Quantitative Economics 14 (2023), 1163–1198 1759-7331/20231163

The importance of supply and demand for oil prices: Evidence from non-Gaussianity

Robin Braun
Board of Governors of the Federal Reserve System

When quantifying the importance of supply and demand for oil price fluctuations, a wide range of estimates have been reported. Models identified via a sharp upper bound on the short-run price elasticity of supply find supply shocks to be minor drivers. In turn, when replacing the upper bound with a weakly informative prior, supply shocks turn out to be substantially more important. In this paper, I revisit the evidence in a model that combines weakly informative priors with identification by non-Gaussianity. For this purpose, a SVAR is developed where the unknown distributions of the structural shocks are modeled nonparametrically. The empirical findings suggest that once identification by non-Gaussianity is incorporated into the model, posterior mass of the short-run oil supply elasticity shifts toward zero and oil supply shocks become minor drivers of oil prices. In terms of contributions to the forecast error variance of oil prices, the model arrives at median estimates of just 6% over a 16-month horizon.

Keywords. Oil market, Structural Vector Autoregression (SVAR), identification by non-Gaussianity, nonparametric Bayes.

JEL classification. C32, Q43.

1. Introduction

Since Kilian (2009), an increasing number of papers have studied the distinct role of supply and demand shocks in driving oil price fluctuations. When quantifying their relative importance, a wide range of estimates have been reported. On the one hand, Kilian (2009, 2022a), Kilian and Murphy (2012, 2014), Juvenal and Petrella (2015), Antolín-Díaz and Rubio-Ramírez (2018), Zhou (2020), and Cross, Nguyen, and Tran (2022) report that oil prices are mainly demand driven and that supply shocks are not important. Their estimates typically attribute less than 10% of the long-term variability in oil prices to supply. On the other hand, recent papers of Baumeister and Hamilton (2019) and Caldara, Cavallo, and Iacoviello (2019) point toward a substantially larger role of supply, estimating variance contributions of up to 37%. Most of the disagreement can be attributed

Robin Braun: robin.a.braun@frb.gov

I thank two anonymous referees who helped to improve this paper substantially. I also thank Knut Aastveit, Christiane Baumeister, Dominik Bertsche, Dario Bonciani, Ralf Brüggemann, Jamie Cross, Gabriele Fiorentini, Luca Gambetti, Sascha Keweloh, Lutz Kilian, Massimiliano Marcellino, Michele Piffer, Mikkel Plagborg Møller, and Giovanni Ricco for useful comments. Part of this work was conducted with financial support by the German Science Foundation (Grant BR 2941/3-1) and the Graduate School of Decision Sciences at the University of Konstanz, which is gratefully acknowledged. The views expressed in this paper are my own and do not reflect those of the Board of Governors of the Federal Reserve System.

 https://doi.org/10.3982/QE2091

to differences in the identification strategy (Herrera and Rangaraju (2020)):[1] imposing a very small upper bounds on the short-run price elasticity of supply yield negligible effects of supply shocks, and vice versa.

In this paper, I revisit the evidence based on a novel identification strategy. The idea is to combine economically motivated prior distributions as in Baumeister and Hamilton (2019) with identification by non-Gaussianity. The latter is based on the assumption that structural shocks are mutually independent and display some degree of non-Gaussianity. As documented in this paper, large deviations from Gaussianity characterize many oil market shocks, and their mutual independence is not an unreasonable assumption. My findings indicate that once non-Gaussianity is exploited, the posterior distribution of the short-run oil-supply concentrates near zero and supply shocks are found to be minor drivers of oil prices.

To build up intuition of the identification strategy, consider a stylized bivariate model for supply and demand:

$$\begin{array}{ll} \text{supply:} & q_t = \alpha p_t + \sigma_1 \varepsilon_t^s \\ \text{demand:} & q_t = \beta p_t + \sigma_2 \varepsilon_t^d \end{array} \begin{pmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{pmatrix} \sim (0, I_2),$$

where q_t and p_t are changes in (log) quantity and (log) prices, α and β are the price elasticities of supply and demand, and $\sigma_{1/2}$ the standard deviations of the supply (ε_t^s) and demand shocks (ε_t^d). The model is not identified from the second moment of the data, as there are four structural parameters but only three reduced- form covariance parameters. The literature typically proceeds imposing identifying restrictions, which reflect priors on the sign and magnitude of the structural parameters. Besides conventional sign restrictions on the slope of the demand ($\beta < 0$) and supply curve ($\alpha > 0$), magnitude restrictions are common to achieve more informative results. For example, Kilian and Murphy (2014) assume a very inelastic supply imposing a tight upper bound on $\alpha \in (0, 0.025)$, while Baumeister and Hamilton (2019) allow for substantially larger values via a truncated student-t prior $\alpha \sim t_{0,\infty}(0.1, 0.2, 3)$ centered at 0.1 and with a scale of 0.2 and 3 degrees-of-freedom. Unfortunately, small differences in these priors have substantial implications for estimates of the relative importance of supply and demand shocks. The reason is that forecast errors in oil production and prices are fairly uncorrelated, yielding a very large set of sign-restricted models equally consistent with the data.

In this paper, I leverage non-Gaussianity as additional source of identifying information. Figure 1 shows a scatterplot of forecast errors from a bivariate VAR for global oil prices and oil production. When the joint distribution is characterized by the second moment ("Gaussian Setting"), many different models are observationally equivalent. Consider two arbitrarily chosen supply and demand schedules (A and B), which yield the same reduced form but imply very different structural dynamics. In model A, supply is inelastic and demand is elastic. Consequently, in such a model oil production would be mainly driven by supply shocks while oil prices would be largely caused by

[1] As noted in Aastveit et al. (2021), the disagreement is much less pronounced once a shorter sample is used for estimation, excluding the large oil price shocks of the 1970s.

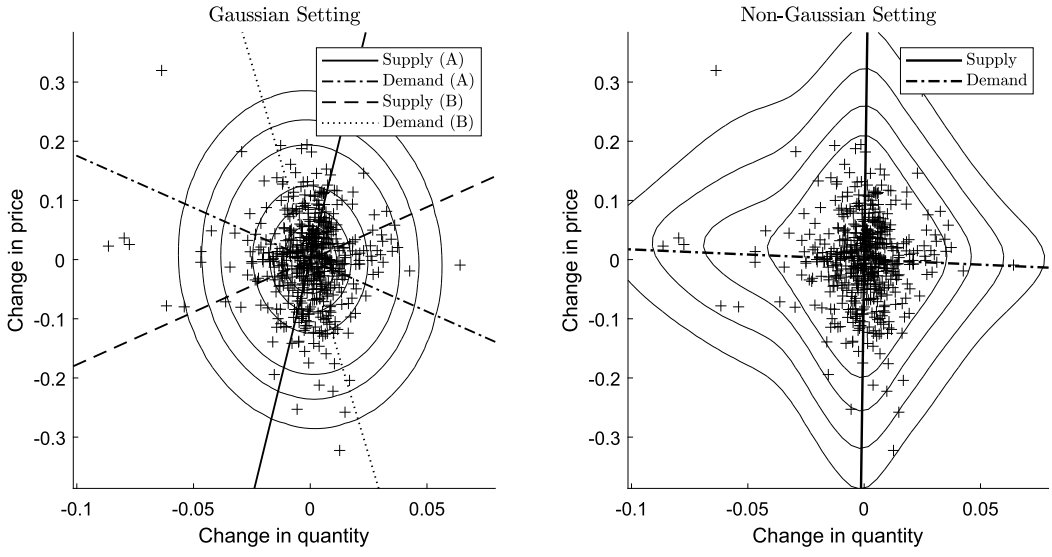

FIGURE 1. Identifying oil demand and supply curves by non-Gaussianity.

demand shocks. In turn, for model B, supply is more elastic and demand is inelastic, implying the exact opposite. Incorporating prior knowledge on elasticities ultimately boils down to picking a range of models from the set of observationally equivalent models, shaping the answer about oil price drivers a priori. While narrative evidence and external estimates have been used to inform the priors (see, e.g., Newell and Prest (2019), Caldara, Cavallo, and Iacoviello (2019), Bjørnland, Nordvik, and Rohrer (2021)), a fairly wide range of estimates suggest that the evidence is not conclusive (Kilian (2022b)).

In the right panel of Figure 1 ("Non-Gaussian Setting"), I illustrate how the joint distribution of the reduced-form errors can help discriminate among observationally equivalent models. The solid lines correspond to contour lines of the estimated joint density implied by the model developed in this paper. The estimator yields a unique supply and demand schedule consistent with the data, rotating the curves such that the forecast errors cluster near the supply and demand schedule, in line with heavy-tailed structural shocks. Hereby, the non-Gaussian shape makes certain shifts of the supply and demand curve more likely than others; this way working as a probabilistic instrument (see also Rigobon (2003) for a similar interpretation for identification by heteroskedasticity). For the data considered in the scatterplot, the proposed identification approach points toward a steep supply and flat demand curve.

Identification by non-Gaussianity yields a set of independent shocks, which per se are not useful for economic analysis. In this paper, I suggest a Bayesian approach where non-Gaussianity is only used in combination with economically meaningful restrictions. Specifically, I incorporate non-Gaussianity into the state-of-the art oil market model of Baumeister and Hamilton (2019) (BH19 henceforth) who identify four structural shocks based on a mix of prior distributions and sign restrictions. There are various benefits from adopting such a combined identification approach. First, the economic structure allows to interpret the shocks at all stages of the analysis, and thereby yields

an unique, order-invariant system of equations. Second, identification of the model remains guaranteed by the economic restrictions, even if there is little identifying information from non-Gaussianity in a given finite sample, or complete absence of non-Gaussianity.

The empirical analysis suggest that at least two out of four structural shocks in BH19 display strong deviations from Gaussianity. Tests for mutual independence suggest that it is reasonable to use this information for identification purposes. I find that non-Gaussianity shifts posterior mass of the supply elasticity toward zero, while demand becomes more elastic. Furthermore, oil supply shocks are found to be less important drivers of oil prices. In terms of forecast error variance decompositions, posterior median estimates suggest a share of 6% for supply shocks, as opposed to 32% obtained by BH19. These findings are supportive of earlier papers in the literature obtained under a strong upper bound on the supply elasticity.

Besides offering new evidence on the drivers of oil prices, this paper develops a new Bayesian semiparametric model where the marginal distributions of each structural shock is left unspecified. This is implemented via the use of univariate Dirichlet process mixture models (DPMM) (Escobar and West (1995)). Much like kernel density estimators, those are the standard model in Bayesian non-parametric statistics to model unknown density functions.[2]

There are various benefits from adopting a semiparametric approach. First of all, the model offers robustness to misspecification of the error term. This is particularly important in non-Gaussian SVAR models. As highlighted in Fiorentini and Sentana (2022), specifying the wrong marginal distribution risks inconsistent estimation of shocks standard deviations, thereby invalidating inference on forecast error variance decompositions. Besides robustness, the estimators flexibility to adapt to the unknown shock distribution can also offer efficiency gains. As illustrated in the empirical application, posterior confidence in the oil market model are more narrow under the semiparametric model than under a parametric alternative that relies on student-t distributions. A further point in favor of the DPPM model is that assessing the amount of non-Gaussianity in the data is a straightforward task. A simple comparison of the posterior predictive density with the kernel of a standard normal gives an indication of how much identifying information one can expect from the statistical properties of each shock.

To conduct inference, I develop a novel MCMC algorithm that iteratively draws from the conditional distributions of the VAR parameters and those of the DPMMs. While most conditionals are well known and straightforward to draw from, the challenging part of the algorithm is drawing the matrix A, which relates VAR forecast errors u_t to structural shocks ε_t via $Au_t = \varepsilon_t$. Here, I make use of the algorithm proposed originally in Waggoner and Zha (2003), already generalized for various settings (Villani (2009), Chan, Koop, and Yu (2023)). I show how the algorithm can be adopted to allow for nonzero normalizing constraints on A. This facilitates prior elicitation on elements in A, as structural

[2]For an earlier use of Dirichlet process priors for nonparametric modeling of error distributions in economics; see, for example, Hirano (2002).

parameters can be separated from the scale of structural shocks. Besides posterior inference, I also discuss evaluation of the marginal likelihood providing a tool to test overidentifying restrictions. Here, I rely on the cross-entropy method of Chan and Eisenstat (2015), evaluating the likelihood via a collapsed sequential importance sampling algorithm (Basu and Chib (2003)).

The methodological part of this paper relates to the literature in various ways. First, the use of nonparametric density estimators for identification of non-Gaussian SVARs is new to the literature.[3] Previous methods either required specifying a functional form for the error distribution (Lanne, Meitz, and Saikkonen (2017), Lanne and Luoto (2020), Anttonen, Lanne, and Luoto (2021)) or a selection of suitable moments or criteria function (Lanne and Luoto (2021), Herwartz (2018)). Large-sample arguments for pseudolikelihood inference are used to discuss robustness to model mis-specification (Gourieroux, Monfort, and Renne (2017), Fiorentini and Sentana (2022)), but without the focus on providing an efficient procedure that adapts flexibly to the error distribution.

Second, the combined identification approach offers some conceptual benefits compared to the currently prevailing empirical strategy for non-Gaussian SVARs. Typically, the model is first estimated based on non-Gaussianity before a labeling exercise follows with the goal to attach an economic meaning to the shocks (Lanne, Meitz, and Saikkonen (2017)). Unfortunately, this may fail in practice, as there may be no combination of shocks that satisfies the researchers economic priors. In contrast, the proposed Bayesian framework allows to express all the economic identifying information via the prior distribution. The analysis leads to a joint posterior, which combines all the information at hand, be it from the prior or the likelihood.

Finally, the paper relates to recent work by Drautzburg and Wright (2023) who propose to exploit non-Gaussianity in order to narrow down the set of admissible models in a sign-restricted SVAR model. This idea is very similar to the combined identification scheme proposed in this paper. However, I adopt a fully Bayesian approach to inference while their analysis leverages frequentist methods. Furthermore, the integrated use of DPMMs exploits the nonparametric density estimator at all stages of inference. This contrasts their two step estimation, which exploits non-Gaussianity only at the stage of identifying suitable rotations of orthogonalized shocks.

The paper is structured as follows. In Section 2, the methodology is covered for a non-Gaussian SVAR model where structural shocks follow Dirichlet process mixture models (DPMM). Section 3 proceeds applying the method revisiting the importance of supply and demand shocks for oil price fluctuations. Section 4 concludes. The Online Supplemental Material (Braun (2023)) contains the Appendix to this article, providing technical details, illustrations, and additional empirical results.

