DIW BERLIN

1940

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2021

Qualitative versus Quantitative External Information for Proxy Vector Autoregressive Analysis

Lukas Boer and Helmut Lütkepohl

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Qualitative versus Quantitative External Information for Proxy Vector Autoregressive Analysis

by

Lukas Boer[1]

DIW Berlin and Humboldt-Univeristät Berlin, Mohrenstr. 58, 10117 Berlin, Germany

email: lboer@diw.de

and

Helmut Lütkepohl

DIW Berlin and Freie Universität Berlin, Mohrenstr. 58, 10117 Berlin, Germany

email: hluetkepohl@diw.de

February 15, 2021

Abstract. A major challenge for proxy vector autoregressive analysis is the construction of a suitable external instrument variable or proxy for identifying a shock of interest. Some authors construct sophisticated proxies that account for the dating and size of the shock while other authors consider simpler versions that use only the dating and signs of particular shocks. It is shown that such qualitative (sign-)proxies can lead to impulse response estimates of the impact effects of the shock of interest that are nearly as efficient as or even more efficient than estimators based on more sophisticated quantitative proxies that also reflect the size of the shock. Moreover, the sign-proxies tend to provide more precise impulse response estimates than an approach based merely on the higher volatility of the shocks of interest on event dates.

Key Words: GMM, heteroskedastic VAR, instrumental variable estimation, proxy VAR, structural vector autoregression.

JEL classification: C32

[1]An earlier version of the paper was titled *A Simple Instrument for Proxy Vector Autoregressive Analysis* (DIW Discussion Papers 1905). We thank Jonathan Wright for making the Matlab code for his analysis available and we are grateful to Max Diegel, Lukas Menkhoff, Malte Rieth, Thore Schlaak, Leopold Zessner-Spitzenberg and two anonymous referees for helpful comments on an earlier draft of the paper. The paper was presented at the Econometric Society European Winter Meeting and at seminars at Humboldt-Universität Berlin and Freie Universität Berlin. We thank the participants for helpful comments.

1 Introduction

In structural vector autoregressive (VAR) analysis, external information in the form of instrumental variables is often used to identify the shocks of interest. The external information is sometimes only available in qualitative form. The goal of this study is to investigate the possible loss in estimation efficiency for the structural parameters due to having only qualitative rather than quantitative information.

Proxy VAR analysis has become quite popular lately, mainly as it does not require economically problematic timing or exclusion restrictions on the behavior of variables and the resulting impulse responses are often more in line with economic theory than responses stemming from timing restrictions. In proxy VAR models, all variables are free to respond simultaneously to all structural shocks (e.g., Stock and Watson (2012), Mertens and Ravn (2013) Gertler and Karadi (2015), Kilian and Lütkepohl (2017, Chapter 15)). However, the construction of a suitable proxy may be a main challenge because it requires additional information from sources external to the model.

Given the difficulties in constructing suitable quantitative instruments, proxies considered in this context sometimes use only qualitative information on the shock of interest. For example, Budnik and Rünstler (2020) construct a qualitative proxy for macroprudential policy shocks that takes on positive or negative values in periods where a change in capital requirements has occurred, depending on the sign of the corresponding shock, and it is zero in periods with no known policy shocks. Likewise, the Romer and Romer (1989) dummy is an indicator for monetary policy shocks that has been used as an instrument variable in an early proxy VAR study by Beaudry and Saito (1998). More generally, Plagborg-Møller and Wolf (2021, Appendix B.3) point out that narrative sign restrictions of the type considered in the literature can be used to construct instrument variables for proxy VARs that assume values ± 1 if a positive or negative shock occurs at the dates of special events, respectively, and that are zero otherwise. All that is needed to construct this type of sign-proxy is knowledge of the dates of the special events and the signs of the possible shocks that may have occurred at the event dates. Such knowledge is available for a number of shocks that have been used in structural VAR analysis. For example, some crises in the Middle East are known to have caused disruptions in oil supply. Such information can be employed to construct a sign-proxy for identifying oil supply shocks. Likewise, there are a number of events such as the 9/11 attacks on the US that have caused increases in economic uncertainty and could be used for constructing a sign-proxy to identify uncertainty shocks (see also Carriero et al. (2015) for a related proposal).

In this study, we compare the estimation precision of impulse responses based on quantitative versus sign-proxies in a frequentist setting.[2] Through simulation, we show that, in terms of root mean squared error (RMSE), sign-proxies may yield more precise estimates of the impact effects of the structural shocks of interest than conventional, more sophisticated quantitative proxies that are not strongly correlated with the shock of interest. Moreover, sign-proxies may yield confidence intervals for impulse responses which have a coverage and width of similar size as quantitative proxies.

If the dates of specific events for the emission of shocks are known but the size and the sign of the shocks is unknown, Wright (2012) proposes to utilize potential changes of the volatility of the shocks on event dates for identifying and estimating the impact effects of the shock. For example, a monetary policy shock may be more volatile at dates of central bank council meetings. Using the additional moment conditions obtained from the heteroskedasticity for estimation, only the dates of the specific events must be known and there is no need to construct a proper instrument variable associated with the specific events. As a drawback, the approach requires more restrictive assumptions for the variances of the structural shocks than the proxy VAR approach, which we discuss in Section 2. We also compare the estimates of the Wright approach to the proxy VAR and sign-proxy estimates for the impact effects of the shock of interest. The Wright estimates turn out to be less efficient in terms of RMSE than estimators based on quantitative proxies or sign-proxies. Moreover, the approach tends to yield wider confidence intervals for impulse responses than its competitors.

In Wright's approach and in the standard proxy VAR approach, the impact effects of the shock of interest and, hence, its impulse responses are typically estimated by the generalized method of moments (GMM). As using more moment conditions may lead to more efficient GMM estimators and since the proxy VAR approach and the heteroskedasticity approach for estimating the impact effects of the shock of interest use different sets of moment conditions, one may conjecture that combining the moment conditions may lead to efficiency gains in estimating the impact effects of the shock. Therefore we also consider this combination approach. It turns out, however, that in the present situation, the combination approach does not lead to uniform improvements of estimation efficiency for the impulse responses. We use a model from Wright (2012) to illustrate the alternative approaches in the context of an empirical example.

The remainder of the paper is structured as follows. In the next section,

[2]Budnik and Rünstler (2020) report related work that compares standard proxies with sign-proxies (sparse narrative instruments in their terminology) in a Bayesian setting.

the general model setup is presented and the different estimators of the impulse responses are discussed. In Section 3, a Monte Carlo experiment is conducted to compare the small sample performance of the estimators. In Section 4, an illustrative example based on a model due to Wright (2012) is presented. Conclusions and extensions are discussed in the final section. Additional simulation results as well as more details on the computational methods used are available in an Online Appendix which accompanies this article.

2 The Structural VAR Setup

A K-dimensional reduced-form VAR model

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t \tag{1}$$

is considered, where u_t is a zero mean, serially uncorrelated white noise process which may be heteroskedastic. A sample y_t, $t \in \mathcal{T} = \{1, \ldots, T\}$, and all required presample values are assumed to be available for inference.

The structural errors, $w_t = (w_{1t}, \ldots, w_{Kt})'$, are obtained from the reduced-form errors by a linear transformation $w_t = B^{-1}u_t$ or $u_t = Bw_t$ such that the structural matrix B represents the impact effects of the structural shocks on the variables y_t. The structural errors are assumed to be instantaneously uncorrelated, i.e., w_t has a diagonal covariance matrix.

Denoting by $\mathbf{b}_i$ the i^{th} column of B, i.e., $B = [\mathbf{b}_1, \ldots, \mathbf{b}_K]$, the vector $\mathbf{b}_i$ contains the impact effects of the i^{th} structural shock. As the ordering of the observed variables and the shocks can be changed arbitrarily, it is assumed that, without loss of generality, the first structural shock, w_{1t}, is the shock of specific interest and it is normalized such that it has a unit impact effect on the first variable. This assumption is just a normalization and does not entail a loss of generality if the impact effect of the first shock on the first variable is nonzero.

If $\mathbf{b}_1$ and the reduced-form VAR parameters from equation (1) are known, the responses of the variables y_t to the first structural shock can be traced over time using the relations

$$\theta_h = \Phi_h \mathbf{b}_1, \quad h = 1, 2, \ldots,$$

where θ_h is a $(K \times 1)$ vector of structural impulse responses for propagation horizon h and $\Phi_i = \sum_{j=1}^{i} \Phi_{i-j} A_j$ can be obtained recursively for $h = 0, 1, \ldots,$ from the VAR slope coefficients starting with $\Phi_0 = I_K$ and defining $A_j = 0$ for $j > p$ (Lütkepohl, 2005, Section 2.1.2). Because estimation of the reduced-form VAR model in expression (1) and, hence, of the estimation of the Φ_i is straightforward, the focus of interest in the following is precise estimation of the structural parameters $\mathbf{b}_1$. Denoting the last $K-1$ elements of $\mathbf{b}_1$ by $\mathbf{b}_{12}$ such that $\mathbf{b}_1' = (1, \mathbf{b}_{12}')$, we focus on estimators of $\mathbf{b}_{12}$ in the following.

2.1 The Proxy VAR Approach

The proxy VAR approach assumes that a suitable instrument is available that can be used for estimating $\mathbf{b}_{12}$. Suppose there is a proxy variable z_t satisfying

$$\mathbb{E}(w_{1t}z_t) = c \neq 0 \quad \text{(relevance)}, \tag{2}$$

$$\mathbb{E}(w_{kt}z_t) = 0, \quad k = 2, \dots, K \quad \text{(exogeneity)}. \tag{3}$$

In that case, a multiple of $\mathbf{b}_1$ can be estimated by using z_t as an instrument and observing that $\mathbb{E}(u_t z_t) = c\mathbf{b}_1$. This setup assumes, however, that the covariance $\mathbb{E}(u_t z_t)$ is time-invariant. In particular, it is not affected by heteroskedasticity of u_t. If $\mathbb{E}(u_t z_t)$ is time-invariant, the relations in (2) and (3) provide moment conditions for estimating a multiple of $\mathbf{b}_1$. An estimator for $\mathbf{b}_1$ may then be obtained by dividing all elements of the estimator of $c\mathbf{b}_1$ by the first element.

More precisely, a GMM estimator is obtained by estimating the reduced-form VAR in expression (1) by equation-wise ordinary least squares (OLS) or some other suitable estimation method and using the proxy VAR estimator

$$\widehat{\mathbf{b}}_{12}^P = \left(\frac{\sum_{t=1}^T \hat{u}_{2t} z_t}{\sum_{t=1}^T \hat{u}_{1t} z_t}, \dots, \frac{\sum_{t=1}^T \hat{u}_{Kt} z_t}{\sum_{t=1}^T \hat{u}_{1t} z_t} \right)' \tag{4}$$

for $\mathbf{b}_{12}$. Here the $\hat{u}_t = (\hat{u}_{1t}, \dots, \hat{u}_{Kt})'$ are the estimated reduced-form errors.

The instrument may be suggested by the subject matter. For example, Piffer and Podstawski (2018) use changes in the price of gold to construct an instrument for uncertainty shocks and Cesa-Bianchi, Thwaites and Vicondoa (2020) construct a time series of intra-day price variation of the 3-month Sterling future contracts around policy decisions of the Monetary Policy Committee of the Bank of England as a proxy for monetary policy shocks.

Note that this approach is also typically used in the proxy VAR literature if the data exhibit changes in volatility (see, e.g., Mertens and Ravn (2013), Piffer and Podstawski (2018), Cesa-Bianchi et al. (2020), Dias and Duarte (2019), Gertler and Karadi (2015), Alessi and Kerssenfischer (2019)). In that case, heteroskedasticity robust methods are often used for inference. This approach implicitly implies that heteroskedasticity does not affect the covariance between the proxy and the reduced-form residuals, as pointed out by Lütkepohl and Schlaak (2020), where also further discussion of the issue can be found. We also make this assumption here whenever heteroskedastic reduced-form errors are considered.

Finding a good proxy that satisfies conditions (2) and (3) for a shock of interest is not always easy. This problem is reflected in the fact that, in some studies, the instrument is only available for a shorter period than the sample period of the other variables in the model (see, e.g., Gertler and Karadi (2015)). In many cases proxies are constructed by considering only certain dates of announcements or special events, where specific shocks were transmitted. Suppose there are M event dates $\mathcal{T}_1 = \{t_{a_1}, \dots, t_{a_M}\}$, then one may construct a simple dummy variable

$$s_t = \begin{cases} \operatorname{sgn}(w_{1t}) \cdot 1 & \text{for} \quad t \in \mathcal{T}_1, \\ 0 & \text{for} \quad t \notin \mathcal{T}_1, \end{cases} \tag{5}$$

where $\operatorname{sgn}(\cdot)$ denotes the sign function which assigns the sign of its argument. Thus, s_t assumes a value of $+1$ or -1 depending on whether the special event induces a positive or negative shock. In periods without known special events, $s_t = 0$.[3] Note that, if the shocks $w_{1t}, \dots, w_{Kt}$ are stochastically independent, e.g., if they are Gaussian as sometimes assumed in the literature, s_t will also be independent of $w_{2t}, \dots, w_{Kt}$ (see, e.g., Mood, Graybill and Boes (1974, p. 151, Theorem 3)). Hence, s_t qualifies as a proxy (see also the discussion of such proxies in Plagborg-Møller and Wolf (2021, Appendix B.3)).

Constructing the sign-proxy, s_t, in this way assumes that the investigator at least knows whether a positive or negative shock was induced in a specific period. It may actually not always be clear which shocks are positive and which ones are negative because the sign of a shock is typically linked to some economic variable. For example, a positive (expansionary) monetary policy shock is often associated with a reduction in the policy interest rate. However, this indicator for monetary policy shocks is not available in times of zero interest rates. For those periods, some other indicator is needed to determine whether a policy shock is positive or negative. For example, expansions in bond purchases may be linked to expansionary monetary policy shocks. Of course, the sign of the shock may be inferred from a set of variables. The important precondition for constructing the sign-proxy s_t

[3]For a coherent stochastic framework one may model the event dates as being generated by i.i.d. Bernoulli random variables D_t which are independent of w_{1t} and for which the probability $\mathbb{P}(D_t = 1) = d$ and $\mathbb{P}(D_t = 0) = 1 - d$, i.e., $D_t \sim B(d)$ with $0 < d \leq 1$. Then the sign-proxy can be represented as $s_t = \operatorname{sgn}(w_{1t})D_t$. In our exposition we condition on a specific realization of event dates because the event dates are typically given in practice.

is that the researcher knows the date and the sign of the structural shock on specific event dates.

It is also important to note that the signs of some proxy z_t and a sign-proxy s_t for the same shock may differ. As both proxies are correlated with the first structural shock, w_{1t}, z_t and s_t will also be correlated with each other. This does not mean, however, that they always share the same sign.

Actually, constructing the dummy with reverse signs makes no difference for our purposes. We may define

$$s_t = \begin{cases} -\text{sgn}(w_{1t}) \cdot 1 & \text{for} \quad t \in \mathcal{T}_1, \\ 0 & \text{for} \quad t \notin \mathcal{T}_1, \end{cases}$$

instead of using the s_t defined in expression (5). The crucial property of s_t is that it has to be correlated with w_{1t}. Whether or not the correlation is positive or negative is not important for its usefulness as an instrument.

