School of Economics Working Paper
2022-02

Heteroskedastic Proxy Vector Autoregressions: Testing for Time-Varying Impulse Responses in the Presence of Multiple Proxies

Martin Bruns *
Helmut Luetkepohl **

*University of East Anglia
**DIW Berlin and Freie Universitaet Berlin

University of East Anglia

SCHOOL OF ECONOMICS

School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/economics

Heteroskedastic Proxy Vector Autoregressions: Testing for Time-Varying Impulse Responses in the Presence of Multiple Proxies

Martin Bruns
University of East Anglia, School of Economics,
Norwich Research Park, NR4 7TJ, Norwich, United Kingdom
email: martin.j.bruns@gmail.com

and

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

This version: May 3, 2022[1]

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. The test 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.

Key Words: Structural vector autoregression, proxy VAR, heteroskedasticity, productivity shocks

JEL classification: C32

[1]The authors thank Robin Braun and Kurt Lunsford for helpful comments on an earlier draft of this paper.

1 Introduction

Using external instruments or proxies to identify shocks of interest in structural vector autoregressive (VAR) analysis has become increasingly popular lately and is now sometimes signified as proxy VAR analysis. In some studies, a set of proxies is used to identify a group of shocks collectively. In that case, it is typically necessary to provide additional information to identify the shocks of interest individually. 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 (2020), or Arias, Rubio-Ramírez and Waggoner (2021)). If such restrictions are controversial or not available, one may also want to take advantage of identifying information extracted from the distributional features of the data. Notably heteroskedasticity has been used in this context (see, e.g., Kilian and Lütkepohl (2017, Chapter 14) or Lütkepohl and Netšunajev (2017)).

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 much of the literature this assumption is used without further investigation. However, some authors question this time-invariance assumption (Bacchiocchi, Castelnuovo and Fanelli (2018), Bacchiocchi and Fanelli (2015)). In a recent article, Lütkepohl and Schlaak (2021) 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 identified by one or more proxies. They note that the situation, where a number of shocks are identified collectively and not individually by a set of proxies, is more difficult to handle because, in order to apply their test, additional information is needed to identify the shocks individually. Hence, their test cannot be used if a set of proxies identifies a collection of shocks jointly but not individually.

In this study we propose a test for time-varying impact effects which can even be applied if a set of shocks is 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. We confirm good small sample size and power properties of the test in a Monte Carlo study.

The test is 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.

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 contains a number of theoretical derivations, 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 u_t is a zero mean white noise process with nonsingular, possibly time-varying covariance matrix Σ_t and ν is a time-invariant intercept vector. In short, $u_t \sim (0, \Sigma_t)$. In other words, there may be heteroskedasticity. The heteroskedasticity is assumed to be such that

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

where $\mathcal{T}_m = \{T_{m-1}+1, \dots, T_m\}$ $(m = 1, \dots, M)$ are M volatility regimes. The volatility changes at the end of time periods T_m for $m = 1, \dots, M-1$, 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.

The structural errors, $\boldsymbol{w}_t = (w_{1t}, \dots, w_{Kt})'$, 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 partition $\boldsymbol{w}_t$ in K_1- and K_2-dimensional subvectors $\boldsymbol{w}_{1t} = (w_{1t}, \dots, w_{K_1 t})'$ and $\boldsymbol{w}_{2t} = (w_{K_1+1,t}, \dots, w_{Kt})'$ 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 $(K \times K_2)$, where $K_2 = K - K_1$. 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 $h = 0, 1, \dots,$ 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 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 $\boldsymbol{w}_{1t}$ 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 $\boldsymbol{w}_{1t}$. 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.

Such additional information can also 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, Lütkepohl and Maciejowska (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, Marcellino and Tornese (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 time-invariant. 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

$$\widehat{D}(m) = \frac{1}{\tau_m T} \sum_{t \in \mathcal{T}_m} \hat{u}_t z_t', \tag{7}$$

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,

$$\sqrt{T}\text{vec}\left(\widehat{D}(m) - D(m)\right) \xrightarrow{d} \mathcal{N}(0, \tau_m^{-1}\Sigma_D(m)), \tag{8}$$

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

$$\mathbb{H}_0 : B_1(m) = B_1(k) \quad \text{versus} \quad \mathbb{H}_1 : B_1(m) \neq B_1(k) \tag{9}$$

for some $m, k \in \{1, \ldots, M\}$, $m \neq k$. The challenge is to derive a test for the pair of hypotheses in (9) 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 $D(m)$ and C_m. In other words, we don't need to assume a time-invariant covariance between proxies and shocks.

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

$$\begin{aligned} &\begin{bmatrix} I_{K_1} \\ B_{12}(m)B_{11}(m)^{-1} \end{bmatrix} = B_1(m)B_{11}(m)^{-1} \\ &= B_1(m)C_m Q C_m'(C_m Q C_m')^{-1}B_{11}(m)^{-1} \\ &= B_1(m)C_m Q C_m' B_{11}(m)'(B_{11}(m)C_m Q C_m' B_{11}(m)')^{-1} \\ &= D(m)QD_1(m)'[D_1(m)QD_1(m)']^{-1} \end{aligned} \tag{10}$$

for any positive definite $(N \times N)$ matrix Q. 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 (10) exist. Note that the left-hand side of expression (10) does not involve elements of C_m but just elements of $B_1(m)$, while the right-hand side of (10) consists of quantities that can be estimated consistently. Hence, we can also estimate the left-hand side consistently. Using this result, we test

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

instead of the pair of hypotheses in (9). If the $\mathbb{H}_1$ in (11) 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 remaining silent about possible time-variation in the covariance of shocks and proxies, C_m.

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 (11) is only a necessary condition for $\mathbb{H}_0$ in (9) to hold. For example, if a change in volatility from regime m to regime k changes $B_1(m)$ to $B_1(k) = cB_1(m)$, that change would cancel in (11). 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 (11) 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 (11) is that we can estimate

$$B_{12}(m)B_{11}(m)^{-1} = D_2(m)QD_1(m)'[D_1(m)QD_1(m)']^{-1}$$

consistently as

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

where we choose

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

in the absence of a more plausible alternative. Given the asymptotic normality of $\widehat{D}(m)$ in (8), Slutsky's theorem implies consistency and asymptotic normality of the estimator $\widehat{B_{12}(m)B_{11}}(m)^{-1}$, i.e.,

$$\sqrt{T}\text{vec}\left(\widehat{B_{12}(m)B_{11}}(m)^{-1} - B_{12}(m)B_{11}(m)^{-1}\right) \xrightarrow{d} \mathcal{N}\left(0, V(m)\right), \tag{13}$$

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 (9) can be tested using the test statistic

$$\eta(m,k) = T\left(\hat{\beta}(m) - \hat{\beta}(k)\right)'\left(\widehat{V}(m) + \widehat{V}(k)\right)^{-1}\left(\hat{\beta}(m) - \hat{\beta}(k)\right) \xrightarrow{d} \chi^2(K_1(K-K_1)). \tag{14}$$

Thus, we can use this statistic for testing the pair of hypotheses (11). In the test statistic in expression (14), the estimators of the covariance matrices may be obtained as

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

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

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 (2021) for testing for time-varying impact effects of the shock identified by the proxy. They mention that an extension of their test to the case of more than one proxy identifying multiple shocks would require additional information to uniquely identify the shocks individually. It is therefore worth emphasizing that our test does not require separately identified individual shocks. It does, however, require that 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 (2021).

