Journal of Monetary Economics 135 (2023) 1–19

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jmoneco

Identification with External Instruments in Structural VARs

Silvia Miranda-Agrippino[a], Giovanni Ricco[b,*]

[a] Bank of England, CfM(LSE) and CEPR. Monetary Analysis, Bank of England, Threadneedle Street, London EC2R 8AH, UK
[b] University of Warwick, OFCE-SciencesPo and CEPR. Department of Economics, University of Warwick, Social Sciences Building, Coventry CV4 7AL, UK

ARTICLE INFO

Article history:
Received 6 September 2021
Revised 11 January 2023
Accepted 11 January 2023
Available online 20 January 2023

Keywords:
Identification with external instruments
Structural VAR
Invertibility
Monetary policy shocks
C36
C32
E30
E52

ABSTRACT

IV methods have become the leading approach to identify the effects of macroeconomic shocks. Conditions for identification generally involve all the shocks in the VAR even when only a subset of them is of interest. This paper provides more general conditions that only involve the shocks of interest and the properties of the instrument of choice. We introduce a heuristic and a formal test to guide the specification of the empirical models, and provide formulas for the bias when the conditions are violated. We apply our results to the study of the transmission of conventional and unconventional monetary policy shocks.

1. Introduction

The study of the dynamic causal effects of structural shocks on economic variables is a central endeavour in empirical macroeconomics. In an important departure from classical statistical identifications based on theory-motivated restrictions, the recent practice has increasingly relied on instrumental variables (IV) for the identification of structural shocks. The instruments, interpreted as noisy measures of the shock of interest, are used either with Structural VARs – as external (SVAR-IV/Proxy-SVARs) or internal instruments (Hybrid VARs) –, or with Local Projections (LP-IV Jordà, 2005) with or without controls. IV-based identification has rapidly become dominant in empirical macro, and has led to important advancements in the research on the transmission of structural shocks.[1]

In stating the conditions for a correct IV-based identification, the SVAR literature has routinely appealed to the notion of full invertibility. This is the requirement that all the structural shocks in the economy can be obtained from linear combinations of the VAR residuals (see Mertens and Ravn, 2013 and Stock and Watson, 2018). Full invertibility is a very strong assumption and likely to be seldom attained in practice (see Canova and Ferroni, 2019). This paper argues that, in the context of SVAR-IV, it is not necessary, and that only invertibility of the shock of interest – or partial invertibility (Sims and Zha, 2006) – is required. To fix ideas, consider the case in which monetary policy is set using a Taylor rule, as a function of past inflation and output. The monetary policy shock can be retrieved from the residuals of the policy rate equation in

* Corresponding author.
E-mail addresses: silvia.miranda-agrippino@bankofengland.co.uk (S. Miranda-Agrippino), G.Ricco@warwick.ac.uk (G. Ricco).

[1] This rapidly expanding research programme has produced a wide array of instruments for the identification of conventional and unconventional monetary policy shocks, tax and government spending shocks, oil and news shocks. See e.g. the survey in Ramey (2016).

https://doi.org/10.1016/j.jmoneco.2023.01.006
0304-3932/

a small VAR that also includes an index of real activity and of price inflation. However, other shocks such as, for example, financial shocks or technology news shocks would not invertible in this system.

The contribution of this paper is to formalise the conditions that the IV must satisfy to achieve correct identification in SVAR-IVs, given partial invertibility. In particular, we show that other than the standard exclusion restrictions, the IV must satisfy a 'limited lead-lag exogeneity condition' that ensures that the VAR innovations and the IV correlate only via the contemporaneous structural shock of interest. Importantly, this condition allows the instrument to correlate with leads and lags of other partially invertible shocks without compromising the correct identification.

We move in three steps. First, we show that under partial invertibility a covariance stationary stochastic vector process admits a 'semi-structural' representation that is the sum of two orthogonal terms. The first one is a moving average of the partially invertible shocks, with structural coefficients. The second is a reduced-form moving average of convolutions of leads and lags of the remaining non-invertible shocks. This result implies that if a VAR correctly captures the autocorrelation structure of the Wold representation (i.e. of the data generating process, or DGP), the IRFs obtained from the partially identified structural moving average are the dynamic causal effects of the shock of interest.

Second, we show that the existence of the semi-structural representation allows for identification with IVs that correlate with leads or lags (but not contemporaneous realisations) of any other invertible shock in the system. We label this as the limited lead-lag exogeneity condition, and derive an explicit formula for the bias in the IRFs that arises when this condition is violated. Extending results in Stock and Watson (2018), we show that the limited lead-lag exogeneity condition is also required in LP-IV with controls.

Third, we analyse the likely cases in which the empirical VAR does not capture the DGP, e.g. due to insufficient lag order, or omitted variables. In this case IRFs at horizon larger than zero are generally biased. However, if partial invertibility of the shock of interest holds in this misspecified VAR, the impact responses are correctly identified provided that the limited lead-lag exogeneity of the instrument holds. We show how to use this result to discriminate between cases of VAR misspecification and contamination of the IV. We formalise this intuition with a Hausman type test of lagged conditional exogeneity, along the lines of Lu and White (2014).

Our conditions for identification focus on the properties that the IV must satisfy, conditional on a given VAR. Conversely, the test provides a useful way to think of the issue of selecting a specific model, conditional on the available IV. In other words, the question of finding a 'core information set' that makes the IV conditionally exogenous. This is the point of view most relevant for empirical researchers. From this perspective, our results provide a theoretical justification to the empirical wisdom of searching for specifications that include 'sufficiently' rich sets of lagged controls, and of proving the robustness of results against 'reasonable' alternative control sets.

We illustrate our results using data simulated from a New-Keynesian DSGE with four shocks, only two of which are invertible from a VAR in the observables, and IVs contaminated to various degrees.

In the empirical application we use our results to review the leading IVs for the identification of monetary policy shocks, for which full exogeneity is typically assumed. The narrative and high-frequency (HF) IVs of Romer and Romer (2004) and Gertler and Karadi (2015) for conventional monetary policy, and those proposed by Swanson (2020) for unconventional policy. The lack of robustness of these instruments has been pointed out in the literature, and the sensitivity of results obtained with these instruments has often been attributed to model misspecification. Instead, through the lens of our results, we show that Romer and Romer (2004)'s narrative IV is likely to fail both the contemporaneous and the lag exogeneity conditions. Hence, whether used as external or internal instrument, it fails to correctly recover the dynamic responses to monetary policy shocks in standard VARs without additional assumptions. Using three alternative HF IVs, all a variant of those introduced in Gertler and Karadi (2015), we show that only those that control for information effects pass the test for lagged conditional exogeneity even in standard small VARs. And recover IRFs that are not sensitive to the specification, even in VARs that are not fully invertible. Hence, the puzzles and instabilities reported in the literature are due to contamination of the HF IVs by shocks that are non-invertible in small VARs, rather than to model misspecification. Finally, Swanson (2020)'s IVs for forward guidance and quantitative easing shocks fail both exogeneity conditions. Absent additional controls for other confounding factors, these IVs are unlikely to recover dynamic responses to monetary policy actions irrespective of the chosen empirical system.

The paper is organised as follows. The reminder of this section summarises the related literature. Section 2 discusses the existence of the semi-structural representation that allows to specify the conditions for identification in Section 3. Section 4 compares the conditions required in SVAR-IV and LP-IV with controls. In Section 5 we discuss misspecification, while the test of conditional lagged exogeneity is in Section 6. The DSGE-based simulations are in Section 7. Section 8 provides a step-by-step 'cookbook' on how to use IVs in SVARs, that is then applied in Section 9 to assess IVs for monetary policy shocks. Section 10 concludes. Technical proofs and additional results are in the Online Supplementary Materials.

Related Literature Proxy SVARs/SVAR-IV techniques were first introduced by Stock (2008), and then explored in Stock and Watson (2012) and Mertens and Ravn (2013). LP-IV with or without controls have been independently proposed in Jordà et al. (2015) and Ramey and Zubairy (2018). The econometric conditions for instruments' validity in LP without control variables have first appeared in lecture notes by Mertens (2014). Stock and Watson (2018) have provided a unified discussion of the use of external instruments in macroeconomics, laid out the conditions for identification in SVARs under full invertibility, and in LP with and without controls. Recently, Arias et al. (2018) have proposed algorithms for exact finite-sample inference for SVAR-IV when multiple instruments are used to identify more than one shock.

This paper adds to the literature that focuses on the link between the conditions for invertibility of structural shocks and the information included in VARs, e.g. Giannone and Reichlin (2006), Forni and Gambetti (2014), and Canova and Hamidi Sahneh (2017). The concepts of partial and 'approximate' invertibility (shocks are revealed only to some degree) were introduced in Sims and Zha (2006), and discussed in Stock and Watson (2018) and Forni et al. (2019). Closest to our approach is the work of Forni et al. (2019) that discuss when an SVAR is informative enough to achieve approximate invertibility and to provide reasonable estimates the dynamic effects of a shock. That paper offers an independently derived expression of the Wold representation that is equivalent to our semi-structural representation in the case of one invertible shock. Differently from Forni et al. (2019), we focus on IV methods under partial invertibility and derive a proof of the existence of the semi-structural representation in a general case with several invertible shocks. Stock and Watson (2018) proposed a first proof of the possibility of IV identification under partial invertibility. Our work provides a new proof of that result and, importantly, specifies the conditions under which IV methods are successful in SVAR under partial invertibility. We also introduce a formal test for some of these conditions and characterise the bias induced by their violation.

2. Partial invertibility

Structural models, such as for example DSGE models, generally have a VARMA solution of the form[2]

$$\Phi(L)Y_t = \Psi(L)u_t \qquad u_t \sim \mathcal{WN}(0, \mathbb{I}_n) \tag{1}$$

where Y_t is a covariance-stationary vector stochastic process, $\Phi(L)$ and $\Psi(L)$ are generic autoregressive (AR) and moving average (MA) components, and u_t are the structural shocks, assumed to be orthogonal processes. The Wold Representation Theorem guarantees that the covariance-stationary Y_t always admits an MA decomposition

$$Y_t = C(L)\nu_t \qquad \nu_t \sim \mathcal{WN}(0, \Sigma_\nu), \tag{2}$$

where ν_t are the Wold innovations, defined as the residuals of the linear projection of Y_t on its past – i.e. $\nu_t = Y_t - Proj(Y_t|Y_{t-1}, Y_{t-2}, \dots)$ –, and the inverse of $C(L)$ is well defined. Hence, Eq. (2) can be rewritten as a VAR

$$A(L)Y_t = \nu_t \qquad A_0 = \mathbb{I}_n \tag{3}$$

where $A(L) = C(L)^{-1}$. The property of invertibility of the structural shocks guarantees the existence of a linear map between ν_t and u_t of the form

$$\nu_t = \Theta_0 u_t \tag{4}$$

where Θ_0 is non-singular. This is equivalent to stating that all the structural disturbances in u_t can be recovered from current and lagged values of Y_t, i.e. $u_t = Proj(u_t|Y_t, Y_{t-1}, \dots)$. Full invertibility is a highly desirable property since it allows to study the dynamic causal effects of all the structural shocks u_t on Y_t. If Eq. (4) holds, we can rewrite the Wold Representation in terms of the true economic shocks, as a Structural MA

$$Y_t = C(L)\Theta_0 u_t = C_0\Theta_0 u_t + C_1\Theta_0 u_{t-1} + \dots + C_n\Theta_0 u_{t-n} + \dots, \tag{5}$$

where the matrices $C_n\Theta_0$ collect the coefficients of the impulse response functions (IRFs) for each shock at horizon n.

While highly desirable, full invertibility rarely holds in the VARs used in empirical macro. Phenomena such as anticipation and foresight, that are often a feature of rational expectation models, can generate non-invertible representations (see e.g. Leeper et al., 2013; Lippi and Reichlin, 1993). More generally, full invertibility is also unlikely to hold in small VARs due to omitted variables and insufficient lag lengths.

In most empirical applications, however, only one or a subset of the structural shocks are of interest. That is, only one or a subset of the columns of Θ_0 need to be identified ('partial identification'). For example, one may want to identify only a monetary policy shock, or an oil price shock. In such settings, the relevant property is that of partial invertibility of the subset of the shocks of interest (see Sims and Zha, 2006).