In fact, we also consider the situation where a researcher does not know the sign of the shock for some event dates for sure and, hence, may occasionally assign signs incorrectly. We denote by $s_t^{(m)}$ a sign-proxy for which m signs are classified incorrectly during the sample period. Clearly, misspecifying some signs may undermine the correlation between the shock of interest and the sign-proxy and, hence, it may weaken the proxy as an instrument for the shock. An alternative strategy would be to drop all event dates where the sign of the shock is uncertain. In other words, the sign-proxy may only assume nonzero values for a subset of $\mathcal{T}_1$. In the following, we denote the sign-proxy by $s_t^{(-)}$ if 10% of the event dates are dropped.

The sign-proxy in (5) is clearly related to the specific shocks induced by the special events and, thus, can be used just like a regular proxy to estimate the impact effects of the first shock, w_{1t}. Hence, the associated sign-proxy estimator is

$$\widehat{\mathbf{b}}_{12}^{SP} = \left(\frac{\sum_{t=1}^{T} \hat{u}_{2t} s_t}{\sum_{t=1}^{T} \hat{u}_{1t} s_t}, \dots, \frac{\sum_{t=1}^{T} \hat{u}_{Kt} s_t}{\sum_{t=1}^{T} \hat{u}_{1t} s_t} \right)'. \tag{6}$$

If s_t is replaced by $s_t^{(m)}$ or $s_t^{(-)}$, the corresponding estimator is denoted by $\widehat{\mathbf{b}}_{12}^{SP(m)}$ or $\widehat{\mathbf{b}}_{12}^{SP(-)}$, respectively.

2.2 Wright's Heteroskedasticity Approach

Wright (2012) proposes another approach to estimate the impact effects of the shock of interest based on the potential change in volatility on event dates. His approach has the advantage that just the dates of special events or announcements have to be known. No proxy variable has to be constructed for the event dates. Wright assumes that there are two volatility regimes associated with covariance matrices Σ_0 and Σ_1. The matrix Σ_0 is the usual covariance and it changes to Σ_1 during the M event periods, i.e., when $t \in \mathcal{T}_1 = \{t_{a_1}, \dots, t_{a_M}\}$. In other words, $\mathbb{E}(u_t u_t') = \Sigma_0$ for $t \notin \mathcal{T}_1$ and $\mathbb{E}(u_t u_t') = \Sigma_1$ for $t \in \mathcal{T}_1$.

Recall that the first structural shock, w_{1t}, is the shock of specific interest and it is normalized such that it has a unit impact effect on the first variable. We now assume in addition that the other shocks, $w_{2t}, \dots, w_{Kt}$, are normalized such that they have unit variances. Then it is easy to see that w_t is a white noise process with zero mean and covariance matrix

$$\Sigma_0^w = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & I_{K-1} \end{bmatrix} \text{ for } t \in \mathcal{T} \backslash \mathcal{T}_1 \quad \text{and} \quad \Sigma_1^w = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & I_{K-1} \end{bmatrix} \text{ for } t \in \mathcal{T}_1. \quad (7)$$

Note that this setup, where only the variance of the first structural shock changes, is not very common in the structural VAR literature. It is, however, the basis for Wright's heteroskedasticity approach to work. In contrast, the proxy VAR approach works under more general assumptions. Specifically, it also works for homoskedastic shocks and for more general heteroskedsticity.

Denoting by $\mathbf{b}_i$ the i^{th} column of B, as before, the relation $u_t = Bw_t = \sum_{k=1}^K \mathbf{b}_k w_{kt}$ implies, under the present assumptions, that

$$\Sigma_0 = \sigma_0^2 \mathbf{b}_1 \mathbf{b}_1' + \sum_{k=2}^K \mathbf{b}_k \mathbf{b}_k' \quad \text{and} \quad \Sigma_1 = \sigma_1^2 \mathbf{b}_1 \mathbf{b}_1' + \sum_{k=2}^K \mathbf{b}_k \mathbf{b}_k'.$$

Hence,

$$\Sigma_1 - \Sigma_0 = c_W \mathbf{b}_1 \mathbf{b}_1', \quad (8)$$

where $c_W = \sigma_1^2 - \sigma_0^2$.

Wright (2012) suggests using the moment conditions related to the variance change in (8) for GMM estimation. More precisely, he proposes to minimize a GMM objective function analogous to

$$J_W(\mathbf{b}_{12}, c_W) =$$
$$\text{vech}(\widehat{\Sigma}_1 - \widehat{\Sigma}_0 - c_W \mathbf{b}_1 \mathbf{b}_1')' \left(\frac{\widehat{\Omega}_0}{T-M} + \frac{\widehat{\Omega}_1}{M} \right)^{-1} \text{vech}(\widehat{\Sigma}_1 - \widehat{\Sigma}_0 - c_W \mathbf{b}_1 \mathbf{b}_1'),$$

with respect to the last $K-1$ elements of $\mathbf{b}_1$ and c_W. Here

$$\widehat{\Omega}_0 = \frac{1}{T-M} \sum_{t \in \mathcal{T} \backslash \mathcal{T}_1} \text{vech}\left(\hat{u}_t \hat{u}_t' - \overline{\hat{u}\hat{u}'}\right) \text{vech}\left(\hat{u}_t \hat{u}_t' - \overline{\hat{u}\hat{u}'}\right)'$$

and

$$\widehat{\Omega}_1 = \frac{1}{M} \sum_{t \in \mathcal{T}_1} \text{vech}\left(\hat{u}_t \hat{u}_t' - \overline{\hat{u}\hat{u}'}\right) \text{vech}\left(\hat{u}_t \hat{u}_t' - \overline{\hat{u}\hat{u}'}\right)'.$$

The resulting estimator of $\mathbf{b}_{12}$ is denoted by $\widehat{\mathbf{b}}_{12}^W$.

It is important to note that, although the moment conditions in equation (8) and in (2)/(3) both specify a multiple of $\mathbf{b}_1$, they may not specify the same multiple of $\mathbf{b}_1$. In other words, c and c_W may be distinct.

2.3 Joint GMM

As more moment conditions may improve the efficiency of GMM estimators, it may make sense to consider the joint moment conditions

$$m(\mathbf{b}_{12}, c, c_W) = \begin{bmatrix} \mathbb{E}(u_t z_t) - c\mathbf{b}_1 \\ \text{vech}(\Sigma_1 - \Sigma_0 - c_W \mathbf{b}_1 \mathbf{b}_1') \end{bmatrix} = 0$$

or to consider these moments relying on the sign-proxy, provided, of course, that the conditions underlying Wright's approach are satisfied and assuming that the proxy satisfies the conditions for heteroskedastic models. Under these conditions, using

$$\overline{\hat{u}z} = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_t z_t \quad \text{and} \quad \overline{\hat{u}s} = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_t s_t$$

as estimators of $\mathbb{E}(u_t z_t)$ and $\mathbb{E}(u_t s_t)$, respectively, the GMM estimator is obtained by minimizing the objective function

$$J_{gen}(\mathbf{b}_{12}, c, c_W) = \widehat{m}(\mathbf{b}_{12}, c, c_W)' \widehat{\Omega}_m^{-1} \widehat{m}(\mathbf{b}_{12}, c, c_W),$$

where $\widehat{\Omega}_m$ is a suitable weighting matrix. For example, for the proxy z_t, using a block-diagonal weighting matrix, a GMM objective function

$$J(\mathbf{b}_{12}, c, c_W) = T\left(\overline{\hat{u}z} - c\mathbf{b}_1\right)' \widehat{\Omega}_{uz}^{-1} \left(\overline{\hat{u}z} - c\mathbf{b}_1\right) + J_W(\mathbf{b}_{12}, c_W), \tag{9}$$

may be considered, where

$$\widehat{\Omega}_{uz} = \frac{1}{T} \sum_{t=1}^{T} \left(\hat{u}_t z_t - \overline{\hat{u}z}\right) \left(\hat{u}_t z_t - \overline{\hat{u}z}\right)'.$$

Equivalently, one could minimize the objective function $J_{gen}(\mathbf{b}_{12}, c, c_W)$ with weighting matrix

$$\widehat{\Omega}_m = \begin{bmatrix} \frac{\widehat{\Omega}_{uz}}{T} & 0 \\ 0 & \frac{\widehat{\Omega}_0}{T-M} + \frac{\widehat{\Omega}_1}{M} \end{bmatrix}.$$

The resulting estimator is denoted by $\widehat{\mathbf{b}}_{12}^{WP}$. Likewise, if the proxy z_t is replaced by the sign-proxy s_t, the corresponding estimator is denoted by $\widehat{\mathbf{b}}_{12}^{WSP}$.

In the next section, the estimators $\widehat{\mathbf{b}}_{12}^{P}$, $\widehat{\mathbf{b}}_{12}^{SP}$, $\widehat{\mathbf{b}}_{12}^{SP(-)}$, $\widehat{\mathbf{b}}_{12}^{SP(m)}$, $\widehat{\mathbf{b}}_{12}^{W}$, $\widehat{\mathbf{b}}_{12}^{WP}$, and $\widehat{\mathbf{b}}_{12}^{WSP}$ of the impact effects of the first structural shock are compared in a simulation study.

3 Monte Carlo Investigation of Estimator Efficiency

As we suspect that the sample size, the lag order and dimension of the VAR process as well as the strength of the proxy and/or the difference in the variances on event dates and non-event dates may have an impact on the small sample properties of the different estimators, we have looked at different data generating processes (DGPs), proxies and event dates to study the small sample properties of the estimators. It turned out that alternative DGPs yielded qualitatively similar results. Therefore we focus on the results obtained for a specific DGP first and then we summarize various other results at the end of this section.

3.1 Monte Carlo Design for DGP1

We focus on representative results for a DGP that is based on the empirical example of Wright (2012). The model is a 6-dimensional VAR(1) process. The precise parameter values are given in the Online Appendix, Section A.2. The eigenvalues of the coefficient matrix have a maximum modulus of 0.994 and, hence, the DGP is stable but very persistent with several autoregressive roots very close to the unit circle. We refer to this DGP as DGP1 and use samples of different sizes generated with this VAR(1) DGP. VAR processes with the true lag order $p = 1$ are fitted in all simulations for DGP1.

The impact effects matrix B is constructed using the Cholesky decomposition of the estimated residual covariance matrix, $T^{-1}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t'$, and dividing all elements by the upper left-hand element such that the first column of B has a first element equal to one, as assumed in Section 2.

Wright (2012, Table 5) also constructs an instrument z_t for a monetary policy shock related to $M = 28$ event days. He uses the first principal component of a set of bond futures traded at the Chicago Mercantile Exchange on the announcement days and constructs a proxy for monetary policy shocks on that basis. We construct our proxy such that it has similar properties using

$$z_t = \begin{cases} w_{1t} + v_t & \text{for} \quad t \in \mathcal{T}_1, \\ 0 & \text{for} \quad t \notin \mathcal{T}_1, \end{cases} \tag{10}$$

where v_t is a measurement error that is assumed to be independent of w_{1t}. This type of proxy is also assumed in other proxy VAR studies (e.g., Caldara and Herbst (2019), Lütkepohl and Schlaak (2020)).

Note that the correlation between the proxy and the first structural shock, $\text{corr}(z_t, w_{1t}) = \sigma_1/\sqrt{\sigma_1^2 + \sigma_v^2}$ for $t \in \mathcal{T}_1$, depends on the variance σ_v^2 of v_t. The implied correlation for the full sample also depends on the fraction of event dates in the sample. More precisely, $\text{corr}(z_t, w_{1t}) = \sqrt{M/T}\sigma_1/\sqrt{\sigma_1^2 + \sigma_v^2}$ for $t \in \mathcal{T}$.[4] Hence, the strength of the instrument will depend on σ_v^2, M and T. In the simulations, we use Gaussian v_t, i.e., $v_t \sim \mathcal{N}(0, \sigma_v^2)$, and choose values for σ_v^2 such that the correlation between z_t and w_{1t} is either 0.9 or 0.7 on event dates, $t \in \mathcal{T}_1$.[5] The former value represents high correlation and is chosen to obtain a strong instrument while the second value of 0.7 results in a weaker instrument, as we will see in the simulations where also tests for strong instruments are reported. Values for T and M used in the simulations are such that $\sqrt{M/T}$ ranges from 0.2 to 0.45. Hence, for a correlation of 0.7 on $\mathcal{T}_1$, we get correlations between 0.14 to 0.315 for the full sample. Such values are rather low and allow us to see the implications of using a weak instrument. The higher correlation of 0.9 on $\mathcal{T}_1$ corresponds to the correlation obtained for the benchmark study discussed in Section 4 (see Table 5).

As we suspect that the performance of the different estimators depends to some extent on the difference between the variances in the two regimes (σ_0^2 and σ_1^2), we vary σ_1^2 in constructing our proxy. More precisely, we construct the first shock as

$$w_{1t} = \begin{cases} \mathcal{N}(0, \sigma_1^2) & \text{for} \quad t \in \mathcal{T}_1, \\ \mathcal{N}(0, \sigma_0^2) & \text{for} \quad t \in \mathcal{T} \setminus \mathcal{T}_1, \end{cases} \tag{11}$$

where we set $\sigma_0^2 = 1$ and assign either the value 4 or the value 10 to σ_1^2. Thus, the standard deviations in the more volatile regimes are 2 and 3.16. In other words, on event dates the shocks are either twice or about three times as volatile as in other periods. The values of σ_1^2 are in the range of what

[4]The correlation is obtained by considering Bernoulli random variables D_t with probabilities $\mathbb{P}(D_t = 1) = M/T$ and $\mathbb{P}(D_t = 0) = 1 - M/T$ and evaluating the correlation between $D_t z_t$ and w_{1t} under the assumption that z_t and D_t are stochastically independent.

[5]Specifically, we use $\sigma_v^2 = 1$ and 2 for $\sigma_1^2 = 4$ and 10, respectively, resulting in a correlation of about 0.9 and we use $\sigma_v^2 = 4$ and 10 for $\sigma_1^2 = 4$ and 10, respectively, which yields a correlation of about 0.7.

we estimated from the data for our benchmark study. More precisely, upon appropriately scaling the process such that $\sigma_0^2 = 1$, we get a value of $\sigma_1^2 = 6.68$ which is in the range of the values we are using in the simulations. The other shocks are simulated as standard normal, i.e., $(w_{2t}, \dots, w_{Kt}) \sim \mathcal{N}(0, I_{K-1})$.

The instruments z_t and s_t are constructed based on w_{1t} as in (10) and (5), respectively. For s_t we also allow for the possibility that some shocks are classified incorrectly as positive or negative. As mentioned earlier, we denote the corresponding estimator by $\widehat{\mathbf{b}}_{12}^{SP(m)}$, where m signifies the number of incorrect sign assignments used for the sign-proxy.

We use different numbers of event dates, M, for different sample sizes T. The precise values of M are given in the tables with simulation results. The event dates are chosen randomly.[6]

In addition, we consider the possibility that, for some event dates, the sign of the shock is unknown and therefore the corresponding date is not treated as an event date. In that case an incorrect zero is assigned to the sign-proxy. In our simulations we consider the possibility that (roughly) 10% of the event dates are ignored and denote the corresponding estimator by $\widehat{\mathbf{b}}_{12}^{SP(-)}$.