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 regime 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, the 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 with just a few sample points.

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.

4.1 DGP1

4.1.1 Setup

The first DGP (DGP1) is based on DGP1 of Lütkepohl and Schlaak (2021). 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 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 (2021), DGP1 follows a three-variate VAR(1). We employ the same parameter values as Lütkepohl and Schlaak (2021) 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

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

As Φ 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 (2021). 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 of Lütkepohl and Schlaak (2021).

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 $\kappa = 3$, they are weaker and for $\kappa = 0.2346$, they are a bit stronger. The precise correlations between the structural shocks and the proxies are provided in Table 1, 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.[2]

Table 1: Correlations of z_t and $\boldsymbol{w}_{1t}$ for DGP1

	κ	z_t	$t \in \mathcal{T}_1$		$t \in \mathcal{T}_2$		$t \in \mathcal{T}_3$	
			w_{1t}	w_{2t}	w_{1t}	w_{2t}	w_{1t}	w_{2t}
$\rho = 0$	0.2346	z_{1t}	0.900	0.000	0.972	0.000	0.900	0.000
		z_{2t}	0.000	0.900	0.000	0.987	0.000	0.972
	1	z_{1t}	0.707	0.000	0.894	0.000	0.707	0.000
		z_{2t}	0.000	0.707	0.000	0.949	0.000	0.894
	3	z_{1t}	0.500	0.000	0.756	0.000	0.500	0.000
		z_{2t}	0.000	0.500	0.000	0.866	0.000	0.756
			w_{1t}	w_{2t}	w_{1t}	w_{2t}	w_{1t}	w_{2t}
$\rho = 0.5$	0.2346	z_{1t}	0.900	0.000	0.972	0.000	0.900	0.000
		z_{2t}	0.410	0.821	0.313	0.938	0.236	0.944
	1	z_{1t}	0.707	0.000	0.894	0.000	0.707	0.000
		z_{2t}	0.333	0.667	0.302	0.905	0.218	0.873
	3	z_{1t}	0.500	0.000	0.756	0.000	0.500	0.000
		z_{2t}	0.243	0.485	0.277	0.832	0.186	0.743

Following Lütkepohl and Schlaak (2021), 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 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 lag orders $p = 1$ and 12 to explore the impact of having to deal with larger lag orders. The number of replications for each simulation design is 5000.

[2]In Table S.1 of the Online Supplement, Lütkepohl and Schlaak (2021) 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.

4.1.2 Results for DGP1

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

Sample size: The sample size has the expected effect. Increasing it generally moves the rejection frequencies closer to the 5% nominal level (see Figure 1) and it improves the power (see Figure 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 Figure 1 and Tables C.1 and C.2 in Appendix C.

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 Figure 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 Figure 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 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 Figure 2 (compare the upper panels to the panels in the lower half). On the other hand, the results in Figure 1 indicate that the rejection frequencies under $\mathbb{H}_0$ are not much affected by the lag order.

Proxy strength: In Figure 2, it can also be seen that weaker proxies (larger κ) lead to a reduction in the 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 Figure 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 Figures 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 consider $\rho = 0.9$ 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.

In summary, the results for DGP1 show that our test has rejection frequencies close to the nominal significance level when the null hypothesis holds, even for relatively small samples. The power of the test increases with the sample size, is reduced for increasing lag order and improves with the strength of the proxies. It depends also on how far apart the relevant parameters are from the parameter space under $\mathbb{H}_0$. The next DGP, in addition, will allow 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.

4.2 DGP2

4.2.1 Setup

The second DGP is informed by an empirical model from Lunsford (2015) which is used as benchmark study for our empirical application in Section 5. This model employs five variables at quarterly frequency from 1948Q2 - 2015Q2, giving $T = 269$ observations. Lunsford (2015) uses $N = 2$ proxies to separately identify $K_1 = 2$ structural shocks. We fit a VAR(1) process with a constant to the data and use the estimated A_1 matrix and intercept term as parameters for our DGP2. The precise parameter values are given in Appendix B. The largest eigenvalue of A_1 is 0.7444, implying a stable process with medium persistence.

We search for a single volatility break point by minimizing the criterion function

$$\psi(T_1) = T_1 \log\det\hat{\Sigma}_u(1) + (T - T_1)\log\det\hat{\Sigma}_u(2) \tag{15}$$

over $T_1 \in \{0.15T, \ldots, 0.85T\}$. Here

$$\hat{\Sigma}_u(m) = \frac{1}{T_m - T_{m-1}} \sum_{t\in\mathcal{T}_m} \hat{u}_t\hat{u}_t',$$

where $\hat{u}_t$ are the OLS residuals of the VAR(1) model. Thereby we find a volatility change point in 1982Q4. The state-dependent covariance matrices $\Sigma_u(1)$ and $\Sigma_u(2)$ corresponding to this change point are used as parameters for our DGP2 and are also provided in Appendix B.

The structural matrix of impact effects is chosen to be time-invariant under $\mathbb{H}_0$ and dependent on the volatility regime under $\mathbb{H}_1$. The actual matrices used are constructed as follows.

Under $\mathbb{H}_0$: We decompose $\Sigma_u(1)$ and $\Sigma_u(2)$ such that

$$\Sigma_u(1) = BB' \quad \text{and} \quad \Sigma_u(2) = B\Lambda_2 B',$$

where B is a (5×5) matrix and $\Lambda_2 = \text{diag}(\lambda_{1,2}, \ldots, \lambda_{5,2})$ is diagonal. We order the $\lambda_{k,2}$ and the columns of B such that $\lambda_{1,2}$ and $\lambda_{2,2}$ are clearly distinct giving

$\Lambda_2 = \text{diag}(0.57, 0.15, 0.18, 0.35, 0.39)$. Thereby the first two columns of B are uniquely identified. The precise values of the elements of B are also given in Appendix B. We choose $B_1(1) = B_1(2)$ as the first 2 columns of B.

