WARWICK
ECONOMICS

Higher-Order Moment Inequality Restrictions for SVARs

Philippe Andrade, Filippo Ferroni & Leonardo Melosi

December 2024 No: 1537

Warwick Economics Research Papers

ISSN 2059-4283 (online)
ISSN 0083-7350 (print)

Higher-Order Moment Inequality Restrictions for SVARs*

Philippe Andrade
FRB Boston

Filippo Ferroni
FRB Chicago

Leonardo Melosi
University of Warwick,
EUI, DNB & CEPR

December 24, 2024

Abstract

We introduce a method that exploits some non-gaussian features of structural shocks to identify structural vector autoregressive models. More specifically, we propose to combine inequality restrictions on the higher-order moments of the structural shocks of interest with other set-identifying constraints, typically sign restrictions. We illustrate how, both in large or small sample settings, higher-moment restrictions considerably narrows the identification of monetary policy shocks compared to what is obtained with minimal sign restrictions typically used in the SVAR literature. The proposed methodology also delivers new insights on the macroeconomic effects of sovereign risk in the Euro Area, and on the transmission of geopolitical risk to the US economy.

Keywords: Shock identification, skewness, kurtosis, sign restrictions, monetary policy, sovereign risk, geopolitical risk.

JEL classification: C32, E27, E32

*An earlier draft of the paper was circulated under the title "Identification Using Higher-Order Moments Restrictions". We would like to thank Christian Wolf, Frank Schorfheide, Diego Känzig, Daniel Lewis, Tincho Almuzara, Fabio Canova, Raffaella Giacomini, Jonathan Wright, Lutz Kilian, 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. We would like to thank Mary Tracy for her outstanding research assistance. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of Boston, the Federal Reserve Bank of Chicago the De Nederlandsche Bank or any other person associated with the Federal Reserve System or the Eurosystem.

1 Introduction

Since the seminal contribution of Sims (1980), numerous structural VAR methods have been developed to study the propagation of unobserved unanticipated shocks to a dynamic system 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 not far from being Gaussian, higher moments restrictions are of little use to identification as they are close to being redundant with the ones imposed on second-order moments. However, macroeconomic and financial data often look non-gaussian with normal-time variations coexisting with unusually large positive or negative changes. Examples are the various types of financial crises potentially leading to sharp recessions, inflation surges induced by large supply disruptions, or abrupt changes in the stance of macroeconomic policies. These large and unpredictable events can be viewed as being triggered by structural shocks that are drawn from a distribution with fat tails of potentially asymmetric mass.

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

We postulate that the DGP of the economy is a VAR with uncorrelated non-gaussian structural shocks, therefore departing from gaussianity while keeping the simplicity of a linear transmission mechanism. In this setup, orthonormal rotations of the underlying shock processes leave their variance-covariance matrix unchanged but have an impact on their third and fourth moments and it is impossible to recover the "true impact" of a structural shock on the observable variables without reproducing its correct higher-moment. As we show, if these higher-order moments were to be 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 moments of structural shocks is challenging. Indeed, these shocks are not directly observable and so on need to work with estimated proxies. Moreover, and related, in-sample estimates of higher-order moments can be very imprecise for sample sizes typically encountered in macroeconomic applications. Overall, as Montiel Olea, Plagborg-Møller and Qian (2022) underlines, inference based on these higher-moments equality restrictions can be misleading. Moreover, 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 moments—typically the third and/or the fourth moments—of structural shocks of interest that we postulate to be non-gaussian. Such postulated non-gaussian features can be motivated by indirect

[1] See Ramey (2016), Nakamura and Steinsson (2018b) for surveys.

empirical evidence or by economic reasoning. Rather than point identification, higher-order moment inequality restrictions are set-identifying conditions. They can thus be combined with other set-identification constraints derived from economic analysis, typically sign restrictions. We show analytically how the identified set shrinks when we impose an inequality restriction on the third moment of the structural shock in a simple illustrative case.

The implementation of our method requires to deal with two difficulties. First, to address the potential small-sample bias issue 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] Second, to estimate and conduct 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 the performances of our method by applying it to simulated data from (i) a calibrated textbook New Keynesian model as in Galí (2015) but in which we assume a Laplace distribution for the monetary policy shock, and (ii) a medium scale DSGE Smets and Wouters (2007)'s model estimated on 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 impose that monetary policy shocks move the interest rate and prices move in opposite directions but leave the impact on real activity unrestricted – with a higher-order moment inequality restriction requiring that monetary policy shocks exhibit minimal excess-kurtosis.

As Wolf (2020) shows, while a monetary policy tightening has a contractionary impact on output in the models used to generate the data, Uhlig (2005)'s sign restrictions lead to estimate that it has no clear effect on output. In contrast, when combining these signs restrictions with an inequality restriction on the excess kurtosis of monetary policy shocks, we obtain that the effect is clearly contractionary. As we document, the reason is that our method strongly reduces the masquerading issue underlined by Wolf (2020), that is the fact that under Uhlig (2005)'s identification scheme combinations of positive supply and demand shocks, that have a positive effect on output, can masquerade as contractionary monetary policy shocks.

We also compare the performances of our approach with those of a method that exploits the independence of structural shocks and relies on exact equality constraints on their third and/or fourth moments. We show that, while this more restrictive approach allows to point identify

[2]Loosely speaking, robust estimators of the kurtosis for example 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 is the distribution around the median. Robust estimators of the skewness compute the distance between mean and median.

the structural shock of interest and performs very 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 getting unbiased in-sample estimates of higher-order moments then proves to be quite challenging as these are very sensitive to outliers and minor perturbation of the data.

We then use our identification method to study three empirical questions. In a first application, 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 documenting that while various measures of US monetary policy shocks proposed in the literature have a correlation that can be quite low, they all feature significant excess kurtosis compared to Gaussian processes. Furthermore, we also uncover that property in various monetary policy shock measures for the euro area and the UK. We thus exploit that robust feature to identify monetary policy shocks by imposing that their excess kurtosis is larger than a lower bound, which is derived from the empirical distributions of the various proxies available. As above, we combine this higher-order moment inequality restrictions with Uhlig (2005)'s agnostic sign restrictions. While the impact of monetary policy shock on output is non-significant when using sign restrictions only, our methodology leads to estimate that a monetary policy tightening induces a significant contraction in output. This is consistent with the results obtained with alternative ways to sharpen the identification of monetary policy derived from minimal sign restrictions, 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 questions that have been less investigated and for which restrictions on higher-order moments of structural shocks are quite natural. 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 Euro Area 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 structural VAR approach spirit, we propose an estimation that uses less specific structural assumptions. We estimate a macro-financial VAR model of the euro area which includes the Italian-German spread on sovereign bonds. We think of sovereign risk shocks as shocks that can 'relatively' frequently be very large and positive, that is which feature a positively skewed distribution with fat-tails. We therefore constraint that sovereign spread shocks meet some inequality constraints on their third and fourth moments. We combine these higher-order moment restrictions with the following minimal sign restrictions: a sovereign spread shock increases on impact and for the following month the 10 year Italian government bond yield, the Italian-German spread and the non-financial corporate borrowing costs. We find that an increase in sovereign risk leads to an immediate pronounced tightening of credit conditions. The real economy of the EA also

responds on impact after the shock and these effects are re-absorbed in less than three years after the shock. The effects on the unemployment rate of the EA are also quite severe and last for longer. In contrast, the macroeconomic impact of the sovereign shocks is found to be non-significant when one relies on sign restrictions to identify them.

Finally, in a third application, we investigate the impact that geopolitical risks have on the US economy. We use the geopolitical risk index developed by Caldara and Iacoviello (2022) which is based on newspaper article 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 that happen relatively frequently compared to the Gaussian case. We propose to identify geopolitical risk shocks by combining third and fourth-order moment inequality restrictions with some minimal sign restrictions which assume that geopolitical risk tighten financial conditions on impact. Consistent with what Caldara and Iacoviello (2022) obtain with the same VAR system but 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 lowers investment and hours. Our estimated effects are significantly larger and more persistent than what Caldara and Iacoviello (2022) obtain 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 and details the estimation strategy. Section 4 shows the role of kurtosis in identifying the effect of monetary policy on output in standard NK models. Sections 5-7 illustrate the usefulness of our approach in three empirical applications: measuring the real effects of the US conventional 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 the 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 rates 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 the structural shocks by minimizing the distance between the VAR

model-implied empirical innovation higher-order moments and sample counterparts. Finally, some studies exploit time variation in the conditional variance of shocks to identify them (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 moments 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, in particular by ruling out situations where shocks can be large at the same time, which may be less justified than restrictions based on economic reasoning. Second, they 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 of these being very sensitive to outliers for sample of limited size. Third, they 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 allow for co-skewness and co-kurtosis. We also treat the higher-order moments properties as necessary restrictions that structural shocks should fulfill. Importantly, and differently from previous contributions, this generates a set-identification of the shock of interest as opposed to the system point identification. This, in turn, allows us to couple these restrictions with other popular macroeconomic assumptions based on the signs (or zero) impact of the shock on the endogenous variables, magnitude or elasticity bounds or narrative restrictions on historical episodes. In this respect our approach does not substitute to economic reasoning. Moreover, our methodology does not rest on in-sample estimates of the empirical innovation higher-order moments. It imposes inequality restrictions on the higher moments of the structural shock itself using non-parametric robust methods based on the distance between different percentiles of the shock empirical distribution.

Our approach is also related to the works aiming at sharpening set-identified structural shocks obtained with restrictions derived from theory. Kilian and Murphy (2012) and Wolf (2020, 2022) underline that imposing minimal and uncontroversial sign restrictions alone is often too weak to provide adequate identification of 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 the recent methods introduced to solve that issue by either combining sign and narrative restrictions (Antolin-Diaz and Rubio-Ramírez 2018), by considering additional sign restrictions (Arias et al. 2019), or by relying on proxy-structural shocks 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).

Strikingly, our method delivers results that are comparable to the ones obtained with several

of these alternative approaches when applied to the identification of monetary policy shocks. In particular, it identifies that monetary policy tightening has a contractionary impact on GDP. This effect is sometimes less clear when using proxy-structural monetary shocks as external instruments because these proxies can also capture other macroeconomic news and therefore fail to meet the orthogonality condition of an instrument (see 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 rates futures around FOMC announcements are leptokurtic. We document that this feature can be found in a wide range of proxies for monetary policy shocks – including the recent measure by Aruoba and Drechsel (2022) who rely 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 be implemented to address questions where these alternatives methods are more difficult or less natural to implement.

Our paper is closely related 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 the higher moment restrictions is different from what we do. They refine the identification set by discarding rotations that are not consistent with statistical independence of the structural shocks and look at the matrix of higher order moments of candidate shocks to enforce that. Specifically, their methodology rules out shocks whose higher moments exhibit co-skewness or co-kurtosis. We do not impose independence. In contrast, we require that a specific structural shock has non-gaussian features. Moreover, their approach requires computing 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 are constructed on a handful of robust and non-parametric higher moments estimates and do not depend on the number of endogenous variables in the VAR.

Finally, our paper is connected to the literature arguing that rare and large shocks help to understand how financial markets price macroeconomic risks (Barro 2006, Gabaix 2012, Gourio 2012) and more generally to the works showing that macroeconomic data favors models featuring non-gaussian shocks (e.g. Cúrdia, Negro and Greenwald 2014). We assume that structural shocks are non-gaussian and use this properties to narrow the estimated effects of those shocks on macroeconomic variables.

2 Identification with higher-order moments

Let us consider that the econometrician observes a $n-$dimensional vector of *empirical innovations*, ι_t, that are *uncorrelated* white-noise processes with unit variance, i.e. $\iota_t \sim (0, I)$.

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 a $n-$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 moments, i.e. $E(\nu_{i,t}^3) = \zeta_i \neq 0$ 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

The higher order moments have been used in Independent Component Analysis (ICA) for reconstructing the original (demixed) sources of variations from a vector of mixed signals. The core result of this literature is that if at most one of the components ν_t is Gaussian, A_o is point identified up to sign change and permutation of its columns; therefore, all the structural shocks can be reconstructed from the empirical innovations. This result is shown by Comon (1994, Theorem 11) and has been discussed and used for identification in the literature, see also Gouriéroux et al. (2019) and Lanne et al. (2017).

Regardless of the number of non-gaussian shocks in the system considered, higher-order moments 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 how in the case where such shock features non-gaussian third or fourth moments.

A spectral decomposition approach The $(n \times n^2)$ matrix collecting the third 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 the $n \times 1$ vector with zeros everywhere except a one in the i^{th} position, J_i the $n \times n$ matrix of zeros everywhere except one in the i^{th} position of the main diagonal. The squared third moment matrix is thus a $(n \times n)$ diagonal matrix

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

[3] This assumption is made for exposure and analytical tractability but our method applies to cases where it is relaxed.

where Λ_ζ is a diagonal matrix collecting the squared third 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 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 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 moments of the structural shock and 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 non-zero third moment and differs from the other structural shock third moments, we can identify its impact on the empirical innovations. Note however that this requires using the full array of empirical innovations third 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 moment, in which case it is associated to the largest eigenvalue in the above spectral decomposition.

Similar arguments carry over to fourth moments. In fact, we can express the $(n^2 \times n^2)$ matrix collecting the fourth 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 (i.e. $x_i = \xi_i - 3$) and $\mathcal{K}^z$ is the matrix of fourth moments of a normal standard multivariate distribution. The latter is a diagonal matrix with non-zero elements only on the positions $j(n+1) - n$ for $j = 1, ..., n$. The matrix collecting the fourth 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 where the first largest n eigenvalues 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 moment, one can derive the column of the original rotation matrix by taking the first n elements of the first eigenvector and divide them by the absolute value of the first elements of the eigenvector, i.e. $P(1:n,1)/\sqrt{|P(1,1)|}$.

Remarks When all the elements of ν_t are non-gaussian, the procedure allows to point identify all the columns in A_o up to a permutation of columns and a sign switch.[4] The other way around, when ν_t is a Gaussian distributed random vector, moments higher than the second are not useful for identification since third moments are zero and fourth 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 which are worth emphasizing. These are reminiscent of the ones made by Montiel Olea et al. (2022). First, it requires that structural shocks have zero cross third and fourth moment, so it is not suited for situations where the economy becomes more volatile no matter the source of shocks. Second, the method requires to know 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, as is well-known, estimates of higher-order moments can be very sensitive to outliers or minor perturbation of the data and 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 compared to the other. Related to that, the procedure does not imply that the shocks have an economic contain, since they are obtained only from statistical properties and not from restrictions implied by economic reasoning.