Definition 1 (Partial Invertibility)**.** Let Y_t be an n-dimensional covariance-stationary vector stochastic process, with rational spectral density, solution to Eq. (1) that admits the representation in Eq. (2). Let u_t^i denote one element of u_t. The structural shock u_t^i is invertible and Y_t–fundamental if

$$u_t^i = Proj(u_t^i|\mathcal{H}_t^Y). \tag{6}$$

where $\mathcal{H}_t^Y$ denotes the Hilbert space generated by all the observations of Y_t up to time t, i.e. $\mathcal{H}_t^Y = \overline{\text{span}}\{Y_{t-j}, j \geq 0\}$. Hence, u_t^i is a linear combination of the innovations ν_t, that is, there exists an n-dimensional vector λ such that

$$u_t^i = \lambda'\nu_t. \tag{7}$$

More generally, partial invertibility holds if a subset $u_t^{1:m} \equiv (u_t^1, \dots, u_t^m)'$ with $m < n$ of the structural shocks can be correctly recovered as a linear combination of the estimated VAR innovations. While seldom acknowledged, partial invertibility

[2] A more formal discussion is provided in the Online Supplementary Material.

is almost always implicitly assumed in the empirical macroeconomic literature concerned with evaluating the effects of a specific shock.

As noted, Eq. (5) is key to study the effects of all the structural shocks in u_t in a VAR. The following proposition guarantees the existence of a 'semi-structural' representation for Y_t that in turn allows for the identification of the IRFs to the partially invertible shocks $u_t^{1:m}$ under the most common identification schemes.

Proposition 1 (Semi-structural Moving Average Representation)**.** *Let the n-dimensional covariance stationary vector process Y_t be solution to Eq. (1), and let $\Psi(L)$ be a non-invertible moving average filter, i.e. $det(\Psi(z)) = 0$ for some ζ_i such that $|\zeta_i| < 1$. Y_t admits the representation in Eq. (2). If the system is partially invertible in the shocks u_t^i, for $i = 1, \dots, m$, i.e. there exist m vectors λ_i such that $\lambda_i'\nu_t = u_t^i$, then Y_t admits a semi-structural moving average representation of the form*

$$Y_t = C(L)\Sigma_\nu \sum_{i=1}^{m} \lambda_i u_t^i + C(L)\Sigma_\nu \tilde{\lambda}\xi_t \tag{8}$$

where ξ_t is an $(n-m) \times 1$ vector of linear combinations of Wold innovations that is orthogonal to all u_t^i for $i = 1, \dots, m$, i.e. $\mathbb{E}(u_t^i \xi_t') = 0$.

Proof. See Online Supplementary Material. □

Proposition 1 is a representation result that guarantees the existence of a semi-structural MA representation for any covariance-stationary process, and for a generic number of invertible shocks.[3] The first term of Eq. (8) depends on the invertible shocks u_t^i for $i = 1, \dots, m$. The second term depends on $(n-m)$ linear combinations of ν_t denoted by ξ_t and orthogonal to the invertible shocks. Due to the presence of Blaschke factors (Lippi and Reichlin, 1994), ξ_t is a convolution of past, current and future non-invertible shocks. It is worth stressing that, while the requirement that ξ_t and the invertible shocks $u_t^{1:m}$ are orthogonal is important, ξ_t does not need to span the space of all the non invertible structural shocks. While the representation in Eq. (8) always exists, it is not unique.[4]

Importantly, Proposition 1 implies that if the VAR correctly captures the autocovariance structure of Y_t, the 'partially' identified SVAR impulse response functions $C(L)\Sigma_\nu\lambda_i u_t^i$ are the dynamic causal effects of the m invertible shocks.

Finally, the following remark generalises the map in Eq. (4) to the case of partial invertibility.

Remark 1. Under partial invertibility, the map between structural shocks and Wold innovations is of the form

$$\nu_t = B(L)u_t = \begin{pmatrix} b_1 & b_2(L) \end{pmatrix} u_t \tag{9}$$

where b_1 is a $n \times m$ matrix, with m the number of partially invertible shocks, and $b_2(L)$ is of dimensions $n \times (n-m)$ and contains Blaschke factors.

Proof. See Online Supplementary Material. □

3. IV identification under partial invertibility

We now focus on the conditions for identification via external instruments in the setting provided by Eq. (8). The intuition that we formalise in what follows is that identification is achieved when the IV, denoted by z_t, correlates with the VAR residuals only via the invertible shock of interest.

Let the structural shock u_t^1 be invertible from a VAR with reduced-form representation as in Eq. (3). Given an external instrument z_t, it is possible to identify u_t^1 and its effects on Y_{t+h}, $h = 0, \dots, H$, under the conditions in the following proposition.

Proposition 2 (Identification in SVAR-IV under Partial Invertibility)**.** *Let $u_t^{1:m}$ denote the m structural shocks invertible from a VAR in Y_t, and $u_t^{m+1:n}$ the remaining $n-m$ non-invertible shocks. Let z_t be a candidate IV for the shock of interest u_t^1 and define $z_t^\perp = z_t - Proj\left(z_t | \mathcal{H}_{t-1}^Y\right)$. The impact effects of u_t^1 on Y_t are identified up to a scaling factor if z_t satisfies the following conditions:*

(i) $\mathbb{E}[u_t^1 z_t^\perp] = \alpha$ *(Relevance)*
(ii) $\mathbb{E}[u_t^{2:n} z_t^\perp] = 0$ *(Contemporaneous Exogeneity)*
(iii) $\mathbb{E}[u_{t-j}^{m+1:n} z_t^\perp] = 0$ *for all $j \neq 0$ for which $\mathbb{E}[u_{t-j}^{m+1:n}\nu_t'] \neq 0$. (Limited Lead-Lag Exogeneity)*

Moreover, if the VAR correctly captures the autocorrelation structure of the Wold representation of Y_t, the horizon h impulse response functions identified via z_t correctly estimate the causal effects of u_t^1 up to a constant, i.e.

$$IRF_{i1}^h \propto [C_h \Sigma_\nu \lambda_1]_i \tag{10}$$

where C_h are the matrix coefficients of the Wold representation at lag h.

[3] In an independently derived result, Forni et al. (2019) propose a moving average equation (their Definition 4) similar to Eq. (8) for the $m = 1$ case.

[4] See the Online Supplementary Material for a more detailed discussion.

Proof. See Online Supplementary Material. □

Conditions (i) and (ii) are the standard validity conditions required for IV identification. Condition (iii) arises because of the dynamics, and requires that if there are any non-invertible shocks, they do not correlate at any leads and lags with the component of the instrument that is orthogonal to past Y_t. Conversely, leads and lags (but not contemporaneous values) of other partial invertible shocks can contaminate the instrument without compromising the identification since they do not enter the VAR residuals.[5]

If the system is invertible in all the structural shocks, Condition (iii) is trivially satisfied, since ν_t are a linear combination of the contemporaneous structural shocks only. These are the conditions discussed by Stock and Watson (2018). Conversely, if u_t^1 is the only invertible shock, Condition (iii) implies a stronger lead-lag exogeneity condition that applies to all the shocks but the invertible one. In the more general case in which only some of the remaining shocks are non-invertible, Proposition 2 ensures that identification with an external instrument is possible as long as the instrument is contaminated only by the past and future realisations of the invertible shocks. It is worth stressing that while Condition (iii) is a relatively stronger condition than that required for a fully invertible SVAR (where lead-lag exogeneity is not required), it is still much weaker than the strong lead-lag exogeneity condition required for identification in LP-IV without controls.

When Conditions (ii) or (iii) are violated, the contamination of the IV induces a bias in the IRFs, as formalised in the following remark.

Remark 2 (Violation of the Exogeneity Conditions)**.** Let z_t be a candidate IV for the invertible shock u_t^1 that satisfies Condition (i) but fails Condition (ii) and Condition (iii) due to contamination by lags, leads or contemporaneous realisations of a non-invertible shock $u_t^{\not 1}$, i.e.

$$z_t = \alpha u_t^1 + \sum_{k \in K} \beta_k u_{t-k}^{\not 1}. \tag{11}$$

Given a VAR that correctly captures the autocorrelation structure of Y_t the IRFs identified via z_t are biased and, up to a constant, of the form

$$\widetilde{IRF}_{i1}^h \propto IRF_{i1}^h + \left[C_h \sum_{j \in J} \sum_{k \in K} b_{2,j,\not 1} \frac{\beta_k}{\alpha} \delta_{jk} \right]_i \tag{12}$$

where IRF_{i1}^h are the IRFs for variable i to the shock u_t^1 at horizon h, C_h are the coefficients of the Wold representation at lag h, $b_{2,j,\not 1}$ is the $\not 1$ column of the matrix of coefficients of the polynomial $b_2(L)$ at lag j and δ_{jk} is the Kronecker's delta.

Proof. See Online Supplementary Material. □

A few elements of Eq. (12) are worth highlighting. First, all else equal, the size of the bias depends on how much the IV correlates with the (leads, lags, or contemporaneous realisations of the) contaminating shock relative to the shock of interest – i.e. on the ratios $\frac{\beta_k}{\alpha}$. Second, the bias depends on the number of lags that are common to those contaminating the instrument (Eq. 11) and those that appear in the Blaschke matrix $b_2(L)$ defined in Remark 1. Finally, and importantly, the bias depends on the relative order of magnitude of the coefficients $b_{2,j,\not 1}$ and b_1 that relate to the variance of variable i that is accounted for by $u_t^{\not 1}$ and u_t^1 respectively. For example, very small values of $b_{2,j,\not 1}$ relative to b_1 imply that shock $u_t^{\not 1}$ explains a small share of the variance of variable i, and hence the distortion is likely to be small.

For comparison, if the IV is contaminated by leads, lags and contemporaneous realisations of another invertible shock, Eq. (12) simplifies to

$$\widetilde{IRF}_{i1}^h = IRF_{i1}^h + \left[C_h b_{1,\not 1} \frac{\beta_0}{\alpha} \right]_i. \tag{13}$$

In fact, only the violation of Condition (ii) matters, i.e. the correlation with the contemporaneous realisation of $u_t^{\not 1}$.

4. SVAR-IV Under partial invertibility and LP-IV

As an alternative to VARs, local projections (LP) can also be used to estimate structural IRFs. As discussed in Stock and Watson (2018), when no control variables are included in the LP, identification is achieved only under stricter conditions that require lead-lag exogeneity of the IV, i.e.

(i) $\mathbb{E}[u_t^1 z_t] = \alpha$ (Relevance)

[5] Interestingly, leads, lags or even contemporaneous realisations of the non-invertible shocks can contaminate z_t, but only via their 'projectable' component $Proj(u_t^{m+1:n} | \mathcal{H}_{t-1}^Y) \neq u_t^{m+1:n}$ that lives in the space spanned by past realisations of Y_t.

Table 1
Estimation of the Dynamic Causal Effects of u_t^1.

	u_t^1 invertible	u_t^1 non-invertible
Strong Lead-Lag Exogeneity		
$\mathbb{E}[u_{t-j}^i z_t] = 0 \ \forall\ i \ \&\ j \neq 0$	LP-IV SVAR-IV SVAR-H	LP-IV SVAR-H
Limited Lead-Lag Exogeneity but Contamination by Past Shocks		
$\mathbb{E}[u_{t-j}^i z_t] \neq 0$ for some $j > 0$ ($= 0$ for $j < 0$) but $\mathbb{E}[u_{t-j}^i z_t^\perp] = 0$ and $\mathbb{E}[u_{t-j}^i \nu_t'] = 0$	LP-IV$^\perp$ SVAR-IV SVAR-H	LP-IV$^\perp$ SVAR-H
Limited Lead-Lag Exogeneity but Contamination by Future Shocks		
$\mathbb{E}[u_{t-j}^i z_t] \neq 0$ for some $j < 0$ but $\mathbb{E}[u_{t-j}^i z_t^\perp] = 0$ and $\mathbb{E}[u_{t-j}^i \nu_t'] = 0$	SVAR-IV	–
Violation of Limited Lead-Lag Exogeneity		
$\mathbb{E}[u_{t-j}^i z_t^\perp] \neq 0$, $j > 0$ and i s.t. $\mathbb{E}[u_{t-j}^i \nu_t'] \neq 0$	–	–

Note: The table reports the methods that are able to correctly estimate the dynamic effects of u_t^1 on a given vector Y_t depending on whether u_t^1 is invertible or not, and on the properties of the instrument z_t (in rows). $^\perp$ denotes orthogonality with respect to $\mathcal{H}_{t-1}^Y$. It is assumed that the conditions of Relevance $\left(\mathbb{E}[u_t^1 z_t] = \alpha\right)$ and Contemporaneous Exogeneity $\left(\mathbb{E}[u_t^{2:n} z_t] = 0\right)$ hold throughout.