The structural shocks are generated as $\boldsymbol{w}_t \sim \mathcal{N}(0, I_5)$ for $t \in \mathcal{T}_1$ and $\boldsymbol{w}_t \sim \mathcal{N}(0, \Lambda_2)$ for $t \in \mathcal{T}_2$ and then we generate the reduced form errors as $u_t = B\boldsymbol{w}_t$ for $t = 1, \ldots, T$. The volatility change point is placed in the middle of the sample.

To generate proxies similar to Lunsford (2015), we compute

$$D_m = \frac{1}{\tau_m T} \sum_{t \in \mathcal{T}_m} \hat{u}_t z_t'$$

and note that

$$C_m = (B_1' B_1)^{-1} B_1' D_m \quad (= \mathbb{E}(\boldsymbol{w}_{1t} z_t') \quad t \in \mathcal{T}_m)$$

is clearly nonzero for $m = 1, 2$, for our dataset (see again Appendix B for details). Hence, the proxies satisfy the relevance condition (3) and are, thus, suitable for our purposes.

As for DGP1, we generate the proxies as

$$z_t = \Phi(m)\boldsymbol{w}_{1t} + v_t$$

with $v_t \sim \mathcal{N}(0, \kappa\Sigma_v)$, $t = 1, \ldots, T$.

For $K_1 = 2$, we choose $\Phi(1)$, $\Phi(2)$, and $\Sigma_v(1) = \Sigma_v(2) = \Sigma_v$, such that, for $\kappa = 1$ the covariance matrix of the proxies,

$$\Sigma_z = \Phi(1)\Phi(1)' + \Sigma_v = \Phi(2)\Lambda_2\Phi(2)' + \Sigma_v$$

is very similar to the empirical covariance matrix of Lunsford (2015),

$$T^{-1} \sum_{t=1}^{T} z_t z_t' = \begin{bmatrix} 9.95 & 5.41 \\ 5.41 & 36.88 \end{bmatrix}.$$

Given our simulation results for DGP1, the matrices $\Phi(1)$ and $\Phi(2)$ are chosen to be diagonal matrices. Based on the results for DGP1, choosing them to be diagonal or not should not make much difference for the simulation results. In contrast, Σ_v is not a diagonal matrix. Precise values and precise details on how they are computed are provided in Appendix B.

For κ we use values 0.1, 0.5, and 1, corresponding to high, intermediate and low strength of the proxies. The resulting correlations between proxies and structural shocks for the different values of κ are shown in Table B.1 in Appendix B. They range from a low 0.368 for $\text{corr}(z_{2t}, w_{2t})$ and $\kappa = 1$ to a high $\text{corr}(z_{1t}, w_{1t}) = 0.897$ for $\kappa = 0.1$.

To investigate the role of the number of shocks being identified, we construct additional proxies, employing one proxy to identify one shock. To do so, we augment the bivariate $\Phi(m)$ and $\Sigma_v(m)$ matrices as specified in Appendix B. We use these settings to explore the small sample properties of our test for $K_1 = 3$ and $K_1 = 4$.

Under $\mathbb{H}_1$: We use a Cholesky decomposition of $\Sigma_u(1)$ for $B(1)$ and choose $B(2)$ as under $\mathbb{H}_0$. Apart from that, we use the same setup as under $\mathbb{H}_0$.

We generate samples of size $T = 150, 300, 600, 1200$ (in addition to all required pre-sample observations). The lag orders of VAR models fitted to the data are $p = 1$, and 12. The nominal significance level is $\alpha = 5\%$, $\kappa = 0.1, 0.5, 1$, and 5000 Monte Carlo replications are employed.

In some of our simulations, the true volatility change point T_1, is again assumed to be known, as for DGP1. We now also investigate the situation where the volatility change point is misspecified. More precisely, we use $T_1 = 0.4T$ as the change point instead of the correct $T_1 = 0.5T$ to investigate the implications of misspecifying this quantity. In addition, we also mimic the situation where the change point is preselected by a statistical criterion. We use the criterion function $\psi(T_1)$ in (15) for that purpose. In each simulation, we compute the criterion function for every $\frac{1}{30}T$-th sample point, e.g., for every fifth sample point for $T = 150$, and choose the change point T_1 for which $\psi(T_1)$ is minimized over this rough grid. Note that this procedure estimates the overall minimizing sample period only very roughly which is likely to be an additional handicap for the test and might work against the test having good properties. Once the volatility change point is chosen, we perform our test for time-varying impact effects conditionally on that change point. Note that we are explicitly not pretesting for the change point. Instead we use a statistical criterion for the determination of the change point which does not involve our test statistic.[3] Our procedure is meant to mimic an approach occasionally encountered in practice and we are interested to see whether it leads to biased test results.

4.2.2 Results for DGP2

Results for DGP2 are depicted in Figures 3 - 5 and precise numbers are given in Tables C.3 - C.5 in Appendix C. In Figure 3 and Table C.3 the correct volatility change point is assumed to be known, while it is misspecified in Figure 4 and Table C.4 and it is estimated in Figure 5 and Table C.5. These results underscore the findings based on DGP1 in that larger samples and stronger proxies (smaller κ) improve power, while larger lag orders undermine power. Thus, these features carry over to the higher-dimensional DGP2. In addition, the following observations emerge.

Number of variables: A larger number of variables, i.e., a higher-dimensional process tends to make the test more conservative in small samples. The first six panels of Figures 3 - 5 show results for a true null hypothesis. For small samples the rejection frequencies are sometimes considerably smaller than 5%. In particular, if the proxies are relatively weak ($\kappa = 1$), the rejection frequencies are even below 1% for $T = 150$ and $p = 12$. Admittedly, this is a rather difficult case, where many parameters have to be estimated from a very small sample. In such situations, the test tends to be conservative.

Number of proxies: In Figures 3 - 5, the number of proxies increases from 2 to 4, respectively, when we move from left to right. Looking at the figures it

[3] Alternatively, one could choose the break point which maximizes our test statistic. In a related context, Andrews (1993) shows that such an approach would change the asymptotic distribution of the test statistic. Therefore we explicitly do not use that strategy.

is apparent that the rejection frequencies under $\mathbb{H}_0$ are not much affected by this increase while the power tends to decline when there are more proxies. Increasing the number of proxies for a VAR model of given dimension affects the power in two ways. First, it reduces the number of restrictions that are to be tested, thereby making the test more focussed and potentially more powerful. A second effect is, however, that it increases the dimension of the parameter space where the test does not have power and, hence, the true parameter values may get closer to the parameter space without power, thereby reducing the power. The latter effect dominates in our simulations.

