Journal of Economic Dynamics & Control 161 (2024) 104837

Contents lists available at ScienceDirect

Journal of Economic Dynamics and Control

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

Heteroskedastic proxy vector autoregressions: An identification-robust test for time-varying impulse responses in the presence of multiple proxies ☆

Martin Bruns [a,*], Helmut Lütkepohl [b]

[a] University of East Anglia, School of Economics, Norwich Research Park, NR4 7TJ, Norwich, United Kingdom
[b] DIW Berlin and Freie Universität Berlin, Mohrenstr. 58, 10117 Berlin, Germany

ARTICLE INFO

Keywords:
Structural vector autoregression
Proxy VAR
Heteroskedasticity
Productivity shocks
Structural change

ABSTRACT

We propose a test for time-varying impulse responses in heteroskedastic structural vector autoregressions that can be used when the shocks are identified by external proxy variables as a group but not necessarily individually. The test is robust to the identification scheme for identifying the shocks individually and can be used even if the shocks are not identified individually. The asymptotic analysis is supported by small sample simulations which show good properties of the test. An investigation of the impact of productivity shocks in a small macroeconomic model for the U.S. illustrates the importance of the issue for empirical work.

1. Introduction

In time series econometrics it is important to account for structural change both for proper inference (e.g., Perron (2006)) and for structural analysis based, for example, on vector autoregressive (VAR) models. In this study we focus on the issue in the context of structural VAR analysis. For such an analysis, using external instruments or proxies to identify shocks of interest has become increasingly popular lately and is now sometimes signified as proxy VAR analysis (see, e.g., Stock and Watson (2012), Mertens and Ravn (2013)). In a number of studies, a set of proxies is used to identify a group of shocks collectively. In that case, if impulse responses or related tools are used, it is typically necessary to provide additional information to identify the shocks of interest individually.[1] We propose a test for time-varying structural impulse responses that is robust to the identification of the individual shocks. Although the test can also be used if the VAR slope coefficients are time-varying, we focus on a heteroskedastic VAR process with time-invariant slope coefficients and thereby consider a scenario assumed in many empirical studies.

☆ The authors thank Robin Braun, Kurt Lunsford, and an anonymous referee for helpful comments on an earlier draft of this paper. They also thank participants at the NBER-NSF Time Series Conference, the Barcelona School of Economics Summer Forum, the Conference on Heavy Tails and Robust Estimation, the European Summer Meeting of the Econometric Society, the 16th International Conference on Computational and Financial Econometrics, and the Virtual Time Series Seminar for comments. An earlier version was circulated under the title "Heteroskedastic Proxy Vector Autoregressions: Testing for Time-Varying Impulse Responses in the Presence of Multiple Proxies". Martin Bruns thanks the British Academy for financial support, grant number TDA22\220046.

* Corresponding author.
E-mail addresses: martin.bruns@uea.ac.uk (M. Bruns), luetkepohlh@gmail.com (H. Lütkepohl).

[1] Such additional information can have the form of zero restrictions on the impact effects (see, e.g., Mertens and Ravn (2013)) or the long-run effects of the shocks. Also sign restrictions may be considered (see, e.g., Piffer and Podstawski (2017), Braun and Brüggemann (2023), or Arias et al. (2021)). Another alternative are restrictions on the forecast error variance decomposition (see Härtl (2022)). If the impact effects are time-invariant, heteroskedasticity can also be used for individually identifying the shocks (see, e.g., Carriero et al. (2021)).

https://doi.org/10.1016/j.jedc.2024.104837
Received 8 November 2023; Received in revised form 17 January 2024; Accepted 10 February 2024

Available online 15 February 2024

In many proxy VAR analyses, it is assumed that the impulse responses of the structural shocks are time-invariant even if there are changes in the volatility of the shocks, that is, even if there is heteroskedasticity. In other words, it is assumed that the responses of the variables are not affected, despite the time-varying variances of the shocks. In much of the literature this assumption is used without further investigation. However, some authors question this time-invariance assumption for the impulse responses (e.g., Angelini et al. (2019), Bacchiocchi et al. (2018), Bacchiocchi and Fanelli (2015)). In a recent article, Lütkepohl and Schlaak (2022) propose a statistical test to explore the validity of such an assumption in the context of heteroskedastic proxy VAR models. Their test works under the premise that a specific single shock is properly identified by one or more proxies. If a number of shocks are identified collectively and not individually by a set of proxies, additional information is needed to identify the shocks individually in order to apply their test.

In this study we propose a test for time-varying impact effects which is robust to the identification scheme for the individual shocks and can even be applied if a set of shocks is only collectively identified by proxies, i.e., a linear transformation of the shocks is identified by the proxies, but the shocks are not individually identified. In other words, we propose a test for time-varying impact effects of the shocks that works even if the impact effects are not point identified and, hence, cannot be estimated consistently. If time-invariance cannot be rejected, the heteroskedasticity can potentially be used to identify the shocks of interest individually, as in the approach of Lanne and Lütkepohl (2008), while this approach is not available if the impact effects vary (see also Carriero et al. (2021)).

We confirm good small sample size and power properties of the test in a Monte Carlo study. The test is then applied to investigate the impact of productivity shocks on a small macroeconomic model for the U.S. economy based on a benchmark study by Lunsford (2015). He considers two shocks to total factor productivity (TFP), one based on the consumption sector without durable goods and the other one based on durable goods and investment. He uses two proxies to identify the two TFP shocks and compares the dynamic effects on the variables of a small U.S. macroeconomic system. Given the volatility change in many U.S. macro data in the middle of the 1980s when the Great Moderation (GM) started, we apply our new test to explore the time-invariance of the impulse responses of the two structural shocks. We find evidence against time-invariance and show that allowing for a change in the dynamic effects of the shocks leads to markedly different dynamic responses in some of the variables in the pre- and post-GM periods, in particular a weaker response of inflation in the post-GM period is observed.

The remainder of this study is structured as follows. In the next section, we present the general model framework. In Section 3, the test for time-varying impact effects of structural shocks at times of volatility change is presented and its small sample properties are investigated by means of a Monte Carlo study in Section 4. The empirical study follows in Section 5 and conclusions are drawn in Section 6. The Appendix presents some related theoretical derivations and an Online Appendix contains details for the Monte Carlo simulations, and additional simulation results.

2. Heteroskedastic proxy VAR models

The basic model is a K-dimensional reduced-form VAR process,

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

where ν is a time-invariant constant term, $A_1, \ldots, A_p$ contain the time-invariant autoregressive slope coefficients, and u_t is a zero mean white noise process with nonsingular, possibly time-varying covariance matrix Σ_t. In short, $u_t \sim (0, \Sigma_t)$. In other words, there may be heteroskedasticity. The heteroskedasticity is assumed to be such that

$$\mathbb{E}(u_t u_t') = \Sigma_t = \Sigma_u(m) \quad \text{for} \quad t \in \mathcal{T}_m, \quad m = 1, \ldots, M, \tag{2}$$

where $\mathcal{T}_m = \{T_{m-1}+1, \ldots, T_m\}$ $(m = 1, \ldots, M)$ are M volatility regimes. The volatility changes at the end of time periods, T_m, for $m = 1, \ldots, M-1$, with $T_0 = 0$ and $T_M = T$, the overall sample size. In our theoretical setup, the change points, T_m, are assumed to occur exogenously and are known to the analyst.[2]

The structural errors, $\boldsymbol{w}_t = (w_{1t}, \ldots, w_{Kt})'$, have a diagonal covariance matrix and are obtained from the reduced-form errors, u_t, by a linear transformation which may depend on the volatility regime, $u_t = B(m)\boldsymbol{w}_t$, such that $B(m)$ is the matrix of impact effects of the structural shocks in volatility regime m. We assume that the first K_1 shocks are of primary interest and partition $\boldsymbol{w}_t$ in K_1- and $(K-K_1)$-dimensional subvectors $\boldsymbol{w}_{1t} = (w_{1t}, \ldots, w_{K_1 t})'$ and $\boldsymbol{w}_{2t} = (w_{K_1+1,t}, \ldots, w_{Kt})'$ such that $\boldsymbol{w}_t' = (\boldsymbol{w}_{1t}', \boldsymbol{w}_{2t}')$ and we partition $B(m) = [B_1(m) : B_2(m)]$ accordingly such that $B_1(m)$ is $(K \times K_1)$ and $B_2(m)$ is $(K \times (K-K_1))$. In other words, $B_i(m)$ contains the impact effects of the shocks in $\boldsymbol{w}_{it}$, $i = 1, 2$, in volatility regime m.

The impact effects, $B(m)$, of the structural shocks are crucial for determining the dynamic effects of the shocks because the structural impulse responses for propagation horizon h are obtained as

$$\Theta_h(m) = \Phi_h B(m),$$

where the $\Phi_i = \sum_{j=1}^{i} \Phi_{i-j} A_j$, with $\Phi_0 = I_K$, may be obtained recursively for $h = 0, 1, \ldots$, from the reduced-form VAR coefficients A_j, with $A_j = 0$ for $j > p$ (see, e.g., Lütkepohl (2005, Sec. 2.1.2)). Thus, the impact effects enter the structural impulse responses at

[2] For the asymptotic analysis in Section 3, we actually need that the sample size in each volatility regime goes to infinity. Hence, one may think of each volatility regime as consisting of a fixed fraction of the full sample. Given that only a single fixed sample is typically available for the actual empirical analysis, we pretend that the volatility change points rather than sample fractions are given to simplify the terminology.

all propagation horizons and make them regime-dependent if the impact effects are time-varying. Identifying and estimating them is therefore of central importance for estimating impulse responses and for the related dynamic analysis.

