Andrade, Philippe; Ferroni, Filippo; Melosi, Leonardo

Working Paper

Higher-order moment inequality restrictions for SVARs

Working Papers, No. 25-3

Provided in Cooperation with:
Federal Reserve Bank of Boston

Suggested Citation: Andrade, Philippe; Ferroni, Filippo; Melosi, Leonardo (2025) : Higher-order moment inequality restrictions for SVARs, Working Papers, No. 25-3, Federal Reserve Bank of Boston, Boston, MA, https://doi.org/10.29412/res.wp.2025.03

This Version is available at:
https://hdl.handle.net/10419/324801

Working Papers

FEDERAL RESERVE BANK OF BOSTON

No. 25-3

Higher-order Moment Inequality Restrictions for SVARs

Philippe Andrade, Filippo Ferroni, and Leonardo Melosi

Abstract:
We introduce a method that exploits some non-Gaussian features of structural shocks to identify structural vector autoregression (SVAR) models. More specifically, we propose combining inequality restrictions on the higher-order moments of the structural shocks of interest with other set-identifying constraints, typically sign restrictions. We illustrate how, in both large and small sample settings, higher-order moment restrictions considerably narrow the identification of monetary policy shocks compared with what is obtained with minimal sign restrictions typically used in the SVAR literature. The proposed methodology also provides new insights into the macroeconomic effects of sovereign risk in the euro area as well as the transmission of geopolitical risk to the US economy.

JEL Classifications: C32, E27, E32
Keywords: Shock identification, skewness, kurtosis, sign restrictions, monetary policy, sovereign risk, geopolitical risk

Philippe Andrade (philippe.andrade@bos.frb.org) is a vice president and economist in the Federal Reserve Bank of Boston Research Department. Filippo Ferroni is a professor of economics at the Università di Bologna. Leonardo Melosi is a professor of economics at the University of Warwick.

An earlier draft of the paper was circulated under the title "Identification Using Higher-order Moments Restrictions."

The authors thank Christian Wolf, Frank Schorfheide, Diego Känzig, Daniel Lewis, Tincho Almuzara, Fabio Canova, Raffaella Giacomini, Jonathan Wright, Luca Gambetti, Giovanni Ricco, and participants at the 2024 NBER Summer Institute, the SED 2024 Annual Meeting, the 2023 Barcelona Summer Forum, and various seminars for comments and suggestions. They also thank Mary Tracy for her outstanding research assistance.

This version: March 2025 **https://doi.org/10.29412/res.wp.2025.03**

1 Introduction

Since the seminal contribution of Sims (1980), numerous structural vector autoregression (SVAR) methods have been developed to study the propagation of unobserved and unanticipated shocks to dynamic systems of observed macroeconomic variables. These identification schemes, which rely on exclusion or sign constraints derived from economic analysis, typically involve restrictions on the second-order moments of the shocks but leave their higher-order moments unrestricted.[1]

When shocks are nearly Gaussian, higher-order moment restrictions are of little use to identification, as they are close to being redundant with restrictions imposed on second-order moments. However, macroeconomic and financial data often appear to be non-Gaussian, with normal-time variations coexisting with unusually large positive or negative changes. Examples include various types of financial crises leading potentially to sharp recessions, inflation surges induced by large supply disruptions, or abrupt changes in macroeconomic policies. These large and unpredictable events can be viewed as being triggered by structural shocks drawn from a distribution with fat tails of potentially asymmetrical mass.

In this paper, we introduce a new identification method that complements second-order moment restrictions, popular in the SVAR literature, with higher-order moment properties of the structural shocks of interest.

We postulate that the data-generating process (DGP) of the economy is a vector autoregression (VAR) with uncorrelated non-Gaussian structural shocks, therefore departing from Gaussianity while retaining the simplicity of a linear transmission mechanism. In this setup, orthonormal rotations of the underlying shock processes leave their variance–covariance matrix unchanged but impact their third- and fourth-order moments, so that it is impossible to recover the "true impact" of a structural shock on the observable variables without reproducing its correct higher-order moment. As we show, if these higher-order moments were known, this property could be used to point-identify the impact of the structural shocks of interest on the observable variables. However, in practice, imposing exact *equality* restrictions on higher-order moments of structural shocks is challenging. Indeed, in-sample estimates of higher-order moments can be very imprecise for sample sizes typically encountered in macroeconomic applications. Moreover, and related, these shocks are not directly observable; thus, one must work with estimated proxies. Overall, as Montiel Olea, Plagborg-Møller and Qian (2022) underscores, inferences based on these higher-order moment equality restrictions can be biased and misleading. In addition, the approach point-identifies structural shocks using statistical assumptions rather than restrictions derived from economic theory.

Our approach is more flexible. We impose *inequality* restrictions on higher-order moments—

[1] See Ramey (2016), Nakamura and Steinsson (2018b) for surveys of the literature on identification in macroeconomics.

typically the third- and/or fourth-order moments—of structural shocks of interest that we postulate to be non-Gaussian. Such non-Gaussian features can be motivated by indirect empirical evidence or by economic reasoning. Higher-order moment inequality restrictions are set-identifying, rather than point-identifying, conditions. They can thus be combined with other set-identification constraints derived from economic analysis, typically sign restrictions. In a simple illustrative case, we show analytically how the identified set shrinks when we impose an inequality restriction on the third-order moment of the structural shock.

Our method poses two practical difficulties. The first involves addressing potential small-sample bias associated with the estimation of higher-order moments. To do so, we treat the distribution of the structural shock of interest non-parametrically and impose restrictions on the distance between different percentiles of the empirical distribution to generate the desired asymmetry and/or fat-tailedness.[2] The second difficulty involves estimating and conducting inference in a VAR with non-Gaussian errors. To do so, we use a Bayesian approach building on the work of Petrova (2022), to which we add a few simulation steps to achieve identification via higher-moment restrictions.

We study our method's performance by applying it to simulated data from (1) a calibrated, textbook New Keynesian model as in Galí (2015) but in which we assume a Laplace distribution for the monetary policy shock, and (2) a Smets and Wouters (2007)'s medium-scale dynamic stochastic general equilibrium (DSGE) model estimated using US data. As we document, the monetary policy shock recovered from that estimated model exhibits significant excess kurtosis. We identify monetary policy shocks by combining the agnostic sign restrictions of Uhlig (2005)—which require that monetary policy shocks move interest rates and prices in opposite directions but leave their impact on real activity unrestricted—with a higher-order moment inequality restriction requiring that monetary policy shocks exhibit minimal excess kurtosis.

As in Wolf (2020), when we implement Uhlig (2005)'s sign restrictions on these simulated data, we find that a monetary policy tightening has no clear effect on output. This is at odds with the models used to generate the data as they both imply that a tightening of monetary policy lowers output. As Wolf (2020) shows, the reason for this inconsistency is that some *combinations* of positive supply and demand shocks that have a positive effect on output also have a negative impact on prices and a positive impact on interest rates. Therefore, when Uhlig (2005)'s sign restrictions are used, these combinations of shocks *masquerade* as contractionary monetary policy shocks. In contrast, when we combine Uhlig (2005)'s sign restrictions with an inequality restriction on the excess kurtosis of monetary policy shocks, we find that a tightening

[2]Robust estimators of kurtosis consider the ratio between the distance of the percentiles in the tails and the distance between the percentiles close to the median. In the case of leptokurtic shocks, the larger the numerator, the thicker the tails; the smaller the denominator, the more clustered the distribution around the median. Robust estimators of skewness compute the distance between mean and median.

of monetary policy clearly lowers output, which is consistent with the models used to generate the data. The reason is that the set of combinations of supply and demand that can masquerade as monetary policy shocks is much smaller in our case, as most of these combinations do not meet the higher-order moment inequality restrictions. In other words, our method strongly mitigates the masquerading issue emphasized in Wolf (2020).

We also study how our identification approach performs compared with an approach that relies on the independence of structural shocks and exact equality constraints on their third- and/or fourth-order moments. We show that, while this more restrictive approach allows us to point-identify the structural shock of interest and performs well for very large samples, our method is much more accurate when applied to samples of the size typical in macroeconomic studies. The reason is that obtaining unbiased in-sample estimates of higher-order moments proves to be quite challenging, as these are highly sensitive to outliers and minor perturbations of the data.

We then apply our identification method to three empirical issues. First, we revisit the question of the effect of monetary policy on output and inflation, this time using real, rather than simulated, data. We start by highlighting that a wide set of measures of US monetary policy shocks proposed in the literature feature significant excess kurtosis compared with Gaussian processes. By contrast, their cross-correlation is quite low. We show that this non-Gaussian feature is also observed in various monetary policy shock measures for the euro area and the United Kingdom. This motivates exploiting this robust feature to identify monetary policy shocks. We thus require that their excess kurtosis be larger than a lower bound which we derive from the empirical distributions of the various monetary policy shock proxies that we consider. We combine this higher-order moment inequality restriction with Uhlig (2005)'s agnostic sign restrictions. We find that the impact of a monetary policy shock on output is non-significant when using minimal sign restrictions only. By contrast, a monetary policy tightening induces a significant output contraction when these minimal sign restrictions are combined with higher-moment inequality restrictions. This is consistent with results obtained through alternative means of sharpening the identification of monetary policy, such as the narrative restrictions of Antolin-Diaz and Rubio-Ramírez (2018) or additional sign restrictions, as in Arias, Caldara and Rubio-Ramírez (2019).

We then turn to two less-investigated questions and for which restrictions on higher-order moments of structural shocks are quite natural to consider. In a second application, we study the macroeconomic impact of sovereign risk shocks in the euro area. Following the 2010–2012 euro-area sovereign debt crisis, these effects were prominent in policymaker discussions; however, there is little quantitative assessment of their impact. Bocola (2016) emphasizes the recessionary impact of these shocks using a structural model. In the spirit of the SVAR approach, we propose

an estimation that uses less-specific structural assumptions. We estimate a macro-financial VAR model of the euro area that includes a measure of non-financial corporate borrowing costs, as well as the Italian and German sovereign bond yields, and the associated sovereign spread. We assume that sovereign risk shocks can relatively frequently be very large and positive; that is, they feature a positively skewed distribution with fat tails. We therefore require that sovereign spread shocks meet some inequality constraints on their third- and fourth-order moments. We combine these higher-order moment restrictions with the following minimal sign restrictions: A sovereign spread shock increases the 10-year Italian government bond yield, the Italian–German spread, and the non-financial corporate borrowing costs, on impact and for the following month. We find that an increase in sovereign risk leads to an immediate tightening of credit conditions, as well as a contraction of output on impact and for the next two to three years. The effects on the euro-area unemployment are also quite severe and last longer. By contrast, we find that the macroeconomic impact of the sovereign shocks is non-significant when relying on sign restrictions to identify them.

Finally, in a third application, we investigate the impact of geopolitical risks on the US economy. We use the geopolitical risk index developed by Caldara and Iacoviello (2022), which is based on newspaper coverage of geopolitical tensions. This index is both right-skewed and fat-tailed, reflecting that geopolitical tensions can lead to very large positive increases in geopolitical risks and can happen relatively frequently compared with the Gaussian case. We identify geopolitical risk shocks by combining third- and fourth-order-moment inequality restrictions with some minimal sign restrictions, which assume that geopolitical risk tightens financial conditions on impact. Consistent with what Caldara and Iacoviello (2022) obtain with the same VAR system but applying a recursive identification scheme, and consistent with the literature emphasizing the macroeconomic effects of uncertainty (Bloom 2009), our results indicate that an increase in geopolitical risk contracts investment and hours. Our estimated effects are significantly larger and more persistent than Caldara and Iacoviello (2022)'s and show that geopolitical risks lead to a persistent tightening of US financial conditions. Importantly, using sign restrictions alone leads to very imprecise and statistically non-significant estimated macroeconomic effects of geopolitical risk shocks.

The paper is organized as follows. The next section (1.1) discusses the existing literature. Section 2 presents our identification approach, detailing the estimation strategy. Section 4 shows the role of kurtosis in identifying the effect of monetary policy on output in standard New Keynesian models. Sections 5-7 illustrate the usefulness of our approach in three empirical applications: measuring the real effects of conventional US monetary policy, the impact of sovereign spread shocks in the euro area, and the transmission of geopolitical risk shocks.

1.1 Literature Review

Our paper contributes to the literature that relies on non-Gaussian features of macroeconomic or financial data to identify structural shocks. Existing methods typically assume a specific non-Gaussian distribution for VAR reduced-form errors and postulate that these are independent (see Lanne, Meitz and Saikkonen 2017, Gouriéroux, Monfort and Renne 2017, 2019, among others). These strong assumptions allow to point identify the system of underlying structural shocks. Jarociński (2021) implements this approach using intraday variations in interest rate futures to identify the effects of conventional and unconventional US monetary policy shocks on financial markets. Other scholars (e.g. Lanne, Liu and Luoto 2022, among others) achieve statistical identification of structural shocks by minimizing the distance between the VAR model-implied empirical innovation higher-order moments and their sample counterparts. Finally, some studies identify shocks by exploiting time variation in the conditional variance of shocks (see Rigobon 2003, Lewis 2021, Brunnermeier, Palia, Sastry and Sims 2021, among many others). Lewis (2024) provides a survey of identification methods using higher-order moment properties.

Montiel Olea et al. (2022) underline three potential drawbacks of these methods. First, they achieve point identification of shocks by relying on strong statistical assumptions, particularly by ruling out situations in which shocks can be large at the same time. Second, existing methods suffer from potentially large in-sample biases, as they rely on point estimates of higher-order moments for all the shocks in the system, each being highly sensitive to outliers for samples of limited size. Third, these methods identify structural shocks using statistical properties rather than constraints derived from economic reasoning. By contrast, we make much less restrictive assumptions, as we merely postulate deviations from Gaussianity for the structural shocks of interest. We do not rely on the independence of shocks, and we allow for co-skewness and co-kurtosis. We also treat the higher-order moment properties as necessary restrictions that structural shocks should fulfill. Importantly, and different from previous contributions, our approach generates a set-identification of the shock of interest as opposed to the system point identification. Working with set-identified shocks allows us to couple these restrictions with other popular macroeconomic assumptions—motivated by economic analysis—on the signs of the (or zero) impact of the shock on the endogenous variables, magnitude or elasticity bounds, or narrative restrictions on historical episodes. Moreover, our methodology does not rest on in-sample estimates of the empirical innovation higher-order moments. It imposes inequality restrictions on the higher-order moments of the structural shock itself using robust non-parametric methods based on the distance between different percentiles of the shock empirical distribution.

Our approach also relates to previous research that aims to sharpen set-identified structural shocks obtained with restrictions derived from theory. Wolf (2020, 2022) underline that imposing minimal and uncontroversial sign restrictions alone is often too weak to adequately iden-

tify structural shocks. We show that exploiting higher-order moment restrictions can sharpen the identification of non-Gaussian shocks achieved by imposing minimal sign restrictions. Our method thus complements recent methods introduced to address that issue by either combining sign and narrative restrictions (Antolin-Diaz and Rubio-Ramírez 2018), considering additional sign restrictions (Arias et al. 2019), or relying on structural shock proxies that can be used as external instruments (Stock and Watson 2012, Mertens and Ravn 2013, Gertler and Karadi 2015, Barnichon and Mesters 2020, Känzig 2021).

When applied to the identification of monetary policy shocks, our method delivers results that are qualitatively comparable to those obtained using several of these alternative approaches. In particular, it identifies that a tightening of monetary policy has a contractionary impact on GDP. This effect is sometimes less clear when using proxies for structural monetary shock as external instruments because these proxies can also capture other macroeconomic news and therefore fail to meet the orthogonality condition of an instrument (see, e.g. Nakamura and Steinsson 2018a, Jarociński and Karadi 2020, Miranda-Agrippino and Ricco 2021, Andrade and Ferroni 2021). Jarociński (2021) shows that US monetary surprises obtained using intraday variations in interest rate futures around Federal Open Market Committee (FOMC) announcements are leptokurtic. We document that this feature is found in a wide range of proxies for monetary policy shocks—including the recent measure by Aruoba and Drechsel (2022), which relies on natural processing language techniques to purge intraday monetary policy surprises from the potential non-Gaussianity of the Fed staff information contained in the Greenbook—and we use that robust characteristic to identify their effects. As we illustrate, our method can also help address questions for which it is more difficult or less natural to apply these alternatives methods.

Finally, our paper relates closely to Drautzburg and Wright (2023), who also exploit non-Gaussianity to sharpen the identified set of structural shocks obtained with sign restrictions. However, their use of higher-order moment restrictions differs from ours. They refine the identification set by discarding rotations that are inconsistent with the statistical independence of structural shocks and examine the matrix of higher-order moments of candidate shocks to enforce that. Their methodology rules out shocks whose higher-order moments exhibit co-skewness or co-kurtosis. By contrast, we do not impose independence; rather, we require that a specific structural shock has non-Gaussian features. Moreover, their approach requires the computing of the higher-order moments (or the marginal empirical distribution) of all the structural shocks in the system, which grows with the dimension of the VAR, therefore increasing the risk of misleading inference. Our restrictions only involve a handful of robust and non-parametric higher-order moment estimates and do not increase with the number of endogenous variables in the VAR.

2 Identification with Higher-order Moments

Let us postulate that the econometrician observes an $n-$dimensional vector of *empirical innovations*, ι_t, that are *uncorrelated* white-noise processes with unit variance, that is, $\iota_t \sim (0, I)$. Typically, these innovations will be obtained as the orthogonalized reduced-form residuals of a VAR model estimated on a long time series of data. In this section, we focus on the identification scheme and abstract momentarily from estimation and inference issues, which we discuss in Section 3. We postulate that these innovations are linear combinations of an $n-$dimensional vector of *structural shocks*, ν_t, namely,

$$\iota_t = A_o \nu_t,$$

where A_o is the true orthonormal rotation matrix and where the elements of ν_t have unit variance and non-Gaussian third- or fourth-order moments, that is, $E(\nu_{i,t}^3) = \zeta_i \neq 0$ or $E(\nu_{i,t}^4) = \xi_i \neq 3$ for some i. In this section, we further assume that $E(\nu_{i,t}\nu_{j,t}\nu_{k,t}) = 0$ for all $i \neq j, k$, $E(\nu_{i,t}\nu_{j,t}\nu_{k,t}\nu_{n,t}) = 0$ for all $i \neq j, k, m$, and $E(\nu_{i,t}^2\nu_{j,t}^2) = 1$ for all $i \neq j$. [3]

2.1 Equality Restrictions

Higher-order moments have been used in independent component analysis (ICA) to reconstruct the original (de-mixed) sources of variation of a vector of mixed signals. The core result from this literature is that if, at most, one of the components ν_t is Gaussian, A_o is point identified up to a sign change and permutation of its columns; therefore, all the structural shocks can be reconstructed from the empirical innovations (see Comon 1994, Theorem 11). This result has beend used for identification in the literature (see, for example, Gouriéroux et al. (2019) and Lanne et al. (2017)).

Regardless of the number of non-Gaussian shocks in the system considered, higher-order moment equality restrictions can be used to point identify the impact of a shock of interest as long as it is non-normally distributed. We illustrate this for the case where such a shock features non-Gaussian third- or fourth-order moments.

A spectral decomposition approach The $(n \times n^2)$ matrix collecting the third-order moments can be expressed as follows:

$$E(\nu_t \nu_t' \otimes \nu_t') = \sum_{i=1}^{n} \zeta_i J_i \otimes \boldsymbol{e}_i',$$

where $\boldsymbol{e}_i$ is a $n \times 1$ vector with a one in the i^{th} position and zeros elsewhere, J_i a $n \times n$ matrix with a one in the i^{th} position (equal to one) of the main diagonal and zeros elsewhere. The

[3] We make this assumption for exposure and analytical tractability, but our method applies to cases in which it is relaxed.

squared third-order moment matrix is thus an $(n \times n)$ diagonal matrix

$$E(\nu_t \nu_t' \otimes \nu_t')E(\nu_t \nu_t' \otimes \nu_t')' = \left(\sum_{i=1}^{n} \zeta_i J_i \otimes \boldsymbol{e}_i\right)\left(\sum_{i=1}^{n} \zeta_i J_i \otimes \boldsymbol{e}_i\right)' = \Lambda_\zeta,$$

where Λ_ζ is a diagonal matrix collecting the squared third-order moments of the structural shocks, ζ_i^2 (see appendix A.2.2 for a formal derivation). Likewise, the matrix collecting the empirical innovation third-order moments can be expressed as $E(\iota_t \iota_t' \otimes \iota_t') = A_o E(\nu_t \nu_t' \otimes \nu_t')(A_o \otimes A_o)'$. This leads to the following eigenvalue/eigenvector or spectral decomposition of the matrix collecting the squared third-order moments of the empirical innovations

$$E(\iota_t \iota_t' \otimes \iota_t')E(\iota_t \iota_t' \otimes \iota_t')' = A_o \Lambda_\zeta A_o'.$$

This spectral decomposition implies that the eigenvalue corresponds to the square of the third-order moments of the structural shock and that the corresponding unit-length eigenvector coincides with the column of the original mixing or impact matrix. Therefore, as long as the shock of interest has a nonzero third-order moment and differs from the other structural shock third-order moments, we can identify its impact on the empirical innovations. Note, however, that this requires using the full array of empirical-innovations third-order moments. Moreover, such an identification is obtained up to a sign switch and permutation of columns. This implies that one needs additional assumptions to point identify the shock of interest. For instance, one can assume that it has the largest third-order moment and therefore is associated with the largest eigenvalue in the preceding spectral decomposition.

Similar arguments carry over to fourth-order moments. In fact, we can express the $(n^2 \times n^2)$ matrix collecting the fourth-order moments *in excess of* the standard normal ones as follows:

$$E(\nu_t \nu_t' \otimes \nu_t' \otimes \nu_t) - \mathcal{K}^z = \sum_{i=1}^{n} x_i J_i \otimes J_i,$$

where x_i is the excess kurtosis of the structural shock i (that is, $x_i = \xi_i - 3$) and $\mathcal{K}^z$ is the matrix of fourth-order moments of a normal standard multivariate distribution. The right-hand side of the equation is a diagonal matrix with nonzero elements only on the positions $j(n+1) - n$ for $j = 1, ..., n$. The matrix collecting the fourth-order moments of the empirical innovations, ι_t, *in excess of* the standard normal ones can be expressed as

$$\begin{aligned} E(\iota_t \iota_t' \otimes \iota_t' \otimes \iota_t) - \mathcal{K}^z &= (A_o \otimes A_o)(E(\nu_t \nu_t' \otimes \nu_t' \otimes \nu_t) - \mathcal{K}^z)(A_o \otimes A_o)' = \\ &= P \Lambda_\xi P', \end{aligned}$$

where Λ_ξ is a diagonal matrix in which the first largest n eigenvalue corresponds to the excess kurtosis of the structural shocks (see appendix A.2.3). If one further assumes, for example, that the shock of interest has the largest fourth-order moment, they can derive the column of the original rotation matrix by dividing the first n elements of the first eigenvector by the absolute value of the first elements of the eigenvector, that is, $P(1:n,1)/\sqrt{|P(1,1)|}$.

Remarks When all the elements of ν_t are non-Gaussian, the procedure allows us to point identify all the columns in A_o up to a permutation of columns and a sign switch.[4] Conversely, when ν_t is a Gaussian distributed random vector, moments higher than second-order ones are not useful for identification, since third-order moments are zero and fourth-order moments are invariant to orthonormal rotations.[5]

Limitations While this approach allows to achieve point identification of some shocks of interest, it is also prone to three main limitations, as emphasized by Montiel Olea et al. (2022). First, the approach requires that structural shocks have zero cross third- and fourth- moments, so it is not suited for situations in which the variables in the system undergo very large movements together. Second, the method requires knowledge of the full set of higher-order moments of the empirical innovations, even when one is interested in identifying the impact of a single shock. The number of such moments grows significantly with the size of the VAR model. This is problematic, as estimates of higher-order moments can be very sensitive to outliers or minor perturbation of the data, so these might be imprecise in short samples. By using several potentially imprecise estimates of these higher-order moments, the methodology reinforces the risk of obtaining corrupted estimates of the shock's impact on observables.[6] Third, the eigenvalue decomposition identifies the shock up to a permutation of columns. This implies that one does not know where to locate one shock of interest in the set of shocks identified. Similarly, the procedure does not always imply that the shocks have an economic interpretation, since they are obtained only from statistical properties, not restrictions implied by economic reasoning.

2.2 Inequality Restrictions

Rather than achieving point-identification from a set of potentially imprecise estimates of higher-order moments of the empirical innovations, our approach relies on inequality restrictions on the higher-order moments of the structural shock itself. These restrictions are set-identifying and can therefore be combined with other set-identification restrictions.

A necessary condition for identification The starting point of our approach is a result that maps the higher-order moments of the structural shock of interest to the column that measures its impact on the empirical innovations. Without loss of generality, assume that we

[4] When using restrictions on skewness for identification, A_o comprises the eigenvectors of the $n \times n$ matrix of squared third-order moments, $E(\iota_t \iota_t' \otimes \iota_t') E(\iota_t \iota_t' \otimes \iota_t')'$. When using restrictions on kurtosis, A_o is reconstructed using the largest eigenvectors of the $n^2 \times n^2$ matrix collecting the fourth-order moments of the empirical innovations, $E(\iota_t \iota_t' \otimes \iota_t' \otimes \iota_t)$, by applying the procedure outlined in Kollo (2008). See appendix A.2.4 for more details.

[5] See Appendix A.2.1

[6] We illustrate this point in section 4.2 using simulations from an estimated standard macro New Keynesian model.

are interested in the last (n^{th}) shock, and let $\boldsymbol{\alpha}_n$ be the rightmost (n^{th}) columns of A_o (the true impact matrix), which measure the impact of the structural shocks on the empirical innovations.

Proposition 1 *Let $\breve{\nu}_{n,t} = \boldsymbol{a}_n' \iota_t$ be a candidate structural shock, with $\boldsymbol{a}_n$ a unit-length vector of weights. Let $\boldsymbol{\alpha}_n$ denote the "true" impact of the structural shock on the empirical innovations. When $\boldsymbol{a}_n = \boldsymbol{\alpha}_n$, then the candidate shock ($\breve{\nu}_{n,t}$) has the same higher-order property as the "true" structural shock.*

The proof is provided in Appendix A.2.