2. Methodology

In the following, I introduce the semiparametric non-Gaussian SVAR model. I start with a quick review of the identification problem and standard results that arise under inde-

[3] In the Independent Component Analysis (ICA) literature, kernel density estimators have been exploited by Boscolo, Pan, and Roychowdhury (2004). However, they do not discuss inference for their estimator, and hence, their method is not directly useful for SVAR analysis.

pendent and non-Gaussian shocks (Section 2.1). Section 2.2 proceeds with an instructive description of the nonparametric methods used to model the marginal distributions of each shock. The multivariate model is outlined in Section 2.3 and Bayesian inference is covered thereafter (Section 2.4). Finally, I complement the methodology with a marginal likelihood estimator that some researchers may find helpful in order to conduct formal model comparison (Section 2.5).

2.1 *Non-Gaussian SVARs*

Consider the following SVAR(p) specification for a K-dimensional time series vector y_t:

$$3y_t = c + \sum_{j=1}^{p} A_j y_{t-j} + u_t, \qquad u_t \sim (0, \Sigma_u), \tag{1}$$

$$\mathrm{A}u_t = \varepsilon_t, \qquad \varepsilon_t \sim (0, \Sigma_\varepsilon), \tag{2}$$

where Σ_u is a full covariance matrix and Σ_ε is diagonal. Motivated by the empirical application, this paper considers an A type of model in the terminology of Lütkepohl (2005), meaning that orthogonal structural shocks (ε_t) are modelled as a linear function of reduced-form errors (u_t). The reduced-form covariance matrix of the VAR forecast errors is linked to the structural parameters by $\Sigma_u = \mathrm{A}^{-1}\Sigma_\varepsilon(\mathrm{A}^{-1})'$. Throughout the paper, stationarity is assumed, that is,

$$\det A(z) = \det\bigl(I_K - A_1 z - \cdots - A_p z^p\bigr) \neq 0 \quad \text{for } |z| \leq 1.$$

It follows that the SVAR(p) has a MA(∞) representation given by $y_t = \mu_y + \sum_{j=1}^{\infty} \Theta_j \varepsilon_{t-j}$ where $\Theta_j = \Phi_j \mathrm{A}^{-1}$, $\Phi_0 = I_K$, $\Phi_j = \sum_{i=1}^{j} \Phi_{j-i} A_i$ for $j \in \mathbb{N}$ with $A_i = 0$ for $i > p$. The ikth entry of matrix Θ_j contains the impulse response, capturing the dynamic effect of structural shock k on the ith variable in y_t, j periods after the shock.

Without additional assumptions, the covariance structure of the forecast errors jointly identifies A and Σ_ε only up to orthogonal rotations. To see this, consider the alternative model $\tilde{\mathrm{A}} = Q'\Sigma_\varepsilon^{-1/2}\mathrm{A}$ and $\tilde{\Sigma}_\varepsilon = I_K$ for any orthogonal matrix Q satisfying $QQ' = I_K$ and $Q^{-1} = Q'$. The implied covariance matrix is equivalent to the original model given that $\tilde{\mathrm{A}}^{-1}\tilde{\Sigma}_\varepsilon(\tilde{\mathrm{A}}^{-1})' = \mathrm{A}^{-1}\Sigma_\varepsilon^{1/2}QQ'\Sigma_\varepsilon^{1/2}(\mathrm{A}^{-1})' = \mathrm{A}^{-1}\Sigma_\varepsilon(\mathrm{A}^{-1})'$. SVAR analysis proceeds by imposing additional restrictions to solve this identification problem and a comprehensive review of different strategies is given in Kilian and Lütkepohl (2017).

In this paper, I will exploit identification by non-Gaussianity. This entails imposing the following distributional assumptions on the structural shocks $\varepsilon_t = [\varepsilon_{1t}, \ldots, \varepsilon_{Kt}]'$:

(i) ε_t is a strictly stationary random vector with $\mathrm{E}[\varepsilon_{it}] = 0$ and $\mathrm{E}[\varepsilon_{it}^2] < \infty$ for $i = 1, \ldots, K$.

(ii) The structural shocks are mutually independent, and at least $K - 1$ have non-Gaussian marginal distributions.

(iii) Each component of ε_t is serially uncorrelated, that is, $\mathrm{Cov}(\varepsilon_{it}, \varepsilon_{i,t+k}) = 0, \forall k \neq 0, i = 1, \ldots, K$.

As established in Lanne, Meitz, and Saikkonen (2017), assumptions (i)–(iii) identify the SVAR model up to permutation, sign and scale. In other words, the set of orthogonal rotation matrices yielding observationally equivalent models reduces to $Q = PD$, where P is a K-dimensional permutation matrix and D a diagonal matrix with elements ± 1.

At this point, it is useful to scrutinize the identifying assumptions. First, identification by non-Gaussianity requires mutual independence of structural shocks. It is important to acknowledge that this is an equally restrictive assumption than conventional SVAR restrictions such as zero or sign restrictions. This is because it rules out higher-order dependence in structural shocks, which can arise if the underlying data-generating process is subject to certain nonlinearities.

An important example is the presence of common stochastic volatility dynamics in the second moment of shocks (see Olea, Plagborg-Møller, and Qian (2022)). Specifically, consider the bivariate SVAR(0) model $\mathrm{A}u_t = \varepsilon_t$, where further it holds that

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} = \sigma_t \begin{bmatrix} \tilde{\varepsilon}_{1t} \\ \tilde{\varepsilon}_{2t} \end{bmatrix},$$
$$\log \sigma_t = \phi^{\sigma} \log \sigma_{t-1} + \varepsilon_t^{\sigma}.$$

Here, $\tilde{\varepsilon}_t$ is independent white noise and σ_t is an AR(1) scalar stochastic volatility factor which is driven by a third structural shock ε_t^{σ}. In this example, ε_t are orthogonal white noise, and the presence of σ_t implies non-Gaussian marginals. However, the elements in ε_t are not mutual independent, invalidating identification by non-Gaussianity. This contrasts to conventional SVAR methods, which only require orthogonality of the shocks and, therefore, can still succeed to identify ε_t depending on the accuracy of the identifying restrictions.[4]

However, unlike conventional restrictions, it is possible to test for the empirical plausibility of the mutual independence assumption in each application. Recently, a series of tests have been developed for this purpose. These include the popular nonparametric test based on distance covariances (Matteson and Tsay (2017)), which is able to test for all forms of dependence between the structural shocks. As an alternative, Olea, Plagborg-Møller, and Qian (2022) propose testing correlations between squared structural shocks, with the goal to direct power against shocks that share a common volatility structure.[5] Unfortunately, these tests have been developed under a frequentist paradigm, where the distribution of the test statistics under the null hypothesis is simulated based on resampling schemes. Ultimately, this means that they are not directly applicable in the Bayesian framework of this paper. However, one may still study the posterior of these test statistics (and resampled versions thereof) in order to defend the empirical plausibility of the mutual independence assumptions.

[4]Other forms of nonlinearities may arise in the conditional mean of the model, thereby inducing higher-order dependence in forecast errors of a linear model. One salient example is the presence of the zero lower bound, which effectively censors the nominal interest rate. However, conventional linear VAR models would also fail to give consistent estimates in this setting; see Mavroeidis (2021).

[5]See also Davis and Ng (2023) and Amengual, Fiorentini, and Sentana (2022) for alternative tests of mutual independence.

Second, note that the identification conditions are compatible with common forms of heteroskedasticity observed in structural shocks induced, for example, by GARCH dynamics (Normandin and Phaneuf (2004), Lanne and Saikkonen (2007)) or stochastic volatility (Bertsche and Braun (2022)). As long as the volatility models are shock specific, mutual independence holds and non-Gaussianity can be applied for identification purposes.

Third, identification by non-Gaussianity yields a set of identified shocks that need to be combined with economic identifying information to be interpretable. In this paper, I propose to use economically motivated prior distributions as in Baumeister and Hamilton (2015) to ensure economic identification.

Finally, note that the invertibility assumption of the model can be relaxed. As pointed out in Gouriéroux, Monfort, and Renne (2020), non-Gaussianity can also aid identification under nonfundamentalness, for example, generated by dynamic stochastic general equilibrium models involving news shocks (Mertens and Ravn (2010)).

2.2 *Dirichlet process mixture models for structural shocks*

To model non-Gaussianity, I make use of nonparametric Dirichlet process mixture models (DPMM). Before introducing the full multivariate model, I start with a review of the univariate DPMM, which is used to model each shock's marginal distribution. Readers familiar with Bayesian nonparametrics may want to skip this part. For ease of exposition, I will drop the i index during this subsection, and reintroduce it for the multivariate model.

Each structural shock is assumed to be independent and follow the following hierarchical model:

$$\begin{aligned} \varepsilon_t|\theta_t &\sim F(\theta_t), \\ \theta_t &\sim G, \\ G &\sim \mathrm{DP}(G_0, \alpha), \end{aligned}$$

for $t = 1, \ldots, T$. Here, $F(\theta_t)$ is a probability distribution parametrized by θ_t with prior $\theta_t \sim G$. In a DPMM, G has the characteristic of being random itself, following a Dirichlet process (DP) $G \sim \mathrm{DP}(G_0, \alpha)$ (Ferguson (1973)). A Dirichlet process is uniquely characterized by a base distribution G_0 and a scalar concentration parameter $\alpha \in \mathbb{R}^+$. Realizations of a DP yield almost surely discrete priors for θ_t, which is why the model can be thought of a countably infinite mixture model. In order to facilitate understanding of the resulting model, I will review two instructive representations of the DPMM.

The first is known as the Pólya Urn representation and goes back to Blackwell and MacQueen (1973). The idea is to marginalize out G, yielding an intuitive and more direct representation of the prior implied for θ. In particular, for $t = 1, \ldots, T$, the distribution can be iteratively constructed as follows:

$$\theta_t|\theta_{t-1}, \ldots, \theta_1 \sim \frac{1}{t-1+\alpha}\sum_{j=1}^{t-1}\delta_{\theta_j} + \frac{\alpha}{t-1+\alpha}G_0,$$

$$\sim \sum_{j=1}^{k} \frac{n_j}{t-1+\alpha} \delta_{\theta_j^\star} + \frac{\alpha}{t-1+\alpha} G_0,$$

where $\delta_{(\cdot)}$ is the Dirac measure and $\{\theta_j^\star, j = 1, \ldots, k\}$ are the distinct values (clusters) of $\{\theta_j, j = 1, \ldots, t\}$, which have cluster size $n_j = \sum_{t=1}^{t-1} \mathbb{1}(\theta_t = \theta_j^\star)$. In words, the first line states that at any point of time t, θ_t may take either the value of a previously drawn parameter or be sampled from the base distribution G_0. The Pólya Urn scheme illustrates the main properties of the DPM prior of θ_t. First, the realizations are almost surely discrete. Second, there is a "richer get richer" property, which leads to heavy clustering of the mixing parameters θ. This is highlighted in the second line, stating that the probability of θ_t joining a certain cluster $\theta_j^\star$ increases in the cluster size n_j. Therefore, the model for ε_t can be interpreted as a flexible yet parsimonious mixture model where the number of components is random and increasing in the sample size. The strength of clustering is governed by the concentration parameter α and lower values are associated with fewer mixture components (clusters) for a given sample size. Finally, the choice of Base distribution G_0 will determine the location of the clusters.

A second convenient representation of the DPMM relates the model to finite mixture models some readers may be more familiar with. As outlined in Neal (2000), a direct link can be established by casting the following model with k mixture components:

$$\varepsilon_t | c_t, \theta^\star \sim F\left(\theta_{c_t}^\star\right), \tag{3}$$

$$c_t | p \sim \text{Discrete}(p_1, \ldots, p_k), \tag{4}$$

$$p \sim \text{Dirichlet}(\alpha/k, \ldots, \alpha/k), \tag{5}$$

$$\theta_j^\star \sim G_0, \quad j = 1, 2, \ldots, \tag{6}$$

where c_t is a discrete assignment variable linking each observation to one of the mixture components. Each component is associated with a unique parameter $\theta_j^\star$, which are drawn from the base distribution G_0. If the mixing proportions $p = (p_1, \ldots, p_k)$ are given a symmetric Dirichlet prior with concentration parameters α/k, a DPMM is obtained letting $k \to \infty$. Exploiting well-known properties of the Dirichlet multinomial distribution, the conditional probability of c_t given the sequence $\{c_{t-1}, \ldots, c_1\}$ can be shown to be (Neal (2000)):

$$P(c_t = c | c_{t-1}, \ldots, c_1) = P(c_{t-1}, \ldots, c_1, c_t = c)/P(c_{t-1}, \ldots, c_1) = \frac{n_{t,c} + \alpha/k}{t-1+\alpha},$$

where $n_{t,c}$ is the number of c_j for $j < t$ equal to c, that is, the size of the active clusters. Hence, when $k \to \infty$,

$$P(c_t = c | c_{t-1}, \ldots, c_1) \to \frac{n_{t,c}}{t-1+\alpha_i}, \tag{7}$$

$$P(c_t \neq c_j \text{ for all } j < t | c_{t-1}, \ldots, c_1) \to \frac{\alpha}{t-1+\alpha}, \tag{8}$$

where the first line gives the probability that the tth shock ε_t is associated with cluster c, while the second line gives the residual probability that ε_t is associated with a cluster

not observed in $\{c_{t-1}, \ldots, c_1\}$. When compared with the Pólya Urn representation, these equations yield the same clustering behavior and model representation.

In order to operationalize the DPMM, one needs to choose a density $F(\theta_t)$ with corresponding base distribution G_0. For this paper, I adopt a simple yet very flexible specification proposed in Escobar and West (1995), where $F(\theta_t)$ is a Gaussian distribution parametrized by mean μ_t and variance σ_t^2; hence, $\theta_t = (\mu_t, \sigma_t^2)'$. For computational convenience, a conjugate base distribution G_0 is chosen, which is the normal inverse gamma: $(\mu, \sigma^2) \sim \mathcal{N}i\mathcal{G}(s/2, S/2, m, \tau) \sim p(\sigma^2)p(\mu|\sigma^2)$, where $p(\sigma^2) \sim i\mathcal{G}(s/2, S/2)$ is inverse gamma and $p(\mu|\sigma^2) \sim \mathcal{N}(m, \tau\sigma^2)$ normal.

