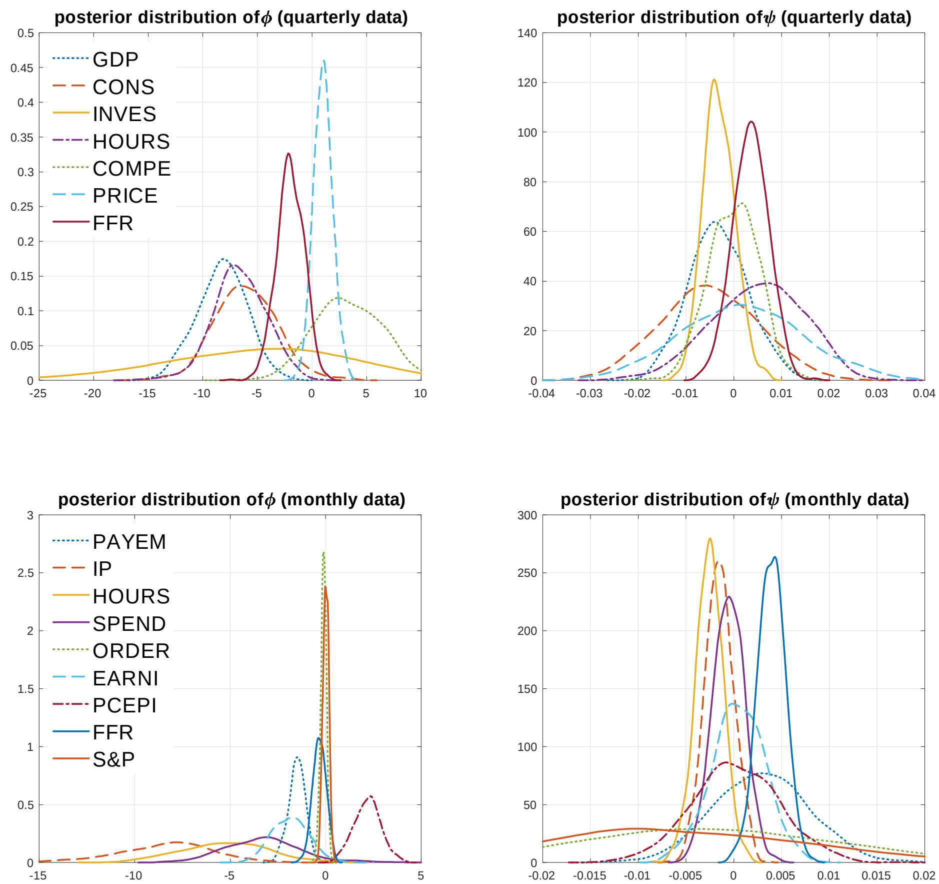

Figure 3: Model with $\beta = 0$

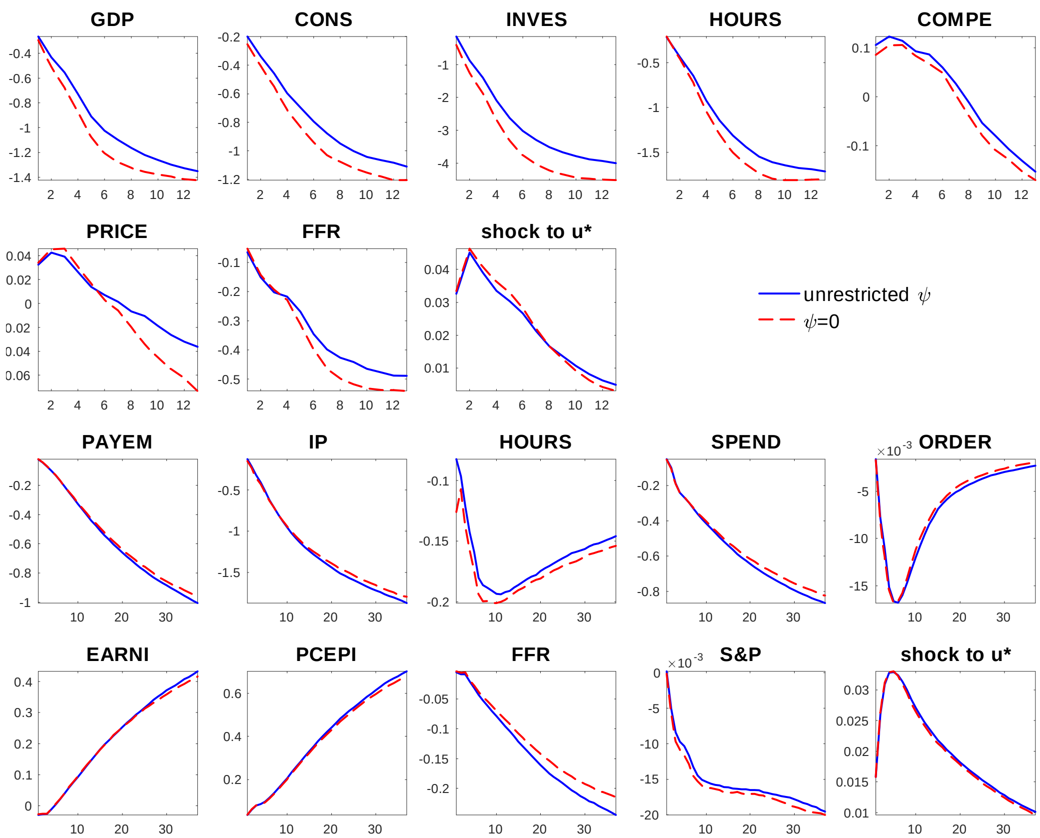

Figure 4: Impulse responses to an uncertainty shock with $\psi = 0$ and $\psi \neq 0$. Rows 1 and 2: quarterly dataset. Rows 3 and 4: monthly dataset. Model without factor structure in the volatilities ($\beta = 0$).

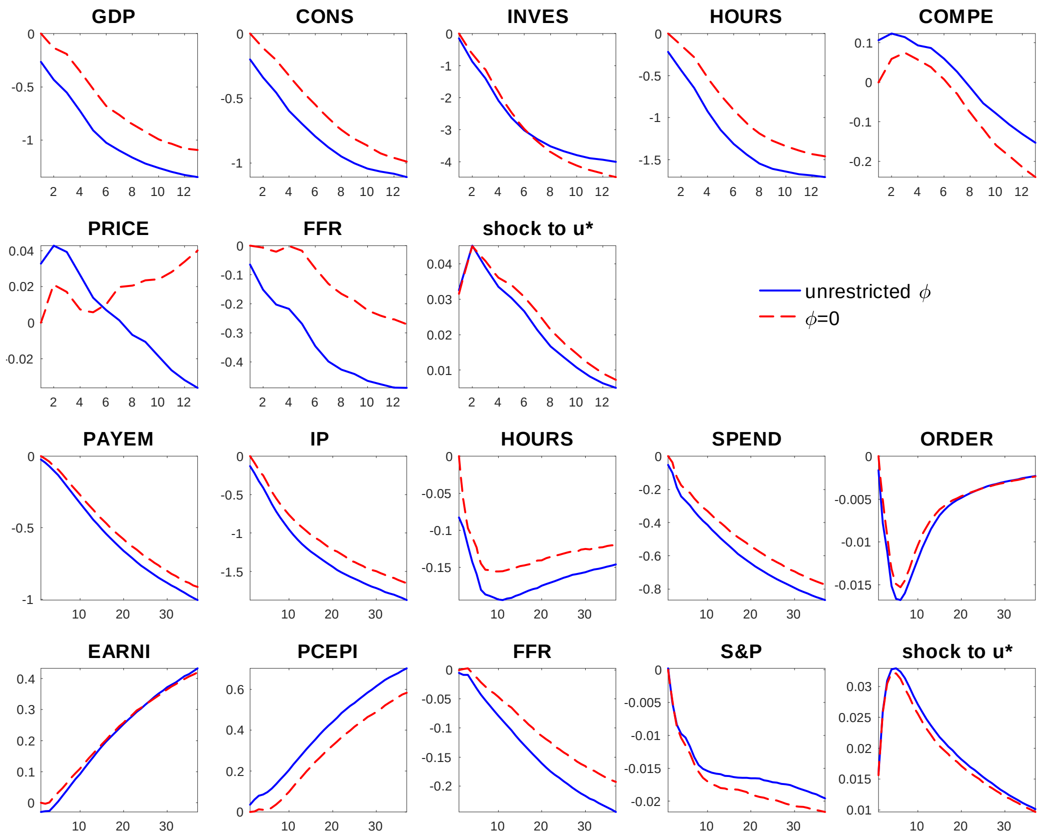

Figure 5: Impulse responses to an uncertainty shock with $\phi = 0$ and $\phi \neq 0$. Rows 1 and 2: quarterly dataset. Rows 3 and 4: monthly dataset. Model without factor structure in the volatilities ($\beta = 0$).