The proposition offers a necessary condition for identification. Consider a rotation matrix and an associated candidate structural shock $\breve{\nu}_{n,t}$ that leads to different higher-order moments than the structural shock; then, according to the proposition, it must be the case that $\boldsymbol{a}_n \neq \boldsymbol{\alpha}_n$. In other words, it is impossible to obtain the true rotation matrix without generating the correct higher-order moments. This implies that one can discard all the rotations that are inconsistent with the conditions imposed on the higher-order moments of the structural shock because these rotations are inconsistent with the true impact of the structural shock on the empirical innovations. In particular, the set of admissible rotations can be reduced by imposing the desired higher-order moments to lie in a preassigned interval. When the underlying DGP is consistent with these inequality restrictions—that is, the true structural shock exhibits deviation from Gaussianity—higher-order moment inequality restrictions shrink the set of rotations.

Conversely, if the true moments of the structural shocks of interest are inconsistent with the restriction, then the set of admissible rotations is empty. For instance, if we assume that monetary policy shocks are leptokurtic while they are not in the true data-generating process, no rotation will meet the restrictions. It is important to highlight that similar considerations also apply to sign, magnitude, or narrative restrictions; if we impose "incorrect" restrictions on, say, the impulse response functions (IRFs) of the structural shocks, either the set of accepted rotation is empty or the identification of the shock is corrupted because it does not reflect the sign pattern of the DGP.

Restrictions considered While we could consider a large set of restrictions, we focus on restrictions on the third- and fourth-order moments of the distribution of the structural shock. Let $\mathcal{S}(x)$ and $K^{\star}(x)$ denote estimators of, respectively, the skewness and the excess kurtosis of a variable x. The first type of restriction is on the shape of the distribution of the structural shock, which can be assumed to feature either asymmetry, fat tails, or both as follows:

Restriction 1 (Asymmetry) *The structural shock x features a longer tail on the right (left) side of the distribution such that its skewness is greater than a positive threshold, that is, $\mathcal{S}(x) > \epsilon_{\mathcal{S}}$ ($\mathcal{S}(x) < \epsilon_{\mathcal{S}}$).*

Restriction 2 (Fat-tailedness) *The structural shock x features tails thicker than the normal distribution such that its excess kurtosis is greater than a positive threshold, that is, $K^\star(x) > \epsilon_\mathcal{K}$.*

These restrictions can be combined in situations in which the structural shock is assumed to have both an asymmetric and leptokurtic distribution. Alternatively, one can postulate that the structural shock is non-Gaussian without specifying whether the deviation stems from its third- or fourth-order moment as follows:

Restriction 3 (Non-gaussianity) *The structural shock x is **not** distributed as a Gaussian distribution if the absolute values of skewness or of excess kurtosis are positive, that is,$|\mathcal{S}(x)| > \epsilon_\mathcal{S} > 0$ or $|K^\star(x)| > \epsilon_\mathcal{K} > 0$.*

Finally, another natural restriction involves ranking the third- or fourth-order moment of the shock of interest as follows:

Restriction 4 (Ranking) *The structural shock x_1 has the largest absolute value of skewness or excess kurtosis (or both) among all the shocks in the system, that is, x_j with $j = 2, .., n$, so that we have for each $j \neq 1$*

$$|\mathcal{S}(x_1)| > |\mathcal{S}(x_j)|$$
$$|K^\star(x_1)| > |K^\star(x_j)|.$$

How to specify higher-order moment inequality restrictions in practice By definition, one does not observe structural shocks. The higher-order restrictions are structural identification assumptions that can be grounded on either theoretical insights or empirical evidence using proxies for unobserved structural shocks. The following applications provide examples using each type of motivation.

Once the type of restriction is chosen, the question is how to choose the thresholds of the inequality restrictions, that is, how much deviation from the normal deviation is postulated. For skewness, a reasonable cutoff point is $\epsilon_\mathcal{S} = 0.1$. This means that median (mode) and mean are at least 10 (30) percent apart. For kurtosis, setting $\epsilon_K = 1.5$ amounts to assuming that realizations larger than three standard deviations should occur at least once every 100 observations (as opposed to roughly one every 700 observations as in the Gaussian case). These numbers provide some arbitrary references and, in general, the choice of the threshold depends on how much non-Gaussianity a researcher is willing to impose. Alternatively, one can rely on empirical estimates of the shock of interest provided in other studies (for example, see section 5.1).

Note that the modeler faces a tradeoff when setting these bounds. Too restrictive bounds (too large deviations from Gaussiantity) can return an empty set, and too loose bounds may lead to very imprecise identification.

Advantages Our method overcomes each of the three limitations outlined earlier. First, it does not require to assume that the shocks are independent. In particular, it does not require to impose zero co-skewness or co-kurtosis across the structural shocks (although we make this assumption for tractability in the previous section). Second, the method only requires to estimate the third- and/or fourth-order moment of the structural shock of interest and allows to use robust estimators of these moments based on the percentile of the empirical distribution of the shock,[7] thus avoiding having to compute a large number of sensitive sample third- or fourth-order moments, which can lead to biased estimates. That risk is further mitigated by the fact that we impose an inequality, rather than equality, restriction, on the third- or fourth-order moment of the shock. Third, inequality restrictions achieve set identification as opposed to point identification. Rather than being a drawback, this allows us to combine these restrictions with other assumptions such as signs, zeros, narrative, magnitude, and/or statistical independence. In this sense, higher-order moment inequality restrictions are *modular*, as they can be used flexibly, either standalone or in combination with other assumptions derived from economic analysis.

2.3 An Analytical Example

We conclude this section with an analytical example illustrating the main point of the paper, that inequality restrictions on higher-order moments of the structural shock can shrink the identified set of the structural parameter. To this end, we assume that $n = 2$ and that $\nu_{1,t} \sim (0, 1)$ and $\nu_{2,t} \sim (0, 1)$. Moreover, we assume that the third-order moments of the structural shocks are zero and one, respectively, that is, $E(\nu_{1,t}^3) = 0$ and $E(\nu_{2,t}^3) = 1$, and cross second- and third-order moments are zero, that is, $E(\nu_{1,t}\nu_{2,t}) = E(\nu_{1,t}\nu_{2,t}^2) = E(\nu_{1,t}^2\nu_{2,t}) = 0$. The structural equations are defined by

$$\iota_{1,t} = \cos\theta_o \nu_{1,t} - \sin\theta_o \nu_{2,t},$$
$$\iota_{2,t} = \sin\theta_o \nu_{1,t} + \cos\theta_o \nu_{2,t},$$

where θ_o is the true unknown angle of rotation with $\theta_o \in (-\pi/2, \pi/2)$ to ensure that the terms of the main diagonal of the impact matrix A_o are positive. We assume $\theta_o \neq 0$, else the problem is trivial. The econometrician does not observe ν_t and only observes ι_t. While first- and second-order moments do not depend on θ_o, third-order moments do. In particular, the population third-order moments of the observed empirical innovations are given by $E(\iota_{1,t}^3) = -\sin^3\theta_o$, $E(\iota_{2,t}^3) = \cos^3\theta_o$, $E(\iota_{1,t}^2\iota_{2,t}) = \sin^2\theta_o\cos\theta_o$, and $E(\iota_{1,t}\iota_{2,t}^2) = -\sin\theta_o\cos^2\theta_o$. Hence, the matrix

[7] See also appendix A.4 for a description of robust estimators of third- and fourth-order moments.

of the third-order moments of the empirical innovation is given by

$$E(\iota_t\iota_t' \otimes \iota_t') = \underbrace{\begin{pmatrix} \sin^2\theta_o & -\sin^2\theta_o\cos\theta_o \\ -\sin\theta_o\cos\theta_o & \cos^2\theta_o \end{pmatrix}}_{\Omega} \otimes \begin{pmatrix} -\sin\theta_o & \cos\theta_o \end{pmatrix}.$$

Since Ω is an idempotent matrix, it holds that $E(\iota_t\iota_t' \otimes \iota_t')E(\iota_t\iota_t' \otimes \iota_t')' = \Omega$. The characteristic polynomial of Ω is $(\sin^2\theta_o - \lambda)(\cos^2\theta_o - \lambda) - \sin^2\theta_o\cos^2\theta_o$, and the associated eigenvalues are zero and one, respectively. This implies that the first structural shock has a third-order moment equals to zero and that the second structural shock has a third-order moment equals to one. The eigenvector associated with the nonzero eigenvalue is $\begin{pmatrix} -\sin\theta_o & \cos\theta_o \end{pmatrix}'$.

Our preferred approach does not directly use the third-order moments of the empirical innovations and imposes weaker restrictions. We assume that the second shock has positive skewness; moreover, we do not make assumptions about the third-order moment of the other shock. Yet, we can considerably restrict the identified set. More formally, we only consider the set of rotations such that the following inequality is verified:

$$E(\breve{\nu}_{2,t}^3) > 0.$$

Let A be a generic rotation with angle θ, that is, $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, and let $\breve{\nu}_t = A'\iota_t$. The corresponding population moment is then

$$\begin{aligned} E(\breve{\nu}_{2,t}^3) &= E(-\sin\theta\iota_{1,t} + \cos\theta\iota_{2,t})^3 = [\sin\theta\sin\theta_o + \cos\theta\cos\theta_o]^3 \\ &= [\text{sign}(\cos\theta_o)\cos(\theta - \theta_o)]^3 > 0. \end{aligned}$$

Since $\text{sign}(\cos\theta_o) = 1$ for all $\theta_o \in (-\pi/2, \pi/2)$, the latter is positive whenever $\cos(\theta - \theta_o) > 0$, which occurs when in the region of points where[8]

$$\max\{-\pi/2, \theta_o - \pi/2\} < \theta < \min\{\pi/2, \theta_o + \pi/2\}.$$

This condition allows to shrink the set of admissible rotations that otherwise would be the set $\theta \in (-\pi/2, \pi/2)$; specifically, when θ_o is positive (negative), the lower (upper) bound shrinks. There is a discontinuity at $\theta_o = 0$. In such a case, the first (second) structural shock coincides with the first (second) empirical innovation, all third-order moments are zero except for the third-order moment of the second empirical innovation, $E\iota_{2,t}^3$, and there is no point in considering them as a system of mixed signals.

3 Bayesian Estimation and Identification

We postulate that the observed data are generated by a $VAR(p)$ model,

$$y_t = \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + \Phi_0 + u_t,$$

[8] More details on the solution can be found in the appendix A.3.1.

where y_t is $n \times 1$ vector of endogenous variables, Φ_0 is a vector of constant and Φ_j are $n \times n$ matrices, and, given initial conditions, $y_0, \ldots, y_{-p+1}$. We assume that u_t is an $n-$dimensional white-noise process with unconditional covariance matrix Σ, which is obtained as a linear combination of the unobserved structural shocks, ν_t, that is,

$$u_t = \Sigma^{1/2} \iota_t = \Sigma^{1/2} \ \Omega \ \nu_t,$$

with Ω an orthonormal matrix, $\Omega\Omega' = \Omega'\Omega = I$.

Inference when shocks are non-Gaussian Standard inference on VAR parameters typically postulates a multivariate normal distribution for the reduced-form innovations. However, such an assumption is not valid in our context. We adopt a robust Bayesian approach that allows to construct posterior credible sets without assuming a specific distribution of the reduced-form residuals.

Our approach builds on the work of Petrova (2022), who proposes a robust and computationally fast Bayesian procedure to estimate the reduced-form parameters of the VAR in the presence of non-Gaussianity. In that configuration, the Quasi Maximum Likelihood (QML) estimator[9] of the VAR parameters is consistent and asymptotically normal. However, inference on the intercept is affected when innovations have a nonsymmetric distribution; and inference on Σ is affected when the innovation distribution shows excess kurtosis relative to the multivariate normal density. Petrova (2022) derives closed-form expressions for corrected Bayesian posterior distributions that rely on the asymptotic covariance matrix of the QML estimator. Combined with a prior, one can draw the VAR reduced-form parameters from such corrected posterior distribution without specifying the shock distribution and allowing for asymmetries and fat tails (or thin tails).[10]

Bayesian procedure Assume that we are interested in identifying the last shock, $\nu_{n,t}$. Let $\Sigma^{(j)}$ and $\Phi^{(j)}$ be the j^{th} draw from the reduced-form parameter posterior distribution obtained using the above approach. The procedure for identifying the structural shock using higher-order restrictions proceeds as follows:

I. Draw $\breve{\Omega}$ from a uniform distribution with the Rubio-Ramírez, Waggoner and Zha (2010) algorithm;

II. Compute the impulse-response function and check if the sign or any other economic restrictions are verified;

[9]The QML is the maximum estimator of the quasi-likelihood, which, in this context, coincides with the likelihood of the VAR when *incorrectly* assuming normality of the reduced-form residuals.

[10]See appendix A.5 for more details.

III. Compute the implied structural shocks

$$\breve{\nu}_t^{(j)} = \breve{\Omega}' \left(\Sigma^{(j)}\right)^{-1/2} (y_t - \Phi_1^{(j)} y_{t-1} - ... - \Phi_p^{(j)} y_{t-p} - \Phi_0^{(j)});$$

IV. Compute $\mathcal{S}(\breve{\nu}_{n,t}^{(j)})$ and/or $\mathcal{K}(\breve{\nu}_{n,t}^{(j)})$ and check if the higher-order moment inequality restrictions are satisfied.

If both [II] and [IV] are satisfied, keep the draw $\Omega^{(j)} = \breve{\Omega}$. Then repeat [I] to [IV].

After a suitable number of iterations, the draws represent the posterior distribution of the impulse responses of interest. The estimation of the reduced-form parameters and the computation of the impulse responses using the higher-order moments are performed using the toolbox described in Ferroni and Canova (2021).[11]

Robust higher-order moment estimators Our procedure requires to estimate third- and fourth-order moments. As illustrated earlier, methods using higher-order moment equality restrictions to identify structural shocks typically require parametric estimates of such moments. A drawback is that these estimates are known to be very sensitive to outliers with short samples (see, e.g., Kim and White 2004). By contrast, our approach allows to rely on robust nonparametric estimators of these moments which mitigates that issue.

These robust estimators are constructed using the empirical distribution of the structural shocks of interest that are obtained in step III of the preceding procedure. For third-order moments or asymmetries, robust measures exploit the standardized distance between median and mean. For fourth-order moments, or tailedness, robust estimators consider the ratio between the distance of the percentiles in the tails and the distance between the percentiles close to the median. For example, the larger the numerator, the thicker the tails; the smaller the denominator, the more clustered the distribution around the median and hence the more leptokurtic the shock. We use the following estimators of skewness ($\mathcal{S}$) and excess kurtosis ($\mathcal{K}^\star$) of a random variable x:

$$\mathcal{S}(x) = \frac{\bar{x} - F^{-1}(0.5)}{\text{std}(x)}, \quad \mathcal{K}^\star(x) = \frac{F^{-1}(0.975) - F^{-1}(0.025)}{F^{-1}(0.75) - F^{-1}(0.25)} - 2.9, \tag{1}$$

where $F^{-1}(\alpha)$ is the α-percentile of the empirical distribution of x, and 2.9 is the value of the kurtosis of the Gaussian distribution. We discuss several alternative robust estimators in the appendix (section A.4).

4 How Higher-Order Moment Inequality Restrictions Help Identify Structural Shocks?

We illustrate how inequality restrictions on higher-order moments can improve the identification of the effects of standard economic shocks on macro variables. In particular, we focus on the

[11] Codes for replication can be found on the Github page.

long-debated impact of monetary policy on output.

Uhlig (2005) finds that monetary policy shocks have no clear effect on output if one uses an agnostic identification procedure. More precisely, he imposes sign restrictions on inflation and the interest rate (moving in opposite directions) but remains *agnostic* about the response on output. Wolf (2020) shows that Uhlig (2005)'s result is consistent with a standard New Keynesian model because, when only sign restrictions on inflation and the interest rate are imposed, supply and demand shocks tend to *masquerade* as monetary policy shocks. He concludes that pure sign restrictions are quite weak in identifying information. Identification can be improved with instruments, as in Gertler and Karadi (2015), or restriction on the reaction coefficients of the policy function, as in Arias et al. (2019).

In this section, we show that higher-order moment inequality restrictions also strongly help resolve the issue of shocks *masquerading* as monetary ones when the latter have a non-Gaussian fourth-order moment that is sufficiently different from supply and demand shocks. We first use simulated data from a standard calibration of the three-equations New Keynesian model, as in Wolf (2020), but where the monetary policy shock is assumed to follow a Laplace distribution and therefore exhibits excess kurtosis compared with the Gaussian distribution. We then use simulations from a more realistic DSGE model featuring a variety of shocks and frictions, in the same spirit as Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007), and that we estimate using US post-World War II data. Strikingly, the distribution of estimated monetary policy shock exhibits excess kurtosis, and more so than the other estimated structural shocks.

4.1 Experimental Evidence Using a Textbook New Keynesian Model

Specification, calibration, and simulation As detailed in Galí (2015), the model in its log-linearized form is described by three equations:

$$\begin{aligned} y_t =& y_{t+1|t} - (i_t - \pi_{t+1|t}) + \sigma_d \nu_t^d, \\ \pi_t =& \beta \pi_{t+1|t} + \kappa y_t - \sigma_s \nu_t^s, \\ i_t =& \phi_\pi \pi_t + \phi_y y_t + \sigma_m \epsilon_t^m, \end{aligned}$$

with y real output, i the nominal interest rate (the federal funds rate), and π inflation. The model has three structural disturbances: a demand shock ν_t^d, a supply shock ν_t^s, and a monetary policy shock ϵ_t^m. The first equation is a standard IS relation (demand block), the second equation is the New Keynesian Phillips curve (supply block), and the third equation is the monetary policy

rule (policy block). This benchmark model yields the following closed-form solution:

$$x_t = \begin{pmatrix} y_t \\ \pi_t \\ i_t \end{pmatrix} = \frac{1}{1+\kappa\phi_\pi+\phi_y} \begin{pmatrix} \sigma_d & \phi_\pi\sigma_s & -\sigma_m \\ \kappa\sigma_d & -(1+\phi_y)\sigma_s & -\kappa\sigma_m \\ (\phi_y+\kappa\phi_\pi)\sigma_d & -\phi_\pi\sigma_s & \sigma_m \end{pmatrix} \begin{pmatrix} \nu_t^d \\ \nu_t^s \\ \epsilon_t^m \end{pmatrix} = A_o\ \nu_t.$$

We consider standard Gaussian supply and demand shocks but leptokurtic monetary policy shocks, which are assumed to follows a Laplace distribution (whose excess kurtosis is three):

$$\nu_t^d \sim N(0,1),\ \nu_t^s \sim N(0,1),\ \epsilon_t^m \sim Laplace(0,1).$$

We assign the following values to the parameters $\sigma_s = \sigma_d = \sigma_m = 1$, $\phi_\pi = 1.5$, $\phi_y = 0.5$, and $\kappa = 0.2$. We then simulate a long time-series of data, $T = 100,000$, compute the sample covariance of the data, and generate candidate rotations using the Haar prior.

Masquerading shocks The masquerading effect of Wolf (2020) is illustrated in the left panel of figure 1. It displays a scatter plot of all the supply and demand shock realizations in our simulations and highlights, in red circles, realizations that generate a negative co-movement between inflation and the interest rate. That latter configuration occurs when supply shocks are relatively small compared with demand shocks and when supply and demand shocks have the same sign. In that area, the combination of, say, positive supply and demand shocks leads to a decline in inflation (as supply dominates), an increase in output (as supply and demand reinforce), and an increase in the interest rate induced by the policy rule as the increase in output dominates the decline in inflation. Overall, this combination of supply and demand shocks generates a pattern that looks like a contractionary monetary policy shock if one only looks at the correlation between inflation and the interest rate. Hence, they masquerade as a monetary policy shock. However, unlike what a contractionary monetary policy shock implies, output goes up due to the expansionary supply and demand shocks. As a result, imposing sign restrictions on only inflation and the interest rate fails to capture the contractionary impact of monetary policy shocks on output that the model implies. This is illustrated in the right panel of figure 1,where the blue bars represent the distribution of the impact on output of monetary policy shocks identified with only sign restriction on our simulated data. Strikingly, although the true impact implied by the model nears $-.6$, the agnostic identification scheme of Uhlig (2005) returns an estimated impact distributed symmetrically around zero in a range of approximately $-.9$ to .9.

Combining sign and excess kurtosis restrictions We now complement the same sign restrictions with a moment inequality restriction on the excess kurtosis of monetary policy shock. In particular, we postulate that the excess kurtosis of monetary policy shocks needs to be larger than a threshold value, that is, $\mathcal{K}_m^\star > \epsilon_\mathcal{K}$, with $\epsilon_\mathcal{K} = 1.5$. As illustrated by the red bars in the

top-right panel of figure 1, the estimated impact distribution changes drastically compared with estimates obtained with sign restrictions only: Almost all the estimated positive responses of output get cut away.

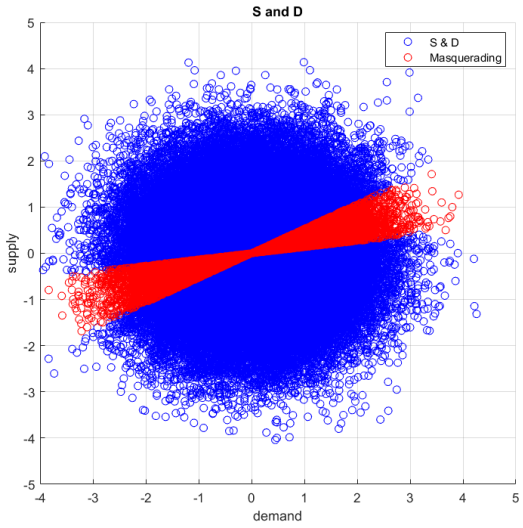
The lower panel of figure 1 shows (blue line) how the probability of obtaining an estimated response of output that is positive (instead of negative in the model) varies with different values of the lower bound imposed on the fourth-order moment of the shock, $\epsilon_{\mathcal{K}}$. The figure also reports the acceptance rate when only considering sign restrictions (grey line) and the acceptance rate when using signs and excess kurtosis restrictions (red line). When $\epsilon_{\mathcal{K}} = 0$, about one-fourth of rotations are accepted with sign only and with sign and kurtosis restrictions, respectively; the probability that the response of output is positive is about one-half. As we increase the minimum value for the excess kurtosis, fewer rotations verify both the signs and the excess kurtosis restrictions, and the probability of a positive response of output to a monetary policy shock declines. It eventually becomes zero when we require the excess kurtosis of monetary policy shocks to be larger than 1.8. Thus, the masquerading effect disappears when we impose a clear deviation from the Gaussian case but only about 60 percent of the true excess kurtosis of the shock.

Identified sets We also compute the identified sets resorting to numerical methods rather than to simulations. The results are very similar to the earlier findings and are presented in section A.6.1.

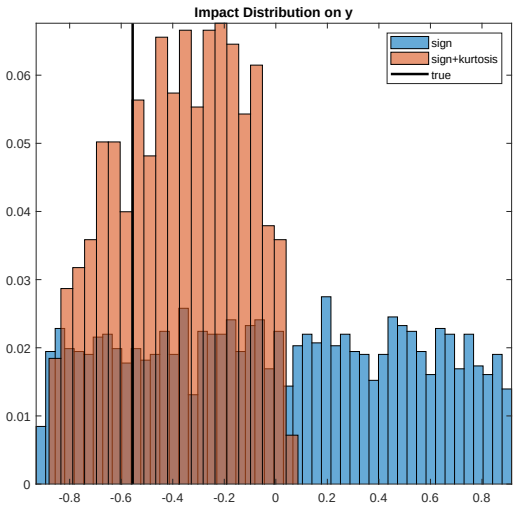
4.2 Experimental Evidence Using a Medium-scale DSGE model

We now study how higher-order moment inequality restrictions perform when applied to data obtained from simulations of a more realistic model that features a number of real and nominal frictions and for which we do not specify the underlying distribution of the structural shocks. More specifically, we work with the Smets and Wouters (2007) (SW) model, perhaps the most well-known example of an empirically successful structural business-cycle model. In this model, the fluctuations in economic activity, labor market variables, prices, and interest rates are explained by handful of shocks: technology, risk premium, investment demand, monetary policy, government/exogenous spending, and price and wage markup shocks. We estimate the model on US post World War II data for output, consumption, investment, real wages, inflation, interest rates, and hours worked, as in Smets and Wouters (2007).[12] We set parameters at their posterior mode estimates and consider smoothed estimates of the structural shocks as realizations of

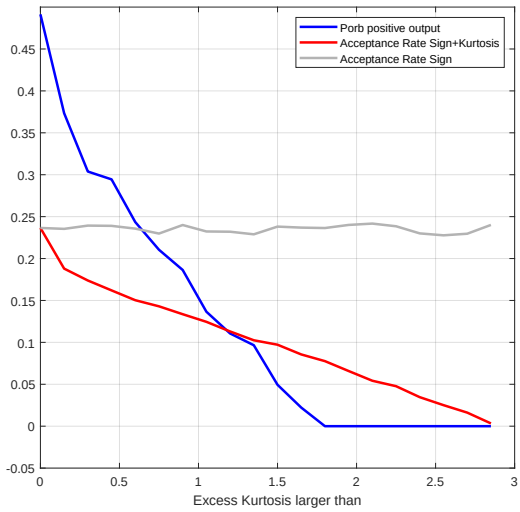
[12]Our implementation of the SW model is based on Dynare's (see Adjemian, Bastani, Juillard, Karamé, Maih, Mihoubi, Perendia, Pfeifer, Ratto and Villemot (2011)) replication code, kindly provided by Johannes Pfeifer. The code is available at https://sites.google.com/site/pfeiferecon/dynare.



(a) Realizations of supply and demand shocks: all (blue circles) and masqueraded monetary policy (red circles).



(b) Distribution of monetary policy impact on output with sign and with moment and sign restrictions.



(c) Probability of positive response of y at different restrictions on monetary policy excess kurtosis.

Figure 1: Masquerading effect: (a) distribution of the impact of monetary policy shock on output; (b) distribution of excess kurtosis for the masquerading shock; and (c) probability of positive response of y after a monetary policy tightening.

these shocks. We then use this model as a laboratory to show how the higher-order moments can sharpen identification.

