Journal of Econometrics 245 (2024) 105876

Contents lists available at ScienceDirect

Journal of Econometrics

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

Testing for strong exogeneity in Proxy-VARs☆

Martin Bruns[a,*], Sascha A. Keweloh[b]

[a] University of East Anglia, School of Economics, Norwich Research Park, NR4 7TJ, Norwich, United Kingdom
[b] TU Dortmund University, Vogelpothsweg 87, 44221 Dortmund, Germany

ARTICLE INFO

JEL classification:
C32

Keywords:
Structural vector autoregression
Proxy VAR
Exogeneity test

ABSTRACT

Proxy variables have gained widespread prominence as indispensable tools for identifying structural VAR models. Analogous to instrumental variables, proxies need to be exogenous, i.e. uncorrelated with all non-target shocks. Assessing the exogeneity of proxies has traditionally relied on economic arguments rather than statistical tests. We argue that the economic rationale underlying the construction of commonly used proxy variables aligns with a stronger form of exogeneity. Specifically, proxies are typically constructed as variables not containing any information on the expected value of non-target shocks. We show conditions under which this enhanced concept of proxy exogeneity is testable without additional identifying assumptions.

1. Introduction

Structural vector autoregressions (SVARs) are routinely combined with external instruments, so-called proxies, to achieve identification. Proxies are valid for identification under two conditions: They need to be relevant, i.e. contemporaneously correlated with the target shock; and they need to be exogenous, i.e. contemporaneously uncorrelated with the non-target shocks.

The traditional conception is that proxy exogeneity cannot be tested using only the proxy variable. Instead, the majority of applications using proxy variables rely on economic arguments to justify the exogeneity condition and not on a statistical test. More recently, several studies propose to test the proxy exogeneity assumption by identifying the model using other identifying assumptions, i.e. using heteroskedasticity in Schlaak et al. (2023), sign restrictions in Braun and Brüggemann (2022), independent and non-Gaussian shocks in Keweloh et al. (2023a,b), changes in unconditional volatility, which can be used to test for proxy exogeneity ex-post in Angelini et al. (2024), or breaks in the simultaneous interaction in Angelini et al. (2024).

In this study, we propose a proxy exogeneity test that exclusively leverages the information embedded in the proxy variable. We argue that the economic reasoning employed to construct proxy variables typically implies a stronger form of exogeneity. Specifically, proxies are typically constructed not merely to exhibit no correlation with non-target shocks but to contain no information at all on the expected value of the non-target shocks, which we denote as a strongly exogenous proxy. If we extend our notion of proxy exogeneity to this stronger exogeneity assumption, a proxy variable can contain information beyond its correlation with the reduced form shocks and this information can be used to detect endogenous proxy variables. Specifically, for a strongly exogenous proxy z_t, we can generate a synthetic proxy $\tilde{z}_t = z_t^2$ which is also exogenous. Therefore, we obtain two proxies and an overidentified system, so that a simple J-test can be used to test the strong exogeneity assumption.

The crux of our study lies in the expansion of the concept of an exogenous proxy beyond mere uncorrelation with non-target shocks to strong exogeneity, meaning the absence of any information on the expected value of non-target shocks. Consider, for

☆ The authors thank Luca Fanelli and Helmut Lütkepohl for helpful comments on an earlier draft of the paper. Martin Bruns thanks the British Academy, United Kingdom for financial support, grant number TDA22\220046.

* Corresponding author.
E-mail addresses: martin.bruns@uea.ac.uk (M. Bruns), sascha.keweloh@tu-dortmund.de (S.A. Keweloh).

https://doi.org/10.1016/j.jeconom.2024.105876
Received 12 December 2023; Received in revised form 11 October 2024; Accepted 13 October 2024
Available online 16 November 2024
0304-4076/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

example, the tax proxy in Mertens and Ravn (2014), constructed as a series of tax shocks based on narrative documents, the monetary policy proxy in Gertler and Karadi (2015) measuring monetary policy shocks based on federal funds futures at FOMC announcements, or the oil supply news proxy in Känzig (2021) capturing supply news shocks based on changes in oil price futures at OPEC announcements. In each case, the economic rationale underlying the construction of the proxy is that the proxy is a function of the target shock but not affected by non-target shocks, which of course justifies uncorrelatedness of the proxy and non-target shocks. However, we show that it also justifies strong exogeneity. If, for instance, the tax proxy is equal to a series of tax shocks and tax shocks contain no information on the expected value of other non-target shocks, then the tax proxy contains no information on the expected value of other non-target shocks and is thus strongly exogenous.

A Strongly exogenous proxy allows to generate additional synthetic proxy variables. Specifically, for a strongly exogenous proxy z_t, it holds that $\tilde{z}_t = h(z_t)$ is uncorrelated with non-target shocks and thus $\tilde{z}_t$ is an exogenous synthetic proxy. Intuitively, if the tax proxy is equal to a series of tax shocks and thus uncorrelated with non-target shocks, the synthetic tax proxy $\tilde{z}_t = z_t^2$ is equal to a series of squared tax shocks, and thus also uncorrelated with non-target shocks. Consequently, for a strongly exogenous proxy, we can use the original and a synthetic proxy to obtain a system of two exogenous proxies and jointly test if both proxies are uncorrelated with the non-target shocks. Rejecting the null hypothesis that both proxies are uncorrelated with the non-target shocks indicates that the proxy is not strongly exogenous.[1]

Having established the economic rationale justifying the exogeneity of the synthetic proxy, the remaining question is under which circumstances does a strongly exogenous proxy contain sufficient information to detect endogeneity? Put differently, what are the conditions to ensure that the synthetic proxy contains additional information not contained in the original proxy? A proxy variable is a function of potentially all structural shocks and additional exogenous variation. If this function is linear and all shocks are Gaussian, a strongly exogenous proxy contains no information beyond its correlation with the reduced form shocks, such that no synthetic proxy can contain any information to detect endogeneity. However, if the proxy does not adhere perfectly to a linear function or if not all shocks influencing the proxy adhere strictly to a Gaussian distribution, a strongly exogenous proxy can contain additional information, which can be leveraged to detect endogeneity of the proxy.

First, consider a proxy z_t equal to a linear function of the skewed target shock and a noise term. In this scenario, the information of the proxy is not entirely contained in its correlation with the reduced form shocks, instead, the synthetic proxy $\tilde{z}_t = z_t^2$ is relevant and can contain information to detect endogeneity. Intuitively, if tax shocks exhibit a left-skewed distribution, implying that large negative tax shocks are more likely than large positive ones, high values of the synthetic tax proxy $\tilde{z}_t = z_t^2$ correlate with negative values of the tax shock, implying that the synthetic tax proxy is correlated with the target shock and hence provides overidentifying restrictions. The same argument can be made if a non-target shock which is correlated with the proxy (a so-called "contaminating shock") exhibits skewness. Second, our approach is not limited to non-Gaussian shocks. Even with Gaussian shocks, the information of the proxy may not be entirely contained in the second moments if the proxy generating function is non-linear. Only in the special case of exactly Gaussian shocks and a perfectly linear proxy model a synthetic proxy cannot contain additional information such that our test has no power and rejects at the nominal level.

We broadly relate to SVAR identification approaches relying on information in higher moments of the shocks, which allow full identification of the SVAR if all shocks are independent and at most one shock is Gaussian, see Matteson and Tsay (2017), Gouriéroux et al. (2017), Keweloh (2021) and Guay (2021). In contrast, we rely on higher moments of the proxy and aim at testing exogeneity of the proxy, which only requires mean independence of the non-target shock and proxy. Moreover, a single (or for a non-linear proxy even no) non-Gaussian shock can be sufficient to detect proxy endogeneity. Additionally, estimates based on higher moments are of course not limited to SVARs. For example, Lewbel (1997) and Erickson and Whited (2002) construct instruments as functions of the data to estimate linear regression models with measurement errors in the variables and Bierens (1982) or Donald et al. (2003) consider estimators for conditional moment restriction models relying on unconditional moment restrictions also involving functions of the data.

We investigate the finite sample properties of the proposed test in a Monte Carlo study using a stylized and a more realistic data-generating process (DGP). We find that our test has a precise nominal level in settings usually encountered in macroeconomic datasets. Its power increases with sample size, with stronger proxies, higher skewness of the target and/ or contaminating shock, stronger forms of proxy non-linearity, and a higher correlation of the proxy with the contaminating shock.

We apply the proposed exogeneity test to three proxies frequently used in fiscal proxy SVARs. These proxies include the narrative tax variable used in Mertens and Ravn (2014) as a proxy for tax shocks, the total factor productivity measure of Fernald (2012) as an output shock proxy, and changes in military spending, which have been used as a government expenditure shock proxy, as seen in Klein and Linnemann (2019), for instance. Our findings reveal a lack of strong exogeneity in the tax proxy, while providing no evidence against exogeneity for the output and spending proxies.

The remainder of this paper is organized as follows: Section 2 discusses the model assumption and develops the proxy exogeneity tests. Section 3 shows the Monte Carlo simulation. Section 4 applies the test in the fiscal SVAR. Section 5 concludes.

[1] Note that in general, rejecting strong exogeneity does not necessarily imply that the proxy is correlated with non-target shocks. However, we show that if the proxy is equal to a function of the target shock and a noise term, both mean independent w.r.t. non-target shocks, rejecting strong proxy exogeneity implies that the proxy is also a function of the non-target shocks. While it is mathematically possible to construct a proxy equal to a function of the target and non-target shocks, yet uncorrelated with the non-target shock, this case is clearly not economically interesting.

2. Model setup

Our model is a K-dimensional SVAR(p) process,

$$y_t = v + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t \quad \text{with} \quad u_t = B_0 \varepsilon_t \tag{1}$$

and a $K \times 1$ vector y_t of endogenous variables, reduced form shocks u_t, and structural shocks ε_t. We normalize the diagonal elements of B_0 to one and w.l.o.g. assume that the first shock, ε_{1t}, is the target shock and $\boldsymbol{\varepsilon}_{2t} = (\varepsilon_{2t}, \dots, \varepsilon_{Kt})'$ contains the non-target shocks. Analogously, we define $\boldsymbol{u}_{2t} = (u_{2t}, \dots, u_{Kt})'$. Moreover, let β_0 be equal to the last $(K-1)$ elements of the first column of B_0, such that β_0 denotes the simultaneous impact of the target shock on the last $(K-1)$ variables and the impact on the first variable is normalized to one.

The reduced form shocks can be estimated via OLS and the simultaneous impact $u_t = B_0 \varepsilon_t$ of the target shock can be estimated using a proxy. For simplicity, we omit the lag structure and focus on the simultaneous interaction. An extension to a VAR(p) with $p > 0$ is straightforward and can be found in Appendix A.

2.1. Proxy SVAR

Typically, the underlying rationale of a proxy is to construct a variable z_t that contains information about the target shock, ε_{1t}, but no information on the non-target shocks, $\boldsymbol{\varepsilon}_{2t}$. These properties are imposed using second-order moments, i.e. a proxy is valid if it is correlated with the target shock and uncorrelated with all non-target shocks.

Assumption 1 (*Valid Proxy*).
The proxy z_t for the shock ε_{1t} is relevant and exogenous:

1. Relevance: $E\left[\varepsilon_{1t} z_t\right] \neq 0$ 2. Exogeneity: $E\left[\boldsymbol{\varepsilon}_{2t} z_t\right] = \mathbf{0}$

Using exogeneity and relevance allows to identify the impact of the target shock with $\beta_0 = \frac{\mathbb{E}(\boldsymbol{u}_{2t} z_t)}{\mathbb{E}(u_{1t} z_t)}$. However, with only a single proxy variable, we cannot statistically test the identifying uncorrelatedness assumption $E\left[\boldsymbol{\varepsilon}_{2t} z_t\right] = 0$. Therefore, it typically remains up to the researcher to find convincing economic arguments why the proxy should not contain information on the non-target shocks.

Assumption 2 (*Strong Exogeneity*).
The proxy z_t for the shock ε_{1t} is strongly exogenous if $E\left[\boldsymbol{\varepsilon}_{2t} | z_t\right] = 0$.

Strong exogeneity is a stronger assumption than uncorrelatedness of the proxy with all non-target shocks, and we propose a test for the strong exogeneity assumption.[2] The question is whether the strong exogeneity assumption is too strong. Meaning, are there proxy variables that are uncorrelated with the non-target shocks but do not meet the criteria for strong exogeneity? While it is statistically possible to define such a variable, we are not aware of any reasonable economic examples where proxy variables are arguably uncorrelated with the non-target shocks yet still contain predictive power on the expected value of the non-target shocks.

The following proposition shows that strong exogeneity of a proxy follows if a proxy is equal to an arbitrary function of the target shock ε_{1t} and a noise term v_t.

Proposition 1. *For $E[\boldsymbol{\varepsilon}_{2t} | v_t, \varepsilon_{1t}] = 0$ and $z_t = g(\varepsilon_{1t}, v_t)$ and a measurable function $g(.)$ it follows that z_t is strongly exogenous, i.e. $E\left[\boldsymbol{\varepsilon}_{2t} | z_t\right] = 0$.*

Proof. The law of iterated expectations implies

$$E[\boldsymbol{\varepsilon}_{2t} | g(v_t, \varepsilon_{1t})] = E[E[\boldsymbol{\varepsilon}_{2t} | v_t, \varepsilon_{1t}, g(v_t, \varepsilon_{1t})] | g(v_t, \varepsilon_{1t})]. \tag{2}$$

Using $z_t = g(v_t, \varepsilon_{1t})$ and $E[\boldsymbol{\varepsilon}_{2t} | v_t, \varepsilon_{1t}] = 0$ yields $E\left[\boldsymbol{\varepsilon}_{2t} | z_t\right] = 0$. □

Therefore, rejecting strong exogeneity indicates that the proxy is not just a function of the target shock and the noise term, but also a function of the non-target shocks. A linear proxy $z_t = \psi_1 \varepsilon_{1t} + v_t$ as considered in large parts of the proxy literature, see, e.g. Angelini and Fanelli (2019), and also internal instruments $z_t = \psi_1 \varepsilon_{1t}$, see, e.g. Jarociński and Karadi (2020), are special cases of $z_t = g(\varepsilon_{1t}, v_t)$.