(ii) $\mathbb{E}[u_t^{2:n} z_t] = 0$ (Contemporaneous Exogeneity)
(iii) $\mathbb{E}[u_{t-j}^{2:n} z_t] = 0$ for all $j \neq 0$ (Strict Lead-Lag Exogeneity.)

In the more likely case in which the instrument correlates with some past shocks, the standard practice is to incorporate lagged macro variables in the LP regression, in order to control for these lagged shocks (LP-IV$^\perp$), i.e.

$$Y_{i,t+h} = \Theta_{h,i1}\widehat{Y}_t^1 + \gamma_h' W_t + \zeta_{i,t+h}^h \tag{14}$$

where W_t denotes a generic set of controls, $\Theta_{h,i1}$ are the causal responses of $Y_{i,t+h}$ to u_t^1 at horizon h, $\widehat{Y}_t^1$ is the fitted value of the instrumented variable Y_t^1 from the first-stage regression on z_t, and $\zeta_{i,t+h}^h$ are serially correlated residuals. The conditions for identification are (see Stock and Watson, 2018)

(i) $\mathbb{E}[u_t^{1,\perp} z_t^\perp] = \alpha$ (Relevance)
(ii) $\mathbb{E}[u_t^{2:n,\perp} z_t^\perp] = 0$ (Contemporaneous Exogeneity)
(iii) $\mathbb{E}[u_{t-j}^\perp z_t^\perp] = 0$ for all $j \neq 0$ (Lead-Lag Exogeneity)

where $x_t^\perp = x_t - Proj(x_t|\mathcal{W}_t)$ for a given x_t, and $\mathcal{W}_t = \overline{\text{span}}\{W_t\}$.

For LP-IV$^\perp$, Stock and Watson (2018) provide a 'no-free lunch' result by showing that, in general, this is equivalent to assuming full invertibility of a VAR with the same information set. The intuition for this result is that, if it is not known which are the lagged shocks that contaminate the instrument, the only way to achieve identification is to ensure that all the possible past shocks are controlled for, which is equivalent to including all the controls that make a VAR in the same variables fully invertible.[6] In this section we generalise this result to the case of partial invertibility, and show that also in this case the conditions required in LP-IV and SVAR-IV are the same.

The following proposition shows that an instrument that correctly identifies the shock of interest (up to a normalisation) in a SVAR-IV under partial invertibility, will also generally identify the same shock in LP-IV$^\perp$ when $\mathcal{W}_t \equiv \mathcal{H}_{t-1}^Y$, and vice versa. Conversely, an instrument that identifies a non-invertible shock in LP-IV$^\perp$ will also identify that same shock in a SVAR if used as an internal instrument, i.e. in a hybrid VAR (SVAR-H) specified on $\left(z_t' \ Y_t'\right)'$ (see also Plagborg-Møller and Wolf, 2018).

Proposition 3 (Relation between SVAR-IV under Partial Invertibility and LP-IV. $^\perp$) *Let Z be the set of scalar stochastic processes z_t that satisfy LP-IV Conditions (i) and (ii) – i.e. $\mathbb{E}[u_t^1 z_t] = \alpha$ and $\mathbb{E}[u_t^{2:n} z_t] = 0$ –, but satisfy Condition LP-IV (iii) $\mathbb{E}[u_{t-j} z_t] = 0$ only for $j < 0$ and not for $j > 0$. Let $\widetilde{Z} \subseteq Z$ be such that any $z_t \in \widetilde{Z}$ satisfies the LP-IV$^\perp$ conditions for $\mathcal{W}_t \equiv \mathcal{H}_{t-1}^Y$. Assume also that $Proj(u_t|\mathcal{H}_{t-1}^Y) = 0$. z_t is an element of $\widetilde{Z}$ if and only if it either (a) satisfies the conditions for SVAR-IV identification under partial invertibility, or (b) satisfies the conditions for SVAR-H identification under non-invertibility.*

Proof. See Online Supplementary Material. □

Table 1 summarises the content of the proposition and compares SVARs and LPs in terms of their ability to correctly estimate the IRFs (up to a normalisation) of the shock of interest u_t^1 on Y_t, for given IV z_t. The rows in the table consider

[6] In other words, when some past shocks are not absorbed by the LP controls, there will exist some pathological IV in the class of IVs contaminated only by past shocks, that satisfies Condition (ii) but fails Condition (iii), even conditional on the controls. In fact, the set of controls that absorbs all past shocks is the one that guarantees that the VAR is invertibile.

different properties of z_t, while the columns of the table distinguish between the cases in which u_t^1 is invertible or not. It is understood that invertibility of any u_t^i, $i = 1, \ldots, n$, is to be intended relative to $\mathcal{H}_{t-1}^Y$.

Conditional on the same information set and IV, SVARs (as SVAR-IV or SVAR-H) and LP-IV with controls generally identify a shock under the same set of conditions. Hence, the choice between LP and VAR should be solely dictated by the specific empirical constraints imposed by the availability of the sample and variables of interest, and in light of the different finite-sample bias-variance properties of the two methods, as observed by Plagborg-Møller and Wolf (2018).[7]

It is worth observing that the three methods – LP-IV with controls, SVAR-IV and SVAR-H – deliver similar responses in most but not all of the relevant empirical cases. Hence, they can be used to empirically gauge violations of the conditions for identification.

5. IV identification and VAR (mis)specification

In Section 3 we discussed how the contamination of the instrument biases both the impact and the dynamic responses. In this section we show that as long as partial invertibility and the conditions in Proposition 2 hold, model misspecification biases the dynamic responses but does not prevent the correct identification of the impact effects.[8] This result can be used to detect potential contamination of the IV.

Let the data generating process for $Y_t = (y_{1,t}' \ \ y_{2,t}')'$ be a generic purely nondeterministic stationary VARMA(p,q)

$$\begin{pmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \Psi_{11}(L) & \Psi_{12}(L) \\ \Psi_{21}(L) & \Psi_{22}(L) \end{pmatrix} \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}. \tag{15}$$

Fitting a VAR(k) to $y_{1,t}$ corresponds to erroneously imposing some or all of the following restrictions

$$\Phi_{11,i} = 0, \ \text{for } i = k+1, k+2, \ldots p; \tag{16}$$

$$\Phi_{12,i} = 0, \ \text{for } i = 1, 2, \ldots p, \tag{17}$$

$$\Psi_{11,i} = 0, \ \text{for } i = 1, 2, \ldots q; \qquad \Psi_{12,i} = 0, \ \text{for } i = 1, 2, \ldots q. \tag{18}$$

When $k < p$, the first restriction (Eq. 16) corresponds to understating the VAR lag order. The second restriction (Eq. 17) omits the variables in $y_{2,t}$; this is also a trivial case of non-invertibility due to the number of variables being smaller than the number of shocks. Finally, the restrictions in Eq. (18) disregard the MA structure of the process.

We now examine what these misspecifications imply for the identification of the shock of interest u_t^1. Assume that partial invertibility for u_t^1 holds on the n_1-dimensional subvector $y_{1,t}$, i.e. $u_t^1 = Proj(u_t^1 | y_{1,t}, y_{1,t-1}, \ldots)$. Hence, u_t^1 can be obtained from the linear projection of $y_{1,t}$ on its lags.

Consider the case of a too short lag order (Eq. 16). In this case, the autoregressive coefficients are biased and inconsistent (see Braun and Mittnik, 1993). But, if k is sufficient to obtain partial invertibility of u_t^1, identification of the impact effects is preserved. Hence, while impact responses are correctly estimated, IRFs at larger horizons are distorted even asymptotically. Exactly the same logic applies to the case of misspecified MA components that can always be mapped into a VAR with infinitely many lags. However, while in the first case more lags trivially resolve the issue, in the second case longer lags approximate the true Wold representation only asymptotically.

Consider now the case of omitted variables (Eq. 17). If partial invertibility holds, the impact effects are correctly retrieved, while the IRFs at longer horizons are distorted. Interestingly, however, also in this case increasing the number of lags recovers the true IRFs asymptotically. To see this, note that the Wold Representation Theorem implies that $y_{1,t}$ has an invertible MA representation. In turn, this guarantees that the dynamics of the system are asymptotically approximated by infinitely many lags of $y_{1,t}$ only.[9]

These observations provide a simple way to gauge the contamination of the available IV versus the misspecification of the chosen model. If one can assume partial invertibility across different VAR specifications, an instrument that satisfies the conditions in Proposition 2 delivers stable impact responses but unstable IRFs across models. In this case, increasing the number of lags and/or selectively adding variables that may be important for the transmission of the shock should help stabilising the IRFs. Conversely, an instrument that violates the limited lead-lag exogeneity condition is likely to also deliver unstable impact responses across different models.[10]

[7] Under partial invertibility SVAR-IV can be identified also when the instrument correlates with future invertible shocks, while this is never possible for LP-IV with or without controls (nor for SVAR-H). However, these cases are empirically unlikely.

[8] Canova and Ferroni (2019) provide a background to and complement our discussion by analysing how VAR misspecification challenges the identification of structural shocks.

[9] If Y_t is covariance-stationary, $y_{1,t} \equiv J Y_t$, where J_t is a selector matrix, is also covariance stationary, with first and second moments respectively equal to $\mathbb{E}(y_{1,t}) = J\mathbb{E}(Y_t)$, and $\Gamma_{y_1}(h) = J\Gamma_Y(h)J'$, where $\Gamma(h)$ is the autocovariance of Y_t at lag h.

[10] Rich information sets can help in this case, since structural shocks are likely invertible in larger models (Giannone and Reichlin, 2006). In turn, this improves the performance of contaminated IVs.

6. A test For conditional (lagged) exogeneity

It is possible to formalise the previous discussion using a statistical test. The i-th equation of the structural model in Eq. (1) can be rewritten as

$$Y_{i,t} = b_{i1}u_t^1 + \gamma_i^{'}W_t + U_{i,t} \qquad i = 1, \dots, n \tag{19}$$

where W_t denotes a generic set of control variables (e.g. lagged values of components of Y_t) and U_t absorbs the structural components of $Y_{i,t}$ not captured by W_t and not due to u_t^1. All the methods discussed – SVAR-IV, SVAR-H, LP-IV$^{\perp}$ – estimate the causal responses of Y_t to u_t^1 by means of linear regressions of the form[11]

$$Y_{i,t} = b_{i1}^{*}z_t + \gamma_i^{*'}W_t + \zeta_{i,t} \qquad i = 1, \dots, n. \tag{20}$$

As is well known from the microeconometrics literature, structural interpretation of the coefficient of the 'predictive regression' in Eq. (20), i.e. $b_{i1}^{*} \propto b_{i1}$, is obtained under the exogeneity of z_t conditional on W_t; that is, $\mathbb{E}[z_t^{\perp}U_{i,t}] = 0$, where $z_t^{\perp} = z_t - Proj(z_t|\mathcal{W}_t)$. Along with the relevance condition, this guarantees that z_t correlates with $Y_{i,t}$ only via the shock of interest u_t^1 at time t, conditional on W_t.