2.2 Inequality restrictions

Rather than achieving point-identification from a set of potentially imprecise estimates of higher moments of the empirical innovations, our approach relies on inequality restrictions on the higher 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 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 are interested in

[4] When using restrictions on skewness for identification, A_o is made of the eigenvectors of the $n \times n$ matrix of squared third-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 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 as laboratory a standard macro New-Keynesian model.

the last (n^{th}) shock, and let $\boldsymbol{\alpha}_n$ bet 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 than the 'true' structural shock.*

The proof is 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 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 get the true rotation matrix without generating the correct higher 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 moments to lie in a preassigned interval. When the underlying data generating process is consistent with these inequality restrictions, that is the true structural shock exhibits deviation from gaussianity, higher-order moment inequality restrictions allows to shrink the set of rotations.

The other way around, if the 'true' moments of the structural shocks of interest are inconsistent with the restriction, the set of admissible rotations is empty. E.g., if we assume that monetary policy shocks are leptokurtic while they are not in the true data generating process, we will not find rotations satisfying the restrictions. It is important to highlight that similar considerations apply also to sign, magnitude or narrative restrictions; if we impose 'incorrect' restriction on –say– the IRF 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 data generating process.

Restrictions considered While a large set of restrictions can potentially be considered, we focus on restrictions on the third and fourth 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 restrictions is on the shape of the distribution of the structural shock which can be assumed to feature either asymmetry or 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 positive threshold, i.e. $\mathcal{S}(x) > \epsilon_{\mathcal{S}}$ ($\mathcal{S}(x) < \epsilon_{\mathcal{S}}$).*

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

These restrictions can be combined together in situations where 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 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 is positive, i.e. $|\mathcal{S}(x)| > \epsilon_\mathcal{S} > 0$ or $|K^\star(x)| > \epsilon_\mathcal{K} > 0$.*

Finally, another natural restriction is to impose a ranking on the third or on the fourth 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, i.e. 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 HoM inequality restrictions in practice By definition, one does not observe structural shocks. The higher order restrictions are structural identification assumptions that can be grounded either on theoretical insights or on empirical evidence using proxy for unobserved structural shocks. Our applications below 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. This amounts to determine a quantification of 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 ten (thirty) percent apart. For kurtosis, setting $\epsilon_K = .1.5$ amounts to assume that realizations larger than 3 standard deviations should occur at least once every hundred observations (as opposed to roughly one every seven hundred observations as in the Gaussian case). These numbers are clearly indicative and in general they depend on how much non-gaussianity a researcher is willing to impose. Alternatively, one can use values available in other studies in the literature that offer empirical estimates of the shock of interest, e.g. see section 5.1.

Remark that the modeler faces a trade-off when setting these bounds. Too restrictive bounds (too large deviations from gaussiantity), can return an empty set. The other way around, too loose bounds may lead to very imprecise identification.

Advantages That method overcomes each of these three limitations outlined above. First, it does not require to impose zero co-skewness or co-kurtosis across the structural shocks. Second, it only requires estimating the third and/or fourth 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 computing a large number of sensitive sample third or fourth moments which can lead to biased estimates. That risk is further mitigated by the fact that we impose an inequality rather than equality restriction, which makes the imprecision in the third or fourth moments estimates having less of an impact on the resulting set of estimates. Third, inequality restrictions achieve set-identification as opposed to point-identification. Rather than a drawback, this allows to combine these restrictions with other assumptions such as signs, zeros, narrative, magnitude and/or statistical independence. In this sense higher-moments inequality restrictions are *modular*, as they can be flexibly used standalone or in combination with other assumptions, in particular some that have economic contain.

2.3 An analytical example

We conclude this section offering an analytical example to illustrate the main point of the paper, that is inequality restrictions on higher moments of the structural shock can lead to shrink the identified set of the structural parameter. To this aim, 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 moments of the structural shocks are zero and one respectively, i.e. $E(\nu_{1,t}^3) = 0$ and $E(\nu_{2,t}^3) = 1$, and and cross second and third moments are zero, i.e. $E(\nu_{1,t}\nu_{2,t}) = E(\nu_{1,t}\nu_{2,t}^2) = E(\nu_{1,t}^2\nu_{2,t}) = 0$. The structural equations are defined by

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

where θ_o is the 'true' unknown angle of rotation with $\theta_o \in (-\pi/2, \pi/2)$ so to restrict to have positive entries in the main diagonal of the impact matrix A_o. 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 moments do not depend on θ_o, third moments do. In particular, the population third 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 of the third moments of the empirical innovation is given by

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

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

Since Ω is an idempotent matrix, we have 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 means that the first structural shock third moment equals zero and the second structural shock third moment equals one. The eigenvector associated with the non-zero eigenvalue is $\begin{pmatrix}-\sin\theta_o & \cos\theta_o\end{pmatrix}'$.

Our preferred approach does not use the third moments of the empirical innovations directly and imposes weaker restrictions. In particular, we assume that the second shock has positive skewness. Moreover, we do not make assumptions about the third moment of the other shock. Yet, we can restrict considerably 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, i.e. $A = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$ 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 which otherwise would be the set $\theta \in (-\pi/2, \pi/2)$; in particular, 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 moments are zero but the third 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

Our estimation and identification approach can be described as follows. We postulate that the observed data are generated by a $VAR(p)$ model,

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

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

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

noise process with unconditional covariance matrix Σ which is obtained as a linear combination of the unobserved structural shocks, ν_t, i.e.

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

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. Such an assumption cannot be considered in our context. We propose to adopt a robust Bayesian approach which allows to construct posterior credible sets which does not request to assume a specific distribution of the reduced form residuals.

Our approach builds on the work of Petrova (2022). She 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 non-symmetric 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 parameters posterior obtained using above. The procedure to identify the structural shock using higher-order restrictions goes 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;

III. Compute the implied structural shocks

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

[9] The QML is the maximum estimator of the quasi-likelihood. The quasi-likelihood 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.

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 are representative of 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 is performed using the toolbox described in Ferroni and Canova (2021).[11]

Robust higher-order moment estimators Our procedure requires to have estimates of third and fourth moments. As illustrated above, methods using higher-order moments 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). In contrast, our approach allows to rely on robust non-parametric estimators, therefore mitigating 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 above procedure. For third moments or asymmetries, robust measures exploit the standardized distance between median and mean. For fourth 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 is the distribution around the median and hence the more leptokurtic the shock is.

In what follows, 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 hom inequality restrictions help identify structural shocks? The impact of monetary policy shocks on output

We illustrate how inequality restrictions on higher moments can improve the identification of the effects of standard economic shocks on macro variables. In particular, we focus on the long debated impact of monetary policy on output.

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

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 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 (NK) model. The reason is that, when only sign restrictions on inflation and interest rate are imposed, supply and demand shocks tend to *masquerade* or disguise as monetary policy shocks. He concludes that pure sign restrictions are quite weak 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 moments inequality restrictions also strongly help resolving the issue of shocks *masquerading* as monetary ones when the latter have a non-gaussian fourth moment that is sufficiently different from supply and demand shocks. We first use simulated data from a standard calibration of the three equations NK model, as in Wolf (2020), but where the monetary policy shock is assumed to follow a Laplace distribution and therefore exhibits excess kurtosis compared to 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) or Smets and Wouters (2007), and that we estimate using US postwar data. Strikingly the distribution of estimated monetary policy shock exhibits excess kurtosis, and more so that the other estimated structural shocks.

4.1 Experimental evidence using a textbook New-Keynesian model

Specification, calibration, and simulations 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 admits 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 is assumed to follows a Laplace distribution, which excess kurtosis is 3:

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

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.

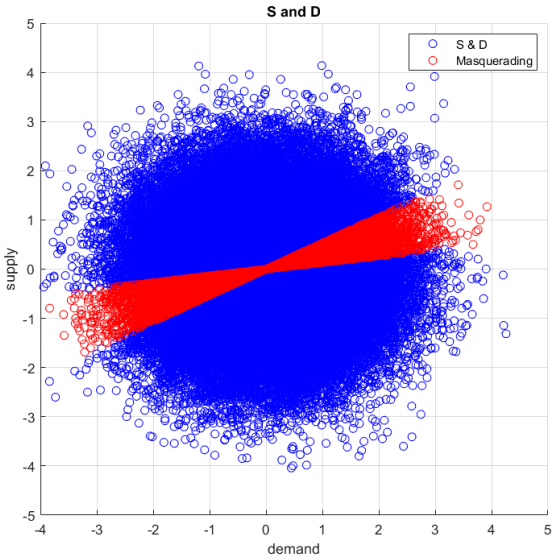
Masquerading shocks The masquerading effect of Wolf (2020) is illustrated in the left panel of figure 1. It displays a scatter plot of all the supply and demand shock realizations in our simulations and highlights in red circles realizations that generate a negative comovement between inflation and the interest rate. That latter configuration occurs when supply shocks are relatively small compared to 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 demand and supply shocks generates a pattern which 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 demand and supply shocks. As a result, imposing sign restrictions on inflation and interest rate only 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 report the distribution of the impact on output of monetary policy shocks identified with sign restriction only 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 symmetrically distributed around zero in a range of approximately $-.9$ and $.9$.

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, i.e. $\mathcal{K}_m^\star > \epsilon_\mathcal{K}$. with $\epsilon_\mathcal{K} = 1.5$. As illustrated by the red-colored bars in the right top panel of figure 1, the estimated impact distribution changes drastically compared to estimates obtained with sign restrictions only: Almost all the estimated positive responses of output gets chopped away.

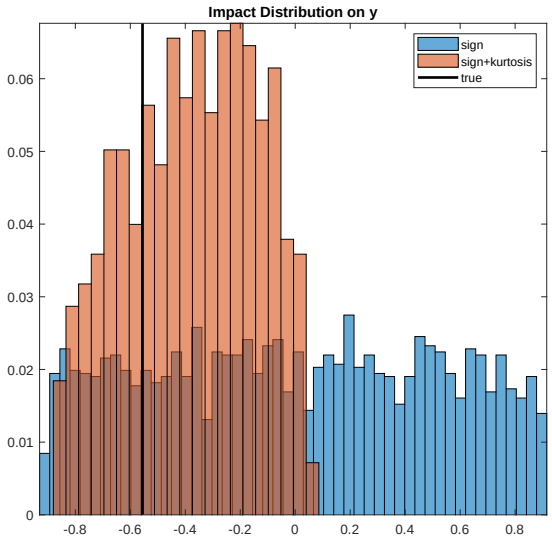
The lower panel of figure 1 shows (blue line) how the probability of obtaining an estimated response of output that is positive (instead of negative in the model) varies with different values

of the lower bound imposed on the fourth 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 $\epsilon_{\mathcal{K}} = 0$, roughly 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 roughly 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. Eventually it becomes zero when we impose that the excess kurtosis of monetary policy shocks needs to be larger than 1.8. So the masquerading effect disappears when on imposes a clear deviation from the gaussian case, but only roughly 60% of the true excess kurtosis of the shock.

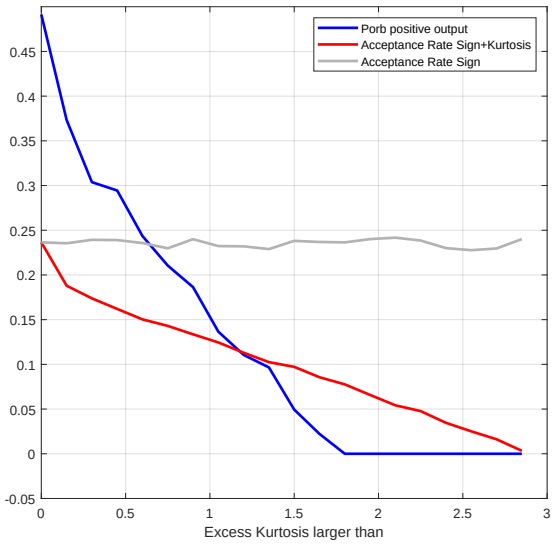
Analytical identified sets We also constructed the identified sets analytically and without resorting to simulations. The results are very similar than the findings above and are presented in section A.6.1.



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



(b) Distribution of MP 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); impact distribution of monetary policy on output (b); distribution of excess-kurtosis for the masquerading shock; and (c) probability of positive response of y after a monetary policy tightening.

4.2 Experimental evidence using a medium-scale DSGE model

We now study how higher-order moments inequality restrictions perform when applied to data obtained from simulations of a more realistic model which 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 rate are explained by handful of shocks: technology, risk premium, investment demand, monetary policy, government/exogenous spending, and price and wage markups shocks. We estimate the model on postwar US data on output, consumption, investment, real wages, inflation, interest rate and hours worked as in the original Smets and Wouters (2007)'s work.[12] We set parameters at their posterior mode estimates and consider smoothed estimates of the structural shocks as realizations of these shocks. We then use this model as a laboratory to show how the higher order moments can sharpen identification.

Supply and demand shocks can masquerade as monetary policy shocks A first property of the estimated SW model is that, as for the textbook NK model, demand and supply 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 rate in the SW model respectively; the last raw displays the dynamic transmission of the sum of supply and demand. Comparing this raw with the monetary policy one shows that the combination of supply and demand shocks can generate a sign pattern for inflation and interest rate that is similar to the one 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 have 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 shocks realizations (top panels) and the empirical probability distribution against the normal one (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), we find an

[12] Our implementation of the Smets-Wouters model is based on Dynare (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.

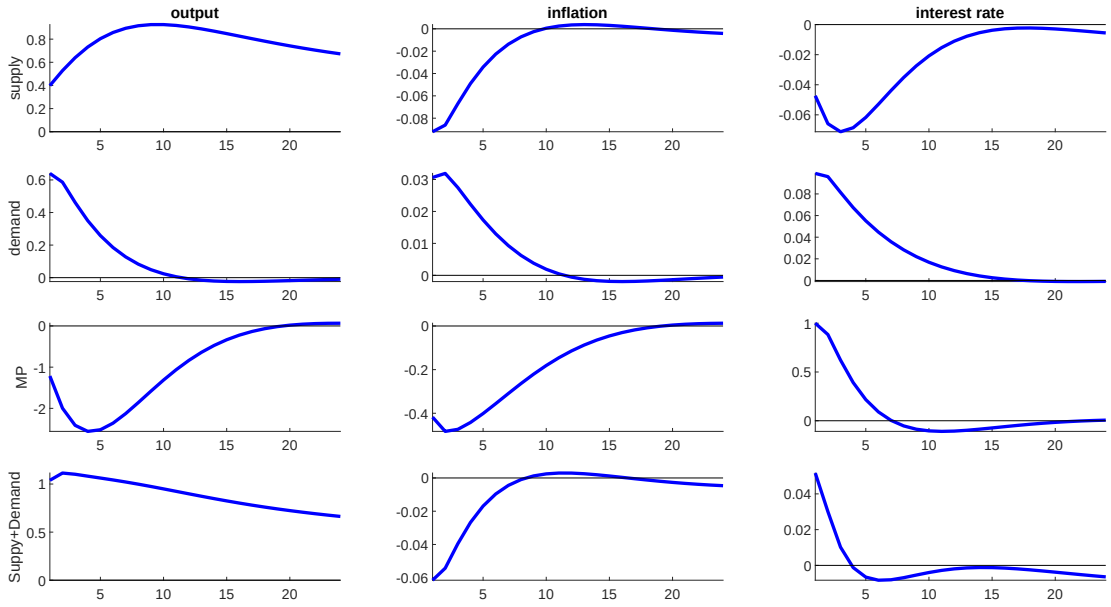


Figure 2: SW estimates of impulse response functions. From top to bottom technology, risk premium and monetary policy shocks and the sum of demand and supply shocks.

excess kurtosis for the SW monetary policy shock equal to 2, well above the excess kurtosis for other shocks. In contrast the robust measure of skewness is very close to zero for every structural shock.