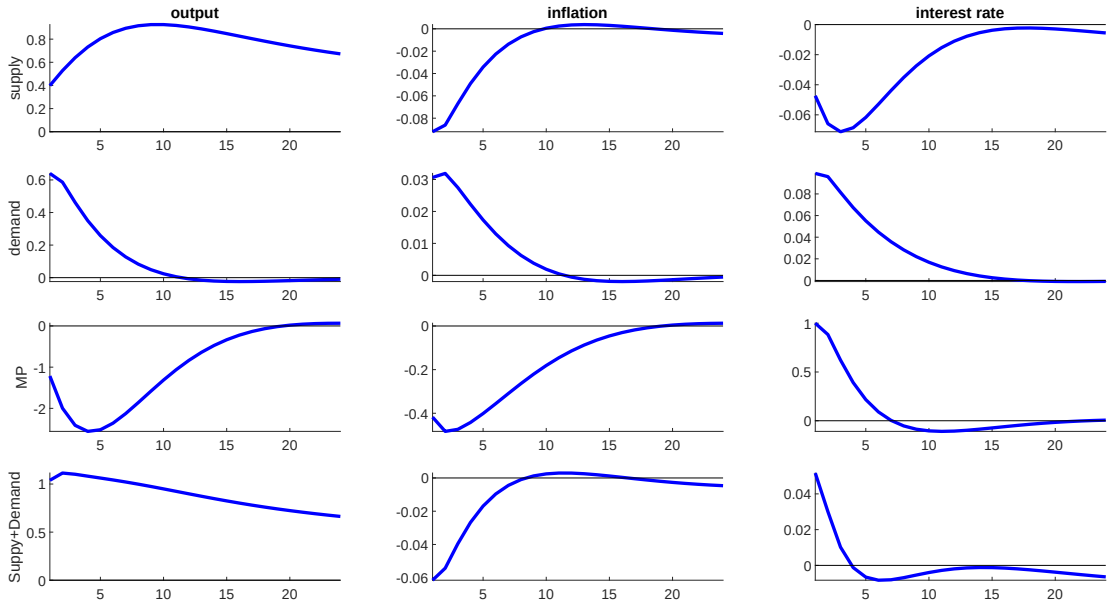


Figure 2: SW model estimates of impulse response functions. From top to bottom: technology, risk-premium, and monetary policy shocks and the sum of supply and demand shocks.

Supply and demand shocks can masquerade as monetary policy shocks A first property of the estimated SW model is that, as for the textbook New Keynesian model, supply and demand shocks can masquerade as a monetary policy shock. This is illustrated in figure 2, which reports—from top to bottom—the impulse response of technology (supply), risk-premium (demand), and monetary policy shocks to output, inflation, and interest rates, respectively, in the SW model; the last row displays the dynamic transmission of the sum of supply and demand. Comparing this row with the monetary policy one shows that the combination of supply and demand shocks can generate a sign pattern for inflation and the interest rate that is similar to the pattern resulting from a monetary policy shock. However, while the output reaction to a monetary policy contraction is negative, the masquerading combination of supply and demand shocks has a positive impact on output. As a result, Uhlig (2005)'s agnostic sign restrictions may not be enough to identify the transmission of monetary policy shocks to output.

Estimated monetary policy shocks are leptokurtic A second striking property of the estimated SW model is that structural shocks display some form of fat-tailedness, and even more so for monetary policy shocks. This is illustrated in figure 3, which reports the estimated shock realizations (top panels) and the empirical probability distribution against the normal distribution (bottom panels). All estimated shocks display some deviations from the Gaussian distribution. In particular, when we compute a robust measure of kurtosis using equation (1),

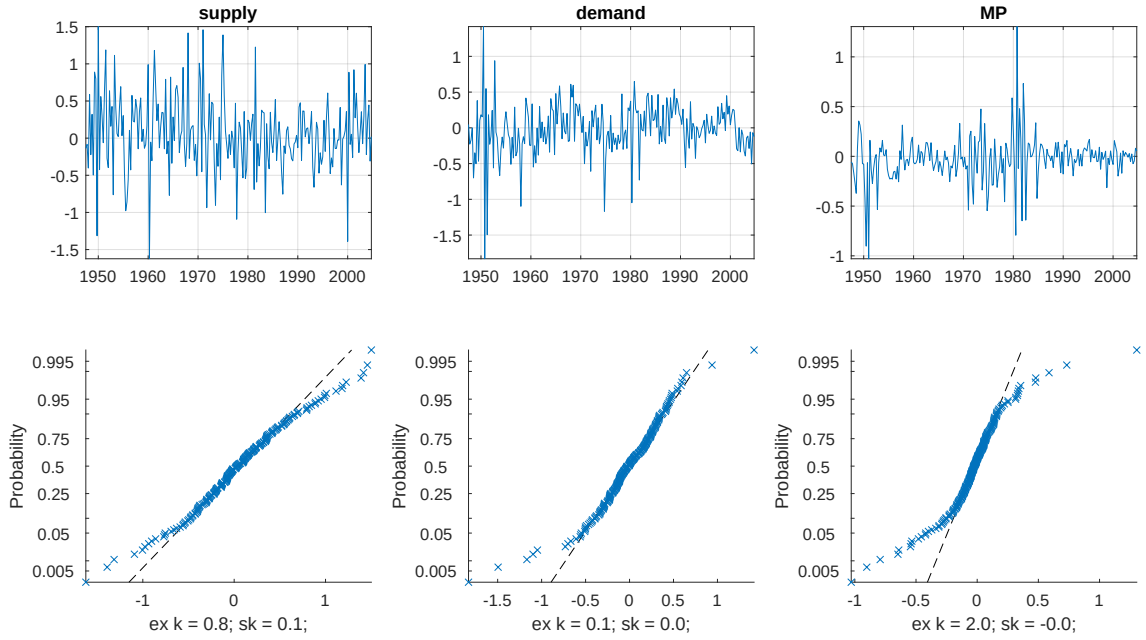


Figure 3: SW-model estimated shocks, from left to right: technology, risk-premium, and monetary policy shocks. Top panels: realizations; bottom panels: probability distribution against the normal distribution.

we find an excess kurtosis for the SW-model monetary policy shock equal to two, well above the excess kurtosis for other shocks. By contrast, the robust measure of skewness is very close to zero for every structural shock.[13]

Identifying monetary policy shocks in large samples We generate a sample of 50,000 observations for output, inflation, and interest rates using random draws from the empirical distribution of the technology, risk-premium, and monetary policy shocks. We then assess how different SVAR restrictions aimed at identifying monetary policy shocks perform on these simulated data. We compare three schemes: *(i)* sign restrictions only, *(ii)* sign and higher-order moment inequality restrictions, as introduced in section 2.2, and *(iii)* higher-order moment equality restrictions using the spectral decomposition approach of section 2.1.

In *(i)*, we postulate that, after a contractionary monetary policy shock, the interest rate increases on impact and for two consecutive quarters, and inflation decreases on impact and for two consecutive quarters. In the spirit of Uhlig (2005), no restriction is imposed on output. In *(ii)*, we impose the additional restriction that monetary policy shocks are leptokrutic, with

[13]These results are consistent with the evidence in Cúrdia, Negro and Greenwald (2014). They estimate a version of the SW model that allows for Student$-t$ distributed shocks and stochastic volatility. They show that the estimation on US data clearly favors shocks with leptokurtic distribution.

a robust measure of excess kurtosis larger than 1.6.[14] We estimate a VAR with 20 lags and rotate the reduced-form VAR innovations so that the sign restrictions in *(i)* are satisfied. We construct 50,000 accepted rotations using the Haar prior (see Rubio-Ramírez et al. 2010). For each of these rotations, we verify if the additional higher-order moment inequality restriction spelled out in *(ii)* is also verified; if so, we keep the rotation. Finally, in *(iii)*, the rotation matrix is derived from the spectral decomposition of the matrix of fourth-order moments of the VAR empirical innovations.

Figure 4 reports the impulse responses obtained with each identification scheme. The first row reports the 90 percent and 99 percent identified sets of impulse responses obtained when using sign restrictions only, and the second row when combining sign and kurtosis inequality restrictions. The red dashed line shows the point-identified impulse responses obtained when using the spectral decomposition of the matrix of fourth-order moments of the VAR empirical innovations. Finally, the blue line displays the true monetary policy impulse responses derived from the model. As in the simple New Keynesian model case, the set of responses identified by sign restrictions is so wide that the impact of monetary policy shock on output is inconclusive. On the contrary, higher-order moments help refine the transmission of monetary policy shocks to output. In the case of inequality restrictions, the 90 percent identified set suggests that most of the output trajectories are negative at least after a few quarters, in line with theoretical predictions.[15] Notably, the identification obtained with the spectral decomposition of the empirical innovation fourth-order moment matrix is very precise. However, with shorter samples, such precision deteriorates, and inequality restrictions perform better, as we show in the next section.

Identified sets We also computed numerically the identified sets obtained when the different identification schemes discussed earlier are applied to the SW model. The results are presented in section A.6.1 and are very similar to the findings obtained with simulations reported previously.

Identifying monetary policy shocks in short samples Identification based on spectral decomposition is very precise when working with very large data samples. However, the precision deteriorates sharply when working with smaller samples of sizes typical of macroeconomic studies. Indeed, this approach requires to compute all the fourth-order moments of the

[14] That threshold value is chosen to be larger than about 80 percent of the value estimated in the empirical distribution. Thresholds lower than one do not alter the identified set compared with imposing sign restrictions only. As appendix figure 15 illustrates, the probability that output is positive one year after a monetary policy tightening is about 40 percent both with sign and with sign and higher-order moment restrictions. This occurs because both supply and demand are somehow leptokurtic; however, their tails are not fat enough to masquerade as a monetary policy shock with leptokurticity larger than 1.3. With such a threshold, we start to see the higher-order moment identified set shrink relative to the sign-restricted identified set.

[15] Higher-order moment inequality restrictions also lead to more precise identified sets when one imposes the sign restrictions that a monetary policy shock generates on the observed variables; appendix figure 14 reports the magnitude of the refinement of the identified set in the context of the SW model.

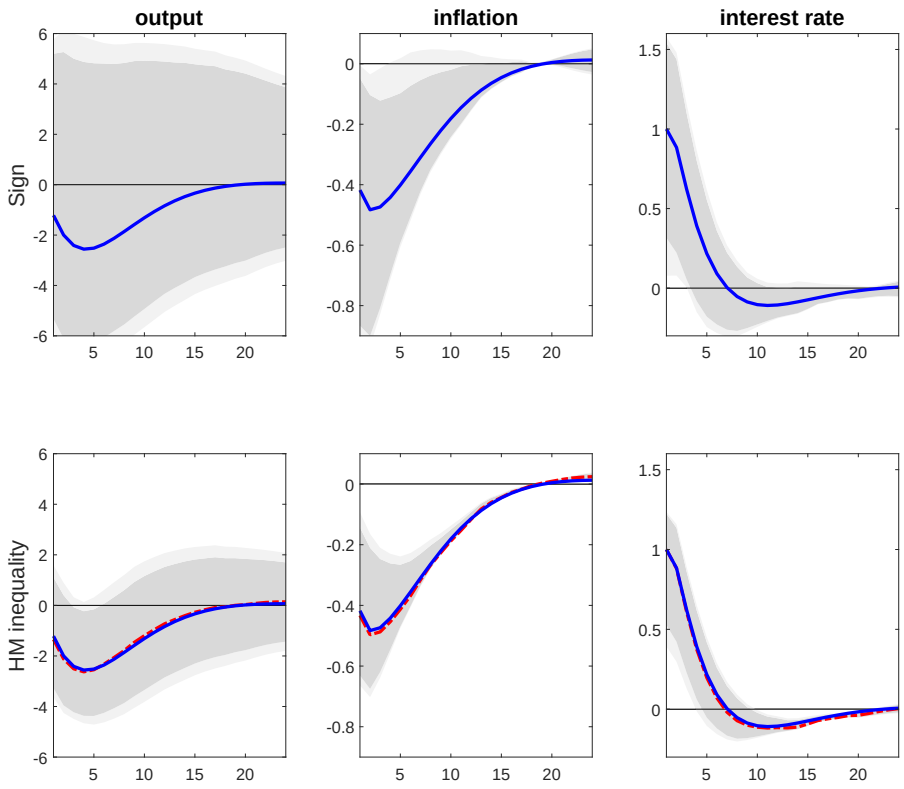


Figure 4: Impulse responses to a monetary policy shock using sign restrictions on prices and interest rates (first row) and sign and higher-order moment inequality restrictions (second row). The blue solid line is the true impulse response. The red dashed line indicates the point identification obtained with spectral decomposition of the fourth-order moments of empirical innovations. The dark (light) gray areas report the 90 percent (99 percent) confidence bands. Reduced-form parameter VAR estimates are based on an artificial sample of 50,000 observations.

VAR empirical innovations, and the sample counterparts of these fourth-order moments can be poorly estimated and very sensitive to in-sample outliers. By contrast, our identification method only involves estimating the fourth-order moment of the shock of interest and allows us to work with robust estimates of that higher-order moment. It also performs well with smaller samples.

To illustrate this point, we simulate 500 artificial samples of 200 observations of output, inflation, and interest rates. For each sample, we estimate a VAR and compute the monetary policy impulse responses using the spectral or eigenvalue decomposition and the sign and higher-order moment inequality restrictions. Figure 5 reports the dispersion of the (median) point estimate across identification schemes. With short samples, the spectral decomposition generates very dispersed point estimates, which include positive responses of output and prices after a monetary policy shock. This does not occur with our identification scheme, in which point

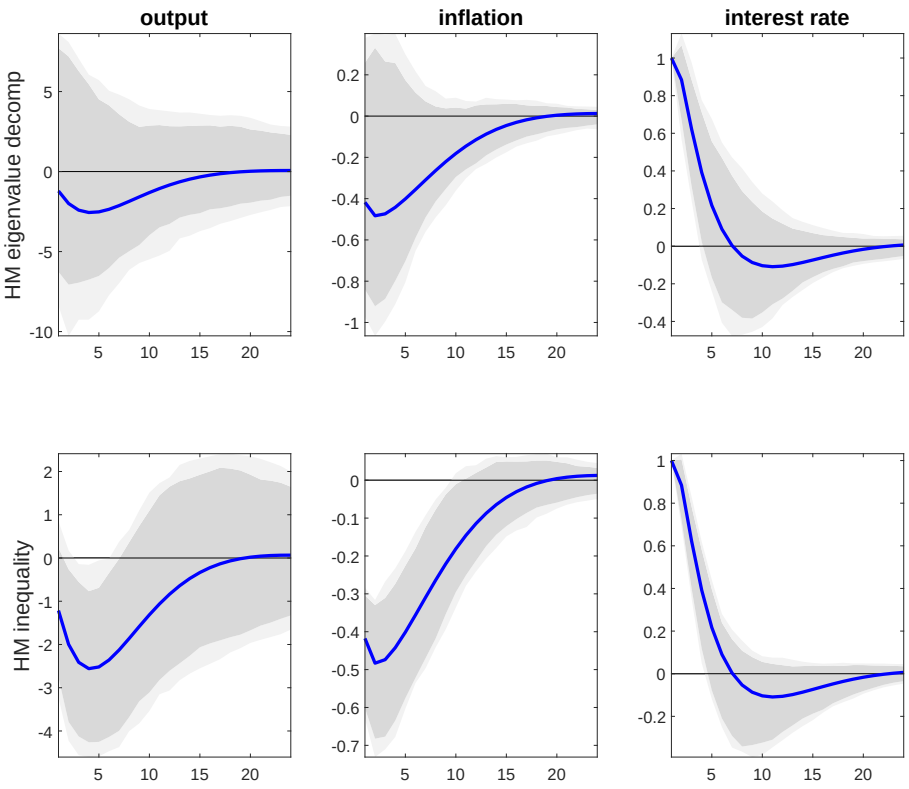


Figure 5: Impulse responses to a monetary policy shock using fourth-order moment spectral decomposition (top row) and sign and higher-order moment inequality restrictions (bottom row). The dark (light) gray areas report the 90 percent (99 percent) dispersion of the median impulse response estimates across 500 artificial samples of 200 observations. The blue line is the true impulse response.

estimates of the responses of output and inflation are estimated to be negative.

Not only are the median impulse response functions better estimated, but the uncertainty around the central tendency is smaller. Figure 6 reports the median value of the upper and lower bounds of the 68 percent high probability density (HPD) sets across artificial samples obtained with sign restrictions only (top row), higher-order moment spectral decomposition (middle row), and sign and higher-order moment inequality restrictions (bottom row). The bounds of the HPD set response of output are negative with the sign and higher-order moment inequality restrictions. This is not the case with the identification obtained through spectral decomposition or sign restrictions only.

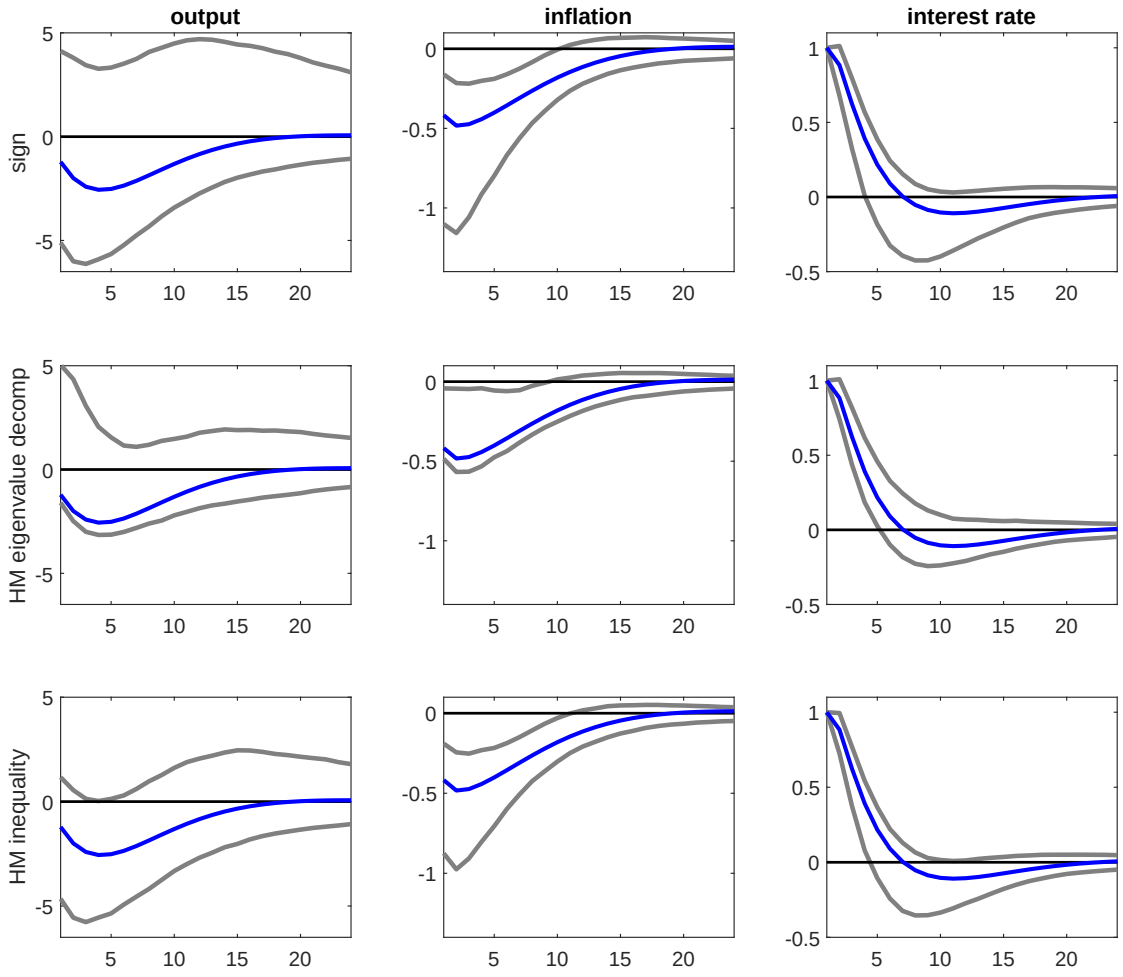


Figure 6: The median upper and lower bounds of the 68 percent HPD set using sign restrictions, fourth-order moment spectral decomposition, and higher-order moment inequality restrictions over 500 artificial samples of 200 observations. The blue line is the true impulse response.

5 Empirical Application I: US Monetary Policy

We now apply our methodology to observed US postwar data and study the transmission mechanism of conventional monetary policy in the United States. We assume that monetary policy shocks are drawn from a leptokurtic distribution. We start by documenting that, indeed, excess kurtosis is a robust feature of monetary policy shocks when examining a large set of proxies or estimates obtained for the United States, the euro area, or the United Kingdom. We show that such higher-order moment inequality restrictions help sharpen the identification of monetary policy shocks when they are combined with a minimal set of sign restrictions that are agnostic about the impact on output, as in Uhlig (2005). In particular, we find that a tightening monetary policy surprise has a contractionary impact on output. This is consistent with the results obtained with recent approaches used in combination with sign restrictions (for example, Arias et al. (2019), Antolin-Diaz and Rubio-Ramírez (2018)) or that rely on external instruments (for example, Gertler and Karadi (2015)).

5.1 Higher-Order Moments of Monetary Policy Shock Proxies

We document the third- and fourth-order moment properties of a wide set of estimated monetary policy shocks. In a recent paper, Jarociński (2021) emphasizes that US monetary policy surprises identified from intraday movements of federal funds-rate futures observed in a narrow window around FOMC announcements exhibit leptokurticity. In this section, we show that this feature is not specific to daily or intraday variations of interest rates around central bank monetary policy decisions and announcements or to US monetary policy. This property is actually observed for a wide range of monetary policy shock proxies obtained with data with different frequencies, for different sample periods, and for different economies.

US data For the United States, we consider the recent monetary policy surprises extracted from high-frequency data sets from Gertler and Karadi (2015) (GK), Miranda-Agrippino and Ricco (2021) (MAR), and Jarociński and Karadi (2020) (JK). Some of these proxies control for the information/Delphic effect of monetary policy (as discussed in Campbell, Evans, Fisher and Justiniano (2012) and Nakamura and Steinsson (2018a)). We also look at the raw (unrotated) first three principal components of the intraday variations of the interest rates term structure around FOMC announcements (that is, USf1–USf3). We also include monetary policy shocks identified with lower-frequency data, in particular the Romer and Romer (2004) (RR)[16] narrative instrument, Sims and Zha (2006) (SZ) monetary policy innovations estimated from an SVAR with regime shifts, and the monetary policy shock estimated in the Smets and Wouters (2007)

[16] We consider here the series constructed in Wieland and Yang (2020) who extend the Romer-Romer (2004) monetary policy shock series.

(SW) DSGE model. Finally, we also consider the natural language measure constructed by Aruoba and Drechsel (2022) (AD), in which they examine the change in the target interest rate, which cannot be predicted by the textual information contained in the documents prepared by Federal Reserve staff in advance of policy decisions.

EA data For the euro area, we consider the monetary policy surprises constructed in Andrade and Ferroni (2021), which exploit high-frequency variations of overnight interest swap (OIS) contracts around the monetary policy decisions and associated press conferences of the European Central Bank (ECB) and distinguish between conventional (AF(target)) and forward-guidance shocks (AF(FWG)), controlling for information/Delphic effects (AF(delphic)). We also look at the unrotated first three principal components of variations of euro-area interest rate futures around ECB decisions and communication constructed in Altavilla, Brugnolini, Gürkaynak, Motto and Ragusa (2019) (EAf1-EAf3).

UK data For the United Kingdom, we consider monetary policy surprises obtained using the high-frequency data from Gerko and Rey (2017), who examines change in three-month sterling future rates observed around the release of the minutes of the Bank of England Monetary Policy Committee (MPC) (GR(minute)) and inflation reports (GR(IR)), from Cesa-Bianchi, Thwaites and Vicondoa (2020) (CBTV), who use changes in three-month Libor around monetary policy events and the methodology introduced in Gürkaynak, Sack and Swanson (2005), and from Kaminska and Mumtaz (2022), who look at variations of the full yield curve of UK government bonds (KM). We also include the monetary policy shocks obtained by Cloyne and Hurtgen (2016), who employ the Romer-Romer identification approach to construct a monthly measure of UK monetary policy shocks (CH).

Results Table 1 reports the robust measures of skewness and excess kurtosis for each of these proxies, along with the 95 percent confidence intervals obtained by bootstrapping the series for all these measures. While there is no clear evidence of skewness of the monetary policy shocks, all the proxies have a statistically significant excess kurtosis, although point estimates are widely dispersed. It is quite striking that all these proxies are leptokurtic, given that they are obtained using different methods, data sets, time spans, and countries. As a result, they span relatively different information sets. For example, in the United States, the high-frequency measures of monetary policy surprises and the Romer-Romer narrative instrument have a correlation of only about 0.2 (see appendix figure 12). Therefore, it is not straightforward to assess which proxy characterizes monetary policy shocks better and should be used as an instrument in proxy VARs. Using higher-order moment restrictions to identify monetary policy shocks rather exploits a robust feature observed for many proxies.