The minimum set W_t for which the IV satisfies conditional exogeneity can be defined as the 'core information set'.[12] This observation offers a different way to frame the identification with IVs. The previous sections formalised the conditions that the instrument must satisfy, conditional on given empirical model and information set. The "dual" problem is to define a core information set such that the exogeneity condition is satisfied, conditional on a given IV.[13]

Crucially, if W_t is correctly identified and included in the empirical model, estimates of b_1^{*} are insensitive to the inclusion of additional controls. This intuition can be formalised with a Hausman type test, following White and Chalak (2010) and Lu and White (2014). Consider the core regression model for Eq. (19)

$$Y_{i,t} = b_{i,1}^{C*}z_t + \gamma_i^{C*'}W_t^C + \zeta_{i,t}^C \qquad i = 1, \dots, n \tag{21}$$

along with an alternative regression model

$$Y_{i,t} = b_{i,1}^{A*}z_t + \gamma_i^{A*'}W_t^C + \kappa_i^{*'}W_t^A + \zeta_{i,t}^A \qquad i = 1, \dots, n \tag{22}$$

where W_t^C denotes the candidate core information set, and W_t^A are additional non-core controls, e.g., other variables and/or other lags of those in the core set. If the instrument is exogenous conditional on W_t^C, the coefficients associated to z_t in Eqs. (21) and (22) have to coincide. That is, testing for the conditional exogeneity of z_t is equivalent to testing the null $\mathbb{H}_0 : b_{i,1}^{C*} = b_{i,1}^{A*}$.[14]

Let the OLS estimates over a sample of size T of the coefficients of the two regression models be, respectively, $\hat{\delta}_i^C = (\hat{b}_{i,1}^C \;\; \hat{\gamma}_i^{C'})'$ and $\hat{\delta}_i^A = (\hat{b}_{i,1}^A \;\; \hat{\gamma}_i^{A'} \;\; \hat{\kappa}_i')'$. Define the joint vector $\hat{\delta}_i \equiv (\hat{\delta}_i^{C'} \;\; \hat{\delta}_i^{A'})'$. Under the mild conditions discussed in Chalak and White (2011), and without assuming correct specification

$$\sqrt{T}(\hat{\delta}_i - \delta_i^{*}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, M_i^{*-1}V_i^{*}M_i^{*-1}) \tag{23}$$

where, accordingly, $\delta_i^{*} \equiv \left(\delta_i^{C*'} \;\; \delta_i^{A*'}\right)'$.[15]

Let now S be a selection matrix such that $S\hat{\delta}_i = (\hat{b}_{i,1}^C \;\; \hat{b}_{i,1}^A)'$ and define the differencing vector Δ such that $\Delta S\hat{\delta}_i = \hat{b}_{i,1}^C - \hat{b}_{i,1}^A$. The test statistic is

$$\mathcal{R}_{i,T} = T\hat{\delta}_i'(\Delta S)'[(\Delta S)\hat{M}_i^{-1}\hat{V}_i\hat{M}_i^{-1}(\Delta S)']^{-1}\Delta S\hat{\delta}_i \tag{24}$$

where $\hat{V}_i$ and $\hat{M}_i$ are consistent estimators of M_i^{*} and V_i^{*} respectively, and $[(\Delta S)\hat{M}_i^{-1}\hat{V}_i\hat{M}_i^{-1}(\Delta S)']$ is assumed to be nonsingular. Under the null, $\mathcal{R}_{i,T} \xrightarrow{\mathcal{D}} \chi_1^2$. As is standard, $\mathbb{H}_0$ is rejected at the α level if $\mathcal{R}_{i,T}$ exceeds the $1-\alpha$ percentile of the χ_1^2 distribution. Being a standard parametric test, it has power against local alternatives at rate $T^{-1/2}$ (Lu and White, 2014).

For a given IV, the test can be used to guide the choice of the VAR core information set by assessing the robustness of the impact responses to the inclusion of additional controls. For LPs, the same logic can be used for the IRF coefficients at all horizons. It is important to note that while necessary, the robustness of the coefficients is not sufficient for valid

[11] In the two-step SVAR-IV, the structural impacts are identified (up to scale) from a regression of the first-stage VAR residuals on the IV. This is equivalent to instrumenting a variable in the VAR and then regressing all other variables on their lags and the instrumented variable. Alternatively, the IV can be included among the endogenous variables (SVAR-H). In the LP-IV framework, the same applies at each horizon.

[12] Also known as the "minimum relevant information set" (Heckman and Navarro-Lozano, 2004) or the set of "core covariates" (Lu and White, 2014) in microeconometrics.

[13] For a perfect IV (or one with classic measurement error), the core set has dimension zero; any regression of $Y_{i,t}$ on the IV delivers the correct impacts. For IVs contaminated by other shocks at time t, the core set can be impossible to obtain, given the observables.

[14] The test is easily generalised to the case of many alternative regression models. See Online Supplementary Material.

[15] The matrices M_i^{*} and V_i^{*} are defined as follows. $M_i^{*} = diag\left(M^{C*}, M^{A*}\right)$, where $M^{C*} = \mathbb{E}\left(X_t^{C'}X_t^C\right)$, $M^{A*} = \mathbb{E}\left(X_t^{A'}X_t^A\right)$ and X denotes the full set of regressors in each model. $V_i^{*} = \begin{pmatrix} V_i^{CC} & V_i^{CA} \\ V_i^{AC} & V_i^{AA} \end{pmatrix}$ where $V_i^{jk} = \mathbb{E}\left[\left(X_t^{j'}\zeta_{i,t}^j\right)'\left(X_t^{k'}\zeta_{i,t}^k\right)\right]$ for $j,k = \{A, C\}$.

causal inference. For example, while rejection of $\mathbb{H}_0$ indicates that conditional exogeneity is not satisfied, failure to reject does not rule out that the IV may violate contemporaneous exogeneity, which the test cannot detect.[16] Hence, our test does not exonerate researchers from providing explicit arguments in support of the relevance and (conditional) exogeneity assumptions of the chosen IV.

7. Monetary policy shocks in a simulated system

We illustrate the theoretical results in the previous sections using a simulated environment. We use a stylised New Keynesian DSGE model that features (*i*) a representative infinitely-lived household that chooses between consumption and leisure; (*ii*) firms that produce a continuum of goods using a Cobb-Douglas technology to aggregate capital and labour; (*iii*) a government that consumes a share of output for wasteful public spending; and (*iv*) a central bank that sets the interest rate using a Taylor rule with smoothing. There are four stochastic disturbances that generate fluctuations in the economy, namely, a monetary policy shock u_t^r, a government spending shock u_t^g, a technology shock u_t^a, and an inflation-specific shock u_t^π.

The processes for technology, spending, inflation, and the policy rate are defined as follows. Log technology a_t evolves with a news component $a_t = \rho_a a_{t-1} + \sigma_a u_{t-4}^a$, where u_t^a is an i.i.d. normally distributed technology news shock that affects the technology process after 4 periods. Similarly, fiscal foresight characterises the spending process g_t, that evolves according to $g_t = \rho_g g_{t-1} + u_{t-4}^g$, where u_t^g is an i.i.d. normally distributed spending shock. The monetary authority sets the nominal interest rate using a Taylor rule with smoothing

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left(\phi_\pi \overline{\pi_t} + \phi_y \overline{\Delta y_t}\right) + \sigma_r u_t^r \tag{25}$$

where $\overline{\pi_t}$ is the average inflation rate over the last four periods, $\overline{\Delta y_t}$ is the average growth rate of output, and u_t^r is a white noise i.i.d. normally distributed monetary policy shock. Finally, price dynamics are governed by a New Keynesian Phillips Curve, as follows

$$\pi_t = \gamma_\pi \pi_{t-1} + \beta \mathbb{E}_t \pi_{t+1} + \frac{(1-\theta_\pi)(1-\theta_\pi \beta)}{\theta_\pi} mc_t + u_t^\pi, \tag{26}$$

where mc_t are marginal costs, and u_t^π is an i.i.d. normally distributed inflation-specific shock. All the model details, including the calibrated parameters, are reported in the Online Supplementary Material.

We consider a VAR(4) in the policy rate, inflation, output, and government spending. Under standard calibration, the model fails the 'poor man's invertibility condition' of Fernandez-Villaverde et al. (2007), hence, the four structural shocks cannot all be recovered from a VAR in the model's observables. However, the specification of the Taylor rule ensures that the monetary policy shock is partially invertible from a VAR(4) in $[r_t, \pi_t, y_t]'$.[17]

The degree of invertibility of each of the structural shocks for the VAR(4) are as follows: $\delta_r = 0.069$, $\delta_a = 0.799$, $\delta_g = 0.494$, $\delta_\pi = 0.343$. The concept was introduced in Sims and Zha (2006) and takes values between 0 and 1. A value of 0 implies that the shock is invertible from the VAR, whereas increasing values of δ imply non-fundamentalness and an increasing degree of non-invertibility.[18] The value of δ for the technology shock is very close to 1, confirming the inability of the VAR to recover it. Similarly, the inflation and spending shocks are also non-invertible, but with a higher degree of invertibility. The monetary policy shock is the only invertible shock in the system. The four shocks play a different role in driving economic fluctuations in the model, notably, the government spending shock plays a negligible role (see Online Supplementary Material).

We simulate from the model 5000 economies each for a sample size of $T = 300$ periods. For each set of simulated data, we estimate a VAR(4) in the four observables, and identify the monetary policy shock using the following four external instruments

$$z_{0,t} = u_t^r \tag{27}$$

$$z_{1,t} = 0.7u_t^r - 0.5u_{t-2}^r + \varsigma_t \tag{28}$$

$$z_{2,t} = 0.7u_t^r - 0.5\left(u_{t-1}^g + u_{t-2}^g + u_{t-3}^g\right) + \varsigma_t \tag{29}$$

$$z_{3,t} = 0.7u_t^r + 0.5\left(u_{t-1}^a + u_{t-2}^a + u_{t-3}^a\right) + \varsigma_t. \tag{30}$$

[16] A possible way around this issue is to introduce among the controls instruments for other shocks that may be contaminating the instrument at test, if available.

[17] A VAR(4) also captures the model's dynamics sufficiently well. The Online Supplementary Material reports results for 1 and 2 lags.

[18] Following Forni et al. (2019) δ_i is calculated as $\delta_i = \text{var}[u_t^i - Proj(u_t^i|\mathcal{H}_t^Y)]/\sigma_{u_i}^2$, where $\sigma_{u_i}^2$ is the variance of u_t^i, and $\mathcal{H}_t^Y$ is the space spanned by Y_t and its lags. δ_i is a deterministic function of the model's parameters, and measures the unexplained variance of the orthogonal projection of each of the structural shocks on the VAR residuals.

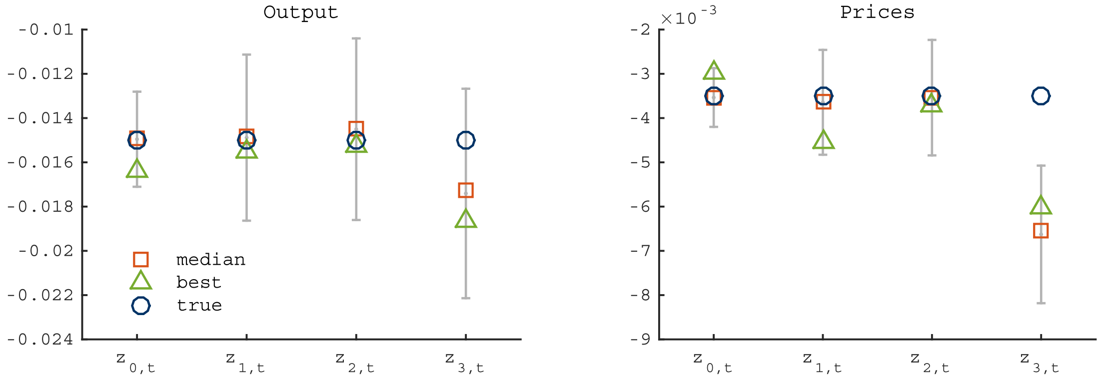


Fig. 1. Impact Responses to Monetary Policy Shock *Note*: Impact responses to monetary policy shock from partially-invertible DSGE identified with external instruments and estimated with a VAR(4) in four observables. $z_{0,t}$: observed shock case; $z_{1,t}$: instrument correlates with monetary policy shock only; $z_{2,t}$: instrument also correlates with past spending shocks; $z_{3,t}$ instrument correlates also with past technology shocks. Grey vertical lines are 2 standard deviations error bars from the distribution of impact responses across 5000 simulated economies of sample size $T = 300$ periods. True impact (blue circle), median across simulations (orange square), minimum distance from median (best) simulation (green triangle). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In Eq. (27) the shock is perfectly observable. The instrument in Eq. (28) is contaminated by classic white noise measurement error, and the second lag of the monetary policy shock. The instruments in Eqs. (29–30) both fail the limited lead-lag exogeneity condition. In fact, while $z_{2,t}$ is contaminated by lagged spending shocks, $z_{3,t}$ correlates with lagged technology shocks. In all cases, ς_t is a normally distributed random measurement error with zero mean and variance equal to that of the structural shocks.

Impact responses for output and inflation recovered from the four instruments are in Figure 1.[19] In each subplot, we use blue circles for the model's responses (true), orange squares for the median across simulations, and green triangles for the simulation that is closest to the median (best).[20] The error bars are two standard deviations intervals constructed from the distribution across simulations.

The results of this simulation validate Proposition 2. Even without full invertibility, the impact effects are correctly recovered both when the shock is observable ($z_{0,t}$) and when the instrument is contaminated with lagged invertible shocks and measurement error ($z_{1,t}$). The introduction of a measurement error in $z_{1,t}$ widens the distribution of impact responses across simulations. Instead, in the case of $z_{3,t}$, the instrument correlates with lagged non-invertible technology shocks that cannot be recovered from the VAR residuals. This results in severely biased impact responses. An interesting case arises when the instrument correlates with lagged spending shocks ($z_{2,t}$). This shock is not invertible, but it is responsible for a negligible share of the variance of the simulated variables. In this case the impact responses are close to the true ones, consistently with what noted in Remark 2. The same considerations apply to the full dynamic responses reported in Figure 2.