The performance criteria for comparing different estimators are linked to the last $K-1$ elements of $\mathbf{b}_1$, i.e., to $\mathbf{b}_{12}$. We consider the root mean squared errors (RMSEs) of the estimators for these elements relative to the RMSE of $\widehat{\mathbf{b}}_{12}^{P}$. To compute the relative RMSE of an estimator such as $\widehat{\mathbf{b}}_{12}^{SP}$, we divide the RMSE of each element of $\widehat{\mathbf{b}}_{12}^{SP}$ by the RMSE of the corresponding element of $\widehat{\mathbf{b}}_{12}^{P}$ and then add these relative RMSEs to get the relative RMSE of the whole estimated vector $\widehat{\mathbf{b}}_{12}^{SP}$. Thereby we control for differences in the RMSEs of the individual elements of a vector. For each simulation design, the experiment is repeated 5,000 times.

As impulse response analysis is a standard tool in structural VAR studies, we also compare coverage rates and interval widths of confidence intervals for impulse responses implied by the alternative estimators for selected simulation designs. The confidence intervals are constructed by a moving-block bootstrap (MBB) as proposed by Jentsch and Lunsford (2019) who show that it works under general conditions for inference for impulse responses in proxy VARs. We use a block length of $\ell = 5.03T^{1/4}$ which is the rule of thumb proposed by Jentsch and Lunsford (2019). As DGP1 is very persistent, we use bias-corrected OLS estimators in the MBB to improve the coverage properties of the bootstrap intervals in small samples (see Kilian (1998), Kilian and Lütkepohl (2017, Section 12.3)). Details of our MBB implementation are

[6] For each repetition of the simulation experiment we reorder the integers $1, \dots, T$ with the Matlab function `randperm(.)` and choose the first M values of the reordered integers for the set of event dates, $\mathcal{T}_1$.

provided in the Online Appendix. In our simulations involving bootstraps, we use 1000 bootstrap replications and 1000 Monte Carlo replications.

3.2 Simulation Results for DGP1

Before looking at the relative RMSEs of the estimators, it may be worth assessing the strengths of the different instruments. This should give a first indication concerning the suitability of the different estimators for empirical exercises.

3.2.1 Instrument Strength

We report the relative frequencies of heteroskedasticity-robust F-statistics for weak instruments being smaller than 10 in our simulations (see Stock, Wright and Yogo (2002) or Kilian and Lütkepohl (2017, Section 15.2.1)).[7] A threshold value of 10 is typically used in the related literature to classify an instrument as being sufficiently strong. Since our data are heteroskedastic, we use a robust F-statistic. In Table 1 the relative frequencies of test values below 10 are reported for the different proxies.

As discussed earlier, the strength of a proxy as an instrument depends on its correlation with the shock of interest. Therefore, it is not surprising that $z_t^{(.9)}$, which denotes the instrument having correlation 0.9 with the first shock on event dates, produces F-values less frequently below 10 than the corresponding instrument $z_t^{(.7)}$, which has only correlation 0.7 with the first shock. It is also not surprising that both proxies tend to yield fewer F-values below 10 when the number of event dates, M, is greater and when the variance σ_1^2 of the first shock on event dates is larger.

Interestingly, there are fewer F-values below 10 for the sign-proxy, s_t, than for the proxy $z_t^{(.7)}$ in all cases considered in Table 1. Taking into account the generating mechanism of the proxy variable in equation (10), this outcome is not implausible because for $\text{corr}(z_t, w_{1t}) = 0.7$, the sign of z_t differs from that of w_{1t} for a number of the simulated values, whereas the sign of s_t is always the same as that of w_{1t} (see equation (5)) and, hence, s_t becomes a stronger instrument than $z_t^{(.7)}$. This holds even if s_t is replaced by $s_t^{(-)}$ which is based on 10% fewer event dates than $z_t^{(.7)}$. Moreover, s_t has only slightly more F-values below 10 than the proxy $z_t^{(.9)}$. For $M = 25$ event dates or more, or a sample size of $T = 500$, both $z_t^{(.9)}$ and s_t never yield F-values

[7] The robust F-statistic is the standard heteroskedasticity-robust statistic for testing the coefficient of z_t to be zero in a regression of the first OLS estimated reduced-form error $\hat{u}_{1t}$ on a constant and the instrument z_t. The statistic corresponds to the effective F-statistic of Montiel Olea and Pflueger (2013).

Table 1: Relative Frequencies of Heteroskedasticity-Robust Weak Instrument F-test Statistics Smaller than 10 (in %)

			Proxy variables						
Sample size T	M	σ_1^2	$z_t^{(.9)}$	$z_t^{(.7)}$	s_t	$s_t^{(-)}$	$s_t^{(1)}$	$s_t^{(3)}$	$s_t^{(5)}$
100	20	4	0.1	15.2	0.1	0.3	5.8	56.3	94.0
		10	0.0	14.5	0.0	0.1	4.9	54.5	94.0
250	10	4	3.7	37.0	6.1	10.2	62.4	99.7	100.0
		10	1.9	36.4	5.4	9.1	61.4	99.8	100.0
	25	4	0.0	7.6	0.0	0.0	0.4	19.0	66.0
		10	0.0	7.3	0.0	0.0	0.3	17.8	65.3
500	25	4	0.0	7.2	0.0	0.0	0.5	18.1	67.7
		10	0.0	7.1	0.0	0.0	0.5	17.7	67.5
	50	4	0.0	0.2	0.0	0.0	0.0	0.0	0.0
		10	0.0	0.2	0.0	0.0	0.0	0.0	0.0

Note: M denotes the number of event dates, σ_1^2 the variance of the structural shock of interest on event dates, $z_t^{(.9)}$ a proxy with a theoretical correlation of 0.9 with the structural shock of interest on event dates, $z_t^{(.7)}$ a proxy with a theoretical correlation of 0.7 with the structural shock on event dates, s_t the sign-proxy, $s_t^{(-)}$ a sign-proxy with 10% omitted signs, and $s_t^{(m)}$ denotes a sign-proxy with m incorrectly specified signs.

below 10. In other words, in a usual proxy VAR analysis, the sign-proxy would be classified as a strong instrument more often than the proxy $z_t^{(.7)}$ and almost as often as the proxy $z_t^{(.9)}$, which is based on quantitative rather than just qualitative information.

The situation changes if some of the signs are assigned incorrectly, as can be seen by looking at the frequencies reported for $s_t^{(1)}$, $s_t^{(3)}$, and $s_t^{(5)}$ in Table 1. For instance, for $T = 250$, even the proxy $z_t^{(.7)}$ produces F-values below 10 in only about 37% of the replications for $M = 10$ event periods, whereas a sign-proxy with 3 incorrect signs out of 10, $s_t^{(3)}$, results in nearly all F-values below the threshold of 10. On the other hand, when there are many event periods with nonzero proxies, the sign-proxy maintains its strong instrument status even with a few incorrect signs. For example, for $T = 250$ and $M = 25$ event dates, $s_t^{(3)}$ is only classified as a weak instrument by the F-test in fewer than 20% of the replications of our simulations. Not surprisingly, the sign-proxy is not a useful instrument if half of the signs are incorrectly specified, as can be seen by looking at the frequencies associated with $s_t^{(5)}$ when there are only $M = 10$ event periods. In that case, the robust F-statistic is always below 10. Comparing the F-test results for s_t and $s_t^{(-)}$ shows that reducing the number of event dates leads to a weaker instrument.

In summary, using standard diagnostics, the sign-proxy may well be classified as a strong instrument as often as a conventional quantitative proxy even if the standard proxy has quite high correlation with the shock of interest. If the proxy is not strongly correlated with the shock, then the sign-proxy clearly dominates under this criterion. Even if some signs are misspecified, the F-test may classify the sign-proxy as a strong instrument more often than the standard proxy if there are sufficiently many event dates. Since the sign-proxy with correctly specified signs typically does not have a weak instrument problem, we do not consider weak instrument robust methods, as discussed by Montiel Olea, Stock and Watson (2020), for any of the estimators for better comparability.[8]

3.2.2 RMSEs

The absolute RMSEs for the two quantitative proxy estimators $\widehat{\mathbf{b}}_{12}^{P(.9)}$ and $\widehat{\mathbf{b}}_{12}^{P(.7)}$ are presented in Table 2, where it can be seen that the estimation precision declines substantially (the RMSEs increase), especially for sample sizes $T = 100$ and 250 if $z_t^{(.7)}$ is used instead of $z_t^{(.9)}$. In fact, some rather large RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.7)}$ indicate that this estimator is not reliable for some of the simulation designs and yields some outlying estimates far away from the true values (e.g., for $T = 250$, $M = 10$, $\sigma_1^2 = 4$).

Table 3 reports the RMSEs of the different estimators for the impact effects of the shocks relative to the corresponding RMSEs of the estimator $\widehat{\mathbf{b}}_{12}^{P(.9)}$. Thus, the relative RMSEs in Table 3 refer to a highly correlated standard proxy, $z_t^{(.9)}$, which is rarely classified as a weak instrument by the F-tests reported in Table 1. Since the quantitative proxy uses the richest information, we expect it to provide potentially more precise estimates and, therefore, use it as a benchmark for the other estimators.

Indeed, all of the relative RMSEs presented in Table 3 for more than 10 event dates ($M > 10$) are larger than 1, meaning that the RMSEs are larger than the corresponding ones of $\widehat{\mathbf{b}}_{12}^{P(.9)}$. However, the sign-proxy estimator $\widehat{\mathbf{b}}_{12}^{SP}$ has relative RMSEs very close to one such that the efficiency loss relative to the quantitative proxy is limited. Only if there are relatively few event dates ($M = 10$) and, hence, even $\widehat{\mathbf{b}}_{12}^{P(.9)}$ is rather unreliable, the relative RMSE of

[8] We have also computed F-tests without heteroskedasticity adjustment because they may be used in practice if heteroskedasticity is not accounted for. The results are shown in Table A.1 of the Online Appendix and convey a similar picture regarding relative instrument strength as the F-values underlying Table 1. In general, however, the non-adjusted F-values are often larger and thus lead to fewer values below 10.

Table 2: Absolute RMSEs of Proxy Estimators $\widehat{\mathbf{b}}_{12}^{P(.9)}$ and $\widehat{\mathbf{b}}_{12}^{P(.7)}$

			absolute RMSEs	
Sample size T	M	σ_1^2	$\widehat{\mathbf{b}}_{12}^{P(.9)}$	$\widehat{\mathbf{b}}_{12}^{P(.7)}$
100	20	4	0.170	0.707
		10	0.105	0.224
250	10	4	0.266	3.181
		10	0.440	0.839
	25	4	0.143	0.200
		10	0.089	0.126
500	25	4	0.141	0.192
		10	0.087	0.121
	50	4	0.097	0.124
		10	0.060	0.079

Note: M denotes the number of event dates and σ_1^2 is the variance of the structural shock of interest on event dates. The RMSE is calculated as $\sqrt{\frac{1}{R}\sum_{r=1}^{R}(\widehat{\mathbf{b}}_{12,r}^{P(.)} - \mathbf{b}_{12})'(\widehat{\mathbf{b}}_{12,r}^{P(.)} - \mathbf{b}_{12})}$, where R denotes the number of Monte Carlo replications.

the sign-proxy estimator is even smaller than or equal to one in two cases. In other words, it happens for cases where the number of event periods is very small and the quantitative proxy is occasionally classified as a weak instrument by the F-tests in Table 1.

Looking at the columns of $\widehat{\mathbf{b}}_{12}^{SP(-)}$, $\widehat{\mathbf{b}}_{12}^{SP(1)}$, $\widehat{\mathbf{b}}_{12}^{SP(3)}$, and $\widehat{\mathbf{b}}_{12}^{SP(5)}$ in Table 3, it is apparent that, if the number of event dates is reduced or some of the signs are assigned incorrectly, the efficiency of the sign-proxy estimator deteriorates, although the increase in relative RMSEs is rather moderate in some cases if 10% of the event dates are omitted or only one sign is incorrect. However, note that, not surprisingly, the sign-proxy is no longer useful as an instrument if 50%, or even 30%, of the signs are incorrect (see the RMSEs for $\widehat{\mathbf{b}}_{12}^{SP(3)}$ and $\widehat{\mathbf{b}}_{12}^{SP(5)}$ when $M = 10$). Additionally, the F-tests in Table 1 indicate that such proxies are not sufficiently strong instruments for proper inference. In other words, if only a small number of event dates is available and only for a subset the sign of the shock of interest is known with certainty, using a sign-proxy is not a good idea.

Looking at the RMSEs of Wright's heteroskedasticity estimator $\widehat{\mathbf{b}}_{12}^{W}$, it is seen that they are usually much worse than those of the estimator based on the quantitative proxy. In some cases, the RMSE of $\widehat{\mathbf{b}}_{12}^{W}$ is more than twice as large as that of $\widehat{\mathbf{b}}_{12}^{P(.9)}$, except for $T = 250$, $M = 10$, $\sigma_1^2 = 10$, where $\widehat{\mathbf{b}}_{12}^{P(.9)}$ is also a rather unreliable estimator. For all other designs reported in Table 3,

Table 3: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.9)}$

			relative RMSEs							
T	M	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(1)}$	$\widehat{\mathbf{b}}_{12}^{SP(3)}$	$\widehat{\mathbf{b}}_{12}^{SP(5)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	20	4	1.09	1.15	1.22	1.66	52.66	2.01	1.62	1.50
		10	1.11	1.18	1.26	1.70	24.03	1.72	1.38	1.34
250	10	4	1.00	1.07	1.41	408.92	881.77	1.66	1.62	1.60
		10	0.50	0.53	0.71	29.58	178.30	0.75	0.73	0.72
	25	4	1.09	1.17	1.19	1.48	1.98	2.18	1.74	1.57
		10	1.12	1.19	1.22	1.51	2.02	1.86	1.49	1.37
500	25	4	1.10	1.17	1.20	1.48	1.99	2.53	2.18	1.96
		10	1.12	1.20	1.23	1.52	2.04	2.34	1.98	1.76
	50	4	1.11	1.17	1.15	1.26	1.40	1.85	1.22	1.23
		10	1.13	1.19	1.18	1.29	1.43	1.37	1.10	1.13

Note: T signifies the sample size, M denotes the number of event dates and σ_1^2 is the variance of the structural shock of interest on event dates.

the Wright estimator is considerably less precise than the quantitative proxy estimator and it is also much less efficient than the sign-proxy estimator. In some cases, its relative RMSE is even larger than those of the corresponding sign-proxy estimators with fewer event dates or some incorrectly assigned signs. Note, however, that the Wright heteroskedasticity estimator improves with increasing variance σ_1^2 of the event dates relative to $\widehat{\mathbf{b}}_{12}^{P(.9)}$.