For the Gaussian DPMM, its instructive to look at the implied predictive density conditional on a realization of the mixture parameters $\theta_{1:T} = \{\theta_T, \ldots, \theta_1\}$:

$$\begin{aligned} p(\varepsilon_{T+1}|\theta_{1:T}) &= \int p(\varepsilon_{T+1}|\theta_{T+1})p(\theta_{T+1}|\theta_{1:T})\,d\theta_{T+1} \\ &= \frac{1}{\alpha+T}\sum_{t=1}^{T}\phi(\varepsilon_{T+1};\mu_t,\sigma_t) + \frac{\alpha}{\alpha+T}T_s(\varepsilon_{T+1};m,M), \\ &= \sum_{j=1}^{k}\frac{n_j}{t-1+\alpha}\phi\big(\varepsilon_{T+1};\mu_j^\star,\sigma_j^\star\big) + \frac{\alpha}{\alpha+T}T_s(\varepsilon_{T+1};m,M), \end{aligned}$$

where $\phi(\cdot;\mu,\sigma)$ denotes the density of the normal distribution and $T_s(\cdot;m,M)$ the density of a student-t with mode m, scale $M^{1/2}$ for $M = (1+\tau)S/s$ and s degrees- of-freedom. At first sight, the predictive density shares some similarities with the popular Gaussian kernel density estimator $p(\varepsilon_{T+1}|\varepsilon_{1:T}) \propto \sum_{t=1}^{T}\phi(\varepsilon_{T+1};\varepsilon_t,H)$ where H is a global smoothing parameter. However, there are a few key differences worth mentioning. First, the fact that the DP induces heavy clustering in θ_t means the predictive is shrunk toward a finite set of $k \ll T$ local modes $\{\mu_j^\star, j = 1, \ldots, k\}$. Furthermore, the component scales $\{\sigma_j^\star, j = 1, \ldots, k\}$ may differ allowing for local smoothing. Finally, the density is shrunk globally toward that of a t-distribution, with decreasing importance as sample size increases. The global smoothing parameter α governs both the strength of clustering (and hence sparsity) in $\theta_{1:T}$ as well as the strength of shrinkage toward the t-density. For more details and theoretical insights including consistency and convergence rates see, for example, Escobar and West (1995), Ghosal et al. (1999), and Ghosh and Ramamoorthi (2003).

With respect to computational simplicity, adopting a conjugate base distribution facilitates MCMC inference on the mixing parameters θ_t. To see this, recall that the structural shocks ε_t are assumed to be independent, and hence exchangeable, which yields the following prior based on the Pólya Urn representation:

$$\theta_t|\theta_{-t} \sim \frac{1}{T-1+\alpha}\sum_{j\neq t}\delta_{\theta_j} + \frac{\alpha}{T-1+\alpha}G_0,$$

where $\theta_{-t} = \{\theta_j, j \neq t\}$. Combined with the likelihood $F(\varepsilon_t|\theta_t)$, the posterior is given by the following mixture:

$$\theta_t|\theta_{-t}, \varepsilon_t \sim \sum_{t\neq j} q_{tj}\delta_{\theta_j} + r_t H_t, \tag{9}$$

where $q_{tj} = bF(\varepsilon_t|\theta_j)$, $r_t = b\alpha \int F(\varepsilon_t|\theta)\, dG_0(\theta)$, and H_t is the posterior of θ based on G_0 and ε_t. Furthermore, b is a normalizing constant such that $\sum_{j\neq t} q_{tj} + r_t = 1$. For the conjugate choice G_0, the posterior is analytically tractable and of known form, implying that r_t can be computed in closed form and a random draw is easily generated.

Cycling through the conditionals in (9) may lead to poor convergence. Hence, in this paper I rely on a refinement developed in Neal (2000) yielding improved posterior mixing. Akin to the finite mixture representation (equations (3)–(6)), the algorithm exploits that $\theta_t = \theta^\star_{c_t}$ can be represented in terms of latent allocation variables c_t and that given conjugacy, we can integrate analytically over the cluster parameters $\theta^\star_j$. Combining the prior for c_t implicit in equations (7)–(8) with the integrated likelihood, this yields the conditional:

$$P(c_t = c_j, j = 1, \ldots, k|c_{-t}, \varepsilon_t) = b\frac{n_{-t,c_j}}{T-1+\alpha}\int F(\varepsilon_t|\theta)\, dH_{-t,c_j}(\theta), \tag{10}$$

$$P(c_t \neq c_j \text{ for all } j \neq t|c_{-t}, \varepsilon_t) = b\frac{\alpha}{T-1+\alpha}\int F(\varepsilon_t|\theta)\, dG_0(\theta), \tag{11}$$

where $c_{-t} = \{c_i, i \neq t\}$, and $c_j, j = 1, \ldots, k$ are unique values in c_{-t} of count n_{-t,c_j}, H_{-t,c_j} is the posterior distribution of θ based on prior G_0 and all shocks of $\varepsilon_{-t} = \{\varepsilon_i, i \neq t\}$ assigned to cluster c_j. Finally, b is a normalizing constant. In a second step, conditional on the assignment variables and exploiting the conjugacy of G_0, the (active) cluster parameters $\theta^\star_j, j = 1, \ldots, k$ can be drawn from known distributions in a straightforward manner. The resulting algorithm is reliable, easy to implement, and widely used.

Besides density $F(\cdot)$ and base distribution G_0, one is required to select a value for the concentration parameter α. To understand the impact of α on the complexity of the model, note Figure 2. For a given value of α, the graph shows the implied distribution for the number of unique clusters k and a set of 50 arbitrary predictive densities obtained conditional on drawing $\theta_{1:T}$.The sample size underlying the figure is set to $T = 200$, reflecting typical time-series lengths in macroeconomics. For $\alpha = 1$ (left column, larger value), most of the prior probability mass for k concentrates at values below 10, with a mode between 5 and 6. The predictive densities illustrate the wide range of distributions that can be generated under the DPMM, displaying all kinds of multimodality, skewness, and fat tails. On the other hand, a smaller value $\alpha = 0.1$ (right column) implies that the prior mass for the number of clusters k concentrates at much lower values, with prior mode at just one component. This translates into the prior predictive to be much more concentrated around unimodal shapes, although the variability remains high.

Given the key role for model complexity, α can be thought of as the global smoothing parameter. In order to facilitate selection of α in practice, it can be useful to relate it to the a priori expected number of clusters and variance thereof. Both moments depend

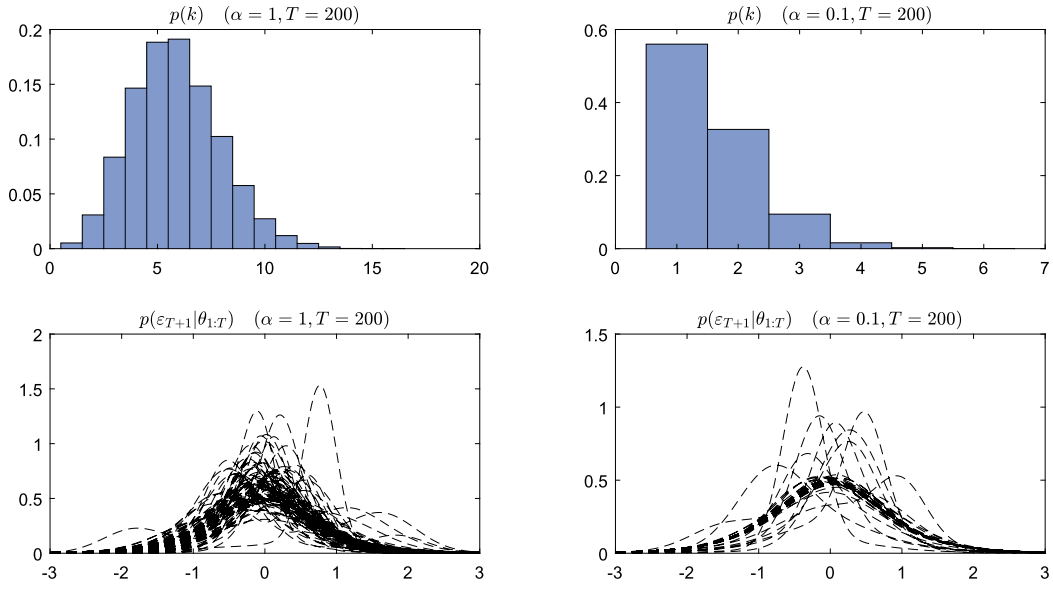

FIGURE 2. Top: implied prior for the number of clusters K. Base distribution is given by $\mathcal{NIG}(s/2 = 5, S/2 = 3/5, m = 0, \tau = 2)$. Bottom: $p(\varepsilon_{T+1}|\theta_{1:T})$ based on 50 prior draws from $\theta_{1:T}$.

on α and the sample size T (Teh et al. (2010)) as follows:

$$\mathrm{E}(k|T,\alpha) = \alpha\big(\psi(\alpha+T) - \psi(\alpha)\big),$$
$$\mathrm{Var}(k|T,\alpha) = \alpha\big(\psi(\alpha+T) - \psi(\alpha)\big) + \alpha^2\big(\psi'(\alpha+T) - \psi'(\alpha)\big),$$

where $\psi(\cdot)$ is the digamma function. In the empirical application, I leverage these formulas to express a prior view on the number of mixture components. The mean-variance Gaussian mixture adopted in this paper is very flexible, able to approximate a wide range of distributions with only few mixture components. For macroeconomic data with relatively low sample size, a reasonable range could therefore be values of α such that $2 \leq \mathrm{E}(k|T,\alpha) \leq 7$. For example, with $T = 500$, setting $\alpha = 0.5$ gives $\mathrm{E}(k|T,\alpha) = 4$ and $\mathrm{Var}(k|T,\alpha) = 2.85$. It is also possible to consider a more conservative choice that puts a larger weight on the identifying information embedded in the economic prior distributions. This can be achieved by setting α very small, effectively centering the prior around a single Gaussian component.

Finally, at the cost of additional computational complexity, it is also possible to spell out a prior distribution for α and learn it from the data. In this case, a convenient prior is the gamma distribution $\alpha \sim \mathcal{G}(a_\alpha, b_\alpha)$ since it allows for simple posterior inference (Escobar and West (1995)). Similarly, the parameters underlying the base distribution τ and m could be treated as random adjusting it to different scales of the data. The conjugate hyperpriors $\tau \sim i\mathcal{G}(a_\tau, b_\tau)$ and $m \sim \mathcal{N}(m_m, V_m)$ are the simplest choice yielding Gibbs steps that can be easily incorporated in the posterior simulator.

2.3 *SVAR–DPMM*

The next step is to embed the DPMM into the multivariate SVAR model, which yields the model proposed in this paper. Let $x_t = [y'_{t-1}, \ldots, y'_{t-p}, 1]'$ and stack the autoregressive coefficients into the $K \times Kp + 1$ matrix $\mathrm{A}_+ = [A_1, A_2, \ldots, A_p, c]$. Assuming the availability of p fixed presample values $y_0, \ldots, y_{-p+1}$, the full model reads:

$$\mathrm{A}(y_t - \mathrm{A}_+ x_t) = \varepsilon_t, \tag{12}$$

$$\varepsilon_{it} | \theta_{it} \sim F(\theta_{it}), \tag{13}$$

$$\theta_{it} \sim G_i, \tag{14}$$

$$G_i \sim \mathrm{DP}(G_{i0}, \alpha_i), \tag{15}$$

for $i = 1, \ldots, K$, $t = 1, \ldots, T$. Here, equation (12) corresponds to the linear SVAR model (Section 2.1) while equations (13)–(15) to the individual DPMM specified for each structural shock. Following Section 2.2, $F(\theta_{it})$ is chosen to be a univariate normal distribution with mean μ_{it} and variance σ^2_{it}, that is, $\theta_{it} = [\mu_{it}, \sigma^2_{it}]$. The base distribution G_{0i} is the conjugate normal inverse gamma distribution $\mathcal{N}i\mathcal{G}(s_i/2, S_i/2, m_i, \tau_i)$.

Denote by $\mathrm{A}_{i\bullet}$ the ith row of A. The following prior distributions are considered for the underlying SVAR model parameters, which completes the specification:

$$\mathrm{A}_{i\bullet} \sim p(\mathrm{A}_{i\bullet}), \tag{16}$$

$$\alpha_+ \sim \mathcal{N}(m_{\alpha_+}, V_{\alpha_+}), \tag{17}$$

for $i = 1, \ldots, K$ and $\alpha_+ = \mathrm{vec}(\mathrm{A}_+)$. Similar to Baumeister and Hamilton (2015), the prior of the structural parameters in $\mathrm{A}_{i\bullet}$ is allowed to take an arbitrary form, enabling the researcher to incorporate identifying information with a high degree of flexibility. To facilitate efficient inference, however, I assume prior independence between different rows of A. As I discuss in Appendix A.1, this allows me to use an extension of the algorithm of Waggoner and Zha (2003) to draw from the conditional posterior of A. For the vectorized reduced-form slope parameters α_+, a Gaussian prior is specified, a fairly common choice, which allows for straightforward inference. The normal prior is widely used in VAR analysis and flexible enough to accommodate both noninformative priors as well as a variety of shrinkage priors including the popular Minnesota prior (Litterman (1986)). Finally, the K concentration parameters $\{\alpha_i, i = 1, \ldots, K\}$ are either treated as fixed or given a prior $\alpha_i \sim \mathcal{G}(a_\alpha, b_\alpha)$, while similar mechanics apply to the parameters underlying the base distribution. In case they are treated as random, $\tau_i \sim i\mathcal{G}(a_\tau, b_\tau)$ and $m_i \sim \mathcal{N}(m_m, V_m)$.

When embedding DPMMs within the SVAR model, some care must be taken with respect to identifiability of location and scale of the shocks. First, unlike Gaussian errors, the marginals arising from DPMMs are not guaranteed to be mean zero. Hence, the intercept of the VAR model is not identified, and can be readily dropped. Alternatively, one may simply ignore the issue as usual quantities important for structural analysis remain unaffected, including impulse response functions or variance decompositions. With respect to scale, a similar problem arises. While in Gaussian SVARs, the scale is often fixed

to unity, doing so within DPMMs is rather involved; see, for example, the approach taken in an earlier version of this paper based on methodology developed in Yang, Dunson, and Baird (2010). For this paper, I follow the model of BH19 and identify the scale of the shocks by normalizing certain elements in A to unity. This is particularly natural if the empirical model can be written as a simultaneous equation system, as is the case for the oil market model considered in this paper. Finally, recall that non-Gaussianity identifies shocks up to an arbitrary permutation (see Section 2.1). In this paper, a unique labeling is obtained through economic restrictions reflected in the prior of A.