Misspecifying the change point: Comparing Figure 4 to Figure 3, it can be seen that misspecifying the volatility change point does not affect the results very much. The rejection frequencies under $\mathbb{H}_0$ as well as under $\mathbb{H}_1$ (power) in Figure 4 are similar to the case of a correctly specified volatility change point in Figure 3. This finding is remarkable because, in Figure 4, we specify the change point at 40% of the sample length while the true change point is in the middle of the sample. In other words, the misspecification is substantial.

Change point preselection: Similarly, comparing Figure 5 to Figure 3 shows that preselecting the change point by our rough statistical procedure does not affect the results very much. Again the rejection frequencies are similar to the case of a correctly specified volatility change point, although we are not even trying to find the optimal change point according to the criterion $\psi(T_1)$ in (15) across the complete sample, but search over a very rough grid only. Thus, at least for DGP2, the preselection of the change point by the statistical criterion does not bias the test results.

Overall our results based on DGP2 confirm the main conclusions from the simulations of DGP1 regarding the impact of the sample size, the 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.

4.3 Summary of Simulation Results

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 relatively weak proxies, 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, see Figure 6), that is, y_t is 5-dimensional. There are $N = 2$ proxies to identify $K_1 = 2$ TPF shocks and the VAR model has $p = 4$ lags[4] and a constant.[5]

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.

Figure 6 plots the endogenous variables and the two proxies. Figure 7 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.[6] 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.

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 neighbouring quarters of 1983Q4 and show the results in Table 2. 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.

We also use the likelihood based criterion $\psi(T_1)$ in (15) for choosing the volatility change point. It has its lowest values in the period 1982-1983 and is minimized for a change in 1982Q4. That change point results in a p-value of our test statistic of 0.08 which is again below 10%.

[4] 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.

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

[6] Using monthly data in monetary studies, Bernanke and Mihov (1998) and Christiano, Eichenbaum and Evans (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).

Table 2: 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

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 2 provide substantial evidence for a change in the impact effects of the TFP shocks 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. 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 instance, one may be interested in a unit shock of some sort which would only require rescaling the impact effects. Hence, we use the estimator $\widehat{D}(m)$ from equation (7) for the impact effects of the two structural shocks in our impulse response analysis. If the true underlying C_m matrices are not diagonal, this may lead to distortions. However, it may still offer insights whether accounting for a possible shift in the impulse responses at the time of the volatility change makes a difference because, if there is a shift in the actual structural impulse responses, the distorted impulse responses are likely to also have a shift.[7]

[7] Alternatively, one could standardize the impact response of one of the variables to 1 (e.g., Paul (2020)). 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. As we are primarily interested in changes in the shape of the dynamic responses of the variables, we prefer our approach which is also more in line with Lunsford (2015).

The impulse responses together with 90% confidence intervals are displayed in Figure 8. The confidence intervals are generated by a residual-based moving-block bootstrap (MBB) as proposed by Brüggemann, Jentsch and Trenkler (2016) and Jentsch and Lunsford (2019, 2021). 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 (2020).[8]

The impulse responses in Figure 8 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.

Thus, the impulse responses in Figure 8 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 obtained under the assumption of time-invariant impulse responses (shown as a solid black line in Figure 8) lie in between the impulse responses 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.[9] 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 as well as a decline in inflation, if the latter variable is affected at all. 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 Figure 8 the investment TFP shock is clearly a negative shock to all quantities and inflation in the pre-GM period, while the response of some variables is 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

[8]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.

[9]Note that the solid black line in Figure 8 represents the impulse responses obtained under time-invariance when both proxies are used jointly as explained earlier. They are very similar but not identical to the impulse responses in Lunsford (2015) who computes them one at a time.

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 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 and, hence, it can also be applied if such information is controversial or absent.

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. Even substantially misspecified 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 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 (2021) 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

$$\frac{\partial \beta}{\partial \text{vec}(D)'} = \frac{\partial \text{vec}\{D_2 Q D_1' [D_1 Q D_1']^{-1}\}}{\partial [\text{vec}(D_1)'\text{vec}(D_2)']} \times \frac{\partial \begin{bmatrix} \text{vec}(D_1) \\ \text{vec}(D_2) \end{bmatrix}}{\partial \text{vec}(D_1')'\text{vec}(D_2')'} \times \frac{\partial \begin{bmatrix} \text{vec}(D_1') \\ \text{vec}(D_2') \end{bmatrix}}{\partial \text{vec}(D)'}$$

$$(K_1(K-K_1) \times KN)$$

(see Sec. 10.7, Rule (2)). Moreover,

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

where

$$\begin{aligned}\frac{\partial \text{vec}\{D_2 Q D_1' [D_1 Q D_1']^{-1}\}}{\partial \text{vec}(D_1)'} =& ([D_1 Q D_1']^{-1} \otimes I_{K-K_1}) \frac{\partial \text{vec}(D_2 Q D_1')}{\partial \text{vec}(D_1)'} \\ &- ([D_1 Q D_1']^{-1} \otimes D_2 Q D_1' [D_1 Q D_1']^{-1}) \frac{\partial \text{vec}(D_1 Q D_1')}{\partial \text{vec}(D_1)'}\end{aligned}$$

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

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

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

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

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

$$\frac{\partial \text{vec}\{D_2 Q D_1' [D_1 Q D_1']^{-1}\}}{\partial \text{vec}(D_2)'} = [D_1 Q D_1']^{-1} D_1 Q \otimes I_{K_2} \quad (K_1(K-K_1) \times (K-K_1)N)$$

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

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

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

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

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

B Additional Details for DGP2

DGP2 is a 5-dimensional VAR(1) process, $y_t = \nu + A_1 y_{t-1} + u_t$, with slope parameter matrix

$$A_1 = \begin{bmatrix} 0.09 & 0.17 & -0.14 & 0.65 & 0.00 \\ 0.11 & 0.61 & -0.05 & 0.35 & -0.02 \\ 0.05 & 0.08 & 0.79 & -0.04 & -0.02 \\ 0.19 & 0.04 & -0.12 & 0.20 & -0.04 \\ 0.75 & 1.03 & -1.32 & 1.12 & -0.28 \end{bmatrix}$$

and constant term

$$\nu = (1.07, -0.52, 0.55, 2.42, 4.24)'.$$

A volatility change is assumed in the middle of the sample. The two reduced form covariance matrices are estimated from the two volatility regimes found for the Lunsford (2015) data as follows:

$$\Sigma_u(1) = \begin{bmatrix} 18.26 & 7.37 & -0.61 & 4.04 & 38.59 \\ 7.37 & 5.55 & 0.78 & 1.95 & 19.35 \\ -0.61 & 0.78 & 3.31 & 0.14 & 3.65 \\ 4.04 & 1.95 & 0.14 & 5.28 & 10.72 \\ 38.59 & 19.35 & 3.65 & 10.72 & 219.91 \end{bmatrix}$$

and

$$\Sigma_u(2) = \begin{bmatrix} 4.69 & 1.02 & 0.02 & 1.52 & 9.62 \\ 1.02 & 0.85 & 0.05 & 0.33 & 3.42 \\ 0.02 & 0.05 & 0.65 & 0.05 & -0.35 \\ 1.52 & 0.33 & 0.05 & 2.17 & 3.06 \\ 9.62 & 3.42 & -0.35 & 3.06 & 69.67 \end{bmatrix}.$$