Suppose there is a set of N instrumental variables (proxies) $z_t = (z_{1t}, \dots, z_{Nt})'$ satisfying, for $t \in \mathcal{T}_m$,

$$\mathbb{E}(\boldsymbol{w}_{1t} z_t') = C_m \neq 0, \quad C_m \ (K_1 \times N), \quad \text{rk}(C_m) = K_1 \quad \text{(relevance)}, \tag{3}$$

$$\mathbb{E}(\boldsymbol{w}_{2t} z_t') = 0 \quad \text{(exogeneity)}. \tag{4}$$

The relevance condition allows the covariance between the structural shocks of primary interest, $\boldsymbol{w}_{1t}$, and the proxies to depend on the volatility regime. It implies that, for $t \in \mathcal{T}_m$,

$$\mathbb{E}(u_t z_t') = B(m)\mathbb{E}(\boldsymbol{w}_t z_t') = B_1(m)C_m. \tag{5}$$

Hence, the proxies contain identifying information for the first $K_1 < K$ structural shocks collectively, but the shocks are not necessarily individually identified by the proxies, z_t, in each of the volatility regimes. Obviously, there must be at least as many proxies as there are identified shocks such that $N \geq K_1$ to satisfy the rank condition for C_m which ensures that the N proxies contain actually identifying information for all shocks in $\boldsymbol{w}_{1t}$. Note that C_m does not depend on the sample size. In other words, we are not considering weak proxies which are asymptotically uncorrelated with the shocks of interest as, e.g., in Montiel Olea et al. (2021).

For individual rather than collective identification of the shocks in $\boldsymbol{w}_{1t}$, further information is required which could take the form of exclusion restrictions on the impact effects or the long-run effects of the shocks or sign restrictions or other types of identifying information that is typically used in this context. Such additional information could come from heteroskedasticity if the impact effects are time-invariant such that $B(m) = B$ for all or some $m \in \{1, \dots, M\}$. To see that, let $\Lambda_m = \text{diag}(\lambda_{1,m}, \dots, \lambda_{K,m})$ $(m = 1, \dots, M)$ be the covariance matrix of $\boldsymbol{w}_t$ for $t \in \mathcal{T}_m$. Then our assumptions imply that

$$\Sigma_u(m) = B(m)\Lambda_m B(m)', \quad m = 1, \dots, M. \tag{6}$$

If the impact effects are invariant across volatility regimes such that $B(m) = B$, then $\Sigma_u(m) = B\Lambda_m B'$ for $m = 1, \dots, M$. These relations uniquely identify the structural parameters B up to column sign and Λ_m, $m = 1, \dots, M$, if the regime dependent variances of the structural shocks are ordered uniquely and are sufficiently heterogeneous (see Lanne et al. (2010) for precise conditions). Thus, if the impact effects of the structural shocks are time-invariant, heteroskedasticity may identify the shocks and the information in the proxies may overidentify the shocks $\boldsymbol{w}_{1t}$, which can potentially be used to sharpen inference (see Carriero et al. (2021)). In fact, it is enough that heteroskedasticity identifies the shocks of interest to combine the information in the proxies with the volatility features to improve inference. For that to be possible, it is important that the impact effects $B_1(m)$ of $\boldsymbol{w}_{1t}$ are invariant across volatility regimes. Hence, having a test for time-varying impact effects of $\boldsymbol{w}_{1t}$ is also of interest in the context of identification through heteroskedasticity. We will propose such a test in the following section.

The previous analysis is also relevant for models with volatility changes driven by a Markov switching process as in Lanne et al. (2010) and Herwartz and Lütkepohl (2014), where also a finite number of volatility regimes is assumed.

3. Testing for time-varying impact effects

We abbreviate the $(K \times N)$ product matrix $B_1(m)C_m$ by $D(m)$ and estimate this matrix as

$$\widehat{D}(m) = \frac{1}{\tau_m T} \sum_{t \in \mathcal{T}_m} \hat{u}_t z_t', \tag{1}$$

where the $\hat{u}_t$ are OLS residuals of the reduced-form VAR model (1). Thus, $\widehat{D}(m)$ is an estimator of the covariance matrix $\mathbb{E}(u_t z_t')$ in volatility regime $m \in \{1, \dots, M\}$. Assuming that $\tau_m = (T_m - T_{m-1})/T$ is a fixed fraction of the sample size such that $T_m - T_{m-1} \to \infty$ with T,

$$\sqrt{T}\text{vec}\left(\widehat{D}(m) - D(m)\right) \xrightarrow{d} \mathcal{N}\left(0, \tau_m^{-1}\Sigma_D(m)\right), \tag{2}$$

where vec denotes the column stacking operator and $\xrightarrow{d}$ signifies convergence in distribution. Under general conditions, this result follows from a central limit theorem. We also assume that $u_t z_t'$ is such that

$$\widehat{\Sigma}_D(m) = \frac{1}{\tau_m T} \sum_{t \in \mathcal{T}_m} \text{vec}(\hat{u}_t z_t' - \widehat{D}(m))[\text{vec}(\hat{u}_t z_t' - \widehat{D}(m))]'$$

is a consistent estimator of $\Sigma_D(m)$ of dimension $(KN \times KN)$.

As explained earlier, for time-invariant impulse responses and for identification through heteroskedasticity, it is essential that $B_1(m)$ is time-invariant and does not depend on the volatility regime. Thus, we would like to test

$$\mathbb{H}_0 : B_1(m) = B_1(k) \quad \text{versus} \quad \mathbb{H}_1 : B_1(m) \neq B_1(k) \tag{3}$$

for some $m, k \in \{1, \dots, M\}$, $m \neq k$. The challenge is to derive a test for the pair of hypotheses in (3) that works although we can only estimate the product matrix $D(m)$ consistently but not the $B_1(m)$ matrix. The test should work regardless of possible time-variation of C_m and, hence, $D(m)$. In other words, the covariance between proxies and shocks may vary across volatility regimes.

If $B_1(m)$ is not fully identified via the proxies, we take advantage of the fact that $B_1(m)$ will be time-varying if a linear transformation is time-varying. We partition the matrix $B_1(m)$ as

$$B_1(m) = \begin{bmatrix} B_{11}(m) \\ B_{12}(m) \end{bmatrix},$$

where $B_{11}(m)$ is $(K_1 \times K_1)$ and $B_{12}(m)$ is $((K-K_1)\times K_1)$. We assume that the variables are arranged such that $B_{11}(m)$ is nonsingular. This is always possible as $B_1(m)$ has rank K_1. Then we consider the transformed matrix

$$\begin{aligned}\begin{bmatrix} I_{K_1} \\ B_{12}(m)B_{11}(m)^{-1} \end{bmatrix} &= B_1(m)B_{11}(m)^{-1} \\ &= B_1(m)C_m H C_m'(C_m H C_m')^{-1}B_{11}(m)^{-1} \\ &= B_1(m)C_m H C_m' B_{11}(m)'(B_{11}(m)C_m H C_m' B_{11}(m)')^{-1} \\ &= D(m)HD_1(m)'[D_1(m)HD_1(m)']^{-1} \end{aligned} \tag{4}$$

for any positive definite $(N \times N)$ matrix H. Here $D_1(m)$ is the upper $(K_1 \times N)$ part of

$$D(m) = \begin{bmatrix} D_1(m) \\ D_2(m) \end{bmatrix}$$

and $D_2(m)$ is a $((K-K_1)\times N)$ matrix. Note that C_m has rank K_1 due to the relevance condition (3) and, hence, all inverses in equation (4) exist. The matrix H cancels on the left-hand side of (4) and will leave our test unaffected. It can be used as a scaling matrix for the right-hand side expression of (4). If scaling is not desired, it may simply be replaced by the identity matrix, i.e., $H = I_N$. Note also that the left-hand side of expression (4) does not involve elements of C_m but just elements of $B_1(m)$, while the right-hand side of (4) consists of quantities that can be estimated consistently. Hence, we can also estimate the left-hand side consistently.

Using the result in (4), we test

$$\mathbb{H}_0 : B_{12}(m)B_{11}(m)^{-1} = B_{12}(k)B_{11}(k)^{-1} \text{ vs. } \mathbb{H}_1 : B_{12}(m)B_{11}(m)^{-1} \neq B_{12}(k)B_{11}(k)^{-1} \tag{5}$$

instead of the pair of hypotheses in (3). If $\mathbb{H}_1$ in (5) holds, then $B_1(m)$ must be regime-dependent and, hence, time-varying as well. It turns out that the latter pair of hypotheses can be tested without individually identifying the shocks in $\boldsymbol{w}_{1t}$ and regardless of C_m which does not show up in the quantities considered under $\mathbb{H}_0$. Hence, a rejection of $\mathbb{H}_0$ indicates time-variation in the identified impact effects, $B_1(m)$, of the $\boldsymbol{w}_{1t}$ shocks while C_m is allowed to vary across volatility regimes even under $\mathbb{H}_0$.

Of course, viewing this test as a test of $\mathbb{H}_0 : B_1(m) = B_1(k)$, it is a test which does not have power against alternatives for which $B_{12}(m)B_{11}(m)^{-1} = B_{12}(k)B_{11}(k)^{-1}$. Hence, $\mathbb{H}_0$ in (5) is only a necessary condition for $\mathbb{H}_0$ in (3) to hold. For example, if a change in volatility from regime m to regime k changes $B_1(m)$ to $B_1(k) = B_1(m)R$ (where R is a $(K_1 \times K_1)$ invertible matrix), that change would cancel in (5). In practice, a change in the impact effects due to a change in volatility that cancels in $B_{12}(k)B_{11}(k)^{-1}$ may not be very likely. Thus, if $\mathbb{H}_0$ in (5) cannot be rejected, this finding is a stronger indication that the assumption of time-invariant impulse responses is reasonable than just taking for granted that time-invariance holds.