8. Assessing IVs in practice

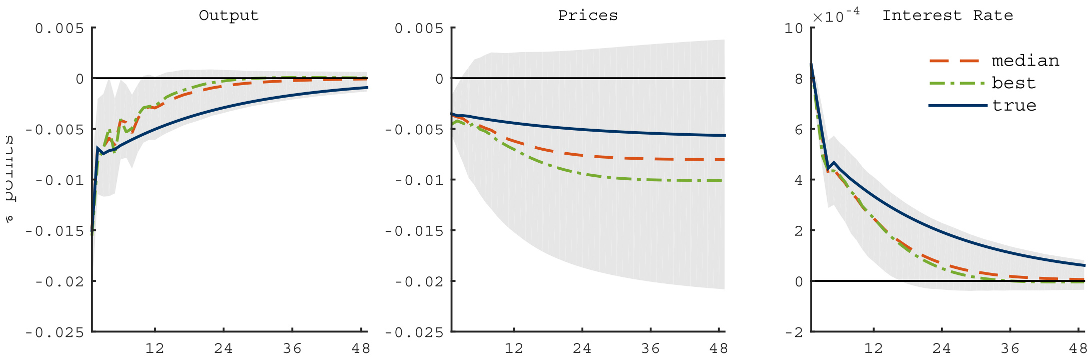
The theoretical results in this paper can be used to guide the empirical practice in the following way.

First, a Granger causality test for the predictability of the IV using past information is useful to assess the potential contamination of the IV. Second, the test of Forni and Gambetti (2014) can be used to gauge whether all or some of the shocks are invertible in the chosen VAR. Invertibility related issues can also be revealed by differences in results obtained using the same IV as an internal or external instrument with the same VAR. Third, conditional on partial invertibility of the shock of interest, the heuristic and test developed in this paper can be used to assess the lagged exogeneity condition for the chosen instrument. Stable impact effects across different VARs point to a valid IV. If concerns arise about contemporaneous contamination with another shock for which an IV is available, this can be used to assess or remove the contamination. Finally, conditional on stable impacts, lack of robustness in the dynamic responses is instead an indication of model misspecification along some dimension.

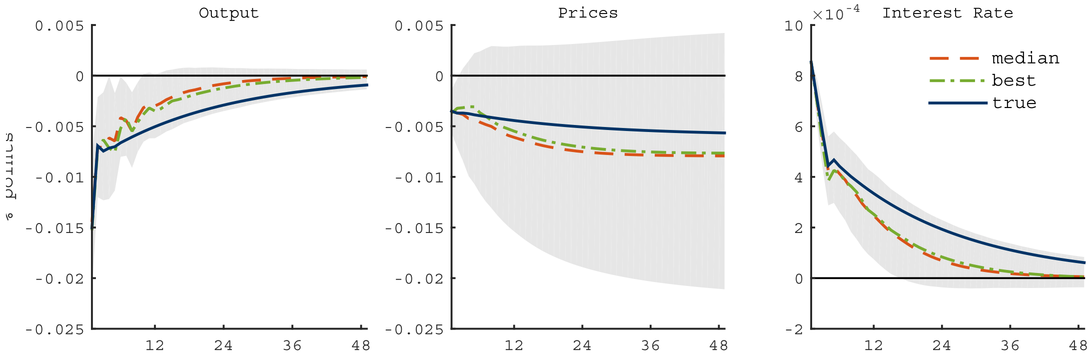
It is worth reminding that while the tests can be used to reject a hypothesis, they do not prove the alternative. Hence, care should be exercised, and explicit arguments on the plausibility of the assumptions made need to be provided.

[19] IRFs are normalised such that the impact response of the policy rate to a monetary policy shock equals that of the model.

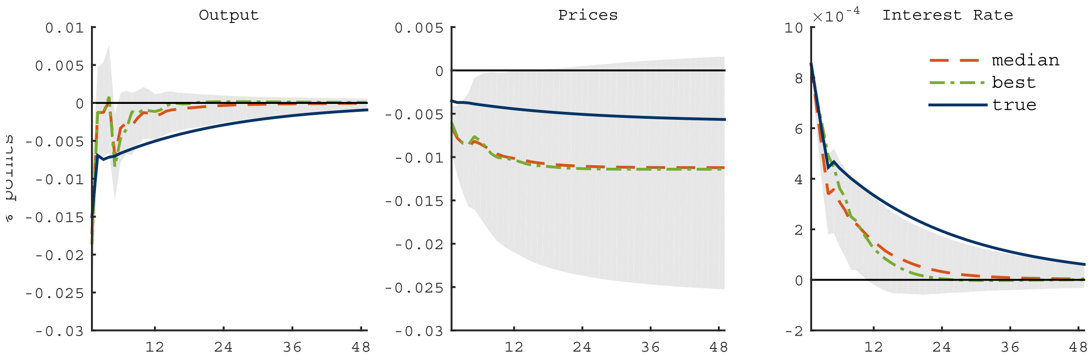
[20] We use the simulation whose IRFs minimise the sum of square deviations from median IRFs over the first 12 periods. This puts more weight on shorter horizons where IRFs have richer dynamics. Changing the truncation horizon yields qualitatively similar results.



(A) $z_{1,t}$: external instrument correlates with monetary policy shock only



(B) $z_{2,t}$: external instrument also correlates with lagged spending shocks



(C) $z_{3,t}$: external instrument also correlates with lagged technology shocks

Fig. 2. Responses to Monetary Policy Shock – Simulation *Notes:* Impulse responses to monetary policy shock from partially-invertible DSGE identified with external instruments and estimated with a VAR(4) in four observables. Instrument correlates with monetary policy shock only (Panel A). Instrument correlates with monetary policy shock and lagged spending shocks (Panel B). Instrument correlates with monetary policy shock and lagged technology shocks (Panel C). Grey shaded areas denote 90th quantiles of the distribution of IRFs across 5000 simulated economies of sample size $T = 300$ periods. Model responses (true, blue solid), median across simulations (orange dashed), minimum distance from median (best) simulation (green dash-dotted). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

9. Instrumental variables for monetary policy shocks

The recent literature on the macroeconomic effects of monetary policy has employed two leading types of IVs to identify conventional shocks. The narrative measure of Romer and Romer (2004), and the high-frequency instruments of Gürkaynak et al. (2005), as proposed by Gertler and Karadi (2015). These were recently extended by Swanson (2020) to unconventional policy actions. Despite having been extensively used, at a closer examination these instruments produce unstable results, and puzzling responses to prices and real variables, as discussed e.g. in Ramey (2016). While these fragilities are known, the debate on their sources is not settled. In this section we review these proxies through the lens of the theoretical results of this paper.

9.1. Conventional monetary policy: Narrative instruments

The seminal work of Romer and Romer (2004) jolted the literature providing a first proxy for U.S. monetary policy shocks, constructed to be orthogonal to endogenous and anticipatory movements. The measure is obtained in two steps. First, a series of federal funds rate (FFR) changes around FOMC meetings is inferred using narrative records. Second, this series is regressed on the Fed's internal forecasts (Greenbook) of output growth, unemployment and inflation to derive a measure free of systematic responses to information about current and expected developments.

In the original paper the proxy was used (along with its lags) in a distributed lag regression that also included the current realisations of output and prices, thus imposing a recursiveness assumption. Since then, the measure has found application in hundreds of papers either as an internal or external instrument in VARs and LPs. Coibion (2012) and Ramey (2016) have noted that results obtained with this series are highly sensitive to the sample, the variables and lags used, and to whether the recursiveness assumption is maintained or not.

Using Section 8 as a guide, we first observe that the narrative IV ($z_{N,t}$) is autocorrelated up to lag 12; several coefficients are individually significant, and the joint null is strongly rejected (F-stat: 2.56, p-val 0.003). We then use six factors extracted from the FRED-MD dataset of McCracken and Ng (2015) in a Granger causality test, as follows,

$$z_{N,t} = \theta_0 + \theta_1 z_{N,t-1} + \sum_{j=1}^{6} \theta_{f_j} f_{j,t-1} + \nu_t. \tag{31}$$

While the joint null is not rejected, some factors are individually significant at the 5% level. These tests suggest that $z_{N,t}$ may correlate with past shocks.

For the empirical setup, we rely on the main VAR specifications discussed in the literature. The benchmark model is Coibion (2012)'s VAR, that includes the log of industrial production, the unemployment rate, the log of the consumer price index, of a commodity price index, and the FFR.[21] The VAR is estimated over the original 1969–1996 sample with 12 lags.

In our reference setup, identification is obtained with $z_{N,t}$ as external instrument in the Coibion (2012) VAR. We then consider a number of variations to this benchmark, and explore the effects of changes to the VAR information set. First, the case in which the IV is the residual of $z_{N,t}$ on its first 12 lags. This to assess the role of the potential interaction between the violation of the lag-exogeneity condition and the non invertibility of the system.[22] Second, the case in which the IV is used as an internal instrument in a Hybrid VAR, achieved by adding $z_{N,t}$ to the VAR, and ordering it first.[23] This is helpful in assessing whether a violation of the partial-invertibility condition may be of consequence, after controlling for the contamination of the IV by past-shocks via the VAR lags. Finally, the case of a Hybrid VAR that imposes the recursiveness assumption, and in which the IV is ordered after the output and price variables.[24] Figure 3 collects the impact and dynamic IRFs across cases, all normalised to increase the FFR by 1% on impact.

A few elements stand out. First, when the recursiveness assumption is relaxed – i.e. when the current realisations of the real and price variables are not in the core information set –, the estimated impact effects are statistically non-zero. The estimated impacts are remarkably stable across cases, but at odds with what predicted by economic theory for both output and prices. This can be seen as a contamination by other contemporaneous shocks, and hence as a violation of the contemporaneous exogeneity condition. Second, the dynamic responses obtained when using the IV as an internal or external instrument are markedly different. The inclusion of $z_{N,t}$ in the VAR delivers much larger peak effects. As noted in Section 4, this points to a violation of the invertibility condition for some of the shocks captured by the IV.[25]

Taken together, these results should alert the reader that the Romer and Romer (2004)'s narrative measure is likely to violate the conditions for identification in standard VARs. Its use as an external IV is also complicated by the difficulty of correctly specifying a core information set that makes it conditionally exogenous. In fact, the contamination by contemporaneous and possibly non-fundamental shocks forces the researcher to control for the contemporaneous realisation of several variables, and hence impose arbitrary recursiveness assumptions. The information content of this measure of monetary policy innovations is an open problem. A conjecture is that the source of contamination may be deviations from the estimated

[21] This is the same system used in Ramey (2016). Data for bond yields, industrial production, and the consumer price index are from the St Louis FRED Database, the commodity price index is from the Commodity Research Bureau.

[22] The test of Forni and Gambetti (2014) indicates that the benchmark system may not be invertible. The FFR residuals are predictable at 5% (F-stat 2.067, p.val 0.047).

[23] Alternative Hybrid VAR specifications in the literature substitute in the VAR the FFR with the cumulative $z_{N,t}$. Results under this alternative are equivalent to those discussed here. See Online Supplementary Material.

[24] In the Hybrid VAR that includes $z_{N,t}$, the FFR innovations are not predictable (F-stat 1.473, p.val 0.176).

[25] If the IV correlates with shocks that are non-invertible in the VAR, the impact effects are still identified when it is used as an external instrument. However, the SVAR-IV would not correctly capture the dynamics due to the presence of the Blashke factor $b(L)$. A detailed discussion is in the proof of Proposition 3, in the Online Supplementary Material.