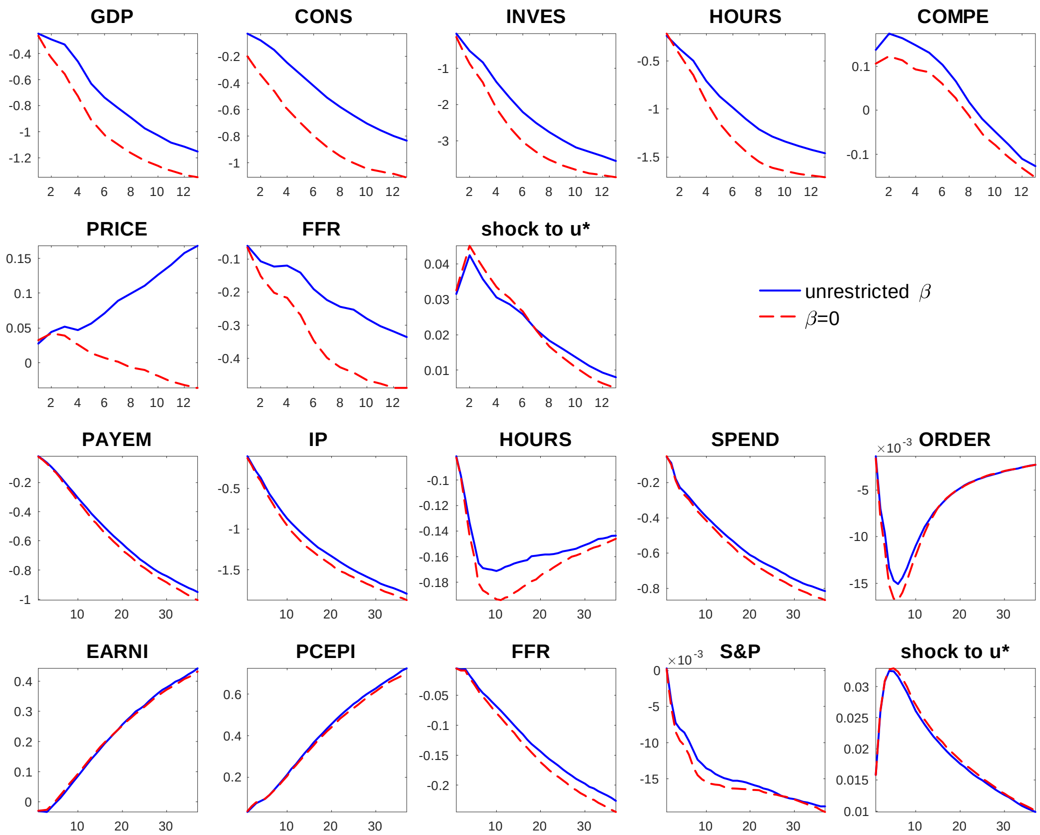

Figure 6: Comparison of models with and without loadings. These are models with ϕ and ψ both different from 0.

3 Robustness to re-orderings within the macro block

In this section we evaluate the robustness of our results to alternative orderings **within** the macro block. This does not imply that there is a different ordering between the macro block and the uncertainty measure block: these two blocks continue to be simultaneous as in the baseline model. We consider the inverted ordering for both the quarterly and monthly model. For the monthly model, we also consider (i) a model in which the S&P 500 is omitted from the cross section of variables, (ii) the ordering of Bloom (2009), and (iii) the inverse of the ordering of Bloom (2009). As seen in the graphs, results are broadly robust to these permutations.

- Figure 7: IRFs for alternative orderings
- Figure 8: Coefficients ϕ for alternative orderings, quarterly data
- Figure 9: Coefficients ψ for alternative orderings, quarterly data
- Figure 10: Coefficients ϕ for alternative orderings, monthly data
- Figure 11: Coefficients ψ for alternative orderings, monthly data

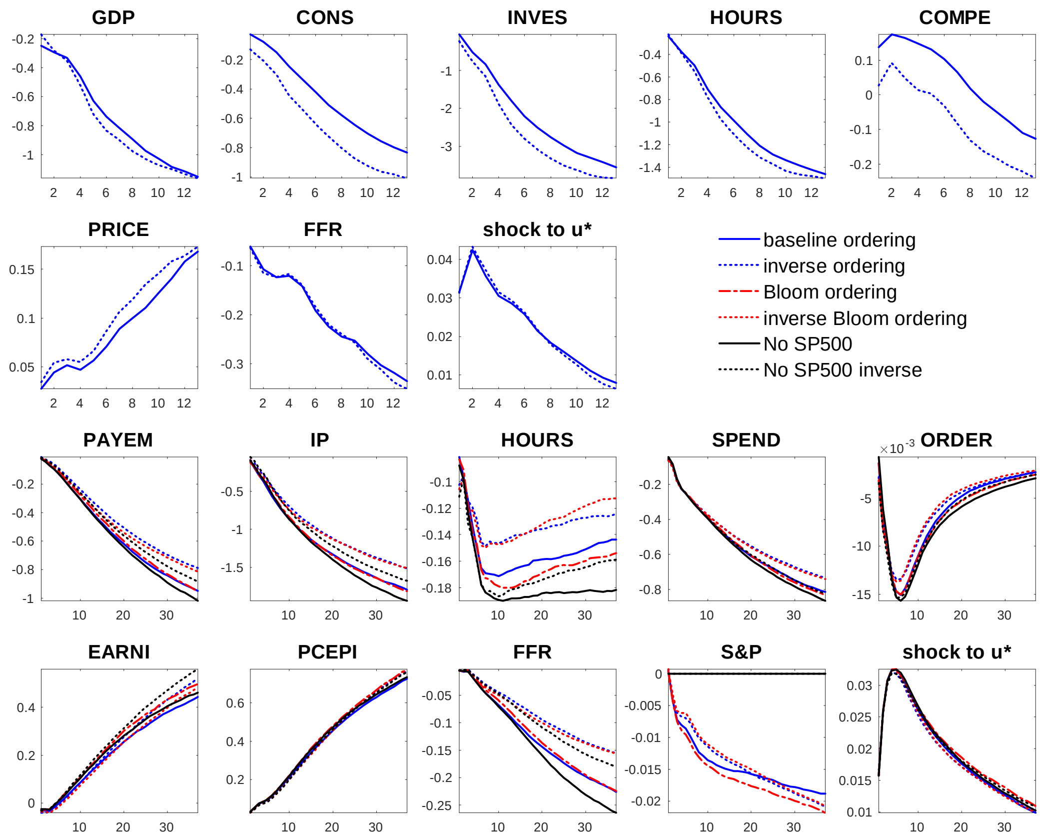

Figure 7: IRFs with alternative orderings within the macro block. Shock to uncertainty. For the monthly dataset, the charts also show results based on Bloom's variable ordering, LMN financial uncertainty measure, and a model with the S&P500 removed from the VAR.