To determine the matrix of impact effects under $\mathbb{H}_0$, we have to find a time-invariant (5×5) matrix B such that

$$\Sigma_u(1) = BB' \quad \text{and} \quad \Sigma_u(2) = B\Lambda_2 B',$$

where $\Lambda_2 = \mathrm{diag}(\lambda_{1,2}, \ldots, \lambda_{5,2})$ is diagonal. We have performed the required decomposition of $\Sigma_u(1)$ and $\Sigma_u(2)$ with an algorithm from Golub and Van Loan (1989, Algorithm 8.7.1). We first computed a Cholesky decomposition of $\Sigma_u(1)$, i.e., we computed $G = \mathrm{chol}(\Sigma_u(1))$, such that $\Sigma_u(1) = GG'$ and define $H = G^{-1}\Sigma_u(2)G'^{-1}$. Then we computed the spectral decomposition $H = U\Lambda U'$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$ and the λ_k are the eigenvalues of H while U is the orthogonal matrix of corresponding eigenvectors. Finally, $B = GU$ and $\Lambda_2 = \Lambda$.

We have ordered the $\lambda_{k,2}$ and the columns of B such that $\lambda_{1,2}$ and $\lambda_{2,2}$ are clearly distinct. Thereby we obtained $\Lambda_2 = \text{diag}(0.57, 0.15, 0.18, 0.35, 0.39)$ and the corresponding matrix

$$B = \begin{bmatrix} 1.76 & -2.99 & 1.87 & 1.25 & -1.08 \\ -0.15 & -2.30 & 0.35 & -0.17 & -0.27 \\ 0.34 & -0.67 & -1.60 & 0.17 & 0.39 \\ 1.29 & -0.69 & 0.05 & -1.32 & -1.18 \\ -0.46 & -7.64 & -1.41 & 5.55 & -11.34 \end{bmatrix}.$$

We choose $B_1(1) = B_1(2)$ as the first 2 columns of this B matrix.

The structural errors $\boldsymbol{w}_t$ are generated as standard normal, $\boldsymbol{w}_t \sim \mathcal{N}(0, I_5)$, for $t \in \mathcal{T}_1$, and $\boldsymbol{w}_t \sim \mathcal{N}(0, \Lambda_2)$ for $t \in \mathcal{T}_2$. Using them we get the reduced-form errors as $u_t = B\boldsymbol{w}_t$ for $t = 1, \dots, T$.

The proxies are generated as

$$z_t = \Phi(m)\boldsymbol{w}_{1t} + v_t,$$

with $v_t \sim \mathcal{N}(0, \kappa\Sigma_v)$, $t = 1, \dots, T$. Clearly, larger κ's imply weaker proxies while smaller κ's lead to stronger proxies (i.e., proxies more strongly correlated with the structural shocks).

To obtain proxies similar to Lunsford (2015) for $\kappa = 1$, we have computed

$$D_m = \frac{1}{\tau_m T} \sum_{t \in \mathcal{T}_m} \hat{u}_t z_t'$$

and note that

$$C_m = (B_1' B_1)^{-1} B_1' D_m \quad (= \mathbb{E}(\boldsymbol{w}_{1t} z_t'), \quad t \in \mathcal{T}_m).$$

For our dataset

$$C_1 = \begin{bmatrix} 0.39 & -0.67 \\ -1.31 & 2.53 \end{bmatrix} \quad \text{and} \quad C_2 = \begin{bmatrix} 0.50 & 0.12 \\ -0.73 & 0.51 \end{bmatrix}$$

which are both clearly different from zero and, hence, the generated z_t qualify as proxies. We then determine $\Phi(1)$, $\Phi(2)$, and $\Sigma_v(1) = \Sigma_v(2) = \Sigma_v$ such that the covariance matrix of the proxies

$$\Sigma_z = \Phi(1)\Phi(1)' + \Sigma_v = \Phi(2)\Lambda_2\Phi(2)' + \Sigma_v$$

is close to the empirical covariance matrix

$$T^{-1} \sum_{t=1}^{T} z_t z_t' = \begin{bmatrix} 9.95 & 5.41 \\ 5.41 & 36.88 \end{bmatrix}.$$

More precisely, $\Phi(m)$ and Σ_ν are obtained as follows:

- Compute $\hat{w}_{it} = D_{\cdot,i}\Sigma^{-1}\hat{u}_t/(D_{\cdot,i}'\Sigma^{-1}D_{\cdot,i})$, where $D_{i,j}$ is the i,j-th element of $D = T^{-1}\sum_{t=1}^{T} \hat{u}_t z_t'$ and $D_{\cdot,j}$ is its j-th column. Let $\hat{\boldsymbol{w}}_{1t}$ consist of the first K_1 elements of $\hat{\boldsymbol{w}}_t = [\hat{w}_{1,t}, \dots, \hat{w}_{K,t}]'$.

- For the first volatility regime, normalise the variance of $\hat{w}_{it}$ to be one and compute $\Phi(1) = \sum_{t=1}^{T} z_t \hat{\boldsymbol{w}}_{1t}' \big(\sum_{t=1}^{T} \hat{\boldsymbol{w}}_{1t} \hat{\boldsymbol{w}}_{1t}' \big)^{-1}$.
- For the second volatility regime, normalise the variance of $\hat{w}_{it}$ to be $\lambda_{i,2}$ and compute $\Phi(2) = \sum_{t=1}^{T} z_t \hat{\boldsymbol{w}}_{1t}' \big(\sum_{t=1}^{T} \hat{\boldsymbol{w}}_{1t} \hat{\boldsymbol{w}}_{1t}' \big)^{-1}$.
- Compute $\Sigma_v = T^{-1} \sum_{t=1}^{T} \hat{v}_t \hat{v}_t'$, where $\hat{v}_t = z_t - \Phi(m) \hat{\boldsymbol{w}}_t$.