Note that in general, rejecting strong exogeneity does not necessarily imply that the proxy is correlated with the non-target shocks. In general, it is possible that $E[\boldsymbol{\varepsilon}_{2t} z_t] = 0$ but $E\left[\boldsymbol{\varepsilon}_{2t} | z_t\right] \neq 0$. However, Proposition 1 implies that in this case, the proxy is also a function of the non-target shocks. For example, a proxy $z_t = \psi_1 \varepsilon_{1t} + \psi_2 \varepsilon_{2t} - \frac{\psi_2}{3} \varepsilon_{2t}^3 + v_t$ with i.i.d. standard normal shocks satisfies $E[\varepsilon_{2t} z_t] = 0$ even though it is not strongly exogenous. Although mathematically possible, these cases are not economically relevant. Furthermore, if we only consider a linear (or internal) proxy, i.e. $z_t = \psi_1 \varepsilon_{1t} + \psi_2 \varepsilon_{2t} + v_t$, strong exogeneity implies $\psi_2 = 0$ with Proposition 1. Therefore, for a linear proxy, strong exogeneity implies uncorrelatedness of the proxy and non-target shocks.

[2] Note that strong exogeneity does not imply independence. For example, strong exogeneity still allows that the proxy variable and the non-target shock are driven by the same volatility process or that the shocks themselves follow an ARCH-type process. Furthermore, it even allows that the non-target shocks can have predictive power for the expected value of the proxy variable, only the opposite is prohibited.

Proposition 1 assumes that the noise and target shock contain no information on the expected value of the non-target shocks. While assuming $E[\varepsilon_{2t}|v_t] = 0$ appears uncontroversial, assuming $E[\varepsilon_{2t}|\varepsilon_{1t}] = 0$ requires justification. Although most identification methods require only uncorrelated shocks, $E[\varepsilon_{2t}\varepsilon_{1t}] = 0$, applications implicitly rest on the mean independence assumption $E[\varepsilon_{2t}|\varepsilon_{1t}] = 0$, see Keweloh (2024). To see this, let $y_{1t} = b_{11}\varepsilon_{1t} + b_{12}\varepsilon_{2t}$. The response b_{11} is typically interpreted as the expected response of y_{1t} to ε_{1t}, i.e. $b_{11} = E\left[y_{1t}|\varepsilon_{1t} = 1\right]$. However, the assumption of uncorrelated shocks is not sufficient to guarantee that this equality holds, but instead requires mean independent shocks, i.e. in general it holds that $E\left[y_{1t}|\varepsilon_{1t} = 1\right] = b_{11}E\left[\varepsilon_{1t}|\varepsilon_{1t} = 1\right] + b_{12}E\left[\varepsilon_{2t}|\varepsilon_{1t} = 1\right]$ and thus $b_{11} = E\left[y_{1t}|\varepsilon_{1t} = 1\right]$ only holds if $E\left[\varepsilon_{2t}|\varepsilon_{1t}\right] = 0$. Consequently, assuming $E[\varepsilon_{2t}|v_t, \varepsilon_{1t}] = 0$ in Proposition 1 is a reasonable assumption, which is required anyway to ensure the interpretation of impulse responses as the expected response to the target shock.

2.2. A strong exogeneity test

In contrast to the commonly used proxy exogeneity Assumption 1, which imposes uncorrelatedness of the proxy with all non-target shocks, strong exogeneity in Assumption 2 is a conditional moment restriction and can be tested, see e.g. Donald et al. (2003) or Bierens (1982) for conditional moment restriction tests in general. Intuitively, a strongly exogenous proxy can be used to construct an additional synthetic proxy, $\tilde{z}_t$. If the synthetic proxy provides overidentifying restrictions, a simple J-test can be used to test for strong exogeneity.

A strongly exogenous proxy contains no information on the expected value of all non-target shocks, $E\left[\varepsilon_{2t}|z_t\right] = 0$. The law of iterated expectations implies $E\left[\varepsilon_{2t}h(z_t)\right] = 0$ for any measurable function $h(\cdot)$. Under strong exogeneity, the variable

$$\tilde{z}_t := h(z_t) \tag{3}$$

is uncorrelated with the non-target shocks and thus, satisfies the traditional proxy exogeneity assumption imposing uncorrelatedness of the proxy with all non-target shocks in Assumption 1. Consequently, we refer to $\tilde{z}_t$ as a synthetic proxy variable. Although Eq. (3) allows to generate infinitely many synthetic proxies, for simplicity, we focus on a straightforward example and use the synthetic proxy $\tilde{z}_t := z_t^2$.

If the original proxy z_t is strongly exogenous, it follows that both proxies are exogenous, i.e. uncorrelated with the non-target shocks. The original proxy z_t yields $(K-1)$ moment conditions

$$E[f_z(\beta, u_t)] = 0 \quad \text{with} \quad f_z(\beta, u_t) = \boldsymbol{u}_{2t}z_t - \beta u_{1t}z_t, \tag{4}$$

which identify β if z_t is valid. The synthetic proxy yields $(K-1)$ additional conditions

$$E[f_{\tilde{z}}(\beta, u_t)] = 0 \quad \text{with} \quad f_{\tilde{z}}(\beta, u_t) = \boldsymbol{u}_{2t}\tilde{z}_t - \beta u_{1t}\tilde{z}_t, \tag{5}$$

which lead to a potentially overidentified system. If the proxy z_t is strongly exogenous, all moment conditions hold[3] and the GMM estimator

$$\hat{\beta}_T := \underset{\beta \in \mathbb{R}^{K-1}}{\text{argmin}} \; g_T(\beta)' W g_T(\beta) \quad \text{with} \quad g_T(\beta) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^T f_z(\beta, u_t) \\ \frac{1}{T}\sum_{t=1}^T f_{\tilde{z}}(\beta, u_t) \end{bmatrix} \tag{6}$$

with a suitable weighting matrix W is consistent, i.e. $\hat{\beta}_T \xrightarrow{p} \beta$, if z_t is relevant. Moreover, since the model is overidentified, the data can provide evidence that the moment conditions do not hold, and thus provide evidence against the strong exogeneity assumption. Therefore, a simple J-test can be used to test the strong exogeneity assumption. The J-test statistic proposed by Hansen (1982) is given by

$$J_T = T g_T(\hat{\beta}_T)' S^{-1} g_T(\hat{\beta}_T), \tag{7}$$

with $S := lim_{T\to\infty} E\left[T g_T(\beta_0) g_T(\beta_0)'\right]$.[4] Under the null hypothesis of a correctly specified model with $E[f_z(\beta, u_t)] = E[f_{\tilde{z}}(\beta, u_t)] = 0$, the distribution of the test statistic is given by $J_T \xrightarrow{d} \chi^2_{r-q}$ where $r = 2(K-1)$ is equal to the number of moment conditions and $q = K-1$ is equal to the number of elements in β. Rejecting the J-test yields evidence that at least one of the moment conditions is not correct, thus providing evidence against strong exogeneity of the proxy variable.

The J-test is robust to weak identification if the efficient weighting matrix is used, see Stock and Wright (2000). If the proxy is relevant, we consistently estimate the efficient weighting matrix, which implies that the exogeneity test is robust w.r.t weak synthetic proxies. However, if the original and synthetic proxy are both irrelevant, the efficient weighting matrix is not consistently

[3] If the proxy is strongly exogenous, both moment conditions hold since $E[f_z(\beta, u_t)] = E[\boldsymbol{u}_{2t}z_t - \beta u_{1t}z_t] \stackrel{*}{=} \beta_0 E[\varepsilon_{1t}z_t] - \beta E[\varepsilon_{1t}z_t] = 0$ where the equality highlighted by $*$ uses the exogeneity assumption and $E[f_{\tilde{z}}(\beta, u_t)] = 0$ follows analogously.

[4] We follow standard practice and use a two-step GMM estimator weighting each moment condition by the inverse of its variance calculated at $\hat{\beta}_T$ obtained from the proxy variable in the first step and in the second step $W = \hat{S}(\hat{\beta}_T)^{-1}$ with $\hat{\beta}_T$ from the first step and $\hat{S}(\hat{\beta}_T) = \frac{1}{T}\sum_{t=1}^T \left[f_t(\tilde{\beta}_T) f_t(\tilde{\beta}_T)'\right]$. The estimator $\hat{S}(\hat{\beta}_T)$ is robust to heteroskedasticity and remains a consistent estimator for S as long as the moment conditions are serially uncorrelated, see Hall (2005). For serially correlated moment conditions the estimator could be replaced by a HAC estimator, see Newey and West (1994). Consistency of $\hat{S}(\hat{\beta}_T)$ ensures that the J-test statistic remains χ^2 distributed, see Hall (2005).

estimated and the test statistic does not follow a χ^2_{r-q} distribution. In this case, robust critical values can be used based on the continuous updating estimator (CUE):

$$\hat{\beta}_T^{CUE} = \underset{\beta}{argmin}\ J_T(\beta) = T g_T(\beta)' S(\beta)^{-1} g_t(\beta). \tag{8}$$

Since $J_T(\beta_0) \xrightarrow{d} \chi^2_r$ and $J_T(\hat{\beta}_T^{CUE}) < J_T(\beta_0)$, we can use critical values of a χ^2_r instead of a χ^2_{r-q} distribution to ensure a conservative test under weak identification, see Stock and Wright (2000). In the appendix, we repeat the Monte Carlo simulations conducted in the main text with irrelevant proxies. We find that tests using the conservative χ^2_r critical value as well as tests using the standard χ^2_{r-q} critical value reject both below the nominal level under the Null, indicating that the tests behave conservatively if the original and the synthetic proxy are both irrelevant. However, the results also show that the conservative χ^2_r critical value leads to a notable power loss. Given that the results indicate that both tests, using the standard and conservative critical value, lead to a conservative test under weak identification, we recommend using the standard critical value χ^2_{r-q}.

2.3. Power properties

The proposed test only leverages information contained in the proxy. This section derives the conditions under which the proxy contains sufficient information to provide evidence against strong exogeneity. Specifically, we show how non-Gaussian shocks and non-linearity of the proxy can lead to informative synthetic proxy variables which allows to provide evidence against strong exogeneity.

The ability to provide evidence against strong exogeneity depends on whether the information of the proxy is contained entirely in the second moments $E(\epsilon_{1t} z_t)$ and $E(\epsilon_{2t} z_t)$. Consider the synthetic proxy moment conditions from Eq. (5) with

$$E[f_{\tilde{z}}(\beta, u_t)] = E[\boldsymbol{u}_{2t} z_t^2 - \beta u_{1t} z_t^2] = (\beta_0 - \beta) E[\epsilon_{1t} z_t^2] + (B_{22,0} - \beta B_{12,0}) E[\epsilon_{2t} z_t^2] = 0,$$

where $B_{12,0}$ is the upper-right $(1 \times (K-1))$ matrix of B_0 and $B_{22,0}$ is the lower-right $((K-1) \times (K-1))$ matrix of B_0. If the synthetic proxy contains no information on the target or non-target shocks, i.e. $E(\epsilon_{1t} z_t^2) = E(\epsilon_{2t} z_t^2) = 0$, this set of moment conditions is equal to zero for any finite β. Moreover, if the synthetic proxy contains no information beyond the information contained in the second moments $E(\epsilon_{1t} z_t)$ and $E(\epsilon_{2t} z_t)$, such that $E(\epsilon_{1t} z_t) = E(\epsilon_{1t} z_t^2)$ and $E(\epsilon_{2t} z_t) = E(\epsilon_{2t} z_t^2)$, the moment conditions are fulfilled by the same β which fulfills the traditional proxy moment conditions in Eq. (4).[5] Therefore, if the proxy information is entirely contained in the second moments and the synthetic proxy contains no additional information, the J-test cannot detect endogeneity and rejects at the nominal level.

Conversely, if the information of the proxy variable is not entirely contained in the second moments, meaning the synthetic proxy is correlated with the target or contaminating shock, the synthetic proxy moment conditions may not be fulfilled for a given β vector, leading to rejection of the J-test. This correlation can result from non-Gaussian shocks or a non-linear proxy process.

First, consider a linear proxy process (see e.g. Bruns and Lütkepohl (2023))

$$z_t = \psi_1 \epsilon_{1t} + \psi_2 \epsilon_{2t} + v_t. \tag{9}$$

In this case, the moments $E(\epsilon_{1t} z_t^2)$ and $E(\epsilon_{2t} z_t^2)$ of the synthetic proxy are equal to

$$E(\epsilon_{1t} z_t^2) = \psi_1^2 E[\epsilon_{1t}^3] \quad \text{and} \quad E(\epsilon_{2t} z_t^2) = \psi_2^2 E[\epsilon_{2t}^3].$$

For a skewed target shock $E[\epsilon_{1t}^3] \neq 0$ or a skewed contaminating shock $E[\epsilon_{2t}^3] \neq 0$, the synthetic proxy contains information which can be used for detecting proxy endogeneity.[6] Specifically, the synthetic proxy moment conditions in Eq. (5) may not be fulfilled by the β vector which fulfills the traditional proxy moment conditions in Eq. (4), such that the J-test rejects.