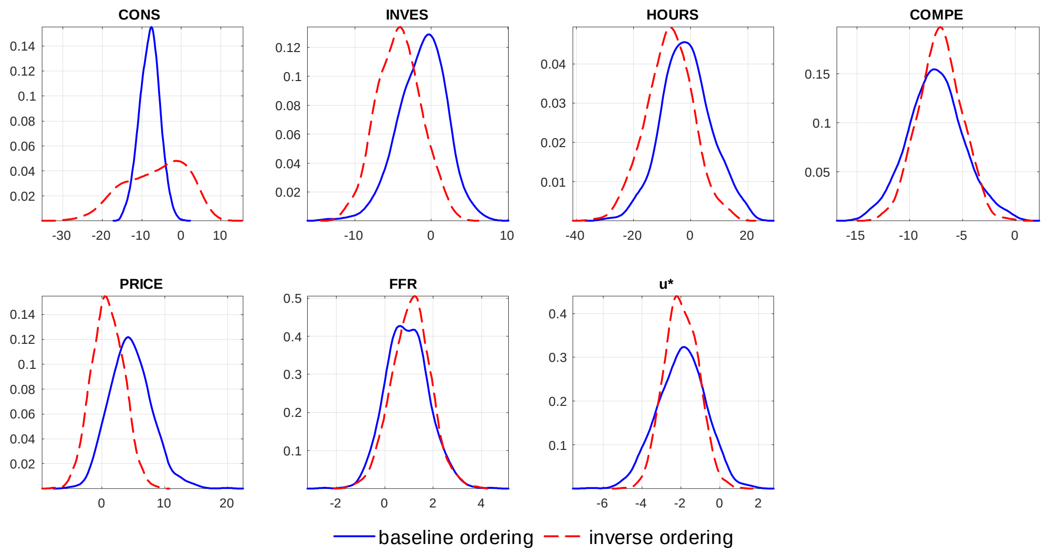

Figure 8: Alternative orderings within the macro block. Quarterly dataset. Coefficients ϕ.

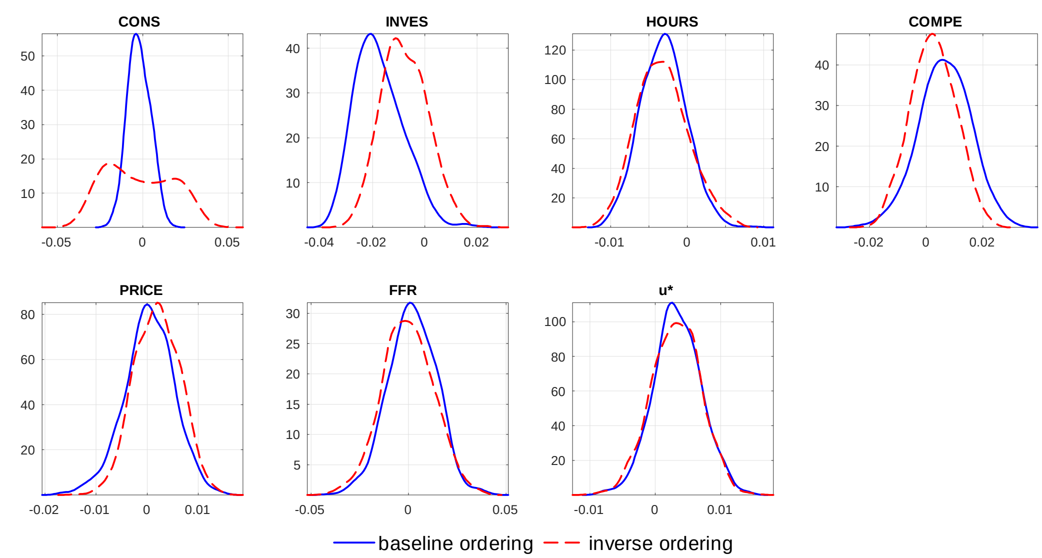

Figure 9: Alternative orderings within the macro block. Quarterly dataset. Coefficients ψ.

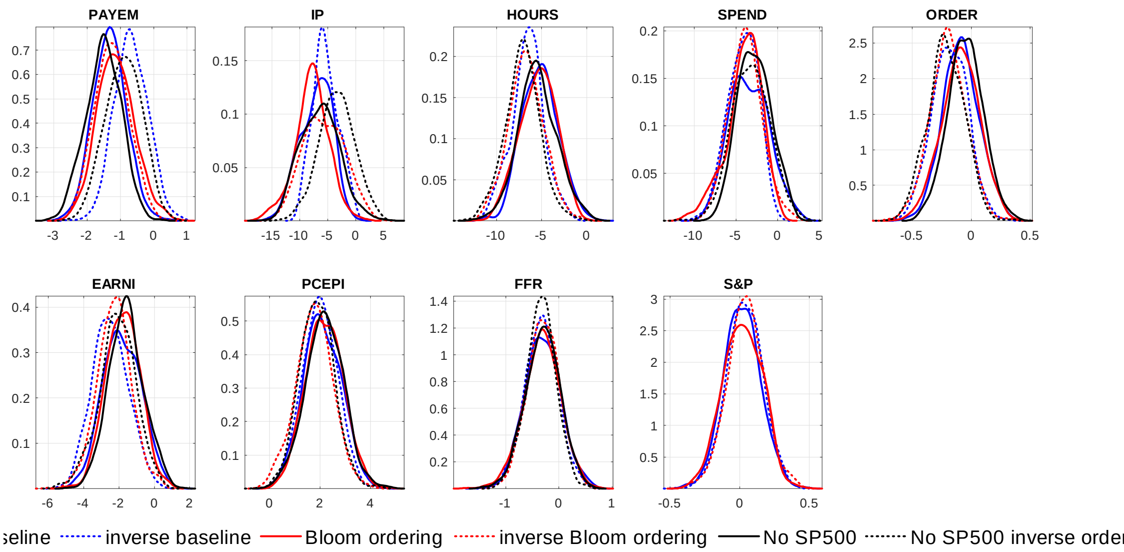

Figure 10: Alternative orderings within the macro block. Monthly dataset. Coefficients ϕ.

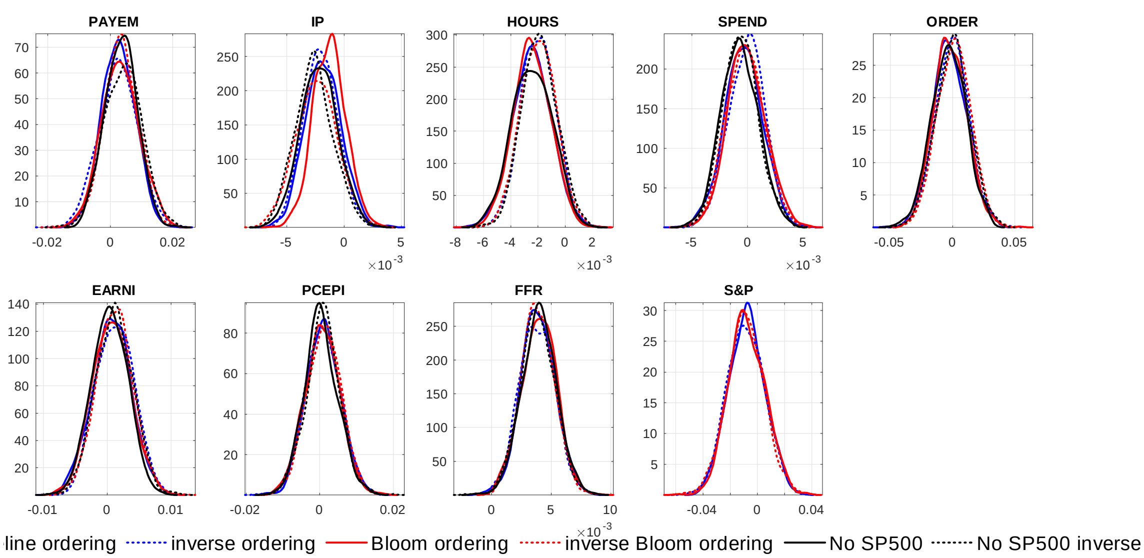

Figure 11: Alternative orderings within the macro block. Monthly dataset. Coefficients ψ.