The estimation precision also improves if the moment conditions of Wright's heteroskedasticity estimator are combined with the moment conditions of the standard proxy or sign-proxy estimator (see the columns for $\widehat{\mathbf{b}}_{12}^{WP}$ and $\widehat{\mathbf{b}}_{12}^{WSP}$ in Table 3). However, the relative RMSEs are typically larger than one and, hence, these estimators are less precise than the standard proxy estimator and the relative RMSEs are usually also greater than those of the sign-proxy estimator. This outcome may seem somewhat surprising, given that $\widehat{\mathbf{b}}_{12}^{WP}$ and $\widehat{\mathbf{b}}_{12}^{WSP}$ are GMM estimators based on more moment conditions than $\widehat{\mathbf{b}}_{12}^{P}$ and $\widehat{\mathbf{b}}_{12}^{SP}$. The reason why $\widehat{\mathbf{b}}_{12}^{WP}$ and $\widehat{\mathbf{b}}_{12}^{WSP}$ are still less precise estimators may be that the GMM objective function for these estimators is quite nonlinear and difficult to optimize. The optimization algorithm may not always find the global minimum. In any case, if the sign-proxy is available, there is nothing or little to be gained from also including the Wright heteroskedasticity moment conditions.

In summary, the results in Table 3 show that, in many scenarios, the sign-proxy estimator is almost as precise as the standard proxy estimator

Table 4: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.7)}$

			relative RMSEs							
T	M	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(1)}$	$\widehat{\mathbf{b}}_{12}^{SP(3)}$	$\widehat{\mathbf{b}}_{12}^{SP(5)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	20	4	0.47	0.50	0.53	0.71	13.73	0.87	0.76	0.65
		10	0.58	0.62	0.65	0.88	11.54	0.89	0.80	0.70
250	10	4	0.08	0.09	0.12	35.62	66.76	0.14	0.14	0.13
		10	0.24	0.26	0.34	15.38	83.30	0.36	0.35	0.35
	25	4	0.79	0.84	0.86	1.07	1.43	1.57	1.34	1.13
		10	0.79	0.84	0.86	1.06	1.43	1.31	1.11	0.96
500	25	4	0.80	0.86	0.88	1.09	1.46	1.86	1.67	1.44
		10	0.81	0.86	0.88	1.09	1.46	1.68	1.51	1.26
	50	4	0.86	0.90	0.89	0.98	1.08	1.43	1.05	0.95
		10	0.86	0.90	0.89	0.98	1.08	1.04	0.90	0.85

Note: T signifies the sample size, M denotes the number of event dates and σ_1^2 is the variance of the structural shock of interest on event dates.

even if the standard proxy is highly correlated with the shock of interest. In contrast, the Wright heteroskedasticity estimator, which uses less information than the sign-proxy estimator, is less precise.

Looking at the results in Table 4 for a quantitative proxy with lower correlation of 0.7 with the shock of interest, it turns out that the sign-proxy estimator in this case is the most precise estimator. For all scenarios presented in Table 4, $\widehat{\mathbf{b}}_{12}^{SP}$ has the smallest or almost the smallest relative RMSEs. Even if 10% of the event dates are omitted or one sign is specified incorrectly, all the relative RMSEs of the corresponding estimators $\widehat{\mathbf{b}}_{12}^{SP(-)}$ and $\widehat{\mathbf{b}}_{12}^{SP(1)}$ are smaller than one and, hence, the estimators are more precise than $\widehat{\mathbf{b}}_{12}^{P(.7)}$. These results are quite plausible, given the finding in Table 1 that s_t, $s_t^{(-)}$ and $s_t^{(1)}$ are classified as strong instruments more often than $z_t^{(.7)}$ for a number of scenarios.

Even the Wright heteroskedasticity estimator still has relative RMSEs smaller than one if there are only $M = 10$ or 20 event dates in short samples. In contrast, the relative RMSEs of $\widehat{\mathbf{b}}_{12}^{W}$ are greater than one for all other scenarios. Thus, even a weaker instrument will typically result in better estimators than the Wright estimator in many cases. Again, combining Wright's moment conditions with those of the proxy or sign-proxy estimator improves the estimation precision but typically does not result in more precise estimators than the sign-proxy estimator.

Overall, the results in Tables 3 and 4 show that the sign-proxy estimator is very competitive with the quantitative proxy estimator even though it uses only qualitative information on the shock of interest. If the quantitative proxy is not strong, the sign-proxy estimator provides even more precise estimates (in terms of RMSE) than the quantitative proxy estimator. If the latter estimator is based on a strong instrument, the sign-proxy estimator is still almost as precise as the quantitative proxy estimator. Thus, given its limited information requirement, the sign-proxy estimator is an excellent choice for applied work if a strong quantitative proxy is difficult to construct. The sign-proxy estimator is clearly preferable to Wright's heteroskedasticity estimator if the event dates and the signs of the shock are known.

As we are using bias-adjusted VAR estimators for the following impulse response analysis, we have also computed RMSEs for estimators based on the corresponding reduced-form residuals and show them in Tables A.2 - A.4 in the Online Appendix. The results are very similar and convey the same message as the RMSEs based on OLS residuals.

3.2.3 Confidence Intervals for Impulse Responses

Although our primary interest is to compare the different estimators of the impact effects of a shock, it is also worth considering the implied impulse response estimators for larger propagation horizons, as they are often used for empirical analysis. In particular, confidence intervals for impulse responses are often considered in empirical studies. Therefore we have also investigated the coverage rates and interval widths of bootstrap confidence intervals for impulse responses up to propagation horizon $H = 20$ periods. Note that the impulse responses for all estimators are based on the same reduced-form parameter estimates and, hence, the same reduced-form impulse responses.

Figure 1 displays empirical coverage rates and interval widths for pointwise MBB confidence intervals with nominal 90% confidence level for the four estimators $\widehat{\mathbf{b}}_{12}^{SP}$, $\widehat{\mathbf{b}}_{12}^{P(.9)}$, $\widehat{\mathbf{b}}_{12}^{P(.7)}$, and $\widehat{\mathbf{b}}_{12}^{W}$. The sample size underlying the results in Figure 1 is $T = 500$, the number of event dates is $M = 50$ and the event date variance $\sigma_1^2 = 4$. The MBB is known to have poor properties for small sample sizes and, thus, a sample size of $T = 500$ may be needed for good coverage rates (see Lütkepohl and Schlaak (2019) for a study of a related issue). In the Online Appendix, Figures A.1, we present results for $T = 250$ which confirm the point.

In Figure 1 the coverage rates of all estimators are relatively similar and reasonably close to 90% for the impact effects for all variables and for three out of six variables this also holds for all propagation horizons. In contrast, for y_{1t}, y_{5t} and y_{6t}, coverage rates a bit below 80% are obtained for larger propagation horizons. Of course, this may be partly due to the MBB which

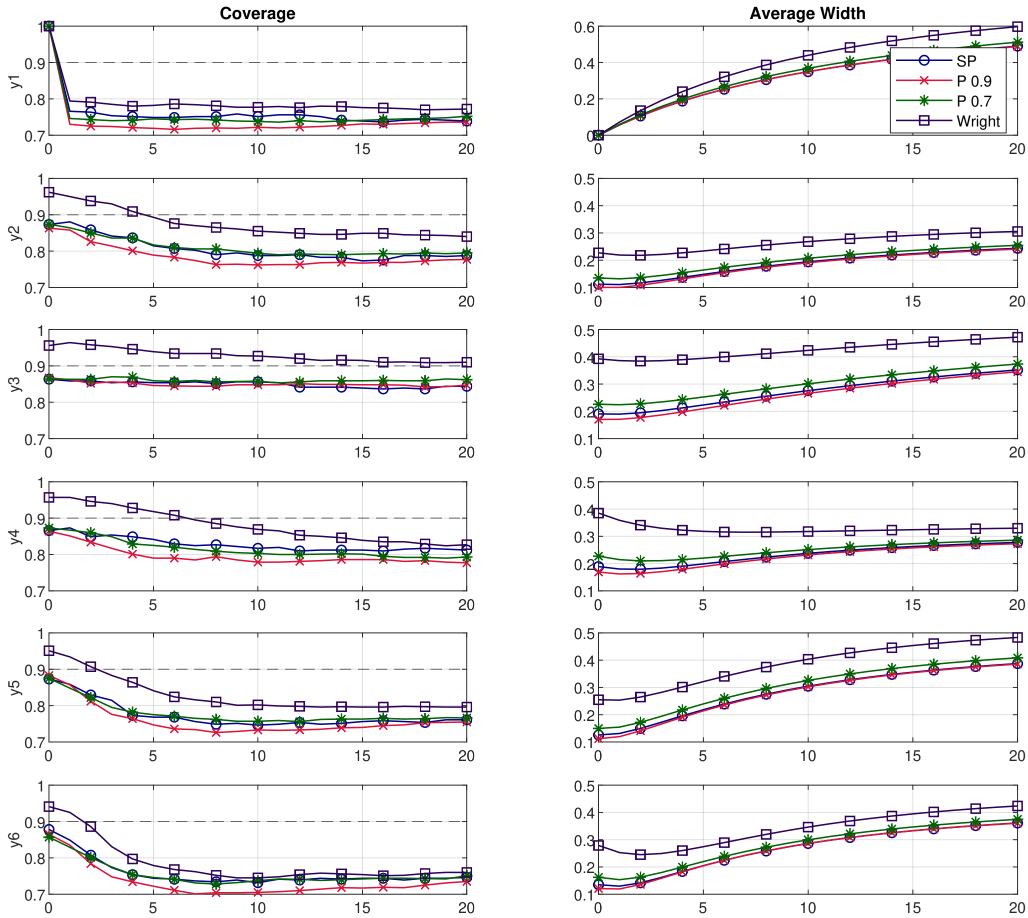

Figure 1: Coverage rates and average interval widths of nominal 90% confidence intervals for sample size $T = 500$, $M = 50$ event dates and event date variance of $\sigma_1^2 = 4$ associated with the four estimators $\widehat{\mathbf{b}}_{12}^{SP}$ (SP), $\widehat{\mathbf{b}}_{12}^{P(.9)}$ (P 0.9), $\widehat{\mathbf{b}}_{12}^{P(.7)}$ (P 0.7), and $\widehat{\mathbf{b}}_{12}^{W}$ (Wright).

is used to generate the confidence intervals. As we use it here to compare confidence intervals, they all share the same handicap.

Interestingly, the average interval widths of the confidence intervals associated with $\widehat{\mathbf{b}}_{12}^{P(.9)}$, $\widehat{\mathbf{b}}_{12}^{P(.7)}$, and $\widehat{\mathbf{b}}_{12}^{SP}$ are rather similar. On the other hand, $\widehat{\mathbf{b}}_{12}^{W}$ yields longer intervals on average than the other estimators but often also coverage rates closer to 90%. The greater interval lengths reflect the larger RMSEs of the latter estimator reported in the previous subsection and makes the estimator less attractive than the competitors with similar coverage rates. For $T = 250$ and $M = 25$, the impulse responses based on the four estimators are shown in the Online Appendix, Figure A.1. They show a similar pattern but tend to yield lower coverage rates and larger average widths. The slightly better coverage rates of $\widehat{\mathbf{b}}_{12}^{W}$ relative to the other estimators go along with larger average widths.

For $T = 500$ and $M = 50$, we have also considered confidence intervals for impulse responses associated with the joint GMM estimator which combines the moment conditions of the sign-proxy and the Wright estimators ($\widehat{\mathbf{b}}_{12}^{WSP}$) and estimators based on sign-proxies with 5 incorrect sign assignments, $\widehat{\mathbf{b}}_{12}^{SP(5)}$, and with 10% fewer event dates, $\widehat{\mathbf{b}}_{12}^{SP(-)}$. Detailed results are shown in Figures A.2 of the Online Appendix, where it can be seen that all four estimators yield similar coverage rates. Again the coverage rates for the impact effects are close to the nominal 90% while for some of the variables coverage rates for larger propagation horizons are a bit smaller than 90%. The average confidence interval lengths of the sign-proxy with 10% fewer event dates is almost identical to its counterpart without missed events. The intervals resulting from the sign-proxy with 5 misspecified signs and the joint GMM estimator are somewhat larger.

Overall it is not surprising that, for larger horizons, the coverage and widths of the confidence intervals based on all estimators become similar, given that the impulse responses for propagation horizons $h \geq 1$ are to some extent determined by the same reduced-form estimates. The results for the impact effects are in line with the RMSEs and confirm that the estimator based on a sign-proxy is a strong competitor to using a quantitative proxy or one of the other estimators in our competition.

3.3 Extensions Based on Other DGPs

In the Online Appendix we also compare the relative performance of the estimators for alternative persistence levels, lag orders, and dimension of the DGP. We do so based on a bivariate DGP that has been used repeatedly in the literature in comparisons of inference procedures for impulse responses. The details of the DGP and a set of tables with RMSEs are given in the Online Appendix (see Section A.3 and Tables A.5 - A.12).

The overall conclusions can be summarized briefly as follows. The relative performance of the estimators is not much affected by the dimension of the DGP, its persistence and the lag order, although the absolute estimation precision is, of course, affected. As the estimators for the impact effects are based on the reduced-form errors and the proxies, the persistence of the VAR process has almost no effect on some of the estimators for larger samples. The absolute RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.9)}$ and $\widehat{\mathbf{b}}_{12}^{P(.7)}$ are very similar across corresponding processes with different persistence levels for sample sizes $T = 250$ and 500 (see Table A.6). Also the lag order has little impact for larger sample sizes (see again Table A.6). Not surprisingly, the RMSEs tend to be a little larger for the larger VAR lag order, as one would expect.

Also the relative RMSEs shown in Tables A.7 - A.12 document that for larger samples and larger numbers of event dates, the relative performance of the estimators does not depend on the persistence and the lag order of the VAR process. Only for $T = 100$ and $T = 250$ in combination with an event date probability $d = 0.1$ there are some larger differences in the relative RMSEs. Clearly, for such small samples or event date probabilities, the reliability of some of the estimators is rather low, as we also observed for DGP1.

As the relative performance of the estimators is the crucial dimension for comparison, it follows that our main conclusions from the six-dimensional DGP1 discussed in Section 3.2 are overall confirmed by the additional results in the Online Appendix. In other words, these results also confirm that the sign-proxy is a serious competitor for a qualitative proxy. Compared to the 6-variable system, the Wright approach and the joint GMM estimators often yield lower RMSEs than the sign-proxy estimator. The smaller number of moment conditions simplifies the minimization of the GMM objective function which may drive this result.

4 Monetary Policy at the Zero Lower Bound

To illustrate the implications of our results for empirical work, we consider the benchmark model of Wright (2012) who investigates the impact of US monetary policy on longer-term interest rates at times when the policy rate is at the zero lower bound. As mentioned in Section 3, Wright considers a VAR(1) model. The following six daily US interest rates are included: (1) the 2-year nominal Treasury zero-coupon yields, (2) 10-year nominal Treasury zero-coupon yields, (3) five-year Treasury Inflation Protected Securities (TIPS) break-even rates, (4) 5-10-year TIPS break-even rates, (5) Moody's index of AAA corporate bond yields, and (6) Moody's index of BAA corporate bond yields. The sample period runs from November 3, 2008, to September 30, 2011, giving a sample of size $T = 729$ plus one presample value.