The resulting $\Phi(m)$ and Σ_v matrices are

$$\Phi(1) = \begin{bmatrix} 1.70 & 0 \\ 0 & 2.24 \end{bmatrix}, \quad \Phi(2) = \begin{bmatrix} 2.26 & 0 \\ 0 & 5.83 \end{bmatrix},$$

and

$$\Sigma_v = \begin{bmatrix} 7.05 & 5.35 \\ 5.35 & 31.89 \end{bmatrix}.$$

The corresponding covariance matrix

$$\Sigma_z = \Phi(1)\Phi(1)' + \Sigma_v = \Phi(2)\Lambda_2\Phi(2)' + \Sigma_v = \begin{bmatrix} 9.95 & 5.35 \\ 5.35 & 36.90 \end{bmatrix},$$

which is very similar to the covariance matrices of Lunsford's proxies given in Section 4.2.1. The resulting correlations between proxies and structural shocks for variations in κ are shown in Table B.1.

Table B.1: Correlations of z_t and $\boldsymbol{w}_{1t}$ for DGP2

		w_{1t}	w_{2t}
$\kappa = 0.1$	z_{1t}	0.897	0.000
	z_{2t}	0.000	0.782
		w_{1t}	w_{2t}
$\kappa = 0.5$	z_{1t}	0.672	0.000
	z_{2t}	0.000	0.489
		w_{1t}	w_{2t}
$\kappa = 1$	z_{1t}	0.540	0.000
	z_{2t}	0.000	0.368

To investigate the impact of the number of shocks being identified on the small sample properties of our test, we generate additional proxies, employing one proxy to identify one shock. To this end, we augment $\Phi(m)$ and $\Sigma_v(m)$ as follows:

$K_1 = 3$**:**

$$\Phi(1) = \begin{bmatrix} 1.70 & 0 & 0 \\ 0 & 2.24 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Phi(2) = \begin{bmatrix} 2.26 & 0 & 0 \\ 0 & 5.83 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\Sigma_v(1) = \Sigma_v(2) = \begin{bmatrix} 7.05 & 5.35 & 0 \\ 5.35 & 31.89 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$K_1 = 4$**:**

$$\Phi(1) = \begin{bmatrix} 1.70 & 0 & 0 & 0 \\ 0 & 2.24 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Phi(2) = \begin{bmatrix} 2.26 & 0 & 0 & 0 \\ 0 & 5.83 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\Sigma_v(1) = \Sigma_v(2) = \begin{bmatrix} 7.05 & 5.35 & 0 & 0 \\ 5.35 & 31.89 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Under $\mathbb{H}_1$ we use a Cholesky decomposition of $\Sigma_u(1)$ for $B(1)$ and maintain the $B(2)$ matrix used under $\mathbb{H}_0$. Apart from that, the same setup as under $\mathbb{H}_0$ is used.

References

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point, *Econometrica* **61**: 821–856.

Arias, J. E., Rubio-Ramírez, J. F. and Waggoner, D. F. (2021). Inference in Bayesian Proxy-SVARs, *Journal of Econometrics* **225**(1): 88–106.

Bacchiocchi, E., Castelnuovo, E. and Fanelli, L. (2018). Gimme a break! Identification and estimation of the macroeconomic effects of monetary policy shocks in the United States, *Macroeconomic Dynamics* **22**: 1613–1651.

Bacchiocchi, E. and Fanelli, L. (2015). Identification in structural vector autoregressive models with structural changes, with an application to US monetary policy, *Oxford Bulletin of Economics and Statistics* **77**: 761–779.

Bernanke, B. S. and Mihov, I. (1998). Measuring monetary policy, *Quarterly Journal of Economics* **113**: 869–902.

Braun, R. and Brüggemann, R. (2020). Identification of SVAR Models by Combining Sign Restrictions With External Instruments, *Technical report*, Department of Economics, University of Konstanz.

Brüggemann, R., Jentsch, C. and Trenkler, C. (2016). Inference in VARs with conditional heteroskedasticity of unknown form, *Journal of Econometrics* **191**: 69–85.

Bruns, M. and Lütkepohl, H. (2020). An alternative bootstrap for proxy vector autoregressions, *Discussion Paper 1913*, DIW, Berlin.

Carriero, A., Marcellino, M. and Tornese, T. (2021). Blended identification in structural VARs, *Technical report*, Queen Mary University of London.

Christiano, L. J., Eichenbaum, M. and Evans, C. (1999). Monetary policy shocks: What have we learned and to what end?, *in* J. B. Taylor and M. Woodford (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. and Gambetti, L. (2009). On the sources of the great moderation, *American Economic Journal: Macroeconomics* **1**: 26–57.

Golub, G. H. and Van Loan, C. F. (1989). *Matrix Computations*, 2nd edn, Johns Hopkins University Press, Baltimore.

Herwartz, H. and Lütkepohl, H. (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks, *Journal of Econometrics* **183**: 104–116.

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.

Jentsch, C. and Lunsford, K. G. (2021). Asymptotically valid bootstrap inference for proxy SVARs, *Journal of Business & Economic Statistics* .

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

Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching, *Journal of Economic Dynamics and 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. and Netšunajev, A. (2017). Structural vector autoregressions with heteroskedasticy: A review of different volatility models, *Econometrics and Statistics* **1**: 2–18.

Lütkepohl, H. and Schlaak, T. (2021). Heteroscedastic proxy vector autoregressions, *Journal of Business and Economic Statistics* (forthcoming).

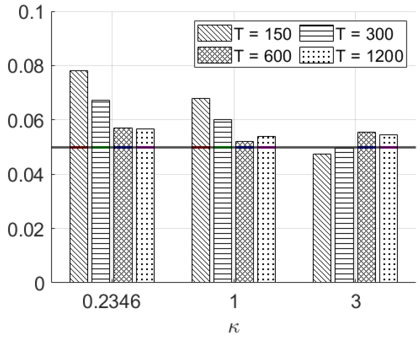
McConnell, M. M. and Perez-Quiros, G. (2000). Output fluctuations in the United States: What has changed since the early 1980's?, *American Economic Review* **90**: 1464–1476.

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.

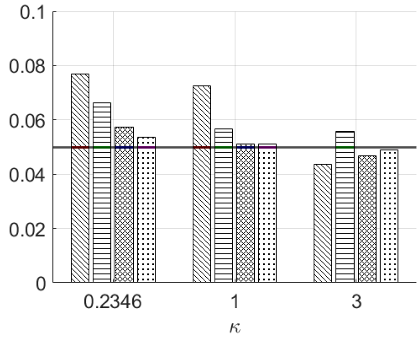
Paul, P. (2020). The time-varying effect of monetary policy on asset prices, *Review of Economics and Statistics* **102**: 690–704.