4 Full distributions of the IRFs

This section contains the posterior distributions of the impulse response functions at selected forecast horizons (h=0, 1, 4, and 12 quarters ahead for the quarterly model and h=0,3,12,36 for the monthly model), for the following cases:

- Figure 12*:* Spec.1, quarterly data, effect of imposing $\psi = 0$
- Figure 13*:* Spec.1, umonthly data, effect of imposing $\psi = 0$
- Figure 14*:* Spec.1, quarterly data, effect of imposing $\phi = 0$
- Figure 15*:* Spec.1, monthly data, effect of imposing $\phi = 0$
- Figure 16*:* Spec.2, quarterly data, effect of imposing $\psi = 0$
- Figure 17*:* Spec.2, monthly data, effect of imposing $\psi = 0$
- Figure 18*:* Spec.2, quarterly data, effect of imposing $\phi = 0$
- Figure 19*:* Spec.2, monthly data, effect of imposing $\phi = 0$

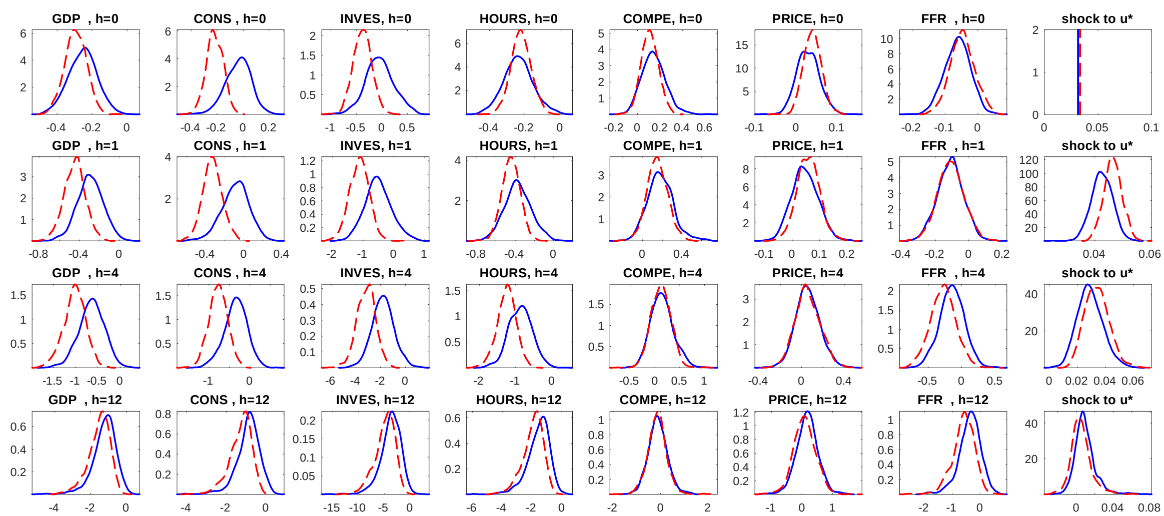

Figure 12: Uncertainty shock, quarterly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\psi \neq 0$, red dashed line denotes the case $\psi = 0$.

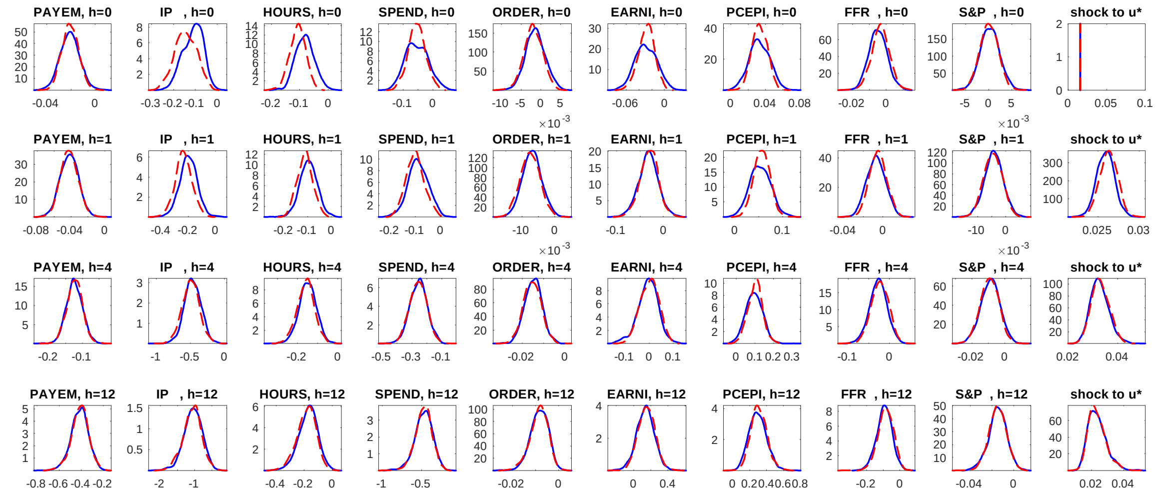

Figure 13: Uncertainty shock, monthly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\psi \neq 0$, red dashed line denotes the case $\psi = 0$.

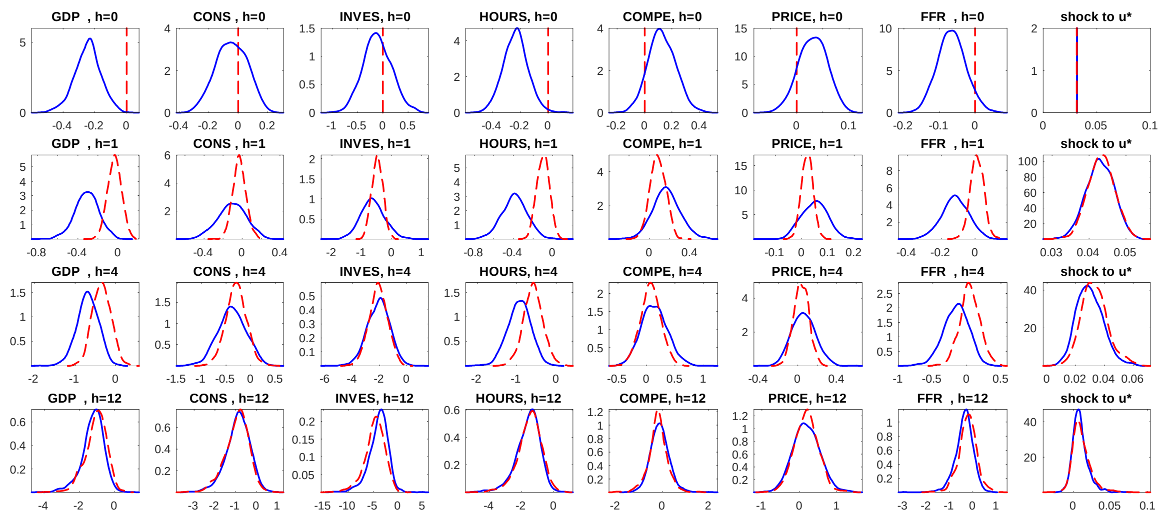

Figure 14: Uncertainty shock, quarterly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\phi \neq 0$, red dashed line denotes the case $\phi = 0$.

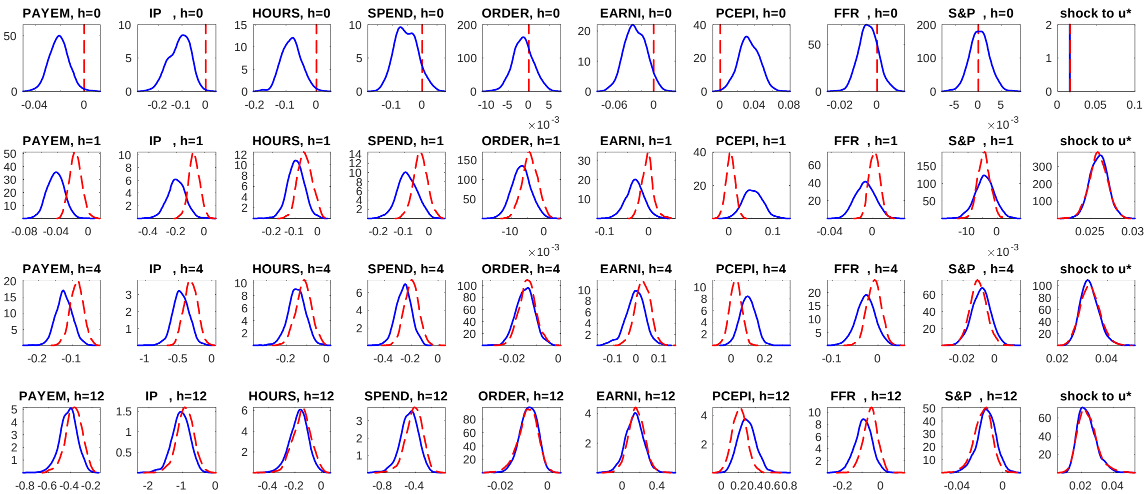

Figure 15: Uncertainty shock, monthly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\phi \neq 0$, red dashed line denotes the case $\phi = 0$.