	Excess Kurtosis	Skewness	Sample Size	Sample Coverage
SW	2.0 [0.4, 3.2]	-0.0 [-0.1, 0.1]	179	1960–2004
SZ	3.8 [1.8, 6.3]	0.0 [-0.0, 0.1]	518	1960–2003
RR	3.2 [1.8, 5.1]	0.0 [-0.1, 0.1]	468	1969–2007
GK	11.3 [5.9, 18.2]	-0.3 [-0.3, -0.2]	269	1990–2012
MAR	3.3 [1.2, 5.9]	-0.1 [-0.2, 0.0]	228	1991–2009
JK	8.8 [5.4, 15.8]	-0.1 [-0.2, -0.0]	323	1990–2016
AD	3.1[1.4, 7.0]	-0.0[-0.1, 0.1]	313	1982–2008
USf1	1.4 [0.3, 3.5]	0.1 [-0.0, 0.2]	204	1994–2017
USf2	3.0 [1.4, 7.1]	0.1 [-0.0, 0.2]	204	1994–2017
USf3	1.9 [0.5, 4.4]	0.0 [-0.1, 0.1]	204	1994–2017
AF(target)	2.5 [0.6, 5.3]	-0.0 [-0.2, 0.1]	134	2004–2015
AF(delphic)	1.3 [0.2, 3.9]	-0.0 [-0.2, 0.1]	134	2004–2015
AF(FWG)	1.4 [0.2, 3.6]	0.0 [-0.1, 0.1]	134	2004–2015
EAf1	3.4 [1.4, 5.6]	-0.0 [-0.2, 0.1]	197	2002–2019
EAf2	1.5 [0.3, 3.9]	-0.1 [-0.2, 0.0]	197	2002–2019
EAf3	1.1 [0.2, 3.4]	0.0 [-0.1, 0.2]	197	2002–2019
CH	13 [5.9, 38]	0 [-0.1, 0.1]	348	1997–2015
GR(minutes)	2.5 [1.3, 4.9]	-0 [-0.2, 0.1]	211	1997–2015
GR(IR)	3.6 [2.8, 4.6]	-0.1 [-0.2, 0.0]	211	1997–2015
CBTV	3.4 [0.6, 7.5]	-0.1 [-0.2, 0.0]	212	1979–2007
KM	5.3 [2.3, 11.2]	-0.0 [-0.1, 0.1]	235	1997–2016

Table 1: Excess kurtosis & skewness for various proxies of monetary policy shocks. Median and, in parenthesis, 95 percent confidence intervals obtained by bootstrap. See Table 2 for details on the proxies and their sources.

5.2 Monetary Policy Transmission with Sign and Kurtosis Restrictions

VAR specification We consider the dataset studied in Uhlig (2005), which consists of real GDP (y), GDP deflator (pi), commodity price index (pcom) and the federal funds rate (FFR) over a 1965:m1 to 2003:m12 sample.[17] We estimate the VAR parameters assuming 12 lags. Since the QML or least-square (LS) estimates of the autoregressive coefficients are asymptotically consistent, the LS residuals are treated as consistent estimators of the reduced-form shocks. Figure 7 plots the least-square estimates of the whitened VAR reduced-form shock, that is, the reduced-form residuals orthogonalized using the Cholesky factorization of the least-square estimates of the residual covariance matrix. The first row of the figure displays the estimated

[17] Uhlig (2005) uses data obtained by interpolating real GDP with industrial production and the deflator with CPI and PPI.

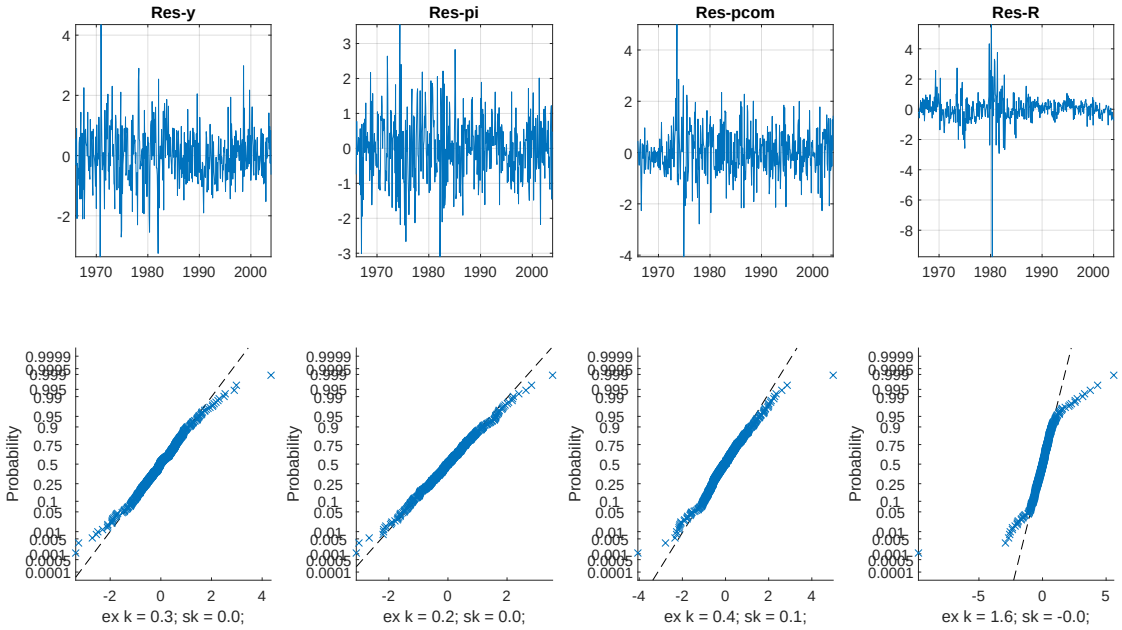


Figure 7: Whitened least-square residuals.

series, and the second row compares the probability plot of the empirical distribution against a standard normal distribution (dashed line); departures from the dashed line indicate departures from the normality assumptions. The are little signs of departures from normality for whitened residuals of output and prices: While there are few tail observations that lie far from the normal distribution, the Kolmogorov-Smirnov (K-S) test fails to reject the null that they could have come from a standard normal distribution; their respective $p-$values are 0.12 and 0.42. More evident departures from normality arise for the whitened residuals of commodity prices and the federal funds rate, where the K-S test $p-$values are 0.07 and 0.00, respectively. In fact, the robust measures of skewness are all very close to zero. Hence, the source of non-Gaussianity comes from the tails' thickness rather than the asymmetry of the distribution. The FFR residuals seems to be characterized by a large negative value at the beginning of the 1980s. But even after removing the extreme values (min and max) of the FFR residuals, we still reject the null of normality with a $p-$value smaller than 1e-8.

Identifying assumptions We combine the sign restrictions of Uhlig (2005) with an inequality restriction on the shock fourth-order moment as follows.

Assumption 1 *An unexpected monetary policy shock induces an increase in the federal funds rate and a decrease in the GDP deflator and in the price of commodities on impact and for the following five months.*

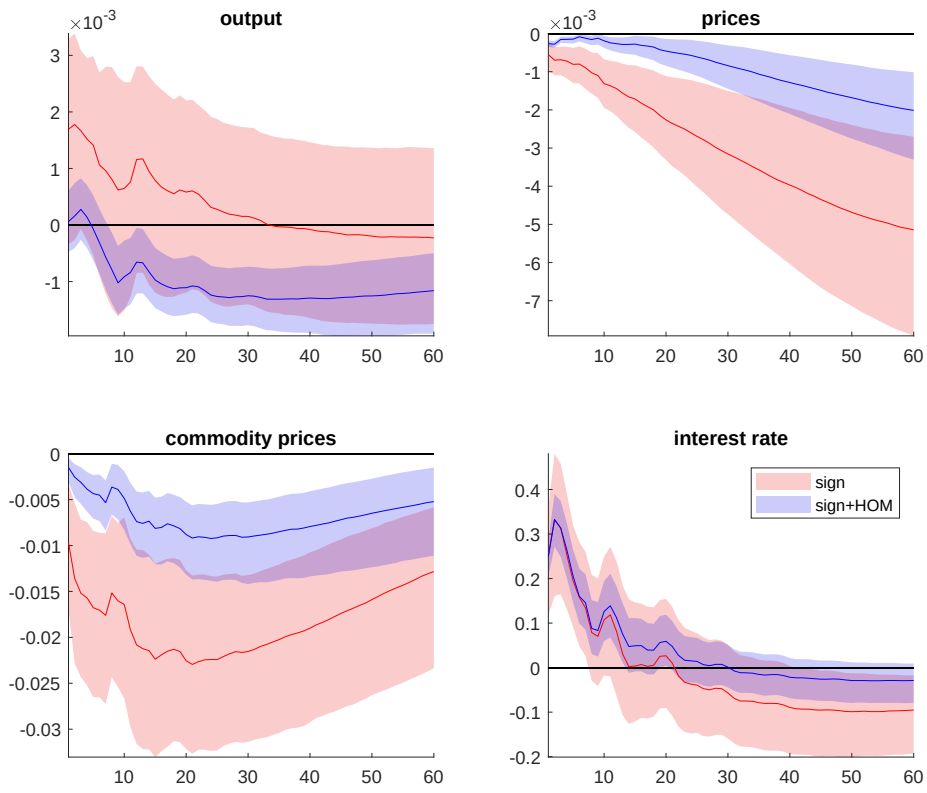


Figure 8: Impulse responses to a monetary policy shock. The colored areas indicate 68 percent credible sets. Sign restrictions are in red. Sign and kurtosis ($\mathcal{K}^{\star}_{mp} > 1.2$) restrictions are in blue.

Assumption 2 *An unexpected monetary policy shock is leptokurtic, and the robust measure of excess kurtosis is larger than 1.2, that is,* $\mathcal{K}^{\star}_{mp} > 1.2$.

The threshold for excess kurtosis is guided by the results of table 1 in section 5.1. We choose a conservative value below the lowest value (1.4) of robust excess kurtosis estimates observed across different proxies of US monetary policy shocks. Still, this threshold implies sizeable deviations from Gaussianity.[18]

Results Figure 8 reports the impulse response function (IRF) for a monetary policy shock. The median IRF obtained with only sign restrictions is in red; the one obtained with both sign and higher-order moment restrictions is in blue. Shaded areas depict the associated 68 percent confidence sets.[19] We find that monetary policy shocks have no clear effect on output when

[18] As a reference, a random variable distributed, say, as mixture of two normal distributions, where the random variable is drawn with a 0.8 (0.2) probability from (three times) the standard normal distribution, has a robust excess kurtosis of 1.5. With this distribution, the probability of observing a realization two (three) standard deviations away from the mean is 25 percent (700 percent) more likely that in the standard normal case.

[19] See figure 16 for the 90 percent confidence bands.

these are identified only with sign restrictions on the interest rate and inflation. When these are complemented with an inequality restriction on the excess kurtosis of monetary policy shocks, almost all the positive output responses are trimmed away, and the distribution tilts toward negative outcomes only. Note that the response of the FFR after an initial 25 basis point increase is almost the same under both identification schemes. Yet, when using inequality restriction on the monetary policy shock kurtosis, prices decline significantly less after a monetary policy shock—that is, the impact is about half the size of what is obtained with sign restriction only. Thus, higher-order moment restrictions also affect the estimated slope of the Phillips curve.

Discussion The confidence bands are obtained using the Haar prior, as in Rubio-Ramírez et al. (2010). There is debate around whether such a prior is overly informative and whether the prior selection vanishes asymptotically (e.g., Baumeister and Hamilton 2015, Arias, Rubio-Ramirez and Waggoner 2020). Giacomini and Kitagawa (2021) propose computing a "robust credible region" that relies on a prior class that specifies an "unrevisable" prior for the structural parameter given the reduced-form estimates. Appendix figure 17 illustrates that our findings remain valid when using such an approach.

We obtain results that are qualitatively similar to those of Arias et al. (2019), who complements Uhlig (2005)'s sign restrictions with the additional constraints that the FFR response to output and to prices must be non-negative, consistent with what a standard monetary policy reaction function implies. As Wolf (2020) illustrates using data simulated from a New Keynesian model, supply and demand shocks masquerading as monetary policy shocks lead to an FFR response to output that is negative, instead of positive, in the model. So, Arias et al. (2019)'s identification scheme helps to eliminate these shocks and to better identify monetary policy. As appendix figure 18 shows, imposing Uhlig (2005)'s sign restrictions on our sample also leads to an estimated policy rate response to output that is significantly negative. As appendix figure 18 also shows, imposing the additional inequality restrictions on the shock's kurtosis significantly moves that estimated reaction which becomes centered on zero, that is closer to Arias et al. (2019)'s identifying restrictions.

6 Empirical Application II: Sovereign Risk in the Euro Area

In the wake of the Great Recession, the euro area has experienced an acute sovereign crisis leading to very large spreads between government bond yields for countries that were exposed to sovereign risk and the other members. These spreads were driven by macroeconomic fundamentals but also by an increase in perceived sovereign risk. In fact, sovereign spreads dropped significantly after Mario Draghi, in 2012, made his "Whatever it takes" statement, which asserted the ECB's commitment to intervening on sovereign debt markets to prevent any risk of

potential breakup of the euro area. These sharp increases in sovereign risks are believed to have had large negative macroeconomic consequences, particularly through their impact on the balance sheet of financial intermediaries and, therefore, lending to the private sector. However, identifying such a macroeconomic impact of euro-area sovereign risk shocks is challenging, as these shocks coincided with other shocks that also had negative macroeconomic consequences—namely the Great Recession, the ensuing deleveraging shocks, and the contractionary monetary policy shocks implied by the zero lower-bond constraint. Some scholars address this issue by using bond market reactions observed in a narrow window around key political events. The assumption is that these market reactions capture policy choices that will only affect debt sustainability, not current or future economic conditions. They then use these market reactions as external instruments for the reduced-form residuals obtained from country-level panel VARs or local projections (e.g. Bahaj 2020, Balduzzi, Brancati, Brianti and Schiantarelli 2023). Other contributions rely on a structural model that they estimate using country-specific data (Bocola 2016) or US regional data and euro-area country data (Martin and Philippon 2017). We propose using sign and higher-order moment inequality restrictions in an SVAR to estimate the impact of sovereign shocks on euro-area aggregates.

VAR specification We use monthly data on euro-area industrial production (IP, in log), euro-area consumer prices measured by core harmonised index of consumer prices (HICP) (Core, in log), the unemployment rate (UNR, in percent), a measure of nonfinancial corporation (NFC) borrowing costs (NFC spreads, in percent),[20] the one-year money market rate (Euribor 1y, in percent), the spread between the five-year Italian and German bond yield (It-De 5y), and the 10-year government bond yield for Italy (It 10y, in percent) and Germany (De 10y, in percent). The sample period goes from 1999m1 to 2019m12. We estimate a VAR model with six lags and uninformative priors for the parameters of interest. Looking at the reduced-form residuals higher-order moment properties, there are salient deviations from normality for several variables in the system: The K-S test rejects the null of Gaussianity for the borrowing costs, the one-year Euribor, and the Italian and German bond yield spread; for the remaining variables, the test fails to reject the null hypothesis of normality.

Identifying assumptions We impose the following restrictions on the sign of the impulse responses to a sovereign risk shock.

Assumption 3 *An unexpected spread shock increases the five-year Italian–German sovereign spread, increases the 10-year Italian government bond yield, and the NFC borrowing costs on impact and for the following month.*

[20] We use the nonfinancial corporation borrowing costs constructed in Gilchrist and Mojon (2017).

This assumption is consistent with models featuring sovereign default risk and financial intermediaries (e.g. Corsetti, Kuester, Meier and Muller 2013, Bocola 2016). We are agnostic about the reaction of other macroeconomic variables to that shock and therefore leave their reaction unrestricted. Other shocks can satisfy these sign restrictions, so we combine them with higher-order moment inequality restrictions capturing what we assume is a feature specific to sovereign risk shocks: They increase sharply relatively frequently.

Assumption 4 *The spread shock has moderate asymmetry to the right and moderately fat tails. In particular, the skewness is larger than 0.2 and the excess kurtosis is larger than 0.5, that is,* $\mathcal{S}_{sp} > 0.2$ *and* $\mathcal{K}^{\star}_{sp} > 0.5$.

Here, we use the robust measure of skewness and kurtosis (see section 3). A skewness value larger than 0.2 only imposes a moderate right-skewness. The latter combined with an excess kurtosis larger than 0.5 implies that the frequency of large positive events is twice (six times) as large for realizations bigger than two (three) standard deviations relative to the normal case.[21]

Results Figure 9 shows the results obtained with our identification scheme (in blue) and compare them with what a standard recursive ordering identification—in which the sovereign risk shock affects the Italian and German sovereign yields as well as the Italian–German sovereign spread on impact but affects the other financial and macroeconomic conditions with a lag—implies (in red). Impulse responses are normalized so that the maximum median effect on the spread over the response horizon is 1 percent. The identification using signs and higher-order moment restrictions produces notable dynamic responses. A sovereign risk shock leading to a 100 basis point peak increase in the five-year Italian–German sovereign spread generates a 50 basis point increase in the Italian 10-year bond yield as well as a 25 basis point decline in the German 10-year bond yield, consistent with flight-to-quality effects. Credit conditions tighten as the NFC spread increases by 70 basis points, showing a large pass-through from the sovereign risk to the corporate sector. Industrial production contracts by nearly 2 percent 16 months after the shock, and the unemployment rate increases by almost 50 basis points 18 months after the shock. The price level for core goods and services increases to reach a tiny (and statistically non-significant) one basis point peak one year after the shock. The one-year Euribor drops as monetary policy eases to stabilize the economy. In less than three years, the shock is absorbed. These effects are qualitatively similar to what Bahaj (2020) obtains, on average, for Ireland, Italy, Portugal, and Spain over a crisis 2007–2013 sample period. Our estimates of the macroeconomic effects are larger, though, especially considering that our estimates are for euro-area aggregates and include non-crisis time. This suggests that country-specific sovereign

[21] Assuming that the distribution is skewed as well as leptokurtic makes the shock differ from monetary policy, which in the previous section was assumed to be leptokurtic but symmetric.

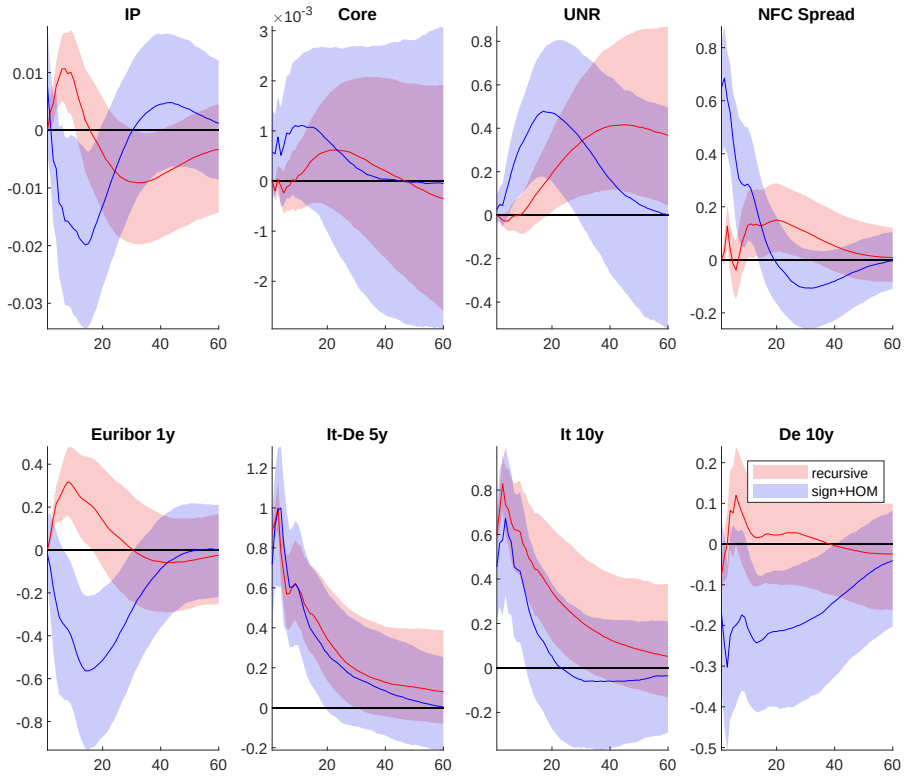


Figure 9: Impulse responses to a sovereign risk shock. The figure displays results obtained with a recursive ordering scheme (in red) and with sign and higher-order moment restrictions (in blue). Responses are normalized so that the maximum median impact on the spread is 1 percent. The solid line indicates the median response, and the colored areas report the 68 percent confidence bands.

risk shocks have negative spillovers into other monetary union members, even if they are not affected by sovereign risk.

The results obtained using a recursive ordering scheme offer a very different picture, with some puzzling patterns. A shock leading to a 1 percent increase in the sovereign spread triggers an increase in the 10-year Italian bond yield. It also increases the German bond yield, so there is no flight-to-quality effect. Credit costs increase but only modestly and with a delay compared with our proposed identification. Industrial production increases in the short run and declines after two years; the unemployment rate increases for several months after the shock, reaching its peak in three to four years after the impulse. The price level only modestly increases but it is not significant. Finally, the one-year money market rate increases, signalling monetary policy tightening. Overall, these effects do not seem consistent with what sovereign risk would produce, implying that the identification mixes several structural shocks. Figure 10 shows the results obtained when only the sign restrictions (that is, assumption 3) are used. The results

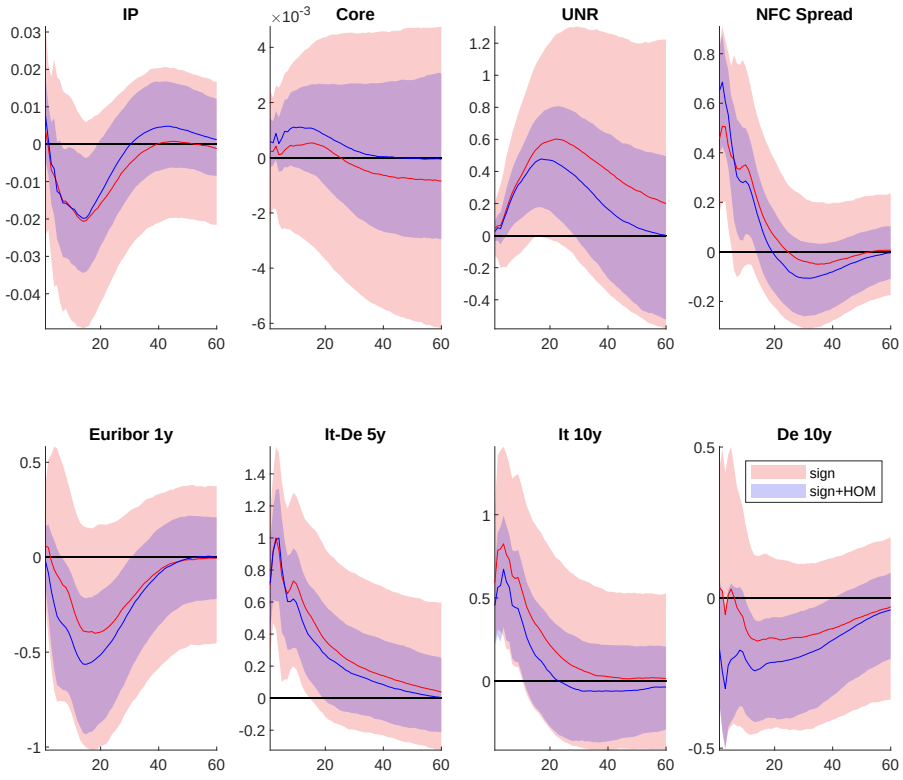


Figure 10: Impulse responses to a sovereign risk shock. The figure displays results obtained with sign restrictions only (in red) and with sign and higher-order moment restrictions (in blue). Responses are normalized so that the maximum median impact on the spread is 1 percent. Solid line median response and colored areas report the 68 percent confidence bands.

show that these constraints are not restrictive enough, as this scheme does not generate any statistically significant response.[22]

[22]Figures 9 and 10 report the 68 percent confidence bands. Appendix figures 19 and 21 also report the 90 percent confidence sets.

7 Empirical Application III: The Macroeconomic Effects of Geopolitical Risk

What are the effects of rising geopolitical tensions on the US economy? Caldara and Iacoviello (2022) develop an index of geopolitical risk (GPR) that can be used to address this question. The index is obtained by computing the share of articles mentioning adverse geopolitical events in leading newspapers published in the United States, the United Kingdom, and Canada. Caldara and Iacoviello (2022)'s GPR index captures exogenous events that are not caused, at least within a quarter, by US macroeconomic performances. However, as Caldara and Iacoviello (2022) discuss, the index correlates with other measures of policy shocks. Indeed, large adverse geopolitical events can also coincide with exceptional and unpredictable policy decisions, either from the fiscal authority (for example, military spending shocks) or from the central bank (for example, the Fed emergency reaction to 9/11). Therefore, some articles discussing geopolitical tensions may also discuss these unforeseen policy decisions. Since these other policy shocks affect macroeconomic aggregates, isolating the macroeconomic impact of GPR shocks raises an identification challenge. We propose to address this challenge by combining sign and higher-order moment restrictions.

VAR specification Caldara and Iacoviello (2022) also provide an SVAR analysis of the macroeconomic effects of geopolitical risk for the US economy. Their VAR consists of eight quarterly variables: (i) the log of the GPR index, (ii) the stock market volatility index (VIX), (iii) the log of real business fixed investment per capita, (iv) the log of private hours per capita, (v) the log of the S&P 500 index, (vi) the log of the West Texas Intermediate (WTI) crude oil price, (vii) the yield on two-year US Treasury securities, and (viii) the Chicago Federal Reserve National Financial Conditions Index (NFCI). The estimation sample is 1986:Q1 to 2019:Q4, and the VAR admits two lags.