Identifying monetary policy shocks in large samples We generate a sample of 50,000 observations for output, inflation and interest rate using random draws from the empirical distribution of the technology, risk premium and monetary policy shocks. We then assess how different structural VAR restrictions aiming at identifying monetary policy shocks perform on these simulated data. We compare three schemes: *(i)* sign restrictions only, *(ii)* sign and higher-order moments inequality restrictions as introduced in section 2.2, and *(iii)* higher-order moments 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 a robust measure of excess kurtosis larger than 1.6.[13] We estimate a VAR with twenty lags and

[13]That threshold value is chosen to be larger than roughly 80% of the value estimated in the empirical distribution. Thresholds lower than one do not alter the identified set compared to imposing only sign restrictions. As Figure 15 in the appendix illustrates, the probability that output is positive one year after a monetary policy tightening is about 40% both with sign and with sign and higher-order moment restrictions. This occurs because both demand and

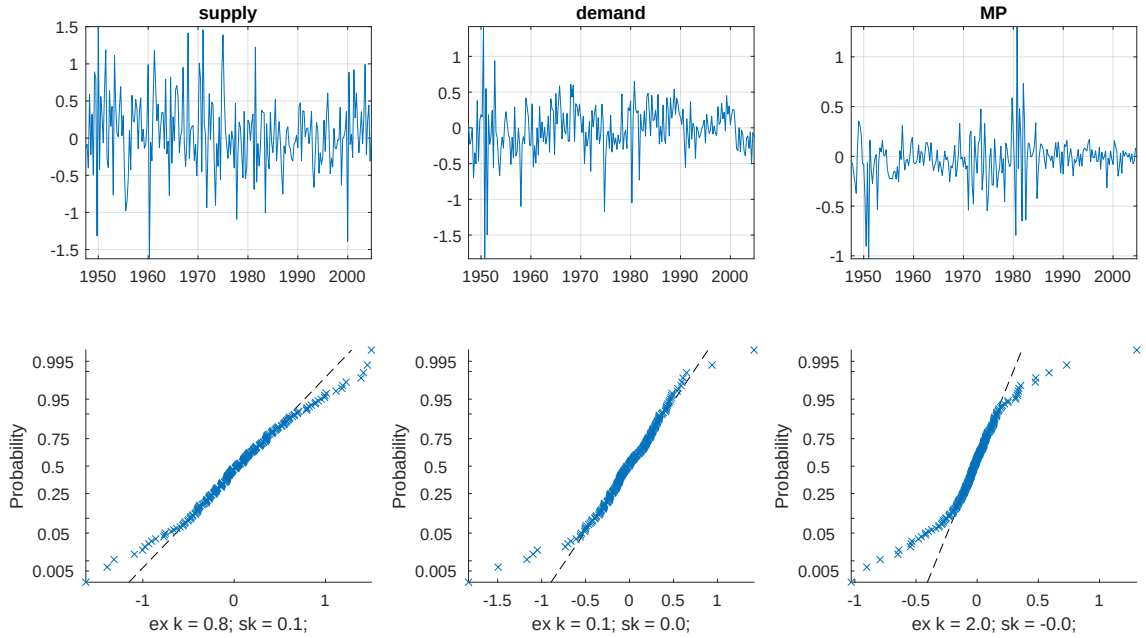


Figure 3: SW estimated shocks: form left to right technology, risk premium and monetary policy shocks. Top panels realizations, bottom panels probability distribution against the normal.

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 moment inequality restriction spelled out in *(ii)* is also verified and keep the rotation if so. Finally, in *(iii)*, the rotation matrix is derived from the spectral decomposition of the matrix of fourth moments of the VAR empirical innovations.

Figure 4 reports the impulse responses obtained with each identification schemes. The first row reports the 90% and 99% identified sets of impulse responses obtained when using sign restrictions only. 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 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 NK 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 refining the transmission of monetary policy shocks to output. In the case of inequality restrictions, the 90% identified set suggests that most of the output trajectories

supply 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 seeing the higher-order moments identified set to shrink relative to the sign restricted identified set.

are negative at least after few quarters in line with what the theory predicts.[14] Interestingly, the identification obtained with the spectral decomposition of the empirical innovation fourth moment matrix is very precise. However, with shorter samples, such precision deteriorates and inequality restrictions perform better as we show in the next section.

Identifying monetary policy shocks in short samples The identification based on the spectral decomposition is very precise when working with very large sample of data. However, the precision deteriorates sharply when working with smaller samples of size typical of macroeconomic studies. Indeed, this approach requires computing all the fourth moments of the VAR empirical innovations and the sample counterparts of these fourth moments can be poorly estimated and very sensitive to outliers in-sample. In contrast, our identification method only involves estimating the fourth moment of the shock of interest and allows working with robust estimates of that higher-order moment. It performs well also with smaller samples.

To illustrate that point, we simulate 500 different set of data each consisting of 200 observations length on output, inflation and interest rate. For each datasets, we estimate a VAR and computed the monetary policy impulse responses using the spectral or eigenvalue decomposition and the sign and higher-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 includes positive responses of output and prices after a monetary policy shock. This does not occur with our identification scheme, where point estimate of the responses of output and inflation are robustly estimated to be negative.

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

Analytical identified sets We also computed analytically the identified sets obtained when the different identification schemes discussed above are applied to the SW model. Results are very similar to the ones obtained with simulations and are reported in section A.6.2 in the appendix.

[14] Higher 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; figure 14 in the appendix report the magnitude of the refinement of the identified set in the context of the SW model.

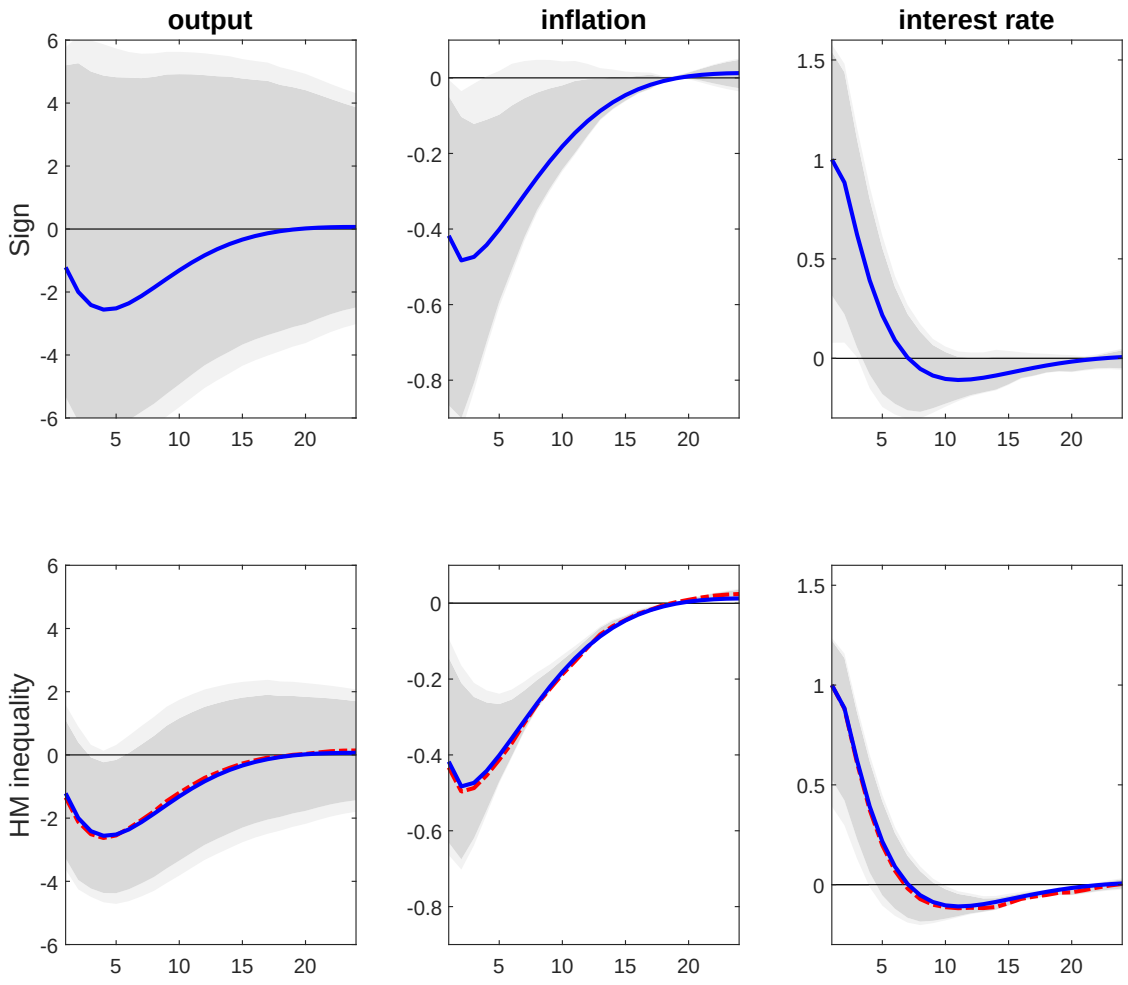


Figure 4: Impulse responses to a monetary policy shock using sign restrictions on prices (−) and interest rate (+) (first row) and sign and higher moment inequality restrictions (second row). The blue solid line is the true impulse response. The red dashed line indicate the point identification obtained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

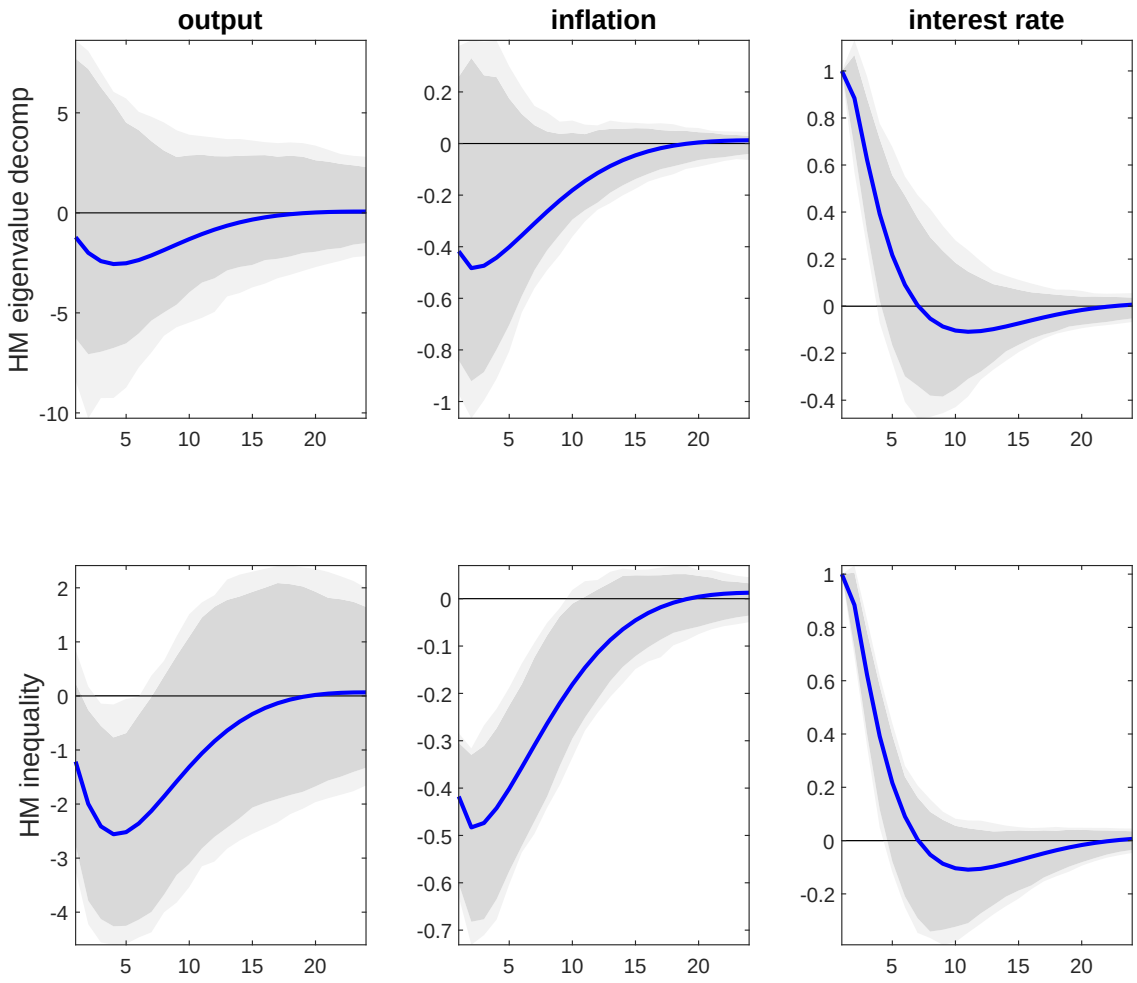


Figure 5: Impulse responses to a monetary policy shock using fourth moments spectral decomposition (top row) and sign and higher-order moment inequality restrictions (bottom row). The dark (light) gray areas report the 90% (99%) dispersion of the point estimates over repeated samples of 200 observation length; the point estimate using the sign and higher-order moment inequality restrictions identification we consider the median impulse response. The blue line is the true impulse response.