The reason for being able to test $\mathbb{H}_0$ in (5) is that we can estimate

$$B_{12}(m)B_{11}(m)^{-1} = D_2(m)HD_1(m)'[D_1(m)HD_1(m)']^{-1}$$

consistently as

$$B_{12}\widehat{(m)B_{11}}(m)^{-1} = \hat{D}_2(m)H\hat{D}_1(m)'[\hat{D}_1(m)H\hat{D}_1(m)']^{-1}, \tag{6}$$

where we choose

$$H = \left(\sum_{t\in\mathcal{T}_m} z_t z_t'\right)^{-1}$$

and emphasize that the matrix just scales the covariance but does not affect the value of the test statistic or its small and large sample properties. Given the asymptotic normality of $\hat{D}(m)$ in (2), Slutsky's theorem implies consistency and asymptotic normality of the estimator $B_{12}\widehat{(m)B_{11}}(m)^{-1}$, i.e.,

$$\sqrt{T}\text{vec}\left(B_{12}\widehat{(m)B_{11}}(m)^{-1} - B_{12}(m)B_{11}(m)^{-1}\right) \xrightarrow{d} \mathcal{N}(0, V(m)), \tag{7}$$

where

$$V(m) = \frac{1}{\tau_m}\frac{\partial \text{vec}[B_{12}(m)B_{11}(m)^{-1}]}{\partial \text{vec}D(m)'}\Sigma_D(m)\frac{\partial \text{vec}[B_{12}(m)B_{11}(m)^{-1}]'}{\partial \text{vec}D(m)}$$

is the $(K_1(K-K_1)\times K_1(K-K_1))$ asymptotic covariance matrix. A closed-form expression of the matrix of partial derivatives $\partial\text{vec}[B_{12}(m)B_{11}(m)^{-1}]/\partial\text{vec}D(m)'$ is derived in Appendix A and can be used to estimate the covariance matrix $V(m)$.

The covariance matrix $V(m)$ is nonsingular. Defining the $K_1(K-K_1)$-dimensional vector

$$\beta(m) = \text{vec}[B_{12}(m)B_{11}(m)^{-1}]$$

and using the asymptotic independence of $\hat{\beta}(m)$ and $\hat{\beta}(k)$ under general conditions, for $m \neq k$, the null hypothesis in (3) can be tested using the test statistic

$$\eta(m,k) = T\left(\hat{\beta}(m)-\hat{\beta}(k)\right)'\left(\hat{V}(m)+\hat{V}(k)\right)^{-1}\left(\hat{\beta}(m)-\hat{\beta}(k)\right) \xrightarrow{d} \chi^2(K_1(K-K_1)). \tag{8}$$

Thus, we can use this statistic for testing the pair of hypotheses (5). Note that the degrees of freedom of the asymptotic χ^2-distribution are equal to the dimension of the vectors $\beta(m)$ and $\beta(k)$ and, hence, correspond to the number of restrictions we are testing. As the $(K \times K_1)$ matrices $B_1(m)$ and $B_1(k)$ are not identified, the number of restrictions is less than KK_1. In the test statistic in expression (8), the estimators of the covariance matrices may be obtained as

$$\hat{V}(m) = \frac{1}{\tau_m}\widehat{\frac{\partial\beta(m)}{\partial \text{vec}\, D(m)'}}\hat{\Sigma}_D(m)\widehat{\frac{\partial\beta(m)'}{\partial \text{vec}\, D(m)}}$$

by replacing the partial derivatives by estimates based on $\hat{D}(m)$ and using $H = \left(\sum_{t\in\mathcal{T}_m} z_t z_t'\right)^{-1}$, as for the estimator of β.[3]

It is important to note that the test statistic can be computed without having the first K_1 shocks identified individually. In fact, if there are restrictions identifying the shocks individually, they are ignored in the computation of the test statistic and, hence, the test is robust to the identification of the shocks. If there is only one proxy identifying a single shock, then the test statistic η reduces to the corresponding test statistic proposed by Lütkepohl and Schlaak (2022) for testing for time-varying impact effects of the shock identified by the proxy. If a set of proxies identifies multiple shocks jointly, applying their test is not possible. In contrast, our test can be applied if the variables are ordered in such a way that B_{11} is a nonsingular matrix which is not necessarily automatically the case. It requires, for example, that the shocks $\boldsymbol{w}_{1t}$ have nonzero impact effects on the first K_1 variables. For the case of a single proxy which identifies a single shock, that condition corresponds to the assumption that the identified shock has a nonzero impact effect on the first variable, as assumed in Lütkepohl and Schlaak (2022).

An advantage of being able to apply our test without the need to identify the shocks of interest, $\boldsymbol{w}_{1t}$, individually is that it may provide an argument for using heteroskedasticity for identifying the shocks individually. If, for $m \neq k$, $\Sigma_u(m) \neq \Sigma_u(k)$ we find that $B_1(m) = B_1(k)$, we have shocks with time-varying variances but time-invariant impact effects. This feature is used in large parts of the literature on identification through heteroskedasticity, provided there is enough heterogeneity in the variance changes of the shocks (see Lanne et al. (2010) or Kilian and Lütkepohl (2017, Sec. 14.3) for precise conditions). This approach for the identification of the shocks is not available if $\mathbb{H}_0$ is rejected and, hence, $B_1(m) \neq B_1(k)$ has to be assumed. Actually, if the impact effects are invariant across three different volatility regimes, say $B(m) = B(k) = B(\ell)$, $m \neq k \neq \ell$, this implies over-identifying restrictions in our setup and can be used for an alternative test for time-varying impact effects that does not rely on proxies but requires at least three distinct volatility regimes (see, e.g., Lanne et al. (2010), Kilian and Lütkepohl (2017, Sec. 14.3), Angelini et al. (2019)).

If there are more than two volatility regimes and we want to test for time-invariance of the impact effects across all volatility regimes by applying our tests to pairs of volatility regimes, we are in a multiple testing situation. The overall significance level of such a procedure will not be the same as the significance levels of the individual tests. An upper bound for the overall significance level can, for instance, be determined by the Bonferroni inequality. The Bonferroni bound would be the sum of the significance levels of the individual tests. However, as is well known, the Bonferroni bound may not be very sharp.

One practical problem is the choice of the number of volatility regimes and the volatility change points. In our derivations of the test statistic we assume that the investigator knows the volatility change points which may not be realistic in practice. The consequences of misspecifying the change points will be explored further in the next section. In practice, one may be inclined to base the choice of the volatility regimes on statistical procedures. However, that strategy may lead to a pretesting issue. Alternatively, volatility models such as the Markov switching model of Lanne et al. (2010) may be of interest to avoid specifying the volatility regimes exogenously. We note that such a model requires alternative asymptotic considerations that may result in a different asymptotic distribution of the test statistic. In any case, our asymptotic analysis requires that the sample size associated with each volatility regime goes to infinity. Hence, the test may not be suitable for volatility regimes associated with just a few sample points.

We emphasize that our test can also be applied if the slope coefficients are time-varying. Obviously, our test is based exclusively on the reduced-form residuals and the proxies. Clearly, suitable reduced-form residuals can also be obtained for VARs with time-varying slope coefficients. We have presented the test in a setup with time-invariant slope coefficients to simplify the notation and to focus attention on the impact effects of the shocks and the fact that time-varying shock transmission is possible even for a VAR with time-invariant slope coefficients, if there is heteroskedasticity. In our view, this point deserves attention because, in empirical work, it is not atypical to assume time-invariant shock transmission for heteroskedastic VAR models with time-invariant slope coefficients.

[3] Note that our asymptotic considerations rely on the assumption that C_m does not depend on the sample size and, hence, its rank is fixed at K_1 throughout the sample. This condition precludes weak proxies which are asymptotically uncorrelated with the shocks of interest. Weak instrument asymptotics is typically used to study the situation where the correlation between the proxies and the shocks is small (see, e.g., Montiel Olea et al. (2021)). In our case such asymptotics is complicated by the possibility that an asymptotic rank reduction of C_m causes identification problems. Therefore we leave weak proxy considerations for future research.

4. Monte Carlo simulations

We set up a Monte Carlo experiment to investigate the small sample properties of our test. As we expect the actual size and power properties of the test to depend on the sample size, the size of the VAR process (number of variables and lag order), the number and strength of the proxies as well as their correlation (among themselves and with the shocks) and the choice of the volatility change points, we use two different data generating processes (DGPs) to investigate the dependence of the small sample properties of our test on all these features.

The first DGP (DGP1) is based on DGP1 of Lütkepohl and Schlaak (2022). It has $M = 3$ volatility regimes and involves three variables ($K = 3$) and two proxies ($N = 2$) by which two ($K_1 = 2$) structural shocks are identified. We keep K, K_1, and N fixed for this DGP and we also assume that the volatility change points are known. These choices enable us to focus attention on changes in the sample size, the VAR lag order, the strength of the proxies and their correlation. The dependence of the small sample properties on other features will be explored in the context of DGP2.