Identifying assumptions Caldara and Iacoviello (2022) identify a GPR shock using a recursive scheme, with the GPR index ordered first. This assumes that any contemporaneous correlation between the GPR index and the other variables in the VAR comes from the causal effect of the GPR shock on the other variables. As reported in Figure 11, which replicates their approach, a one-standard-deviation increase in geopolitical risk has a significant contractionary impact on financial conditions and macroeconomic aggregates such as investment and hours. However, as Caldara and Iacoviello (2022) note, the GPR index can also capture other shocks such as policy shocks. If these shocks also affect financial and macroeconomic conditions contemporaneously, a recursive ordering may, in part, capture these as well as the GPR shocks. This could explain why, under a recursive ordering identification scheme, although the geopo-

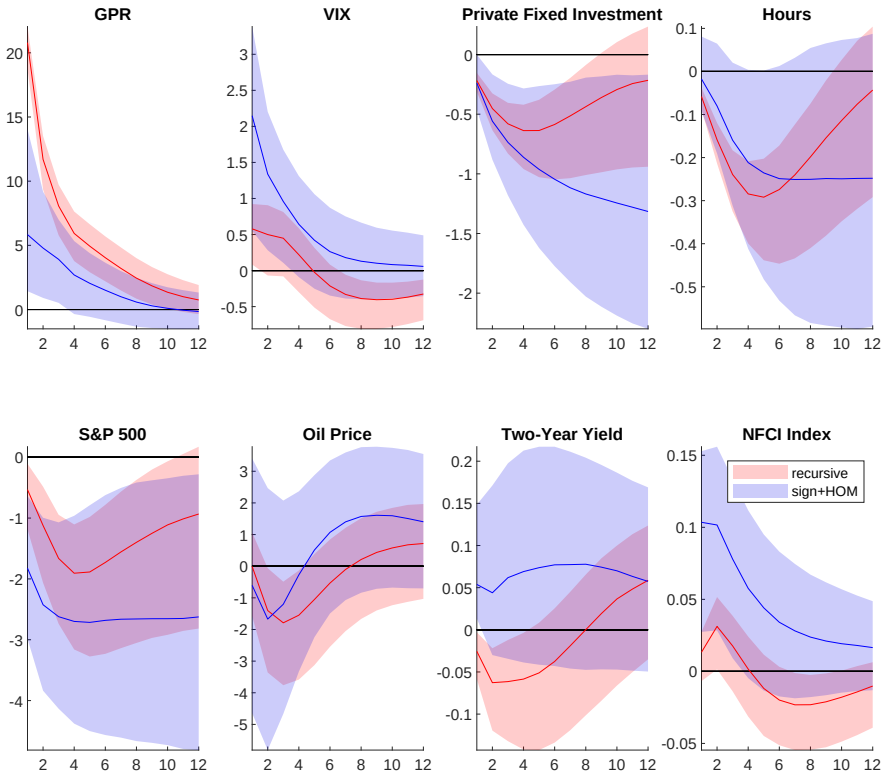


Figure 11: Impulse responses to a geopolitical risk shock. The figure displays results obtained with a recursive ordering scheme (in red) and with sign and higher-order moment restrictions (in blue). The responses are normalized to a one-standard-deviation shock. The solid line indicates the median response, and the colored areas report the 68 percent confidence bands.

litical tensions last well over one year after the initial GPR shock, the VIX starts to decline, the two-year bond yield declines, and overall financial conditions, as captured by the NFCI, loosen. One way to address this potential issue is to use sign restrictions requiring that, on impact, the GPR shock increases the GPR index and to tighten financial conditions by lowering the SP 500 and increasing the two-year US Treasury yield. However, unreported results show that these sign restrictions are too loose and lead to a set of admissible rotations that are too large and median IRFs that are not significantly different from zero. We thus investigate an alternative identification scheme that combines these sign restrictions with restrictions on higher-order moments of the GPR shock. Indeed, the GPR index is skewed to the right and leptokurtic, consistent with the intuition that geopolitical risks are characterized by relatively frequent spikes in international tensions. We assume that GPR shocks have a distribution that is also skewed to the right and leptokurtic. More specifically, we postulate the following:

Assumption 5 *A GPR shock initially contracts financial conditions that is it lowers the SP*

500 and increases the two-year US Treasury yield on impact.

Assumption 6 *The GPR shock distribution has small asymmetry to the right and fat tails. In particular, the skewness is larger than 0.1 and excess kurtosis is larger than 1.5, that is,* $\mathcal{S}_{gpr} > 0.1$ *and* $\mathcal{K}^{\star}_{gpr} > 1.5$.

Results As Figure 11 illustrates, consistent with the results in Caldara and Iacoviello (2022), we find evidence of a persistent contractionary effect of GPR shocks on financial conditions and macroeconomic aggregates. The estimated effects obtained with our scheme are less precise compared with the recursive ordering, as we achieve set identification rather than point identification. However, they are larger and more persistent, consistent with the possibility that the recursive identification scheme can capture both the effects of geopolitical surprises and of coincident expansionary fiscal and monetary policy shocks resulting from discretionary decisions made in the wake of some geopolitical events. Notably, the reaction of the two-year US Treasury yield is positive on impact, consistent with an increase in risk premium, but then non-significantly different from zero soon after the shock. This may reflect the offsetting effects that an expansionary fiscal and monetary policy implemented after the GPR shock would have on that yield. Finally, oil prices do not react to geopolitical shock either. Indeed, depending on the shock, oil prices may in some cases decline due to decreased demand associated with an increase in uncertainty, and, in other cases, prices may increase if geopolitical events occur in oil-producing countries. Overall, the evidence aligns with models in which uncertainty shocks lower investment (Bloom 2009) and employment (Basu and Bundick 2017) and tighten financial conditions (Gilchrist, Sim and Zakrajsek 2014).

References

Adjemian, S., Bastani, H., Juillard, M., Karamé, F., Maih, J., Mihoubi, F., Perendia, G., Pfeifer, J., Ratto, M. and Villemot, S.: 2011, Dynare: Reference manual version 4, *Dynare Working Papers 1*, CEPREMAP.

Altavilla, C., Brugnolini, L., Gürkaynak, R. S., Motto, R. and Ragusa, G.: 2019, Measuring Euro Area Monetary Policy, *Journal of Monetary Economics* **108**, 162–179.

Andrade, P. and Ferroni, F.: 2021, Delphic and odyssean monetary policy shocks: Evidence from the euro area, *Journal of Monetary Economics* **117**, 816–832.

Antolin-Diaz, J. and Rubio-Ramírez, J. F.: 2018, Narrative Sign Restrictions for SVARs, *American Economic Review* **108**(10), 2802–2829.

Arias, J. E., Caldara, D. and Rubio-Ramírez, J. F.: 2019, The systematic component of monetary policy in SVARs: An agnostic identification procedure, *Journal of Monetary Economics* **101**, 1–13.

Arias, J. E., Rubio-Ramirez, J. F. and Waggoner, D.: 2020, Uniform priors for impulse responses, *Working Papers 22-30*, Federal Reserve Bank of Philadelphia.

Aruoba, B. and Drechsel, T.: 2022, Identifying Monetary Policy Shocks: A Natural Language Approach, *CEPR Discussion Papers 17133*, C.E.P.R. Discussion Papers.

Bahaj, S.: 2020, Sovereign spreads in the euro area: Cross border transmission and macroeconomic implications, *Journal of Monetary Economics* **110**, 116–135.

Balduzzi, P., Brancati, E., Brianti, M. and Schiantarelli, F.: 2023, Political risk, populism, and the economy, *The Economic Journal* **133**(653), 1677–1704.

Barnichon, R. and Mesters, G.: 2020, Identifying Modern Macro Equations with Old Shocks, *The Quarterly Journal of Economics* **135**(4), 2255–2298.

Basu, S. and Bundick, B.: 2017, Uncertainty Shocks in a Model of Effective Demand, *Econometrica* **85**, 937–958.

Baumeister, C. and Hamilton, J. D.: 2015, Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information, *Econometrica* **83**(5), 1963–1999.

Bloom, N.: 2009, The Impact of Uncertainty Shocks, *Econometrica* **77**(3), 623–685.

Bocola, L.: 2016, The pass-through of sovereign risk, *Journal of Political Economy* **124**(4), 879–926.

Bowley, A. L.: 1926, *Elements of statistics*, number 8, King.

Brunnermeier, M., Palia, D., Sastry, K. A. and Sims, C. A.: 2021, Feedbacks: Financial Markets and Economic Activity, *American Economic Review* **111**(6), 1845–1879.

Caldara, D. and Iacoviello, M.: 2022, Measuring Geopolitical Risk, *American Economic Review* **112**(4), 1194–1225.

Campbell, J. R., Evans, C. L., Fisher, J. D. and Justiniano, A.: 2012, Macroeconomic Effects of Federal Reserve Forward Guidance, *Brookings Papers on Economic Activity* (Spring), 1–80.

Cesa-Bianchi, A., Thwaites, G. and Vicondoa, A.: 2020, Monetary policy transmission in the united kingdom: A high frequency identification approach, *European Economic Review* **123**, 103375.

Christiano, L. J., Eichenbaum, M. and Evans, C. L.: 2005, Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, *Journal of Political Economy* **113**(1), 1–45.

Cloyne, J. and Hurtgen, P.: 2016, The macroeconomic effects of monetary policy: A new measure for the united kingdom, *American Economic Journal: Macroeconomics* **8**(4), 75–102.

Comon, P.: 1994, Independent Component Analysis, a new concept?, *Signal Processing* **36**, 287–314.

Corsetti, G., Kuester, K., Meier, A. and Muller, G. J.: 2013, Sovereign risk, fiscal policy, and macroeconomic stability, *The Economic Journal* **123**(566), F99–F132.

Crow, E. L. and Siddiqui, M.: 1967, Robust estimation of location, *Journal of the American Statistical Association* **62**(318), 353–389.

Cúrdia, V., Negro, M. D. and Greenwald, D. L.: 2014, Rare Shocks, Great Recessions, *Journal of Applied Econometrics* **29**(7), 1031–1052.

Drautzburg, T. and Wright, J.: 2023, Refining Set-Identification in VARs through Independence, *Journal of Econometrics* **Forthcoming**.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J. and Watson, M. W.: 2007, ABCs (and Ds) of understanding VARs, *American Economic Review* **97**(3), 1021–1026.

Ferroni, F. and Canova, F.: 2021, A Hitchhiker's Guide to Empirical Macro Models, *Working Paper Series WP-2021-15*, Federal Reserve Bank of Chicago.

Galí, J.: 2015, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications Second edition*, number 10495 in *Economics Books*, Princeton University Press.

Gerko, E. and Rey, H.: 2017, Monetary policy in the capitals of capital, *Journal of the European Economic Association* **15**(4), 721–745.

Gertler, M. and Karadi, P.: 2015, Monetary Policy Surprises, Credit Costs, and Economic Activity, *American Economic Journal: Macroeconomics* **7**(1), 44–76.

Giacomini, R. and Kitagawa, T.: 2021, Robust bayesian inference for set-identified models, *Econometrica* **89**(4), 1519–1556.

Gilchrist, S. and Mojon, B.: 2017, Credit Risk in the Euro Area, *Economic Journal* **128**, 118–158.

Gilchrist, S., Sim, J. W. and Zakrajsek, E.: 2014, Uncertainty, Financial Frictions, and Investment Dynamics, *NBER Working Papers 20038*, National Bureau of Economic Research, Inc.

Gouriéroux, C., Monfort, A. and Renne, J.-P.: 2017, Statistical inference for independent component analysis: Application to structural var models, *Journal of Econometrics* **196**(1), 111–126.

Gouriéroux, C., Monfort, A. and Renne, J.-P.: 2019, Identification and Estimation in Non-Fundamental Structural VARMA Models, *The Review of Economic Studies* **87**(4), 1915–1953.

Groeneveld, R. A. and Meeden, G.: 1984, Measuring skewness and kurtosis, *Journal of the Royal Statistical Society: Series D (The Statistician)* **33**(4), 391–399.

Gürkaynak, R. S., Sack, B. and Swanson, E.: 2005, Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements, *International Journal of Central Banking* **1**(1), 55–93.

Hogg, R. V.: 1972, More light on the kurtosis and related statistics, *Journal of the American Statistical Association* **67**(338), 422–424.

Jarociński, M.: 2021, Estimating the Fed's unconventional policy shocks, *Working Paper Series 20210*, European Central Bank.

Jarociński, M. and Karadi, P.: 2020, Deconstructing Monetary Policy Surprises: The Role of Information Shocks, *American Economic Journal: Macroeconomics* **12**(2), 1–43.

Kaminska, I. and Mumtaz, H.: 2022, Monetary policy transmission during QE times: role of expectations and term premia channels, *Staff Working Paper 978*, Bank of England.

Känzig, D. R.: 2021, The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements, *American Economic Review* **111**(4), 1092–1125.

Kendall, M. G. and Stewart, A.: 1977, The advanced theory of statistics. vols. 1., *The advanced theory of statistics. Vols. 1.* **1**(Ed. 4).

Kim, T.-H. and White, H.: 2004, On more robust estimation of skewness and kurtosis, *Finance Research Letters* **1**(1), 56–73.

Kollo, T.: 2008, Multivariate skewness and kurtosis measures with an application in ica, *Journal of Multivariate Analysis* **99**(10), 2328–2338.

Lanne, M., Liu, K. and Luoto, J.: 2022, Identifying structural vector autoregression via leptokurtic economic shocks, *Journal of Business and Economic Statistics* .

Lanne, M., Meitz, M. and Saikkonen, P.: 2017, Identification and estimation of non-gaussian structural vector autoregressions, *Journal of Econometrics* **196**(2), 288–304.

Lewis, D.: 2024, Identification based on higher moments, *CeMMAP working papers 03/24*, Institute for Fiscal Studies.

Lewis, D. J.: 2021, Identifying Shocks via Time-Varying Volatility [First Order Autoregressive Processes and Strong Mixing], *Review of Economic Studies* **88**(6), 3086–3124.

Martin, P. and Philippon, T.: 2017, Inspecting the Mechanism: Leverage and the Great Recession in the Eurozone, *American Economic Review* **107**(7), 1904–1937.

Mertens, K. and Ravn, M. O.: 2013, The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, *American Economic Review* **103**(4), 1212–47.

Miranda-Agrippino, S. and Ricco, G.: 2021, The Transmission of Monetary Policy Shocks, *American Economic Journal: Macroeconomics* **13**(3), 74–107.

Montiel Olea, J. L., Plagborg-Møller, M. and Qian, E.: 2022, Svar identification from higher moments: Has the simultaneous causality problem been solved?, *AEA Papers and Proceedings* **112**, 481–85.

Moors, J.: 1988, A quantile alternative for kurtosis, *Journal of the Royal Statistical Society: Series D (The Statistician)* **37**(1), 25–32.

Nakamura, E. and Steinsson, J.: 2018a, High Frequency Identification of Monetary Non-Neutrality, *Quarterly Journal of Economics* **133**(3), 1283–1330.

Nakamura, E. and Steinsson, J.: 2018b, Identification in Macroeconomics, *Journal of Economic Perspectives* **32**(3), 59–86.

Petrova, K.: 2022, Asymptotically valid bayesian inference in the presence of distributional misspecification in var models, *Journal of Econometrics* **230**(1), 154–182.

Ramey, V. A.: 2016, Macroeconomic Shocks and Their Propagation, *in* J. B. Taylor and H. Uhlig (eds), *Handbook of Macroeconomics*, Vol. 2A, North Holland, chapter 3, pp. 71–162.

Rigobon, R.: 2003, Identification Through Heteroskedasticity, *The Review of Economics and Statistics* **85**(4), 777–792.

Romer, C. D. and Romer, D. H.: 2004, A New Measure of Monetary Shocks: Derivation and Implications, *American Economic Review* **94**(4), 1055–1084.

Rubio-Ramírez, J. F., Waggoner, D. F. and Zha, T.: 2010, Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference, *Review of Economic Studies* **77**(2), 665–696.

Schott, J.: 2016, *Matrix Analysis for Statistics*, Wiley Series in Probability and Statistics (Third Edition), Wiley.

Sims, C. A.: 1980, Macroeconomics and reality, *Econometrica* **48**(1), 1–48.

Sims, C. A. and Zha, T.: 2006, Were There Regime Switches in U.S. Monetary Policy?, *American Economic Review* **96**(1), 54–81.

Smets, F. and Wouters, R.: 2007, Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach, *American Economic Review* **97(3)**, 586–606.

Stock, J. H. and Watson, M. W.: 2012, Disentangling the Channels of the 2007-09 Recession, *Brookings Papers on Economic Activity* (Spring), 81–156.

Uhlig, H.: 2005, What are the effects of monetary policy on output? Results from an agnostic identification procedure, *Journal of Monetary Economics* **52**(2), 381–419.

Wieland, J. F. and Yang, M.-J.: 2020, Financial dampening, *Journal of Money, Credit and Banking* **52**(1), 79–113.

Wolf, C.: 2020, SVAR (Mis-)Identification and the Real Effects of Monetary Policy Shocks, *American Economic Journal: Macroeconomics* **12**(4), 1–32.

Wolf, C.: 2022, What Can We Learn From Sign-Restricted VARs?, *AEA Papers and Proceedings* **122**, 471–475.

A Appendix

A.1 Additional Tables and Figures

tag	Paper	Description	Frequency	country
SZ	Sims and Zha (2006)	SVAR zero restrictions	M	US
RR	Romer and Romer (2004)	narrative	M	US
GK	Gertler and Karadi (2015)	HF	M	US
MAR	Miranda-Agrippino and Ricco (2021)	HF corrected for info-effect	M	US
JK	Jarociński and Karadi (2020)	HF corrected for info-effect	M	US
SW	Smets and Wouters (2007)	DSGE	Q	US
AD	Aruoba and Drechsel (2022)	Text based and LLM	M	US
AF	Andrade and Ferroni (2021)	HF corrected for info-effect	M	EA
GK(M)	Gerko and Rey (2017)	HF around minutes	M	UK
GK(IR)	Gerko and Rey (2017)	HF around the inflation report	M	UK
CBTV	Cesa-Bianchi et al. (2020)	HF around monetary policy events	M	UK
KM	Kaminska and Mumtaz (2022)	HF around monetary policy events	M	UK
CH	Cloyne and Hurtgen (2016)	narrative	M	UK

Table 2: Various monetary policy surprises, estimates, and sources.

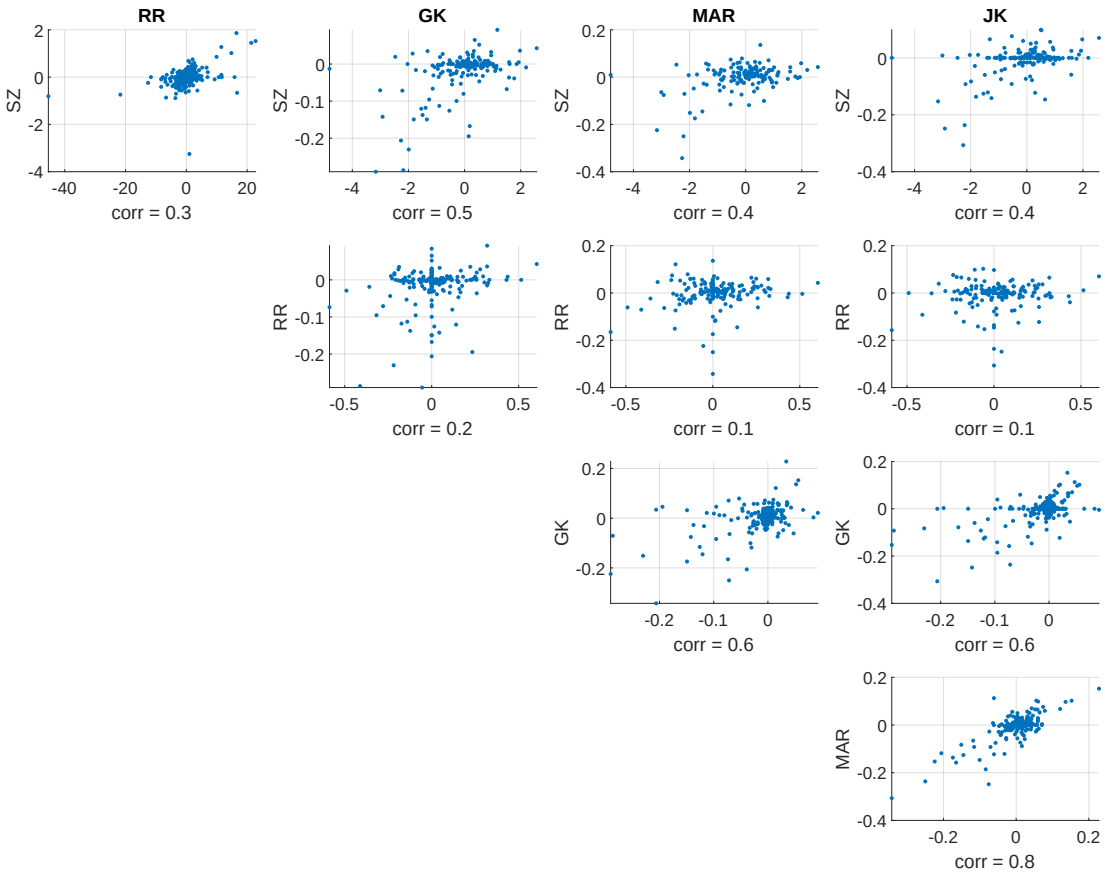


Figure 12: Correlations across measures of US monetary policy shocks.

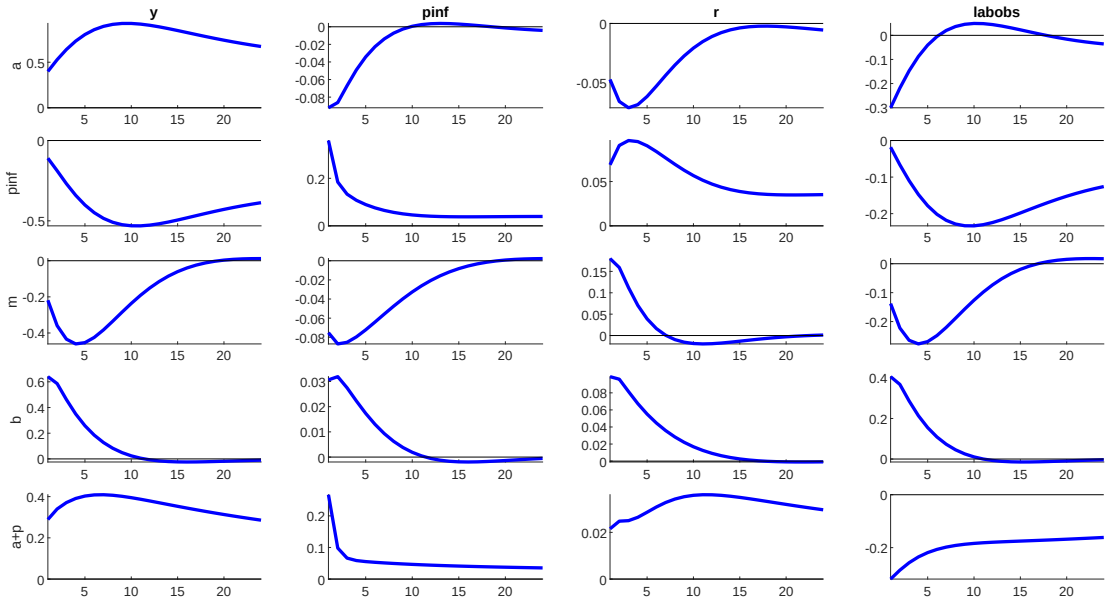


Figure 13: SW estimates of impulse response functions. From top to bottom: technology, price markup, monetary policy shocks and risk premium, and the sum of technology and price markup shocks.

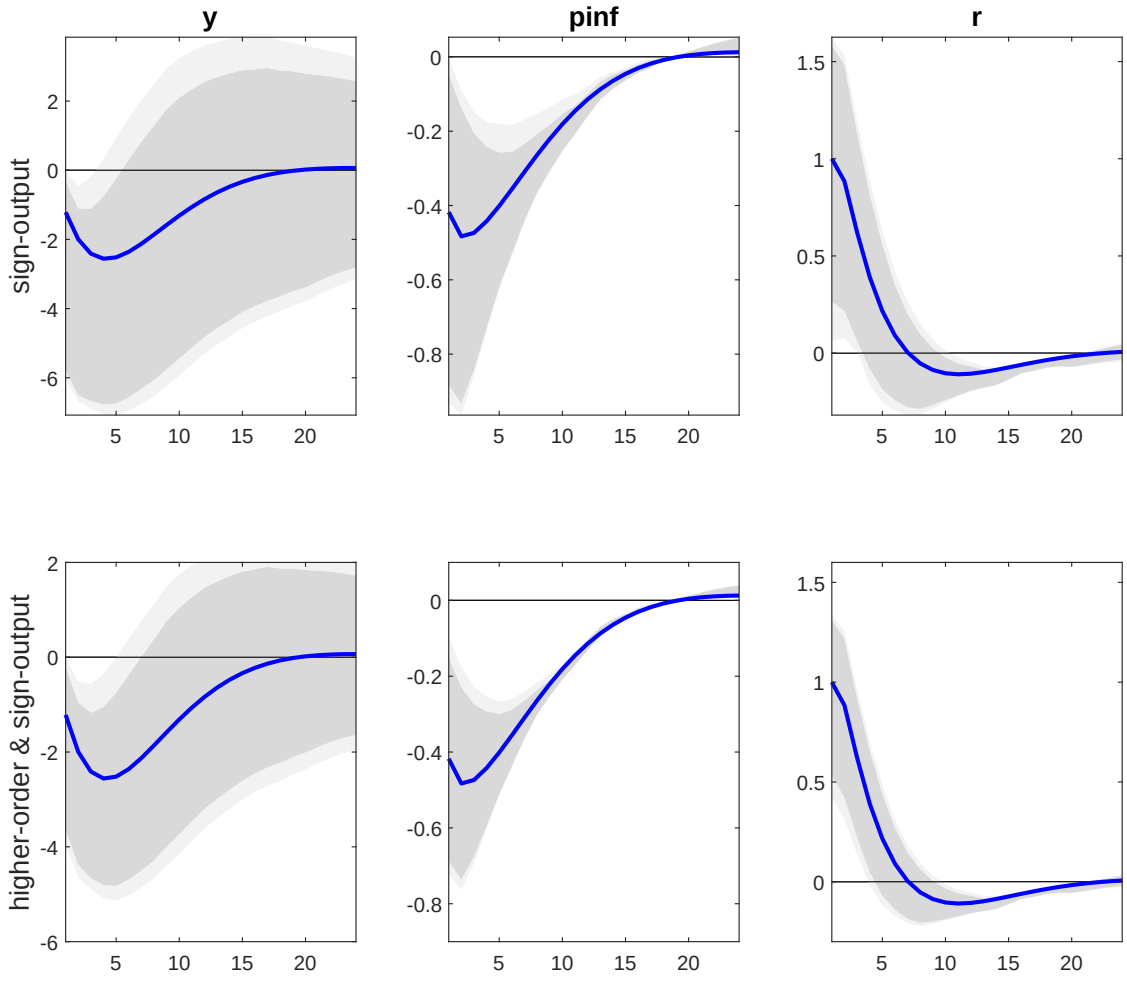


Figure 14: Impulse responses to a monetary policy shock using sign restrictions on output (−), prices (−) and interest rate (+) (first row) and the same sign restrictions combined with higher-order moment inequality (second row) restrictions. The blue solid line is the true impulse response; the dark (light) gray areas report the 90 percent (99 percent) confidence bands.