Second, consider an SVAR with exclusively Gaussian shocks and a non-linear proxy DGP $z_t = g(\epsilon_{1t}, \epsilon_{2t}, v_t)$. For this proxy, the covariances are given by

$$E(\epsilon_{1t} z_t^2) = E[\epsilon_{1t} g(\epsilon_{1t}, \epsilon_{2t}, v_t)^2] \quad \text{and} \quad E(\epsilon_{2t} z_t^2) = E[\epsilon_{2t} g(\epsilon_{1t}, \epsilon_{2t}, v_t)^2].$$

Depending on the type of non-linearity, either or both correlations can be non-zero even in the presence of fully Gaussian shocks. As a simple example, consider the process $z_t = \epsilon_{1t} + \epsilon_{1t}^2 + \eta_t$ where the proxy is affected by the volatility of the target shock and let ϵ_{1t} and η_t be i.i.d. standard normal shocks. In this case, the moment $E(\epsilon_{1t} z_t^2)$ is equal to $E(\epsilon_{1t}(\epsilon_{1t} + \epsilon_{1t}^2 + \eta_t)^2) = 2E(\epsilon_{1t}^4) \neq 0$, such that synthetic proxy moment conditions can again provide evidence against strong exogeneity.

In summary, the proposed test only uses information contained in the proxy variable, which contains sufficient information to detect endogeneity if the synthetic proxy is informative about the target or contaminating shock. Potential sources for such information are skewed shocks or non-linear proxy variables. Only for the special case of exactly Gaussian shocks and an exactly linear proxy model, the synthetic proxy cannot provide additional information and the test has no power.

[5] This case, for example, occurs for a binary proxy variable.

[6] Note that if only the non-contaminating non-target shock ϵ_{3t} or the proxy noise term v_t displays skewness, then the synthetic proxy does not contain information which can be used to test for strong proxy exogeneity.

2.4. Generalizations

The exogeneity test can easily be extended to the case of multiple proxies with multiple target shocks. For example, consider two proxy variables z_{1t} and z_{2t} correlated with the target shocks ε_{1t} and ε_{2t}, and exogenous w.r.t. the remaining shocks summarized in $\boldsymbol{\varepsilon}_{3t}$, i.e. $E[\boldsymbol{\varepsilon}_{3t} z_{1t}] = E[\boldsymbol{\varepsilon}_{3t} z_{2t}] = 0$. In this case, the proxies identify the target shocks up to a linear transformation, see e.g. Bruns and Lütkepohl (2024). Analogously to the exogeneity test for a single proxy variable, we can test the strong exogeneity assumption $E[\boldsymbol{\varepsilon}_{3t}|z_{1t}] = E[\boldsymbol{\varepsilon}_{3t}|z_{2t}] = 0$. Specifically, if both proxies are strongly exogenous w.r.t. $\boldsymbol{\varepsilon}_{3t}$, we can construct synthetic proxies $h_1(z_{1t})$ and $h_2(z_{2t})$ and derive a J-test analogously to Section 2.2, see Appendix B.

Moreover, synthetic proxies can be useful beyond the proposed exogeneity test. For example, using a synthetic proxy in addition to the original proxy in the SVAR estimation can lead to asymptotic efficiency gains compared to the estimation using only the original proxy. In a general setup not related to SVAR models, Chamberlain (1987) derives an asymptotic efficiency bound for conditional moment restrictions models and Donald et al. (2003) propose estimators using a series of unconditional moment restrictions where the number of moment conditions increases with the sample size, which can achieve the efficiency bound. However, it is not clear whether synthetic proxies yield efficiency gains in typical macroeconomic applications with small samples, as distortions introduced by many moments in small samples may outweigh potential asymptotic efficiency gains in small samples.

Lastly, the focus of this section is on a simple illustrating example of a synthetic proxy, i.e. $\tilde{z}_t = z_t^2$. The simplicity allows to easily trace the power of the exogeneity test to the shocks' skewness or non-linearity of the proxy. Furthermore, the application in Section 4 illustrates that the simple synthetic proxy is capable of providing evidence against exogeneity of a commonly used fiscal proxy. At the same time, the previous subsection highlights cases where the simple synthetic proxy $\tilde{z}_t = z_t^2$ yields no power to detect endogeneity, while a different synthetic proxy might have power. In general, strong exogeneity $E\left[\varepsilon_{2t}|z_t\right] = 0$ implies an infinite number of synthetic proxies $\tilde{z}_t = h(z_t)$. Donald et al. (2003) propose to employ approximating functions to generate a set of unconditional moment conditions, where the number of moment conditions goes to infinity with the sample size. Therefore, in the limit, any function $h(z_t)$ can be approximated.

3. Monte Carlo simulation

We set up a Monte Carlo experiment to investigate the small sample performance of the proposed exogeneity test. We find that the size and power depend on sample size, degree of proxy endogeneity, proxy strength, and the degree of the synthetic proxy informativeness, which is affected by the skewness of the structural shocks and the non-linearity in the proxy equation. We use two data-generating processes (DGPs).

3.1. DGP1: Stylized SVAR simulation

First, we use a simplification of the DGP in Lütkepohl and Schlaak (2022) and Bruns and Lütkepohl (2024). The SVAR has three variables, one proxy to identify one shock, and follows a VAR(0)

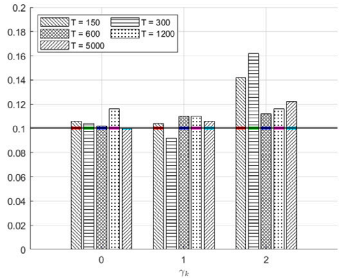
$$u_t = B_0 \varepsilon_t \quad \text{with} \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 4 \\ 4 & 6 & 6 \end{bmatrix}.$$

Structural shocks are generated from the Pearson family of distributions, i.e. $\varepsilon_{kt} \sim \mathcal{P}(\mu_k, \sigma_k^2, \gamma_k, \kappa_k)$, where μ_k is the mean, σ_k^2 the variance, γ_k the skewness, and κ_k the kurtosis. We assume $\mu_k = 0$, $\sigma_k^2 = 1$ and vary $\gamma_k = [0, 1, 2]$ for all shocks jointly. We keep $\kappa_k = 6$ constant across simulations. Proxies are generated using $z_t = \psi_1 \varepsilon_{1t} + \psi_2 \varepsilon_{2t} + v_t$, with $v_t \sim N(0, 1)$. We choose (ψ_1, ψ_2) to simulate samples with three different levels of proxy strength with $corr(\varepsilon_{1t}, z_t) = (0.5, 0.7, 0.9)$ and three levels of proxy exogeneity with $corr(\varepsilon_{2t}, z_t) = (0, -0.2, -0.5)$. We generate $T = [150, 300, 600, 1200, 5000]$ observations (plus pre-sample values) and use 500 repetitions for each simulation. We use $\tilde{z}_t = z_t^2$ as a synthetic proxy.

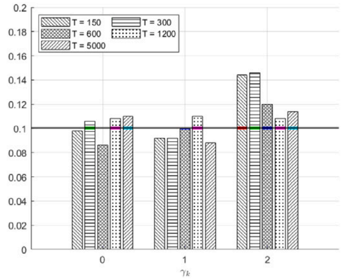
Fig. 1 panels (a)–(c) show the rejection frequencies when data are generated under the Null of proxy exogeneity, i.e. $corr(\varepsilon_{2t}, z_t) = 0$, with a nominal level of 10%. While an increase in the sample size, T, leads to a better matching of the nominal level, we note that even for the smallest sample size, $T = 150$ the rejection frequency does not exceed 16% for any setup. Variations in proxy strength, $corr(\varepsilon_{1t}, z_t)$ do not seem to affect the nominal level much. Moreover, the skewness, and consequently the strength of the synthetic proxy, has little effect on the nominal level.

Fig. 1 panels (d)–(i) show the rejection frequencies when data are generated under proxy endogeneity, i.e. $corr(\varepsilon_{2t}, z_t) \neq 0$. When shocks have zero skewness ($\gamma_k = 0$) the test has no power and rejects close to the nominal level of 10%. Intuitively, in this case z_t^2 is not informative and cannot provide evidence against strong exogeneity. A larger skewness, i.e. increasing γ_k leads to more power. Stronger proxies, i.e. higher $corr(\varepsilon_{1t}, z_t)$ as well as larger deviations from the Null in the form of a higher proxy endogeneity, i.e. higher $corr(\varepsilon_{2t}, z_t)$, also increase the test's power.

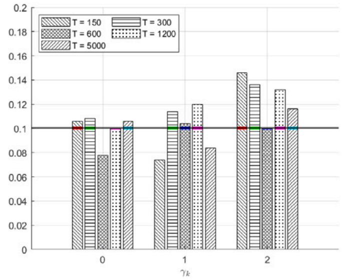
In Appendix A, we introduce two approaches for incorporating lags into our testing methodology. First, we propose a one-step approach, where the J-test incorporates supplementary moment conditions corresponding to the lags of the SVAR. Second, we adopt a two-step approach, where the VAR is estimated in the first step and the J-test is conducted similarly to the previous section using the reduced form shocks from the VAR. A similar two-step procedure is typically used to test proxy strength using an F-test, see, Stock et al. (2002). Our simulations in the next section and within the appendix show that the two-step approach yields results close to the nominal level for all simulations with lags and exogenous proxy variables.



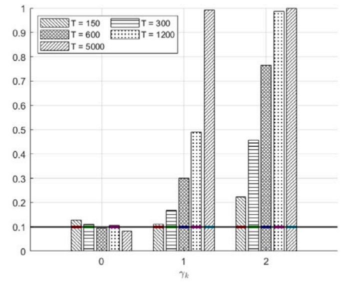
(a) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = 0$



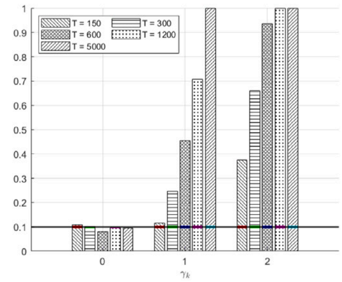
(b) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = 0$



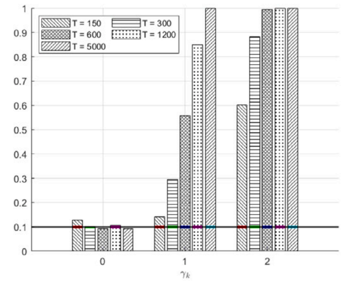
(c) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = 0$



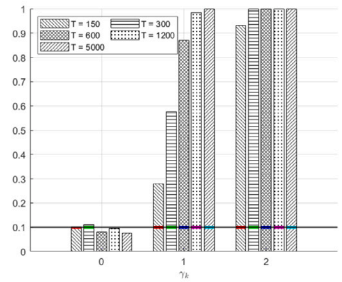
(d) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.2$



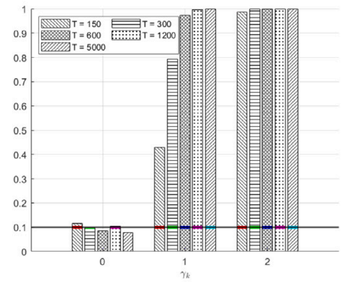
(e) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.2$



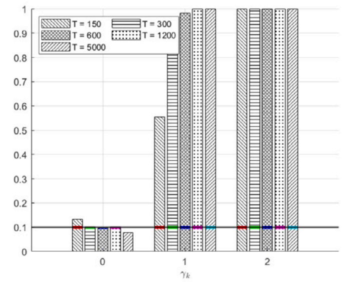
(f) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.2$



(g) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.5$



(h) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.5$



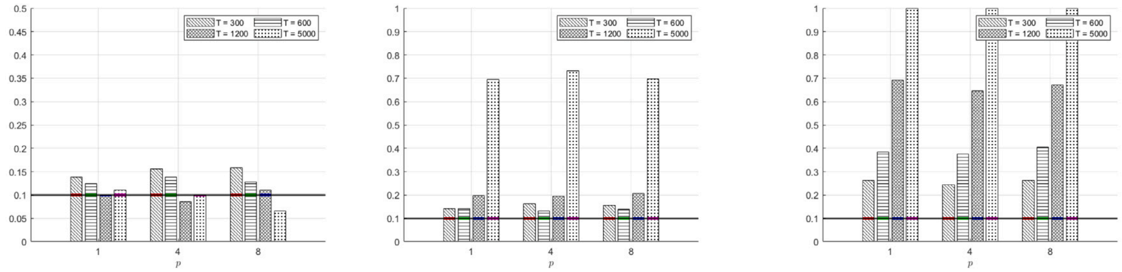
(i) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.5$

Fig. 1. Relative rejection frequencies for DGP1. Nominal significance level 10%. $p = 0$.

The simulation above is based on a very stylized VAR(0) model. In Appendix D we conduct a wide range of alternative simulations based on a VAR(p) including models estimated with $p = (1, 12)$ lags (Figs. D.1 and D.2). We also show that for our test to have power, it is enough for either the target (ε_{1t}) or the contaminating shock (ε_{2t}) to be skewed, while a skewed non-contaminating non-target shock (ε_{3t}) does not provide power (Fig. D.3). Moreover, we investigate a fully Gaussian set of shocks and three different types of non-linearities in the proxy equation and find that the test has power in these scenarios as well (Fig. D.4). Lastly, we show results for an SVAR with two proxies for two shocks (Fig. D.5), simulations using different functions of synthetic proxies (Figs. D.6 and D.7), results for an irrelevant proxy (Fig. D.8), and results using a continuous updating estimator together with weak-instrument robust critical values (Figs. D.9 and D.10).

3.2. DGP2: Fiscal SVAR simulation

Second, we use a more realistic DGP based on the VAR in Mertens and Ravn (2014). The exact VAR can be found in the appendix. We estimate models with $p = [1, 4, 8]$ lags and intercept for $T = [300, 600, 1200, 5000]$ and 500 repetitions per simulation.



(a) $corr(\varepsilon_{2t}, z_t) = 0$ (b) $corr(\varepsilon_{2t}, z_t) = -0.1$ (c) $corr(\varepsilon_{2t}, z_t) = -0.2$

Fig. 2. Relative rejection frequencies for DGP2 (2-step test). Nominal significance level 10%.