2.4 *Posterior inference*

Denote the SVAR parameters $\varphi = \{\mathrm{A}, \alpha_+\}$ and define the collection of auxiliary mixing parameters as $\theta = \{\theta_{it}, i = 1, \ldots, K, t = 1, \ldots, T\}$. The posterior distribution of φ based on observed data Y is proportional to prior times likelihood $p(\varphi|Y) \propto p(Y|\varphi)p(\varphi)$. Note that for DPMM models, the likelihood itself is not directly available, but must be obtained by integrating out the auxiliary parameters θ, that is, $p(Y|\varphi) = \int p(Y|\theta, \varphi) \times p(\theta|\varphi)\, d\theta$. Since both likelihood and posterior are intractable, a full-scale MCMC algorithm is used in this paper to conduct posterior inference on the augmented set of parameters $\xi = \{\varphi, \theta\}$. In the following, I will quickly sketch the algorithm at a high level, and refer to Appendix A.1 for a detailed description.

Let ξ_{-x} be all parameters in ξ but x, and initialize the parameters at some arbitrary initial values. Then the algorithm draws from the posterior by iterating through the following blocks of conditionals:

(1) For each row $\mathrm{A}_{i\bullet}$, draw the SVAR structural parameters from $p(a_i|Y, \xi_{-\{a_i\}})$ via an extension of the algorithm proposed in Waggoner and Zha (2003). Denote by $\mathrm{A}'_{i\bullet} = w_i + W_i a_i$ where a_i is a vector of r_i free elements, W_i a $K \times r_i$ selection matrix, and w_i an $K \times 1$ vector containing constrained values. In Appendix A.1, I show how a random draw can be generated from $p(a_i|Y, \xi_{-\{a_i\}})$ when $w_i \neq 0$, using either a uniform or Gaussian prior for a_i. Under a more general prior, such as priors on nonlinear functions of a_i, a Metropolis–Hastings step can be added to correct for the difference in prior density between proposed and current value of a_i. For the priors considered in the empirical application, the MH acceptance probabilities are very high and vary between 0.6 and 0.99 depending on the row of A.

(2) Draw the VAR regression parameters from $p(\alpha_+|Y, \xi_{-\{\alpha_+\}})$, which is a multivariate normal distribution.

(3) Draw the DPMM parameters as proposed in Algorithm 3 of Neal (2000) (Section 2.2). This includes the hyperparameters $(\alpha_i, \tau_i, m_i, i = 1, \ldots, K)$ in case they are treated as random variables (Escobar and West (1995)).

In order to compute variance and historical decompositions in the SVAR–DPMM model, it is necessary to back out the unconditional variance of structural shocks. Within the MCMC algorithm, it is straightforward to recover these moments from the predictive

density. Conditional on a draw of θ, it is

$$p(\varepsilon_{i,T+1}|\theta_{i,1:T}) = \sum_{j=1}^{k} \frac{n_{ij}}{T-1+\alpha_i} \phi\big(\varepsilon_{i,T+1}; \mu_{ij}^{\star}, \sigma_{ij}^{\star}\big) + \frac{\alpha_i}{\alpha_i + T} T_{s_i}(\varepsilon_{i,T+1}; m_i, M_i),$$

where $M = (1+\tau_i)S_i/s_i$. Effectively, this is a mixture of $k_i + 1$ distributions with component weights given by $w_{ij} = \frac{n_{ij}}{T-1+\alpha_i}$, $j \leq k_i$ and $w_{i,k_i+1} = \frac{\alpha_i}{\alpha_i+T}$. Corresponding component means are $\mu_{ij}^{c} = \mu_{ij}^{\star}$, $j \leq k_i$ and $\mu_{i,k_i+1}^{c} = m_i$, while variances are given by $(\sigma_{ij}^{c})^2 = (\sigma_{ij}^{\star})^2$, $j \leq k_i$ and $(\sigma_{i,k_i+1}^{c})^2 = M_i \frac{s_i}{s_i/(s_i-2)}$. Hence, mean and variance of the predictive can be backed out by standard formulas for mixture models:

$$\mathrm{E}(\varepsilon_{i,T+1}|\theta_{i,1:T}) = \mu_i = \sum_{j=1}^{k_i+1} w_{ij}\mu_{ij}^{c}, \tag{18}$$

$$\mathrm{Var}(\varepsilon_{i,T+1}|\theta_{i,1:T}) = \sigma_i^2 = \sum_{j=1}^{k_i+1} w_{ij}\big(\big(\sigma_{ij}^{c}\big)^2 + \big(\mu_{ij}^{c}\big)^2 - \mu_i^2\big). \tag{19}$$

It might also be useful to study higher-order moments such as skewness and kurtosis. Here, a generic formula for the mth central moment can be used given by $E[(\varepsilon_{i,T+1} - \mu_i)^m] = \sum_{j=1}^{k_i+1} w_{ij} \sum_{l=0}^{m} \binom{m}{l} (\mu_{ij} - \mu_i)^{m-l} C_j(l)$, where $C_j(l)$ denotes the lth central moment of the jth mixture component density. Given posterior draws of ξ, inference for predictive moments is a straightforward byproduct to obtain from the algorithm.

2.5 *Model comparison*

The Bayesian paradigm adopted in this paper postulates to reflect any available identifying information in the prior distributions. In some cases, however, a researcher might be interested in testing competing economic restrictions. In the following, I will provide formal tools to test such overidentifying restrictions in the SVAR–DPMM via Marginal Likelihoods (ML) and Bayes factors.

Consider two models M_1 and M_0, each defined by a likelihood function $p(Y|\varphi_i, M_i)$ and a prior $p(\varphi_i|M_i)$ (for $i = 0, 1$). In the context of testing overidentifying restrictions, think of M_0 as the more restrictive model subject to overidentifying constraints while M_1 is the less restrictive model. A popular metric to quantify the support of the overidentifying restrictions is the Bayes factor, defined as

$$\mathrm{BF}_{10} = \frac{p(Y|M_1)}{p(Y|M_0)},$$

where $p(Y|M_1)$ and $p(Y|M_0)$ are the marginal likelihoods $P(Y|M_i) = \int p(Y|\varphi_i, M_i) \times p(\varphi_i|M_i)\, d\varphi_i$ (for $i = 1, 2$), that is, the probabilities that the data Y has been generated according to models M_1 and M_0, respectively. Under equal prior probability of M_1 and M_0, the Bayes factor has the natural interpretation of posterior odds of M_1 over M_0. High values of the Bayes factor then suggest strong evidence in favor of the less restrictive model and can be interpreted as evidence against the overidentifying restrictions. For

a comprehensive guide on how applied researchers typically interpret magnitudes of Bayes factors, see Kass and Raftery (1995).

For the model considered in this paper, neither $p(Y|M_i)$ nor the likelihood $p(Y|\varphi_i, M_i)$ can be evaluated analytically. Therefore, I suggest to rely on the simulation based cross-entropy (CE) method of Chan and Eisenstat (2015), which has successfully been used to estimate marginal likelihoods in VARs of similar complexity, such as VARs with stochastic volatility (Chan and Eisenstat (2018)). The core of the method exploits an importance sampling estimator for the integral underlying the marginal likelihood. Simplifying the notation by dropping dependence on the model, this estimator is

$$\widehat{p_{IS}(Y)} = \frac{1}{M}\sum_{j=1}^{M}\frac{p\big(Y|\varphi^{(j)}\big)p\big(\varphi^{(j)}\big)}{f\big(\varphi^{(j)}\big)}, \tag{20}$$

where $\varphi^{(1)}, \ldots, \varphi^{(J)}$ are independent draws from an importance density $f(\cdot)$ that dominates the posterior, that is, $f(x) = 0 \Rightarrow p(Y|x)p(x) = 0$. Note that a hypothetical zero variance estimator is given by the (intractable) posterior $f^{\star}(\varphi) = \frac{p(Y|\varphi)p(\varphi)}{p(Y)}$. In order to obtain a feasible estimator with low variance, the CE approach involves selecting a density that is close to the posterior in the sense of the cross-entropy distance $\mathcal{D}(f_1, f_2) = \int f_1(x)\log\frac{f_1(x)}{f_2(x)}\,dx$. A density $f(\cdot; v) \in \mathcal{F}$ with parameters v is chosen such that it minimizes $\mathcal{D}(f^{\star}, f(\cdot; v))$ with respect to v. Algebraic manipulation show that the optimal parameters v_{ce} maximize

$$v_{ce} = \arg\max_{v}\int p(Y|\varphi)p(\varphi)\log f(\varphi; v),$$

which given posterior draws $\varphi^{(1)}, \ldots, \varphi^{(M)}$ can be approximated by

$$v_{ce} = \arg\max_{v} M^{-1}\sum_{j=1}^{M}\log f(\varphi; v).$$

This closely resembles a maximum likelihood estimation problem for v for the density $f(\varphi; v)$, and posterior draws acting as observations. I suggest to choose the following parametric family for the DPMM–SVAR parameters $\varphi = \{A(a), \alpha_+\}$, where a is the vector of free parameters underlying the matrix A:

$$\mathcal{F} = \big\{f_N(\alpha_+; v_{1,\alpha_+}, v_{2,\alpha_+}) \times f_N(a; v_{1,a}, v_{2,a})\big\},$$

where $f_N(\cdot; v_1, v_2)$ is the multivariate normal with mean v_1 and covariance matrix v_2. This simple choice allows to compute v_{ce} in closed form using posterior means and covariance matrices. It is also straightforward to work with a truncated version of $f_N(a; v_{1,a}, v_{2,a})$ if a lot of posterior probability mass lies near the boundary of sign restricted parameters in a.

In order to operationalize the IS estimator in equation (20), one also is required to estimate the likelihood ordinate $p(Y|\varphi)$. Given mutual independence of the error terms

$\varepsilon_t = \mathrm{A}(y_t - \mathrm{A}_+ x_t)$, the likelihood can be factored as $p(Y|\varphi) = |\mathrm{A}|^T \prod_{i=1}^K p(\varepsilon_i|\varphi)$, where

$$p(\varepsilon_i|\varphi) = \int \left\{ \prod_{t=1}^{T} \int f(\varepsilon_{it}|\varphi, \theta_{it})\, dG(\theta_{it}) \right\} dP(G|\alpha_i, G_{i0}),$$

and $dP(G|\alpha_i, G_{i0})$ the DP measure. Also, for the Gaussian mean-variance mixture recall that $\theta_{it} = [\mu_{it}, \sigma_{it}^2]'$ and $f(\varepsilon_{it}|\varphi, \theta_{it}) = \mathcal{N}(\varepsilon_{it}; \mu_{it}, \sigma_{it}^2)$ (see Section 2.2). Evaluating the likelihood requires to integrate over the latent auxiliary variables θ_{it} and it's random prior distribution G_i, which is analytically infeasible. However, given the conjugate base distribution, it is possible to follow Basu and Chib (2003) and exploit a collapsed sequential importance sampler for numerical evaluation. For each shock ($i = 1, \ldots, K$) and $g = 1, \ldots, G$ runs, the underlying algorithm proceeds as follows. First, evaluate $u_{i1}^{(g)} = p(\varepsilon_{i1}|\varphi) = \int p(\varepsilon_{i1}|\varphi, \theta_i)\, dG_{i0}(\theta_i)$ and set $c_{i1}^{(g)} = 1$. Then, for $t = 2, \ldots, T$ sequentially iterate as follows:

1. Compute the predictive probability:

$$\begin{aligned} u_{it}^{(g)} &= p\big(\varepsilon_{it}|\varepsilon_{i,(t-1)}, c_{i,(t-1)}^{(g)}, \varphi, G_{i0}\big), \\ &= \frac{\alpha_i}{\alpha_i + t - 1} \int p(\varepsilon_{it}|\varphi, \theta_i)\, dG_{i0}(\theta_i) + \sum_{j=1}^{k_{t-1}} \frac{n_{j,t-1}}{\alpha_i + t - 1} \int p(\varepsilon_{it}|\varphi, \theta_i)\, dH_{j,t-1}(\theta_i), \end{aligned}$$

   where $\varepsilon_{i,(t-1)} = \{\varepsilon_{il} : l \leq t-1\}$, $H_{j,t-1}(\theta_i)$ is the posterior of θ_i based on prior G_{i0} and all $n_{j,t-1}$ shocks assigned to the jth cluster, that is, $\{\varepsilon_{il} : l \leq t-1, c_l = j\}$.

2. Draw $c_{it}^{(g)}$ from the following categorical distribution:

$$p\big(c_{it} = j|\varepsilon_{i,(t)}, c_{i,(t-1)}^{(g)}, \varphi\big) = \begin{cases} b\dfrac{n_{j,t-1}}{a_i + t - 1} \displaystyle\int p(\varepsilon_{it}|\varphi, \theta_i)\, dH_{j,t-1}(\theta_i) & \text{for } 1 \leq j, \leq k_{t-1}, \\ b\dfrac{\alpha_i}{\alpha_i + t - 1} \displaystyle\int p(\varepsilon_{it}|\varphi, \theta_i)\, dG_{i0}(\theta_i) & \text{for } j = k_{t-1} + 1. \end{cases}$$

   Computing $\omega_i^{(g)} = u_{i1}^{(g)} \prod_{t=2}^{T} u_{it}^{(g)}$ for each run, the ith shock likelihood estimate is then given by $\widehat{p(\varepsilon_i|\varphi)} = \frac{1}{G} \sum_{g=1}^{G} \omega_i^{(g)}$.

The multivariate joint likelihood estimate is then simply given by $\widehat{p(Y|\varphi)} = |\mathrm{A}|^T \times \prod_{i=1}^K \widehat{p(\varepsilon_i|\varphi)}$, which can be used to evaluate the marginal likelihood estimator of equation (20). Note that the resulting procedure resembles the Importance Sampling Squared (IS^2) approach of Tran, Scharth, Pitt, and Kohn (2013). To trade off computational costs of estimating the likelihood and accuracy of the ML estimator, one may follow their analysis in setting G adaptively such that the variance of $\log \widehat{p(Y|\varphi)}$ is about one. Given that the ML estimator is not used in the empirical application of this paper, I use Appendix C to illustrate the reliability of the procedure based on simulated data.

3. The importance of supply and demand for oil prices

In the following, I will use the model to revisit the importance of supply and demand shocks for oil price fluctuations. The empirical strategy is kept simple. Throughout the

analysis, I revisit the four variable oil market model considered in many papers (Kilian and Murphy (2014), Baumeister and Hamilton (2019)) and recover structural shocks with two different identification strategies. The first strategy closely follows Baumeister and Hamilton (2019) (**BH19** henceforth) by imposing a set of sign restrictions combined with weakly informative prior distributions on structural parameters. Combined with a Gaussian (pseudo)likelihood, the resulting posterior distribution reflects information of the prior updated by the covariance structure of the data. The second identification strategy relies on the same identifying information for structural parameters, but in addition, assumes mutual independence and non-Gaussianity of the shocks (**BH19+NG**). Hence, any difference in the posteriors between the two identification approaches will reflect the additional identifying information from non-Gaussianity.