As assumed by Lütkepohl and Schlaak (2022), DGP1 follows a three-variate VAR(1). We employ the same parameter values as Lütkepohl and Schlaak (2022) for A_1, $B(m)$, and Λ_m, i.e.,

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

$B(m) = I_3$ under $\mathbb{H}_0$, and

$$B(1) = I_3, \quad B(2) = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 4 \\ 4 & 6 & 6 \end{bmatrix}, \quad B(3) = \begin{bmatrix} 4 & 2 & 1 \\ -2 & 2 & 8 \\ 2 & 1 & 10 \end{bmatrix}$$

under $\mathbb{H}_1$. Hence,

$$B_{12}(1)B_{11}(1)^{-1} = [0,0], \quad B_{12}(2)B_{11}(2)^{-1} = [-8,6], \quad B_{12}(3)B_{11}(3)^{-1} = [0.5,0],$$

under $\mathbb{H}_1$. Clearly, $\beta(1) = (0,0)'$ is distinctly different from $\beta(2) = (-8,6)'$ and the latter vector is clearly distinct from $\beta(3) = (0.5,0)'$, while $\beta(1)$ and $\beta(3)$ are much closer together and we expect to have low power for testing $\mathbb{H}_0 : \beta(1) = \beta(3)$ given also our other parameter settings. Thereby we may get some insight in the power properties of our test under difficult scenarios where little power can be expected.

The other parameter settings are

$$\Lambda_1 = I_3, \quad \Lambda_2 = \text{diag}(4,9,12), \quad \Lambda_3 = \text{diag}(1,4,9),$$

and the proxies are generated as

$$z_t = \Xi \boldsymbol{w}_{1t} + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v), \quad \Xi = \begin{bmatrix} 1 & 0 \\ \rho & 1 \end{bmatrix}.$$

As Ξ is the covariance matrix $\text{cov}(z_t, \boldsymbol{w}_{1t})$, a nonzero ρ implies nonzero correlation between the second proxy, z_{2t}, and the first structural shock, w_{1t}. A value of $\rho = 0$ corresponds to the settings used by Lütkepohl and Schlaak (2022). We also consider $\rho = 0.5$ to investigate the implications of a single proxy being correlated with more than one structural shock and thereby violating the conditions underlying the test of Lütkepohl and Schlaak (2022).

The error v_t of the process generating the proxies is independent of the structural shocks and we choose covariance matrices

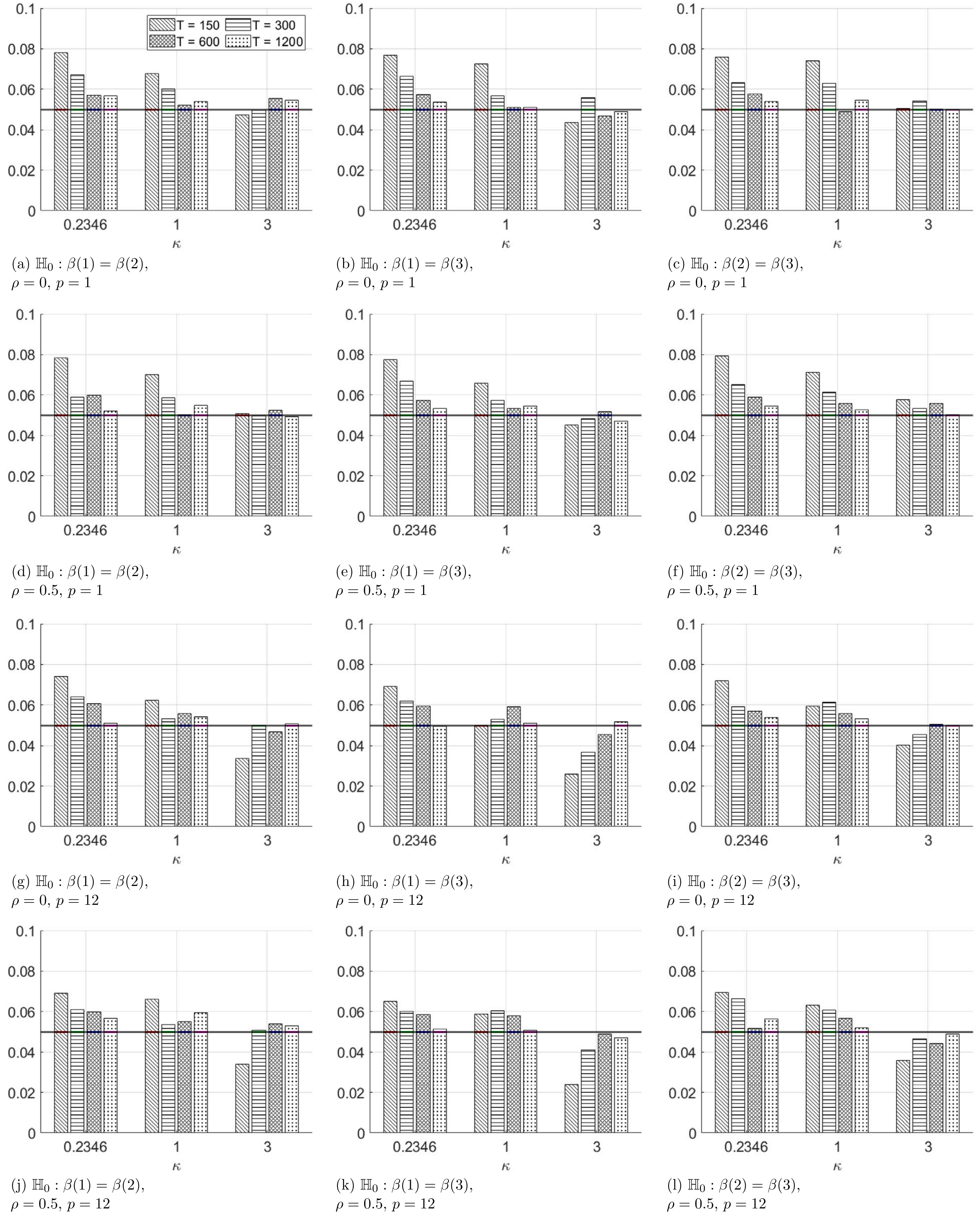
$$\Sigma_v = \kappa \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

Thus, the components of v_t are correlated and so are the proxies. For $\kappa = 1$ the proxies are of intermediate strength, while for $\kappa = 3$, their correlation to the shocks is lower and for $\kappa = 0.2346$, it is larger. The precise correlations between the structural shocks and the proxies are provided in the Online Appendix, where it can be seen that the correlations range from below 0.5 to more than 0.9 and allow us to explore the impact of the strength of the proxies on the properties of our test.[4]

Following Lütkepohl and Schlaak (2022), we generate samples of size $T = 150$, 300, 600, 1200. In the first volatility regime, we generate $T_1 = T/3 + p$ observations, where p is the VAR lag length used for estimation. We set $T_2 = 2T/3$. The estimated model includes a constant term, although the term is zero in the DGP. We use VAR lag orders $p = 1$ and 12 to explore the impact of having to deal with VAR models with larger lag orders. The number of replications for each simulation design is 5000.

Some results of our simulations based on DGP1 are summarized graphically in Figs. 1 and 2. The corresponding numerical results are presented in the Online Appendix. From the figures, the following observations emerge:

[4] In Table S.1 of their Online Supplement, Lütkepohl and Schlaak (2022) present the estimated correlations between proxies and structural shocks for a number of proxy VAR studies. These correlations range from 0.4 to 0.76. Thus, the correlations used for DGP1 are in the range of models used in applied work. Note that here we are not using the terminology 'weak proxies' for proxies with low correlation to the shocks to avoid confusion with a 'weak proxy setup', where the correlation between the shocks and the proxies vanishes asymptotically.



(a) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0$, $p = 1$

(b) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0$, $p = 1$

(c) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0$, $p = 1$

(d) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0.5$, $p = 1$

(e) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0.5$, $p = 1$

(f) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0.5$, $p = 1$

(g) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0$, $p = 12$

(h) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0$, $p = 12$

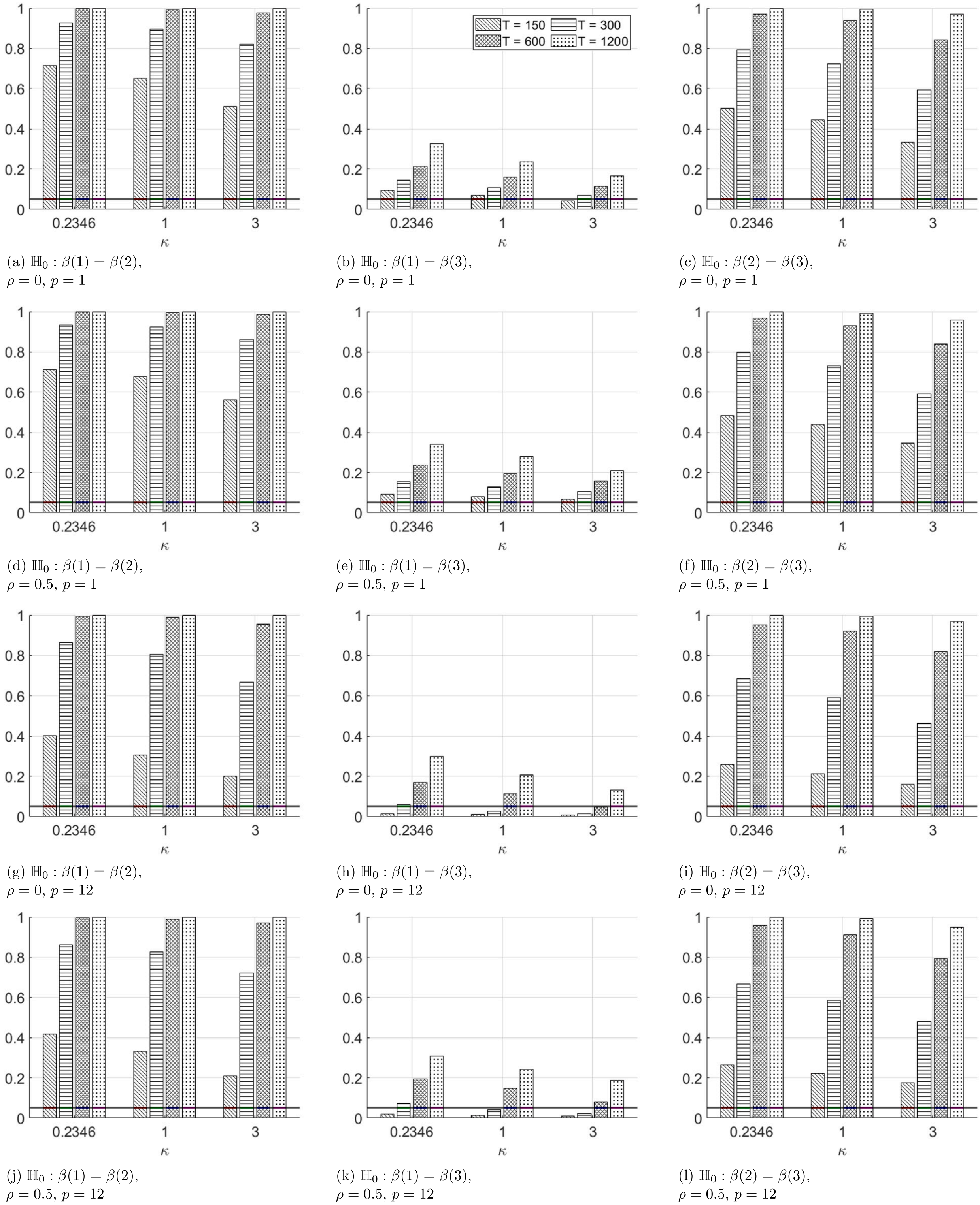
(i) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0$, $p = 12$