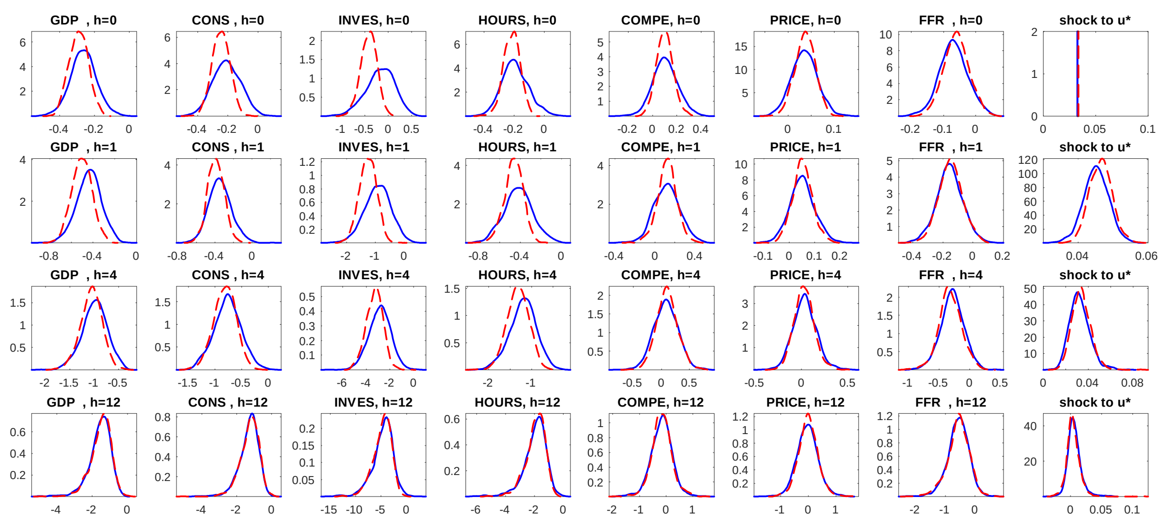

Figure 16: Specification 1 ($\beta = 0$). Specification 1 ($\beta = 0$). Uncertainty shock, quarterly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\psi \neq 0$, red dashed line denotes the case $\psi = 0$.

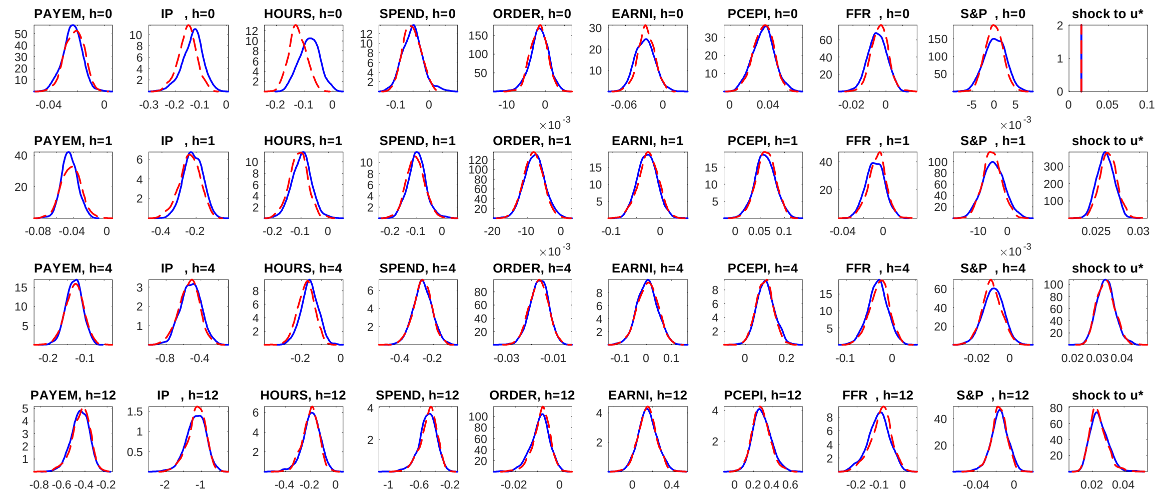

Figure 17: Specification 1 ($\beta = 0$). Uncertainty shock, monthly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\psi \neq 0$, red dashed line denotes the case $\psi = 0$.

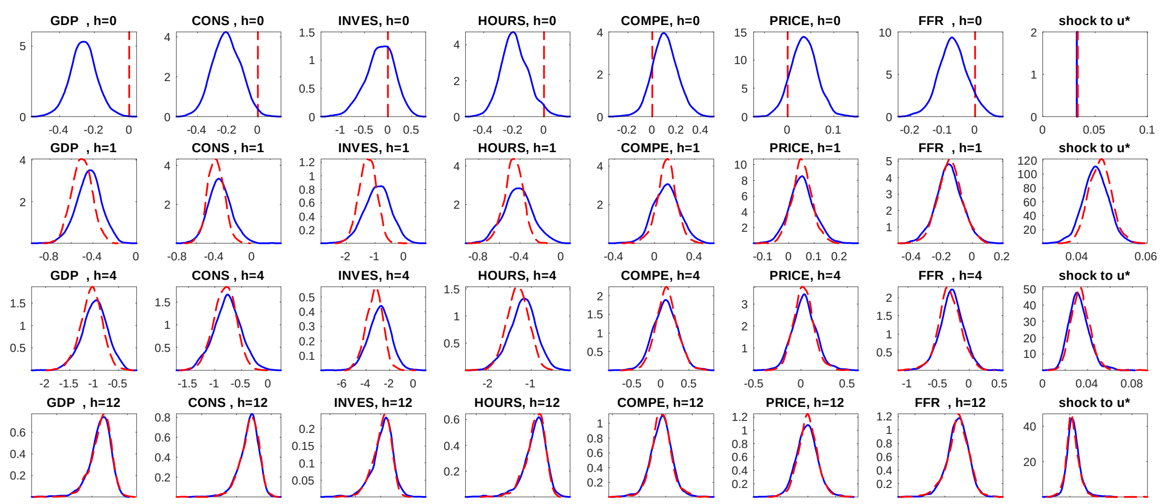

Figure 18: Specification 1 ($\beta = 0$). Uncertainty shock, quarterly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\phi \neq 0$, red dashed line denotes the case $\phi = 0$.

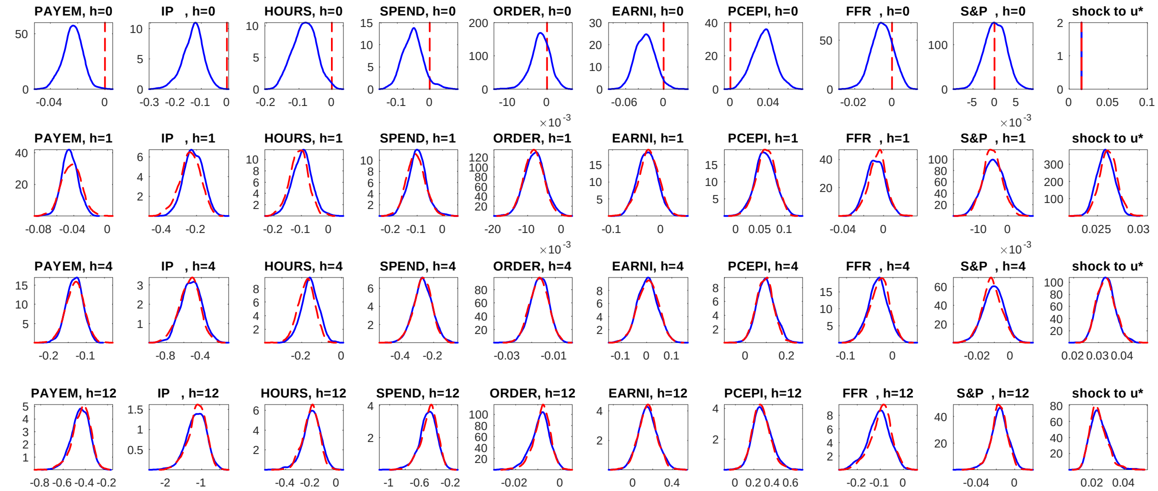

Figure 19: Specification 1 ($\beta = 0$). Uncertainty shock, monthly data. Posterior distributions of impulse responses at selected horizons. Blue solid line denotes the case $\phi \neq 0$, red dashed line denotes the case $\phi = 0$.