As mentioned in Section 3, Wright also constructs a proxy variable for a monetary policy shock related to $M = 28$ announcement days based on the first principal component of a set of Treasury futures (2, 5, 10 and 30-year) traded at the Chicago Mercantile Exchange (see Wright's Table 1). His identification of the measure is similar to Gürkaynak, Sack and Swanson (2005) and relies on tick data in a short window around FOMC announcements.

However, Wright does not use the resulting measure of monetary policy surprises as a proxy for the monetary policy shock but instead directly as the shock. Specifically, he regresses the reduced-form VAR residuals on the surprise measure and interprets the regression coefficients as the impact effects in $\mathbf{b}_1$. Here, in contrast, we use his measure as a proxy that is merely assumed to be correlated with the monetary policy shock. We construct a sign-proxy with values ± 1 on event dates using the signs of Wright's quantitative proxy. All other elements are fixed at 0.

In his analysis, Wright also considers the possibility of using only 13 especially important event dates around new phases of quantitative easing. Since the number of event dates, M, was found to be an important determinant of the estimation precision in the simulations, we also consider a quantitative proxy and associated sign-proxy with only the 13 major event dates written in italics in Wright's Table 1.

We present the robust F-values of the tests of the strength of both proxies and corresponding sign-proxies in Table 5, where it can be seen that all proxies, apart from the sign-proxy for $M = 28$, come with robust F-values well above the threshold of 10. Thus, they would be classified as strong instruments in a standard proxy VAR analysis. The sign-proxy for $M = 28$ yields a robust F-statistic of 9.64, slightly missing the cut-off to be regarded as a strong instrument. Our simulation results in Section 3 suggest that the sign-proxy based on 28 event dates may be a weaker instrument than the sign-proxy based on 13 event dates because the former sign-proxy may include more incorrect sign assignments or a smaller variance difference between the variances associated with dates with and without events.

We also present the empirical correlations between the proxies and the estimated first shocks on event dates in Table 5. Clearly, with -0.55 the correlation between the sign-proxy and the shock of interest is relatively small for $M = 28$ event dates. All other proxies have stronger (negative) correlations with their respective shocks. This result indicates that with $M = 28$, we may be in a situation where, based on our simulation results of Section 3, the corresponding quantitative proxy can be expected to yield more precise impulse response estimates than the sign-proxy. On the other hand, with $M = 13$, where both the quantitative proxy and the sign-proxy are classified as strong instruments, the relative performance of the estimates $\widehat{\mathbf{b}}_{12}^{P}$ and $\widehat{\mathbf{b}}_{12}^{SP}$ is not clear a priori, although the correlation between the sign-proxy and the shock is weaker than the correlation between the quantitative proxy and the shock which might suggest a superior performance of the quantitative proxy.

We estimate impulse responses of the monetary policy shock and, as in the Monte Carlo simulations, use the MBB based on bias-corrected OLS

Table 5: Diagnostics for Proxy Strength

	$M = 28$ event dates		$M = 13$ event dates	
	quant. proxy	sign-proxy	quant. proxy	sign-proxy
Robust F-statistic	51.31	9.64	52.62	19.25
Corr(proxy, shock)	-0.90	-0.55	-0.91	-0.75

Note: F-statistics and estimated shocks are based on bias-corrected OLS estimators for the reduced-form VAR parameters.

estimators to construct confidence intervals around the impulse responses estimated with the alternative estimators discussed in Section 2.[9]

In Figure 2, we compare pointwise 90% MBB confidence intervals associated with the quantitative proxy and the corresponding sign-proxy estimator for $M = 13$. The shock is standardized such that it reduces the 10-year Treasury yields by 25 basis points on impact. In most cases, the two estimators yield very similar point estimates and confidence intervals. In some cases, the confidence intervals based on the quantitative proxy are slightly smaller than the corresponding intervals of the sign-proxy (e.g., the short-horizon responses of the 2-year Treasury rates) and, in other cases, the situation is reversed (e.g., the longer horizon responses of the break even rates). Overall, there is not much to choose between the quantitative proxy and the sign-proxy estimates. Thus, if there is any additional value from using the more sophisticated quantitative proxy over the simple sign-proxy, it is very limited.

The impulse responses are also largely qualitatively the same as in Wright (2012) (see, e.g., his Figure 1). The monetary policy shock does not have much of an effect on 5-year break-even rates (i.e., short to medium-term inflation expectations) and lowers BAA and AAA yields by less than 25 basis points. However, the confidence intervals in Wright's Figure 1 are partly considerably larger than in our Figure 2. A more systematic comparison of Wright's heteroskedasticity approach and the sign-proxy estimates is given below.

In Figure 3, we present the impulse response estimates obtained with proxies based on $M = 28$ event dates. As expected on the basis of our simulation results in Section 3 and the F-statistics in Table 5, in this case the

[9] Wright (2012) uses a slightly different bootstrap. As in our Monte Carlo study, we use the MBB because Jentsch and Lunsford (2019) show that it yields asymptotically correct confidence intervals under general conditions for inference for impulse responses in proxy VARs. For the example we use a block length of $\ell = 25$, which corresponds roughly to the rule of thumb of $\ell = 5.03T^{1/4}$ of Jentsch and Lunsford (2019). The displayed confidence intervals are based on 2,000 bootstrap draws and look very similar for a block length of 50 or using a residual wild bootstrap instead.

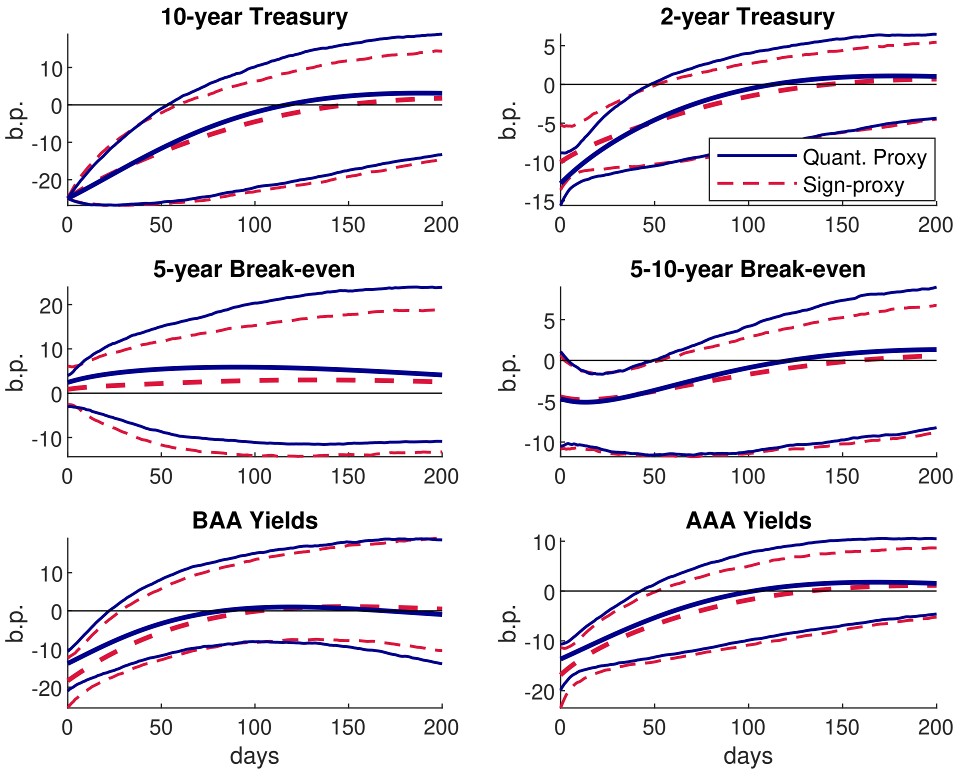

Figure 2: Quantitative proxy and sign-proxy estimates of responses to 25 basis points monetary policy shock with 90% MBB confidence intervals for $M = 13$ event dates.

quantitative proxy estimator yields overall smaller confidence intervals than the sign-proxy estimator. However, in most cases, the conclusions regarding the responses of the variables are again qualitatively the same. Exceptions are the initial response of the AAA yields which turns significantly negative only after the impact when the sign-proxy is considered and the first few responses of the 5-10-year break-even rate. For the break-even rate, the quantitative proxy yields a prolonged significantly negative effect, although also not on impact. Overall, the impulse response bands in Figure 3 reflect what we find in the simulations in Section 3, namely that the decline in estimation precision can be rather limited even if a sign-proxy is used instead of a strong quantitative proxy, which is constructed based on additional knowledge of the market structure.

In Figure 4, we show the impulse responses and 90% pointwise confidence intervals of the Wright heteroskedasticity estimator and the sign-proxy es-

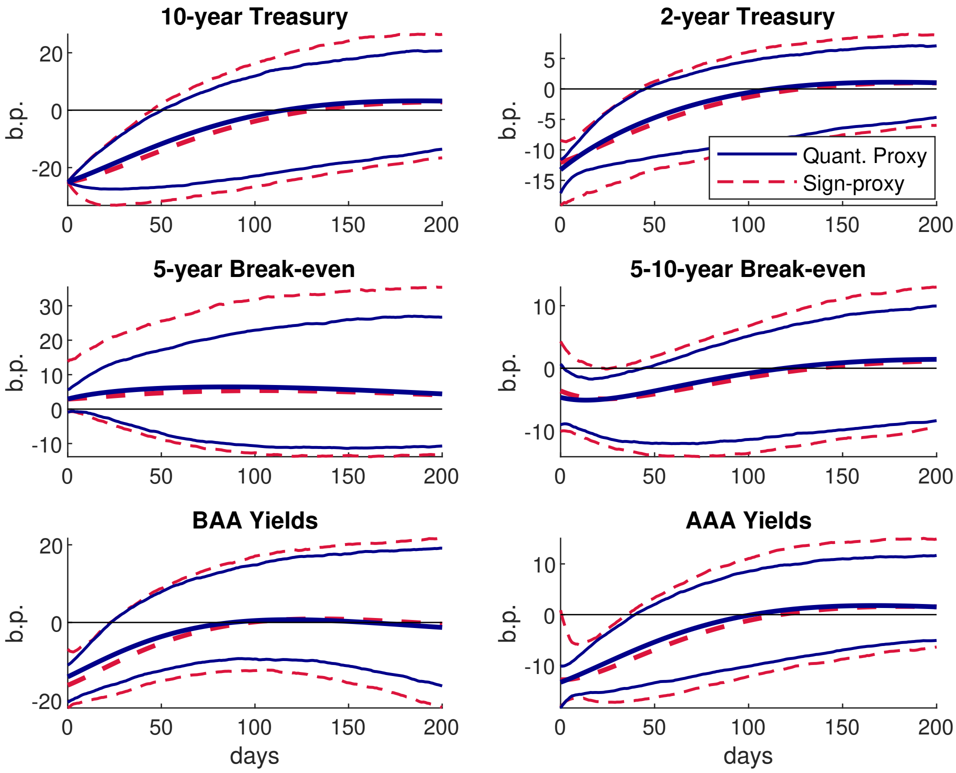

Figure 3: Quantitative proxy and sign-proxy estimates of responses to 25 basis points monetary policy shock with 90% MBB confidence intervals for $M = 28$ event dates.

timator based on $M = 13$ event dates.[10] Although we obtain the same qualitative results from both estimators for all the variables which are also largely in line with the results in Wright (2012), the sign-proxy intervals are a bit smaller. Thus, although the qualitative conclusions of Wright's study are confirmed, the precision of the inference can be improved by using the sign-proxy estimator if we assume that the size of the confidence intervals properly reflects the uncertainty in the estimates in the presently considered model.

In Figure 5, the MBB confidence intervals of the two GMM estimators $\widehat{\mathbf{b}}_{12}^{WP}$ and $\widehat{\mathbf{b}}_{12}^{WSP}$ which combine the moment conditions of Wright's quantitative proxy and the sign-proxy, respectively, with the moment conditions of Wright's heteroskedasticity estimator are compared for $M = 28$ event dates.

[10]Note that the Wright heteroskedasticity estimates differ slightly from those in Wright (2012, Figure 2) because we use a different algorithm for minimizing the GMM objective function (see also Online Appendix A.1.1).

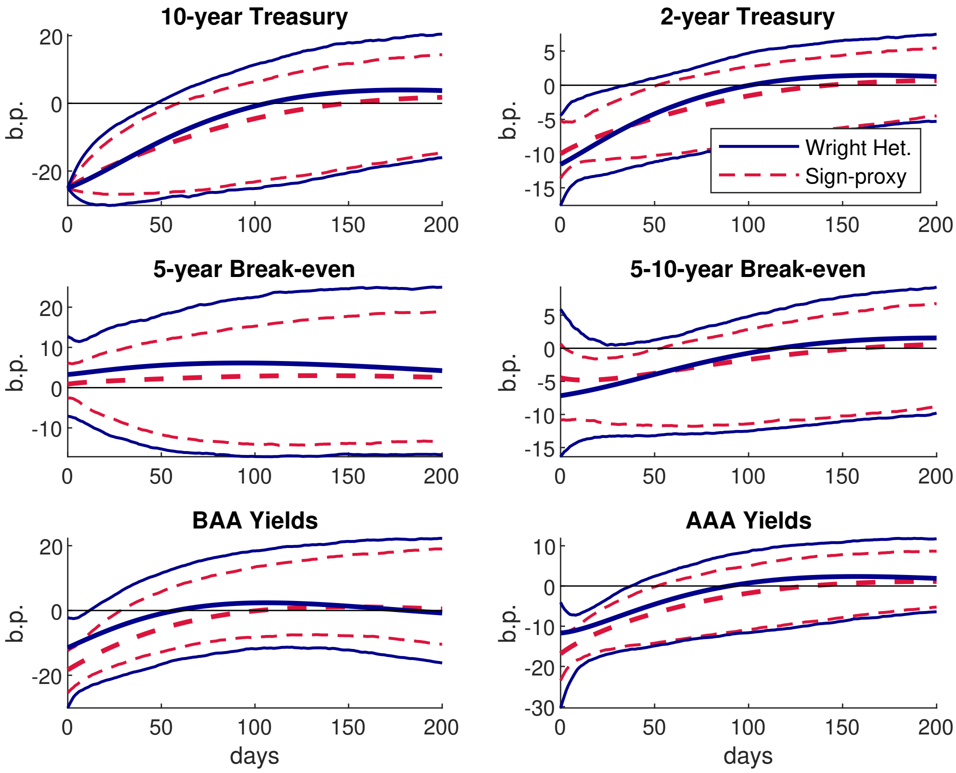

Figure 4: Sign-proxy and Wright heteroskedasticity estimates of responses to a 25 basis points monetary policy shock with 90% MBB confidence intervals for $M = 13$ event dates.

To compare them to the quantitative proxy, we also show 90% confidence intervals for the quantitative proxy estimator in Figure 5. Across all estimators the point estimates of the impulse responses are very similar. For lucidity, we omit them in Figure 5 and focus on the 90% pointwise confidence intervals. The bootstrap confidence intervals of the two GMM combination estimators are practically identical and they are also quite similar to the confidence intervals associated with the quantitative proxy estimator. In line with our simulation results in Section 3, the intervals of the quantitative proxy estimator are slightly smaller in some cases. Judged by the size of the confidence intervals, the GMM estimation precision of the sign-proxy and the conventional proxy are practically the same when they are used in combination with Wright's heteroskedasticity moment conditions. However, using the conventional quantitative proxy estimator without the additional moment conditions tends to improve the precision slightly. Overall, the example reflects what we find in our Monte Carlo simulations in Section 3.