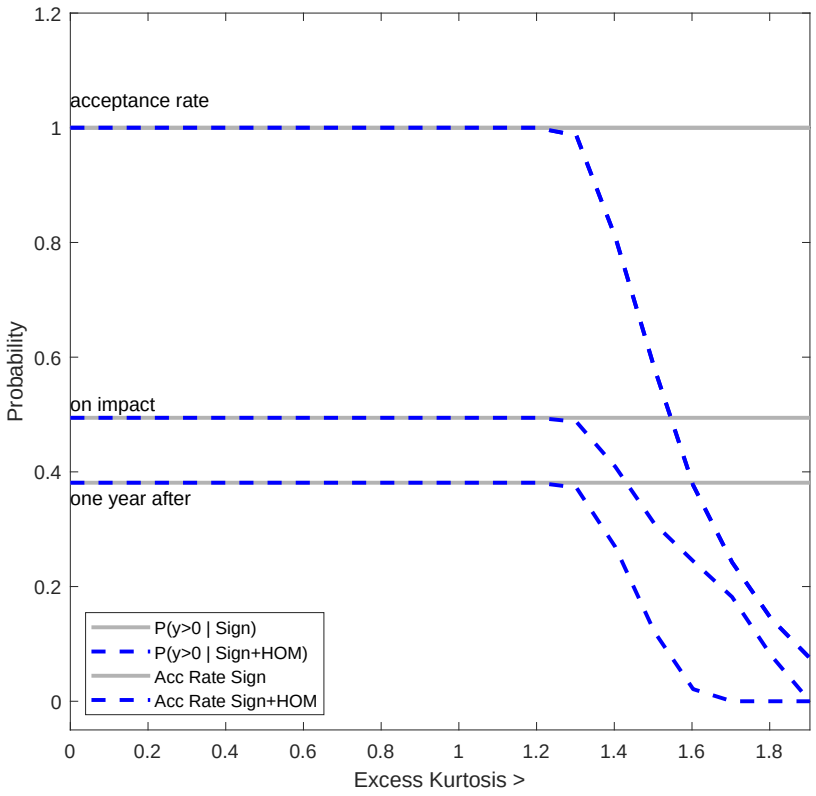


Figure 15: Probability of positive response of output after a monetary policy tightening as a function of the size of the interval.

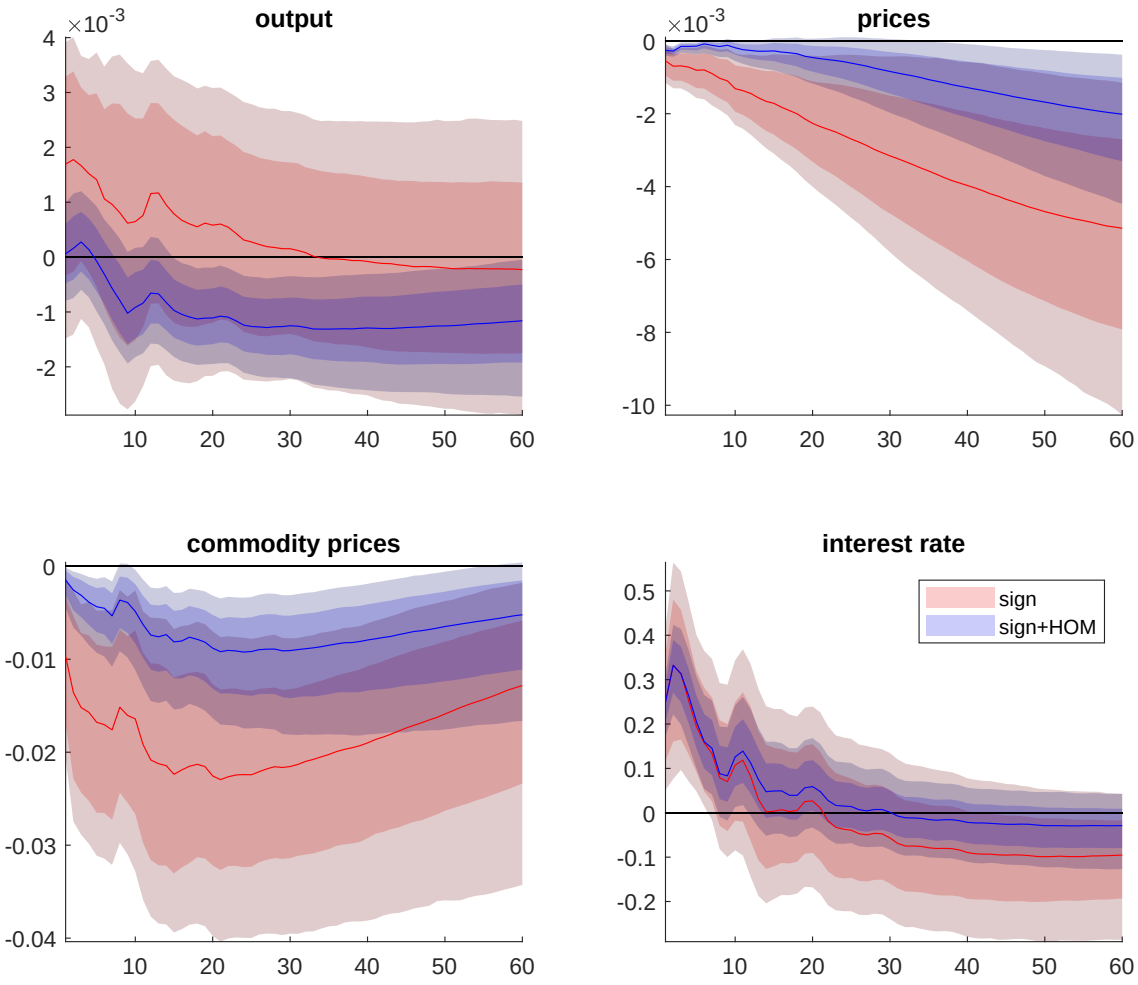


Figure 16: Impulse responses to a monetary policy shock. The colored areas indicate 68 percent and 90 percent credible sets. The figure displays results obtained with sign restrictions only (in red) and with sign and kurtosis ($\mathcal{K}_{mp} > 1.2$) restrictions (in blue).

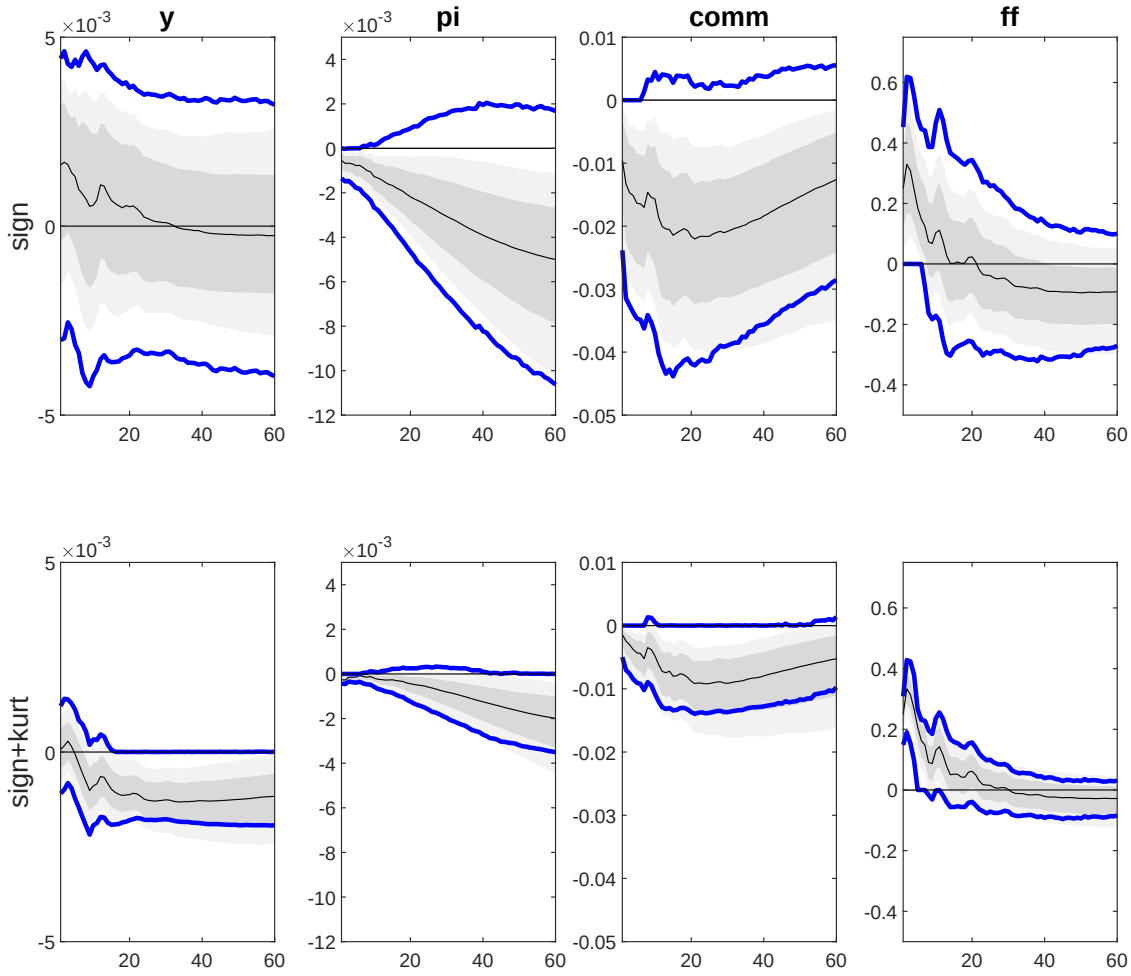


Figure 17: Impulse responses to a monetary policy shock. The colored gray areas indicate 68 percent and 90 percent credible sets obtained using draws from the Haar prior, as in Rubio-Ramírez et al. (2010). The blue lines report the robust confidence bands at the 68 percent confidence level discussed in Giacomini and Kitagawa (2021). Sign restrictions are in the top panels; sign and kurtosis ($\mathcal{K}_{mp} > 1.2$) restrictions are in the bottom panels.

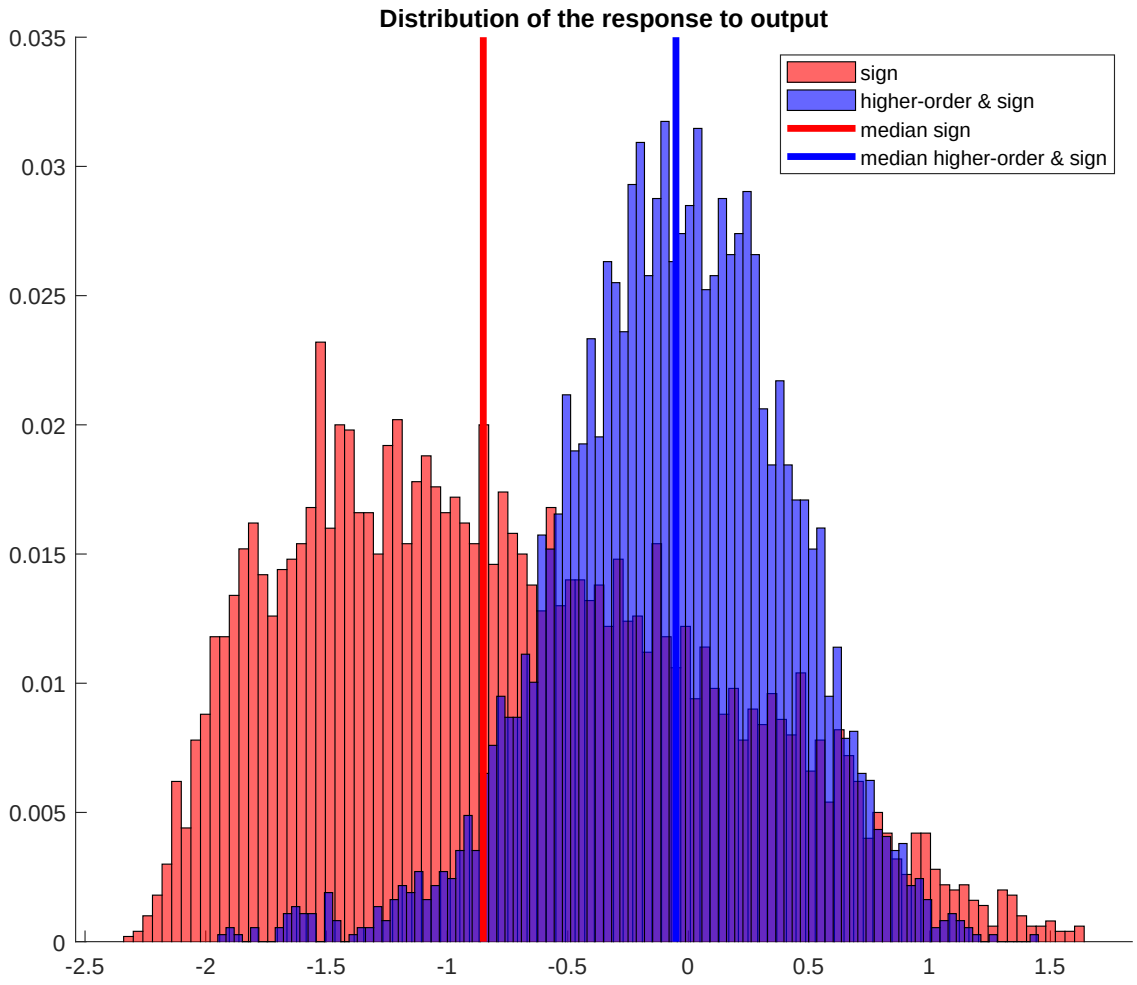


Figure 18: Distribution of the response of the policy rate to output. Results obtained with sign restrictions only are in red. Results obtained with sign and kurtosis ($\mathcal{K}_{mp} > 1.2$) restrictions are in blue.

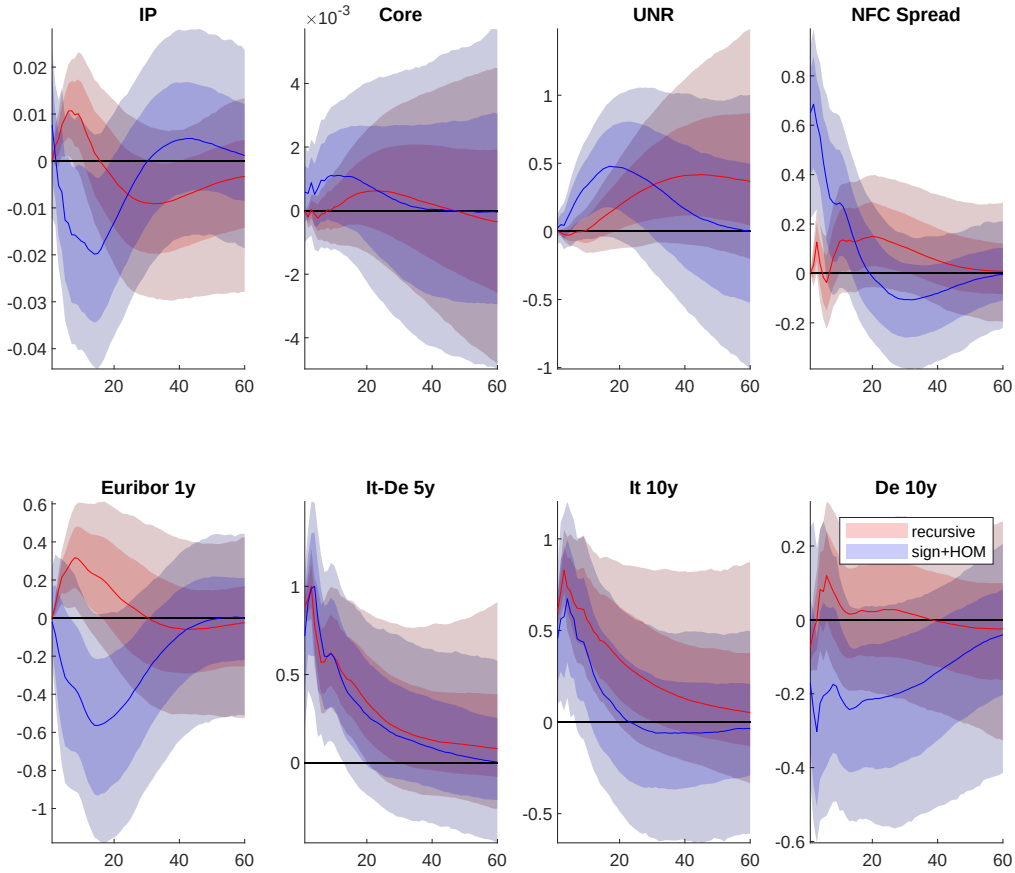


Figure 19: Impulse responses to a sovereign risk shock. Results obtained with a recursive ordering are in red. Results obtained with sign and higher-order moment restrictions are in blue. Impulse responses are normalized so that the maximum median impact on the spread is 1 percent. The solid line indicates the median response, and the colored areas report the 68 percent and the 90 percent confidence bands.

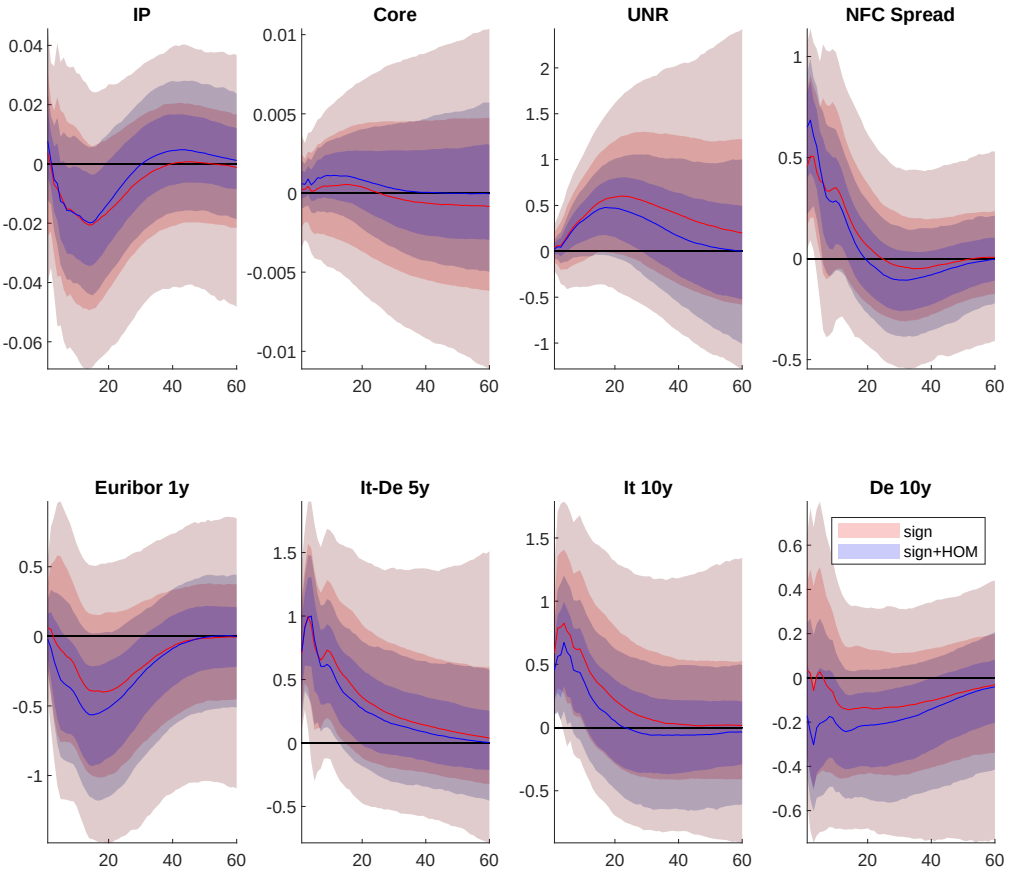


Figure 20: Impulse responses to a sovereign risk shock. Results obtained with sign restrictions only are in red. Results obtained with sign and higher-order moment restrictions are in blue. Impulse responses are normalized so that the maximum median impact on the spread is 1 percent. The solid line indicates the median response, and the colored areas report the 68 percent and 90 percent confidence bands.

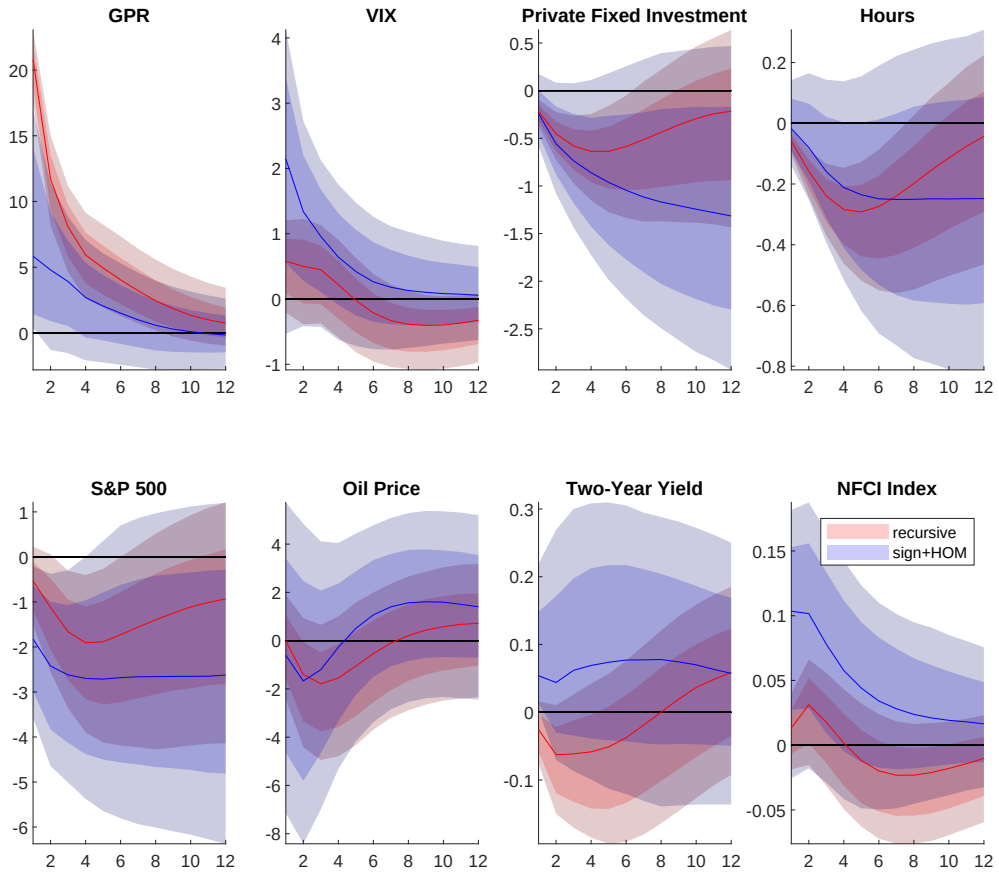


Figure 21: Impulse responses to a GPR shock. Results obtained with a recursive ordering are in red. Results obtained with sign and higher-order moment restrictions are in blue. Impulse responses are normalized to one standard deviation. The solid line indicates the median response, and the colored areas report the 68 percent and 90 percent confidence bands.

A.2 Derivations

Notation First, define $\boldsymbol{e}_k$ as the $n \times 1$ vector with zeros everywhere except a one in the k^{th} position, J_k the $n \times n$ matrix of zeros everywhere except one in the k^{th} position of the main diagonal, and J_{jk} the $n \times n$ matrix of zeros everywhere except one in the $(j,k)^{th}$ element. For example, when $n = 3$, $k = 2$, and $j = 3$, then

$$\boldsymbol{e}_2' = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, J_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } J_{3,2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Notice that

$$\begin{aligned} \boldsymbol{e}_k' \otimes \boldsymbol{e}_k =& J_{k,k} = J_k, \\ \boldsymbol{e}_j' \otimes \boldsymbol{e}_k =& J_{k,j}, \\ \boldsymbol{e}_j' \otimes \boldsymbol{e}_k =& \boldsymbol{e}_k \boldsymbol{e}_j'. \end{aligned}$$

Denote with I_n the identity matrix of size n and with $\text{vec}(X)$ the column-wise vectorization of X. The following identities hold:

$$\begin{aligned} \text{vec}(I_n) &= \sum_{i=1}^{n} \boldsymbol{e}_i \otimes \boldsymbol{e}_i, \\ \text{vec}(I_n)\text{vec}(I_n)' &= \left(\sum_{i=1}^{n} \boldsymbol{e}_i \otimes \boldsymbol{e}_i \right) \left(\sum_{i=1}^{n} \boldsymbol{e}_i \otimes \boldsymbol{e}_i \right)' = \left(\sum_{i=1}^{n} \boldsymbol{e}_i \otimes \boldsymbol{e}_i \right) \left(\sum_{i=1}^{n} \boldsymbol{e}_i' \otimes \boldsymbol{e}_i' \right) = \\ &= \sum_{k,j=1}^{n} (\boldsymbol{e}_k \otimes \boldsymbol{e}_k)(\boldsymbol{e}_j' \otimes \boldsymbol{e}_j') = \sum_{k,j=1}^{n} \boldsymbol{e}_k \boldsymbol{e}_j' \otimes \boldsymbol{e}_k \boldsymbol{e}_j' = \\ &= \sum_{k,j=1}^{n} J_{k,j} \otimes J_{k,j}. \end{aligned}$$

Define the commutation matrix, $K_{n,n}$, the $(n^2 \times n^2)$ matrix consisting of $n \times n$ blocks where the $(j,i)-$element of the (i,j) block equals one; elsewhere there are all zeros. Notice that

$$K_{n,n} = \sum_{k,j=1}^{n} J_{k,j} \otimes J_{j,k}.$$

Assumptions about the structural shocks ν: We assume that the structural shocks are independent and identically distributed over time. Moreover, we postulate that

- $E(\nu_{i,t}^2) = 1$ and $E(\nu_{i,t}\nu_{j,t}) = 0$ for all i, j;
- $E(\nu_{i,t}^3) = \zeta_i$ and $E(\nu_{i,t}\nu_{j,t}\nu_{k,t}) = 0$ for all $i \neq j, k$;
- $E(\nu_{i,t}^4) = \xi_i$, $E(\nu_{i,t}^2\nu_{j,t}^2) = 1$ for all $i \neq j$, and $E(\nu_{i,t}\nu_{j,t}\nu_{k,t}\nu_{m,t}) = 0$ for all $i \neq j, k, m$.