3.1 *Model and identification*

Following Baumeister and Hamilton (2019), the model includes the following four observables:

$$y_t = \left[100 \times \Delta q_t,\, 100 \times \Delta y_t^a,\, 100 \times \Delta p_t,\, \Delta i_t\right]',$$

where q_t is the log of global crude oil production (in million barrels per day) and y_t^a is a measure of world economic activity proxied by industrial production of the OECD plus 6 major countries. Furthermore, p_t is the log of the real oil price, defined as the US Refiner's Acquisition Cost of oil deflated with the US consumer price index, and Δi_t is a proxy for OECD oil inventories expressed as a fraction of global crude oil production. The data set is monthly and covers the period from 1974m1 until 2019m12. For a detailed description of the data set, I refer to the paper of BH19.

Abstracting from dynamics and the Δ notation, the structural oil market model takes the form of the following simultaneous equations:

$$\textit{Supply}: \quad q_t = \alpha_{qp} p_t + \varepsilon_t^s, \tag{21}$$

$$\textit{Economic activity}: \quad y_t^a = \alpha_{yp} p_t + \varepsilon_t^{ad}, \tag{22}$$

$$\textit{Consumption demand}: \quad q_t - i_t^\star = \beta_{qy} y_t^a + \beta_{qp} p_t + \varepsilon_t^{cd}, \tag{23}$$

$$\textit{Inventory demand}: \quad i_t^\star = \psi_1 q_t + \psi_3 p_t + \varepsilon_t^{id}, \tag{24}$$

$$\textit{Measurement error}: \quad i_t = \chi i_t^\star + \varepsilon_t^{me}, \tag{25}$$

where $\varepsilon_t = [\varepsilon_t^s, \varepsilon_t^{ea}, \varepsilon_t^{cd}, \varepsilon_t^{id}, \varepsilon_t^{me}]' \sim (0, \Sigma_\varepsilon)$ are uncorrelated structural shocks, which implies that $\Sigma_\varepsilon = \text{diag}(\sigma_1^2, \ldots, \sigma_5^2)$ is diagonal. Note that there are five equations that summarize the contemporaneous relations across four observables. First, consider equation (25), which reflects an assumption about additive measurement error in the observed inventories variable i_t. Specifically, it decomposes the variable into an unobserved "true" inventory series $i_t^\star$ and a measurement error ε_t^{me}. BH19 rationalize this approach by noting that inventory data is only available for OECD countries, which is arguably only a fraction χ of world inventories. Equation (21) characterizes the behavior of global oil

supply, relating production to real oil prices via the coefficient α_{qp}. Given that both variables are expressed in log deviations, the coefficient α_{qp} can be interpreted as the (short-run) price elasticity of oil supply. The third equation (22) characterizes global economic activity (EA), decomposing world industrial production into a component driven by oil prices and an EA shock ε_t^{ea}. Equation (23) models consumption demand, relating quantity consumed $q_t - i_t^\star$ to world output and oil prices. Here, β_{qp} is the oil price elasticity of demand while β_{qy} characterizes the response of demand to increased economic activity. Finally, equation (24) captures residual demand for oil inventory, which is related to quantity and prices via coefficients ψ_1 and ψ_3.

Adding dynamics, the simultaneous equation model can be written as an A- type structural VAR as described in Section 2:

$$y_t = c + \sum_{j=1}^{p} A_j y_{t-j} + u_t, \tag{26}$$

$$\underbrace{\begin{pmatrix} 1 & 0 & -\alpha_{pq} & 0 & 0 \\ 0 & 1 & -\alpha_{yp} & 0 & 0 \\ 1 & -\beta_{qy} & -\beta_{qp} & 0 & -1 \\ -\psi_1 & 0 & -\psi_3 & 0 & 1 \\ 0 & 0 & 0 & 1 & -\chi \end{pmatrix}}_{\mathrm{A}} \underbrace{\begin{pmatrix} u_t^q \\ u_t^y \\ u_t^p \\ u_t^i \\ u_t^{i^\star} \end{pmatrix}}_{\tilde{u}_t} = \underbrace{\begin{pmatrix} \varepsilon_t^s \\ \varepsilon_t^{ea} \\ \varepsilon_t^{cd} \\ \varepsilon_t^{id} \\ \varepsilon_t^{me} \end{pmatrix}}_{\varepsilon_t}. \tag{27}$$

Equation (26) models the reduced-form dynamics of the observables with a VAR, yielding prediction errors $u_t = [u_t^q, u_t^y, u_t^p, u_t^i]'$. In order to allow for sufficient dynamics, the model includes $p = 12$ lags. The structural model is given in equation (27) and written in terms of augmented errors $\tilde{u}_t = [u_t, u_t^{i^\star}]'$, which includes $u_t^{i^\star}$, the unobserved prediction error for the (latent) global inventory series. A simple counting exercise reveals that this model cannot be identified from second moments of the data, given that $u_t^{i^\star}$ is unobserved. In particular, there are 12 structural parameters (7 elements in A plus 5 elements in Σ_ε) but there are only 10 reduced-form parameters available in the covariance matrix of the observable prediction errors u_t.

Two identification strategies are considered throughout this section, differing in the specification of the error term. In the first model (**BH19**), a Gaussian (pseudo)likelihood is used, that is, $\varepsilon_t \sim \mathcal{N}(0, \Sigma_\varepsilon)$. Here, identification is obtained via a set of sign-restricted prior distributions for elements underlying the A matrix. In the second specification, the same sign restrictions and prior information holds, with the additional identifying assumption that all shocks but the measurement error (ε_t^{me}) are mutually independent and non-Gaussian (**BH19+NG**). It is important to note that this model is the more restrictive model, ruling out any higher-order dependence across shocks (see Section 2.3). For the latter model, each marginal is modeled in a Bayesian nonparametric way, that is, for $i = 1, \ldots, 4$, I assume Dirichlet process mixture models of the form $\varepsilon_{it}|\theta_{it} \sim F(\theta_{it})$, $\theta_{it} \sim G_i$, $G_i \sim \mathrm{DP}(G_{i0}, \alpha_i)$. Following the methodology outlined in Section 2, I set $F(\cdot)$ as the univariate normal distribution parameterized by mean and variance $\theta_{it} = [\mu_{it}, \sigma_{it}^2]$.

The priors for the parameters (G_i) follow Dirichlet processes with conjugate normal inverse gamma base distributions $G_{0i} \sim \mathcal{NIG}(s_i/2, S_i/2, m_i, \tau_i)$. Finally, in both models the latent measurement error is assumed to be Gaussian, that is, $\varepsilon_t^{me} \sim \mathcal{N}(0, \sigma_5^2)$.

The priors and sign restrictions used for each parameter are set out in Table 1. First, consider the structural parameters underlying A. Regarding the oil price elasticities of supply α_{qp} and demand β_{qp}, BH19 make use of truncated student-t distributions concentrated around 0.1 and -0.1, respectively. While these values seem to be small, the prior mean for α_{qp} is typically restricted to much lower values (Zhou (2020)). However, with scales of 0.2 and 3 degrees-of-freedom, the distributions are only weakly informative. As for the income elasticity β_{qy}, BH19 draw on external evidence from the literature to elicit a positively truncated student-t distribution with mode around 0.7, scale 0.2 and 3 degrees-of-freedom. The effect of oil prices on economic activity α_{yp} is judged to be rather small, reflected in a (negatively) truncated t-distribution with mode at just -0.05. A smaller scale of 0.1 reflects more prior certainty than for the other parameters endowed with a student-t prior, but the degrees-of-freedom are still set to 3; hence, the distribution is relatively spread out. For the parameters of the inventory equation ψ_1 and ψ_3, no prior knowledge is available, so uninformative student-t priors are used concentrated around 0 with scale of 0.5 and 3 degrees-of-freedom. With respect to χ, the fraction of inventories held by OECD countries, BH19 specify a Beta prior concentrated around 0.6, matching roughly the share of OECD countries in world oil consumption. The prior parameters are set in such a way that the standard deviation is equal to 0.1, reflecting a moderate degree of uncertainty for this number.

The diagonal elements of Σ_ε are given weakly informative inverse gamma priors in the Gaussian model, for all shocks but the measurement error. In the non-Gaussian model, shock variances are indirectly parameterized by the Dirichlet process mixture model. To keep the model simple, the same inverse gamma prior is used as the base distribution of the scale, while a weakly informative specification is used for the location. The global smoothing parameters are chosen such that for the sample size at hand, the a priori expected number of mixture components (k_i) is $\mathrm{E}(k_i|T, \alpha_i) = 3$ for each shock, which can be obtained for $\alpha_i = 0.3$.[6]

Finally, consider the variance of the measurement error. Instead of a direct prior for σ_5^2, BH19 express a prior belief on the nonlinear transformation $\rho = \frac{\chi^{-1}\sigma_5^2}{\sigma_3^2 + \chi^{-2}\sigma_5^2}$. This is motivated by the fact that, since $u_t^{i\star}$ is unobserved, the algorithm developed in their previous paper (Baumeister and Hamilton (2015)) cannot be readily applied. To get around this issue, BH19 rewrite the first four equations of (27) using observables. Algebraic manipulations yield $\mathrm{A}^\dagger u_t = \varepsilon_t^\dagger$, for

$$\mathrm{A}^\dagger = \begin{pmatrix} 1 & 0 & -\alpha_{qp} & 0 \\ 0 & 1 & -\alpha_{yp} & 0 \\ 1 & -\beta_{qy} & -\beta_{qy} & -\chi^{-1} \\ -\tilde{\psi}_1 & 0 & -\tilde{\psi}_3 & 1 \end{pmatrix},$$

[6]For this value of α_i, the a priori variance of k_i is $\mathrm{Var}(k_i|T, \alpha) = 1.9$. Note that the results are not sensitive to using a more involved model where α_i and the parameters underlying the base distribution are treated as random variables (Section 2.4).

TABLE 1. Summary of prior distributions.

Student-t Distribution					
		Location	Scale	dof	Sign Restriction
α_{qp}	Oil supply elasticity	0.1	0.2	3	$\alpha_{qp} > 0$
α_{yp}	Effect of p on activity	−0.05	0.1	3	$\alpha_{yp} < 0$
β_{qy}	Income elasticity of oil demand	0.7	0.2	3	$\beta_{qy} > 0$
β_{qp}	Oil demand elasticity	−0.1	0.2	3	$\beta_{qp} < 0$
ψ_1	Effect of q on inventories	0	0.5	3	none
ψ_3	Effect of p on inventories	0	0.5	3	none
Beta Distribution					
		Mean	Standard Deviation		Sign Restriction
χ	Fraction of inventories	0.6	0.1		none
$\rho^\star$	Importance of measurement error in $u_t^{i\star}$	0.25	0.12		none
Normal Distribution					
		Mean	Variance		Sign Restriction
α_+	vector of autoregressive parameters	0	$100 \times I$		none
Inverse Gamma Distribution (Only Gaussian Model)					
		Mean	Variance		Sign Restriction
σ_i^2, $i = 1, 2, 3, 4$	shock variances	2	2		none
Normal Inverse Gamma Distribution (Only Non-Gaussian Model)					
		Mean	Variance		Sign Restriction
σ^2	scale of DPMM base distributions	2	2		none
μ	location of DPMM base distributions	0	σ^2		none
Other Fix Parameters (Only Non-Gaussian Model)					
		Value			
α_i, $i = 1, 2, 3, 4$	concentration parameter	0.3			none

$\psi_1^\dagger = \chi\psi_1$, $\psi_3^\dagger = \chi\psi_3$, and $\varepsilon_t^\dagger = [\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t} - \chi^{-1}\varepsilon_{5t}, \chi\varepsilon_{4t} + \varepsilon_{5t}]$. BH19 then show that premultiplying the system further by

$$\Gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \rho & 1 \end{pmatrix}$$

yields orthogonal shocks $\varepsilon_t^\star = \Gamma \varepsilon_t^\dagger$, and hence allows the use of their standard algorithm for $A^\star = \Gamma A^\dagger$. Note that ρ can be thought of the as the negative coefficient from a regression of $\varepsilon_{4t}^\dagger$ on $\varepsilon_{3t}^\dagger$. Since by construction $\rho \in (0, \chi)$, BH19 use a Beta prior centered around 0.25χ as to reflect a moderate importance of the measurement error.

Unfortunately, in the non-Gaussian model this simplification strategy cannot be applied. First, although the residuals in $\varepsilon_t^\star$ are orthogonal by construction, they are no longer independent. This means, it is necessary to infer the latent inventory prediction errors $u_t^{i\star}$ and measurement error ε_t^{me} during estimation, and keep them apart from the other non-Gaussian structural shocks. Second, a prior on ρ instead of σ_5^2 is not compatible with the inference algorithm proposed for the DPMM–SVAR of this paper, given that it would induce prior dependence between structural parameters and the DPMM model of the third shock, invalidating the MCMC algorithm proposed in this paper. To see this, first solve for $\sigma_5^2 = \frac{\rho\sigma_3^2}{\chi^{-1} - \rho\chi^{-2}}$. Hence, the implicit prior for σ_5^2 not only depends on χ and ρ, but also σ_3^2, which is parameterized by the DPMM of the third shock (see Section 2.4). To break the dependence, I instead express a prior belief on the fraction of variance in the latent inventories explained by the measurement error. Noting that $u_t^i = \chi^{-1} u_t^{i\star} + \varepsilon_{5t}$, the resulting coefficient is given by $\rho^\star = \frac{\sigma_5^2}{\chi^{-2}\operatorname{var}(u_t^i) + \sigma_5^2} \in (0, 1)$, where I use a training sample (pre-1974 inventory data) to set $\operatorname{var}(u_t^i) \approx 1.3$. Reflecting a moderate degree of importance, I set the Beta prior such that $E(\rho^\star) = 0.25$ with standard deviation 0.12.