(j) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0.5$, $p = 12$

(k) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0.5$, $p = 12$

(l) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0.5$, $p = 12$

Fig. 1. Relative rejection frequencies for DGP1 under $\mathbb{H}_0$. Nominal significance level 5%.



(a) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0$, $p = 1$

(b) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0$, $p = 1$

(c) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0$, $p = 1$

(d) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0.5$, $p = 1$

(e) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0.5$, $p = 1$

(f) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0.5$, $p = 1$

(g) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0$, $p = 12$

(h) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0$, $p = 12$

(i) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0$, $p = 12$

(j) $\mathbb{H}_0 : \beta(1) = \beta(2)$, $\rho = 0.5$, $p = 12$

(k) $\mathbb{H}_0 : \beta(1) = \beta(3)$, $\rho = 0.5$, $p = 12$

(l) $\mathbb{H}_0 : \beta(2) = \beta(3)$, $\rho = 0.5$, $p = 12$

Fig. 2. Relative rejection frequencies for DGP1 under $\mathbb{H}_1$ (power). Nominal significance level 5%.

Sample size: The sample size has the expected effect. Increasing it generally moves the rejection frequencies closer to the 5% nominal level (see Fig. 1) and it improves the power (see Fig. 2). Note, however, that the relative rejection frequencies are close to 5% already for relatively small sample sizes if the null hypothesis holds. Clearly, although $T = 150$ is a rather standard sample size in macroeconometric analysis, it is a small sample size for our test. It leaves only 50 observations for each of the three volatility regimes, which is not very much for a 3-dimensional VAR(12) model. Even for such small samples, the simulated rejection frequencies are close to the 5% nominal level. For sample size $T = 600$ the relative rejection frequencies under $\mathbb{H}_0$ are all between 0.044 and 0.061 in Fig. 1.

Distance from $\mathbb{H}_0$: As one would expect, the distance of parameter values under the alternative from the parameter space under the null hypothesis is crucial for the power of the test. In Fig. 2 it is obvious that the power for testing $\mathbb{H}_0 : \beta(1) = \beta(3)$ is much lower than for the other null hypotheses. As discussed in the previous subsection, under $\mathbb{H}_1$, $\beta(1)$ and $\beta(3)$ are relatively close together and this is very clearly reflected in the low relative rejection frequencies in Fig. 2. In fact, in panels (h) and (k) of the figure, the rejection frequencies are less than the nominal significance level of 5% for small sample sizes. In other words, the test may have very low power in small samples if the transformed impact effects to be tested are close to each other in both volatility regimes under test.

Lag order: Larger VAR lag orders lead to more parameters in the model and, hence, increase the overall estimation uncertainty. Thus, it is not surprising that our test tends to have reduced power for longer lag orders, as can be seen in Fig. 2 (compare the upper panels to the panels in the lower half). On the other hand, the results in Fig. 1 indicate that the rejection frequencies under $\mathbb{H}_0$ are not much affected by the lag order.

Proxy strength: In Fig. 2, it can also be seen that stronger proxies (smaller κ) lead to larger power of our test in smaller samples, as one would expect. The effect is reduced or vanishes for the larger sample sizes reported in the figure. Again, the relative rejection frequencies under $\mathbb{H}_0$ in Fig. 1 are not much affected by the proxy strength. Although there are differences in the rejection frequencies for varying κ under $\mathbb{H}_0$, the empirical frequencies are still all close to the nominal 5%.

Proxy-shock correlation: Comparing the situation where each proxy is correlated only to a single shock ($\rho = 0$) to a simulation design where the second proxy is correlated with both shocks ($\rho = 0.5$) in Figs. 1 and 2, shows that there is not much difference in the outcomes of the tests. In other words, the panels in the second and fourth rows (corresponding to $\rho = 0.5$) of the two figures are qualitatively similar to the first and third rows (corresponding to $\rho = 0$), respectively. We conclude that, whether or not a proxy is correlated with more than one structural shock does not seem to matter much for the small sample properties of our test. In further simulations we also considered $\rho = 0.9$ and got similar results which are therefore not presented. Given these results, we consider the situation where each proxy is correlated only with a single shock in the scenarios studied for DGP2.

The details of the second DGP are presented in the Online Appendix. It allows us to study the dependence of the small sample properties of the test on the number of variables in the model, the number of proxies, and the choice of the volatility change points. The results confirm the conclusions from the simulations of DGP1 regarding the impact of the sample size, the VAR lag order, and the proxy strength and also show that increasing the dimension of the process or the number of proxies tends to reduce the power of the test for time-varying impact effects. Preselecting the volatility change points by the likelihood-based statistical criterion presented in the Online Appendix has little impact on the properties of the test. Also, misspecifying the change point slightly does not affect the properties of the test much. For the details see the Online Appendix.

In summary, our simulations based on two types of DGPs suggest that the test for time-varying impulse responses has small sample rejection frequencies close to the nominal significance level under $\mathbb{H}_0$ and good power under $\mathbb{H}_1$ for moderate and large samples. For large models, small samples, and proxies of lower strength, the test tends to be conservative. To ensure good power, the parameter values under $\mathbb{H}_1$ have to be clearly distinct from the values under $\mathbb{H}_0$. Generally the strength of the proxies (their correlation with the structural shocks to be identified) is important for achieving good power properties. A higher-dimensional model with more parameters tends to reduce the power and also having more proxies to identify more shocks may reduce the power of the test. Finally, the correct choice of the volatility change point is of very limited importance for the small sample properties of the test. Moreover, preselecting the change point by our statistical criterion has very little impact on the properties of the test. The latter finding is potentially important for empirical work because, in practice, it is often uncertain where exactly a volatility change has occurred so that the actual change point is estimated by a statistical procedure or placed not exactly in the correct period. In the next section, we discuss an empirical example which illustrates the virtue of the test for applied work.

5. The impact of TFP shocks on the U.S. economy

We use a benchmark study of Lunsford (2015) to illustrate the benefits of applying our test in a proxy VAR analysis. Lunsford (2015) investigates the dynamic effects of two types of total factor productivity (TFP) shocks. The quarterly SVAR model uses data from 1948Q2 to 2015Q2 which implies a total of 269 observations. The number of endogenous variables is $K = 5$ (GDP growth, employment growth, inflation, consumption growth, investment growth), that is, y_t is 5-dimensional. There are $N = 2$ proxies to identify $K_1 = 2$ TFP shocks and the VAR model has $p = 4$ lags[5] and a constant.[6]

[5] Lunsford (2015) states that he uses 3 lags. However, in a later revised version of the paper, $p = 4$ is claimed. Thus, we use the latter lag length.

[6] The dataset is available at https://sites.google.com/site/kurtglunsford/research.