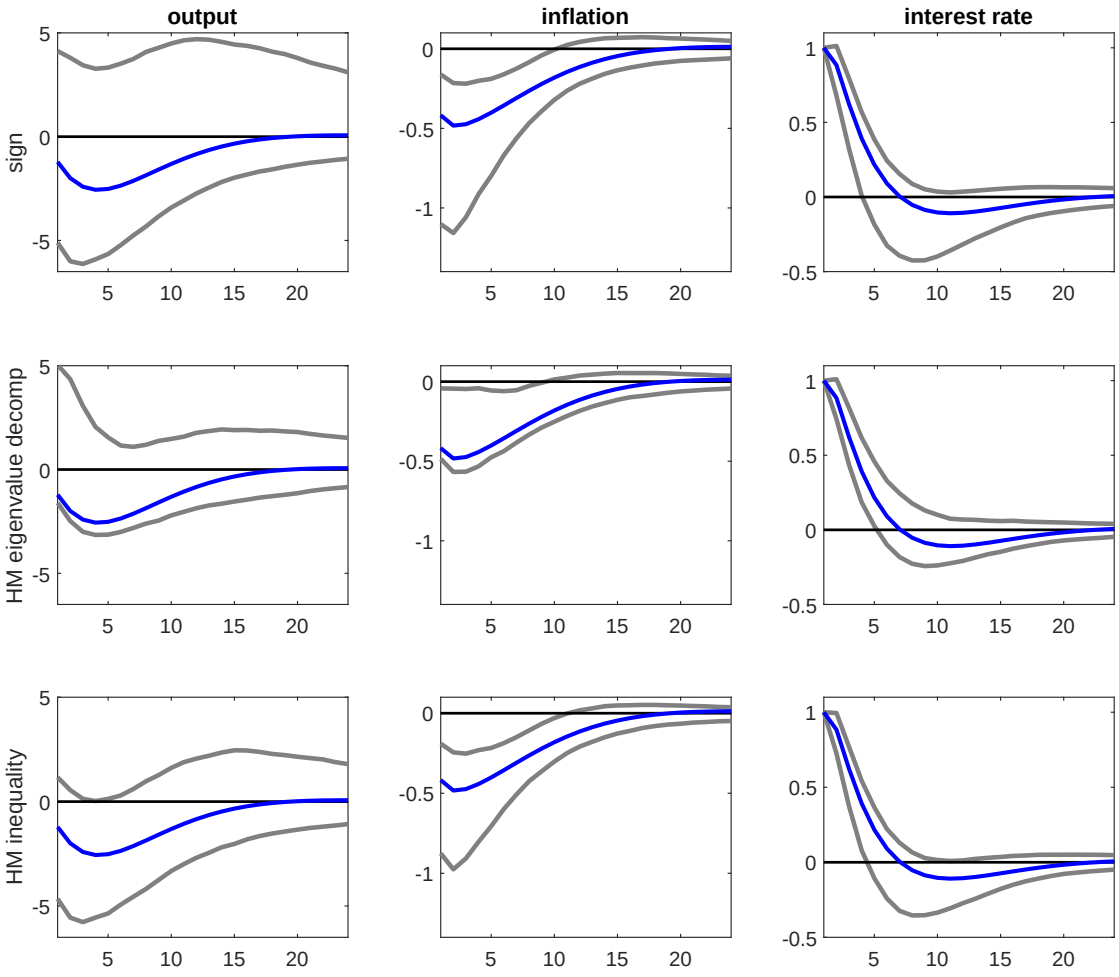


Figure 6: The median upper and lower bounds of the 68% HPD set using sign restrictions, fourth moments spectral decomposition and higher-moment inequality restrictions over repeated samples of 200 observation length. The blue line is the true impulse response.

5 Empirical Application I: US Monetary Policy

We now apply our methodology to observed post-war US data and study the transmission mechanism of conventional monetary policy in the US. We assume that monetary policy shocks are drawn from a leptokurtic distribution, where most realizations are tiny but large deviations are more likely than with normal distributed shock. We start by documenting that indeed excess kurtosis is a robust feature of monetary policy shocks when looking at a large set of proxies or estimates obtained for the US, the euro area, or the UK. We show that such higher-order 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, e.g. Arias et al. (2019), Antolin-Diaz and Rubio-Ramírez (2018), or relying on external instruments, e.g. Gertler and Karadi (2015).

5.1 Higher moments of monetary policy shock proxies

We document the third and fourth 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, nor to US monetary policy. This property is actually observed for a wide range of monetary policy shock proxies obtained with data of different frequency, for different sample-periods, and for different economies.

US data For the US, we consider the recent monetary policy surprises extracted from high frequency datasets 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) or 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 (i.e. USf1-USf3). We also include monetary policy shocks identified with lower frequency data, in particular, the Romer and Romer (2004) (RR)[15] narrative instrument, Sims and Zha (2006) (SZ) monetary policy innovations estimated from a SVAR with regimes shifts, and the monetary policy shock estimated in the Smets and Wouters (2007) (SW) DSGE model.

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

Finally, we also consider the natural language measure constructed by Aruoba and Drechsel (2022) (AD) where they look at 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 exploits high-frequency variations of future OIS contracts around the ECB monetary policy decisions and press conference 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 the Euro Area interest rates futures around the ECB decision and communication constructed in Altavilla, Brugnolini, Gürkaynak, Motto and Ragusa (2019) (EAf1-EAf3).

UK data For the UK, we consider monetary policy surprises obtained using the high-frequency data from Gerko and Rey (2017) who looks at change in three-month sterling future rates observed around the release of the minutes of the MPC (GR(minute)) and inflation report (GR(IR)), from Cesa-Bianchi, Thwaites and Vicondoa (2020) (CBTV) who use changes in three months 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% confidence intervals obtained by bootstrapping the series for all these various measures. While there is no clear evidence about skewness of the monetary policy shocks, all the proxies have a statistically significant excess kurtosis, although point estimates are very dispersed. It is quite striking that all of these proxies are leptokurtic given that they are obtained using different methods, different datasets, different time spans and countries. As a result, they span relatively different information sets. For example, in the US the HF measures of monetary policy surprises and the narrative instrument constructed by RR have a correlation of about 0.2 only (see figure 12 in the appendix). So it is not straightforward to assess which proxy better characterizes monetary policy shocks and should be used to instrument the reduced form VAR residuals. Our proposal to use higher-order moment restrictions to identify monetary policy shocks rather leverage on a robust feature observed for various proxies.

	Ex-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: Proxies of Monetary Policy shocks/surprises: Ex-Kurtosis & Skewness - Bootstrap; median and in parenthesis 95% confidence intervals. Proxies descriptions, sources and tags are in Table 2.

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 1965m1 to 2003m12 sample.[16] We estimate the VAR parameters assuming 12 lags. As the QML or least square (LS) estimates of the autoregressive coefficients are asymptotically consistent, so 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, i.e. the reduced form residuals orthogonalized using the Cholesky factorization of the least square estimates of the residuals covariance matrix. The first row of the figure displays the estimated series and the second row compares the probability plot of the empirical distribution against

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

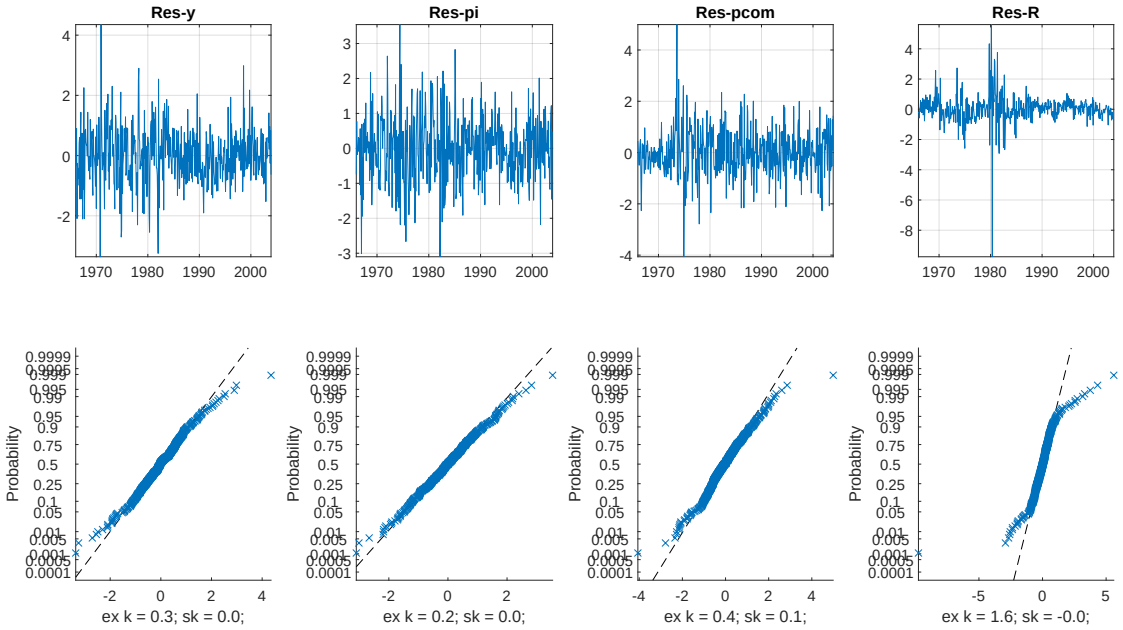


Figure 7: Whitened Least Square residuals.

a standard normal distribution (dash line); departures from the dash line indicate departures from the normality assumptions. Departures from normality are not fully evident for the output and prices whitened residuals; while there are few observations in the tails 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 the commodity prices and of the federal funds rate where the K-S test $p-$values are 0.07 and 0.00 respectively. The FFR residuals seems to be characterized by a large negative value at the beginning of the 80's. 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. As a matter of fact the robust measure of skewness are all very close to zero, whereas the robust measures of kurtosis of the federal funds rate residual is significantly larger than those implied by the Gaussian distribution. Hence, the source of non-gaussianity seems more evident in the tails thickness rather than in the asymmetry of the distribution.

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

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

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

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.[17]

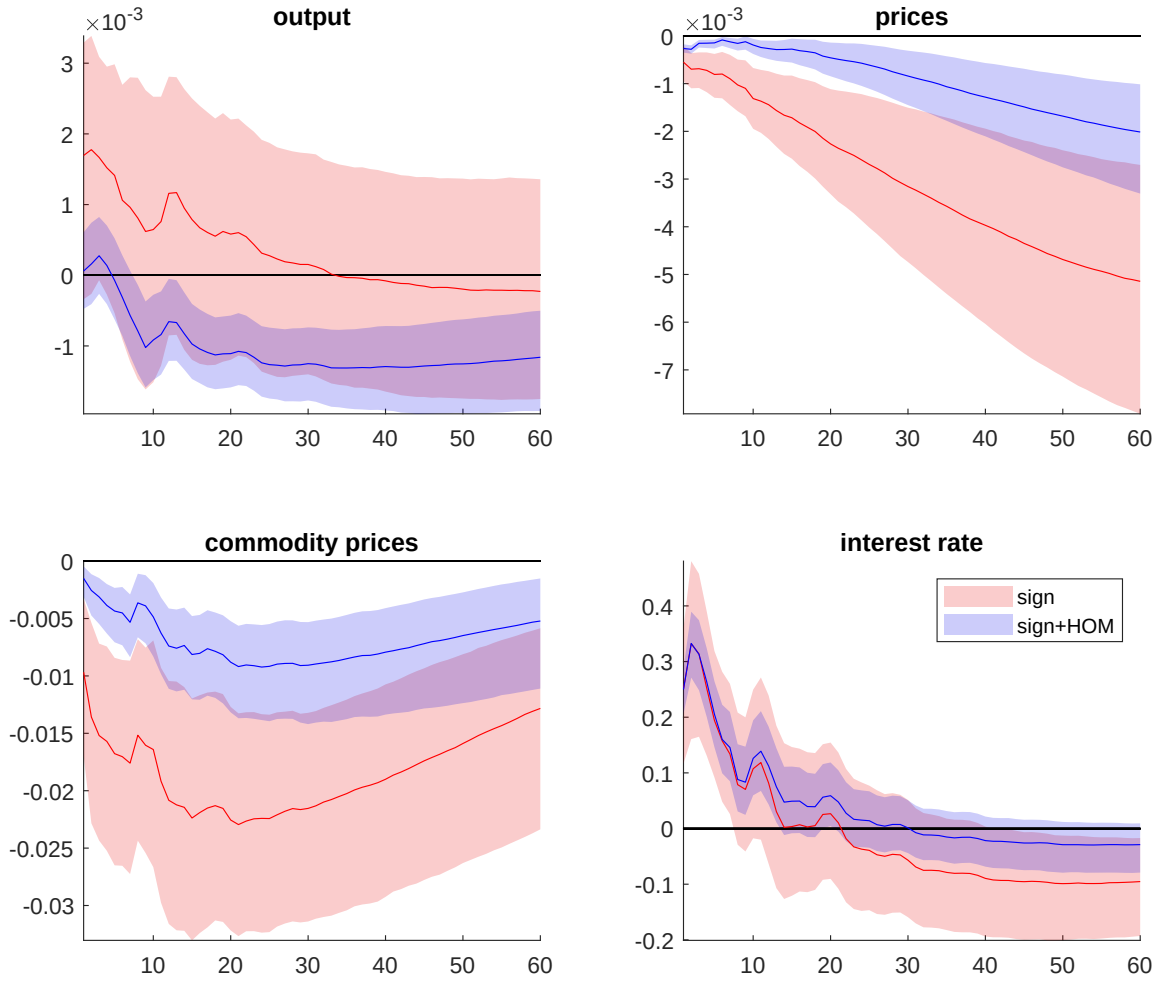


Figure 8: Impulse responses to a monetary policy shock. The colored areas indicate 68% credible sets. Sign restrictions red. Sign and kurtosis ($\mathcal{K}^{\star}_{mp} > 1.2$) restrictions blue.

[17] 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 distant two (three) standard deviations away from the mean is 25% (700 %) more likely that in the standard normal case.

Results Figure 8 reports the impulse response function (IRF) to a monetary policy shock. The median IRF obtained with sign restrictions only is in red. The one obtained with both sign and higher moments restrictions is in blue. Shaded areas depict the associated 68 % confidence sets.[18] We find that monetary policy shocks have no clear effect on output when these are identified only with sign restrictions on interest rate and inflation. When these are complemented with an inequality restriction on the excess kurtosis of monetary policy shocks, almost all the positive responses of output are chopped away and the distribution tilts towards negative outcomes only. Note that the response of the FFR after an initial 25bps 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 – the impact is about half of the size than what is obtained with sign restriction only. So higher-order moments 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 a debate as to whether such a prior is overly informative—and the prior selection does not vanish asymptotically (e.g. Baumeister and Hamilton 2015)—or not (e.g. Arias, Rubio-Ramirez and Waggoner 2020). Giacomini and Kitagawa (2021) propose to compute a 'robust credible region' that relies on a prior class that specifies an 'unrevisable' prior for the structural parameter given the reduced-form estimates. Figure 17 in the appendix illustrates that our findings remain valid when using such an approach.

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

[18] See figure 16 for the 90% confidence bands.

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. As a matter of fact, sovereign spreads dropped very significantly after Mario Draghi made his infamous "Whatever it takes" statement in 2012 which asserted the ECB commitment to intervene on sovereign debt markets to prevent any risk of potential break-up of the euro area. These sharp increases in sovereign risks are believed to have had large negative macroeconomic consequences in particular 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 they coincided with other shocks — the Great Recession, the ensuing deleveraging shocks, and the contractionary monetary policy shocks implied by the zero lower bond constraint — that also had negative macroeconomic consequences. Some scholars address that issue by using bond market reactions observed in a tight window around key political events. The assumption is that these market reactions capture policy choices that will only affect the debt sustainability but not current and future economic conditions. They then use these market reaction as instruments for the reduced form residuals obtained from country level panel VAR 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 to use sign and higher-order moment inequality restrictions in a structural VAR to estimate the impact of sovereign shocks on euro area aggregates.