5 Specification tests

This section contains diagnostic tests for the Gaussianity of the (posterior means of the) innovations contained in ε_t^*. Results are organized in the following figures:

- Figure 20: Empirical distribution, Specification 2 ($\beta \neq 0$)
- Figure 21: Empirical distribution, Specification 1 ($\beta = 0$)

The 10% and 5% critical values for the Jarque-Bera tests are, respectively, 4.605 and 5.991.

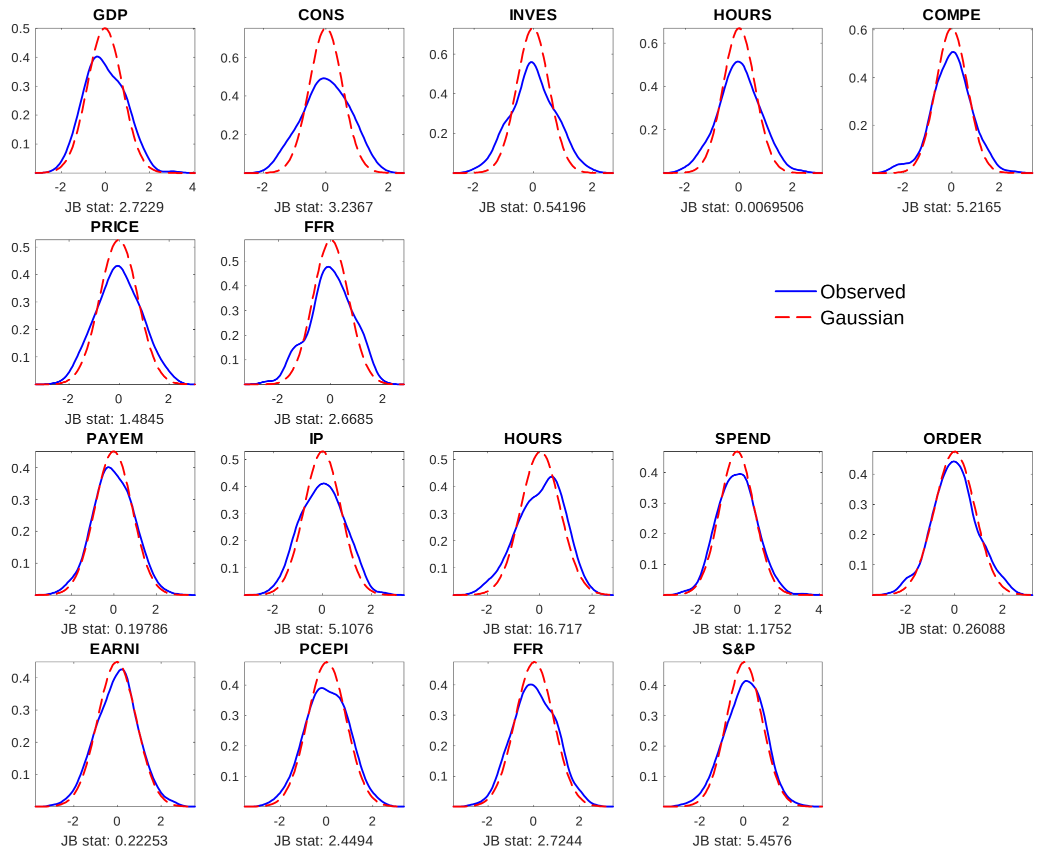

Figure 20: Normality tests, Specification 2 ($\beta \neq 0$).

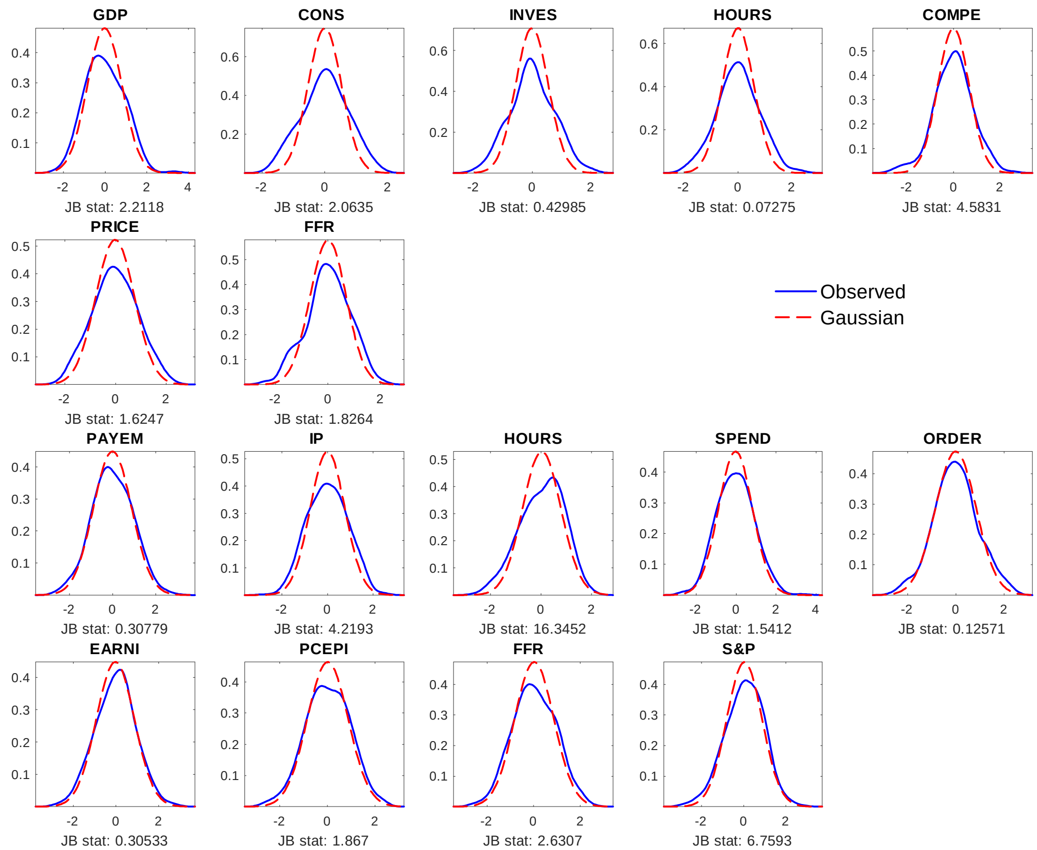

Figure 21: Normality tests, Specification 1 ($\beta = 0$).

6 Full results of Monte Carlo experiments

- Figure 22: Design 1. DGP from univariate application, model estimated with $\psi \neq 0$ and $\psi = 0$.
- Figure 23: Design 2. DGP from univariate applition, but with $\psi = 0$. Model estimated with $\psi \neq 0$ and $\psi = 0$
- Figure 24: Design 3. DGP from univariate application, model estimated with $\phi \neq 0$ and $\phi = 0$
- Figure 25: Design 4. DGP for the volatility is either σ_η^2 or $\sigma_\eta^2/100$, where σ_η^2 is the value found in the univariate application. Model estimated with ϕ and ψ unrestricted.
- Figure 26: Design 5. DGP from univariate application, and with shocks to y_t and m_t distributed either as Gaussian or as student t. The estimated model assumes Gaussian shocks.