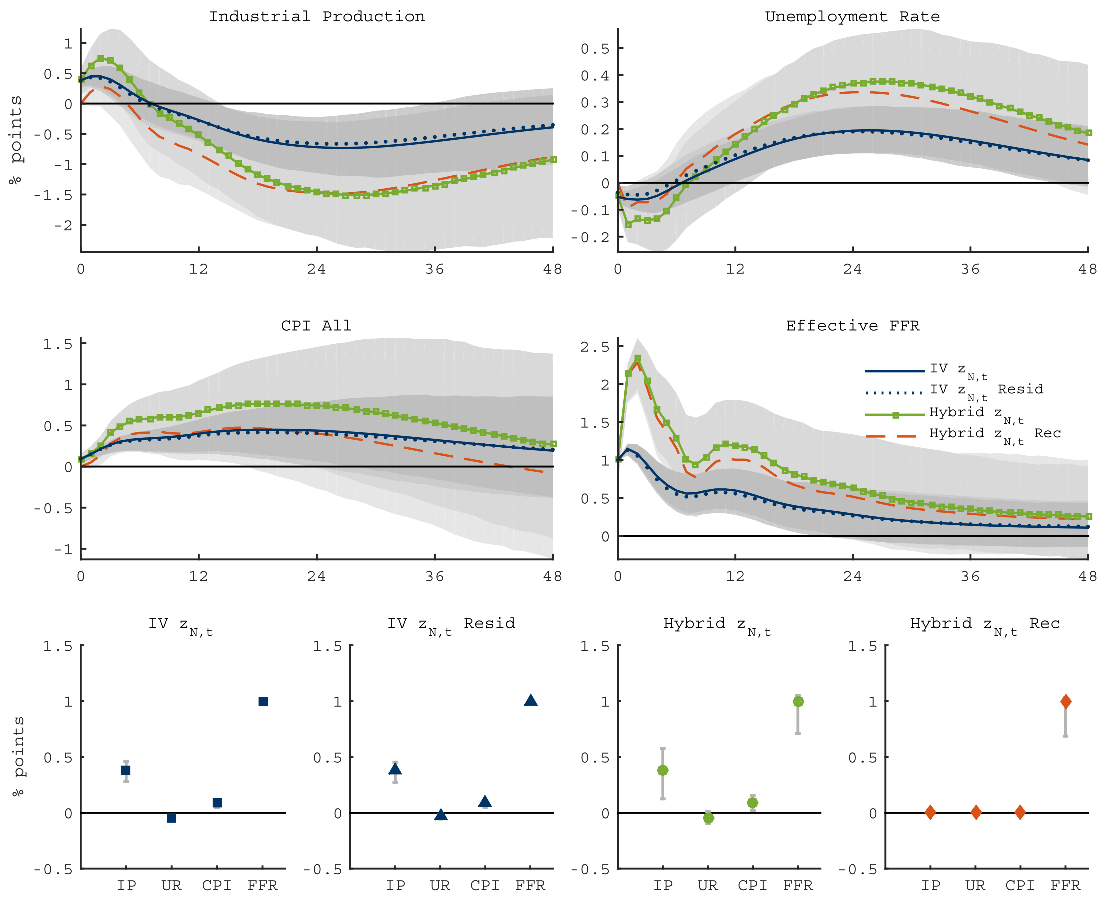


Fig. 3. Responses to Conventional MP Shocks: Narrative Instrument *Notes:* Blue solid line, SVAR-IV with $z_{N,t}$. Blue dotted line, SVAR-IV with $z_{N,t}$ regressed on its first 12 lags. Green line with markers, Hybrid VAR that adds $z_{N,t}$ as first variable in the VAR. Orange dashed line, Hybrid VAR that adds $z_{N,t}$ as last variable in the VAR. Sample 1969–1996, VAR(12). Shaded areas denote 90% posterior coverage bands. Bottom subplots highlight the impact IRFs.

Fed's reaction function that are due to shocks not captured by the selection of Greenbook forecasts that the IV controls for, but are likely anticipated by market participants.

9.2. Conventional monetary policy: High-Frequency instruments

We now examine high-frequency instruments, arguably the state of the art in the identification of monetary policy shocks at the short end of the yield curve. Price and output puzzles, and sample instabilities, are often obtained when using these IVs (Ramey, 2016). These are typically explained as the result of model misspecification, for example due to the omission of financial variables, as conjectured by Caldara and Herbst (2019). Alternatively, as due to contamination of the IVs by other shocks, owed to information effects, as pointed out by a growing literature.[26]

We consider three versions of this proxy, all a variant of the monetary surprises of Gürkaynak et al. (2005). The first IV is constructed by measuring high-frequency surprises in the fourth federal funds futures (FF4) around all the FOMC announcements between 1990 and 2012. This is similar to the instrument used in Stock and Watson (2018) and Caldara and Herbst (2019), and we denote it by $z_{A,t}$. The second instrument is a monthly moving average of high-frequency FF4 surprises around all FOMC announcements from 1990 to 2012. This was originally proposed in Gertler and Karadi (2015), denoted by $z_{B,t}$. The third IV is the residual of a projection of high-frequency FF4 surprises around all FOMC announcements on their lags and on Fed Greenbook forecasts from 1990 to 2009. This is the instrument in Miranda-Agrippino and Ricco (2021), denoted by $z_{C,t}$. The projection is intended to control for non-policy and past shocks that may affect market surprises due to the information channel of monetary policy (see Melosi, 2017) and for autocorrelation in expectation revisions due to imperfect information (see Coibion and Gorodnichenko, 2012; 2015).

[26] Among others, Cieslak (2018), Nakamura and Steinsson (2018), Cieslak and Schrimpf (2019), Jarociński and Karadi (2020), Lunsford (2020), Miranda-Agrippino and Ricco (2021), Karnaukh (2020), Bauer and Swanson (2020), and Sastry (2021).

Table 2
Test of Invertibility/Information Sufficiency.

	IP	UNRATE	CRBPI	CPI	EBP	1YR
Baseline VAR						
VAR(12), n=6						
	2.642	1.642	0.597	0.800	0.358	0.270
	(0.004)	(0.093)	(0.817)	(0.629)	(0.964)	(0.987)
Coibion VAR						
VAR(12), n=5						
	3.615	1.687	0.408	1.232		0.476
	(0.000)	(0.082)	(0.943)	(0.269)		(0.905)
Coibion VAR						
VAR(2), n=5						
	6.290	8.203	0.780	1.675		0.956
	(0.000)	(0.000)	(0.648)	(0.085)		(0.482)

Note: F-statistic for the joint null that the lagged state variables do not granger cause the VAR residuals. IP: Industrial Production; UNRATE: Unemployment Rate; CRBPI: Commodity Price Index; CPI: Consumer Price Index; EBP: Excess Bond Premium; 1YR: 1-Year Interest Rate. Top panel: Baseline VAR(12), n=6; Middle panel: Coibion VAR(12), n=5; Bottom panel: Coibion VAR(2), n=5. All VARs are estimated over the sample 1979:1-2012:12. Regressions include a constant and 10 lagged macro-financial factors. Robust standard errors. p-values in parentheses.

Table 3
Test for Conditional Exogeneity of the High-Frequency IV: Short-Horizon MP.

	$z_{A,t}$		$z_{B,t}$		$z_{C,t}$	
	(1)	(2)	(1)	(2)	(1)	(2)
output	8.01	6.45	4.94	7.09	1.11	0.05
	(0.005)	(0.011)	(0.026)	(0.008)	(0.293)	(0.824)
prices	0.31	4.80	1.01	7.21	0.87	0.08
	(0.579)	(0.029)	(0.315)	(0.007)	(0.351)	(0.775)
N	270	270	270	270	228	228

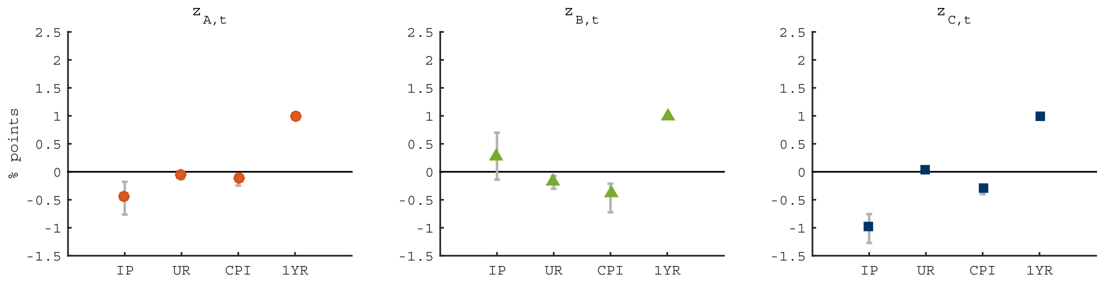
Note: Test $\sim \chi_{(1)}$, p-values in parentheses. Simple model: $y_t = bz_t + \gamma w_{1,t-1} + \nu_t$, richer model: $y_t = bz_t + \kappa w_{2,t-1} + \nu_t$, where y_t is either output or prices, $w_{1,t}$ includes IP, UNRATE, CPI, CRBPI, GS1, EBP, and $w_{2,t}$ excludes the EBP. (1): the instrument is included as one of the endogenous variables. (2): residuals are estimated in a first stage, and the regression on the instrument run in a second stage. $z_{A,t}$: sum of high-frequency surprises within the month; $z_{B,t}$: moving average of high-frequency surprises within the month; $z_{C,t}$: residuals of $z_{A,t}$ on Fed Greenbook forecasts.

For each instrument, we run a Granger causality test as in Eq. 31, again using factors extracted from the FRED-MD dataset. The Wald test statistics associated to the joint null are equal to 2.31 (p-value 0.013) for $z_{A,t}$, 3.52 (p-value 0.0002) for $z_{B,t}$, and 1.77 (p-value 0.067) for $z_{C,t}$ over the common sample 1990-1:2009-12. This suggests possible contamination by past shocks for the first two instruments.

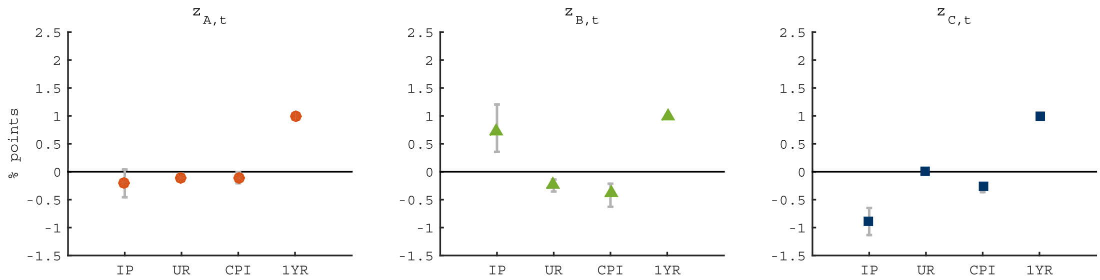
We then consider three monthly VARs estimated from 1979 to 1 to 2012-12. A VAR(12) that includes the log of industrial production, of the consumer price index, and of a commodity price index, the unemployment rate, the one-year rate as the policy variable, and the excess bond premium (EBP) of Gilchrist and Zakrajšek (2012). A second VAR(12) that excludes the EBP; this is equivalent to the Coibion (2012) VAR but with the one year rate in lieu of the FFR. And a third VAR that has composition identical to the latter but only includes 2 lags. The information set of this VAR is equivalent to that of the LP in Ramey (2016). This VAR is likely to understate the lag order, but it is compatible with a central bank that sets the interest rate using a simple Taylor rule.

Table 2 reports the results of the test of Forni and Gambetti (2014). Full invertibility is rejected. However, in all cases we do not find evidence that past information Granger causes the residuals of the policy-rate equation over this sample. This suggests partial invertibility of conventional monetary policy shocks in these VARs.

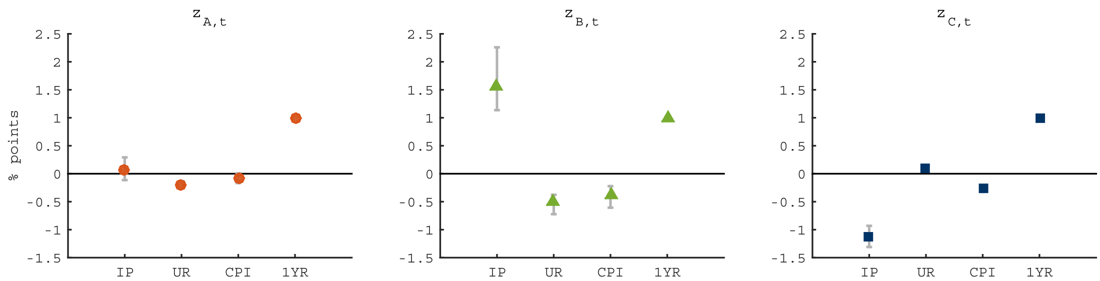
Table 3 reports two specifications for the test of conditional exogeneity for the impact responses of output and prices. One in which the instrument is included among the endogenous variables as in a Hybrid VAR, and one in which the test is run on first-stage residuals, as in the SVAR-IVs presented here. This latter specification does not take into account parameters uncertainty and hence tends to over-reject the null. Conditional exogeneity is rejected for $z_{A,t}$ and $z_{B,t}$ conditional on the information set in the 5-variable VAR. Conversely, we cannot reject that the same VAR is a correctly specified core information set for $z_{C,t}$. This suggests that impact effects under $z_{C,t}$ are likely to be stable across the models once these core variables are included.