VAR specification We use monthly data on the EA industrial production (IP, in log), the EA consumer prices measured by core HICP (Core, in log), the unemployment rate (UNR, in percent), a measure of non-financial corporation borrowing costs (NFC spreads, in percent),[19] the one year money market rate (Euribor 1y, in percent), the spread between the 5 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 moments properties, there are salient departures for normality for some variables: The K-S test rejects the null that the borrowing costs, the one year Euribor and the spread could have come from a standard normal distribution; for the remaining variables the test fails to reject the null hypothesis of normality.

[19] We use the non-financial corporation borrowing costs constructed in Gilchrist and Mojon (2017).

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 5 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.*

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 potentially satisfy these sign restrictions, so we combine them with higher order moments inequality restrictions capturing what we assume to be a feature specific of sovereign risk shocks: They increase sharply relatively frequently.

Assumption 4 *The spread shock has moderate asymmetry to the right and moderate fat tails, In particular, the skewness larger than 0.2 and excess kurtosis larger than 0.5, i.e.* $\mathcal{S}_{sp} > 0.2$ *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.[20]

Results Figure 9 shows the results obtained with our identification scheme (in blue) and compare them with what a standard recursive ordering identification — where the sovereign risk shock affects the Italian, German sovereign yields and their spread on impact, but 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 one percent. The identification using signs and higher moments restrictions produces interesting dynamic responses. A sovereign risk shocks leading to a 100 basis points peak increase in the 5-year italian-german sovereign spread generates a 50 basis points increase in the Italian 10-year bond yield as well as a 25 basis points decline in the German 10-year bond yield, consistent with flight to quality effects. Credit conditions tighten with a NFC spread increasing 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 point 18 months after the shock. The price level for core goods and services increase to reach by a tiny (and non statistically significant) 1 basis points peak

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

one year after the shock. The 1-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) obtain on average for Ireland, Italy, Portugal, and Spain over a crisis 2007-13 sample period. Our estimates of the macroeconomic effects are larger though. This is even more the case as our estimates are for euro area aggregates and includes no-crisis time. This suggests that country specific sovereign risk shocks have negative spillovers on other monetary union members even if they are not affected by sovereign risk.

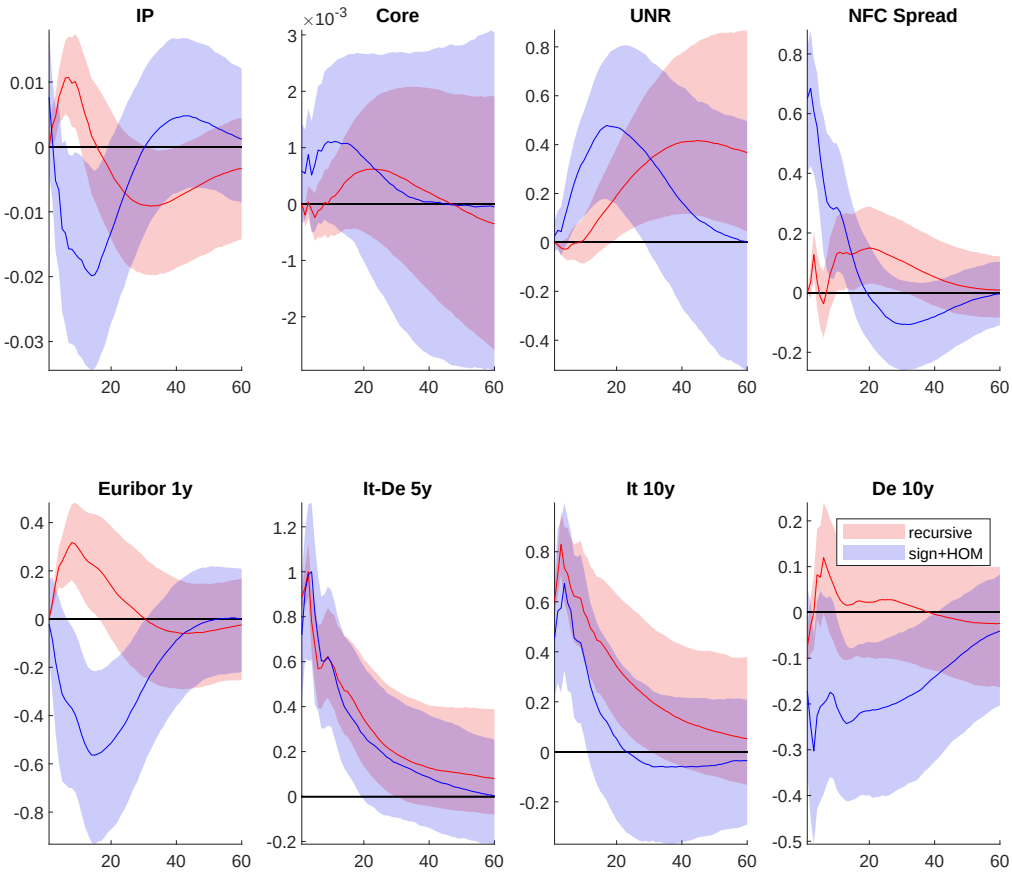


Figure 9: Impulse responses to a spread shock. Recursive ordering in red. Sign and higher-order moment restrictions in blue. Impulse responses are normalized so that the maximum median impact on the spread is 1 percent. The solid line indicates the median response and the colored areas report the 68% confidence bands.

The results obtained with a recursive ordering scheme offer a very different picture, with some puzzling patterns. A shock leading to a one percent increase in the sovereign spread triggers an increase in the 10-year Italian bond yield. It also increases the German ones, so there is no flight to quality effect. Credit costs increase but only modestly and with a delay compared

to 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 1-year money market rate increase, 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 (i.e. assumption 3) are used. The results show that these constraints are not restrictive enough as this scheme does not generate any statistically significant response.[21]

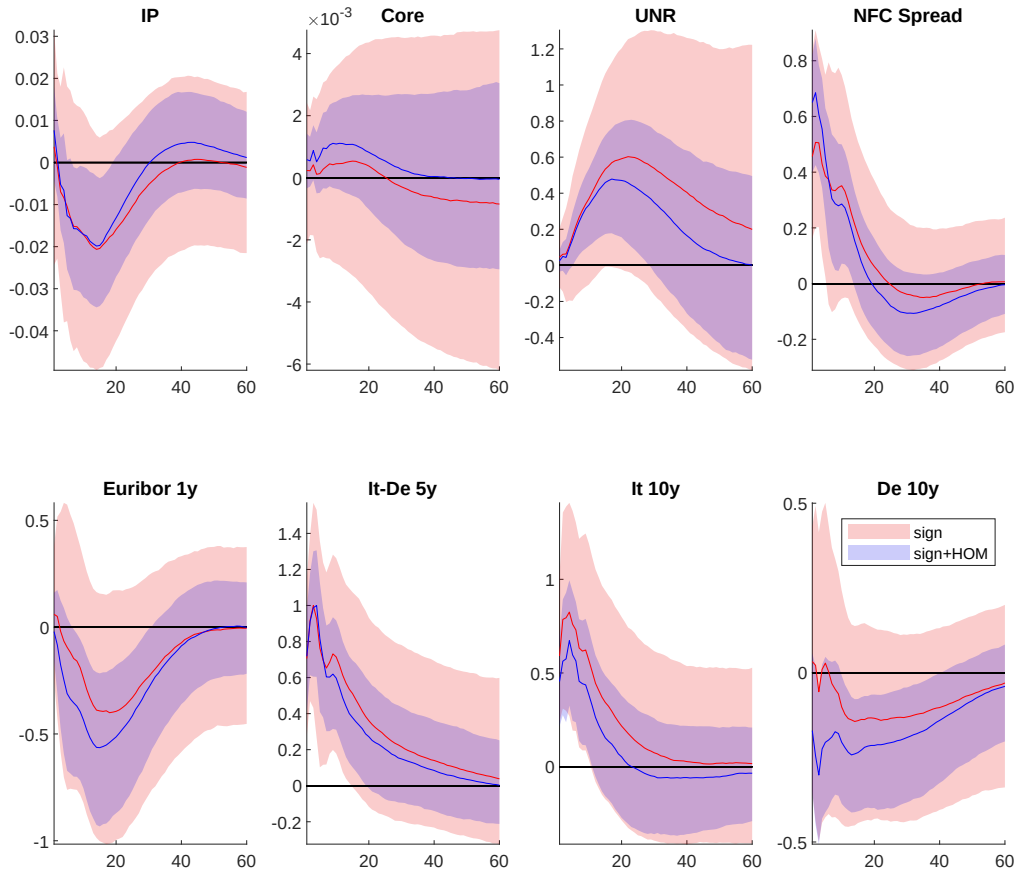


Figure 10: Impulse responses to a spread shock. Sign restrictions in red. Sign and higher-order moment restrictions in blue. Impulse responses are normalized so that the maximum median impact on the spread is 1 percent. Solid line median response and colored areas report the 68% confidence bands.

[21]Figures 9 and 10 report report the 68% confidence bands. Figures 19 and 21 in the appendix report also the 90% 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 risks (GPR) that can be used to address that question. Such an 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. The GPR index captures exogenous events that are not caused, at least within a quarter, by US macroeconomic performances. However, large adverse geopolitical events can also lead to unforeseen policy decisions, from either the fiscal authority, think of military spending shocks, or from the central bank, think of the Fed emergency reaction to 09/11, and some articles discussing geopolitical tensions can also be articles discussing these unforeseen policy decisions. As noticed by Caldara and Iacoviello (2022), the GPR index indeed correlates with measures of policy shocks, for instance the military spending shock of Ramey (2011). Isolating these shocks therefore raises an identification issue. We propose to address it by combining sign and higher order moments restrictions.

VAR specification Caldara and Iacoviello (2022) also provide a structural VAR analysis of the macroeconomic effects of geopolitical risk (GPR) for the US economy. Their VAR consists of eight quarterly variables: (i) the log of the GPR index: (ii) the 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 WTI price of oil; (vii) the yield on two-year US Treasuries; (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 unforeseen 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 captures 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 geopolitical risk ones. This could explain why, under a recursive ordering identification scheme, although the geopolitical tensions last well over one year after the initial GPR shock, the VIX starts to decline, the 2-year bond yield declines, and overall financial conditions, as captured by

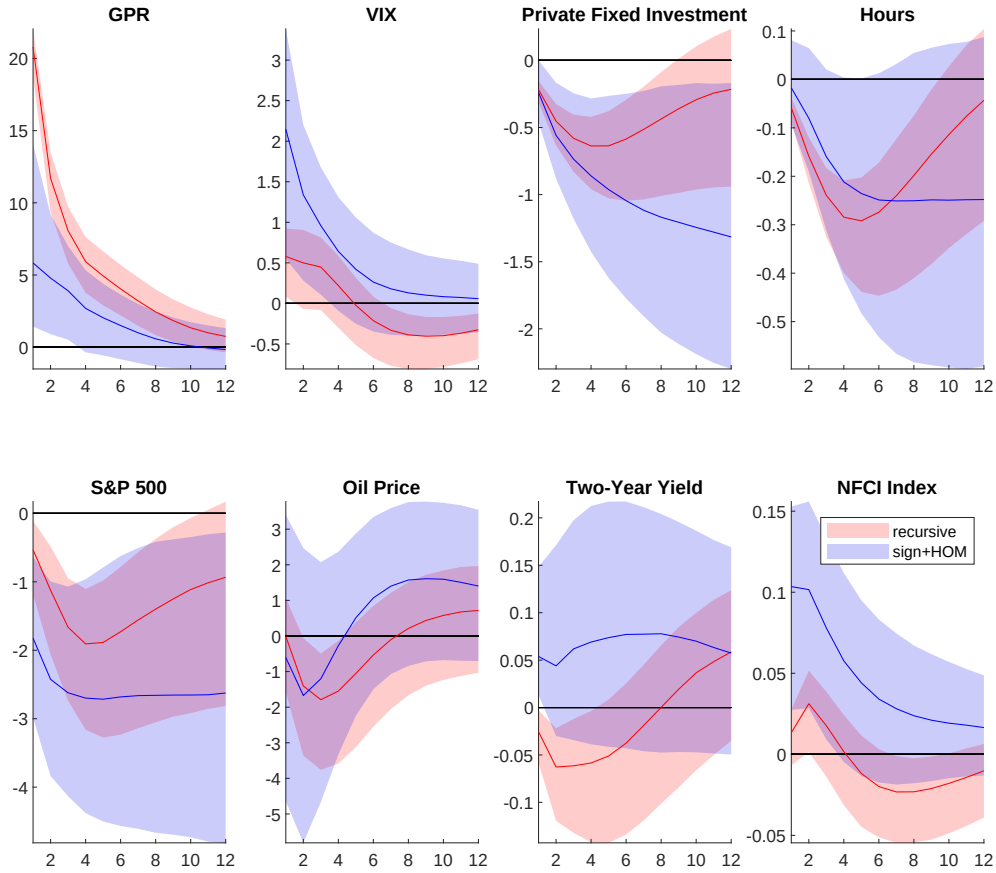


Figure 11: Impulse responses to a GPR shock. Recursive restrictions in red. Sign and higher-order moment restrictions in blue. For each identification scheme we impulse the shocks with a one standard deviations. Solid line median response and colored areas report the 68% and 90% confidence bands.

the NFCI index, 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 tighten financial conditions by lowering the SP500 and increasing the 2-year US treasury yield. However, unreported results show that these sign-restrictions are too loose and lead to a set of admissible rotations that is too large and median IRFs that are not significantly different from zero. We thus investigate an alternative identification scheme which 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 spikes in international tensions which are large and relatively frequent. We assume that geopolitical risk shocks have a distribution that is also skewed to the right and leptokurtic. More specifically, we postulate:

Assumption 5 *An unexpected GPR shock lowers the SP500 and increase the 2-year US treasury yield on impact.*

Assumption 6 *The GPR shock has small asymmetry to the right and fat tails, In particular, the skewness larger than 0.1 and excess kurtosis larger than 1.5, i.e.* $\mathcal{S}_{gpr} > 0.1$ *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 to the recursive ordering as we achieve set-identification rather than point identification. But 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 taken in the wake of some geopolitical events. Interestingly, the reaction of the 2-year 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 geopolitical risk shock would have on that yield. Finally, oil prices do not significantly react to geopolitical shock either. Indeed, depending on the shock, oil prices may in some cases decline due to drop in demand associated with an increased uncertainty, and, in other cases, increase if geopolitical events happen in oil producing countries. Overall, the evidence is in line with models where 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.