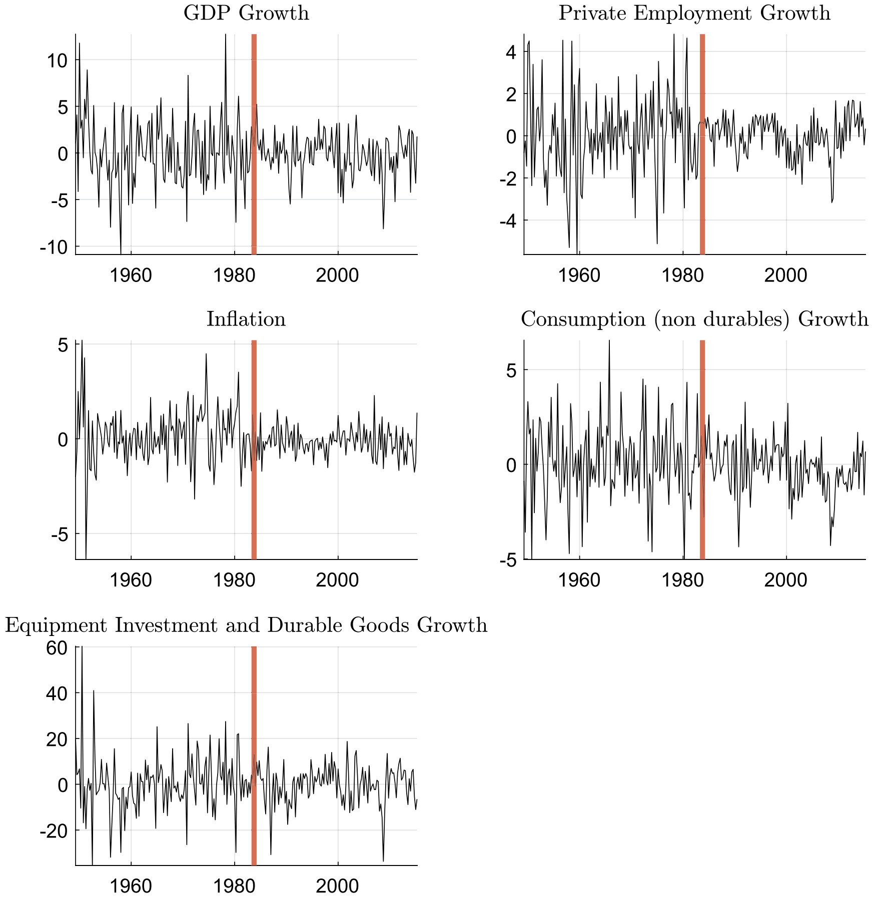


Fig. 3. OLS residuals from 1949Q2 to 2015Q2 for the VAR(4) U.S. macro model. The solid red line indicates the possible variance change point in 1983Q4. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

The two proxies are based on two utilization-adjusted TFP measures constructed by Fernald (2014), one for the consumption sector excluding durable goods (consumption TFP) and the second one for durable goods and equipment investment (investment TFP). The proxies are constructed by regressing consumption TFP and investment TFP on four lags of the y_t and a constant and using the resulting residuals as proxies.

Fig. 3 plots the OLS residuals of the VAR(4) model. Clearly, the residuals display some change in volatility around the time when the Great Moderation (GM) started. The GM is typically assumed to have started in the middle of the 1980s, but the exact date is not clear and different authors make different assumptions regarding the starting date. McConnell and Perez-Quiros (2000) and Galí and Gambetti (2009) place it at the beginning of 1984.[7] Therefore, in the following, the period up to 1983Q4 will be referred to as the pre-GM period and the period from 1984Q1 will be called the post-GM regime.[8]

Computing the test statistic $\eta(1,2)$ for the possible volatility change point 1983Q4, the resulting test value is $\eta(1,2) = 12.364$, which corresponds to a p-value of 0.054 for a χ^2-distribution with 6 degrees of freedom and, thus, the test rejects the null hypothesis at a significance level of 10%. As it is not clear where exactly the GM started, we also compute the test statistic $\eta(1,2)$ for neighboring quarters of 1983Q4 and show the results in Table 1. Obviously, most p-values are below 10% and the p-values for 1984Q1-1985Q1 are even below 5%. These results are consistent with our simulation results which indicate that slight misspecifications of the actual change in volatility do not make much of a difference for the test outcome in small samples.

Given that the sample size for the present example is relatively small and taking into account our simulation findings regarding the power of the test, the results in Table 1 provide substantial evidence for a change in the impact effects of the TFP shocks

[7] Using monthly data in monetary studies, Bernanke and Mihov (1998) and Christiano et al. (1999) consider a potential regime change in 1984M2, while Stock and Watson (2003) assume that the GM started in 1983-1985. Specifically they point out that tests on quarterly GDP growth suggest a change in 1982Q4 to 1985Q3 (see p. 161 of their article).

[8] We have also used the likelihood based statistical criterion presented in the Online Appendix to determine the volatility change point. The criterion places the change in 1982Q4 if only one change point is considered. If we allow for a second volatility change point, the criterion suggests 1976Q2. We have also investigated the possibility of changing impact effects of the shocks in that period and did not find evidence for it. Therefore, for expository purposes, we discuss the results for a volatility change in 1983Q4 only and account for potential further volatility changes by using heteroskedasticity robust inference.

Table 1
Tests for time-varying impact effects.

T_1	test statistic	p-value
1982Q3	11.004	0.088
1982Q4	11.282	0.080
1983Q1	10.980	0.089
1983Q2	10.600	0.102
1983Q3	11.953	0.063
1983Q4	12.364	0.054
1984Q1	13.013	0.043
1984Q2	13.730	0.033
1984Q3	13.712	0.033
1984Q4	13.679	0.033
1985Q1	12.987	0.043
1985Q2	12.533	0.051
1985Q3	12.512	0.051
1985Q4	12.234	0.057

around 1983Q4. Despite this evidence, we compute impulse responses under both alternative scenarios, with time-varying as well as time-invariant impact effects, to compare time-invariant to time-varying dynamic responses of the variables.

Computing impulse responses raises the question of how to separately identify the two TFP shocks. Recall that the standard procedure for identification through heteroskedasticity is not available here because we have rejected time-invariance of the impact effects at the time of the volatility change. Lunsford (2015) uses the two proxies individually to identify one shock at a time. He justifies his approach by pointing out the low empirical correlation of each proxy with the structural shock estimated with the other proxy. If each of the two TFP proxies is only correlated with one of the two TFP shocks, such an approach is indeed justified. In that case, $\mathbb{E}(\boldsymbol{w}_{1t} z_t') = C_m$ would be a diagonal matrix which in turn implies that the columns of $D(m) = B_1(m)C_m$ are scalar multiples of the impact effects of the two TFP shocks. In other words, using $D(m)$ as matrix of impact effects of the two shocks would be justified if we are only interested in the shape of the impulse responses but not in the size of the shock or if the shock size is fixed by some other consideration anyway. For the TFP proxies and shocks, Lunsford (2015, Table 3) provides evidence that C_m is indeed diagonal. Thus, we use this device for computing impulse responses of the two shocks.

We present results for shocks of size one standard deviation, where the standard deviation refers to the period for which the impulse response function is computed. In other words, for the impulse responses for the whole sample, the standard deviation for the full sample is used, while the associated subsample standard deviations are considered for the impulse responses for the two subperiods. Alternatively, one could standardize the impact response of one of the variables to 1 (e.g., Stock and Watson (2018)). A drawback of that approach is that we have to take a stand on which variable has a nonzero impact response across all volatility regimes. For the system under consideration, there is no obvious candidate variable that clearly reacts on impact to both shocks in all regimes. Therefore we prefer our approach which is also more in line with Lunsford (2015).

The impulse responses together with 90% confidence intervals are displayed in Fig. 4. The confidence intervals are generated by a residual-based moving-block bootstrap (MBB) as proposed by Brüggemann et al. (2016) and Jentsch and Lunsford (2019, 2022). These authors show the asymptotic validity of the method for inference in structural VAR analysis even under time-varying volatility. We implement the MBB exactly as in Bruns and Lütkepohl (2022).[9]

The impulse responses in Fig. 4 show that some of the dynamic effects of the two TFP shocks clearly depend on the regime if we allow for time-varying impulse responses. For example, the inflation responses to a consumption TFP shock in the pre- and post-GM regimes have non-overlapping confidence intervals. In the pre-GM regime, inflation declines in response to a positive consumption TFP shock while it is not clear that inflation responds at all if a consumption TFP shock hits in the post-GM period. Likewise, the initial effects of an investment TFP shock on GDP growth and employment growth are clearly different if we allow for a change in 1983Q4. In the pre-GM period, the initial effects of investment TFP shocks are clearly stronger than in the post-GM regime.

In Tables 2 and 3 we present the corresponding point estimates of the forecast error variance contributions of the two TFP shocks. They are also clearly distinct pre- and post-GM. For example, the contributions of the two shocks to the forecast error variance of inflation are much smaller post-GM than pre-GM, reflecting the weaker post-GM response of inflation to the TFP shocks observed in Fig. 4. Moreover, in Tables 2 and 3 the consumption TFP shock contributes much less to GDP growth pre-GM than post-GM, while the situation is reversed for the investment TFP shock. Thus, the forecast error variance decompositions reinforce the conclusion that the transmission of the shocks has changed around the onset of the GM.

Thus, the impulse responses in Fig. 4 and the forecast error variance decompositions in Tables 2 and 3 support the conclusion drawn from our test for time-varying impulse responses that there may have been a change in the dynamic responses at the time of the onset of the GM. Interestingly, the impulse responses and forecast error variance contributions obtained under the assumption of time-invariant impulse responses (shown as a solid black line in Fig. 4 and in the upper panels of Tables 2 and 3) lie by and large

[9] We use the MBB separately for the pre- and post-GM periods. The block length is chosen according to the rule of thumb from Jentsch and Lunsford (2019), $\ell_m \approx 5.03(\tau_m T)^{0.25}$, i.e., we use $\ell_1 = 18$ and $\ell_2 = 17$ for the first and second volatility regime, respectively.