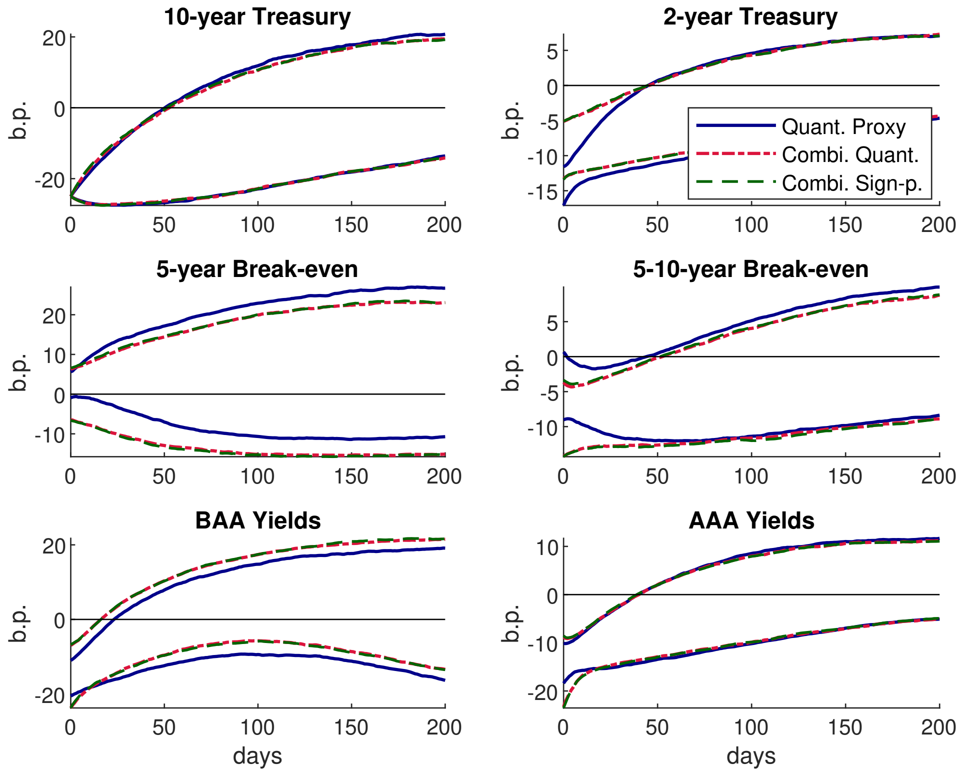

Figure 5: Pointwise 90% MBB confidence intervals of quantitative proxy and GMM combination estimators of responses to a 25-basis-points monetary policy shock for $M = 28$ event dates. (Point estimates are omitted for lucidity. They are all three very similar and well within the confidence intervals.)

5 Conclusions and Extensions

This study contributes to the rapidly growing literature on proxy VAR analysis. It compares quantitative to qualitative proxy variables that may be easier to construct. Obtaining a suitable instrument for estimating the initial responses of a shock of interest is crucial in a proxy VAR analysis. The precision of the estimates depends on the quality of the proxy, which has to be well correlated with the structural shock and uncorrelated with all other shocks. Such a proxy variable may be difficult to find in practice.

Based on an extensive Monte Carlo experiment, we find that an estimator has considerable merit that is based on a qualitative sign-proxy that assigns a $+1$ for positive shocks and a -1 for negative shocks on special event dates only. Thus, to construct the sign-proxy, it is enough to know the timing and direction of the shock for special dates. An estimator based on the sign-proxy tends to provide more precise inference than the Wright heteroskedasticity estimator. It can even be as precise as the conventional quantitative proxy estimator, which requires the construction of a suitable and possibly controversial strong quantitative proxy. As the simulations show, the sign-proxy may even dominate the quantitative proxy if the latter is a weak instrument. Thus, the sign-proxy estimator is a competitive alternative for empirical structural VAR analyses.

Moreover, we point out that the moment conditions implied by the proxy variable can be supplemented by moment conditions from possible volatility changes due to the special events. The combined moment conditions can be used for GMM estimation of the impact effects of the shocks. While this GMM estimator does not improve the estimation precision in some of our simulations, it may have merit if the resulting GMM objective function can be minimized with sufficient reliability. In any case, the sign-proxy or combination estimator tends to be more precise than an estimator that is based only on the volatility changes induced by the special events.

We illustrate the benefits of using the sign-proxy by investigating the impact of US monetary policy shocks on longer-term interest rates in times of a zero policy interest rate. It is shown that bootstrap confidence intervals of the impulse responses based on the sign-proxy estimator or the combination estimator tend to be smaller than their competitor based only on the volatility changes. Thus, the simulation conclusions are reinforced. The sign-proxy estimator is an attractive choice for empirical structural VAR analyses if information on the shocks of interest is scarce because it requires only qualitative information external to the VAR model.

There are a number of extensions of our work which may be of interest for future research. First of all, local projections are sometimes applied in the proxy VAR context. Although they were found to be less efficient in small samples when the assumed VAR process is indeed the DGP, they may have advantages when there is uncertainty about the model. Therefore it may be of interest that the sign-proxies can be combined straightforwardly with local projection estimators because they can be used as regular instruments. It may be an interesting topic for future research to compare more recent proposals for using external instruments in a local projection setting with an estimator based on a sign-proxy. Note, however, that we have compared the alternative estimators using the same reduced-form VAR estimates. Using local projections would extend the study to different reduced-form estimators.

In some recent proxy VAR studies, a number of structural shocks are identified jointly with a set of proxy variables. Again it may be interesting in future research to investigate the implications of replacing or combining quantitative proxies with sign-proxies in such a context.

References

Alessi, L. and Kerssenfischer, M. (2019). The response of asset prices to monetary policy shocks: Stronger than thought, *Journal of Applied Econometrics* **34**: 661–672.

Beaudry, P. and Saito, M. (1998). Estimating the effects of monetary shocks: An evaluation of different approaches, *Journal of Monetary Economics* **42**: 241–260.

Budnik, K. and Rünstler, G. (2020). Identifying SVARs from sparse narrative instruments: dynamic effects of U.S. macroprudential policies, *Working Paper Series 2353*, European Central Bank, Frankfurt.

Caldara, D. and Herbst, E. (2019). Monetary policy, real activity, and credit spreads: Evidence from bayesian proxy svars, *American Economic Journal: Macroeconomics* **11**: 157–192.

Carriero, A., Mumtaz, H., Theodoridis, K. and Theophilopoulou, A. (2015). The Impact of Uncertainty Shocks under Measurement Error: A Proxy SVAR Approach, *Journal of Money, Credit and Banking* **47**(6): 1223–1238.

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

Dias, D. A. and Duarte, J. B. (2019). Monetary policy, housing rents, and inflation dynamics, *Journal of Applied Econometrics* **34**: 673–687.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity, *American Economic Journal: Macroeconomics* **7**: 44–76.

Gürkaynak, R. S., Sack, B. and Swanson, E. (2005). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements, *International Journal of Central Banking* **1**(1): 55–93.

Jentsch, C. and Lunsford, K. G. (2019). The dynamic effects of personal and corporate income tax changes in the United States: Comment, *American Economic Review* **109**: 2655–2678.

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions, *Review of Economics and Statistics* **80**: 218–230.

Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*, Cambridge University Press, Cambridge.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. and Schlaak, T. (2019). Bootstrapping impulse responses of structural vector autoregressive models identified through GARCH, *Journal of Economic Dynamics and Control* **101**: 41–61.

Lütkepohl, H. and Schlaak, T. (2020). Heteroskedastic proxy vector autoregressions, *Discussion Paper 1876*, DIW Berlin, Berlin.

Mertens, K. and Ravn, M. O. (2013). The dynamic effects of personal and corporate income tax changes in the United States, *American Economic Review* **103**: 1212–1247.

Montiel Olea, J. L. and Pflueger, C. (2013). A Robust Test for Weak Instruments, *Journal of Business & Economic Statistics* **31**(3): 358–369.

Montiel Olea, J. L., Stock, J. H. and Watson, M. W. (2020). Inference in structural vector autoregressions identified with an external instrument, *Journal of Econometrics* .

Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to the Theory of Statistics*, Vol. Auckland, 3rd edn, McGraw-Hill.

Piffer, M. and Podstawski, M. (2018). Identifying uncertainty shocks using the price of gold, *The Economic Journal* **128**(1549): 3266–3284.

Plagborg-Møller, M. and Wolf, C. K. (2021). Local projections and VARs estimate the same impulse responses, *Econometrica* **89**(2): 955–980.

Romer, C. and Romer, D. (1989). Does monetary policy matter? A new test in the spirit of Friedman and Schwartz, *NBER Macroeconomics Annual* **4**: 121–170.

Stock, J. H. and Watson, M. W. (2012). Disentangling the channels of the 2007-09 recession, *Brookings Papers on Economic Activity* pp. 81–135.

Stock, J. H., Wright, J. H. and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business & Economic Statistics* **20**(4): 518–529.

Wright, J. H. (2012). What does monetary policy do to long-term interest rates at the zero lower bound?, *Economic Journal* **122**: F447–F466.

Online Appendix for

Qualitative versus Quantitative External Information for Proxy Vector Autoregressive Analysis

by

Lukas Boer
DIW Berlin and Humboldt-Univeristät Berlin, Mohrenstr. 58, 10117 Berlin, Germany
email: lboer@diw.de

and

Helmut Lütkepohl
DIW Berlin and Freie Universität Berlin, Mohrenstr. 58, 10117 Berlin, Germany
email: hluetkepohl@diw.de

February 15, 2021

A.1 Details on Computations

A.1.1 Minimization of Nonlinear Objective Functions

The Wright heteroskedasticity estimator $\widehat{\mathbf{b}}_{12}^{W}$ and the two GMM combination estimators $\widehat{\mathbf{b}}_{12}^{WP}$ and $\widehat{\mathbf{b}}_{12}^{WSP}$ using the heteroskedasticity moment conditions and the quantiative proxy and sign-proxy moment conditions, respectively, are estimated with the Matlab constrained optimization routine fmincon(.). For the Wright heteroskedasticity estimator we have also used unconstrained Matlab minimization routines where we set $c_W = 1$ in equation (8) following the procedure in Wright (2012). Specifying instead a lower bound larger but close to 0 for the difference, c_W, in variances of the first structural shock and then solving for c_W and $\mathbf{b}_{12}$ jointly leads to much improved RMSEs in the simulation parts for the heteroskedasticity estimator. As initial values for the optimization algorithms we have specified the true c_W and $\mathbf{b}_{12}$ values. For the combination estimators we additionally set the initial value of the covariance between shock and proxy, c, to 1/6 without any bounds.

A.1.2 Implementation of the Moving-Block Bootstrap

Bootstrap samples are generated by a moving-block bootstrap (MBB), as in Jentsch and Lunsford (2019). In the following, $\hat{u}_1, \dots, \hat{u}_T$ are the estimated residuals of bias-adjusted OLS estimation of the reduced-form VAR model and $z_1, \dots, z_T$ are the observed proxy values.

Let $\ell < T$ be the block length for the MBB. Denoting by $[\cdot]$ the smallest number greater than or equal to the argument such that $\ell s \geq T$, $s = [T/\ell]$ is the number of blocks required for constructing a bootstrap sample of y_t. Blocks of ℓ consecutive residuals and proxies are arranged as follows:

$$\begin{bmatrix} \begin{pmatrix} \hat{u}_1 \\ z_1 \end{pmatrix} & \begin{pmatrix} \hat{u}_2 \\ z_2 \end{pmatrix} & \dots & \begin{pmatrix} \hat{u}_\ell \\ z_\ell \end{pmatrix} \\ \begin{pmatrix} \hat{u}_2 \\ z_2 \end{pmatrix} & \begin{pmatrix} \hat{u}_3 \\ z_3 \end{pmatrix} & \dots & \begin{pmatrix} \hat{u}_{1+\ell} \\ z_{1+\ell} \end{pmatrix} \\ \vdots & \vdots & & \vdots \\ \begin{pmatrix} \hat{u}_{T-\ell+1} \\ z_{T-\ell+1} \end{pmatrix} & \begin{pmatrix} \hat{u}_{T-\ell+2} \\ z_{T-\ell+2} \end{pmatrix} & \dots & \begin{pmatrix} \hat{u}_T \\ z_T \end{pmatrix} \end{bmatrix}.$$

From these blocks, s are drawn with replacement and these draws are joined end-to-end, retaining only the first T residuals and proxies,

$$\begin{pmatrix} u_t^* \\ z_t^{MBB} \end{pmatrix}, \quad t = 1, \dots, T.$$

The u_t^* are recentered as

$$u_{j\ell+i}^{MBB} = u_{j\ell+i}^* - \frac{1}{T - \ell + 1} \sum_{r=0}^{T-\ell} \hat{u}_{i+r}$$

for $i = 1, 2, \dots, \ell$ and $j = 0, 1, \dots, s - 1$ and the bootstrap residuals and proxies are obtained as

$$\begin{pmatrix} u_t^{MBB} \\ z_t^{MBB} \end{pmatrix}, \quad t = 1, \dots, T.$$

The bootstrap samples are generated sequentially as $y_t^{MBB} = \hat{\nu} + \hat{A}_1 y_{t-1}^{MBB} + \cdots + \hat{A}_p y_{t-p}^{MBB} + u_t^{MBB}$, $t = 1, \dots, T$, starting from p randomly chosen consecutive sample values, $y_{-p+1}^{MBB}, \dots, y_0^{MBB}$.

Bootstrap impulse response estimates are obtained from N bootstrap samples $y_{-p+1}^{(n)}, \dots, y_0^{(n)}, y_1^{(n)}, \dots, y_T^{(n)}$ and $z_1^{(n)}, \dots, z_T^{(n)}$, $n = 1, \dots, N$, as follows:

1. Fitting a VAR(p) model to the sample yields bootstrap estimates $\hat{A}^{(n)}$,

   $$\hat{\Phi}_i^{(n)} = \sum_{j=1}^{i} \hat{\Phi}_{i-j}^{(n)} \hat{A}_j^{(n)}, \quad i = 1, \dots, H,$$

   and residuals $\hat{u}_t^{(n)}$.

2. Bootstrap estimates $\widehat{\mathbf{b}}_1^{(n)}$ are computed with all the alternative estimation methods using $\mathcal{T}_1 = \{t | z_t \neq 0\}$ for the Wright heteroskedasticity estimator and associated combination estimators.

3. Bootstrap estimates of the impulse responses of interest are computed as

$$\widehat{\Theta}(H)^{(n)} = [\widehat{\mathbf{b}}_1^{(n)}, \hat{\Phi}_1^{(n)}\widehat{\mathbf{b}}_1^{(n)}, \dots, \hat{\Phi}_H^{(n)}\widehat{\mathbf{b}}_1^{(n)}]$$

and stored.