Barro, R. J.: 2006, Rare Disasters and Asset Markets in the Twentieth Century, *The Quarterly Journal of Economics* **121**(3), 823–866.

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), 1021â1026.

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.

Gabaix, X.: 2012, Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance, *The Quarterly Journal of Economics* **127**(2), 645–700.

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.

Gourio, F.: 2012, Disaster Risk and Business Cycles, *American Economic Review* **102**(6), 2734–2766.

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).

Kilian, L. and Murphy, D. P.: 2012, Why Agnostic Sign Restrictions Are Not Enough: Understanding The Dynamics Of Oil Market Var Models, *Journal of the European Economic Association* **10**(5), 1166–1188.

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.: 2011, Identifying Government Spending Shocks: It's all in the Timing, *The Quarterly Journal of Economics* **126**(1), 1–50.

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.

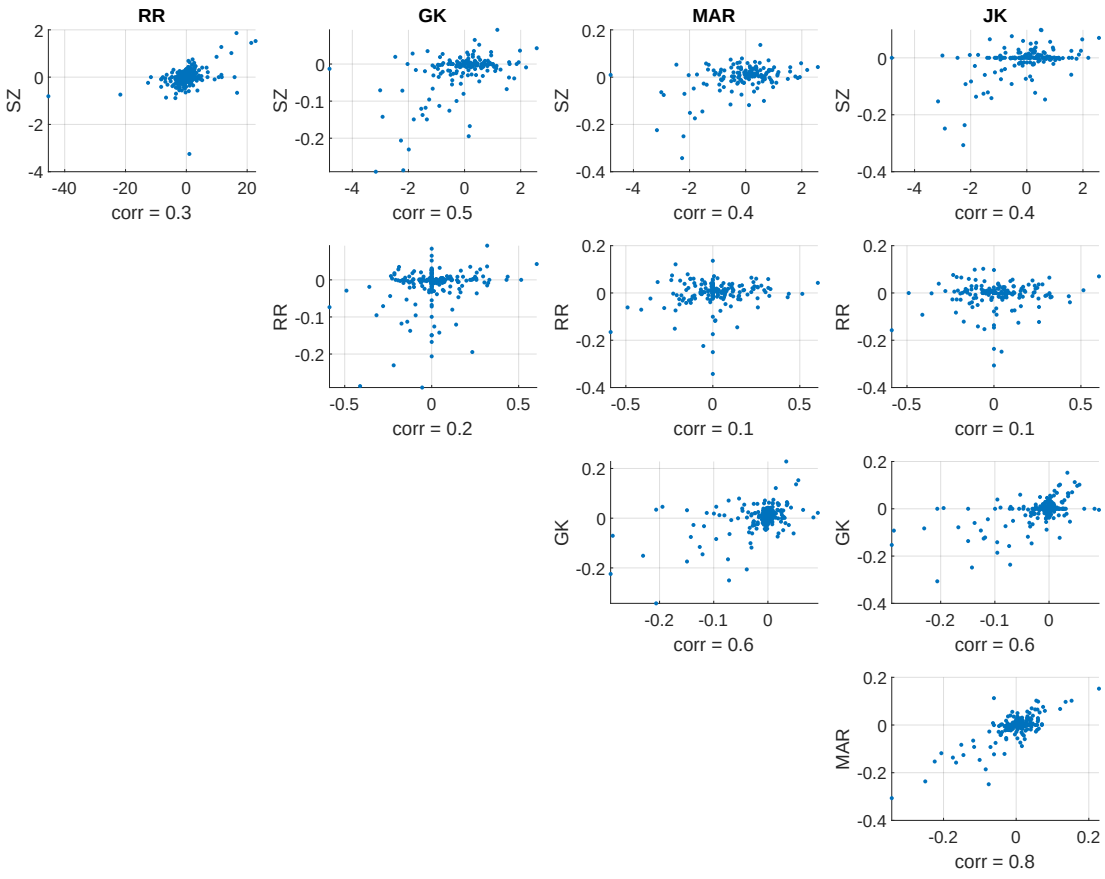


Figure 12: Correlations across measures of U.S. monetary policy shocks.

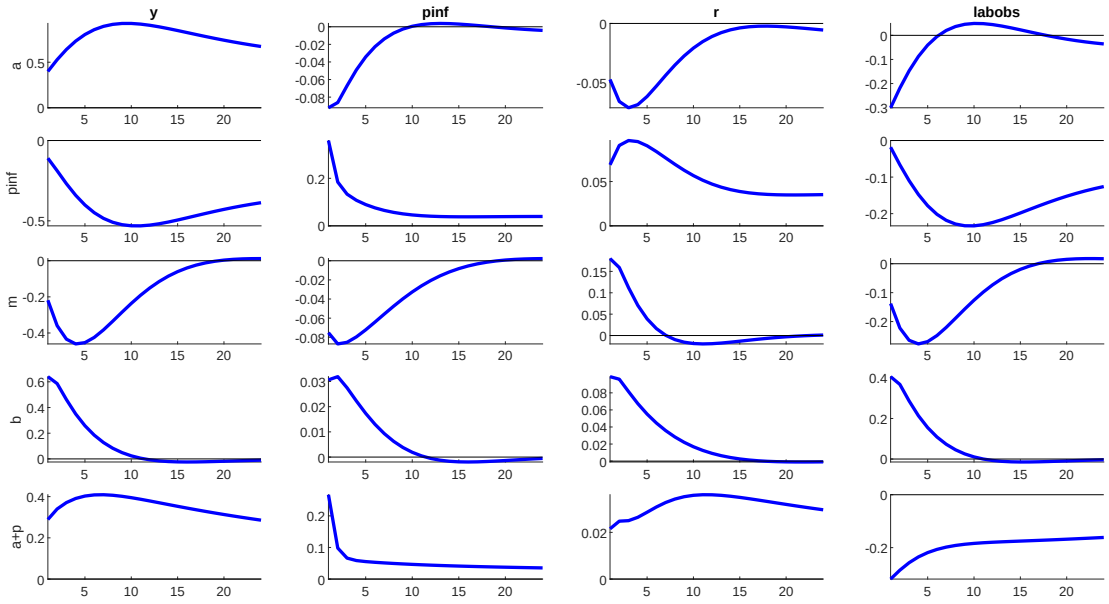


Figure 13: SW estimates of impulse response functions. From top to bottom technology, price markup, monetary policy shocks and risk premium and the sum of technology and price markup shocks.

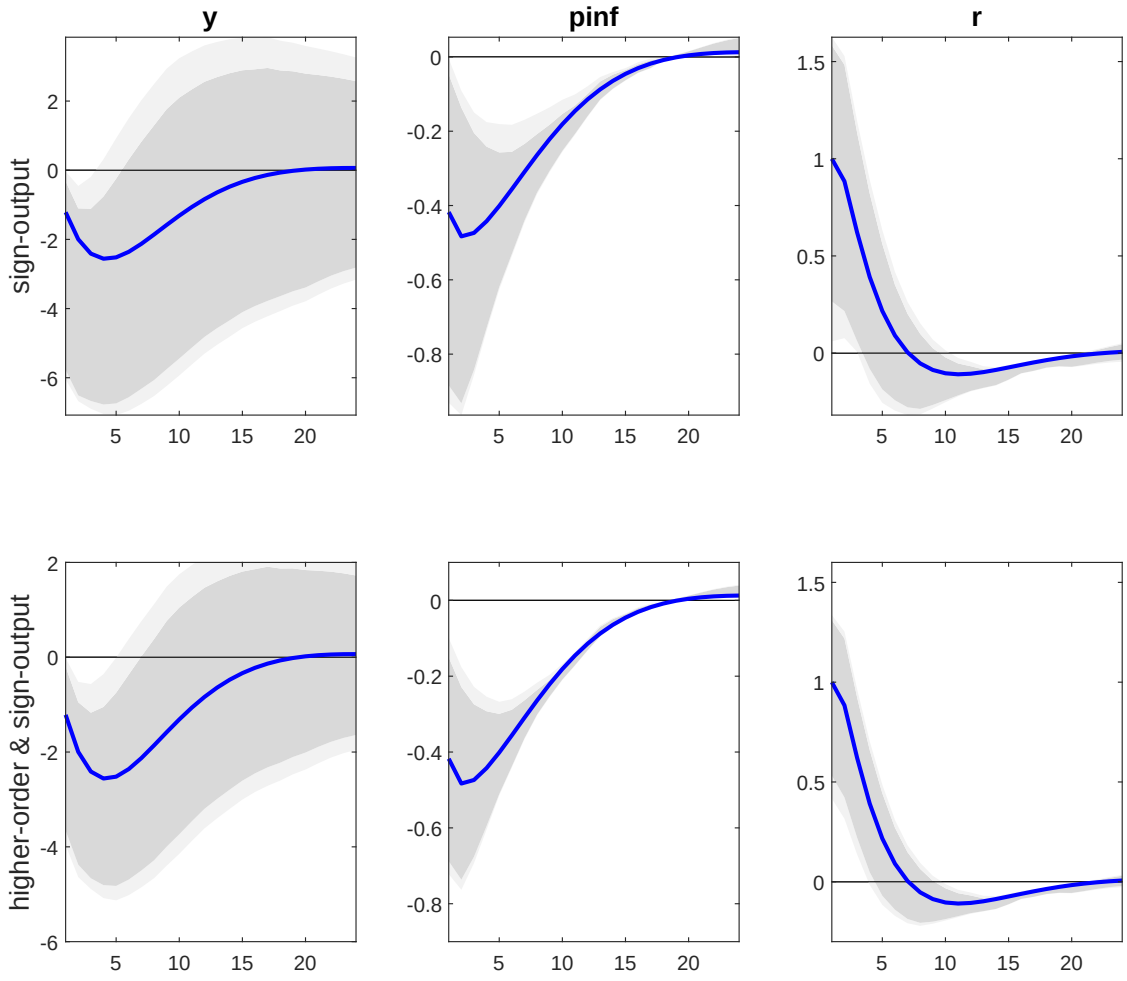


Figure 14: Impulse responses to a monetary policy shock using sign restrictions on output (−), prices (−) and interest rate (+) (first row) and the same sign restrictions combined with higher moment inequality (second row) restrictions. The blue solid line is the true impulse response. The dark (light)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

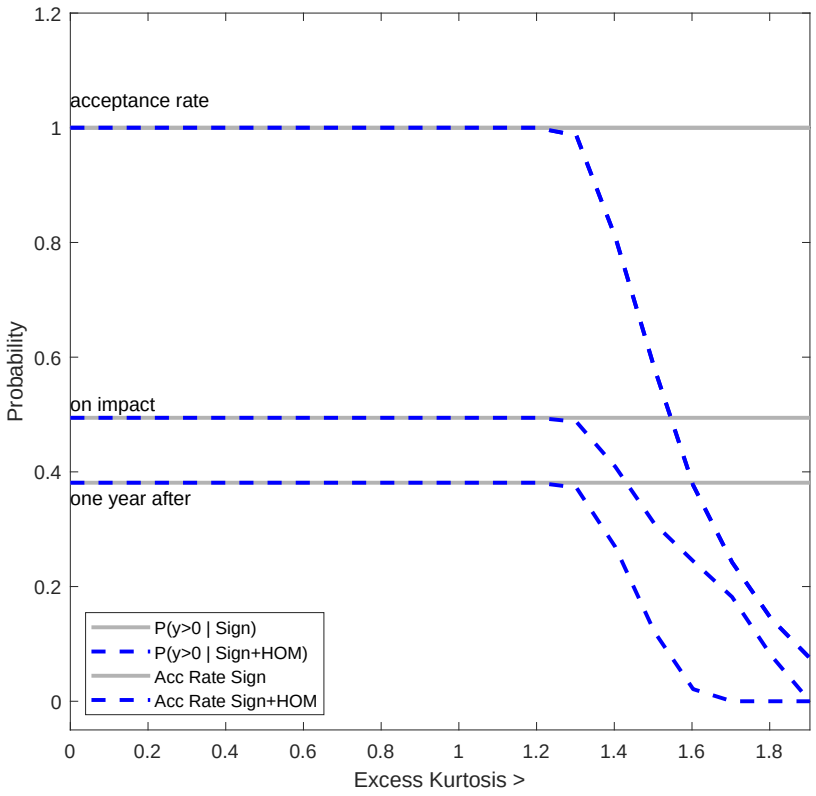


Figure 15: Probability of positive response of output after a monetary policy tightening as a function of the size of the interval.

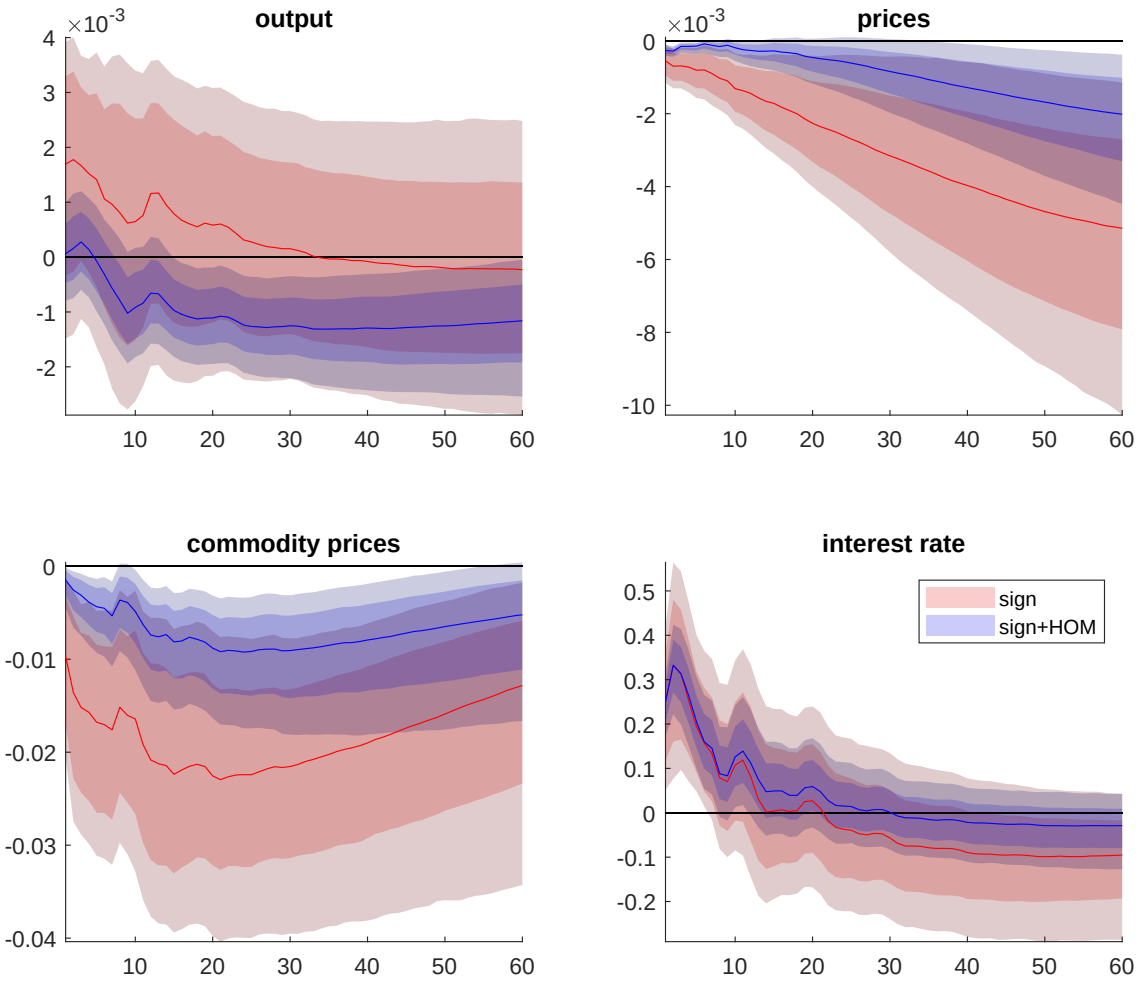


Figure 16: Impulse responses to a monetary policy shock. The colored areas indicate 68% and 90% credible sets. Sign restrictions in red. Sign and kurtosis ($\mathcal{K}_{mp} > 1.2$) restrictions in blue.