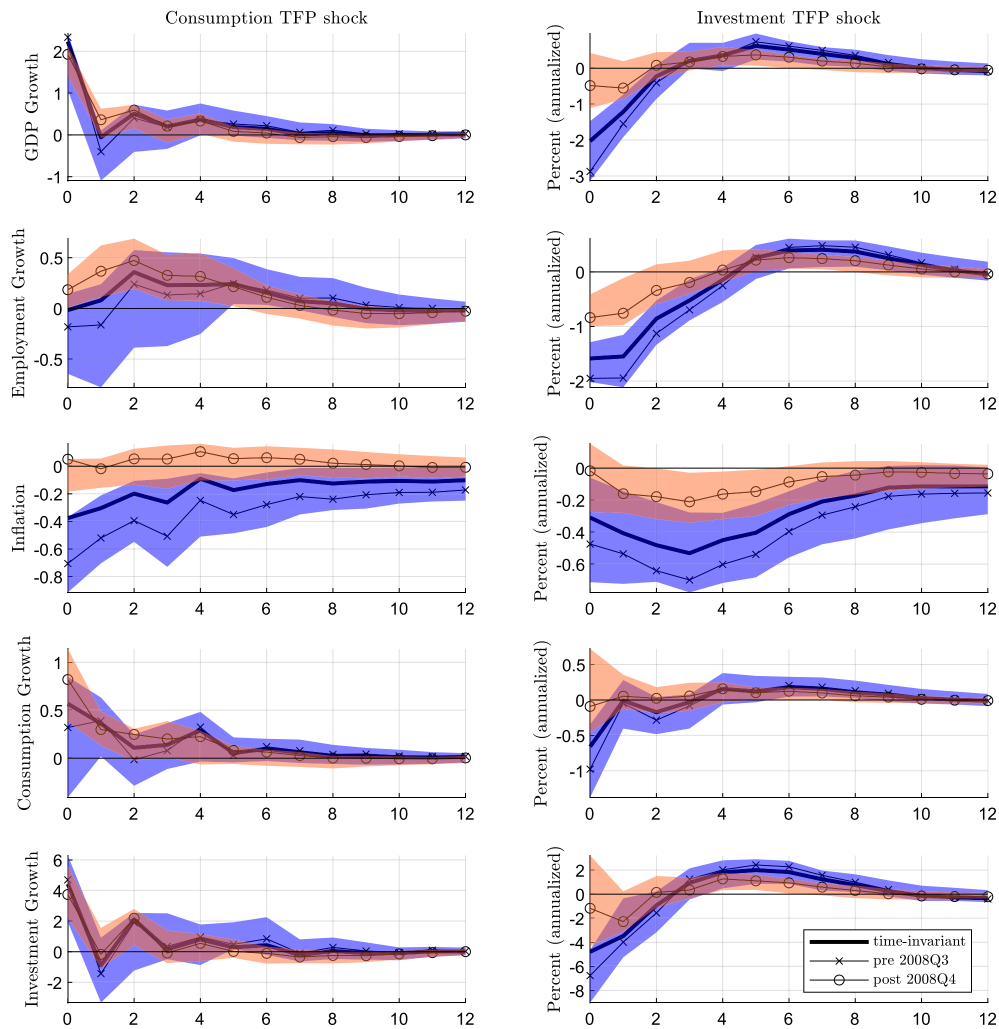


Fig. 4. Comparison of responses to TFP shocks assuming time-varying or, alternatively, time-invariant dynamic responses with 90% pointwise confidence intervals based on a MBB. (Volatility regime 1: 1949Q2 - 1983Q4 (blue); Volatility regime 2: 1984Q1 - 2015Q2 (red); Time-invariant impulse responses: solid black line.)

in between the impulse responses and forecast error variance contributions estimated for the pre- and post-GM regimes. Thus, they may just average the responses in the two different regimes which leads to a gross distortion in some cases.

Despite some differences in the impulse responses associated with the two different volatility regimes, some of the main features observed by Lunsford (2015) are maintained if we allow for a change in the dynamic responses of the variables to the TFP shocks in 1983Q4.[10] A main feature that is maintained from Lunsford (2015) is that the consumption TFP shock can be interpreted as a supply shock in the post-GM regime in that a positive consumption shock leads to an increase in investment growth, GDP growth, and employment growth. The variables respond in a similar manner in the pre-GM regime. In that regime the inflation response is clearly negative while GDP growth and investment growth increase initially. The responses of consumption and employment are less certain, as reflected in their larger confidence intervals.

Like in Lunsford (2015), the investment TFP shock is not consistent with a supply shock because inflation moves in the same direction as the other growth rates. In Fig. 4 the investment TFP shock is clearly a negative shock to all quantities and inflation in the pre-GM period, while the responses of some variables are not clear in the post-GM regime. In fact, only employment growth declines clearly on impact, whereas the 90% confidence intervals of all other variables include zero in the post-GM period. Hence, the impact response to an investment TFP shock is not clear for these variables. Lunsford discusses a rationalization of the responses to this shock in the context of a DSGE model.

In summary, it is important to notice that allowing for time-varying impulse responses leads to evidence of quite distinct dynamic effects before and after the onset of the GM. Some of these differences are quite relevant for economic policy. For example, for a

[10] Note that the solid black line in Fig. 4 represents the impulse responses obtained under time-invariance when both proxies are used jointly as explained earlier. Apart from their scale, they are very similar but not identical to the impulse responses in Lunsford (2015) who computes them one at a time.

Table 2
Share of h-step ahead forecast error variance explained by consumption TFP shock.

time-invariant					
h	1	3	6	9	12
GDP Growth	0.53	0.40	0.38	0.36	0.36
Employment Growth	0.00	0.02	0.04	0.04	0.04
Inflation	0.10	0.10	0.09	0.09	0.09
Consumption Growth	0.10	0.11	0.13	0.13	0.13
Investment Growth	0.17	0.15	0.14	0.14	0.14
pre-GM					
h	1	3	6	9	12
GDP Growth	0.40	0.30	0.29	0.28	0.28
Employment Growth	0.01	0.01	0.02	0.02	0.02
Inflation	0.22	0.22	0.21	0.20	0.20
Consumption Growth	0.02	0.05	0.06	0.06	0.06
Investment Growth	0.14	0.12	0.12	0.12	0.12
post-GM					
h	1	3	6	9	12
GDP Growth	0.77	0.62	0.59	0.58	0.58
Employment Growth	0.04	0.15	0.19	0.19	0.19
Inflation	0.00	0.01	0.02	0.02	0.02
Consumption Growth	0.36	0.35	0.35	0.35	0.35
Investment Growth	0.21	0.21	0.20	0.19	0.19

Table 3
Percent of h-step ahead forecast error variance explained by investment TFP shock.

time-invariant					
h	1	3	6	9	12
GDP Growth	0.44	0.43	0.43	0.44	0.44
Employment Growth	0.98	0.84	0.75	0.75	0.75
Inflation	0.06	0.18	0.27	0.27	0.26
Consumption Growth	0.14	0.11	0.11	0.13	0.13
Investment Growth	0.20	0.22	0.25	0.27	0.27
pre-GM					
h	1	3	6	9	12
GDP Growth	0.61	0.57	0.55	0.56	0.56
Employment Growth	0.92	0.84	0.77	0.77	0.76
Inflation	0.10	0.22	0.31	0.31	0.30
Consumption Growth	0.22	0.19	0.18	0.19	0.19
Investment Growth	0.29	0.28	0.31	0.33	0.33
post-GM					
h	1	3	6	9	12
GDP Growth	0.05	0.08	0.11	0.13	0.13
Employment Growth	0.80	0.53	0.44	0.46	0.46
Inflation	0.00	0.06	0.10	0.09	0.09
Consumption Growth	0.00	0.00	0.02	0.03	0.03
Investment Growth	0.02	0.07	0.10	0.11	0.11

central bank it makes a difference whether a shock to consumption leaves inflation untouched or moves it up or down. Thus, if there is heteroskedasticity, it is worth investigating its impact on the impulse responses. In other words, it makes sense to use our test in this situation in a proxy VAR analysis.

6. Conclusions

In structural VAR analysis based on heteroskedastic models, the dynamic effects of the structural shocks may change with the volatility of the shocks. We propose a test for time-varying impulse responses for heteroskedastic proxy VAR models, where a set of proxies identifies a set of shocks collectively but not necessarily individually. In such a situation, it is typically necessary to provide further information to identify the structural shocks of interest individually. The proposed test for time-varying impact effects does not require such additional information for individually identifying the shocks. It is robust to the scheme used for individually identifying the shocks and, hence, it can also be applied if no information is available for identifying the shocks individually.

We have derived the asymptotic properties of the test and also present simulation results to investigate the performance of the test in small samples. The Monte Carlo simulations show that larger samples and stronger proxies improve the small sample power of the test, while larger lag orders, larger dimensions of the underlying VAR process as well as larger numbers of proxies and shocks to be identified tend to reduce power. Misspecifying the volatility change points as well as using estimated instead of true change points has very little impact on the power. The latter property is, of course, particularly helpful for empirical work, where uncertainty regarding the volatility change points is not uncommon.

Our results suggest that our test is useful for applied work whenever there are different volatility regimes in a proxy VAR model, provided the volatility regimes are long enough for reliable estimation of the regime dependent quantities in each regime. Our simulation results show that the test works even if the volatility regimes are relatively short. In that case, the power of the test may be low, however. We have also emphasized that exact knowledge of the volatility change points is not essential for the performance of the test.

We have applied our tests to investigate the time-invariance of two TFP shocks in a U.S. macroeconomic model. We have found that the dynamic effects of the shocks may have changed during the GM period. The impulse responses and forecast error variance components of some of the variables to the TFP shocks are clearly distinct if we allow for a change in 1983Q4. For example, a positive consumption TFP shock is found to reduce inflation in the pre-GM period, while it may have little or no impact on inflation post-GM. Clearly, such differences would be relevant for economic policy action. Thus, it is important to explore the time-invariance of the dynamic effects of structural shocks in heteroskedastic proxy VAR models.

In our study we have assumed a rather simple model for the underlying heteroskedasticity in that we assume that there is a finite number of volatility regimes during the sample period. In practice, more complicated mechanisms may generate the volatility change. Examples are smooth changes of the volatility over a number of periods or conditional heteroskedasticity. Investigations by Lütkepohl and Schlaak (2022) for the case of a single proxy variable suggest that our test may even work under such more general volatility models. This issue would be an interesting topic for future research.