Piffer, M. and Podstawski, M. (2017). Identifying uncertainty shocks using the price of gold, *The Economic Journal* **128**(616): 3266–3284.

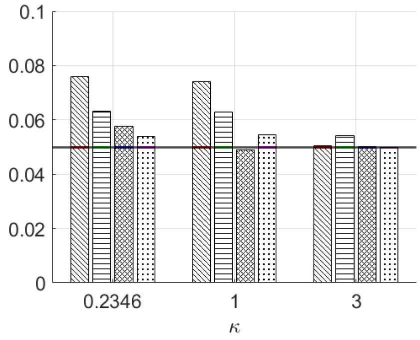
Stock, J. H. and Watson, M. W. (2003). Has the business cycle changed and why?, *NBER Macroeconomics Annual 2002, Volume 17*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 159–230.



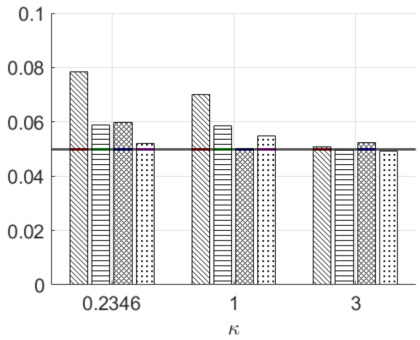
(a) $\mathbb{H}_0 : \beta(1) = \beta(2)$,
$\rho = 0$, $p = 1$



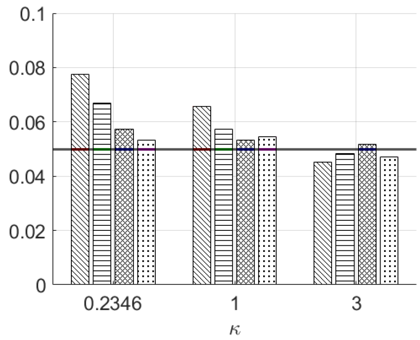
(b) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0$, $p = 1$



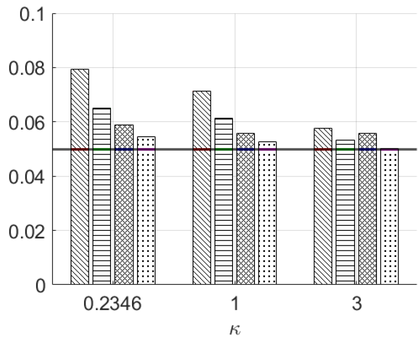
(c) $\mathbb{H}_0 : \beta(2) = \beta(3)$,
$\rho = 0$, $p = 1$



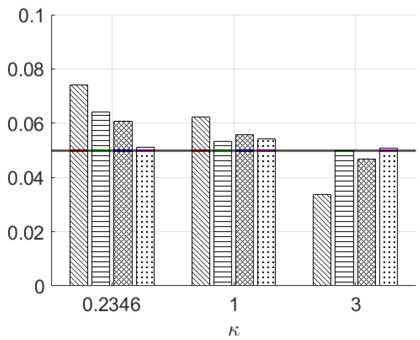
(d) $\mathbb{H}_0 : \beta(1) = \beta(2)$,
$\rho = 0.5$, $p = 1$



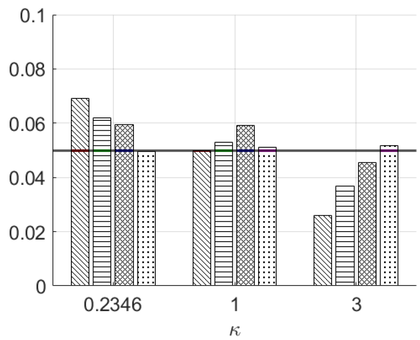
(e) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0.5$, $p = 1$



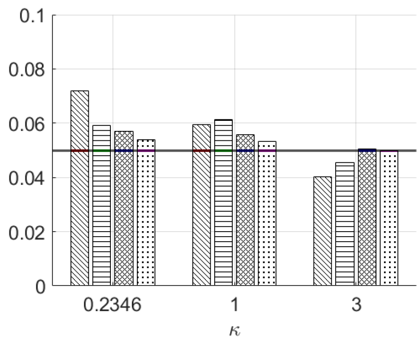
(f) $\mathbb{H}_0 : \beta(2) = \beta(3)$,
$\rho = 0.5$, $p = 1$



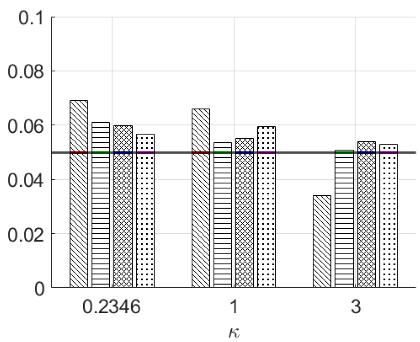
(g) $\mathbb{H}_0 : \beta(1) = \beta(2)$,
$\rho = 0$, $p = 12$



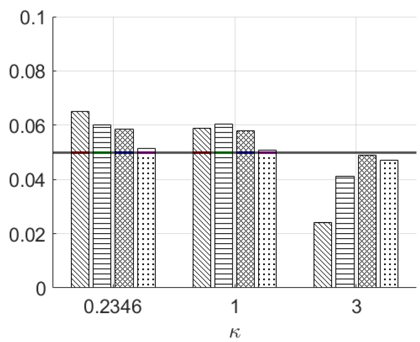
(h) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0$, $p = 12$



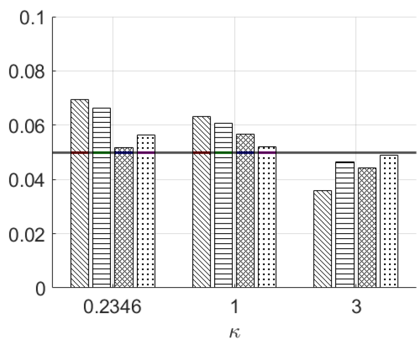
(i) $\mathbb{H}_0 : \beta(2) = \beta(3)$,
$\rho = 0$, $p = 12$