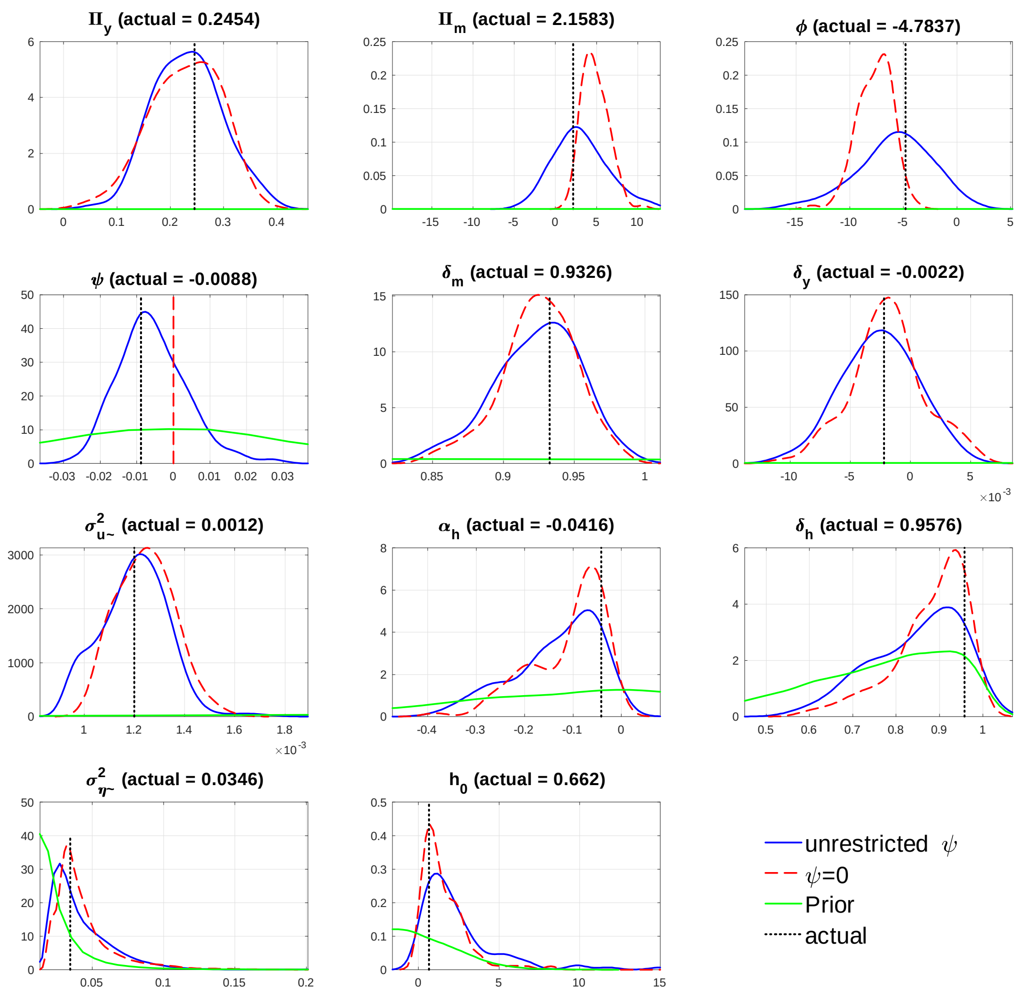

Figure 22: Design 1. DGP from univariate application, model estimated with $\psi \neq 0$ and $\psi = 0$.

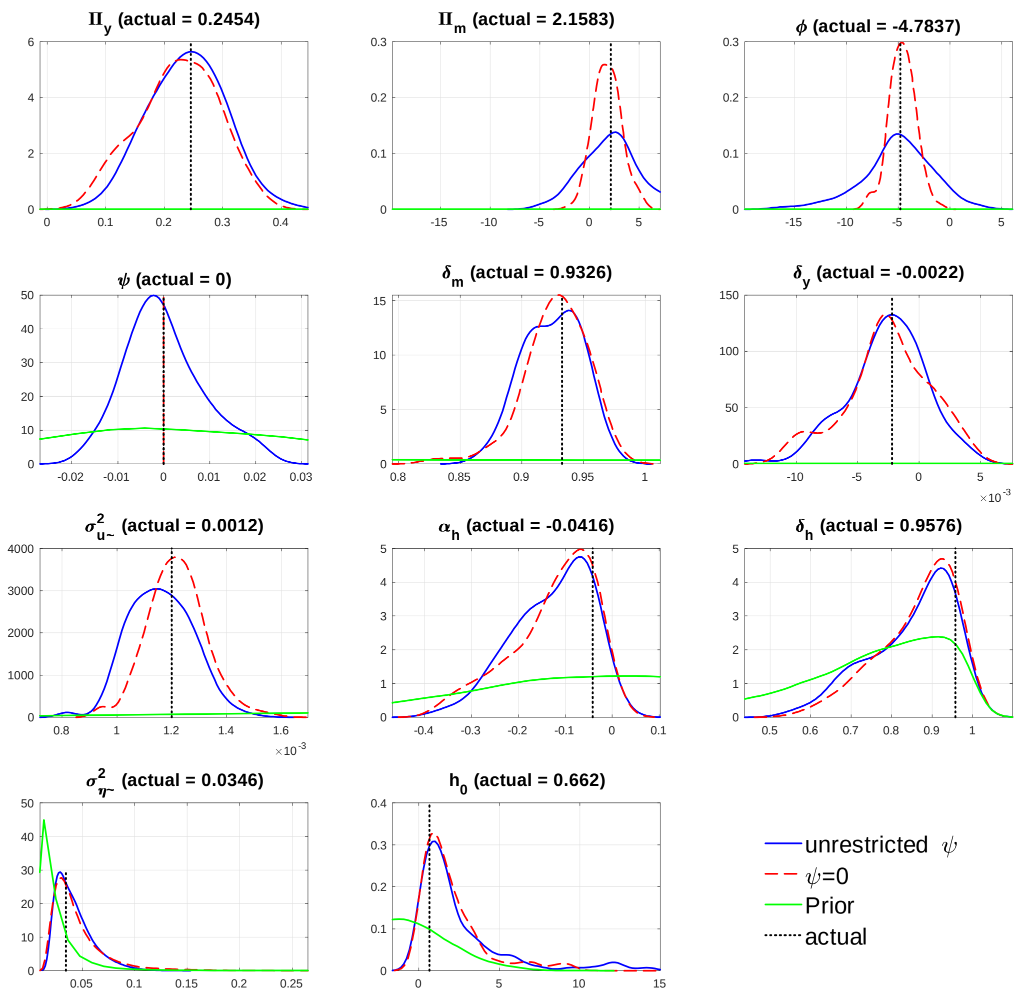

Figure 23: Design 2. DGP from univariate applition, but with $\psi = 0$. Model estimated with $\psi \neq 0$ and $\psi = 0$