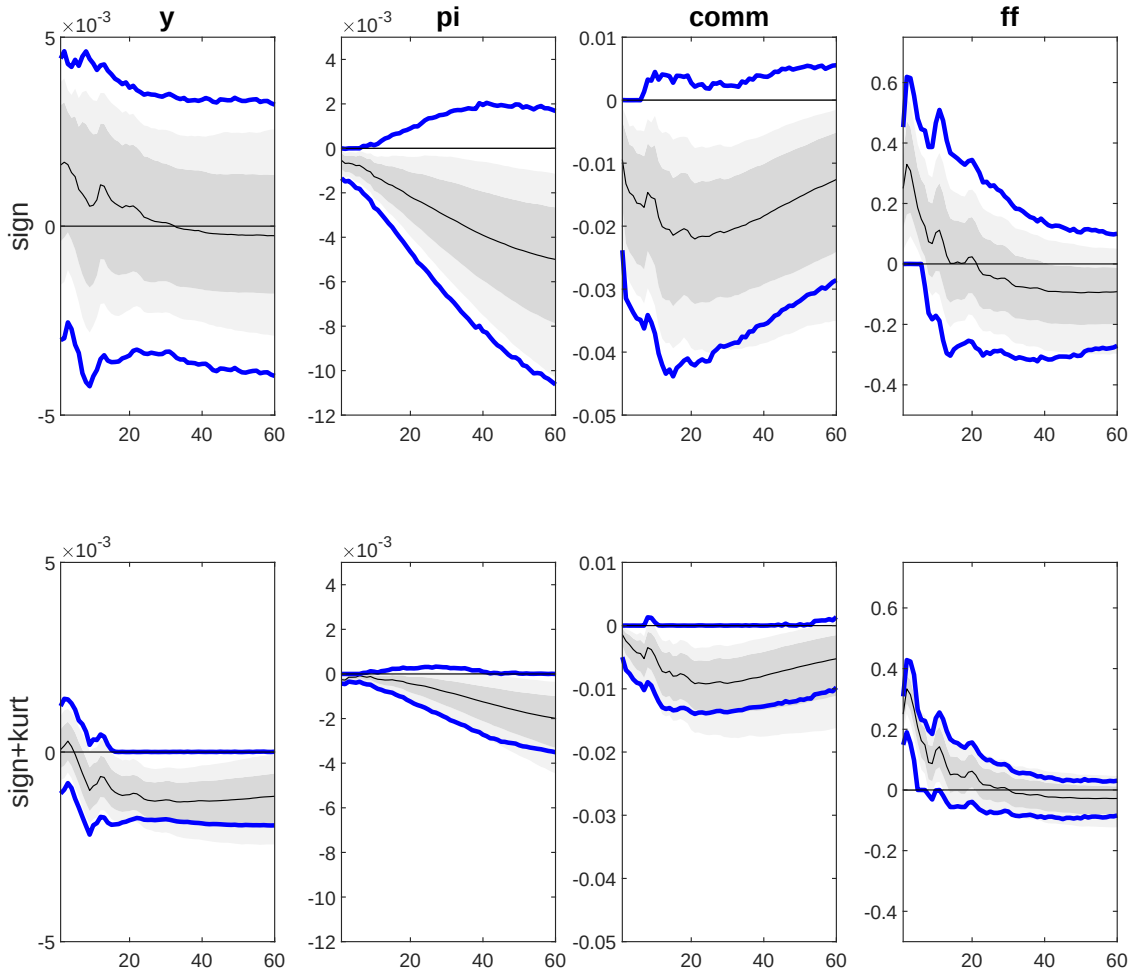


Figure 17: Impulse responses to a monetary policy shock. The colored gray areas indicate 68% and 90% credible sets obtained using draws from the Haar prior as in Rubio-Ramírez et al. (2010). The blue lines report the robust confidence bands at 68% confidence level discussed in **?**. Sign restrictions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

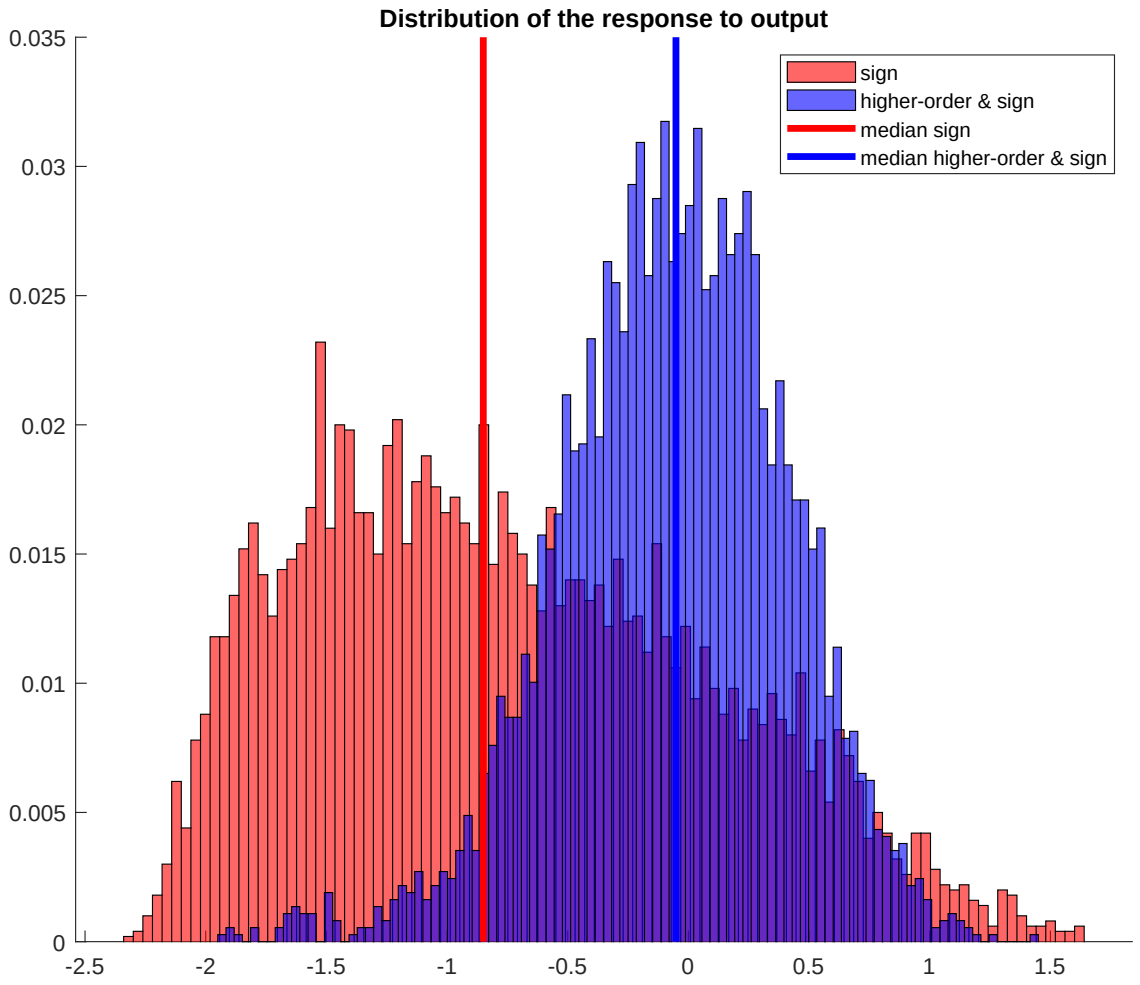


Figure 18: Distribution of the response of the policy rate to output. Sign restrictions in red. Sign and kurtosis ($\mathcal{K}_{mp} > 1.2$) restrictions in blue.

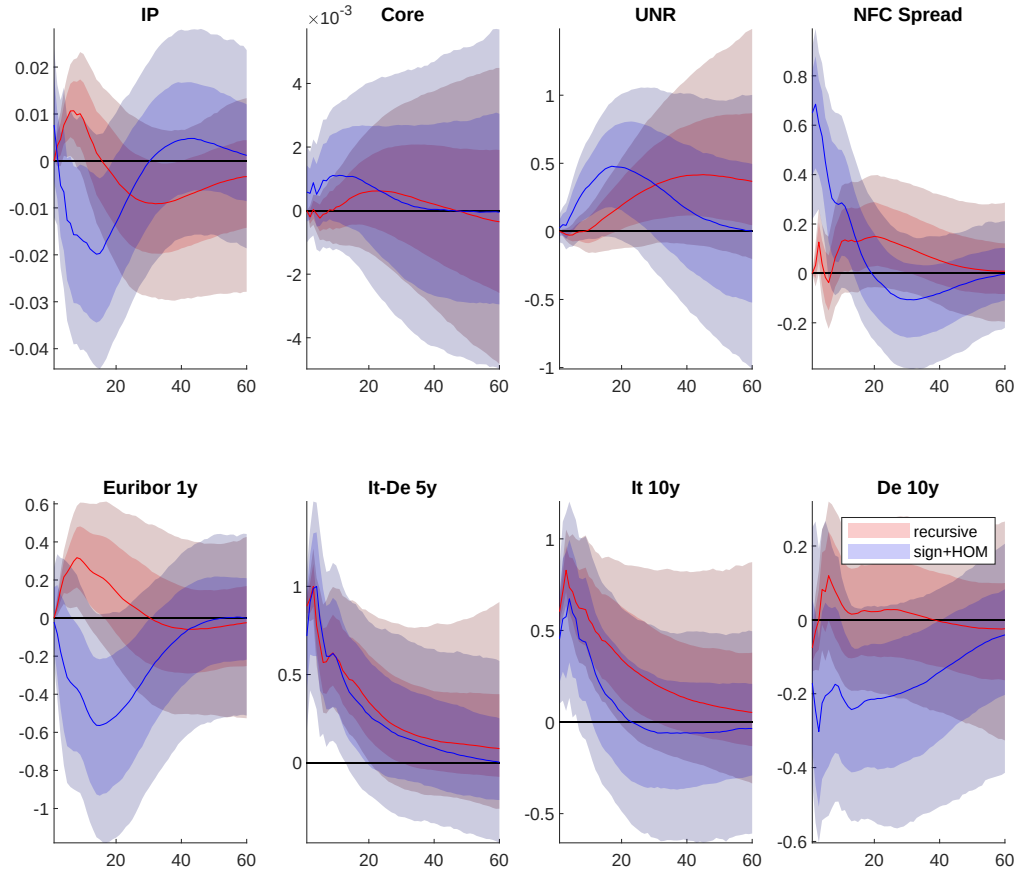


Figure 19: Impulse responses to a spread shock. Recursive ordering in red. Sign and higher-order moment restrictions in blue. Impulse responses are normalized so that the maximum median impact on the spread is 1 percent. The solid line indicates the median response and the colored areas report the 68%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

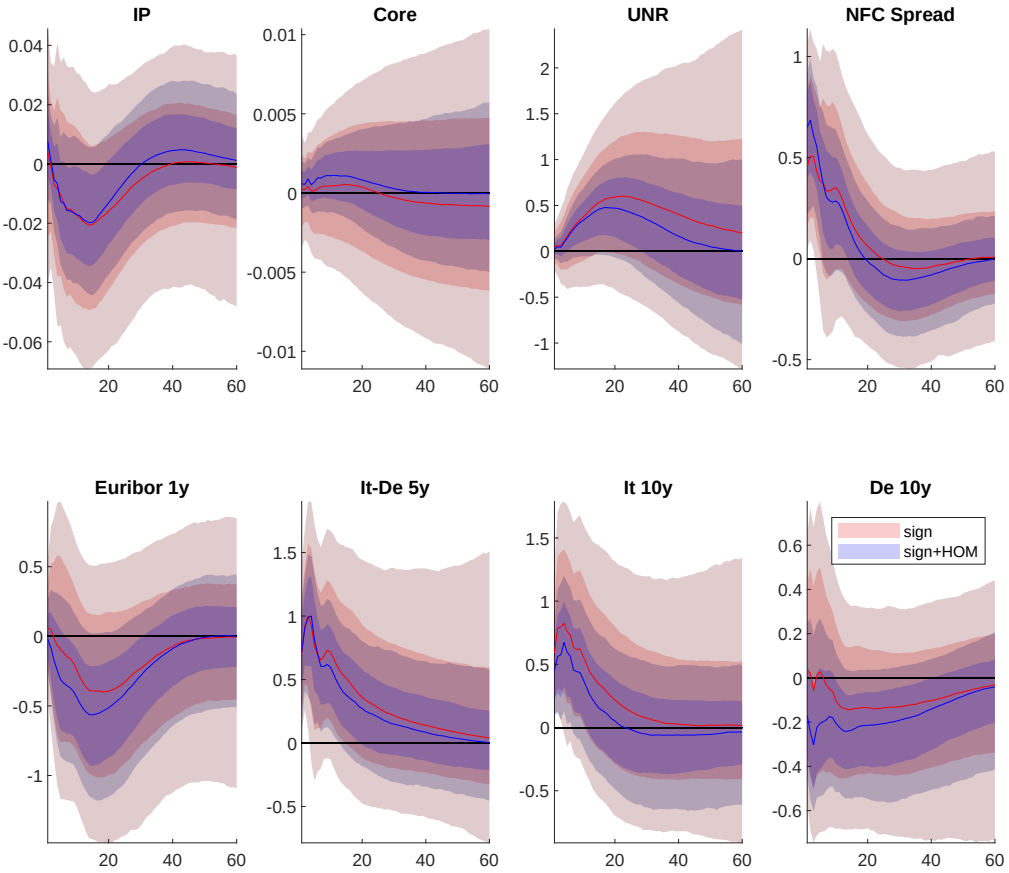


Figure 20: Impulse responses to a spread shock. Sign restrictions in red. Sign and higher-order moment restrictions in blue. Impulse responses are normalized so that the maximum median impact on the spread is 1 percent. Solid line median response and colored areas report the 68% and 90% confidence bands.