3.2 *Inference*

Posterior inference for the model is conducted via the Markov Chain Monte Carlo algorithm described in Appendix A.1, iterating through the conditional distributions of α_+, each row of A, and the auxiliary parameters underlying each of the Dirichlet process mixture models. Note that due to the additional complexity induced by the measurement error equation, the algorithm needs to be modified in two ways. First, an additional block is a necessary draw from the conditional distribution of $u_t^{i\star}$, the latent "true" inventory series. The full conditional for $[u_1^{i\star}, \ldots, u_T^{i\star}]'$ is Gaussian and described in Appendix A.2. The second modification concerns drawing the parameters underlying the measurement error equation (25). For this purpose, note that σ_5^2 can be absorbed into $A_{5\bullet} = [0, 0, 0, \sigma_5^{-1}, -\chi\sigma_5^{-1}]'$, normalizing the the variance to unity ($\varepsilon_{5t} \sim \mathcal{N}(0, 1)$). Furthermore, recall that a prior is imposed on $\rho^\star = \frac{\sigma_5^2}{\chi^{-2}1.3 + \sigma_5^2}$ rather than σ_5^2, and algebraic manipulations yield $A_{5\bullet} = [0, 0, 0, a_{54}, a_{55}]'$ for $a_{54} = (1.3\frac{\rho^\star}{(1-\rho^\star)}\chi^{-2})^{1/2}$ and $a_{55} = -\chi(1.3\frac{\rho^\star}{(1-\rho^\star)}\chi^{-2})^{1/2}$. Since both elements a_{54} and a_{55} are nonlinear functions of χ and $\rho^\star$, it is necessary to compute the Jacobian of transformation at the stage of drawing $A_{5\bullet}$. This takes into account that the proposal distribution is developed under a uniform prior for elements in $A_{5\bullet}$, which is not necessarily uniform for χ and $\rho^\star$. As described in Appendix A.1, it is straightforward to account for the Jacobian during the Metropolis–Hastings step. Finally, Appendix A.3 also describes how the algorithm is adjusted when the errors are Gaussian (BH19). Here, I simply replace the block responsible

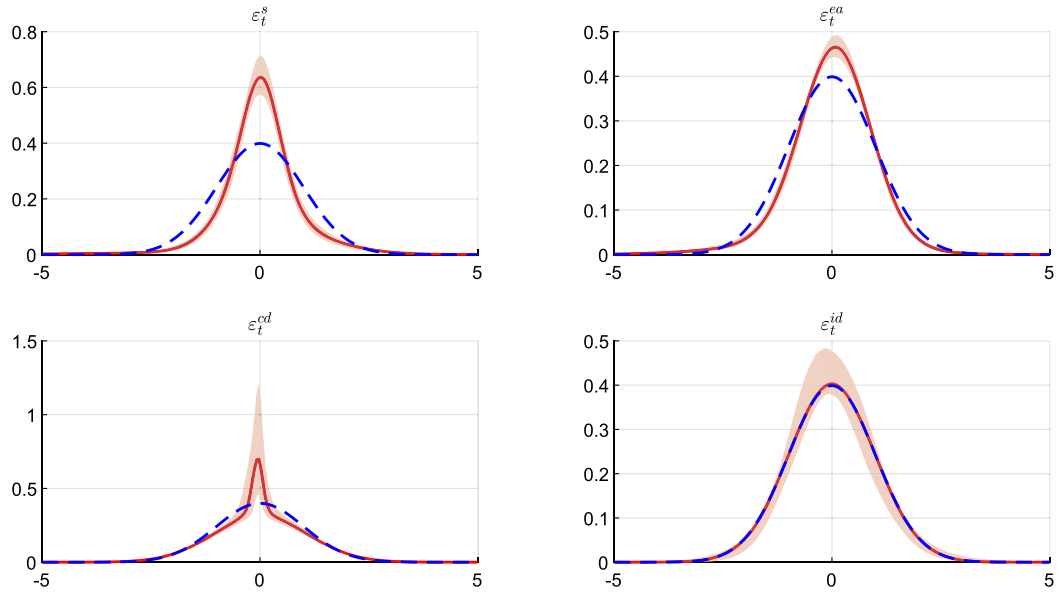

FIGURE 3. Posterior predictive densities (90% credible interval) of standardized structural shocks $\tilde{\varepsilon}_{i,T+1} = \sigma_i^{-\frac{1}{2}}(\varepsilon_{i,T+1} - \mu_i)$. The blue dashed line indicates the density of a standard normal distribution.

for drawing the DPMM parameters with a block that draws from the conditional distribution of the shock variances in Σ_ε.

3.3 *Empirical results*

Before discussing results from structural analysis obtained under the two identification schemes, it is useful to assess the empirical plausibility of assuming non-Gaussianity and mutual independence in the oil market context.

In Figure 3, posterior median estimates of the predictive densities are provided for standardized structural shocks, alongside 90% posterior confidence sets (shaded area). Furthermore, for comparison, the density of a standard normal distribution is drawn in as blue dashed line. The results suggest that three out of four structural shocks display large degrees of non-Gaussianity in some regions of the predictive density. Table 2 further contains summary statistics for skewness and kurtosis of the posterior predictive distributions. The supply and economic activity shock display strong excess kurtosis and left skewness, while for the consumption demand shock there is evidence for a minor degree of excess kurtosis. In line with the visual analysis of the posterior predictive, there is no evidence of skewness or excess kurtosis for the inventory demand shock. Overall, the evidence suggests that there is potential for considerable identifying information that can be exploited in the context of the oil-market application.

Next, I study the plausibility of the mutual independence assumption, which is important for the non-Gaussian model. Given that the structural shocks are endowed with

TABLE 2. Skewness and kurtosis of the identified oil market shocks (non-Gaussian model).

Moment	ε_t^s	ε_t^{ea}	ε_t^{cd}	ε_t^{id}	ε_t^{me}
Skewness	−1.08	−0.68	0.03	0.01	0
	(−1.51, −0.22)	(−0.99, −0.33)	(−0.02, 0.12)	(−0.58, 0.91)	
Kurtosis	12.76	6.53	3.53	3.38	3
	(9.68, 16.75)	(5.10, 8.65)	(3.24, 3.85)	(2.86, 6.73)	

Note: The table gives posterior median estimates of the skewness and kurtosis of structural shocks in the non-Gaussian model. Values in brackets indicate corresponding 90% posterior credibility sets.

independent DPMMs, popular testing strategies such as Bayes factors cannot be pursued to assess independence within the current framework. However, it is straightforward to report posterior predictive distributions of popular frequentist test statistics. I follow this route and report the posterior of two popular statistics used previously in non-Gaussian SVAR analysis. The first is a nonparametric test developed in Matteson and Tsay (2017). Denote by $E = [\varepsilon_1 : \ldots : \varepsilon_T]'$ the $T \times K$ structural shocks, then the statistic is given by $U(E) = T \sum_{j=1}^{K-1} \mathcal{I}_T(\hat{U}_j, \hat{U}_{j+})$, where $j+ = \{l : j < l \leq K\}$ denotes the indices $(j+1, \ldots, K)$, $\hat{U}_j$ has elements defined as $\hat{u}_{i,k} = \frac{1}{T}\text{rank}\{\varepsilon_{ij} : \varepsilon_{ij} \in E_j\}$, and $\mathcal{I}_T$ is the empirical distance covariance as defined in Matteson and Tsay (2017). Distance covariances are a multivariate dependence measure for two random vectors, which is zero only under mutual independence. The test statistic converges to a nondegenerate distribution, which can be easily approximated by a bootstrap. While this test is consistent against all types of dependence, others may have higher power against certain alternatives. One alternative is discussed in Olea, Plagborg-Møller, and Qian (2022), testing for shared volatility factors in structural shocks. They propose the test statistic $S(E) = \sqrt{\frac{1}{K(K-1)} \sum_{i=1}^{K} \sum_{j \neq i} \text{Corr}(\varepsilon_{it}^2, \varepsilon_{jt}^2)^2}$, which measures the root mean squared sample cross-correlations of squared structural shocks. Again, a bootstrap is used to approximate the distribution of the test under the null hypothesis of independence.

Figure 4 plots the posterior of these two test statistics, for the non-Gaussian (top row) and Gaussian model (bottom row) separately. For comparison, I overlap each distribution with that of the same statistic computed for randomly repermuted shocks, denoted by $U_0(E)$ and $S_0(E)$.[7] This helps to get an indication of how the posterior of the test statistic would look like under the null of mutual independence.

With respect to the non-Gaussian model (top row), the distribution of $U(E)$ is virtually indistinguishable from that based on resampled shocks ($U_0(E)$), suggesting no evidence against mutual independence. The test of Olea, Plagborg-Møller, and Qian (2022) indicates that some differences are present between the posteriors of $S(H)$ and $S_0(H)$. However, the distributions overlap to a large extend hinting at non-conclusive evidence. I find that the posterior median of the test statistic $S(H)$ roughly aligns with the 90% quantile of the posterior obtained for resampled shocks ($S_0(H)$). A frequentist interpretation would suggest that under mutual independence 10 out of 100 random samples

[7]Consistent with mutual independence, each shock $\hat{\varepsilon}_{j,t}$ is resampled independently of the other shocks instead of resampling all components in the vector ε_t jointly. This is repeated at each iteration of the posterior inference algorithm.

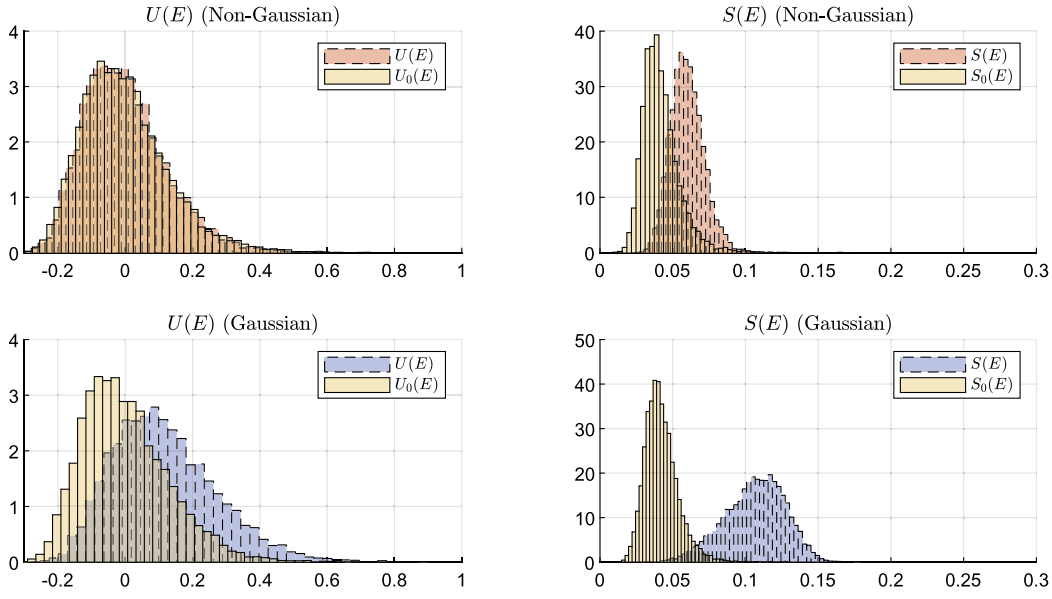

FIGURE 4. Posterior distribution of the test statistic proposed by Olea, Plagborg-Møller, and Qian (2022) ($S(E)$) and Matteson and Tsay (2017) ($U(E)$) for the non-Gaussian (top row) and Gaussian (bottom row) model. For comparison, $S_0(H)$ and $H_0(H)$ denote the corresponding test statistics under the null hypothesis, obtained using shocks that are randomly repermuted under the mutual independence assumption.

end up yielding such a difference, which is fairly weak evidence against the null. In the light of these results, I proceed under the assumption that the data does not object exploiting the mutual independence assumption for identification purposes.

The results for the conventionally identified SVAR are given in the bottom row. The posterior distributions of $U(E)$ and $U_0(E)$ still largely overlap, although a discrepancy starts to become visible. For the test of Olea, Plagborg-Møller, and Qian (2022), however, the evidence against mutual independence is very strong. The posterior of the test statistic $S(E)$ is substantially larger than in BH19+NG, and a comparison with the distribution of the resampled statistic $S_0(H)$ suggests very little overlap. It is fair to say that this evidence does not invalidate the results of BH19 per se, as the success of conventional identification strategies does not depend on mutual independence. On the other hand, readers comfortable maintaining the assumption of mutual shock independence may view this as evidence against the posterior estimates obtained by Baumeister and Hamilton (2019).

Given the strong deviations from Gaussianity in oil market shocks, one can expect that the posterior distributions of structural parameters differ across the two identification schemes. In Figure 5, I plot prior and posterior distributions for key structural parameters in the model and indeed find considerable differences. The first column compares prior and posterior distribution of α_{qp}, the short-run oil price elasticity of supply. Under a Gaussian likelihood (top), the prior distribution is peaking close to the prior mode, although uncertainty decreased substantially. In contrast, in the non-Gaussian

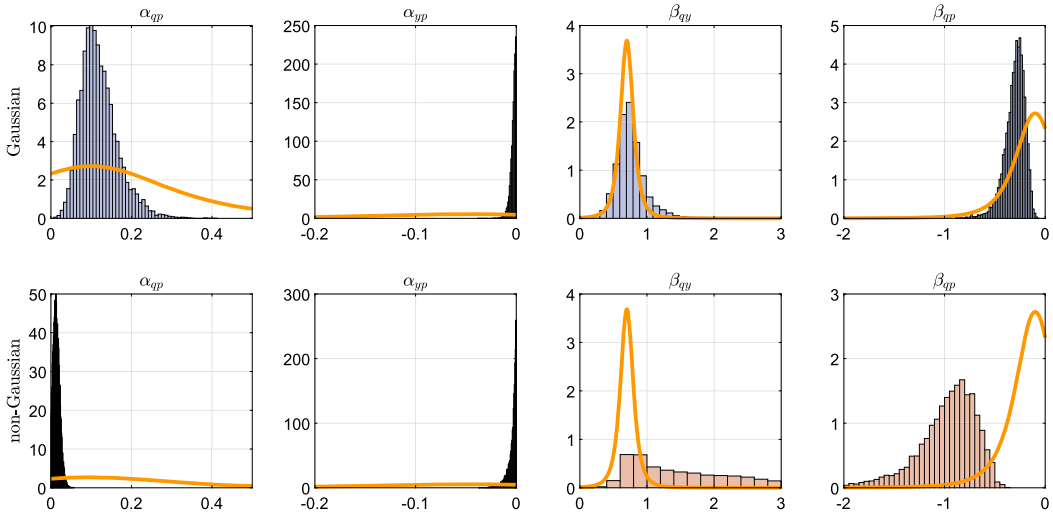

FIGURE 5. Prior (orange solid line) and posterior density of key structural parameters. Top panel: Gaussian model. Bottom panel: non-Gaussian model.

model, posterior mass is concentrated very close to zero. With respect to the effect of oil prices on activity, α_{yp}, the posterior distributions are very similar across both identification approaches. Compared to the prior, both posteriors concentrate strongly around values close to zero. Stronger differences in the posteriors are visible in the parameters underlying the consumption demand equation. With respect to the income elasticity of oil demand (β_{qy}), prior and posterior coincide in the Gaussian model, which may reflect that there is very little information in the covariance structure of the data to learn about this parameter. In the non-Gaussian model, however, the prior distribution is updated to some extent. While the modal value is still below one, a remarkable degree of posterior mass is attached to larger values. A similar picture arises for the oil demand elasticity (β_{qp}). In the Gaussian model, the prior is only slightly revised but posterior mass still concentrates around high density regions of the prior. Instead, in the non-Gaussian model, the posterior is revised to a much larger extent. The posterior mode indicates that the demand elasticity is estimated to be larger than indicated by the prior, with a modal value of -0.9. However, posterior uncertainty remains high. Posteriors of the remaining structural parameters of the simultaneous equation model do not differ, and I refer the interested reader to Appendix D for the details.