(A) Baseline VAR(12)



(B) Coibion VAR(12)



(C) Coibion VAR(2)

Fig. 4. Impact Responses to Conventional MP Shocks – 1979:2012 *Notes:* Baseline VAR(12) in all variables, top panel (A). Coibion VAR(12), middle panel (B). Coibion VAR(2), bottom panel (C). VARs estimated with standard macroeconomic priors. Identification in all cases uses the full length of the instruments. $z_{A,t}$: high-frequency surprises at scheduled FOMC meetings; $z_{B,t}$: moving average of high-frequency surprises within the month; $z_{C,t}$: residuals of $z_{A,t}$ on Fed Greenbook forecasts. Shaded areas denote 90% posterior coverage bands.

This is indeed the case, as visible from Figure 4, where all responses are normalised to a 1% impact increase in the policy rate. The top row of the figure collects results for the larger VAR, while the VAR(12) and VAR(2) that exclude the EBP are in the middle and bottom rows respectively. Consistent with the outcome of the conditional exogeneity test, impact responses recovered under either $z_{A,t}$ or $z_{B,t}$ vary across models, and are statistically different. Median impact responses of output to a contractionary monetary policy shock of the same size go from -0.5% to essentially zero under $z_{A,t}$, and from being non-significant to strongly positive at around 1% under $z_{B,t}$.

Figure 5 plots the full IRFs identified by the three IVs across all VARs. Given stable impacts, $z_{C,t}$ also recovers similar dynamic IRFs across models. Increasing the number of lags marginally stabilises the IRFs. This suggests that misspecification due to omitted variables or lags is minor. It follows that the large variability in the dynamic IRFs under $z_{A,t}$ or $z_{B,t}$ is predominantly due to the unstable impacts. That is, to contamination of the IVs.[27]

Taking stock, we can interpret these results as follows. Given partial invertibility of the monetary policy shocks in the VARs considered, $z_{C,t}$ recovers both stable impacts, and stable dynamic IRFs across models. Combined with the Granger causality and lagged-exogeneity tests, this suggests that the IV satisfies the conditions for identification laid out in this

[27] Results in this section are robust to the sample specification and to using a version of $z_{A,t}$ that is measured only around scheduled FOMC announcements. See Online Supplementary Material.

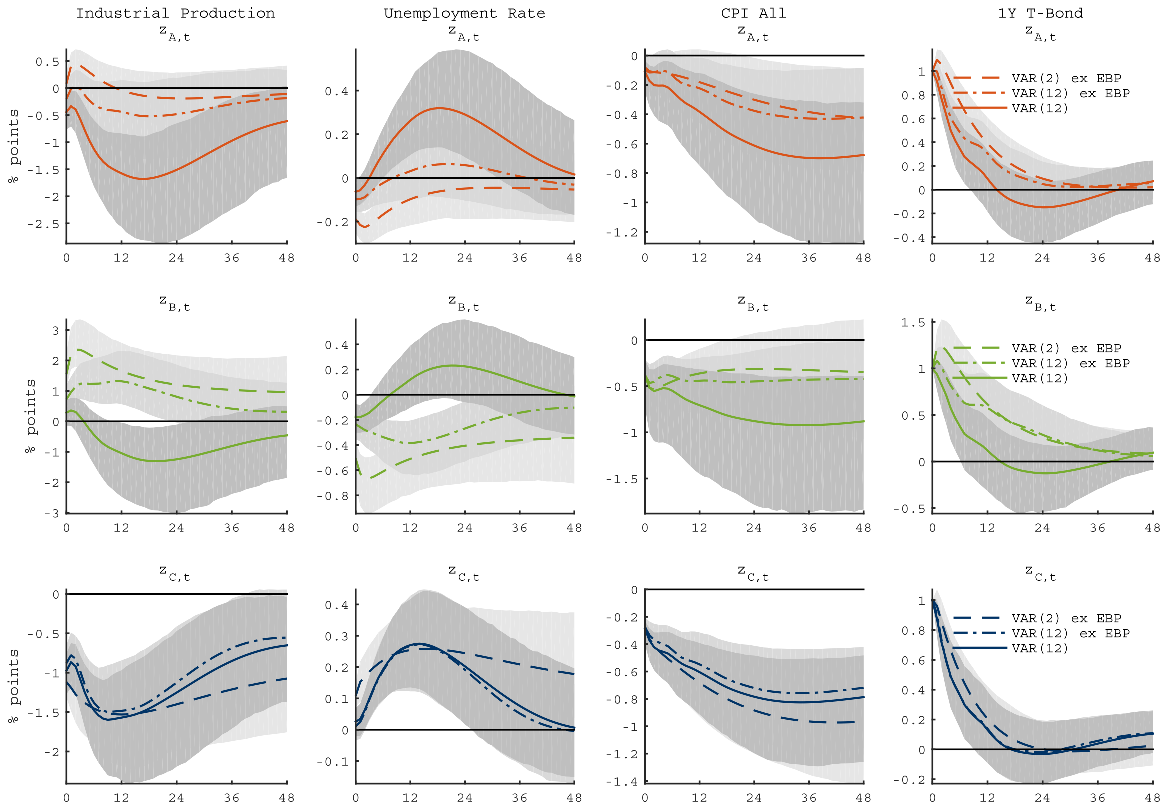


Fig. 5. Responses to Conventional MP Shocks – 1979:2012 *Notes:* SVAR-IVs. Top panels: $z_{A,t}$: sum of high-frequency surprises within the month; Middle panels: $z_{B,t}$: moving average of high-frequency surprises within the month; Bottom panels: $z_{C,t}$: residuals of $z_{A,t}$ on Fed Greenbook forecasts. In each subplot, the solid line is for the baseline VAR(12), the dash-dotted line for the Coibion VAR(12) and the dashed line for the Coibion VAR(2). Identification in all cases uses the full length of the instruments. Shaded areas denote 90% posterior coverage bands.

paper. Moreover, the sign of the responses is compatible with theoretical priors on the effects of monetary policy on output and prices.

The same does not hold true for $z_{A,t}$ and $z_{B,t}$, for which the inclusion of the EBP in the VAR alters the IRFs materially, while at the same time reducing the extent of the puzzles. This finding had led Caldara and Herbst (2019) to attribute the instability to model misspecification. Our results suggest a different interpretation instead. They indicate that both IVs correlate with other lagged and possibly contemporaneous shocks that are non-invertible in VARs without financial variables. The inclusion of the EBP helps obtaining IRFs that are in line with theoretical priors by absorbing and making invertible in the larger system shocks that contaminate the IVs, e.g. due to information effects. In turn, this reduces the bias, as discussed in Remark 2 and in Section 5. In other words, the EBP helps approximating the core information set that makes these IVs conditionally exogenous.

These results provide a justification to the intuition in Barakchian and Crowe (2013) and Gertler and Karadi (2015), that the inclusion of forward looking variables in otherwise standard VARs can stabilise the IRFs and avoid large puzzles. Our results also offer an alternative interpretation to the findings in Miranda-Agrippino and Ricco (2021). Seen through these lens, central bank forecasts are the controls needed to obtain the core information set conditional on which high-frequency instruments are exogenous. Hence, an alternative to the approach of Miranda-Agrippino and Ricco (2021) is to include the Greenbook in the VAR. Abstracting from estimation uncertainty and risks of overfitting, the two procedures should deliver identical results.

It is interesting to note that the Greenbook alone may not be sufficient to control for the endogenous/anticipated component of monetary policy. This was evident in the case of the narrative IV discussed earlier. Yet, they do seem to be effective in controlling for residual information effects when market surprises are used to isolate the unanticipated component of policy announcements, i.e. conditional on market participants' information sets.

9.3. Unconventional monetary policy: High-Frequency instruments

High-frequency instruments for forward guidance and large-scale asset purchases (LSAP) have been proposed by Swanson (2020). These instruments summarise monetary surprises at the medium and long end of the yield curve. Specifically, for-

Table 4
Test for Conditional Exogeneity of the High-Frequency IV: Longer-Horizon MP.

	$z_{FWG,t}$		$z_{LSAP,t}$	
	(1)	(2)	(1)	(2)
output	2.77	3.56	7.55	21.54
	(0.096)	(0.059)	(0.006)	(0.000)
prices	0.78	0.57	4.77	21.57
	(0.378)	(0.449)	(0.029)	(0.000)
N	134	134	134	134

Note: Test $\sim \chi_{(1)}$, p-values in parentheses. Simple model: $y_t = bz_t + \sum_{\ell=1}^{2} \gamma_\ell w_{1,t-\ell} + \nu_t$, richer model: $y_t = bz_t + \sum_{\ell=1}^{2} \gamma_\ell w_{2,t-\ell} + \nu_t$. y_t is either output or prices, $w_{1,t}$ includes IP, CPI, GS2, GS10, $w_{2,t}$ includes IP, UNRATE, CPI, PPIACO, GS2, GS10, EBP. (1): the instrument is included as one of the endogenous variables. (2): residuals are estimated in a first stage, and the regression on the instrument run in a second stage. $z_{FWG,t}$: Forward Guidance component of FOMC announcements; $z_{LSAP,t}$: LSAP component of FOMC announcements. Sample: 2008-1:2019-4.

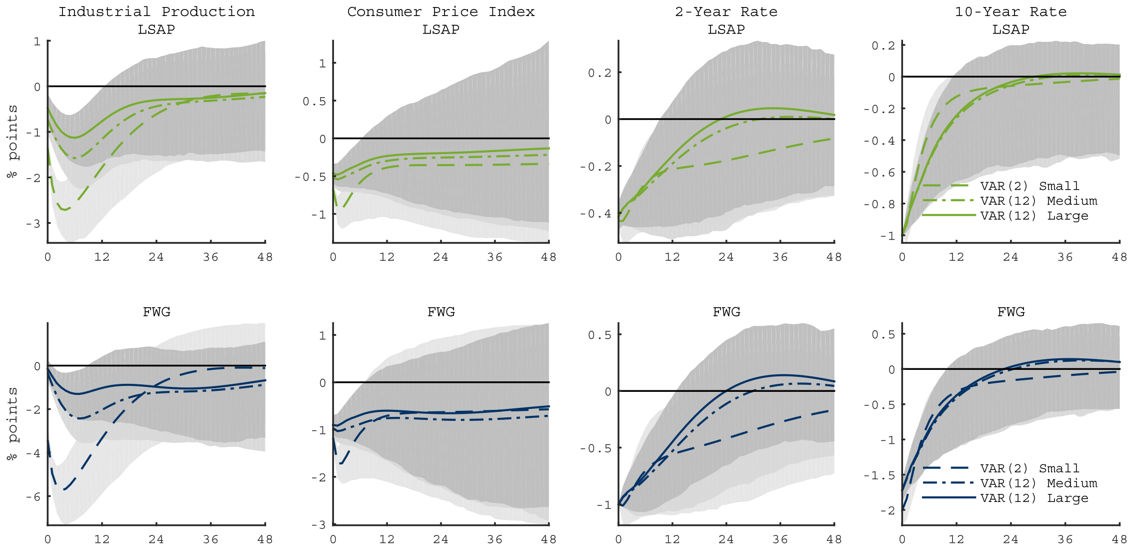


Fig. 6. Responses to Unconventional MP Shocks – 2008:2019 *Notes:* Top row: $z_{LSAP,t}$: high-frequency LSAP factor; Bottom row: $z_{FWG,t}$: high-frequency forward guidance factor. In each subplot: baseline VAR(12) solid lines, alternative VAR(12) dash-dotted line, small VAR(2) dashed line. Identification in all cases uses the full length of the instruments. Sample 2008–2019. Shaded areas denote 90% posterior coverage bands.

ward guidance surprises ($z_{FWG,t}$) have higher loadings on 1 to 2-year maturity rates, while the LSAP ones ($z_{LSAP,t}$) mostly captures variation in longer term (10-year) rates.

As in the previous section, we evaluate the properties of these IV with the aid of three VARs. A baseline VAR(12) with the log of IP, of the CPI index, and of a commodity price index, the unemployment rate, the two-year and ten-year government bond rates, and the EBP. A second VAR(12) that excludes the EBP. And a VAR(2) that also excludes the unemployment rate and the commodity price index. To focus on unconventional monetary policy, we estimate the VARs on the sample 2008-1 to 2019-12. The IRFs identified with $z_{FWG,t}$ and $z_{LSAP,t}$ are normalised to a 1% impact decrease in the 2-year and 10-rate respectively. Hence, they can be interpreted as the effects of an easing of the monetary policy stance through either tool.