Fig. 2 shows the results for the two-step testing procedure. In panel (a) data are generated under the Null of strong proxy exogeneity, i.e. $corr(\varepsilon_{2t}, z_t) = 0$. The empirical rejections frequencies are close to the nominal level of 10% for all sample sizes and lag lengths shown. Panels (b) and (c) show that the power of the test increases with sample size and distance from the Null, i.e. higher correlation of the proxy with the non-target shock, but is only marginally affected by the lag length, p. In the Appendix we show that using a one-step testing procedure taking into account the moment conditions for the autoregressive coefficients leads to higher power at the cost of a less precise nominal level in smaller samples (Fig. D.11). Overall, DGP2 shows that the test has good size and power properties in a realistic environment.

4. Application

This section applies the proposed exogeneity test to three proxy variables used in the fiscal proxy SVAR literature. Specifically, we test exogeneity of the tax proxy used in Mertens and Ravn (2014), the Fernald (2012) total factor productivity measure as an output shock proxy, and military spending changes as a government spending shock proxy (see e.g. Klein and Linnemann (2019)). We find evidence against strong exogeneity of the tax proxy, but no evidence against exogeneity of the output and spending proxies. However, we find that the synthetic output proxy is weak, which may limit our ability to detect exogeneity violations of the output proxy.

We consider the fiscal SVAR as proposed by Mertens and Ravn (2014) for the US. The variables are federal tax revenues (τ_t), federal government consumption (g_t), and output (y_t), all in log real per capita terms for the sample 1950Q2 to 2006Q4, leading to $T = 228$ observations. The SVAR has four lags a constant, linear and quadratic trends, and a dummy for $1975Q2$ all contained in X_t with

$$\begin{bmatrix} \tau_t \\ g_t \\ y_t \end{bmatrix} = \gamma X_t + \sum_{i=1}^{4} A_i \begin{bmatrix} \tau_{t-i} \\ g_{t-i} \\ y_{t-i} \end{bmatrix} + \begin{bmatrix} u_{\tau,t} \\ u_{g,t} \\ u_{y,t} \end{bmatrix} \text{ and } \begin{bmatrix} u_{\tau,t} \\ u_{g,t} \\ u_{y,t} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_{\tau,t} \\ \varepsilon_{g,t} \\ \varepsilon_{y,t} \end{bmatrix}, \tag{10}$$

tax shocks $\varepsilon_{\tau,t}$, government spending shocks $\varepsilon_{g,t}$, and output shocks $\varepsilon_{y,t}$.

We test the exogeneity of the following three proxies. First, the narrative tax proxy $z_{\tau,t}$ for tax shocks $\varepsilon_{\tau,t}$ constructed by Mertens and Ravn (2014) based on the Romer and Romer (2010) tax shocks identified by studying narrative records of tax policy decisions. Second, the Fernald (2012) TFP measure as a proxy $z_{y,t}$ for output shocks $\varepsilon_{y,t}$ used by Caldara and Kamps (2017) as a non-fiscal proxy. Third, military spending changes as a proxy $z_{g,t}$ for government spending shocks $\varepsilon_{g,t}$ used in Klein and Linnemann (2019).

Table 1 shows evidence for the relevance of the original and synthetic proxies. Since proxies can be thought of as noisy shock measurements, skewed proxies indicate skewed shocks, which lead to relevant synthetic proxies, see Section 2.3. Therefore, Table 1 displays the skewness of the three proxy variables as well as the F-statistic for the original and synthetic proxies. The F-statistic indicates that the original tax proxy may be weak. However, the tax proxy exhibits a strong negative skewness (see also Fig. E.12 panel (b) in Appendix E), potentially driven by left skewed structural tax shocks, which leads to a F-statistic of the synthetic tax proxy close to 10. The spending proxy is right skewed, potentially driven by right skewed spending shocks, and the F-statistic indicates relevant original and synthetic spending proxies. The F-statistic also indicates a relevant output proxy. However, the output proxy has a skewness close to zero, which leads to an irrelevant synthetic output proxy. Based on the results presented in the previous section, we expect that the exogeneity test may have low power for the tax and output proxy.

Table 2 shows the results of the proxy exogeneity tests. First, we find evidence that the tax proxy is endogenous. Specifically, we reject the null hypothesis that the proxy and synthetic proxy moment conditions (4) and (5) hold at the 10% level, indicating that the tax proxy is not strongly exogenous. This outcome aligns with the findings of Lewis (2021) and Keweloh et al. (2023b), both of whom provide evidence against exogeneity of the tax proxy. However, these studies rely on additional identifying assumptions, i.e. time-varying volatility or non-Gaussian and independent shocks. Our contribution to this body of literature is the provision of additional evidence challenging the exogeneity of the tax proxy, without relying on any supplementary identifying assumptions, such as heteroskedasticity or the assumption of independent and non-Gaussian shocks.

Table 1
Skewness and strength of proxy variables.

	Proxy	Proxy		Synth proxy	
	Skewness	F-stat	F-stat (rob)	F-stat	F-stat (rob)
Tax proxy	−4.47	4.22	1.6	9.6	9.3
Spending proxy	3.22	131	69.8	25.7	82.5
Output proxy	−0.02	56.5	38.9	0.01	0.01

Note: Robust F-statistics allow for heteroskedasticity.

Table 2
Proxy exogeneity test.

	J-statistic	p-Value
Tax proxy	5.48	0.06
Spending proxy	0.68	0.71
Output proxy	1.59	0.45

Note: The table presents the results of the two-stage proxy exogeneity test, which involves estimating the VAR in the first step and subsequently conducting the exogeneity test in the second step using the original proxy variable z_t and its corresponding synthetic proxy $\tilde{z}_t = z_t^2$.

Second, we find no evidence against exogeneity of the spending and output proxy. Specifically, we cannot reject the null hypothesis that the proxy and synthetic proxy moment conditions (4) and (5) for the spending and output proxy hold at the 10% level. However, we stress that the output proxy displays almost no skewness and the synthetic output proxy appears to be not relevant. Therefore, the exogeneity test may have little power to detect exogeneity violations of the output proxy. In contrast, the spending proxy has a positive skewness and both, the original and the synthetic spending proxy, are found to be relevant, indicating that our exogeneity test may have power to detect exogeneity violations of the spending proxy.

Consequently, we present evidence against strong exogeneity of the tax proxy. Our findings suggest that the tax proxy carries information related to the expected value of the non-target shocks. It is essential to note that while it is theoretically conceivable that the tax proxy is uncorrelated with all non-target shocks, while the squared tax proxy, i.e. the synthetic tax proxy, is correlated with a non-target shock. In such a scenario, our test would indeed reject the strong exogeneity of the tax proxy, even when the tax proxy itself shows no correlation with non-target shocks. Nevertheless, the critical inquiry does not revolve around the technical possibility of this scenario but rather its economic plausibility. Mertens and Ravn (2014) construct the tax proxy as a series of unanticipated tax shocks based on the Romer and Romer (2010) narrative tax shocks. If, indeed, the tax proxy represents a series of unforeseen tax shocks, then the squared tax proxy equates to a series of squared unforeseen tax shocks. Therefore, the same reasoning used to motivate exogeneity of the tax proxy can also be applied to argue for exogeneity of the synthetic tax proxy. Consequently, in line with the economic rationale underlying the proxy's construction, we expect both the original and synthetic proxy variables to be exogenous. However, the data provide evidence against this hypothesis. Taking into account the economic rationale of the proxy's construction, our results provide evidence against the validity of the proxy itself.

5. Conclusions

Our study addresses the issue of proxy exogeneity in structural vector autoregressions. Traditionally, asserting the exogeneity of a proxy has rested on economic justifications rather than statistical assessments. We introduce a novel proxy exogeneity test based on the strong proxy exogeneity assumption, which implies that the proxy variable is not just uncorrelated with non-target shocks but contains no information at all on the expected value of non-target shocks. This extension allows for direct testing of the enhanced notion of proxy exogeneity, offering a more robust framework for evaluating the reliability of proxy-based SVAR models than just relying on some narrative justification for the exogeneity of the proxy. Importantly, the proposed framework only exploits information contained in the proxy itself. We show that the proxy itself can contain the necessary information to provide evidence against exogeneity if the proxy is a non-linear function of the shocks or affected by non-Gaussian shocks. By applying our approach to widely-used proxy variables in the fiscal SVAR literature, we demonstrate its effectiveness in uncovering deviations from strong exogeneity.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Extension to $p > 0$

For the realistic setting of a VAR(p) model with $p > 0$ one needs to decide how to account for the estimation uncertainty in the residuals, $\hat{u}_t$. Here, we suggest two approaches: As a baseline, the J- test presented in the paper can be directly applied conditioning on OLS-residuals, $\hat{u}^{OLS}$, without further changes. We refer to this as the "two-step" testing procedure since it involves first estimating the model via OLS to obtain residuals and then applying our test. Alternatively, the moment conditions in (B.5) can be augmented by the $K(Kp+3)$ moment conditions relating to the autoregressive slope coefficients, an intercept term as well as a linear and quadratic time trend as follows:

$$\hat{\theta}_T := \underset{\theta \in \mathbb{R}^{n-1+K(Kp+3)}}{\operatorname{argmin}} \ g_T(\theta)' W g_T(\theta) \quad \text{with} \quad g_T(\theta) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} x_t(y_t - \Pi x_t) \\ \frac{1}{T}\sum_{t=1}^{T} f_z(\beta, u_t) \\ \frac{1}{T}\sum_{t=1}^{T} f_{\tilde{z}}(\beta, u_t) \end{bmatrix}, \tag{A.11}$$

with $\theta = [vec(\Pi), \beta]'$, $\Pi = [A_1, \ldots, A_p]$, and $x_t = [1, t, t^2, y_{t-1}, \ldots, y_{t-p}]'$. We label this the "one-step" testing procedure since the VAR slope coefficients, Π, and the impact effects of the identified structural shock, β, are jointly rather than sequentially estimated. Compared to the "two-step" procedure it has the advantage that the weighting matrix, W, can be chosen to be consistent so that the implied J-statistic has an asymptotic χ^2-distribution. In finite samples, and depending on the model's dimensions, the "one-step" procedure is likely to be less precise. The two approaches are numerically identical for $p = 0$.

Appendix B. Extension to multiple proxies

This section generalizes the proposed strong exogeneity test to multiple proxies with multiple target shocks. For simplicity, we only consider the case with two proxies and two target shocks. A generalization to an arbitrary number of proxies is straightforward.

Consider two proxy variables z_{1t} and z_{2t} for the target shocks ε_{1t} and ε_{2t}. Both proxies are assumed to be strongly exogenous w.r.t. the non-target shocks summarized in $\boldsymbol{\varepsilon}_{3t}$ such that $E[\boldsymbol{\varepsilon}_{3t}|z_{1t}] = E[\boldsymbol{\varepsilon}_{3t}|z_{2t}] = 0$. Strong exogeneity again yields synthetic proxies $\tilde{z}_{1t} = h(z_{1t})$ and $\tilde{z}_{2t} = h(z_{2t})$.

Let $\boldsymbol{u}_{3t} = [u_{3t}, \ldots, u_{nt}]'$, $A_0 = B^{-1}$, and rewrite the SVAR as $\boldsymbol{u}_{3t} = -A_{0,31}u_{1t} - A_{0,32}u_{2t} + B_{33}\boldsymbol{\varepsilon}_{3t}$ to obtain the moment conditions. For $\beta = [\beta_1, \beta_2]$, the proxy moment conditions are now equal to

$$E[f_{z_1}(\beta, u_t)] = 0 \quad \text{with} \quad f_{z_1}(\beta, u_t) = \boldsymbol{u}_{3t}z_{1t} - \beta_1 u_{1t}z_{1t} - \beta_2 u_{2t}z_{1t} \tag{B.1}$$

and

$$E[f_{z_2}(\beta, u_t)] = 0 \quad \text{with} \quad f_{z_2}(\beta, u_t) = \boldsymbol{u}_{3t}z_{2t} - \beta_1 u_{1t}z_{2t} - \beta_2 u_{2t}z_{2t} \tag{B.2}$$

and the synthetic proxy moment conditions are now equal to

$$E[f_{\tilde{z}_1}(\beta, u_t)] = 0 \quad \text{with} \quad f_{\tilde{z}_1}(\beta, u_t) = \boldsymbol{u}_{3t}\tilde{z}_{1t} - \beta_1 u_{1t}\tilde{z}_{1t} - \beta_2 u_{2t}\tilde{z}_{1t} \tag{B.3}$$

and

$$E[f_{\tilde{z}_2}(\beta, u_t)] = 0 \quad \text{with} \quad f_{\tilde{z}_2}(\beta, u_t) = \boldsymbol{u}_{3t}\tilde{z}_{2t} - \beta_1 u_{1t}\tilde{z}_{2t} - \beta_2 u_{2t}\tilde{z}_{2t}. \tag{B.4}$$

The GMM estimator is given by

$$\hat{\beta}_T := \underset{\beta \in \mathbb{R}^{K-1}}{\operatorname{argmin}} \ g_T(\beta)' W g_T(\beta) \quad \text{with} \quad g_T(\beta) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} f_{z_1}(\beta, u_t) \\ \frac{1}{T}\sum_{t=1}^{T} f_{\tilde{z}_1}(\beta, u_t) \\ \frac{1}{T}\sum_{t=1}^{T} f_{z_2}(\beta, u_t) \\ \frac{1}{T}\sum_{t=1}^{T} f_{\tilde{z}_2}(\beta, u_t) \end{bmatrix} \tag{B.5}$$

and J-test statistic is equal to

$$J_T = T g_T(\hat{\beta}_T)' S^{-1} g_T(\hat{\beta}_T). \tag{B.6}$$

Under the null hypothesis of a correctly specified model, the distribution of the test statistic is given by $J_T \overset{d}{\to} \chi^2_{r-q}$ where $r = 2(K-1)$ is equal to the number of moment conditions and $q = K - 1$ is equal to the number of elements in β. Rejecting the J-test yields evidence that at least one of the moment conditions is not correct, thus providing evidence against strong exogeneity of the proxy variables.

A setup with multiple proxies, which are exogenous w.r.t. the target shock, would require additional identifying assumptions to achieve point identification of the impact effects of the shocks. The proposed test is robust to such identifying assumptions so that is can be employed in circumstances where multiple proxies are used to identify multiple target shocks, but no further identifying information is available to disentangle these shocks.