The N bootstrap estimates $\widehat{\Theta}(H)^{(1)}, \dots, \widehat{\Theta}(H)^{(N)}$ are used to construct pointwise confidence intervals from the relevant quantiles of the bootstrap distributions of the individual elements.

A.2 Additional Information for DGP1

A.2.1 Parameter Values of DGP1

DGP1 is a VAR(1) process with intercept term obtained by estimating a 6-dimensional model based on a dataset from Wright (2012). The estimated parameters are

$$\nu = (0.156, 0.059, 0.030, 0.128, 0.184, 0.225)'$$

for the constant term and the slope coefficient matrix is

$$A_1 = \begin{bmatrix} 1.028 & -0.003 & 0.023 & -0.014 & 0.032 & -0.090 \\ 0.040 & 0.947 & 0.011 & -0.008 & 0.022 & -0.058 \\ 0.029 & -0.041 & 1.001 & -0.003 & 0.016 & -0.039 \\ 0.019 & 0.000 & 0.006 & 0.947 & -0.002 & -0.008 \\ 0.008 & 0.007 & 0.003 & -0.027 & 0.998 & -0.028 \\ 0.046 & -0.008 & 0.017 & -0.011 & 0.031 & 0.886 \end{bmatrix}.$$

The reduced-form errors that are added to create the different DGP1 samples are obtained via the relationship $u_t = Bw_t$ where the structural shocks are constructed as

$$w_{1t} = \begin{cases} \mathcal{N}(0, \sigma_1^2) & \text{for} \quad t \in \mathcal{T}_1, \\ \mathcal{N}(0, \sigma_0^2) & \text{for} \quad t \in \mathcal{T} \setminus \mathcal{T}_1, \end{cases}$$

and $(w_{2t}, \dots, w_{Kt}) \sim \mathcal{N}(0, I_{K-1})$. The structural impact effects matrix B resulting from the Cholesky decomposition of the estimated residual covariance matrix and a normalization such that the upper left-hand element equals one is given by

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.454 & 0.386 & 0 & 0 & 0 & 0 \\ 0.322 & 0.146 & 0.633 & 0 & 0 & 0 \\ 0.457 & -0.160 & 0.313 & 0.538 & 0 & 0 \\ 0.795 & -0.074 & -0.035 & -0.034 & 0.417 & 0 \\ 0.843 & 0.077 & 0.021 & 0.001 & 0.318 & 0.328 \end{bmatrix}.$$

A.2.2 Additional Simulation Results for DGP1

Table A.1: Relative Frequencies of Weak Instrument F-test Statistics Smaller than 10 (in %) (without heteroskedasticity adjustment)

			Proxy variables						
Sample size T	M	σ_1^2	$z_t^{(.9)}$	$z_t^{(.7)}$	s_t	$s_t^{(-)}$	$s_t^{(1)}$	$s_t^{(3)}$	$s_t^{(5)}$
100	20	4	0.2	7.1	0.1	0.4	1.2	19.7	66.0
		10	0.0	2.6	0.0	0.0	0.1	6.9	42.5
250	10	4	6.1	27.1	6.1	10.1	28.5	87.5	99.4
		10	0.7	10.1	0.3	0.5	5.7	62.1	94.9
	25	4	0.0	1.7	0.0	0.0	0.0	1.3	14.8
		10	0.0	0.4	0.0	0.0	0.0	0.0	3.2
500	25	4	0.0	1.3	0.0	0.0	0.0	0.8	11.5
		10	0.0	0.3	0.0	0.0	0.0	0.1	2.1
	50	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		10	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Note: M denotes the number of event dates, σ_1^2 the variance of the structural shock of interest on event dates, $z_t^{(.9)}$ a proxy with a theoretical correlation of 0.9 with the structural shock of interest on event dates, $z_t^{(.7)}$ a proxy with a theoretical correlation of 0.7, s_t the sign-proxy, $s_t^{(-)}$ a sign-proxy with 10% omitted event dates, and $s_t^{(m)}$ denotes a sign-proxy with m incorrectly specified signs.

Table A.2: Absolute RMSEs of Proxy Estimators $\widehat{\mathbf{b}}_{12}^{P(.9)}$ and $\widehat{\mathbf{b}}_{12}^{P(.7)}$ (with bias-adjustment of VAR estimators)

			absolute RMSEs	
Sample size T	M	σ_1^2	$\widehat{\mathbf{b}}_{12}^{P(.9)}$	$\widehat{\mathbf{b}}_{12}^{P(.7)}$
100	20	4	0.167	0.338
		10	0.103	0.193
250	10	4	0.276	1.028
		10	0.309	3.069
	25	4	0.142	0.199
		10	0.088	0.126
500	25	4	0.140	0.192
		10	0.087	0.121
	50	4	0.097	0.124
		10	0.060	0.079

Note: M denotes the number of event dates and σ_1^2 is the variance of the structural shock of interest on event dates. The RMSE is calculated as $\sqrt{\frac{1}{R}\sum_{r=1}^{R}(\widehat{\mathbf{b}}_{12,r}^{P(.)} - \mathbf{b}_{12})'(\widehat{\mathbf{b}}_{12,r}^{P(.)} - \mathbf{b}_{12})}$, where R denotes the number of Monte Carlo simulations.

Table A.3: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.9)}$ (with bias-adjustment of VAR estimators)

			relative RMSEs							
T	M	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(1)}$	$\widehat{\mathbf{b}}_{12}^{SP(3)}$	$\widehat{\mathbf{b}}_{12}^{SP(5)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	20	4	1.09	1.15	1.22	1.65	7.96	1.98	1.61	1.49
		10	1.11	1.18	1.25	1.69	6.84	1.66	1.35	1.32
250	10	4	0.96	1.03	1.67	47.25	186.50	1.62	1.57	1.56
		10	0.64	0.68	1.02	66.20	90.84	0.95	0.93	0.93
	25	4	1.09	1.17	1.19	1.48	1.97	2.20	1.73	1.57
		10	1.12	1.19	1.22	1.51	2.02	1.83	1.47	1.37
500	25	4	1.10	1.17	1.20	1.48	1.99	2.56	2.18	1.98
		10	1.12	1.20	1.23	1.51	2.04	2.33	1.96	1.75
	50	4	1.10	1.17	1.15	1.26	1.39	1.84	1.21	1.23
		10	1.13	1.19	1.18	1.29	1.42	1.37	1.10	1.13

Note: T signifies the sample size, M denotes the number of event dates and σ_1^2 is the variance of the structural shock of interest on event dates.

Table A.4: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.7)}$ (with bias-adjustment of VAR estimators)

			relative RMSEs							
T	M	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(1)}$	$\widehat{\mathbf{b}}_{12}^{SP(3)}$	$\widehat{\mathbf{b}}_{12}^{SP(5)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	20	4	0.59	0.62	0.66	0.89	4.35	1.07	0.94	0.81
		10	0.62	0.66	0.70	0.94	3.59	0.93	0.82	0.74
250	10	4	0.26	0.28	0.46	12.98	55.02	0.44	0.43	0.42
		10	0.18	0.19	0.27	24.27	26.26	0.26	0.26	0.26
	25	4	0.79	0.84	0.86	1.07	1.43	1.59	1.35	1.14
		10	0.79	0.84	0.86	1.06	1.42	1.29	1.11	0.96
500	25	4	0.81	0.86	0.88	1.09	1.47	1.88	1.68	1.46
		10	0.81	0.86	0.88	1.09	1.47	1.67	1.49	1.25
	50	4	0.86	0.90	0.89	0.98	1.08	1.43	1.05	0.95
		10	0.86	0.90	0.89	0.98	1.08	1.04	0.90	0.85

Note: T signifies the sample size, M denotes the number of event dates and σ_1^2 is the variance of the structural shock of interest on event dates.

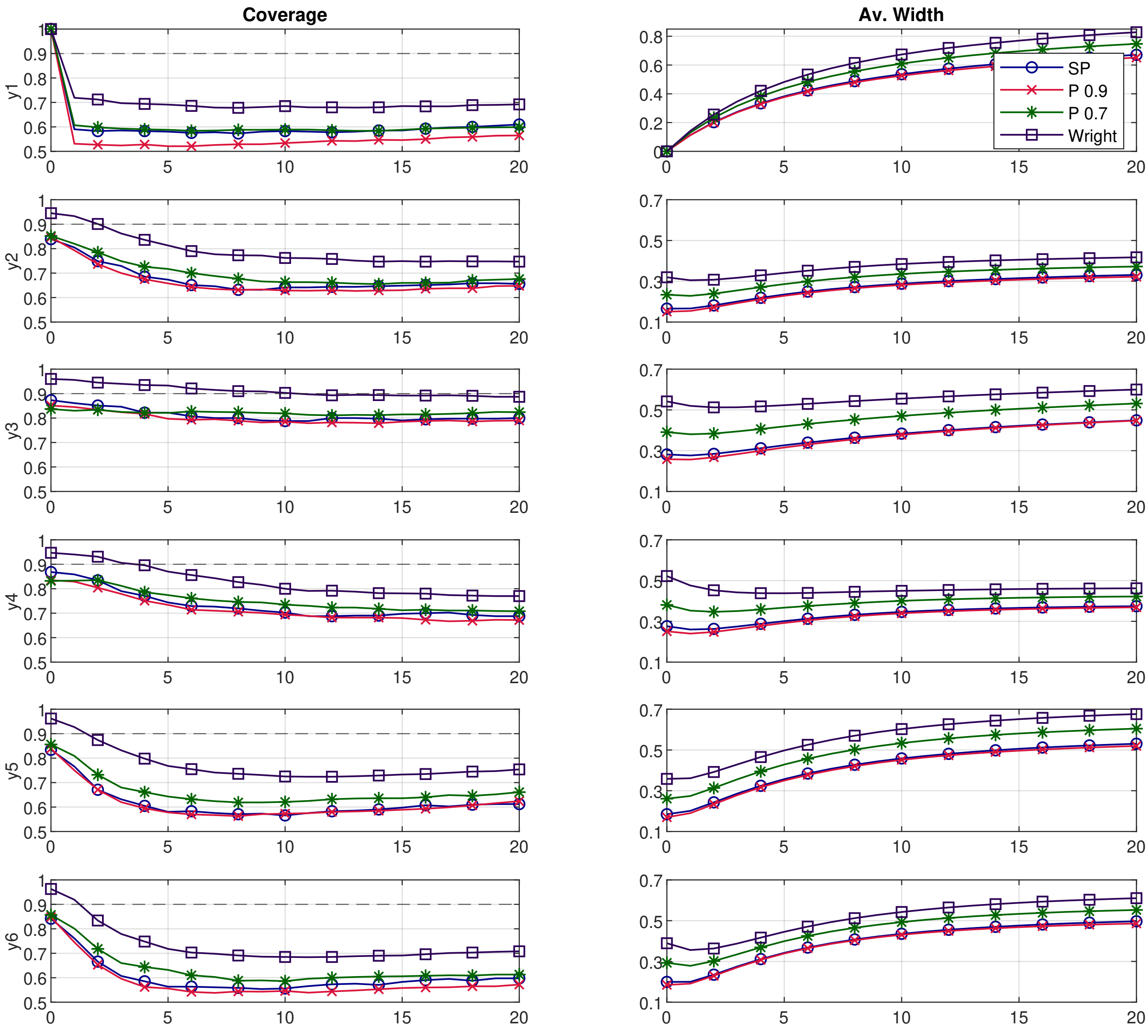

Figure A.1: Coverage rates and avearge interval widths of nominal 90% confidence intervals for sample size $T = 250$, $M = 25$ event dates and event date variance of $\sigma_1^2 = 4$ associated with the four estimators $\widehat{\mathbf{b}}_{12}^{SP}$ (SP), $\widehat{\mathbf{b}}_{12}^{P(.9)}$ (P 0.9), $\widehat{\mathbf{b}}_{12}^{P(.7)}$ (P 0.7), and $\widehat{\mathbf{b}}_{12}^{W}$ (Wright).

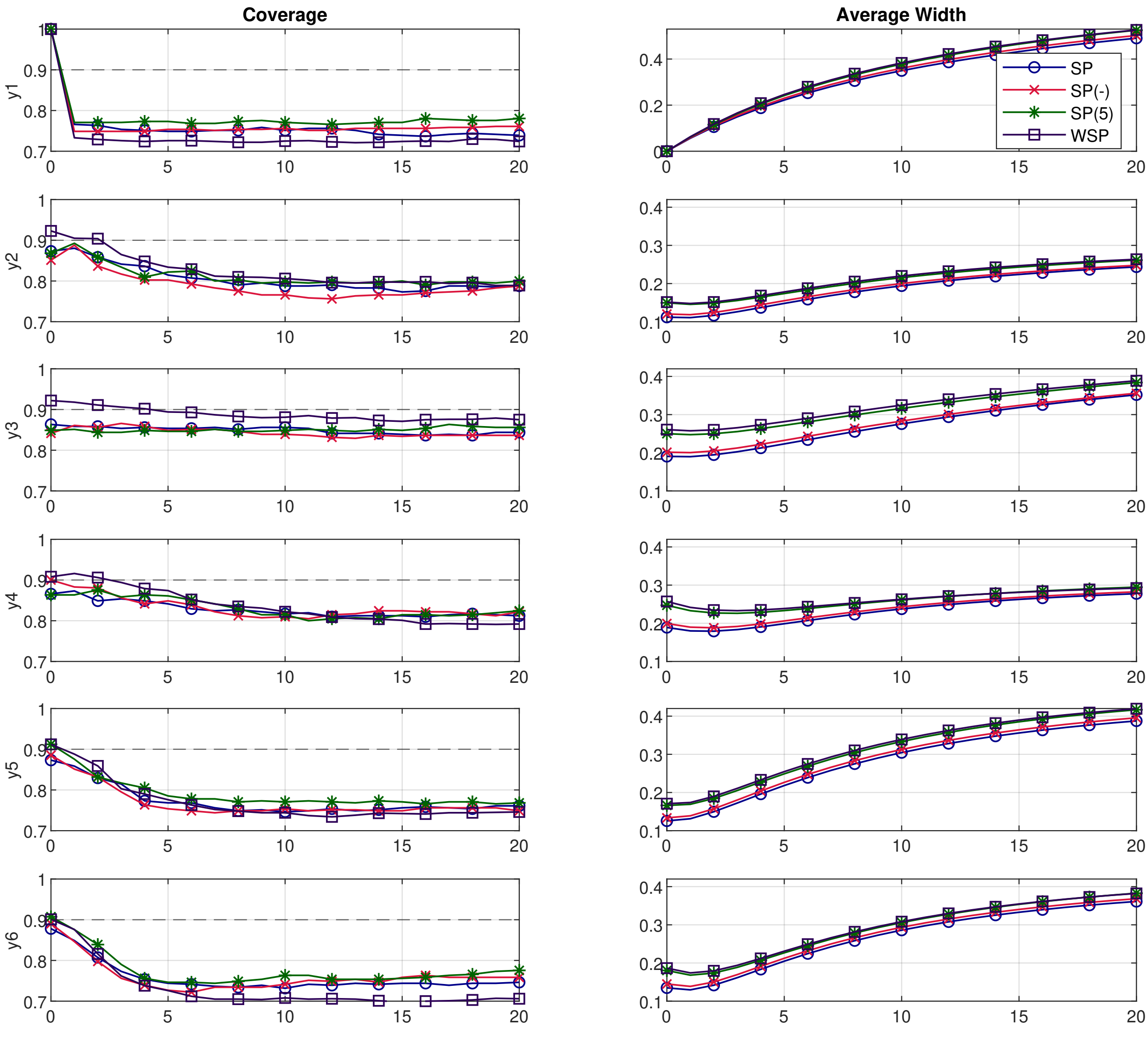

Figure A.2: Coverage rates and average interval widths of nominal 90% confidence intervals for sample size $T = 500$, $M = 50$ event dates and event date variance of $\sigma_1^2 = 4$ associated with the four estimators $\widehat{\mathbf{b}}_{12}^{SP}$, $\widehat{\mathbf{b}}_{12}^{SP(-)}$, $\widehat{\mathbf{b}}_{12}^{SP(5)}$, and $\widehat{\mathbf{b}}_{12}^{WSP}$.