The information sufficiency test indicates that full invertibility is not rejected in the larger VARs, but there is evidence of predictability of the 2-year rate residuals in the smaller VAR (results are available upon request). Table 4 reports the test for conditional exogeneity for the two IVs, calculated using a specification that includes all the variables in the larger VAR, with one that only includes the controls in the smallest one. Similar to what noted in the previous section, conditional exogeneity is rejected for both high-frequency instruments and, consistently, this translates into model-dependent impact responses.[28] What is arguably more relevant in this case is that, however unstable, the impact responses consistently deliver severe output and price puzzles across all VARs. This is visible in Figure 6, where IRFs identified with $z_{LSAP,t}$ and $z_{FWG,t}$ are

[28] See Online Supplementary Material. Using a specification similar to Eq. (31) with seven factors, the Wald test statistics associated to the joint null of no forecastability are equal to 0.59 (p-value 0.7603) and 1.84 (p-value 0.0843) for $z_{FWG,t}$ and $z_{LSAP,t}$ respectively.

in the top and bottom panels respectively. Both IVs identify shocks that lead to the 10-year rate falling by more than the 2-year, and a contraction in output and prices.

We can interpret these results as follows. The conditional exogeneity test indicates that the small VAR is not a valid core information set for either IV. Hence, the large change in the impact effects from the small to the larger VARs suggests that the IVs correlate with other past shocks. However, while the size of the impact varies, their sign does not. The two shocks lead to impact inversions of the yield curve typically associated with business cycle recessions, and at odds with the theoretical priors on the effects of monetary policy. Similar to the narrative IV, this suggest correlation with other contemporaneous shocks. For $z_{FWG,t}$, it can be reconciled with the notion of Delphic guidance, i.e. with a public signal on the economic outlook. But, in general, these results suggest that absent controls for confounding factors such as e.g. information effects and risk premia shocks, these are not valid IVs for unconventional policy shocks even in VARs that include forward-looking variables and many lags. More research on their information content is needed for them to be used as external IVs in VAR or LP.[29]

10. Conclusions

Correct identification of Structural VARs with external instruments requires that only the shock of interest is invertible in the VAR of choice, and that the IV satisfies a limited lead-lag exogeneity condition in addition to the standard relevance and contemporaneous exogeneity conditions. Our results broaden the scope of IV methods in SVARs, and relieve the empirical researcher from discussing the invertibility of all the shocks in the systems that are not of direct interest to the study. They also provide a formal characterisation of the bias that arises when the conditions for identification are not met, and introduce a testing procedure to assess the lag exogeneity of the IV of interest.

Data availability

Data will be made available on request.

Acknowledgements

We thank Guido Ascari, Matteo Barigozzi, Fabio Canova, Marco Del Negro, Luca Gambetti, Domenico Giannone, Diego Kaenzig, Peter Karadi, Michael McMahon, Mikkel Plagborg-Møller, Giuseppe Ragusa, Valerie Ramey, Lucrezia Reichlin, Juan Rubio-Ramírez, James Stock, Paolo Surico, Christian Wolf, Francesco Zanetti, and an anonymous referee for helpful discussions and suggestions. We also thank participants at the Nowcasting Meeting 2018, the 2019 IAAE, the WBS Workshop on ‘Advances in Empirical Macroeconomics’, Oxford University, and the 2019 Meeting of the RES for comments. Silvia Miranda-Agrippino gratefully acknowledges hospitality and support from Northwestern University. The authors acknowledge support from the British Academy: Leverhulme Small Research Grant SG170723. The views expressed in this paper are those of the authors and do not represent those of the Bank of England or any of its Committees.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jmoneco.2023.01.006

References

Arias, J.E., Rubio-Ramirez, J.F., Waggoner, D.F., 2018. Inference in Bayesian Proxy-SVARs. Working Papers. Federal Reserve Bank of Philadelphia. 10.21799/
Barakchian, S.M., Crowe, C., 2013. Monetary policy matters: evidence from new shocks data. J. Monet. Econ. 60 (8), 950–966. doi:10.1016/j.jmoneco.2013.09.
Bauer, M.D., Swanson, E.T., 2020. The Fed’s Response to Economic News Explains the ”Fed Information Effect”. Working Paper Series. National Bureau of Economic Research.
Braun, P.A., Mittnik, S., 1993. Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions. J. Econom. 59 (3), 319–341.
Caldara, D., Herbst, E., 2019. Monetary policy, real activity, and credit spreads: evidence from bayesian proxy svars. Am. Econ. J. 11 (1), 157–192. doi:10.1257/mac.20170294.
Canova, F., Ferroni, F., 2019. Mind the gap! stylized dynamic facts and structural models.Mimeo.
Canova, F., Hamidi Sahneh, M., 2017. Are small-scale svars useful for business cycle analysis? revisiting nonfundamentalness. J. Eur. Econ. Assoc. 16 (4), 1069–1093.
Chalak, K., White, H., 2011. Viewpoint: an extended class of instrumental variables for the estimation of causal effects. Canad. J. Economics/Revue canadienne d’économique 44 (1), 1–51. doi:10.1111/j.1540-5982.2010.01622.x.
Cieslak, A., 2018. Short-Rate expectations and unexpected returns in treasury bonds. Rev. Financ. Stud. 31 (9), 3265–3306. doi:10.1093/rfs/hhy051.
Cieslak, A., Schrimpf, A., 2019. Non-monetary news in central bank communication. J. Int. Econ. 118 (C), 293–315. doi:10.1016/j.jinteco.2019.01.
Coibion, O., 2012. Are the effects of monetary policy shocks big or small? Am. Econ. J. 4 (2), 1–32.
Coibion, O., Gorodnichenko, Y., 2012. What can survey forecasts tell us about information rigidities? J. Polit. Econ. 120 (1), 116–159. doi:10.1086/665662.

[29] Swanson (2020) provides compelling evidence on the effectiveness of FWG and LSAP on asset prices using $z_{FWG,t}$ and $z_{LSAP,t}$ in event studies. However, conditions for identification in VARs are more stringent, as also noted in Gürkaynak et al. (2020).

Coibion, O., Gorodnichenko, Y., 2015. Information rigidity and the expectations formation process: a simple framework and new facts. Am. Econ. Rev. 105 (8), 2644–2678. doi:10.1257/aer.20110306.
Fernandez-Villaverde, J., Rubio-Ramirez, J.F., Sargent, T.J., Watson, M.W., 2007. ABCs (And ds) of understanding VARs. Am. Econ. Rev. 97 (3), 1021–1026.
Forni, M., Gambetti, L., 2014. Sufficient information in structural VARs. J. Monet. Econ. 66 (C), 124–136. doi:10.1016/j.jmoneco.2014.04.
Forni, M., Gambetti, L., Sala, L., 2019. Structural vars and noninvertible macroeconomic models. J. Appl. Econ. 34 (2), 221–246. doi:10.1002/jae.2665.
Gertler, M., Karadi, P., 2015. Monetary policy surprises, credit costs, and economic activity. Am. Econ. J. 7 (1), 44–76. doi:10.1257/mac.20130329.
Giannone, D., Reichlin, L., 2006. Does information help recovering structural shocks from past observations? J. Eur. Econ. Assoc. 4 (2–3), 455–465.
Gilchrist, S., Zakrajšek, E., 2012. Credit spreads and business cycle fluctuations. Am. Econ. Rev. 102 (4), 1692–1720. doi:10.1257/aer.102.4.1692.
Gürkaynak, R.S., Kara, A.H., Kısacıkoğlu, B., Lee, S.S., 2020. Monetary Policy Surprises and Exchange Rate Behavior. CEPR Discussion Papers. C.E.P.R. Discussion Papers.
Gürkaynak, R.S., Sack, B., Swanson, E., 2005. Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. Int. J. Central Bank. 1 (1).
Heckman, J., Navarro-Lozano, S., 2004. Using matching, instrumental variables, and control functions to estimate economic choice models. Rev. Econ. Stat. 86 (1), 30–57. doi:10.1162/003465304323023660.
Jarociński, M., Karadi, P., 2020. Deconstructing monetary policy surprises– The role of information shocks. Am. Econ. J. 12 (2), 1–43.
Jordà, Ò., 2005. Estimation and inference of impulse responses by local projections. Am. Econ. Rev. 95 (1), 161–182.
Jordà, Ò., Schularick, M., Taylor, A.M., 2015. Betting the house. J. Int. Econ. 96 (S1), 2–18. doi:10.1016/j.jinteco.2014.12.
Karnaukh, N., 2020. Growth Forecasts and News about Monetary Policy. Working Paper Series. Ohio State University, Charles A. Dice Center for Research in Financial Economics.
Leeper, E.M., Walker, T.B., Yang, S.-C. S., 2013. Fiscal foresight and information flows. Econometrica 81 (3), 1115–1145. ECTA8311
Lippi, M., Reichlin, L., 1993. The dynamic effects of aggregate demand and supply disturbances: comment. Am. Econ. Rev. 83 (3), 644–652.
Lippi, M., Reichlin, L., 1994. VAR Analysis, nonfundamental representations, blaschke matrices. J. Econom. 63 (1), 307–325.
Lu, X., White, H., 2014. Robustness checks and robustness tests in applied economics. J. Econom. 178, 194–206. doi:10.1016/j.jeconom.2013.08.016. Annals Issue: Misspecification Test Methods in Econometrics
Lunsford, K.G., 2020. Policy language and information effects in the early days of federal reserve forward guidance. Am. Econ. Rev. 110 (9), 2899–2934. doi:10.1257/aer.20181721.
McCracken, M.W., Ng, S., 2015. FRED-MD: A Monthly Database for Macroeconomic Research. Working Papers. Federal Reserve Bank of St. Louis.
Melosi, L., 2017. Signaling effects of monetary policy. Rev. Econ. Stud. forthcoming.
Mertens, K., 2014. Advances in Empirical Macroeconomics, Lecture 2. Bonn Summer School.
Mertens, K., Ravn, M.O., 2013. The dynamic effects of personal and corporate income tax changes in the united states. Am. Econ. Rev. 103 (4), 1212–1247. doi:10.1257/aer.103.4.1212.
Miranda-Agrippino, S., Ricco, G., 2021. The transmission of monetary policy shocks. Am. Econ. J. 13 (3), 74–107. doi:10.1257/mac.20180124.
Nakamura, E., Steinsson, J., 2018. High-Frequency identification of monetary non-Neutrality: the information effect*. Q. J. Econ. 133 (3), 1283–1330. doi:10.1093/qje/qjy004.
Plagborg-Møller, M., Wolf, C.K., 2018. Local Projections and VARs Estimate the Same Impulse Responses. Working Paper.
Ramey, V.A., 2016. Macroeconomic Shocks and Their Propagation. In: Taylor, J.B., Uhlig, H. (Eds.), Handbook of Macroeconomics. In: Handbook of Macroeconomics, Vol. 2. Elsevier, pp. 71–162. doi:10.1016/bs.hesmac.2016.03.003.
Ramey, V.A., Zubairy, S., 2018. Government spending multipliers in good times and in bad: evidence from us historical data. J. Polit. Econ. 126 (2), 850–901. doi:10.1086/696277.
Romer, C.D., Romer, D.H., 2004. A new measure of monetary shocks: derivation and implications. Am. Econ. Rev. 94 (4), 1055–1084.
Sastry, K., 2021. Disagreement about monetary policy. Available at SSRN https://doi.org/10.2139/ssrn.3421723.
Sims, C.A., Zha, T., 2006. Does monetary policy generate recessions? Macroecon. Dyn. 10 (02), 231–272.
Stock, J.H., 2008. What's New in Econometrics: Time Series, Lecture 7. NBER Summer Institute, Short course lectures.
Stock, J.H., Watson, M.W., 2012. Disentangling the channels of the 2007-09 recession. Brookings Pap. Econ. Act. 44 (1 (Spring), 81–156.
Stock, J.H., Watson, M.W., 2018. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. Econ. J. 128 (610), 917–948. doi:10.1111/ecoj.12593.
Swanson, E.T., 2020. Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. J. Monet. Econ. doi:10.1016/j.jmoneco.2020.09.003.
White, H., Chalak, K., 2010. Testing a conditional form of exogeneity. Econ. Lett. 109 (2), 88–90.