Appendix C. Extension to multiple synthetic proxies

From the conditional moment restriction implied by strong proxy exogeneity, i.e. $E\left[\varepsilon_{2t}|z_t\right] = 0$, we extract two unconditional moment conditions, namely

$$E\left[\varepsilon_{2t}z_t\right] = 0$$

$$E\left[\varepsilon_{2t} z_t^2\right] = 0,$$

call z_t^2 a "synthetic proxy", and show that this single additional proxy leads to good size and power properties in a large number of contexts typically encountered in empirical macroeconomics. Of course, additional synthetic proxies (and functions thereof) could be constructed to exploit higher moments of the structural shocks. However, this construction gives rise to a trade-off between the potentially limited additional information contained in such additional synthetic proxies, and the additional noise introduced through more moment conditions. The choice of the number of synthetic proxies is therefore application-specific.

Alternatively, Donald et al. (2003) propose an approach to approximate conditional moments using a function of unconditional moment conditions. Specifically, they propose to use a vector of spline approximating functions

$$q^{\eta}(z_t) = (1, z_t, \dots, z_t^s, \mathbb{1}(z_t - t_1 > 0)(z_t - t_1)^s, \dots, \mathbb{1}(z_t - t_{\eta-s-1} > 0)(z_t - t_{\eta-s-1})^s),$$

where in our context $t_1, \dots, t_{\eta-s-1}$ are evenly-spaced percentiles of z_t. They then construct the vector of moment conditions.

$$E\left[(\boldsymbol{u}_{2t} - \beta u_{1t}) \otimes q^{\eta}(z_t)\right] = 0. \tag{C.1}$$

We follow Donald et al. (2003) and use $s = 3$. We investigate a fixed $\eta = 3$ (see Fig. D.6 bottom row) and a spline order which increases with the sample size using $\eta = T^{1/3}$ (see Fig. D.7). This choice ensures the correct rate restrictions for convergence are satisfied (see Table 1 in Donald et al. (2003)). In both cases the resulting test statistic approximately follows an χ^2-distribution, as in our baseline.

Appendix D. Additional simulation results

D.1. DGP1 with lags

To investigate the test's performance when including lags in the model, we generate data y_t from a VAR(1) following Lütkepohl and Schlaak (2022) and Bruns and Lütkepohl (2024) and augment DGP1 by the following autoregressive parameters:

$$A_1 = \begin{bmatrix} 0.79 & 0.00 & 0.25 \\ 0.19 & 0.95 & -0.46 \\ 0.12 & 0.00 & 0.62 \end{bmatrix}.$$

The largest Eigenvalue of A_1 is 0.95, implying a persistent but stable process. We generate data recursively without intercept starting from $y_1 = [0, 0, 0]'$. The results are shown in Figs. D.1, D.2, D.8 (panel b and c) and D.9 (panel b and c).

D.2. DGP1 with linear proxy model and only one skewed shock

To investigate the test's performance when the proxy model is linear and not all, but only shock ε_{kt} exhibit skewness, we modify DGP1 to allow for only one skewed shock. The results are shown in Fig. D.3. If shocks ε_{1t} (target shock) or ε_{2t} (contaminating shock) are skewed, then the proxy still contains information beyond its first two moments, leading to power to detect a false Null. If only shock ε_{3t} is skewed, then the proxy does not contain such information since the third shock does not enter the proxy Eq. (9).

D.3. DGP1 with non-linear proxy model and Gaussain shocks

To investigate the test's performance when all shocks are exactly Gaussian, but the proxy equation is non-linear, we draw from $\varepsilon_{kt} \sim N(0, 1), \forall k$ and modify (9) as

NL 1:

$$z_t = \begin{cases} \psi_1 \varepsilon_{1t} + v_t & \varepsilon_{1t} + \varepsilon_{2t} < abs(\Phi^{-1}(\psi_3)), \\ \psi_1 \varepsilon_{1t} + \psi_2 \varepsilon_{2t} + v_t & \varepsilon_{1t} + \varepsilon_{2t} > abs(\Phi^{-1}(\psi_3)) \end{cases}, \tag{D.1}$$

where $\Phi^{-1}(x)$ is the cdf of a normal distribution with standard deviation 2. In words, the proxy is contaminated by shock ε_{2t} if the sum $\varepsilon_{1t} + \varepsilon_{2t}$ exceeds a threshold. The non-linearity is stronger for higher values of ψ_3. We investigate $\psi_1 = 1$, $\psi_2 = [0, 0.25, 0.5]$, and $\psi_3 = [0.8, 0.7, 0.5]$. The results are shown in Fig. D.4 (top panel).

Another non-linearity in the proxy is introduced as

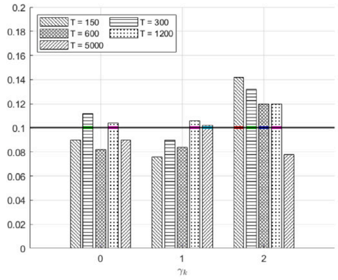
NL 2:

$$z_t = \psi_1 \varepsilon_{1t} + \psi_2 \varepsilon_{2t} + \psi_3 \varepsilon_{1t}^2 + v_t \tag{D.2}$$

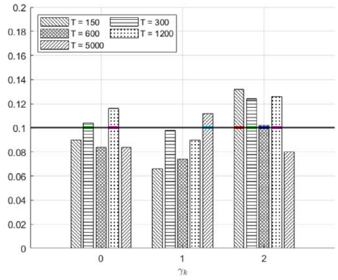
we set $\psi_1 = 1$, vary $\psi_2 = [0, 0.1, 0, 25]$, and $\psi_3 = [0.01, 0.05, 0.1]$. In this case, the non-linearity arises from the squared term, ε_{1t}^2. Again, the non-linearity is stronger for higher values of ψ_3. The results are shown in Fig. D.4 (middle panel).

A third non-linearity in the proxy is motivated by an interest rate zero lower bound. The time series i_t is a truncated AR(1) process with $i_t = \psi_3 + 0.9 i_{t-1} + u_t$ and

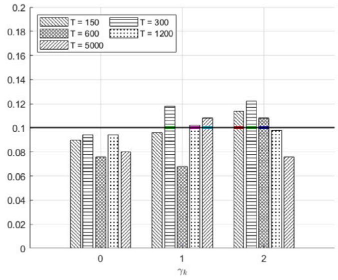
$$w_t = \begin{cases} \psi_1 \varepsilon_{1t} & \psi_3 + 0.9 i_{t-1} + \psi_1 \varepsilon_{1t} > 0 \\ max(\psi_1 \varepsilon_{1t}, -(\psi_3 + 0.9 i_{t-1})\lambda_t) & \psi_3 + 0.9 i_{t-1} + \psi_1 \varepsilon_{1t} \leq 0 \end{cases} \tag{D.3}$$



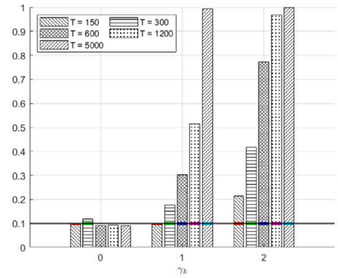
(a) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = 0$



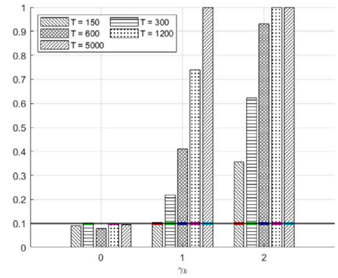
(b) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = 0$



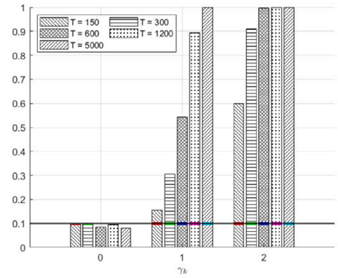
(c) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = 0$



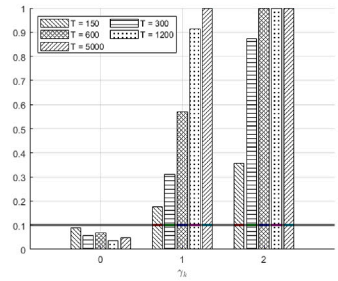
(d) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.2$



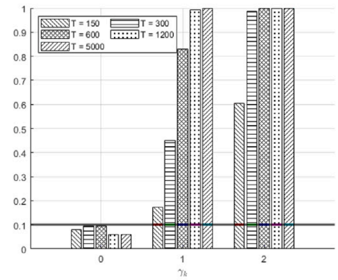
(e) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.2$



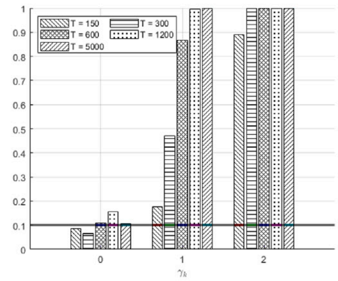
(f) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.2$



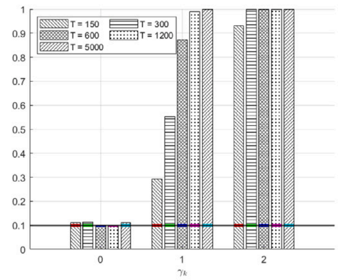
(g) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.3$



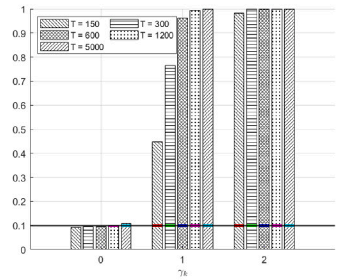
(h) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.3$



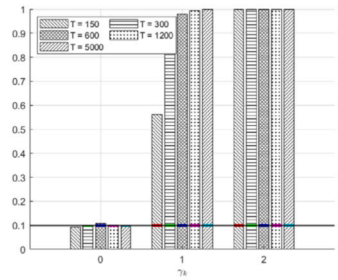
(i) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.3$



(j) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.5$



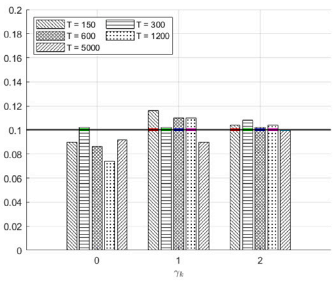
(k) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.5$



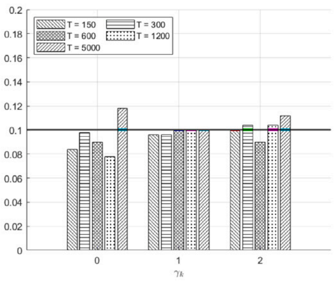
(l) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.5$

Fig. D.1. Relative rejection frequencies for DGP1 (2-step test). Nominal significance level 10%. $p = 1$.

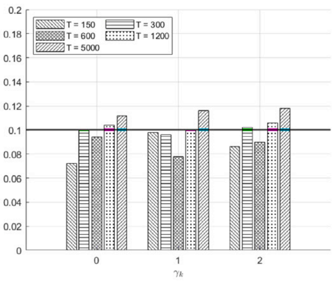
and $\lambda_t \sim \mathcal{U}_{[0,1]}$ such that if a shock ε_{1t} would drive i_t below the zero lower bound, only a random fraction of the shock is realized ensuring that i_t remains above the zero lower bound. The proxy z_t is given by



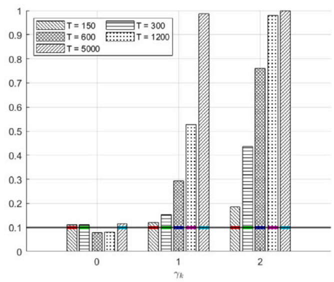
(a) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = 0$



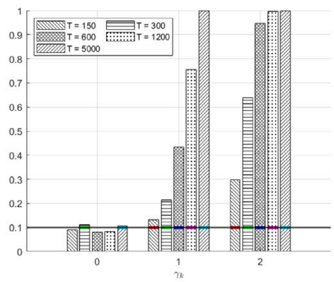
(b) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = 0$



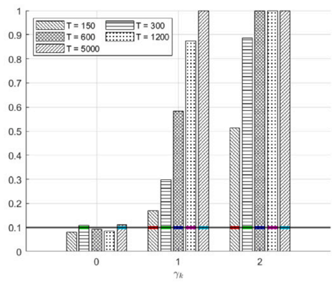
(c) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = 0$



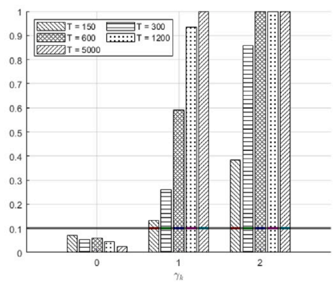
(d) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.2$



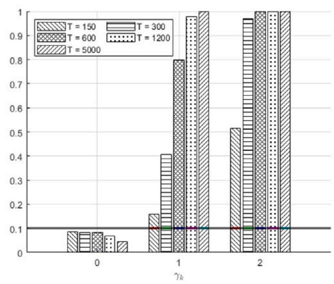
(e) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.2$



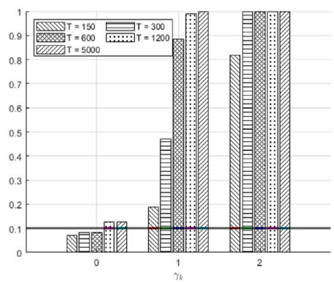
(f) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.2$



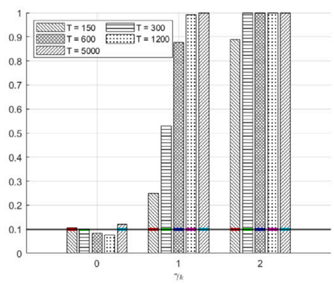
(g) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.3$



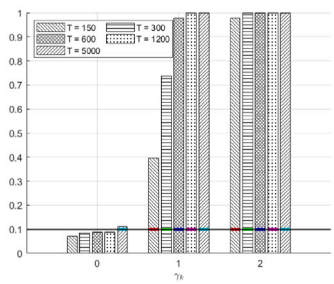
(h) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.3$



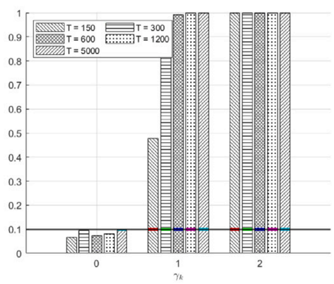
(i) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.3$



(j) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.5$

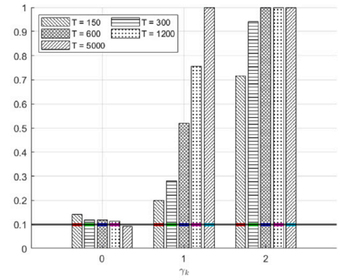


(k) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.5$

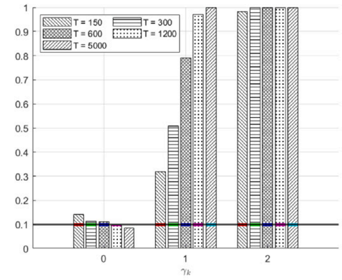


(l) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.5$

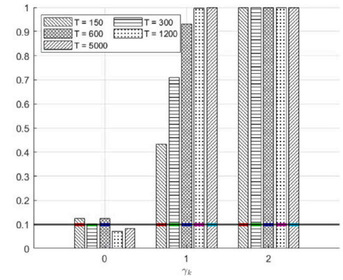
Fig. D.2. Relative rejection frequencies for DGP1 (2-step test). Nominal significance level 10%. $p = 12$.