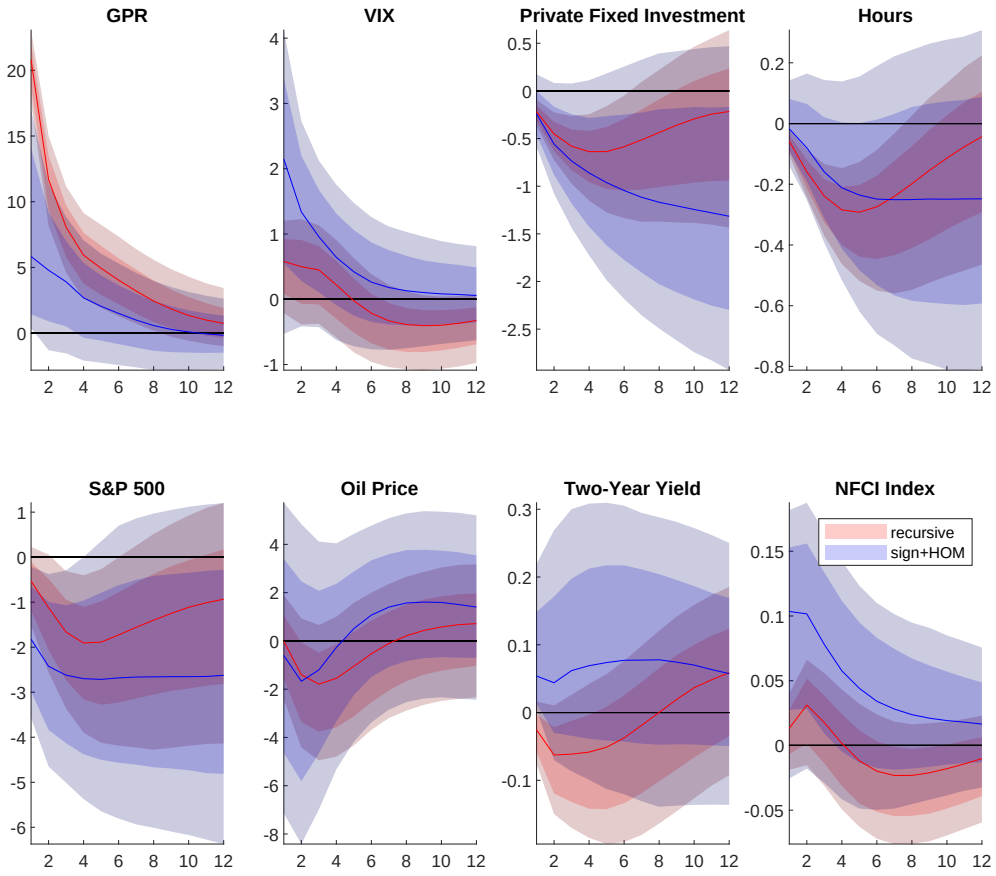


Figure 21: Impulse responses to a GPR shock. Recursive restrictions in red. Sign and higher-order moment restrictions in blue. Impulse responses are normalized to one standard deviation. Solid line median response and colored areas report the 68% and 90% confidence bands.

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 $j = 3$ we have

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

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 vecotrization of X. The following identities hold

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

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 strucutral 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 $\boldsymbol{a}_k$ are the k^{th} column of A_o (the true impact matrix) and A (a candidate rotation) respectively.

A.2.1 Normal distribution fourth moments

In this section we show that the fourth moments of a multivariate normal distribution are invariant to orthonormal rotation matrix. Denote with $\mathcal{K}^z$ the matrix that collects the fourth 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[22] that $K_{n,n}(A\otimes B) = (B\otimes A)K_{n,n}$, we have that the commutation matrix is invariant to any orthonormal rotation Ω, i.e. $(\Omega\otimes\Omega)'K_{n,n}(\Omega\otimes\Omega) = K_{n,n}(\Omega\otimes\Omega)'(\Omega\otimes\Omega) = K_{n,n}(\Omega'\otimes\Omega')(\Omega\otimes\Omega) = K_{n,n}(\Omega'\Omega\otimes\Omega'\Omega) = K_{n,n}$. Moreover, using the relationship between the vectorization and Kronecker product, i.e. $\text{vec}(ABC) = (C'\otimes A)\text{vec}(B)$, we have 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, we have that

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

Denote with $\mathcal{K}^n$ the matrix that collects the fourth 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

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

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

A.2.2 Third moments

Assume that $E(\nu_{n,t}^3) = \zeta_n \neq 0$. The $(n \times n^2)$ matrix collecting the structural shocks third 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, i.e. $(A \otimes B)(C \otimes D) = AC \otimes BD$, 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 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 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 we have 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 can focus on the $(n, n)-$element of $\Phi^{(m)}$, i.e.

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

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 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 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}^2E\nu_{n,t}^2 \end{pmatrix} \otimes \boldsymbol{e}_1' \otimes \boldsymbol{e}_1 + ... + \begin{pmatrix} E\nu_{1,t}^2E\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 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 moments coincide with the standard

normal distribution ones. We can rewrite the fourth 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 matrix and $\mathcal{K}^z$ denotes the matrix that collects the fourth moments of the standard normal distribution. Therefore, we obtain that the the excess fourth 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 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 get

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

In this case as well, with no departure from the normality (i.e. $x_i = 0$ 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 we have that $\mathcal{K}^\star_{(n^2,n^2)} + \mathcal{K}^z_{(n^2,n^2)} = x_n + 3 = \xi_n$, 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 moments

The eigenvector decomposition of fourth moments is based on Kollo (2008) and it requires more notation. The approach computes first the sum of the n^2 blocks of $(n \times n)$ sub-matrices of the fourth moment matrix and then take the eigenvalue/vector decomposition of the resulting matrix. The eigenvectors associated to non-zero 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 of the 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 elementwise (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 moments of the structural shocks, i.e. $E(\nu_{i,t}^4) - 3$; we have that

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

We then have

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

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 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-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 fourth and third sample moments can be very sensitive to outliers with short samples, see Kim and White (2004). When shocks are distributed with t-student with low degrees of freedom, not even a million draws are enough to have median unbiased estimates of kurtosis, see Ferroni and Tracy (2022). A number of robust measures have been proposed in the literature and they all perform well for the sample size typically considered 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 is the distribution around the median. Hence, with distributions centered around zero, realizations are often very small but sometime quite big. Robust measures of skewness exploit the distance between median and mean. More formally, let 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, i.e. $p_j = F^{-1}(j/8)$ 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 the estimation and identification strategy that we use to estimate a VAR with non-gaussian errors and to construct IRF 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 i.i.d. zero mean random vectors with unconditional covariance matrix Σ. We assume that the VAR reduced form shocks are linear combination of the unobserved structural shocks, ν_t, i.e.

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

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

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 to adopt a robust Bayesian approach which allows to construct posterior credible sets without the need for distributional assumptions of the reduced form residuals. The Bayesian approach we use builds on the work by Petrova (2022), where she propose 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, inference 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[23] 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 in the consistent estimator of Σ, i.e. the QML estimator $\hat{\Sigma}$, and with 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 moment of the VAR reduced form shocks. Importantly, as previously mentioned, sample counterparts of the fourth moments can be poorly estimated and be sensitive to outliers. We then follow Petrova (2022) and consider a shrinkage approach which consists in tilting the sample fourth moments estimates of the reduced form orthogonalized errors towards 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 moments of the orthogonalized reduced form residuals, i.e. $\widehat{\mathcal{K}}_T = 1/T\sum \text{vech}(\iota_t\iota_t')\otimes\text{vech}(\iota_t\iota_t')$ 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, elsewhere there are all zeros; and D_n^{+} is the generalized inverse of the duplication matrix D_n.[24] The first bit of the equation (2) represents the sample fourth moments and the second bit the fourth moments implied by a standard normal distribution; τ is the amount of shrinkage that 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 at this point that the sample fourth 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. Fourth sample 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 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 consistently estimating third moments. As for the fourth moments, we consider the consistent shrinkage estimator given by $\widehat{\mathcal{S}}_T^{\star} = \frac{T}{T+\tau}\widehat{\mathcal{S}}_T$

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

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

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 (SUR) format. Let $k = np + 1$, we have

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

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[25], the estimation identification procedure can be then summarized as follows. Let $\widehat{S} = (Y - X\widehat{\Phi})'(Y - X\widehat{\Phi})$ 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 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 check if the higher moment inequality restrictions are satisfied.

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

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

After a suitable number of iterations, the draws are representative of 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 is performed using the toolbox described in Ferroni and Canova (2021).[26]

A.5.1 Informative priors

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

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

A.6 Identified set

In this section, we discuss how we construct the identified set with sign restrictions and with sign and higher-order moment restrictions when the DGP is the three equation static NK model and for the SW model respectively (without simulating the data from the DGP). In what follows the identified set is defined as the superior and inferior bounds of the impulses responses over the set of possible accepted rotations. For a formal definition, see Wolf (2020) Appendix A.1 Definition 2.

A.6.1 NK model

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

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

$$= \Sigma^{1/2}\, A_o\, \epsilon_t.$$

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

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

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

Notice that the true rotation matrix is difficult to characterize analytically because of the Choleski decomposition which is 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 $\theta_j \in [0, \pi]$ 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 (i.e. 63^3) different candidate rotations. For notation purposes I will drop the argument of A.

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

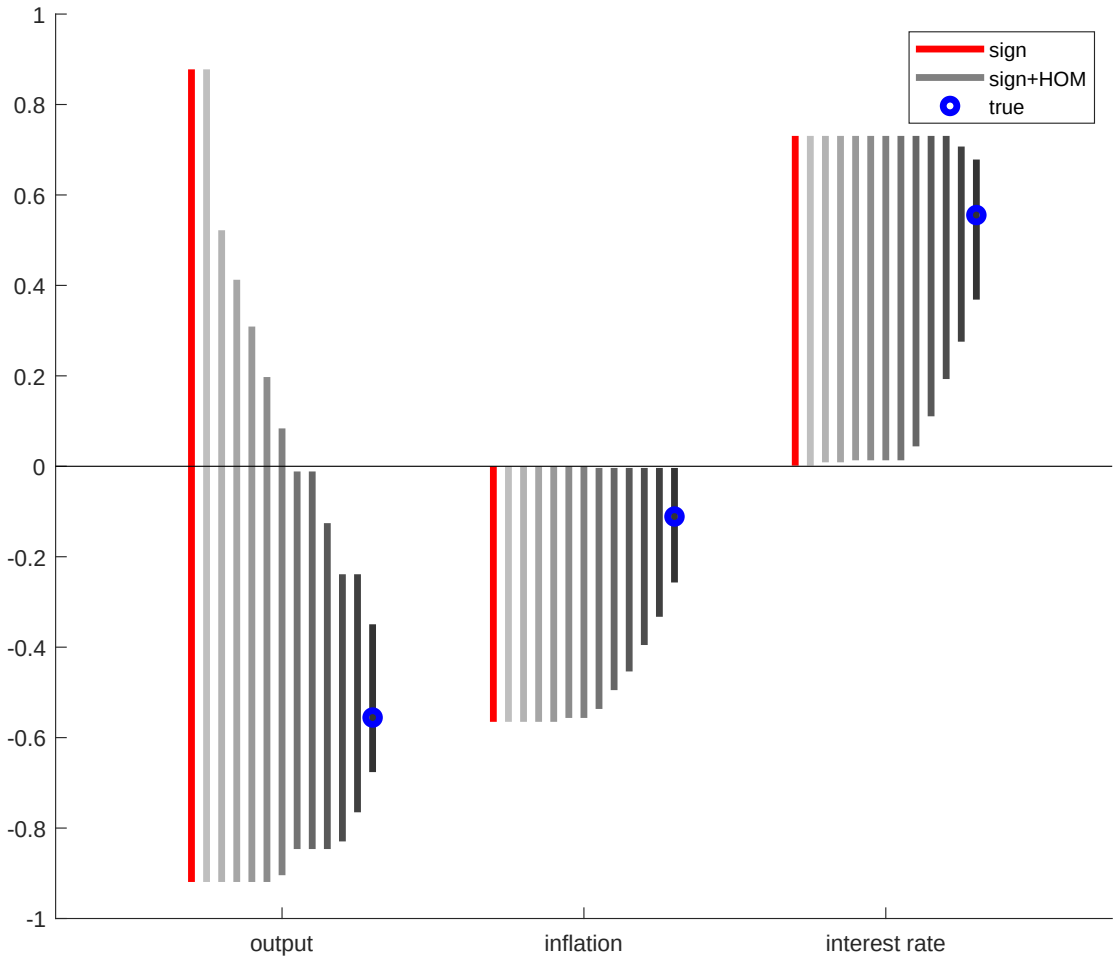
$$\mathcal{K}^\star = (A' \otimes A')(A_o \otimes A_o)\; \mathcal{K}^\star_\epsilon \; (A'_o \otimes A'_o)(A \otimes A) \tag{3}$$

where $\mathcal{K}^\star_\epsilon$ is the excess kurtosis matrix of the structural shocks (i.e. a matrix of zero except a 3 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 in case of the sign restriction identification only with the red lines and the sign and higher-order moment restrictions with 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: NK model identified set: sign restriction identification red lines; sign and higher-order moment inequality restriction identification in gray lines with darker shades of gray indicating a larger threshold. True impact in blue.



A.6.2 SW model

In this section we consider the SW model equations at the posterior mode parameterization. As in the NK example, we assume that the excess kurtosis of monetary policy shock is 3 and the remaining shocks have zero excess kurtosis. Once the model is solved it can be cast in the $ABCD$ representation, i.e.

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

where in this case as well the shocks are renormalized to have a unit variance, i.e. $\tilde{B} = B\Sigma^{1/2}$ and $\tilde{D} = D\Sigma^{1/2}$. The SW model admits the following $VMA(\infty)$ representation, see e.g. 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 we have, $\Sigma_w = \tilde{D}\tilde{D}'$. 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 characterize 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, i.e.

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

For each of the candidate rotations constructed as in the previous sections, A, we verify that

- after a monetary policy shock we have $F(3,3) > 0$, $F(6,3) > 0$ ($i_t > 0$, $i_{t+1} > 0$) 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 (i.e. a matrix of zero except a 3 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 in case of the sign restriction identification with red areas and the sign and higher-order moment restrictions with gray areas 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 is represented by the dashed black lines.

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