Differences in the posterior distribution of key structural parameters will have implications for structural analysis. In Figure 6, I provide (pointwise) posterior medians and 90% credible sets for impulse response functions (IRFs) up to 16 months, each standardized to increase oil prices by 1% on impact.[8] The IRFs track the dynamic response of structural innovations on the level of the four endogenous variables. First, consider the effects of the oil supply shock (first row). In the non-Gaussian model, a much larger disruption in supply is required to achieve a price increase of the same magnitude. In turn,

[8]For an alternative approach involving joint inference on impulse response functions, see Inoue and Kilian (2022).

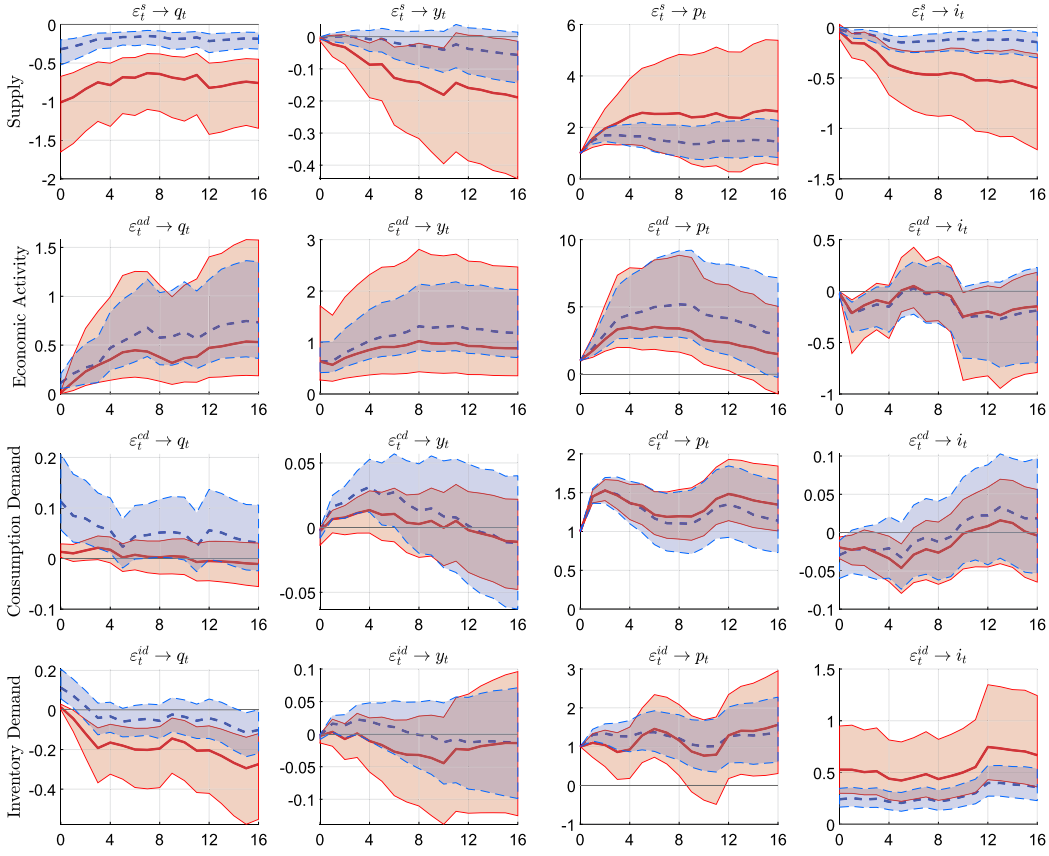

FIGURE 6. Posterior median IRFs with 90% credible intervals (shaded areas). Blue (dashed lines): Gaussian model. Red (solid lines): non-Gaussian model.

this leads to a considerably stronger response of global economic activity and drawdown of inventories compared to the Gaussian specification. The opposite can be found for the consumption demand shock. Here, the estimated increase in oil produced is considerably more muted in the non-Gaussian model. No significant difference can be observed across the identification schemes for the response of global economic activity and oil inventories. In line with the literature, economic activity may slightly increase while inventories are drawn-up to mitigate some of the price increase.

IRFs to an Economic Activity (EA) shock are virtually indistinguishable across both identification approaches (BH19 and BH19+NG). An EA shock that increases oil prices by 1% is associated with a slowly increasing production, increase in global activity, and decrease in inventories. With respect to the inventory demand shock, some subtle differences are found. First, note that the response of world activity, oil price, and inventory are quite similar across specifications. For both models, oil prices and inventories display a positive co-movement, while global activity is barely affected. However, the impact response of global oil production differs. While in BH19, oil production is estimated to increase for a few months before gradually decreasing, the impact in the

TABLE 3. Forecast Error Variance Decomposition (FEVD) of the real oil price growth.

Horizon	ε_t^s	ε_t^{ea}	ε_t^{cd}	ε_t^{id}	ε_t^{me}
			Gaussian Model		
4	**0.32**	0.06	**0.57**	0.03	0.01
	(0.15, 0.55)	(0.03, 0.11)	(0.34, 0.75)	(0.01, 0.08)	(0, 0.03)
16	**0.29**	0.08	**0.53**	0.03	0.06
	(0.15, 0.5)	(0.04, 0.12)	(0.32, 0.68)	(0.01, 0.08)	(0.03, 0.1)
			Non-Gaussian Model		
4	**0.05**	0.05	**0.87**	0.01	0.02
	(0.02, 0.1)	(0.01, 0.12)	(0.79, 0.93)	(0, 0.02)	(0, 0.03)
16	**0.06**	0.06	**0.8**	0.01	0.06
	(0.03, 0.11)	(0.03, 0.12)	(0.72, 0.86)	(0, 0.03)	(0.03, 0.1)

Note: The table gives posterior median estimates of the contribution of each shock to the forecast error variance of the real oil price at the 4- and 16-months horizon. Values in brackets indicate corresponding 90% posterior credibility sets.

non-Gaussian model is virtually zero and is estimated to decrease afterwards. Hence, in BH19+NG, the shock behaves similar to the oil supply news shock discussed in Känzig (2021). These shocks reflect an anticipated decrease in oil production, which is associated with a sudden precautionary build-up of oil inventories and strong increase in oil prices. Note, however, that the effect on oil prices is more muted than documented in Känzig (2021).

Finally, Table 3 contains the forecast error variance decomposition of the real price of oil at the 4- and 16-months horizon. Once more, the main difference across the two identification approaches is found along the effects of supply and consumption demand shocks (highlighted in bold). As for the supply shocks, they are found to be more important in the Gaussian model than the non-Gaussian model. In particular, posterior median estimates indicate that in the BH19 model, supply shocks explain around one-third of the variance observed in real oil prices, with 90% credible sets covering anything between 15% and 55%. On the contrary, if non-Gaussianity is exploited as an additional identification device, posterior median estimates suggest that supply shocks explain only a very small fraction of oil price movements, with median estimates at just 6%. In this identification scheme, posterior credibility sets are substantially more narrow and indicate that supply shocks are unlikely to explain more than 11% of the variation. As for demand shocks, the opposite effect can be documented. Here, 90% posterior credible sets suggest that in the Gaussian model, consumption demand shocks explain between 32% and 75% of the variation. This contrasts sharply with much larger estimates associated with the non-Gaussian model. Specifically, posterior credible sets cover values between 72% and 93%.

Summing up, the results from the empirical analysis yield two main findings. First, when non-Gaussianity is exploited for identification, oil supply is estimated to be more inelastic, while (consumption) demand is more elastic than reported in Baumeister and Hamilton (2019). Second, under non-Gaussianity, supply shocks are found to be only minor drivers of oil price fluctuations and demand shocks explain most of the variation. These results are in line with early papers in the literature that impose a very inelastic supply as identifying restriction, leaving the demand elasticity relatively unrestricted

(Kilian and Murphy (2012, 2014), Zhou (2020)). However, in contrast to these papers, the estimates obtained above are not a result of strong priors but of identifying information contained in assuming mutual independence and non-Gaussianity of shocks.

In Appendix E, I offer more insights on the relationship between restricting α_{pq} in a conventionally identified model (BH19), and resulting posterior estimates of the demand elasticity β_{qp} and variance decomposition of the real price of oil. First, the findings of BH19 are obtained when restricting α_{pq} near the prior mean of 0.1, yielding relatively inelastic demand and supply shocks explaining around 30% of oil price fluctuations. On the other hand, the results of BH19+NG are obtained when restricting α_{qp} to a very small value (0.01), yielding a more elastic demand and little role of supply shocks.

3.4 *Robustness*

It is fair to say that for many experts in the oil market literature, the posterior of the short-run price elasticity of demand (β_{qp}) obtained under non-Gaussianity is too large in absolute terms to be economically plausible. Typically, values below -0.8 have been considered unreasonable, which corresponds to often cited estimates of the long-run elasticity of demand (see, e.g., Hausman and Newey (1995)). As discussed in Appendix E, models that find (or directly impose) a small oil supply elasticity, end up with a fairly large coefficient for β_{qp}, and the model of this paper is no exception in this regard.

While there is question on what statistic is the most appropriate to measure demand elasticity (Kilian (2022b)), I conduct two robustness exercises to test how much the results are driven by the posterior of β_{qp}. First, if $\beta_{qp} < -0.8$ is considered unreasonable a priori, I can follow Kilian and Murphy (2012) and Baumeister and Peersman (2013) by directly truncating the prior of β_{qp} over the interval $(-0.8, 0)$. Table 4 shows the results of such an exercises, labeled as R1. To maintain comparability, I provide results for both the Gaussian and non-Gaussian model, reporting posterior estimates for α_{qp}, β_{qp}, and the contribution of ε_t^s to the FEVD of the real price of oil. First, note that the estimates for the Gaussian model are not affected under the alternative prior. This is not a surprise given that the bulk of the posterior mass of β_{qp} already lies above -0.8 in the baseline results (see Figure 5). In the non-Gaussian model, 90% posterior credibility sets of β_{qp} are now between -0.6 and -0.79 reflecting the additional hard constraint. Other than that, imposing the alternative prior does not materially affect the posterior of α_{qp} nor estimates of the supply shocks contribution to the forecast error variance of the oil price. While estimates are slightly higher than reported under the baseline results, supply shocks still play a minor role. Point estimates suggest that only 10% of price variation is driven by supply shocks.

In a second robustness check, I assess the sensitivity of the results to using data only starting from January 1985. This is motivated by the fact that previous studies document a possible break in oil-market dynamics around that date; see, for example, Baumeister and Peersman (2013). Corresponding results are labeled as R2 in Table 4. Under the shorter sample, the posteriors of the Gaussian model point toward smaller elasticities (in absolute terms) of supply and demand, which resembles findings of Aastveit et al. (2021). A similar pattern applies to the results obtained by the non-Gaussian model.

TABLE 4. Robustness analysis for the main empirical findings.

	Gaussian			Non-Gaussian		
	5%	50%	95%	5%	50%	95%
			Panel A: Posterior α_{qp}			
R1	0.06	0.12	0.21	0.01	0.02	0.03
			Panel A: Posterior β_{qp}			
R1	−0.48	−0.28	−0.16	−0.79	−0.70	−0.52
R2	−0.29	−0.17	−0.09	−0.36	−0.24	−0.16

	Gaussian		Non-Gaussian	
	$h=4$	$h=16$	$h=4$	$h=16$
	Panel C: Contribution of ε_t^s to the FEVD of the real price of oil			
R1	0.31 (0.14, 0.54)	0.29 (0.14, 0.49)	0.08 (0.05, 0.12)	0.08 (0.05, 0.13)
R2	0.32 (0.15, 0.55)	0.29 (0.15, 0.48)	0.2 (0.11, 0.29)	0.19 (0.12, 0.27)

Note: For robustness check R1, the model is reestimated based on a student-t prior of β_{qy} truncated on the interval $(-0.8, 0)$. For robustness check R2, the model is reestimated based on a shortened sample covering January 1985 to December 2019. In panel C, the values in brackets give 90% posterior confidence sets.

While the posterior for α_{qp} is still very close to zero, 90% credibility sets of β_{qp} lie between -0.16 and -0.36. This is much lower than observed in the full sample. Regarding the contribution of ε_t^s to the FEVD of the real price, point estimates of around 19–20% obtained under the non-Gaussian model are considerable higher than in the baseline specification. However, the same qualitative pattern is observed in that once non-Gaussianity is introduced into the model; supply shocks become less important drivers of oil prices.

Another interesting exercise is to check if the results obtained under non-Gaussianity are sensitive to a more informative prior for α_{qp}. To this end, I re-estimate the model with a tighter prior $\alpha \sim t_{0,\infty}(0.1, 0.05, 100)$, that is, a positively truncated student-t distribution with 100 degrees-of-freedom, scale of 0.05, and the same mode 0.1. Interestingly, the results are not affected. For the oil market model studied in this paper, the information from non-Gaussianity seems to clearly dominate weakly informative prior distributions. In order to materially alter the estimates obtained under BH19-NG, one would have to tighten the prior a lot more, which becomes difficult to justify economically.

In Appendix F, I provide further robustness analysis with respect to the error specification used in the combined identification scheme (BH19+NG). First, I study the use of parametric student-t distributions instead of nonparametric DPMMs. I find that point estimates turn out to be fairly similar, but that posterior uncertainty is larger in the parametric alternative. This points toward a higher efficiency of the nonparametric estimator in the empirical application. Second, I study if the results are sensitive to the choice of α_i, the global smoothing parameter of the density estimators. Specifically, α_i is set such that a priori, the model strongly favors only one mixture component, and hence Gaussian shocks. In this model, only the supply and economic activity shock display

non-Gaussian posterior predictive distributions. However, the main empirical results are very similar to those reported by the baseline DPMM model.