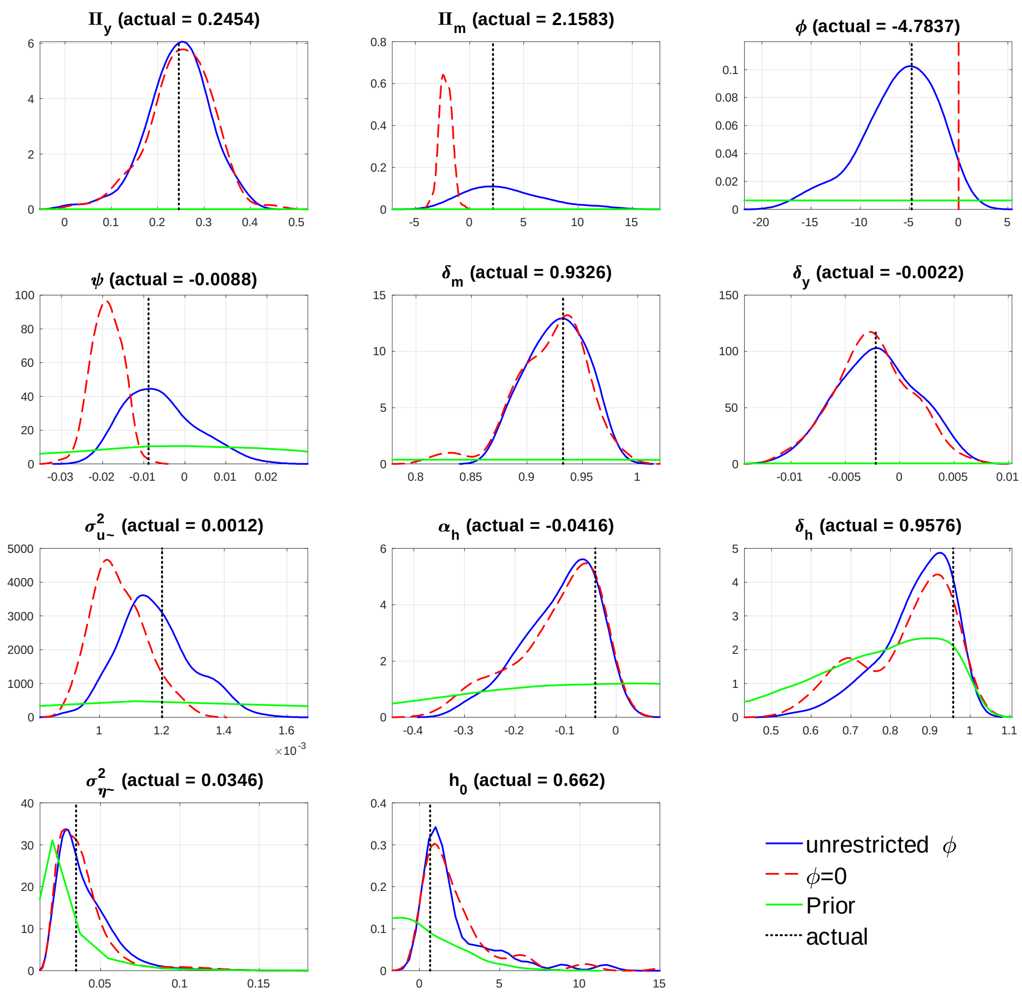

Figure 24: Design 3. DGP from univariate application, model estimated with $\phi \neq 0$ and $\phi = 0$

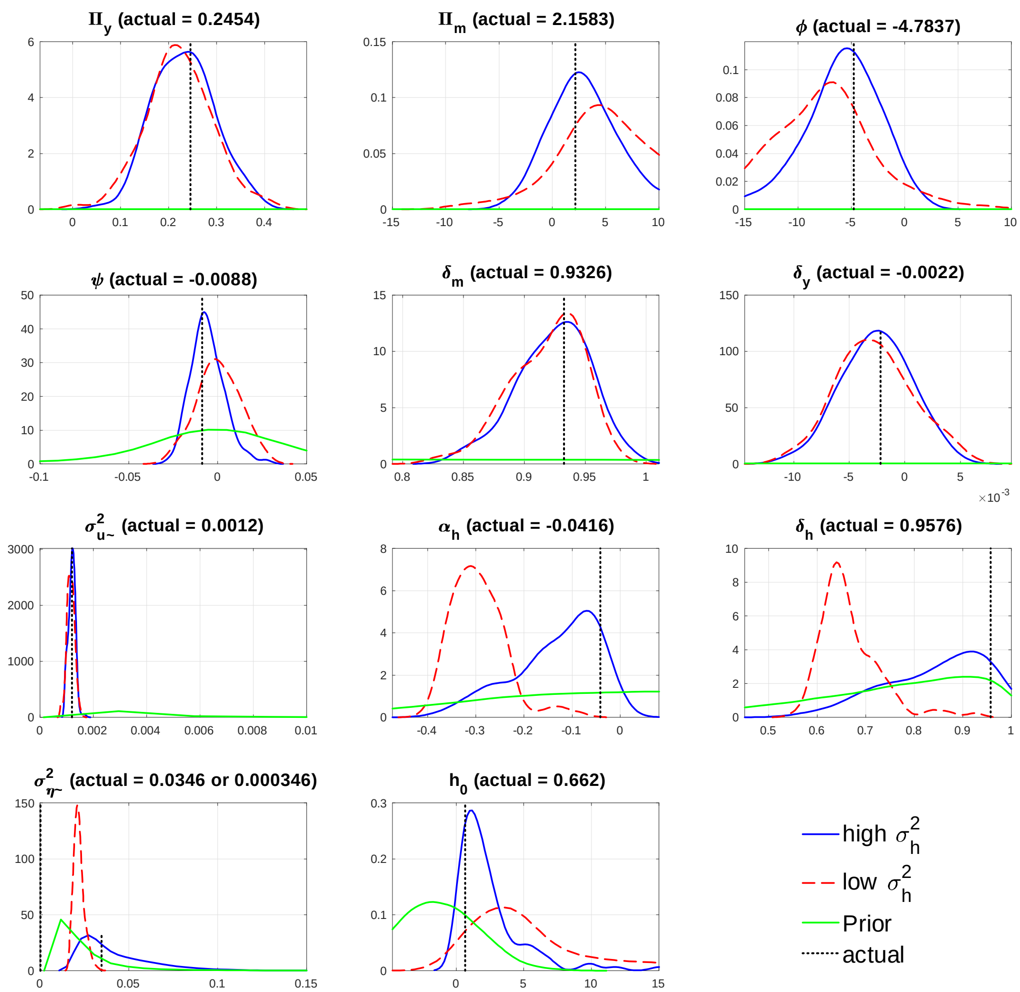

Figure 25: Design 4. DGP for the volatility is either σ^2_η or $\sigma^2_\eta/100$, where σ^2_η is the value found in the univariate application. Model estimated with ϕ and ψ unrestricted.

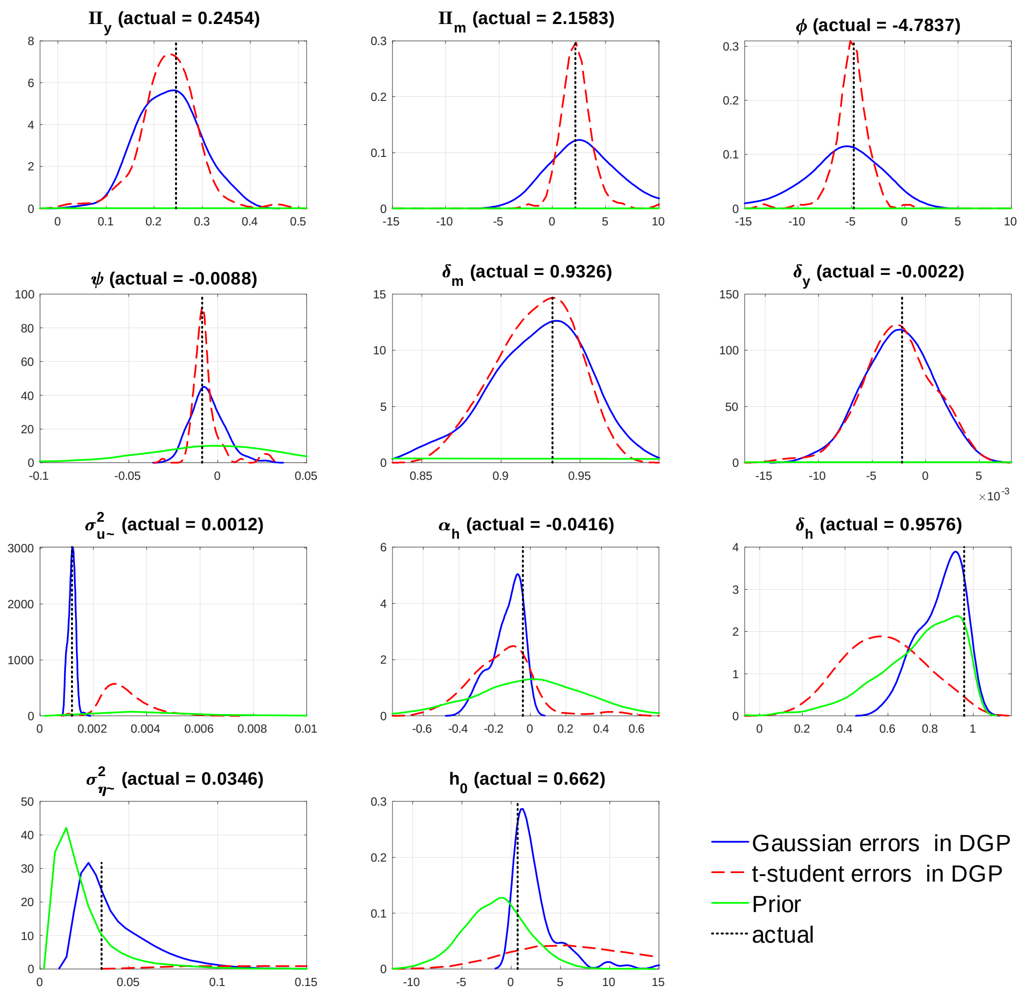

Figure 26: Design 5. DGP from univariate application, and with shocks to y_t and m_t distributed either as Gaussian or as student t. The estimated model assumes Gaussian shocks.