Appendix A. Construction of the covariance matrices $V(m)$

In this appendix we provide the derivatives needed for the construction of the covariance matrices $V(m), V(k)$ which are part of the test statistic $\eta(m,k)$ and we drop the arguments m or k for simplicity. We will make use of commutation matrices $\mathbf{K}_{ij}$ defined such that $\text{vec}(A') = \mathbf{K}_{ij}\text{vec}(A)$ for any $(i \times j)$ matrix A. We use rules for vector and matrix differentiation from (Lütkepohl, 1996, Chapter 10) and provide precise numbers of rules from that source, where they are used in the following.

Note that

$$\frac{\partial \beta}{\partial \text{vec}(D)'} = \frac{\partial \text{vec}\{D_2 H D_1' [D_1 H D_1']^{-1}\}}{\partial[\text{vec}(D_1)'\text{vec}(D_2)']} \times \frac{\partial \begin{bmatrix} \text{vec}(D_1) \\ \text{vec}(D_2) \end{bmatrix}}{\partial \text{vec}(D_1')'\text{vec}(D_2')'} \times \frac{\partial \begin{bmatrix} \text{vec}(D_1') \\ \text{vec}(D_2') \end{bmatrix}}{\partial \text{vec}(D)'}$$
$$(K_1(K-K_1) \times KN) \tag{9}$$

(see Sec. 10.7, Rule (2)). Closed-form expressions for the three matrices of partial derivatives on the right-hand side of this equation will be derived in the following.

For the first matrix of partial derivatives we get

$$\frac{\partial \text{vec}\{D_2 H D_1' [D_1 H D_1']^{-1}\}}{\partial[\text{vec}(D_1)'\text{vec}(D_2)']} = \left[\frac{\partial \text{vec}\{D_2 H D_1' [D_1 H D_1']^{-1}\}}{\partial \text{vec}(D_1)'} : \frac{\partial \text{vec}\{D_2 H D_1' [D_1 H D_1']^{-1}\}}{\partial \text{vec}(D_2)'} \right],$$

where

$$\frac{\partial \text{vec}\{D_2 H D_1' [D_1 H D_1']^{-1}\}}{\partial \text{vec}(D_1)'} = ([D_1 H D_1']^{-1} \otimes I_{K-K_1}) \frac{\partial \text{vec}(D_2 H D_1')}{\partial \text{vec}(D_1)'} - ([D_1 H D_1']^{-1} \otimes D_2 H D_1' [D_1 H D_1']^{-1}) \frac{\partial \text{vec}(D_1 H D_1')}{\partial \text{vec}(D_1)'}$$

(see Sec. 10.6.3, Rule (3)) with

$$\frac{\partial \text{vec}(D_1 H D_1')}{\partial \text{vec}(D_1)'} = [(D_1 H \otimes I_{K_1}) + (I_{K_1} \otimes D_1 H)\mathbf{K}_{K_1 N}] \quad (K_1^2 \times K_1 N)$$

(see Sec. 10.5.1, Rule (6)) and

$$\frac{\partial \text{vec}(D_2 H D_1')}{\partial \text{vec}(D_1)'} = (I_{K_1} \otimes D_2 H)\mathbf{K}_{K_1 N} \quad (K_1(K-K_1) \times K_1 N)$$

(see Sec. 10.4.1, Rule (4)).

$$\frac{\partial \text{vec}\{D_2 H D_1' [D_1 H D_1']^{-1}\}}{\partial \text{vec}(D_2)'} = [D_1 H D_1']^{-1} D_1 H \otimes I_{K-K_1} \quad (K_1(K-K_1) \times (K-K_1)N)$$

(see Sec. 10.4.1, Rule (3)).

For the second matrix on the right-hand side of equation (9) we obtain

$$\frac{\partial \begin{bmatrix} \text{vec}(D_1) \\ \text{vec}(D_2) \end{bmatrix}}{\partial [\text{vec}(D_1')' \text{vec}(D_2')']} = \begin{bmatrix} \mathbf{K}_{NK_1} & 0_{K_1 N \times (K-K_1)N} \\ 0_{(K-K_1)N \times K_1 N} & \mathbf{K}_{N(K-K_1)} \end{bmatrix} \quad (KN \times KN)$$

(see Sec. 10.4.1, Rule (1)).

Finally, for the last term in (9) we get

$$\frac{\partial \begin{bmatrix} \text{vec}(D_1') \\ \text{vec}(D_2') \end{bmatrix}}{\partial \text{vec}(D)'} = \frac{\partial \text{vec}(D')}{\partial \text{vec}(D)'} = \mathbf{K}_{KN} \quad (KN \times KN)$$

(see Sec. 10.4.1, Rule (1)).

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jedc.2024.104837.

References

Angelini, G., Bacchiocchi, E., Caggiano, G., Fanelli, L., 2019. Uncertainty across volatility regimes. J. Appl. Econom. 34, 437–455.
Arias, J.E., Rubio-Ramírez, J.F., Waggoner, D.F., 2021. Inference in Bayesian Proxy-SVARs. J. Econom. 225 (1), 88–106.
Bacchiocchi, E., Castelnuovo, E., Fanelli, L., 2018. Gimme a break! Identification and estimation of the macroeconomic effects of monetary policy shocks in the United States. Macroecon. Dyn. 22, 1613–1651.
Bacchiocchi, E., Fanelli, L., 2015. Identification in structural vector autoregressive models with structural changes, with an application to US monetary policy. Oxf. Bull. Econ. Stat. 77, 761–779.
Bernanke, B.S., Mihov, I., 1998. Measuring monetary policy. Q. J. Econ. 113, 869–902.
Braun, R., Brüggemann, R., 2023. Identification of SVAR models by combining sign restrictions with external instruments. J. Bus. Econ. Stat. 41, 1077–1089.
Brüggemann, R., Jentsch, C., Trenkler, C., 2016. Inference in VARs with conditional heteroskedasticity of unknown form. J. Econom. 191, 69–85.
Bruns, M., Lütkepohl, H., 2022. An alternative bootstrap for proxy vector autoregressions. Comput. Econ. https://doi.org/10.1007/s10614-022-10323-w. forthcoming.
Carriero, A., Marcellino, M., Tornese, T., 2021. Blended identification in structural VARs. Technical report. Queen Mary University of London.
Christiano, L.J., Eichenbaum, M., Evans, C., 1999. Monetary policy shocks: what have we learned and to what end? In: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics, vol. 1A. Elsevier, Amsterdam, pp. 65–148.
Fernald, J., 2014. A quarterly, utilization-adjusted series on total factor productivity. Working paper. Federal Reserve Bank of San Francisco.
Galí, J., Gambetti, L., 2009. On the sources of the great moderation. Am. Econ. J. Macroecon. 1, 26–57.
Härtl, T., 2022. Identifying Proxy VARs with Restrictions on the Forecast Error Variance. Technical Report, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics. ZBW-Leibniz Information Centre for Economics, Kiel, Hamburg.
Herwartz, H., Lütkepohl, H., 2014. Structural vector autoregressions with Markov switching: combining conventional with statistical identification of shocks. J. Econom. 183, 104–116.
Jentsch, C., Lunsford, K.G., 2019. The dynamic effects of personal and corporate income tax changes in the United States: comment. Am. Econ. Rev. 109, 2655–2678.
Jentsch, C., Lunsford, K.G., 2022. Asymptotically valid bootstrap inference for proxy SVARs. J. Bus. Econ. Stat. 40, 1876–1891.
Kilian, L., Lütkepohl, H., 2017. Structural Vector Autoregressive Analysis. Cambridge University Press, Cambridge.
Lanne, M., Lütkepohl, H., 2008. Identifying monetary policy shocks via changes in volatility. J. Money Credit Bank. 40, 1131–1149.
Lanne, M., Lütkepohl, H., Maciejowska, K., 2010. Structural vector autoregressions with Markov switching. J. Econ. Dyn. Control 34, 121–131.
Lunsford, K., 2015. Identifying structural VARs with a proxy variable and a test for a weak proxy. Technical report. Federal Reserve Bank of Cleveland.
Lütkepohl, H., 1996. Handbook of Matrices. John Wiley & Sons, Chichester.
Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin.
Lütkepohl, H., Schlaak, T., 2022. Heteroscedastic proxy vector autoregressions. J. Bus. Econ. Stat. 40, 1268–1281.
McConnell, M.M., Perez-Quiros, G., 2000. Output fluctuations in the United States: what has changed since the early 1980's? Am. Econ. Rev. 90, 1464–1476.
Mertens, K., Ravn, M.O., 2013. The dynamic effects of personal and corporate income tax changes in the United States. Am. Econ. Rev. 103, 1212–1247.
Montiel Olea, J.L., Stock, J.H., Watson, M.W., 2021. Inference in structural vector autoregressions identified with an external instrument. J. Econom. 225, 74–87.
Perron, P., 2006. Dealing with structural breaks. In: Mills, T.C., Patterson, K. (Eds.), Palgrave Handbook of Econometrics, Volume 1, Econometric Theory. Palgrave Macmillan, Houndmills, Basingstoke, pp. 278–352.
Piffer, M., Podstawski, M., 2017. Identifying uncertainty shocks using the price of gold. Econ. J. 128 (616), 3266–3284.
Stock, J.H., Watson, M.W., 2003. Has the business cycle changed and why? NBER Macroeconomics Annual 2002, Volume 17. NBER Chapters. National Bureau of Economic Research, Inc., pp. 159–230.
Stock, J.H., Watson, M.W., 2012. Disentangling the channels of the 2007-09 recession. Brookings Pap. Econ. Act., 81–135.
Stock, J.H., Watson, M.W., 2018. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. Econ. J. 128 (610), 917–948.