Define the empirical innovation as

$$\begin{aligned}\iota_{1,t} &= \alpha_{1,1}\nu_{1,t} + ... + \alpha_{1,m}\nu_{n,t},\\ &\vdots\\ \iota_{n,t} &= \alpha_{n,1}\nu_{1,t} + ... + \alpha_{n,n}\nu_{n,t},\end{aligned}$$

and we define with A_o the matrix collecting the structural coefficients $\alpha_{i,j}$. Finally, define the candidate structural shocks, $\breve{\nu}_t$, as $\breve{\nu}_t = A'\iota_t$, and denote with $\boldsymbol{\alpha}_k$ and $\boldsymbol{a}_k$ the k^{th} column of A_o (the true impact matrix) and A (a candidate rotation), respectively.

A.2.1 Normal distribution fourth-order moments

In this section, we show that the fourth-order moments of a multivariate normal distribution are invariant to an orthonormal rotation matrix. Denote with $\mathcal{K}^z$ the matrix that collects the fourth-order moments of the *standard* normal distribution, which are given by

$$\mathcal{K}^z = I_{n^2} + K_{n,n} + \text{vec}(I_n)\text{vec}(I_n)';$$

see Kollo (2008) for more details. First, we show that $\mathcal{K}^z$ is invariant to orthonormal rotations. Using the property of the commutation matrix[23] that $K_{n,n}(A\otimes B) = (B\otimes A)K_{n,n}$, it holds that the commutation matrix is invariant to any orthonormal rotation Ω, that is, $(\Omega\otimes\Omega)'K_{n,n}(\Omega\otimes\Omega) = K_{n,n}(\Omega\otimes\Omega)'(\Omega\otimes\Omega) = K_{n,n}(\Omega'\otimes\Omega')(\Omega\otimes\Omega) = K_{n,n}(\Omega'\Omega\otimes\Omega'\Omega) = K_{n,n}$. Moreover, using the relationship between the vectorization and Kronecker product, that is,$\text{vec}(ABC) = (C'\otimes A)\text{vec}(B)$, it holds that $(\Omega\otimes\Omega)'(\text{vec}(I_n)\text{vec}(I_n)')(\Omega\otimes\Omega) = (\Omega'\otimes\Omega')\text{vec}(I_n)\ ((\Omega'\otimes\Omega')\text{vec}(I_n))' = \text{vec}(\Omega' I_n\Omega)\text{vec}(\Omega' I_n\Omega)' = \text{vec}(I_n)\text{vec}(I_n)'$. Therefore,

$$(\Omega\otimes\Omega)'\ \mathcal{K}^z\ (\Omega\otimes\Omega) = \mathcal{K}^z.$$

Denote with $\mathcal{K}^n$ the matrix that collects the fourth-order moments of the multivariate normal distribution with covariance Σ, which is given by

$$\mathcal{K}^n = (I_{n^2} + K_{n,n})(\Sigma\otimes\Sigma) + \text{vec}(\Sigma)\text{vec}(\Sigma)'.$$

Using the same properties of matrices and Kronecker products, it is straightforward to show that

$$\mathcal{K}^n = (\Sigma^{1/2}\otimes\Sigma^{1/2})'\ \mathcal{K}^z\ (\Sigma^{1/2}\otimes\Sigma^{1/2}).$$

[23] See e.g. Schott (2016) (Theorem 8.26 at page 342)

A.2.2 Third-order moments

Assume that $E(\nu_{n,t}^3) = \zeta_n \neq 0$. The $(n \times n^2)$ matrix collecting the structural shocks' third-order moments can be written as

$$\begin{aligned}
E(\nu_t\nu_t' \otimes \nu_t') &= E(\nu_t\nu_t' \otimes [\begin{pmatrix}\nu_{1,t} & ... & 0\end{pmatrix} + \cdots + \begin{pmatrix}0 & ... & \nu_{n,t}\end{pmatrix}]) = \\
&= E(\nu_{1,t}\nu_t\nu_t' \otimes \boldsymbol{e}_1' + ... + \nu_{n,t}\nu_t\nu_t' \otimes \boldsymbol{e}_n') = \\
&= E(\nu_{1,t}\nu_t\nu_t') \otimes \boldsymbol{e}_1' + ... + E(\nu_{n,t}\nu_t\nu_t') \otimes \boldsymbol{e}_n' = \\
&= \zeta_1 J_1 \otimes \boldsymbol{e}_1' + ... + \zeta_n J_n \otimes \boldsymbol{e}_n' = \\
&= \sum_{k=1}^{n} \zeta_k J_k \otimes \boldsymbol{e}_k'.
\end{aligned}$$

Using the properties of the Kronecker product, that is, $(A \otimes B)(C \otimes D) = AC \otimes BD$, notice that

$$\begin{aligned}
&\left(\sum_{i=1}^{n} \zeta_i J_i \otimes \boldsymbol{e}_i\right)\left(\sum_{i=1}^{n} \zeta_i J_i \otimes \boldsymbol{e}_i\right)' = \\
&\qquad = \sum_{i=1}^{n} \zeta_i^2 \left(J_i \otimes \boldsymbol{e}_i'\right)\left(J_i \otimes \boldsymbol{e}_i'\right)' + \underbrace{\sum_{i \neq k}^{n} \zeta_i\zeta_k \left(J_i \otimes \boldsymbol{e}_i'\right)\left(J_k \otimes \boldsymbol{e}_k'\right)'}_{=0} = \\
&\qquad = \sum_{i=1}^{n} \zeta_i^2 \left(J_i \otimes \boldsymbol{e}_i'\right)\left(\boldsymbol{e}_i \otimes J_i'\right) = \sum_{i=1}^{n} \zeta_i^2 \left(J_i \otimes \boldsymbol{e}_i'\right)\left(\boldsymbol{e}_i \otimes J_i'\right) = \\
&\qquad = \sum_{i=1}^{n} \zeta_i^2 J_i\boldsymbol{e}_i \otimes (J_i\boldsymbol{e}_i)' = \sum_{i=1}^{n} \zeta_i^2 \boldsymbol{e}_i \otimes \boldsymbol{e}_i' = \\
&\qquad = \sum_{i=1}^{n} \zeta_i^2 J_i' = \begin{pmatrix}\zeta_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \zeta_m^2\end{pmatrix} = \Lambda_\zeta,
\end{aligned}$$

where Λ_ζ is a diagonal matrix collecting the squared third-order moments of the structural shocks. Notice that the cross product are zero, since $(J_i \otimes \boldsymbol{e}_i')(\boldsymbol{e}_k \otimes J_k') = J_k\boldsymbol{e}_i \otimes \boldsymbol{e}_k'J_i' = 0$. Using again the property of the Kroeneker product, the third-order moments of the candidate structural shocks are given by

$$\begin{aligned}
E(\breve{\nu}_t\breve{\nu}_t' \otimes \breve{\nu}_t') &= A'E(\iota_t\iota_t'A \otimes \iota_t'A) = A'E(\iota_t\iota_t' \otimes \iota_t')(A \otimes A) = \\
&= A'E(A_o\nu_t\nu_t'A_o' \otimes \nu_t'A_o')(A \otimes A) = A'A_oE(\nu_t\nu_t \otimes \nu_t')(A_o' \otimes A_o')(A \otimes A) = \\
&= A'A_oE(\nu_t\nu_t \otimes \nu_t')(A_o'A \otimes A_o'A) = \\
&= A'A_o[\zeta_1 J_1 \otimes \boldsymbol{e}_1' + ... + \zeta_n J_n \otimes \boldsymbol{e}_n'](A_o'A \otimes A_o'A) = \\
&= \zeta_1(A'A_oJ_1 \otimes \boldsymbol{e}_1')(A_o'A \otimes A_o'A) + ... + \zeta_n(A'A_oJ_n \otimes \boldsymbol{e}_n')(A_o'A \otimes A_o'A) = \\
&= \zeta_1 A'A_oJ_1A_o'A \otimes \boldsymbol{e}_1'A_o'A + ... + \zeta_n A'A_oJ_nA_o'A \otimes \boldsymbol{e}_n'A_o'A.
\end{aligned}$$

Notice that for all $k = 1, ..., n$,

$$\begin{aligned}
\boldsymbol{e}_k' A_o' A &= \begin{pmatrix} \boldsymbol{\alpha}_k' \boldsymbol{a}_1 & \boldsymbol{\alpha}_k' \boldsymbol{a}_2 & \dots & \boldsymbol{\alpha}_k' \boldsymbol{a}_n \end{pmatrix}; \\
A' A_o J_k A_o' A &= \begin{pmatrix} \boldsymbol{a}_1' \boldsymbol{\alpha}_1 & \dots & \boldsymbol{a}_1' \boldsymbol{\alpha}_n \\ & \ddots & \\ \boldsymbol{a}_n' \boldsymbol{\alpha}_1 & \dots & \boldsymbol{a}_n' \boldsymbol{\alpha}_n \end{pmatrix} \begin{pmatrix} 0 & \dots & 0 \\ & 1 & \\ 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_1' \boldsymbol{a}_1 & \dots & \boldsymbol{\alpha}_1' \boldsymbol{a}_n \\ & \ddots & \\ \boldsymbol{\alpha}_n' \boldsymbol{a}_1 & \dots & \boldsymbol{\alpha}_n' \boldsymbol{a}_n \end{pmatrix} = \\
&= \begin{pmatrix} (\boldsymbol{a}_1' \boldsymbol{\alpha}_k)(\boldsymbol{\alpha}_k' \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_k' \boldsymbol{\alpha}_k)(\boldsymbol{\alpha}_k' \boldsymbol{a}_1) \\ & \ddots & \\ (\boldsymbol{a}_1' \boldsymbol{\alpha}_k)(\boldsymbol{\alpha}_k' \boldsymbol{a}_n) & \dots & (\boldsymbol{a}_n' \boldsymbol{\alpha}_k)(\boldsymbol{\alpha}_k' \boldsymbol{a}_n) \end{pmatrix} = \\
&= A' \boldsymbol{\alpha}_k \boldsymbol{\alpha}_k' A.
\end{aligned}$$

Therefore, for all $k = 1, ..., n$,

$$\begin{aligned}
E(\breve{\nu}_t \breve{\nu}_t' \otimes \breve{\nu}_t') &= \sum_k \zeta_k A' A_o J_k A_o' A \otimes \boldsymbol{e}_k' A_o' A = \\
&= \sum_k \zeta_k \begin{pmatrix} \boldsymbol{\alpha}_k' \boldsymbol{a}_1 \times A' A_o J_k A_o' A & \boldsymbol{\alpha}_k' \boldsymbol{a}_2 \times A' A_o J_k A_o' A & \dots & \boldsymbol{\alpha}_k' \boldsymbol{a}_n \times A' A_o J_k A_o' A \end{pmatrix} = \\
&= \begin{pmatrix} \sum_k \zeta_k \boldsymbol{\alpha}_k' \boldsymbol{a}_1 \times A' A_o J_k A_o' A & \sum_k \zeta_k \boldsymbol{\alpha}_k' \boldsymbol{a}_2 \times A' A_o J_k A_o' A & \dots & \sum_k \zeta_k \boldsymbol{\alpha}_k' \boldsymbol{a}_n \times A' A_o J_k A_o' A \end{pmatrix} = \\
&= \begin{pmatrix} \Phi^{(1)} & \Phi^{(2)} & \dots & \Phi^{(n)} \end{pmatrix}.
\end{aligned}$$

As $E(\breve{\nu}_{n,t}^3)$ occupies the (n, n^2) position in the matrix $E(\breve{\nu}_t \breve{\nu}_t' \otimes \breve{\nu}_t')$, we focus on the $(n, n)-$element of $\Phi^{(m)}$, that is,

$$\begin{aligned}
\Phi^{(n)} &= \sum_k \zeta_k \boldsymbol{\alpha}_k' \boldsymbol{a}_n \times A' A_o J_k A_o' A = \\
&= \zeta_1 \boldsymbol{\alpha}_1' \boldsymbol{a}_n \times A' A_o J_1 A_o' A + \dots + \zeta_n \boldsymbol{\alpha}_n' \boldsymbol{a}_n \times A' A_o J_n A_o' A.
\end{aligned}$$

If $\boldsymbol{a}_n = \boldsymbol{\alpha}_n$, then $\boldsymbol{\alpha}_j' \boldsymbol{a}_n = \boldsymbol{\alpha}_j' \boldsymbol{\alpha}_n = 0$ with $j \neq n$ and $\boldsymbol{\alpha}_n' \boldsymbol{a}_n = \boldsymbol{\alpha}_n' \boldsymbol{\alpha}_n = 1$. Hence,

$$\Phi^{(n)} = \zeta_n A' A_o J_n A_o' A.$$

The $(n, n)-$element of $\Phi^{(n)}$ equals $\zeta_m (\boldsymbol{a}_n' \boldsymbol{\alpha}_n)(\boldsymbol{\alpha}_n' \boldsymbol{a}_n) = \zeta_m$. Finally, notice that if $\zeta_n = 0$, the impact column vector $\boldsymbol{a}_n$ of the matrix A is not identified.

A.2.3 Fourth-order moments

Assume that $E(\nu_{n,t}^4) = \xi_n \neq 3$. The $(n^2 \times n^2)$ matrix collecting the full set of structural shocks' fourth-order moments can be written as

$$
\begin{aligned}
\mathcal{K} = E(\nu_t\nu_t' \otimes \nu_t' \otimes \nu_t) &= E\left((\nu_{1,t}\nu_t\nu_t' \otimes \boldsymbol{e}_1' + ... + \nu_{n,t}\nu_t\nu_t' \otimes \boldsymbol{e}_n') \otimes \begin{pmatrix} \nu_{1,t} \\ \vdots \\ \nu_{n,t} \end{pmatrix} \right) = \\
&= E\left((\nu_{1,t}\nu_t\nu_t' \otimes \boldsymbol{e}_1' + ... + \nu_{n,t}\nu_t\nu_t' \otimes \boldsymbol{e}_n') \otimes (\nu_{1,t}\boldsymbol{e}_1 + ... + \nu_{n,t}\boldsymbol{e}_n) \right) = \\
&= E(\nu_{1,t}^2\nu_t\nu_t') \otimes \boldsymbol{e}_1' \otimes \boldsymbol{e}_1 + ... + E(\nu_{n,t}^2\nu_t\nu_t') \otimes \boldsymbol{e}_n' \otimes \boldsymbol{e}_n + \sum_{j\neq k} E(\nu_{j,t}\nu_{k,t}\nu_t\nu_t') \otimes \boldsymbol{e}_j' \otimes \boldsymbol{e}_k = \\
&= \begin{pmatrix} E\nu_{1,t}^4 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & E\nu_{1,t}^2 E\nu_{n,t}^2 \end{pmatrix} \otimes \boldsymbol{e}_1' \otimes \boldsymbol{e}_1 + ... + \begin{pmatrix} E\nu_{1,t}^2 E\nu_{n,t}^2 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & E\nu_{n,t}^4 \end{pmatrix} \otimes \boldsymbol{e}_n' \otimes \boldsymbol{e}_n \\
&\quad + \sum_{j\neq k} \sum_{j,k=1} \begin{pmatrix} E(\nu_{1,t}^2\nu_{j,t}\nu_{k,t}) & \cdots & E(\nu_{1,t}\nu_{n,t}\nu_{j,t}\nu_{k,t}) \\ & \ddots & \\ E(\nu_{1,t}\nu_{n,t}\nu_{j,t}\nu_{k,t}) & \cdots & E(\nu_{n,t}^2\nu_{j,t}\nu_{k,t}) \end{pmatrix} \otimes J_{k,j} = \\
&= \begin{pmatrix} \xi_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & 1 \end{pmatrix} \otimes \boldsymbol{e}_1' \otimes \boldsymbol{e}_1 + ... + \begin{pmatrix} 1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \xi_n \end{pmatrix} \otimes \boldsymbol{e}_n' \otimes \boldsymbol{e}_n + \sum_{j\neq k} (J_{j,k} + J_{k,j}) \otimes J_{k,j} = \\
&= ((\xi_1 - 1)J_1 + I_n) \otimes J_1 + ... + ((\xi_n - 1)J_n + I_n) \otimes J_n + \sum_{j\neq k} (J_{j,k} + J_{k,j}) \otimes J_{k,j} = \\
&= \sum_{i=1}^{n} (\xi_i - 1)J_i \otimes J_i + I_n \otimes (J_1 + \cdots + J_n) + \sum_{j\neq k} (J_{j,k} + J_{k,j}) \otimes J_{k,j} = \\
&= \sum_{i=1}^{n} (\xi_i - 1)J_i \otimes J_i + I_{n^2} + \sum_{j\neq k} \sum_{j,k=1}^{n} (J_{j,k} + J_{k,j}) \otimes J_{k,j}.
\end{aligned}
$$

Since $\boldsymbol{e}_i' \otimes \boldsymbol{e}_i = J_i$ and $\boldsymbol{e}_j' \otimes \boldsymbol{e}_k = J_{k,j}$, for any i, j, k; and $I_n = J_1 + \cdots + J_n$. It is convenient to express the fourth-order moments in deviation from the standard normal distribution analogs. To this end, define $\xi_i = x_i + 3$; when $x_i = 0$, then the fourth-order moments coincide with the standard normal distribution ones. We can rewrite the fourth-order moments of the structural

shock as follows:

$$
\begin{aligned}
\mathcal{K} &= \sum_{i=1}^{n}(x_i+2)J_i\otimes J_i + I_{n^2} + \sum_{j\neq k}^{n}\sum_{j,k=1}(J_{j,k}+J_{k,j})\otimes J_{k,j} = \\
&= \sum_{i=1}^{n} x_i J_i\otimes J_i + I_{n^2} + 2\sum_{i=1}^{n} J_i\otimes J_i + \sum_{j\neq k}^{n}\sum_{j,k=1}[J_{j,k}\otimes J_{k,j} + J_{k,j}\otimes J_{k,j}] = \\
&= \sum_{i=1}^{n} x_i J_i\otimes J_i + I_{n^2} + \left[\sum_{i=1}^{n} J_i\otimes J_i + \sum_{j\neq k}^{n}\sum_{j,k=1} J_{j,k}\otimes J_{k,j}\right] + \left[\sum_{i=1}^{n} J_i\otimes J_i + \sum_{j\neq k}^{n}\sum_{j,k=1} J_{k,j}\otimes J_{k,j}\right] = \\
&= \sum_{i=1}^{n} x_i J_i\otimes J_i + I_{n^2} + \sum_{j,k=1}^{n} J_{j,k}\otimes J_{k,j} + \sum_{j,k=1}^{n} J_{k,j}\otimes J_{k,j} = \\
&= \sum_{i=1}^{n} x_i J_i\otimes J_i + \underbrace{I_{n^2} + K_{n,n} + \mathrm{vec}(I_n)\mathrm{vec}(I_n)'}_{\mathcal{K}^z} = \\
&= \sum_{i=1}^{n} x_i J_i\otimes J_i + \mathcal{K}^z,
\end{aligned}
$$

where $K_{n,n}$ is the commutation matrix, vec is the column-wise vectorization of the matrix, and $\mathcal{K}^z$ denotes the matrix that collects the fourth-order moments of the standard normal distribution. Therefore, we obtain that the excess fourth-order moments of the structural shocks are given by the following sum of diagonal matrices:

$$
E(\nu_t\nu_t'\otimes\nu_t'\otimes\nu_t) - \mathcal{K}^z = \sum_{i=1}^{n} x_i J_i\otimes J_i.
$$

It is straightforward now to derive the fourth-order moments of the candidate structural shocks:

$$
\begin{aligned}
E(\breve{\nu}_t\breve{\nu}_t'\otimes\breve{\nu}_t'\otimes\breve{\nu}_t) &= E(A'\iota_t\iota_t'A\otimes\iota_t'A\otimes A'\iota_t) = E(A'[(\iota_t\iota_t'\otimes\iota_t')(A\otimes A)]\otimes A'\iota_t) = \\
&= E((A'\otimes A')[(\iota_t\iota_t'\otimes\iota_t')(A\otimes A)\otimes\iota_t]) = \\
&= E((A'\otimes A')[(\iota_t\iota_t'\otimes\iota_t')(A\otimes A)\otimes\iota_t\cdot 1]) = \\
&= E((A'\otimes A')[(\iota_t\iota_t'\otimes\iota_t'\otimes\iota_t)((A\otimes A)\otimes 1)]) = \\
&= (A'\otimes A')E(\iota_t\iota_t'\otimes\iota_t'\otimes\iota_t)(A\otimes A).
\end{aligned}
$$

With similar algebra, we derive the following:

$$
\begin{aligned}
E(\breve{\nu}_t\breve{\nu}_t'\otimes\breve{\nu}_t'\otimes\breve{\nu}_t) &= (A'\otimes A')(A_o\otimes A_o)E(\nu_t\nu_t'\otimes\nu_t'\otimes\nu_t)(A_o'\otimes A_o')(A\otimes A) = \\
&= (A'\otimes A')(A_o\otimes A_o)\left(\sum_{i=1}^{n} x_i J_i\otimes J_i + \mathcal{K}^z\right)(A_o'\otimes A_o')(A\otimes A) = \\
&= \mathcal{K}^z + \underbrace{(A'\otimes A')(A_o\otimes A_o)\left(\sum_{i=1}^{n} x_i J_i\otimes J_i\right)(A_o'\otimes A_o')(A\otimes A)}_{\mathcal{K}^\star}.
\end{aligned}
$$

In this case as well, with no departure from the normality (that is, $x_i = 0$ for all i), the rotation matrix cannot be identified. Since we focus on the kurtosis of the shock ordered last—we are interested in the $(n^2, n^2)-$element of the matrix $E(\breve{\nu}_t\breve{\nu}_t' \otimes \breve{\nu}_t' \otimes \breve{\nu}_t)$—we can disregard the first matrix $\mathcal{K}^z$ and only consider the $(n^2, n^2)-$ element of the matrix $\mathcal{K}^\star$.

$$\begin{aligned}
\mathcal{K}^\star &= (A' \otimes A')(A_o \otimes A_o)\left(\sum_{i=1}^{n} x_i J_i \otimes J_i\right)(A_o' \otimes A_o')(A \otimes A) = \\
&= (A'A_o \otimes A'A_o)\left(\sum_{i=1}^{n} x_i J_i \otimes J_i\right)(A_o'A \otimes A_o'A) = \\
&= \sum_{i=1}^{n} x_i (A'A_o J_i A_o' A) \otimes (A'A_o J_i A_o' A) = \\
&= \sum_{i=1}^{n} x_i (A'\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' A) \otimes (A'\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' A).
\end{aligned}$$

The $(n^2, n^2)-$element of $\mathcal{K}^\star$ is the sum of the square of the $(n, n)-$elements of the matrices $A'\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' A$ times x_i for $i = 1, \ldots, n$, which is given by

$$\begin{aligned}
\mathcal{K}^\star_{(n^2,n^2)} &= x_1(\boldsymbol{a}_n'\boldsymbol{\alpha}_1)^2(\boldsymbol{\alpha}_1'\boldsymbol{a}_n)^2 + \cdots + x_m(\boldsymbol{a}_n'\boldsymbol{\alpha}_n)^2(\boldsymbol{\alpha}_n'\boldsymbol{a}_n)^2 = \\
&= x_1(\boldsymbol{a}_n'\boldsymbol{\alpha}_1\boldsymbol{\alpha}_1'\boldsymbol{a}_n)^2 + \cdots + x_m(\boldsymbol{a}_n'\boldsymbol{\alpha}_n\boldsymbol{\alpha}_n'\boldsymbol{a}_n)^2.
\end{aligned}$$

Hence, $\mathcal{K}^\star_{(n^2,n^2)} + \mathcal{K}^z_{(n^2,n^2)} = x_n + 3 = \xi_n$, if $\boldsymbol{a}_n = \boldsymbol{\alpha}_n$ (since $\boldsymbol{\alpha}_j'\boldsymbol{a}_n = 0$ if $j \neq n$ and $\boldsymbol{\alpha}_n'\boldsymbol{a}_n = 1$).

A.2.4 Eigenvector decomposition of fourth-order moments

The eigenvector decomposition of fourth-order moments is based on Kollo (2008), and it requires more notation. The approach first computes the sum of the n^2 blocks of $(n \times n)$ sub-matrices of the fourth-order moment matrix and then takes the eigenvalue/vector decomposition of the resulting matrix. The eigenvectors associated with nonzero eigenvalues coincide with the columns of the original rotation matrix up to a sign switch and permutation of columns.

Definition Let A be an $m \times n$ matrix and B an $mr \times ns$ partitioned matrix consisting of $r \times s$ blocks $B_{i,j}$, $i = 1, ..., m$ and $j = 1, ..., n$. The star product $A \star B$ of A and B is an $r \times s$ matrix such that

$$A \star B = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}B_{i,j}.$$

Define with $\mathcal{I}_n$ the $n \times n$ matrix of ones. Notice that

$$\begin{aligned}
\mathcal{I}_n \star E(\iota_t\iota_t' \otimes \iota_t' \otimes \iota_t) &= \mathcal{I}_m \star (A_o \otimes A_o)(\mathcal{K} - \mathcal{K}_z)(A_o \otimes A_o)' = \\
&= \sum_{i,j=1}^{n} (A_o \otimes \alpha_i)(\mathcal{K} - \mathcal{K}_z)(\alpha_j \otimes A_o)'.
\end{aligned}$$

We use Lemma 1 in Kollo (2008), which we report here

Lemma Let M, N, U and V be $n \times n$ matrices and a, b be $n \times 1$-vectors. Then,

$$(M \otimes a')diag(K_{n,n})(U \otimes V)diag(K_{n,n})(b \otimes N) = Mdiag(a)(U \circ V)diag(b)N,$$

where $\circ$ denotes the element-wise (Hadamard) product of matrices, and *diag* is the diagonal matrix constructed on the vector.