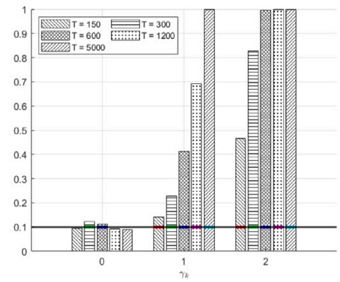
(a) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_2 = \gamma_3 = 0$



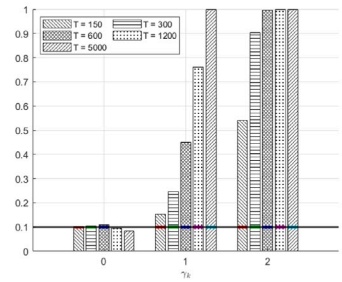
(b) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_2 = \gamma_3 = 0$



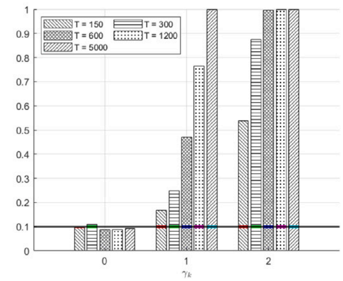
(c) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_2 = \gamma_3 = 0$



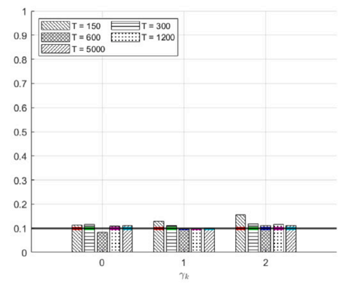
(d) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_1 = \gamma_3 = 0$



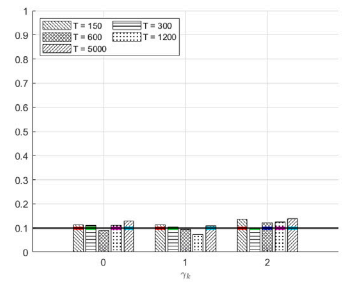
(e) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_1 = \gamma_3 = 0$



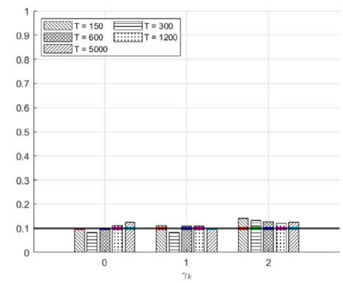
(f) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_1 = \gamma_3 = 0$



(g) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_2 = \gamma_3 = 0$



(h) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_2 = \gamma_3 = 0$



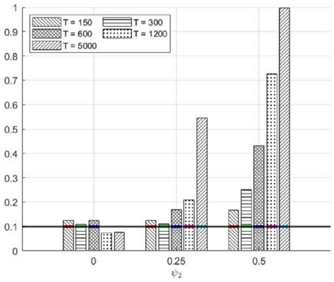
(i) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = 0.5$, $\gamma_2 = \gamma_3 = 0$

Fig. D.3. Relative rejection frequencies for DGP1 when only one shock is skewed. Nominal significance level 10%. $p = 0$. Shock skewness is non-zero only for shock w_{1t} (panels (a)–(c)), only for shock w_{2t} (panels (d)–(f)) or only for shock w_{3t} (panels (g)–(i)).

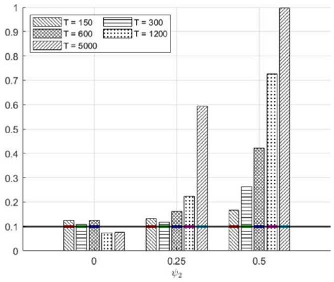
NL 3:

$$z_t = w_t + \psi_2 \varepsilon_{2t} + v_t \tag{D.4}$$

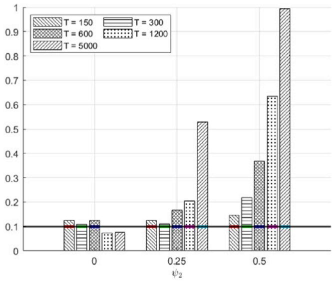
with $\psi_1 = 1$, $\psi_2 = [0, 0.25, 0.5]$ determining the degree of endogeneity and proxy strength, and $\psi_3 = [1, 0.1, 0.]$ governing the degree of non-linearity where higher values lead to a larger distance to the zero lower bound and thus less non-linearity.



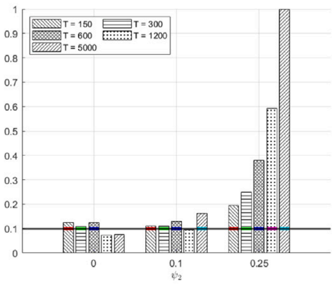
(a) $\psi_3 = 0.5$



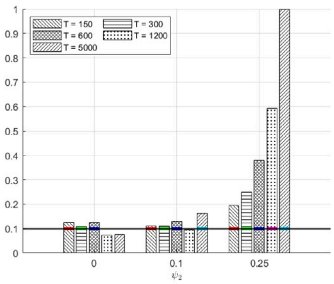
(b) $\psi_3 = 0.7$



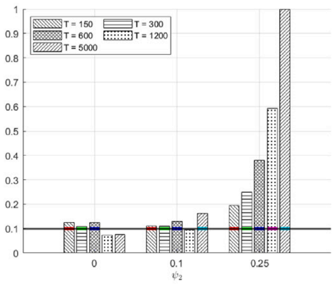
(c) $\psi_3 = 0.8$



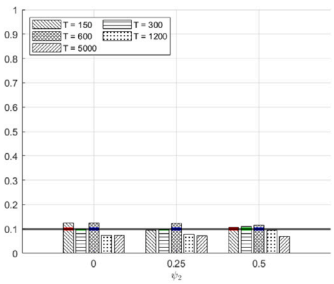
(d) $\psi_3 = 0.01$



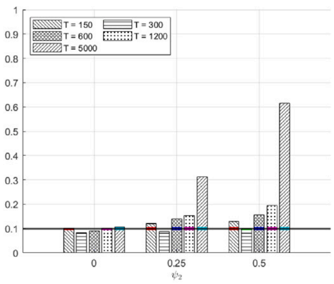
(e) $\psi_3 = 0.1$



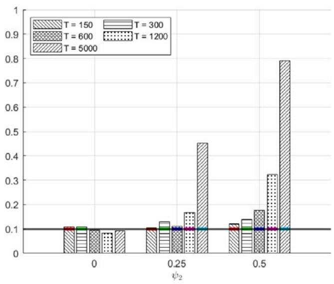
(f) $\psi_3 = 0.05$



(g) $\psi_3 = 1$



(h) $\psi_3 = 0.1$



(i) $\psi_3 = 0$

Fig. D.4. Relative rejection frequencies for DGP1 when proxy equation is non-linear according to *NL 1* (top panel), *NL 2* (middle panel), and *NL 3* (bottom panel). Nominal significance level 10%. $p = 0$.

The results are shown in Fig. D.4 (bottom panel). The figure shows that our test leads to the correct nominal level if the proxy is exogenous, i.e. $\psi_2 = 0$, irrespective of the non-linearity of the proxy. For endogenous proxy variables with $\psi_2 \neq 0$, we find that the power increases with the degree of non-linearity, the sample size, and proxy endogeneity.

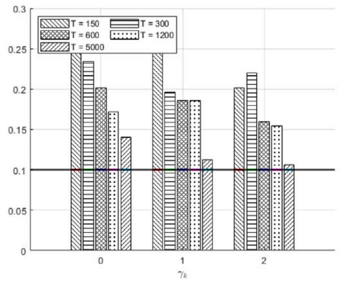
D.4. DGP1 with multiple proxies for multiple shocks

To investigate the test's performance when multiple shocks are identified using multiple proxies, the following Monte Carlo simulation uses DGP1 from the main text. However, we simulate two proxy variables with

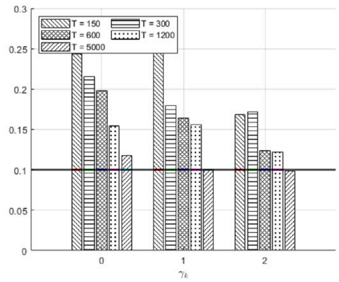
$$z_{1t} = \psi_1 \varepsilon_{1t} + \frac{1}{2}\psi_1 \varepsilon_{2t} + \psi_2 \varepsilon_{3t} + v_{1t} \quad \text{and} \quad z_{2t} = \frac{1}{2}\psi_1 \varepsilon_{1t} + \psi_1 \varepsilon_{2t} + \psi_2 \varepsilon_{3t} + v_{2t}$$

where $v_{1t}, v_{2t} \sim N(0, 1)$.

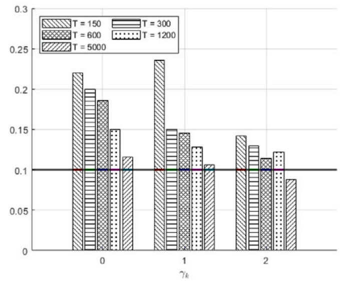
As before, ψ_1, ψ_2 are chosen to achieve a desired correlation of the proxies with the target and contaminating shocks, i.e. $corr(z_{1t}, \varepsilon_{1t}) = corr(z_{2t}, \varepsilon_{2t}) = [0.4, 0.5, 0.6]$ and $corr(z_{1t}, \varepsilon_{3t}) = corr(z_{2t}, \varepsilon_{3t}) = [0, -0.5, -0.6]$. Note that z_{1t} is designed to be more



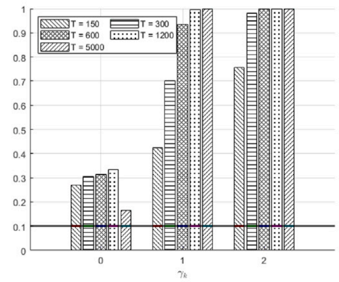
(a) $corr(z_t, \varepsilon_{1t}) = 0.4$, $corr(z_t, \varepsilon_{3t}) = 0$



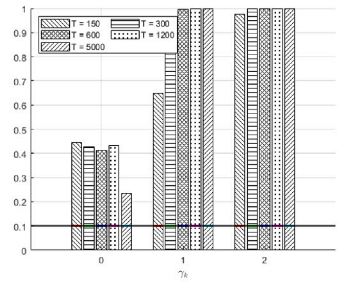
(b) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{3t}) = 0$



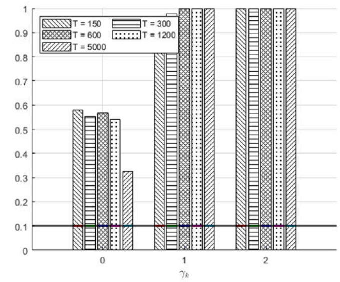
(c) $corr(z_t, \varepsilon_{1t}) = 0.6$, $corr(z_t, \varepsilon_{3t}) = 0$



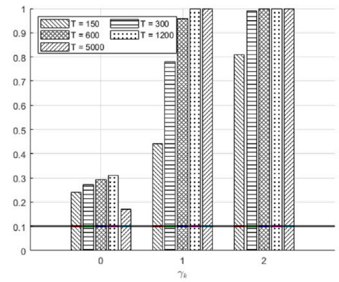
(d) $corr(z_t, \varepsilon_{1t}) = 0.4$, $corr(z_t, \varepsilon_{3t}) = -0.5$



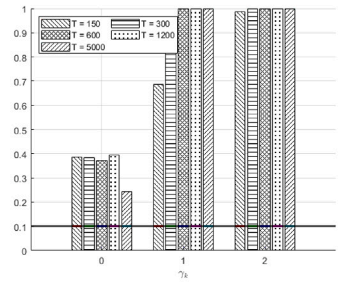
(e) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{3t}) = -0.5$



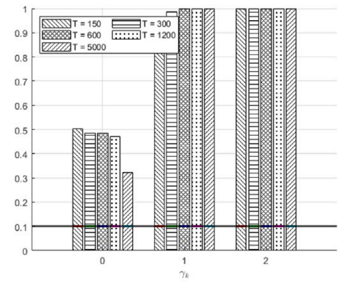
(f) $corr(z_t, \varepsilon_{1t}) = 0.6$, $corr(z_t, \varepsilon_{3t}) = -0.5$



(g) $corr(z_t, \varepsilon_{1t}) = 0.4$, $corr(z_t, \varepsilon_{3t}) = -0.6$



(h) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{3t}) = -0.6$



(i) $corr(z_t, \varepsilon_{1t}) = 0.6$, $corr(z_t, \varepsilon_{3t}) = -0.6$

Fig. D.5. Relative rejection frequencies for DGP1 with two proxies. Nominal significance level 10%. $p = 0$.