A.3 Simulation Results Based on Additional DGP

A.3.1 Setup of DGP2

We have also performed simulations with a DGP that has been used in a number of other investigations of inference methods for impulse responses (e.g., Kilian (1998), Kilian and Kim (2011), Lütkepohl, Staszewska-Bystrova and Winker (2015a, 2015b)). Our DGP2 is a two-dimensional VAR(1) process,

$$y_t = \begin{bmatrix} a_{11} & 0 \\ 0.5 & 0.5 \end{bmatrix} y_{t-1} + u_t, \tag{A.1}$$

with $a_{11} = 0.5$ and 0.95. The parameter a_{11} determines the persistence of the process. The process is more persistent if a_{11} is closer to one. The value $a_{11} = 0.5$ yields a process with moderate persistence and $a_{11} = 0.95$ yields a rather persistent process.

The structural errors, w_t, are Gaussian, with zero mean and covariance matrices

$$\Sigma_0^w = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{for } t \in \mathcal{T} \setminus \mathcal{T}_1 \quad \text{and} \quad \Sigma_1^w = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{for } t \in \mathcal{T}_1,$$

and the reduced-form errors are obtained as $u_t = Bw_t$, where

$$B = \begin{bmatrix} 1 & 0 \\ 0.5 & 1.5 \end{bmatrix}.$$

The initial values of the y_t series are $y_0 \sim \mathcal{N}(0, I_2)$. VAR models of order $p = 1$ and $p = 12$ with constant term are fitted to the data. Thus, the full sample has length $T + p$ and $\forall\, t \geq 1$, y_t is constructed following the DGP in (A.1).

The proxy z_t is generated as

$$z_t = D_t(w_{1t} + v_t),$$

where we choose the error v_t to be Gaussian, $v_t \sim \mathcal{N}(0, \sigma_v^2)$, in our simulations. We use the same values for σ_v^2 as for DGP1. They are given in Table A.5.

The quantity D_t is a Bernoulli distributed random variable with parameter d, $B(d)$, which is stochastically independent of v_t and w_{1t}. Thus, $\mathbb{E}(D_t) = d$ and, thus, d is the expected fraction of event dates. For d we use values 0.1 and 0.2 which correspond to the sample fractions used for event dates for DGP1. Now, however, the number of event dates is stochastic and so is $\mathcal{T}_1$, and the correlation between the proxy z_t and the first shock w_{1t} has to be assessed for the whole sample using

$$\operatorname{corr}(z_t, w_{1t}) = \sigma_1 \sqrt{d} \Big/ \sqrt{\sigma_1^2 + \sigma_v^2}\,.$$

Thus, the parameters d, σ_1^2 and the error term v_t determine the strength of the correlation between z_t and w_{1t}. The correlations used in the Monte Carlo simulations are given in Table A.5. For comparability with the results for DGP1, where we condition on a fixed number of event dates, we denote estimators based on the quantitative proxies by $\widehat{\mathbf{b}}_{12}^{P(.9)}$, if $\sigma_v^2 = 1$ or 2, and by $\widehat{\mathbf{b}}_{12}^{P(.7)}$, if $\sigma_v^2 = 4$ or 10. Clearly, for a given value of d, the proxy underlying $\widehat{\mathbf{b}}_{12}^{P(.9)}$ has a stronger correlation with w_{1t} than the proxy underlying $\widehat{\mathbf{b}}_{12}^{P(.7)}$ which justifies the notation because it makes the estimates comparable to the corresponding ones for DGP1.

Instead of the sign-proxies with a fixed number of incorrect signs, we now consider sign-proxies with 10%, 20%, and 25% incorrect signs and denote the corresponding estimators by $\widehat{\mathbf{b}}_{12}^{SP(10\%)}$, $\widehat{\mathbf{b}}_{12}^{SP(20\%)}$, and $\widehat{\mathbf{b}}_{12}^{SP(25\%)}$, respectively. The estimator $\widehat{\mathbf{b}}_{12}^{SP(-)}$ is again based on a sign-proxy for which 10% of the event dates are omitted.

Table A.5: Correlations Between Shock and Proxies

d	σ_1^2	σ_v^2	$\text{corr}(z_t, w_{1t})$
0.1	4	1	0.2828
	4	4	0.2236
	10	2	0.2887
	10	10	0.2236
0.2	4	1	0.4000
	4	4	0.3162
	10	2	0.4082
	10	10	0.3162

Whenever a sample of straight zeros is drawn for D_t, it is deleted and replaced by a new one. The number of replications for each Monte Carlo design is $R = 5{,}000$.

A.3.2 Results for DGP2

Table A.6: Absolute RMSEs of Proxy Estimators $\widehat{\mathbf{b}}_{12}^{P(.9)}$ and $\widehat{\mathbf{b}}_{12}^{P(.7)}$ for DGP2

			absolute RMSEs					
			$p=1$ $a_{11}=0.5$		$p=1$ $a_{11}=0.95$		$p=12$ $a_{11}=0.95$	
Sample size T	d	σ_1^2	$\widehat{\mathbf{b}}_{12}^{P(.9)}$	$\widehat{\mathbf{b}}_{12}^{P(.7)}$	$\widehat{\mathbf{b}}_{12}^{P(.9)}$	$\widehat{\mathbf{b}}_{12}^{P(.7)}$	$\widehat{\mathbf{b}}_{12}^{P(.9)}$	$\widehat{\mathbf{b}}_{12}^{P(.7)}$
100	0.2	4	0.214	0.379	0.215	1.036	0.247	0.386
		10	0.132	0.247	0.132	0.861	0.150	0.317
250	0.1	4	0.185	0.268	0.185	0.265	0.196	0.270
		10	0.114	0.168	0.114	0.166	0.120	0.172
	0.2	4	0.124	0.159	0.124	0.159	0.132	0.172
		10	0.077	0.100	0.077	0.101	0.081	0.108
500	0.1	4	0.125	0.161	0.125	0.161	0.124	0.164
		10	0.077	0.102	0.077	0.102	0.077	0.104
	0.2	4	0.087	0.111	0.087	0.111	0.087	0.112
		10	0.054	0.070	0.054	0.070	0.054	0.071

Note: d denotes the probability of event dates and σ_1^2 is the variance of the structural shock of interest on event dates. The RMSE is calculated as $\sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\widehat{\mathbf{b}}_{12,r}^{P(.)} - \mathbf{b}_{12}\right)^2}$, where R denotes the number of Monte Carlo replications.

Table A.7: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.9)}$ for DGP2 with $p = 1$ and $a_{11} = 0.5$

			relative RMSEs							
T	d	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(10\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(20\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(25\%)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	0.2	4	1.09	1.14	1.41	3.48	7.65	1.71	1.10	1.13
		10	1.12	1.17	1.44	24.68	4.81	1.13	1.01	1.04
250	0.1	4	1.10	1.17	1.43	2.29	11.07	1.50	1.09	1.13
		10	1.12	1.20	1.47	2.33	8.08	1.09	1.01	1.04
	0.2	4	1.10	1.17	1.39	1.91	2.38	1.31	1.07	1.11
		10	1.13	1.19	1.43	1.95	2.43	1.05	0.99	1.03
500	0.1	4	1.09	1.16	1.39	1.91	2.39	1.29	1.07	1.10
		10	1.11	1.18	1.42	1.95	2.44	1.05	0.99	1.02
	0.2	4	1.11	1.17	1.38	1.86	2.27	1.26	1.06	1.10
		10	1.13	1.19	1.41	1.90	2.32	1.04	0.99	1.02

Table A.8: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.9)}$ for DGP2 with $p = 1$ and $a_{11} = 0.95$

			relative RMSEs							
T	d	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(10\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(20\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(25\%)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	0.2	4	1.08	1.14	1.41	2.65	7.46	1.79	1.10	1.13
		10	1.12	1.17	1.45	2.62	640.73	1.13	1.01	1.04
250	0.1	4	1.09	1.17	1.43	2.20	4.89	1.44	1.09	1.13
		10	1.12	1.20	1.46	2.27	5.68	1.09	1.01	1.04
	0.2	4	1.10	1.17	1.39	1.91	2.38	1.31	1.07	1.11
		10	1.13	1.19	1.43	1.95	2.43	1.05	0.99	1.02
500	0.1	4	1.09	1.16	1.39	1.91	2.39	1.29	1.07	1.10
		10	1.11	1.18	1.42	1.95	2.45	1.05	0.99	1.02
	0.2	4	1.11	1.17	1.38	1.86	2.27	1.25	1.06	1.10
		10	1.13	1.19	1.41	1.90	2.32	1.04	0.99	1.02

Table A.9: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.9)}$ for DGP2 with $p = 12$ and $a_{11} = 0.95$

			relative RMSEs							
T	d	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(10\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(20\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(25\%)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	0.2	4	1.08	1.15	1.57	4.22	17.27	1.88	1.12	1.16
		10	1.11	1.19	1.52	3.50	11.83	1.39	1.02	1.07
250	0.1	4	1.08	1.14	1.40	2.31	5.51	1.55	1.07	1.11
		10	1.11	1.18	1.44	2.58	16.55	1.07	0.99	1.02
	0.2	4	1.11	1.17	1.40	1.96	2.44	1.35	1.07	1.12
		10	1.14	1.20	1.43	2.00	2.50	1.08	1.00	1.04
500	0.1	4	1.12	1.19	1.44	1.96	2.49	1.31	1.07	1.12
		10	1.15	1.22	1.47	2.00	2.54	1.06	1.00	1.03
	0.2	4	1.10	1.16	1.38	1.92	2.34	1.29	1.07	1.11
		10	1.12	1.18	1.42	1.97	2.39	1.06	1.00	1.02

Table A.10: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.7)}$ for DGP2 with $p = 1$ and $a_{11} = 0.5$

			relative RMSEs							
T	d	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(10\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(20\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(25\%)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	0.2	4	0.61	0.65	0.80	1.96	4.31	0.96	0.74	0.64
		10	0.59	0.62	0.77	13.14	2.56	0.60	0.58	0.56
250	0.1	4	0.75	0.81	0.98	1.57	7.61	1.03	0.86	0.78
		10	0.76	0.81	1.00	1.58	5.49	0.74	0.73	0.71
	0.2	4	0.86	0.91	1.09	1.50	1.86	1.03	0.92	0.87
		10	0.86	0.91	1.09	1.50	1.86	0.81	0.80	0.79
500	0.1	4	0.85	0.90	1.08	1.48	1.86	1.01	0.91	0.86
		10	0.85	0.90	1.08	1.48	1.86	0.80	0.80	0.78
	0.2	4	0.87	0.91	1.08	1.46	1.78	0.99	0.91	0.87
		10	0.87	0.91	1.08	1.46	1.78	0.80	0.81	0.79

Table A.11: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.7)}$ for DGP2 with $p = 1$ and $a_{11} = 0.95$

			relative RMSEs							
T	d	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(10\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(20\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(25\%)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	0.2	4	0.23	0.24	0.29	0.55	1.55	0.37	0.27	0.24
		10	0.17	0.18	0.22	0.40	98.43	0.17	0.17	0.16
250	0.1	4	0.77	0.82	1.00	1.54	3.43	1.01	0.88	0.79
		10	0.77	0.82	1.01	1.56	3.91	0.75	0.74	0.72
	0.2	4	0.86	0.91	1.09	1.49	1.86	1.02	0.92	0.87
		10	0.86	0.91	1.09	1.50	1.86	0.81	0.80	0.78
500	0.1	4	0.85	0.90	1.08	1.49	1.87	1.01	0.91	0.86
		10	0.85	0.90	1.08	1.49	1.86	0.80	0.80	0.78
	0.2	4	0.87	0.92	1.08	1.46	1.78	0.98	0.91	0.87
		10	0.87	0.92	1.08	1.46	1.78	0.80	0.80	0.78

Table A.12: RMSEs of Estimators for Impact Effects of the First Shock Relative to the Corresponding RMSEs of $\widehat{\mathbf{b}}_{12}^{P(.7)}$ for DGP2 with $p = 12$ and $a_{11} = 0.95$

			relative RMSEs							
T	d	σ_1^2	$\widehat{\mathbf{b}}_{12}^{SP}$	$\widehat{\mathbf{b}}_{12}^{SP(-)}$	$\widehat{\mathbf{b}}_{12}^{SP(10\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(20\%)}$	$\widehat{\mathbf{b}}_{12}^{SP(25\%)}$	$\widehat{\mathbf{b}}_{12}^{W}$	$\widehat{\mathbf{b}}_{12}^{WP}$	$\widehat{\mathbf{b}}_{12}^{WSP}$
100	0.2	4	0.69	0.74	1.00	2.70	11.08	1.20	0.90	0.75
		10	0.53	0.57	0.72	1.66	5.62	0.66	0.57	0.51
250	0.1	4	0.79	0.83	1.02	1.68	4.01	1.13	0.88	0.81
		10	0.78	0.82	1.00	1.80	11.58	0.74	0.73	0.71
	0.2	4	0.85	0.90	1.07	1.50	1.87	1.04	0.92	0.86
		10	0.85	0.89	1.07	1.49	1.87	0.80	0.80	0.77
500	0.1	4	0.85	0.90	1.09	1.49	1.89	0.99	0.90	0.85
		10	0.85	0.90	1.09	1.49	1.89	0.79	0.79	0.77
	0.2	4	0.85	0.90	1.07	1.49	1.82	1.01	0.91	0.86
		10	0.85	0.90	1.08	1.49	1.82	0.81	0.81	0.78

References

Jentsch, C. and Lunsford, K. G. (2019). The dynamic effects of personal and corporate income tax changes in the United States: Comment, *American Economic Review* **109**: 2655–2678.

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions, *Review of Economics and Statistics* **80**: 218–230.

Kilian, L. and Kim, Y. (2011). How reliable are local projection estimators of impulse responses?, *Review of Economics and Statistics* **93**: 1460–1466.

Lütkepohl, H., Staszewska-Bystrova, A. and Winker, P. (2015a). Comparison of methods for constructing joint confidence bands for impulse response functions, *International Journal for Forecasting* **31**: 782–798.

Lütkepohl, H., Staszewska-Bystrova, A. and Winker, P. (2015b). Confidence bands for impulse responses: Bonferroni versus Wald, *Oxford Bulletin of Economics and Statistics* **77**: 800–821.