Let Λ_ξ be the diagonal matrix collecting the excess fourth-order moments of the structural shocks, that is, $E(\nu_{i,t}^4) - 3$; we have that

$$\mathcal{K} - \mathcal{K}_z = diag(K_{n,n})(\Lambda_\xi \otimes I_n)diag(K_{n,n}).$$

Then,

$$\begin{aligned}
\mathcal{I}_m \star \Xi_4 &= \sum_{i,j=1}^{m} (A_o \otimes \alpha_i)(\mathcal{K} - \mathcal{K}_z)(\alpha_j \otimes A_o)' = \\
&= \sum_{i,j=1}^{m} (A_o \otimes \alpha_i)diag(K_{n,n})(\Lambda_\xi \otimes I_n)diag(K_{n,n})(\alpha_j \otimes A_o)' = \\
&= \sum_{i,j=1}^{m} A_o(\Lambda_\xi \circ I_n)diag(\alpha_i)diag(\alpha_j)A_o' = \\
&= A_o \left(\sum_{i,j=1}^{m} (\Lambda_\xi \circ I_n)diag(\alpha_i)diag(\alpha_j) \right) A_o',
\end{aligned}$$

where the matrix in square brackets is diagonal. Therefore, we can compute the eigenvalue/vetor decomposition of the $n \times n$ matrix $\mathcal{I}_m \star E(\iota_t \iota_t' \otimes \iota_t' \otimes \iota_t)$ and retrieve the impact matrix.

A.3 Examples

A.3.1 Bivariate case: Third-order moments

To derive the identified set, we use the following trigonometric identities:

$$\begin{aligned}
a\cos x + b\sin x &= \operatorname{sgn}(a)\sqrt{a^2+b^2}\cos\left(x + \arctan\left(-\frac{b}{a}\right)\right); \\
x &= \arctan(\tan(x)); \quad \tan x = \frac{\sin x}{\cos x}; \quad \tan(-x) = -\tan x.
\end{aligned}$$

The higher-order moment restrictions is then

$$\begin{aligned}
(\cos\theta\cos\theta_o + \sin\theta\sin\theta_o)^3 &> 0 \\
(\operatorname{sgn}(\cos\theta_o)\cos(\theta - \theta_o))^3 &> 0.
\end{aligned}$$

A.4 Robust estimates of kurtosis and skewness

Estimates of kurtosis and skewness based on sample fourth- and third-order moments can be very sensitive to outliers with short samples (see Kim and White (2004)). When shocks follow Student-t distributions with low degrees of freedom, not even a million draws are enough to deliver unbiased median estimates of kurtosis. A number of robust measures have been proposed in the literature, and they all perform well for sample sizes typical in macroeconomics. These robust estimators are constructed using ratios of distance between different percentiles of the empirical distribution. Loosely speaking, robust measures of kurtosis consider the ratio between the distance of the percentiles in the tails and the distance between the percentiles close to the median. The larger the numerator, the thicker the tails; the smaller the denominator, the more clustered the distribution around the median. Hence, with distributions centered on zero, realizations are often very small but sometimes quite large. Robust measures of skewness exploit the distance between median and mean. More formally, letting x be a random variable with cumulative density function F, the robust measures of kurtosis considered are:

- Moors (1988) kurtosis:

$$\mathcal{K}_m(x) = \frac{(p_7 - p_5) + (p_3 - p_1)}{p_2 - p_4},$$

  where p_j represents the octile of the empirical distribution of x, that is, $p_j = F^{-1}(j/8)$, with $j = 1, ..., 7$. If x follows a Gaussian distribution, then $\mathcal{K}_m(x)$ equals 1.23. Therefore, the excess kurtosis is given by $\mathcal{EK}_m(x) = \mathcal{K}_m(x) - 1.23$.

- Hogg (1972) kurtosis:

$$\mathcal{K}_h(x) = \frac{u_{0.05} - l_{0.05}}{u_{0.5} - l_{0.5}},$$

  where $u_\alpha(l_\alpha)$ is the average of the upper (lower) α percentile of the distribution of x. If x follows a Gaussian distribution, then $\mathcal{K}_h(x)$ equals 2.59. Therefore, the excess kurtosis is given by $\mathcal{EK}_h(x) = \mathcal{K}_h(x) - 2.59$.

- Crow and Siddiqui (1967) kurtosis:

$$\mathcal{K}_{cs}(x) = \frac{F^{-1}(0.975) - F^{-1}(0.025)}{F^{-1}(0.75) - F^{-1}(0.25)},$$

  where $F^{-1}(\alpha)$ is the α percentile of the distribution of x. If x follows a Gaussian distribution, then $\mathcal{K}_{cs}(x)$ equals 2.91. Therefore, the excess kurtosis is given by $\mathcal{EK}(x)_{cs} = \mathcal{K}_{cs}(x) - 2.91$.

Robust measures of skewness considered are:

- Bowley (1926) skewness:

$$\mathcal{S}_b(x) = \frac{p_3 + p_1 - 2 \times p_2}{p_3 - p_1}$$

where p_j represents the quartiles of the empirical distribution of the candidate shock, i.e. $p_j = F^{-1}(j/4)$ with $j = 1, 2, 3$.

- Groeneveld and Meeden (1984) skewness:

$$\mathcal{S}_{gm}(x) = \frac{\text{mean} - \text{median}}{E|x - \text{median}|},$$

where $E|x - \text{median}|$ represents the average of the absolute deviation from the median.

- Kendall and Stewart (1977) skewness:

$$\mathcal{S}_{ks}(x) = \frac{\text{mean} - \text{median}}{\sigma},$$

where σ is the standard deviation of the empirical distribution of the candidate shock.

Clearly, all these robust measures of skewness are zero with the normal distribution.

A.5 Estimation and Identification

In this section, we briefly describe our estimation and identification strategy for estimating a VAR with non-Gaussian errors and for constructing IRFs using higher-order moment restrictions. Let a $VAR(p)$ be

$$y_t = \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + \Phi_0 + u_t,$$

where y_t is $n \times 1$ vector of endogenous variables, Φ_0 is a vector of constant, and Φ_j are $n \times n$ matrices. We assume $y_0, \ldots, y_{-p+1}$ are fixed. We assume that u_t are independently and identically distributed (i.i.d) zero mean random vectors with unconditional covariance matrix Σ. We assume that the VAR reduced-form shocks are linear combinations of the unobserved structural shocks, ν_t, that is,

$$u_t = \Sigma^{1/2} \iota_t = \Sigma^{1/2}\ \Omega\ \nu_t,$$

where $\Sigma^{1/2}$ is the Cholesky factorization of Σ, and Ω is an orthonormal matrix, that is, $\Omega\Omega' = \Omega'\Omega = I$. The structural shocks, ν_t, are zero mean orthogonal shocks with unitary variance, that is, $\nu_t \sim (0, I)$.

Standard inference on VAR parameters typically postulates a multivariate normal distribution for the reduced-form innovations. Such an assumption cannot be considered in our context. We propose adopting a robust Bayesian approach, which allows us to construct posterior credible sets without needing distributional assumptions of the reduced-form residuals. Our Bayesian approach builds on the work by Petrova (2022), who proposes a robust and computationally fast Bayesian procedure to estimate the reduced-form parameters of the VAR in the presence of non-Gaussianity. While Bayesian inference about the autoregressive coefficients is asymptotically unaffected by the distribution of the error terms, inferences on the intercept and the covariance matrix are invalid in the presence of skewness and kurtosis.

The robust approach relies on the asymptotic normality of the quasi-maximum likelihood (QML) estimator[24] of reduced-form parameters, autoregressive coefficients, and covariances; Petrova (2022) derives the closed-form expression for the asymptotic covariance matrix of the QML estimator allowing for fast simulation from its asymptotic distribution. In the case of symmetric distribution (no skewness), she shows that asymptotic valid inference for Σ can be performed by drawing from the asymptotic normal distribution centered on the consistent estimator of Σ, that is, the QML estimator $\hat{\Sigma}$, and with the covariance matrix equal to $\frac{1}{T}\left(\widehat{\mathcal{K}} - \text{vech}(\hat{\Sigma})\text{vech}(\hat{\Sigma})'\right)$, where $\widehat{\mathcal{K}}$ is a consistent estimator of the fourth-order moment of the VAR reduced-form shocks. Importantly, as mentioned previously, sample counterparts of the fourth-order moments can be poorly estimated and sensitive to outliers. Following Petrova (2022), we consider a shrinkage approach that tilts the sample fourth-order moment estimates of the reduced-form orthogonalized errors toward the normal distributed counterparts. In particular, the shrinkage estimator for the kurtosis is defined as

$$\widehat{\mathcal{K}}^{\star} = \frac{T}{T+\tau}\widehat{\mathcal{K}}_T + \frac{\tau}{T+\tau}D_n^{+}(I_n + K_{n,n} + \text{vec}(I_n)\text{vec}(I_n)')D_n^{+'}, \tag{2}$$

where $\widehat{\mathcal{K}}_T$ represents the sample fourth-order moments of the orthogonalized reduced-form residuals, that is, $\widehat{\mathcal{K}}_T = 1/T\sum \text{vech}(\iota_t\iota_t') \otimes \text{vech}(\iota_t\iota_t')$ with $\iota_t = \hat{\Sigma}^{-1/2}u_t$; $K_{n,n}$ is a commutation matrix, which is a $(n^2 \times n^2)$ matrix consisting of $n \times n$ blocks where the $(j,i)-$element of the (i,j) block equals one, and elsewhere there are all zeros; and D_n^{+} is the generalized inverse of the duplication matrix D_n.[25] The first part of the equation (2) represents the sample fourth-order moments and the second part the fourth-order moments implied by a standard normal distribution, with τ being the amount of shrinkage we assign to the normal implied moments; the larger this value, the more weight we give to the normality assumption.

It is important to highlight that the sample fourth-order moments are used only to construct the asymptotic covariance matrix of the reduced-form VAR errors' volatility matrix, which measures the asymptotic uncertainty around the consistent estimator of Σ. Sample fourth-order moments are not used for the shock's identification.

The posterior distribution of the autoregressive parameters conditional on Σ is standard, and any prior can be used for the purpose. When there are important departures from symmetry, the posterior distributions for the intercept term and the covariance matrix are not independent, even for large samples, so robust Bayesian inference requires consistent estimates of third-order moments. As for fourth-order moments, we consider the consistent shrinkage estimator given by $\widehat{\mathcal{S}}_T^{\star} = \frac{T}{T+\tau}\widehat{\mathcal{S}}_T$, where $\widehat{\mathcal{S}}_T = (1/T\sum \text{vech}(u_t u_t') \otimes u_t)$.

For inferential purposes, it is useful to rewrite the VAR in a seemingly unrelated regression

[24] The QML is the maximum estimator of quasi-likelihood, which, in this context, coincides with the likelihood of the VAR when *incorrectly* assuming normality of the reduced form residuals.

[25] For more details on the notation for the multivariate kurtosis, see Kollo (2008).

(SUR) format. Letting $k = np + 1$, we obtain

$$\underbrace{Y}_{T\times n} = \underbrace{X}_{T\times k}\underbrace{\Phi}_{k\times n} + \underbrace{E}_{T\times n},$$

with

$$Y = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_T' \end{pmatrix} = \begin{pmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,n} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,n} \\ \vdots & & & \\ y_{T,1} & y_{T,2} & \cdots & y_{T,n} \end{pmatrix}, \quad X = \begin{pmatrix} \mathbf{x}_0' & 1 \\ \mathbf{x}_1' & 1 \\ \vdots & \\ \mathbf{x}_{T-1}' & 1 \end{pmatrix}, \quad \underset{(np\times 1)}{\mathbf{x}_t} = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix},$$

$$\Phi = \begin{pmatrix} \Phi_1' \\ \vdots \\ \Phi_p' \\ \Phi_0' \end{pmatrix}, \quad E = \begin{pmatrix} u_1' \\ \vdots \\ u_T' \end{pmatrix}.$$

Assuming a flat prior[26], the estimation identification procedure can then be summarized as follows. Letting $\widehat{S} = (Y - X\widehat{\Phi})'(Y - X\widehat{\Phi})$ and $\widehat{\Phi} = (X'X)^{-1}X'Y$, the steps of the Gibbs sampler are for $j = 1, ..., J$:

1. Draw $\Sigma^{(j)}$ from

$$N\left(\text{vech}(\widehat{S}), \widehat{\mathcal{C}}\right),$$

where $\widehat{\mathcal{C}} = \frac{1}{T}D_n^+\left(\widehat{S}^{1/2}\otimes\widehat{S}^{1/2}\right)D_n\left(\widehat{\mathcal{K}}^\star - \text{vech}(I_n)\text{vech}(I_n)'\right)D_n'\left(\widehat{S}^{1/2}\otimes\widehat{S}^{1/2}\right)'D_n^{+\prime}$ captures the fourth-order moments.

2. Conditional on $\Sigma^{(j)}$, draw $\Phi^{(j)}$ from

$$N\left(\widehat{\Phi}, \Sigma^{(j)}\otimes(X'X)^{-1}\right);$$

   i. in case of an asymmetric distribution, the intercept, Φ_0, is drawn from

$$N(\widehat{\Phi}_0 + \widehat{\mathcal{S}}^\star\widehat{\mathcal{C}}^{-1}\text{vech}(\Sigma^{(j)} - \widehat{S}), \Sigma^{(j)} - 1/T\widehat{\mathcal{S}}_T^\star\widehat{\mathcal{C}}^{-1}\widehat{\mathcal{S}}_T^{\star\prime}).$$

3. Draw $\breve{\Omega}$ from a uniform distribution with the Rubio-Ramírez et al. (2010) algorithm and

   I. compute the impulse response function and check if the sign restrictions are verified,

   II. compute the implied structural shocks

$$\breve{\nu}_t^{(j)} = \breve{\Omega}'\left(\Sigma^{(j)}\right)^{-1/2}(y_t - \Phi_1^{(j)}y_{t-1} - ... - \Phi_p^{(j)}y_{t-p} - \Phi_0^{(j)}),$$

   and determine if the higher-order moment inequality restrictions are satisfied.

If both [I] and [II] are satisfied, keep the draw $\Omega^{(j)} = \breve{\Omega}$; otherwise, repeat [I] and [II].

[26]Se the appendix for extending the Gibb sampler to informative priors.

After a suitable number of iterations, the draws represent the posterior distribution of interest. The estimation of the reduced-form parameters and the computation of the impulse responses and of the higher-order moments are performed using the toolbox described in Ferroni and Canova (2021).[27]

A.5.1 Informative priors

We assume a multivariate normal (MN) prior for the autoregressive parameters:

$$\Phi \sim N(\Phi_0, \Sigma \otimes V) \quad = (2\pi)^{-nk/2}|\Sigma|^{-k/2}|V|^{-n/2}\exp\left\{-\frac{1}{2}tr\left[\Sigma^{-1}(\Phi-\Phi_0)'V^{-1}(\Phi-\Phi_0)\right]\right\}.$$

Modify the second step of the Gibbs sample with the following:

$$\Phi|\Sigma, Y, X, \sim N(\overline{\Phi}, \Sigma \otimes (X'X + V^{-1})^{-1}),$$
$$\overline{\Phi} = (X'X + V^{-1})^{-1}(X'Y + V^{-1}\Phi_0).$$

A.6 Identified set

In this section, we construct the identified set obtained with sign restrictions and with sign and higher-order moment restrictions respectively without resorting to simulating the data from the DGP. We define the identified set as the superior and inferior bounds of the impulse responses over the set of possible accepted rotations. For a formal definition, see Wolf (2020) Appendix A.1, Definition 2. We consider the cases where the DGP is a simple three-equation New Keynesian model and the case where the DGP is an SW DSGE model.

A.6.1 New Keynesian model

Consider the simple NK example of section 4.1. The solution of the model is a VMA of order zero, that is,

$$\begin{pmatrix} y_t \\ \pi_t \\ i_t \end{pmatrix} = \underbrace{\frac{1}{1+\kappa\phi_\pi+\phi_y}\begin{pmatrix} \sigma_d & \phi_\pi\sigma_s & -\sigma_m \\ \kappa\sigma_d & -(1+\phi_y)\sigma_s & -\kappa\sigma_m \\ (\phi_y+\kappa\phi_\pi)\sigma_d & -\phi_\pi\sigma_s & \sigma_m \end{pmatrix}}_{B} \begin{pmatrix} \nu_t^d \\ \nu_t^s \\ \epsilon_t^m \end{pmatrix} =$$
$$= \Sigma^{1/2}\ A_o\ \epsilon_t,$$

where $\Sigma^{1/2} = \text{chol}(BB')$ and the **true** rotation matrix A_o is given by $A_o = \Sigma^{-1/2}B$. It follows that

$$\Sigma^{-1/2}\begin{pmatrix} y_t \\ \pi_t \\ i_t \end{pmatrix} = \iota_t = \ A_o\ \epsilon_t.$$

[27] Codes for replication can be found on the Github page.

Notice that the true rotation matrix is difficult to characterize analytically because of the Choleski decomposition needed to construct the orthogonalized innovations. Therefore, we characterize the identified set numerically. We characterize a candidate orthonormal rotation using the Givens rotations (which do not require simulation). In particular, a candidate rotation $A(\theta_1, \theta_2, \theta_3)$ is constructed as follows:

$$A(\theta_1, \theta_2, \theta_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 \\ 0 & \sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} \cos\theta_3 & -\sin\theta_3 & 0 \\ \sin\theta_3 & \cos\theta_3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

for some $\theta_j \in [0, \pi]$ and for $j = 1, 2, 3$. We discretize the latter interval with step of size 0.1 and construct a three-dimensional grid. We constructed 250,047 (that is, 63^3) different candidate rotations. For notation purposes, we drop the j argument in A.

Let A be a candidate rotation and $\breve{\nu} = A\iota$ a candidate structural shock. First, we check that the sign constraints are verified, that is, $i_t > 0$ and $\pi_t < 0$, on impact. Second, we determine if the higher-order moment restriction is verified by computing the excess kurtosis matrix of the candidate shocks, $\mathcal{K}^\star$, which is given by (see section A.2 for a derivation)

$$\mathcal{K}^\star = (A' \otimes A')(A_o \otimes A_o)\, \mathcal{K}_\epsilon^\star\, (A_o' \otimes A_o')(A \otimes A), \tag{3}$$

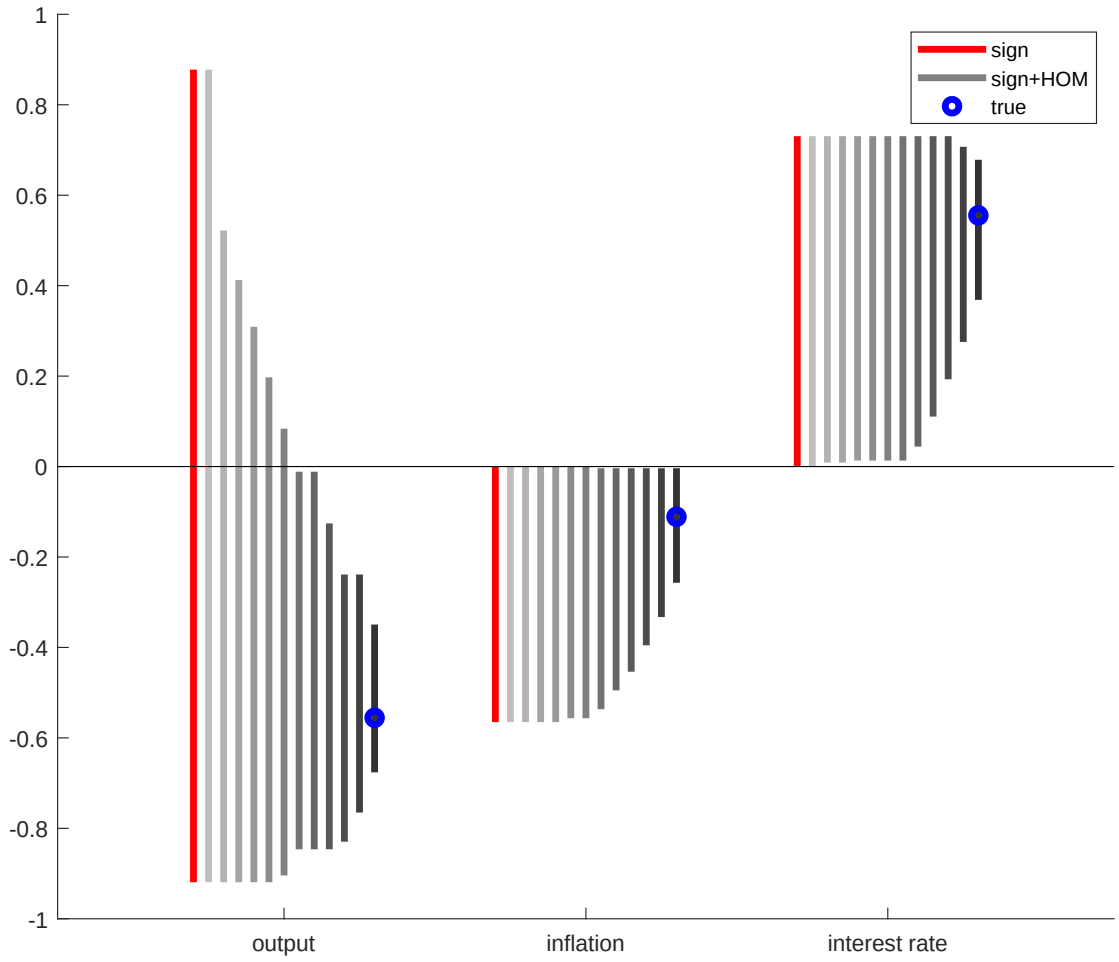
where $\mathcal{K}_\epsilon^\star$ is the excess kurtosis matrix of the structural shocks (that is, a matrix of zeros except a three in the $(9,9)^{th}$ position). To enforce the restriction, we keep the rotation if

$$\mathcal{K}^\star(9,9) > \texttt{threshold}.$$

If both requirements are verified, then the candidate rotation is accepted. We consider the `threshold` to vary between `[0:0.25:3]`.

Figure 22 reports the identified set, with sign restriction identification only displayed in red lines and the sign and higher-order moment restrictions in gray lines, where darker shades of gray indicate a larger `threshold`. When we set `threshold=0`, the identified sets with the sign restrictions and with sign and higher-order moment inequality restrictions coincide.

Figure 22: New Keynesian model identified sets. Identification obtained with sign restrictions only is displayed in red lines; identification obtained with both sign and higher-order moment inequality restrictions is in gray lines, with darker shades of gray indicating a larger threshold. True impact is in blue.



A.6.2 SW model

In this section, we consider the SW model equations at the posterior mode parameterization. As in the simple New Keynesian model example, we assume that the excess kurtosis of monetary policy shock is three and that the remaining shocks have zero excess kurtosis. Once the model is solved, it can be cast in the $ABCD$ representation, that is,

$$x_t = Ax_{t-1} + \tilde{B}\begin{pmatrix}\nu_t^d\\ \nu_t^s\\ \epsilon_t^m\end{pmatrix}$$
$$\begin{pmatrix}y_t\\ \pi_t\\ i_t\end{pmatrix} = Cx_{t-1} + \tilde{D}\begin{pmatrix}\nu_t^d\\ \nu_t^s\\ \epsilon_t^m\end{pmatrix},$$

where, in this case as well, the shocks are renormalized to have a unit variance, that is, $\tilde{B} = B\Sigma^{1/2}$ and $\tilde{D} = D\Sigma^{1/2}$. The SW model admits the following $VMA(\infty)$ representation (see, for example, Fernández-Villaverde, Rubio-Ramírez, Sargent and Watson (2007)):

$$y_t = \tilde{D}\epsilon_t + C\tilde{B}\epsilon_{t-1} + CA\tilde{B}\epsilon_{t-2} + \dots$$

Let Σ_w be the covariance of one-step-ahead prediction error, $w_t = y_t - E_{t-1}(y_t)$, which is a linear function of the structural shocks, $w_t = \tilde{D}\epsilon_t$. By construction, $\Sigma_w = \tilde{D}\tilde{D}'$. In this setup, the **true** rotation matrix A_o is given by

$$A_o = \Sigma_w^{-1/2}\tilde{D}.$$

The VMA of the SM model can be expressed as

$$\begin{aligned} y_t &= \underbrace{\Sigma_w^{1/2}}_{L_0}A_o\epsilon_t + \underbrace{C\tilde{B}A_o'}_{L_1}A_o\epsilon_{t-1} + \underbrace{CA\tilde{B}A_o'}_{L_2}A_o\epsilon_{t-2} + \dots \\ &= L_0\iota_t + L_1\iota_{t-1} + L_2\iota_{t-2} + \dots, \end{aligned} \tag{4}$$

where the last equation characterizes the impulse responses of the model to recursive representation of the one-step-ahead prediction error, w_t. Let F be the matrix of the impulse responses of interest for a candidate rotation A, that is,

$$F = \begin{pmatrix}L_0A\\ L_1A\end{pmatrix}.$$

As in the previous sections, for each of the candidate rotations constructed, A, we verify that

- after a monetary policy shock, $F(3,3) > 0$, $F(6,3) > 0$ ($i_t > 0$, $i_{t+1} > 0$) and $F(2,3) < 0$, $F(5,3) < 0$ ($\pi_t < 0$, and $\pi_{t+1} < 0$), and
- $\mathcal{K}^\star(9,9) >$ `threshold`, using the matrix in equation (3), where $\mathcal{K}_\epsilon^\star$ is the excess kurtosis matrix of the structural shocks (that is, a matrix of zero except a three in the $(9,9)^{th}$ position).

If both restrictions are satisfied, then the candidate rotation is accepted. In this setup as well, we consider different values of the `threshold`; in particular, we allow it to vary between `[0:0.25:3]`. Figure 23 reports the identified set, with the sign restriction identification displayed in red and the sign and higher-order moment restrictions in gray, where darker shades of gray indicate a larger `threshold`. When we set `threshold=0`, the identified sets with the sign restrictions and with sign and higher-order moment inequality restrictions coincide. When we set `threshold=1.5`, the identified sets are represented by the dashed black lines.

Figure 23: Smets and Wouters (2007) model identified set. Identification with sign restrictions only is displayed with red areas; identification with sign and higher-order moment inequality restriction identification is displayed with gray areas, with darker shades of gray indicating a larger threshold for kurtosis. Dashed black lines represent the identified set when the excess kurtosis of the candidate monetary policy shock is larger than 1.5. True impact and propagation are in blue.