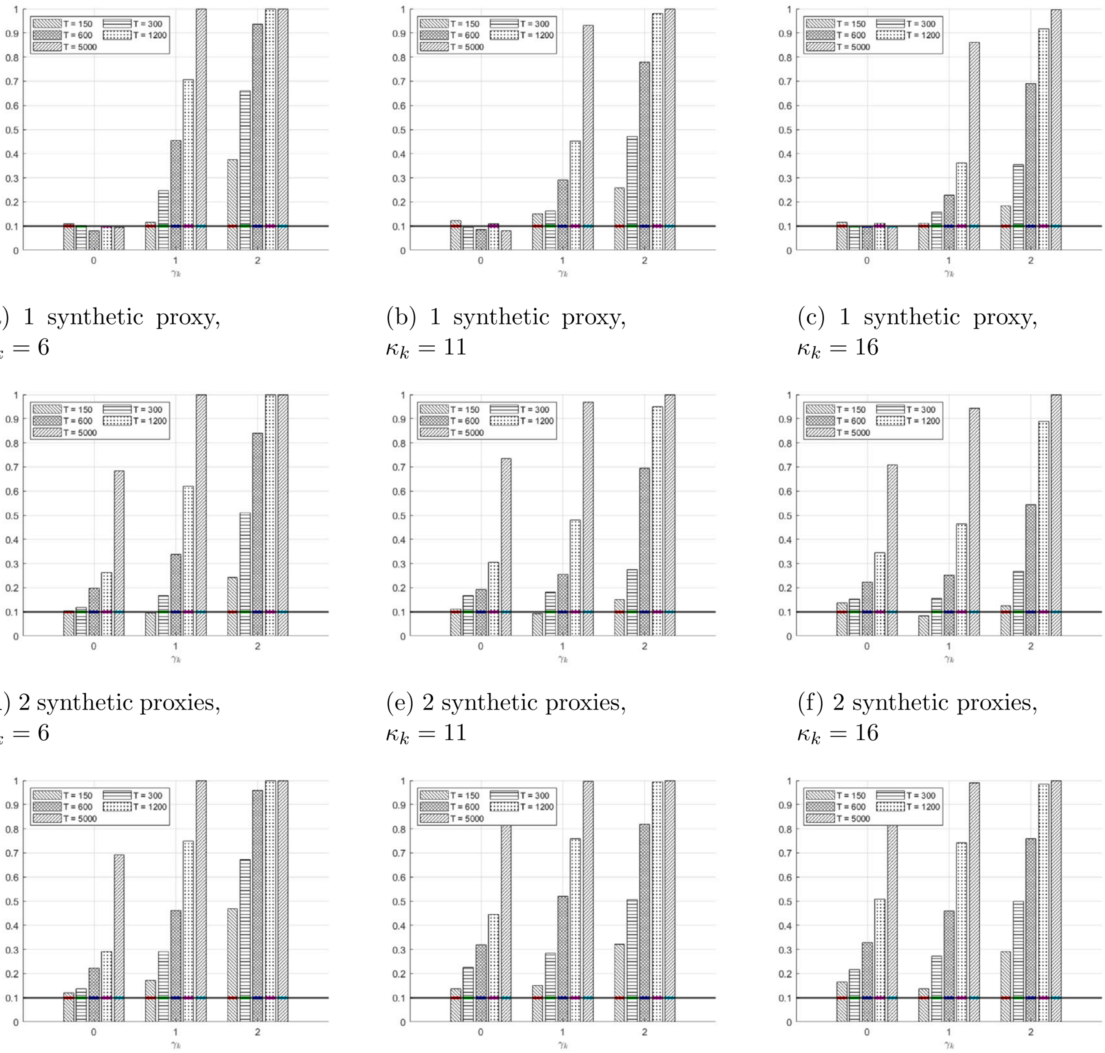
strongly correlated with ε_{1t}, while z_{1t} is designed to be more strongly correlated with ε_{1t}. Both proxies are contaminated for non-zero values of ψ_2.

Fig. D.5 shows that for the case of two proxies and two shocks our test performs equally well as in the baseline in terms of level and power.

D.5. DGP1 with multiple synthetic proxies and approximations by Donald et al. (2003)

To investigate the test's performance when a single shock is identified using multiple synthetic proxies, we repeat the simulation for DGP1 with varying degrees of kurtosis $\kappa_k = (6, 11, 16)$. We investigate the original setup (Fig. D.6(a)–(c)), a setup with two synthetic proxies, z_t^2 and z_t^3, (Fig. D.6(d)–(f)), the approximation by Donald et al. (2003) using a fixed approximation order (Fig. D.6(g)–(h)), and the approximation by Donald et al. (2003) using an approximation order, which rises with the sample size at an appropriate rate (Fig. D.7(a)–(c)).

We confirm that in these cases, the test has power even for non-skewed shocks, as long as they have excess kurtosis compared to a Gaussian distribution. The power rises with the degree of kurtosis, as expected. In this sense, using higher-order synthetic proxies can be beneficial. On the other hand, the additional moment conditions will add noise so that a trade-off arises and the choice of the number of synthetic proxies is application-specific. None of the methods is clearly dominated, but we conclude that the Donald et al.

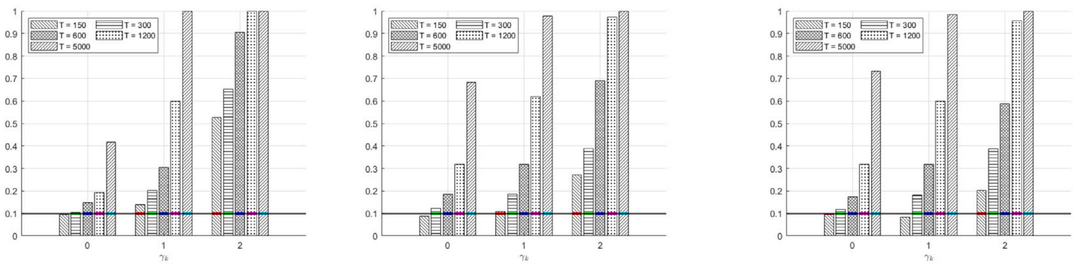


(a) 1 synthetic proxy, $\kappa_k = 6$ (b) 1 synthetic proxy, $\kappa_k = 11$ (c) 1 synthetic proxy, $\kappa_k = 16$

(d) 2 synthetic proxies, $\kappa_k = 6$ (e) 2 synthetic proxies, $\kappa_k = 11$ (f) 2 synthetic proxies, $\kappa_k = 16$

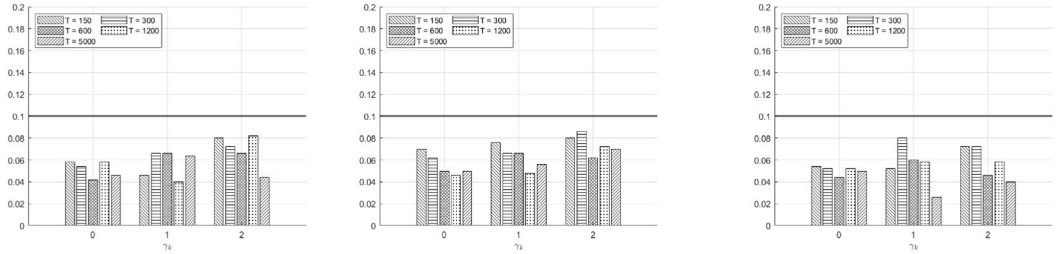
(g) Donald et al. approximation, $\kappa_k = 6$ (h) Donald et al. approximation, $\kappa_k = 11$ (i) Donald et al. approximation, $\kappa_k = 16$

Fig. D.6. Relative rejection frequencies for DGP1 with one synthetic proxy (top row), two synthetic proxies (middle row) and Donald et al. approximation with fixed approximation order (bottom row). $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.2$. Nominal significance level 10%. $p = 0$.



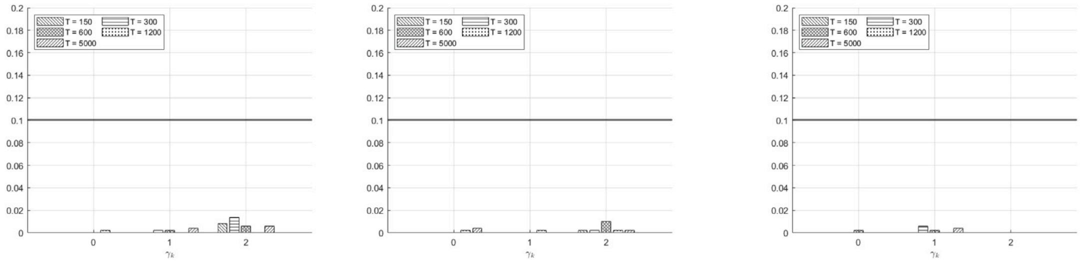
(a) $\kappa_k = 6$ (b) $\kappa_k = 11$ (c) $\kappa_k = 16$

Fig. D.7. Relative rejection frequencies for DGP1 with Donald et al. approximation and flexible approximation order. $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.2$. Nominal significance level 10%. $p = 0$.



(a) $p = 0$ (b) $p = 1$ (c) $p = 12$

Fig. D.8. Relative rejection frequencies for DGP1 with irrelevant proxy, $corr(z_t, \varepsilon_{1,t}) = corr(z_t, \varepsilon_{2,t}) = 0$. Nominal significance level 10%.



(a) $p = 0$ (b) $p = 1$ (c) $p = 12$

Fig. D.9. Relative rejection frequencies for DGP1 with irrelevant proxy, $corr(z_t, \varepsilon_{1,t}) = corr(z_t, \varepsilon_{2,t}) = 0$. Nominal significance level 10%. Continuous updating estimator with conservative critical values.

(2003) approximation with sample size-dependent approximation leads to slightly lower power in our simulation setup compared to the alternatives.

D.6. Irrelevant proxies and weak proxy robust critical values

Fig. D.8 displays the rejection rates of the exogeneity test in setups with an irrelevant proxy. The results indicate a conservative test behavior under irrelevant proxies. Fig. D.9 shows the rejection rates of the test based on the CUE with weak proxy robust critical values. The test again behaves conservatively, which is in line with the theoretical results in Section 2.2. Fig. D.10 uses DGP1 from the main text, however, the test now uses the CUE with weak proxy robust critical values. We find that the conservative test leads to a notable loss of power in small samples.

D.7. DGP2 details

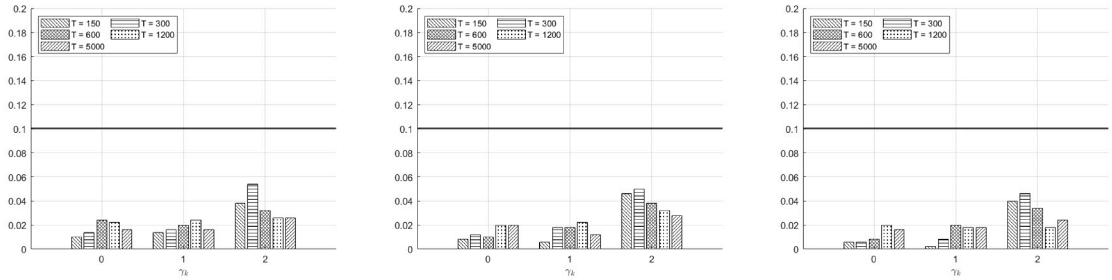
To obtain parameters for DGP2 we estimate a VAR(1) model with constant, but without trends or time dummies for the variables in Mertens and Ravn (2014) to obtain the following parameters:

$$A_1 = \begin{bmatrix} 0.930 & 0.053 & -0.055 \\ -0.063 & 0.770 & 0.250 \\ -0.014 & -0.043 & 1.045 \end{bmatrix}$$

The maximum Eigenvalue of the associated companion form is 0.9975 indicating a very persistent but stable process. We then follow Mertens and Ravn (2014)'s identification approach and assume that all shocks are uncorrelated with unit variance and restrict the simultaneous impact of output shocks on government spending to zero to obtain $B = \begin{bmatrix} 0.004 & 0 & 0.028 \\ 0.013 & 0.026 & 0.000 \\ -0.005 & 0.008 & 0.002 \end{bmatrix}$.

Structural shocks are generated with the strongest non-Gaussian specification from DGP1 using $\gamma_k = -2\ \forall\ k$ and $\kappa_k = \gamma_k^2 + 2 = 6$ $\forall\ k$, mimicking moments of the proxy in Mertens and Ravn (2014).

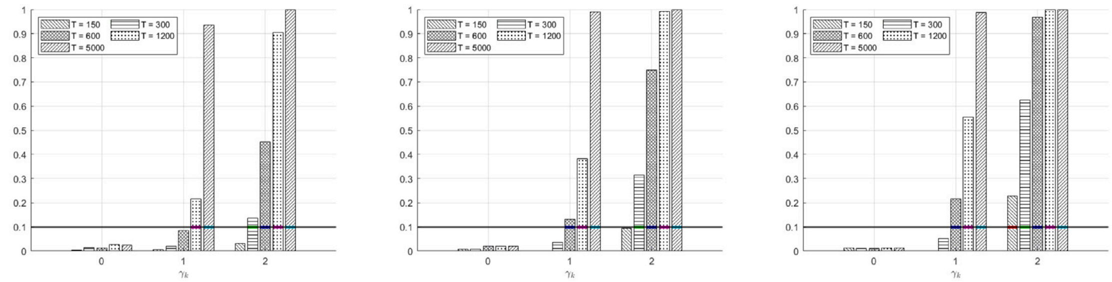
The proxy z_t is generated from Eq. (9) with $corr(\varepsilon_{1t}, z_t) = 0.26$ as in Mertens and Ravn (2014). We investigate $corr(\varepsilon_{2t}, z_t) = (0, -0.1, -0.2)$, which is informed by the finding in Keweloh et al. (2023b) that the tax proxy in Mertens and Ravn (2014) has a correlation of -0.2 with the output shock, suggesting some degree of endogeneity.