Finally, note that Carriero, Marcellino, and Tornese (2023) also revisit the model of Baumeister and Hamilton (2019), but based on identification by heteroskedasticity instead of non-Gaussianity. Unlike this paper, the identification strategy allows for mutual dependence of shocks via the presence of cross-sectional volatility clustering. It is encouraging that their findings on the importance of supply shocks are very similar to those presented in this paper, pointing toward further robustness of the results to an alternative set of statistically motivated identifying assumptions. Earlier are also provided in Lütkepohl and Netšunajev (2014) and Herwartz and Plödt (2016).

4. Conclusion

In this paper, new evidence is provided on the relative importance of supply and demand shocks for oil price fluctuations. To disentangle their effects, identification by non-Gaussianity is exploited in addition to a set of sign restrictions and weakly informative prior distributions for structural parameters (Baumeister and Hamilton (2019)). The empirical findings indicate that under this identification strategy, oil supply shocks become minor drivers of oil prices. The results are compatible with estimates obtained previously in the literature (Kilian (2009), Kilian and Murphy (2012, 2014), Zhou (2020)), however, without the need of very strong identifying restrictions for structural parameters.

From an econometric point of view, this paper offers a novel Bayesian estimator for non-Gaussian SVAR models. Specifically, each structural shocks marginal density is modeled nonparametrically using Bayesian infinite mixture models. The benefit from pursuing a nonparametric approach is that estimates are robust to misspecification of the error term, and that the procedure requires no prior knowledge on the form of non-Gaussianity. The flexible density estimators are perfectly able to exploit deviations from normality at any region of the sample space, and hence, flexibly capture excess kurtosis, skewness, or other type of non-Gaussianity often documented in structural shocks.

It is important to acknowledge that identification via non-Gaussianity is not cost-free, and requires strong assumptions on the independence of structural shocks. Higher-order dependence can arise easily in rotations of orthogonalized linear VAR prediction errors, for example, under nonlinearities in the second moment of shocks. To test the empirical plausibility of the mutual independence assumption, this paper simply reports posterior distributions of frequentist test statistics and resampled versions thereof (Matteson and Tsay (2017), Olea, Plagborg-Møller, and Qian (2022)). Future work should focus on developing coherent Bayesian tests with the goal to provide powerful devices to test the mutual independence assumption.

References

Aastveit, Knut Are, Hilde C. Bjørnland, and Jamie L. Cross (2021), "Inflation expectations and the pass-through of oil prices." *The Review of Economics and Statistics*, 1–26. [1164, 1191]

Amengual, Dante, Gabriele Fiorentini, and Enrique Sentana (2022), "Moment tests of independent components." *SERIEs*, 13 (1), 429–474. [1169]

Antolín-Díaz, Juan and Juan F. Rubio-Ramírez (2018), "Narrative sign restrictions for SVARs." *American Economic Review*, 108 (10), 2802–2829. [1163]

Anttonen, Jetro, Markku Lanne, and Jani Luoto (2021), "Statistically identified svar model with potentially skewed and fat-tailed errors." Available at SSRN 3925575. [1167]

Basu, Sanjib and Siddhartha Chib (2003), "Marginal likelihood and Bayes factors for Dirichlet process mixture models." *Journal of the American Statistical Association*, 98 (461), 224–235. [1167, 1179]

Baumeister, Christiane and James D. Hamilton (2015), "Sign restrictions, structural vector autoregressions, and useful prior information." *Econometrica*, 83 (5), 1963–1999. [1170, 1175, 1182]

Baumeister, Christiane and James D. Hamilton (2019), "Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks." *American Economic Review*, 109 (5), 1873–1910. [1163, 1164, 1165, 1180, 1187, 1190, 1193]

Baumeister, Christiane and Gert Peersman (2013), "The role of time-varying price elasticities in accounting for volatility changes in the crude oil market." *Journal of Applied Econometrics*, 28 (7), 1087–1109. [1191]

Bertsche, Dominik and Robin Braun (2022), "Identification of structural vector autoregressions by stochastic volatility." *Journal of Business & Economic Statistics*, 40 (1), 328–341. [1170]

Bjørnland, Hilde C., Frode Martin Nordvik, and Maximilian Rohrer (2021), "Supply flexibility in the shale patch: Evidence from North Dakota." *Journal of Applied Econometrics*, 36 (3), 273–292. [1165]

Blackwell, David and James B. MacQueen (1973), "Ferguson distributions via Pólya urn schemes." *The Annals of Statistics*, 1 (2), 353–355. [1170]

Boscolo, Riccardo, Hong Pan, and Vwani P. Roychowdhury (2004), "Independent component analysis based on nonparametric density estimation." *IEEE Transactions on Neural Networks*, 15 (1), 55–65. [1167]

Braun, Robin (2023), "Supplement to 'The importance of supply and demand for oil prices: Evidence from non-Gaussianity'." *Quantitative Economics Supplemental Material*, 14, https://doi.org/10.3982/QE2091. [1167]

Caldara, Dario, Michele Cavallo, and Matteo Iacoviello (2019), "Oil price elasticities and oil price fluctuations." *Journal of Monetary Economics*, 103, 1–20. [1163, 1165]

Carriero, Andrea, Massimiliano Giuseppe Marcellino, and Tommaso Tornese (2023), "Blended identification in structural vars." *BAFFI CAREFIN Centre Research Paper*, (200). [1193]

Chan, Joshua, Gary Koop, and Xuewen Yu (2023), "Large order-invariant Bayesian vars with stochastic volatility." *Journal of Business & Economic Statistics*. [1166]

Chan, Joshua C. C. and Eric Eisenstat (2015), "Marginal likelihood estimation with the cross-entropy method." *Econometric Reviews*, 34 (3), 256–285. [1167, 1178]

Chan, Joshua C. C. and Eric Eisenstat (2018), "Bayesian model comparison for time-varying parameter vars with stochastic volatility." *Journal of Applied Econometrics*, 33 (4), 509–532. [1178]

Cross, Jamie L., Bao H. Nguyen, and Trung Duc Tran (2022), "The role of precautionary and speculative demand in the global market for crude oil." *Journal of Applied Econometrics*, 37 (5), 882–895. [1163]

Davis, Richard and Serena Ng (2023), "Time series estimation of the dynamic effects of disaster-type shocks." *Journal of Econometrics*, 235 (1), 180–201. [1169]

Drautzburg, Thorsten and Jonathan H. Wright (2023), "Refining set-identification in vars through independence." *Journal of Econometrics*, 235 (2), 1827–1847. [1167]

Escobar, Michael D. and Mike West (1995), "Bayesian density estimation and inference using mixtures." *Journal of the American Statistical Association*, 90 (430), 577–588. [1166, 1172, 1174, 1176]

Ferguson, Thomas S. (1973), "A Bayesian analysis of some nonparametric problems." *The Annals of Statistics*, 1 (2), 209–230. [1170]

Fiorentini, Gabriele and Enrique Sentana (2022), "Discrete mixtures of normals pseudo maximum likelihood estimators of structural vector autoregressions." *Journal of Econometrics*, 235 (2), 643–665. [1166, 1167]

Ghosal, Subhashis, Jayanta K. Ghosh, R. V. Ramamoorthi et al. (1999), "Posterior consistency of Dirichlet mixtures in density estimation." *The Annals of Statistics*, 27 (1), 143–158. [1172]

Ghosh, Jayanta K. and R. V. Ramamoorthi (2003), *Bayesian Nonparametrics*. Springer Science & Business Media. [1172]

Gourieroux, Christian, Alain Monfort, and Jean-Paul Renne (2017), "Statistical inference for independent component analysis: Application to structural VAR models." *Journal of Econometrics*, 196 (1), 111–126. [1167]

Gouriéroux, Christian, Alain Monfort, and Jean-Paul Renne (2020), "Identification and estimation in non-fundamental structural varma models." *The Review of Economic Studies*, 87 (4), 1915–1953. [1170]

Hausman, Jerry A. and Whitney K. Newey (1995), "Nonparametric estimation of exact consumers surplus and deadweight loss." *Econometrica: Journal of the Econometric Society*, 1445–1476. [1191]

Herrera, Ana Maria and Sandeep Kumar Rangaraju (2020), "The effect of oil supply shocks on us economic activity: What have we learned?" *Journal of Applied Econometrics*, 35 (2), 141–159. [1164]

Herwartz, Helmut (2018), "Hodges–Lehmann detection of structural shocks: An analysis of macroeconomic dynamics in the euro area." *Oxford Bulletin of Economics and Statistics*, 80 (4), 736–754. [1167]

Herwartz, Helmut and Martin Plödt (2016), "The macroeconomic effects of oil price shocks: Evidence from a statistical identification approach." *Journal of International Money and Finance*, 61, 30–44. [1193]

Hirano, Keisuke (2002), "Semiparametric Bayesian inference in autoregressive panel data models." *Econometrica*, 70 (2), 781–799. [1166]

Inoue, Atsushi and Lutz Kilian (2022), "Joint Bayesian inference about impulse responses in VAR models." *Journal of Econometrics*, 231 (2), 457–476. [1188]

Juvenal, Luciana and Ivan Petrella (2015), "Speculation in the oil market." *Journal of Applied Econometrics*, 30 (4), 621–649. [1163]

Känzig, Diego R. (2021), "The macroeconomic effects of oil supply news: Evidence from OPEC announcements." *American Economic Review*, 111 (4), 1092–1125. [1190]

Kass, Robert E. and Adrian E. Raftery (1995), "Bayes factors." *Journal of the american statistical association*, 90 (430), 773–795. [1178]

Kilian, Lutz (2009), "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market." *American Economic Review*, 99 (3), 1053–1069. [1163, 1193]

Kilian, Lutz (2022a), "Facts and fiction in oil market modeling." *Energy Economics*, 110, 105973. [1163]

Kilian, Lutz (2022b), "Understanding the estimation of oil demand and oil supply elasticities." *Energy Economics*, 107, 105844. [1165, 1191]

Kilian, Lutz and Helmut Lütkepohl (2017), *Structural Vector Autoregressive Analysis. Themes in Modern Econometrics*. Cambridge University Press. [1168]

Kilian, Lutz and Daniel P. Murphy (2012), "Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market VAR models." *Journal of the European Economic Association*, 10 (5), 1166–1188. [1163, 1191, 1193]

Kilian, Lutz and Daniel P. Murphy (2014), "The role of inventories and speculative trading in the global market for crude oil." *Journal of Applied Econometrics*, 29 (3), 454–478. [1163, 1164, 1180, 1191, 1193]

Lanne, Markku and Jani Luoto (2021), "GMM estimation of non-Gaussian structural vector autoregression." *Journal of Business & Economic Statistics*, 39 (1), 69–81. [1167]

Lanne, Markku and Jani Luoto (2020), "Identification of economic shocks by inequality constraints in Bayesian structural vector autoregression." *Oxford Bulletin of Economics and Statistics*, 82 (2), 425–452. [1167]

Lanne, Markku, Mika Meitz, and Pentti Saikkonen (2017), "Identification and estimation of non-Gaussian structural vector autoregressions." *Journal of Econometrics*, 196 (2), 288–304. [1167, 1169]

Lanne, Markku and Pentti Saikkonen (2007), "A multivariate generalized orthogonal factor GARCH model." *Journal of Business & Economic Statistics*, 25 (1), 61–75. [1170]

Litterman, Robert (1986), "Forecasting with Bayesian vector autoregressions—five years of experience." *Journal of Business & Economic Statistics*, 4 (1), 25–38. [1175]

Lütkepohl, Helmut (2005), *New Introduction to Multiple Time Series Analysis*. Springer Science & Business Media. [1168]

Lütkepohl, Helmut and Aleksei Netšunajev (2014), "Disentangling demand and supply shocks in the crude oil market: How to check sign restrictions in structural VARs." *Journal of Applied Econometrics*, 29 (3), 479–496. [1193]

Matteson, David S. and Ruey S. Tsay (2017), "Independent component analysis via distance covariance." *Journal of the American Statistical Association*, 112 (518), 623–637. [1169, 1186, 1187, 1193]

Mavroeidis, Sophocles (2021), "Identification at the zero lower bound." *Econometrica*, 89 (6), 2855–2885. [1169]

Mertens, Karel and Morten O. Ravn (2010), "Measuring the impact of fiscal policy in the face of anticipation: A structural var approach." *The Economic Journal*, 120 (544), 393–413. [1170]

Montiel Olea, José Luis, Mikkel Plagborg-Møller, and Eric Qian (2022), "Svar identification from higher moments: Has the simultaneous causality problem been solved?" In *AEA Papers and Proceedings*, Vol. 112, 481–485. [1169, 1186, 1187, 1193]

Neal, Radford M. (2000), "Markov chain sampling methods for Dirichlet process mixture models." *Journal of computational and graphical statistics*, 9 (2), 249–265. [1171, 1173, 1176]

Newell, Richard G. and Brian C. Prest (2019), "The unconventional oil supply boom: Aggregate price response from microdata." *The Energy Journal*, 40 (3). [1165]

Normandin, Michel and Louis Phaneuf (2004), "Monetary policy shocks: Testing identification conditions under time-varying conditional volatility." *Journal of Monetary Economics*, 51 (6), 1217–1243. [1170]

Rigobon, Roberto (2003), "Identification through heteroskedasticity." *Review of Economics and Statistics*, 85 (4), 777–792. [1165]

Teh, Yee Whye et al. (2010), "Dirichlet process." *Encyclopedia of machine learning*, 1063, 280–287. [1174]

Tran, Minh-Ngoc, Marcel Scharth, Michael K. Pitt, and Robert Kohn (2013), "Importance sampling squared for Bayesian inference in latent variable models." ArXiv preprint. [1179]

Villani, Mattias (2009), "Steady-state priors for vector autoregressions." *Journal of Applied Econometrics*, 24 (4), 630–650. [1166]

Waggoner, Daniel F. and Tao Zha (2003), "A Gibbs sampler for structural vector autoregressions." *Journal of Economic Dynamics and Control*, 28 (2), 349–366. [1166, 1175, 1176]

Yang, Mingan, David B. Dunson, and Donna Baird (2010), "Semiparametric Bayes hierarchical models with mean and variance constraints." *Computational Statistics & Data Analysis*, 54 (9), 2172–2186. [1176]

Zhou, Xiaoqing (2020), "Refining the workhorse oil market model." *Journal of Applied Econometrics*, 35 (1), 130–140. [1163, 1182, 1191, 1193]

Co-editor Tao Zha handled this manuscript.

Manuscript received 8 February, 2022; final version accepted 9 June, 2023; available online 27 June, 2023.