(a) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = 0$

(b) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = 0$

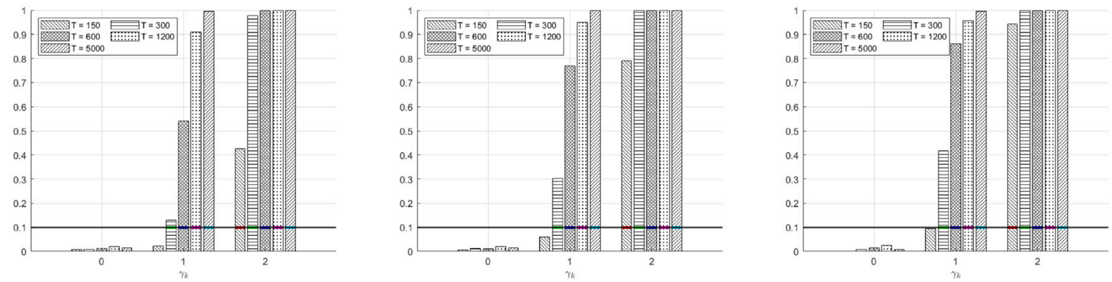
(c) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = 0$



(d) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.2$

(e) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.2$

(f) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.2$

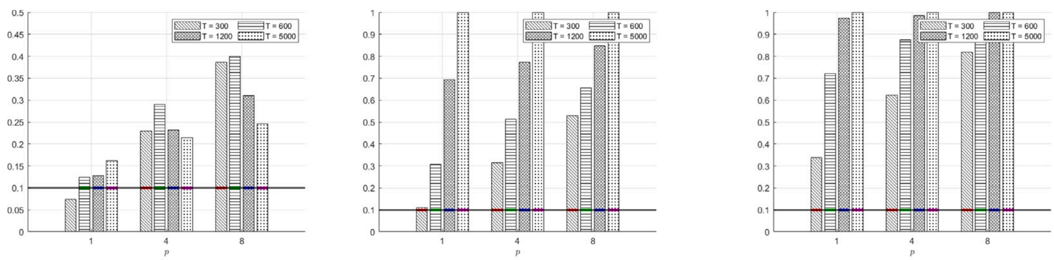


(g) $corr(z_t, \varepsilon_{1t}) = 0.5$, $corr(z_t, \varepsilon_{2t}) = -0.5$

(h) $corr(z_t, \varepsilon_{1t}) = 0.7$, $corr(z_t, \varepsilon_{2t}) = -0.5$

(i) $corr(z_t, \varepsilon_{1t}) = 0.9$, $corr(z_t, \varepsilon_{2t}) = -0.5$

Fig. D.10. Relative rejection frequencies for DGP1. Nominal significance level 10%. $p = 0$. Continuous updating estimator with robust critical values.

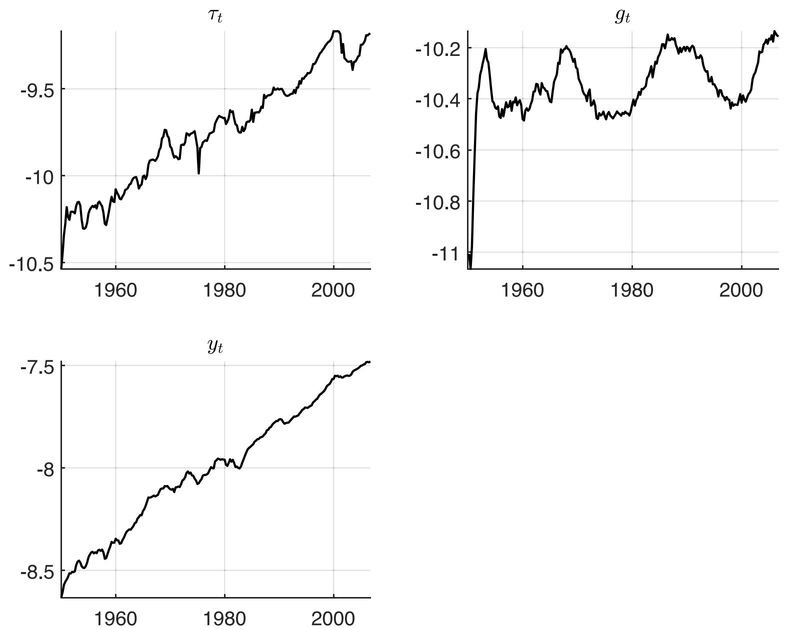


(a) $corr(\varepsilon_{2t}, z_t) = 0$

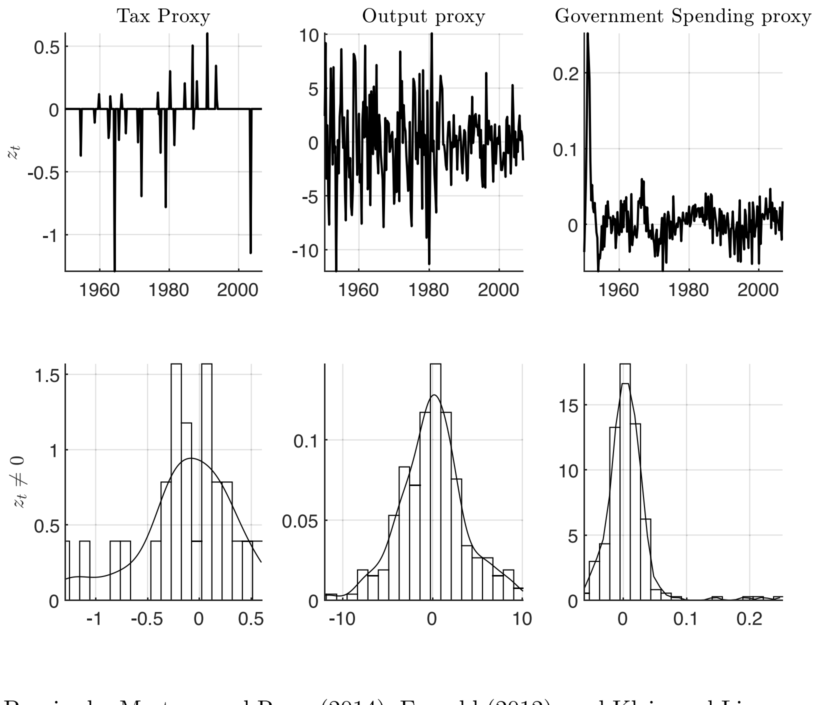
(b) $corr(\varepsilon_{2t}, z_t) = -0.1$

(c) $corr(\varepsilon_{2t}, z_t) = -0.2$

Fig. D.11. Relative rejection frequencies for DGP2 (1-step test). Nominal significance level 10%.



(a) Mertens and Ravn (2014) variables



(b) Proxies by Mertens and Ravn (2014), Fernald (2012), and Klein and Linnemann (2019). Time series plot (top panels) together with histogram and kernel density estimate (bottom panel).

Fig. E.12. Data.

Table E.1
Proxy exogeneity test (extended).

Proxies	Estimator	J-statistic	p-Value	p-Value robust
$z_{\tau,t}, z_{\tau,t}^2$	GMM	5.48	0.06	–
$z_{\tau,t}, z_{\tau,t}^2$	CUE	6.44	0.04	0.17
$z_{\tau,t}, z_{\tau,t}^2, z_{\tau,t}^3$	GMM	7.25	0.12	–
$z_{\tau,t}, z_{\tau,t}^2, z_{\tau,t}^3$	CUE	7.57	0.11	0.27
$z_{g,t}, z_{g,t}^2$	GMM	0.68	0.71	–
$z_{g,t}, z_{g,t}^2$	CUE	0.76	0.68	0.94
$z_{g,t}, z_{g,t}^2, z_{g,t}^3$	GMM	3.42	0.49	–
$z_{g,t}, z_{g,t}^2, z_{g,t}^3$	CUE	3.67	0.45	0.72
$z_{y,t}, z_{y,t}^2$	GMM	1.59	0.45	–
$z_{y,t}, z_{y,t}^2$	CUE	1.60	0.45	0.81
$z_{y,t}, z_{y,t}^2, z_{y,t}^3$	GMM	2.84	0.58	–
$z_{y,t}, z_{y,t}^2, z_{y,t}^3$	CUE	2.65	0.62	0.85

Note: The table shows the results of different two-stage proxy exogeneity tests. The J-statistics are based on the GMM estimator or the CUE estimator. The column denoted by p-Value robust uses the conservative critical values from a χ_r^2 distribution which are robust to weak proxies.

Appendix E. Application

See Table E.1 and Fig. E.12

Appendix F. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2024.105876.

References

Angelini, G., Fanelli, L., 2019. Exogenous uncertainty and the identification of structural vector autoregressions with external instruments. J. Appl. Econometrics 34 (6), 951–971.
Angelini, G., Fanelli, L., Neri, L., 2024. Invalid proxies and volatility changes. arXiv preprint arXiv:2403.08753.
Bierens, H.J., 1982. Consistent model specification tests. J. Econometrics 20 (1), 105–134.
Braun, R., Brüggemann, R., 2022. Identification of SVAR models by combining sign restrictions with external instruments. J. Bus. Econom. Statist. 41 (4), 1077–1089.
Bruns, M., Lütkepohl, H., 2023. An alternative bootstrap for proxy vector autoregressions. Comput. Econ. 62 (4), 1857–1882.
Bruns, M., Lütkepohl, H., 2024. Heteroskedastic proxy vector autoregressions: An identification-robust test for time-varying impulse responses in the presence of multiple proxies. J. Econom. Dynam. Control 104837.
Caldara, D., Kamps, C., 2017. The analytics of SVARs: a unified framework to measure fiscal multipliers. Rev. Econ. Stud. 84 (3), 1015–1040.
Chamberlain, G., 1987. Asymptotic efficiency in estimation with conditional moment restrictions. J. Econometrics 34 (3), 305–334.
Donald, S.G., Imbens, G.W., Newey, W.K., 2003. Empirical likelihood estimation and consistent tests with conditional moment restrictions. J. Econometrics 117 (1), 55–93.
Erickson, T., Whited, T.M., 2002. Two-step GMM estimation of the errors-in-variables model using high-order moments. Econometric Theory 18 (3), 776–799.
Fernald, J.G., 2012. A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. Working Paper Series 2012–19, Federal Reserve Bank of San Francisco.
Gertler, M., Karadi, P., 2015. Monetary policy surprises, credit costs, and economic activity. Am. Econ. J.: Macroecon. 7 (1), 44–76.
Gouriéroux, C., Monfort, A., Renne, J.-P., 2017. Statistical inference for independent component analysis: Application to structural VAR models. J. Econometrics 196 (1), 111–126.
Guay, A., 2021. Identification of structural vector autoregressions through higher unconditional moments. J. Econometrics 225 (1), 27–46.
Hall, A.R., 2005. Generalized Method of Moments. Oxford University Press.
Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econ.: J. Econ. Soc. 1029–1054.
Jarociński, M., Karadi, P., 2020. Deconstructing monetary policy surprises—the role of information shocks. Am. Econ. J.: Macroecon. 12 (2), 1–43.
Känzig, D.R., 2021. The macroeconomic effects of oil supply news: Evidence from OPEC announcements. Amer. Econ. Rev. 111 (4), 1092–1125.
Keweloh, S.A., 2021. A generalized method of moments estimator for structural vector autoregressions based on higher moments. J. Bus. Econom. Statist. 39 (3), 772–782.
Keweloh, S.A., 2024. Uncertain short-run restrictions and statistically identified structural vector autoregressions. arXiv preprint arXiv:2303.13281v2.
Keweloh, S.A., Hetzenecker, S., Seepe, A., 2023a. Monetary policy and information shocks in a block-recursive SVAR. J. Int. Money Finance 137, 102892.
Keweloh, S.A., Klein, M., Prüser, J., 2023b. Estimating fiscal multipliers by combining statistical identification with potentially endogenous proxies. arXiv preprint arXiv:2302.13066.
Klein, M., Linnemann, L., 2019. Tax and spending shocks in the open economy: Are the deficits twins? Eur. Econ. Rev. 120, 103300.
Lewbel, A., 1997. Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D. Econ.: J. Econ. Soc. 1201–1213.
Lewis, D.J., 2021. Identifying shocks via time-varying volatility. Rev. Econ. Stud. 88 (6), 3086–3124.
Lütkepohl, H., Schlaak, T., 2022. Heteroscedastic proxy vector autoregressions. J. Bus. Econom. Statist. 40 (3), 1268–1281.
Matteson, D.S., Tsay, R.S., 2017. Independent component analysis via distance covariance. J. Amer. Statist. Assoc. 112 (518), 623–637.
Mertens, K., Ravn, M.O., 2014. A reconciliation of SVAR and narrative estimates of tax multipliers. J. Monetary Econ. 68, S1–S19.
Newey, W.K., West, K.D., 1994. Automatic lag selection in covariance matrix estimation. Rev. Econ. Stud. 61 (4), 631–653.
Romer, C.D., Romer, D.H., 2010. The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. Amer. Econ. Rev. 100, 763–801.
Schlaak, T., Rieth, M., Podstawski, M., 2023. Monetary policy, external instruments, and heteroskedasticity. Quant. Econ. 14 (1), 161–200.
Stock, J.H., Wright, J.H., 2000. GMM with weak identification. Econometrica 68 (5), 1055–1096.
Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. J. Bus. Econom. Statist. 20 (4), 518–529.