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

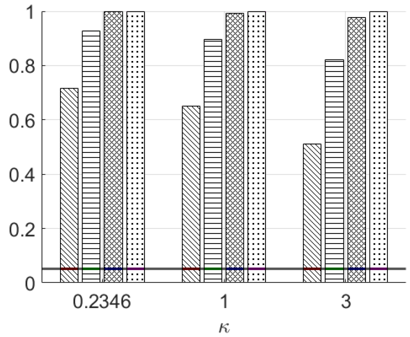


(k) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0.5$, $p = 12$

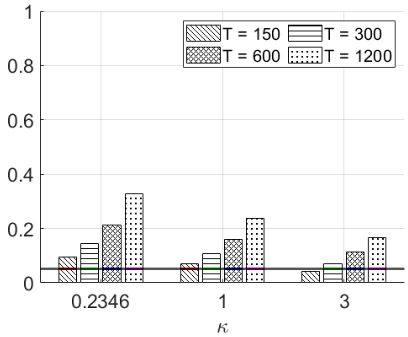


(l) $\mathbb{H}_0 : \beta(2) = \beta(3)$,
$\rho = 0.5$, $p = 12$

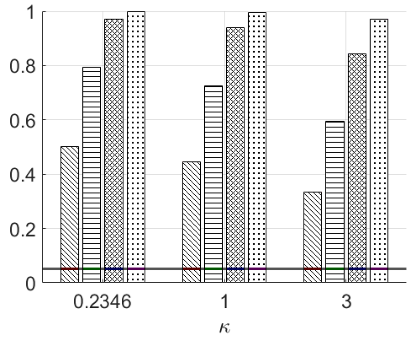
Figure 1: Relative rejection frequencies for DGP1 under $\mathbb{H}_0$. Nominal level 5%.



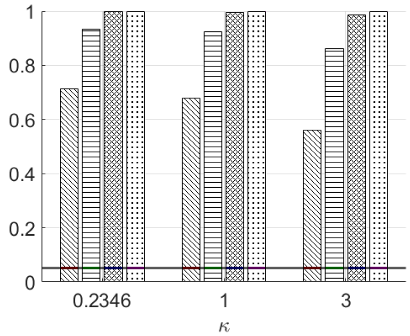
(a) $\mathbb{H}_0 : \beta(1) = \beta(2)$,
$\rho = 0$, $p = 1$



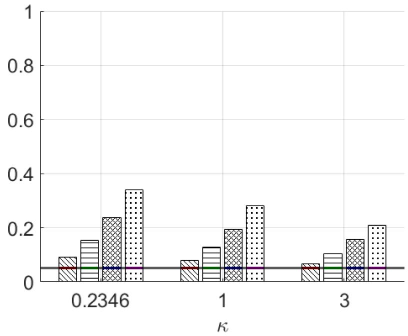
(b) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0$, $p = 1$



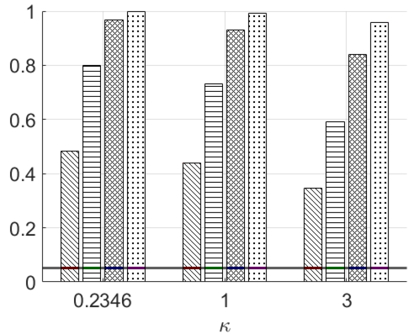
(c) $\mathbb{H}_0 : \beta(2) = \beta(3)$,
$\rho = 0$, $p = 1$



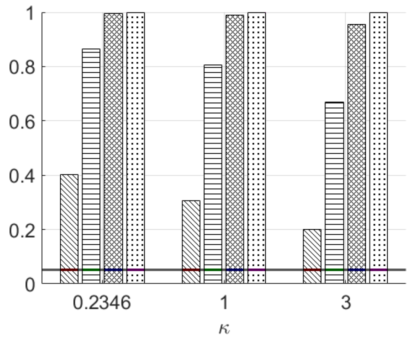
(d) $\mathbb{H}_0 : \beta(1) = \beta(2)$,
$\rho = 0.5$, $p = 1$



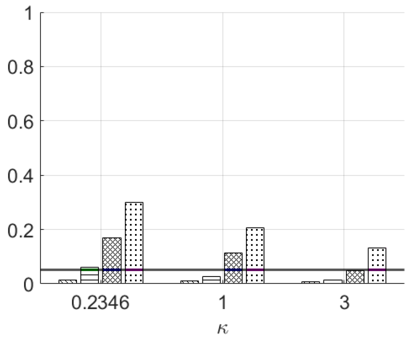
(e) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0.5$, $p = 1$



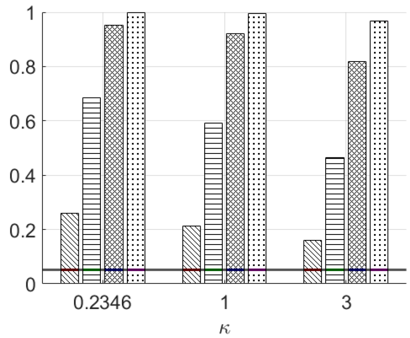
(f) $\mathbb{H}_0 : \beta(2) = \beta(3)$,
$\rho = 0.5$, $p = 1$



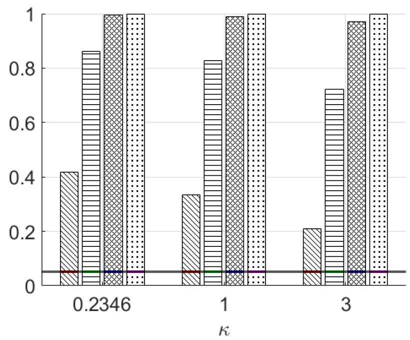
(g) $\mathbb{H}_0 : \beta(1) = \beta(2)$,
$\rho = 0$, $p = 12$



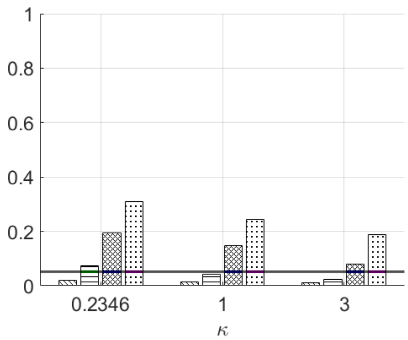
(h) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0$, $p = 12$



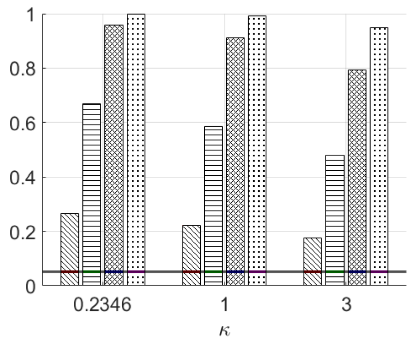
(i) $\mathbb{H}_0 : \beta(2) = \beta(3)$,
$\rho = 0$, $p = 12$



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

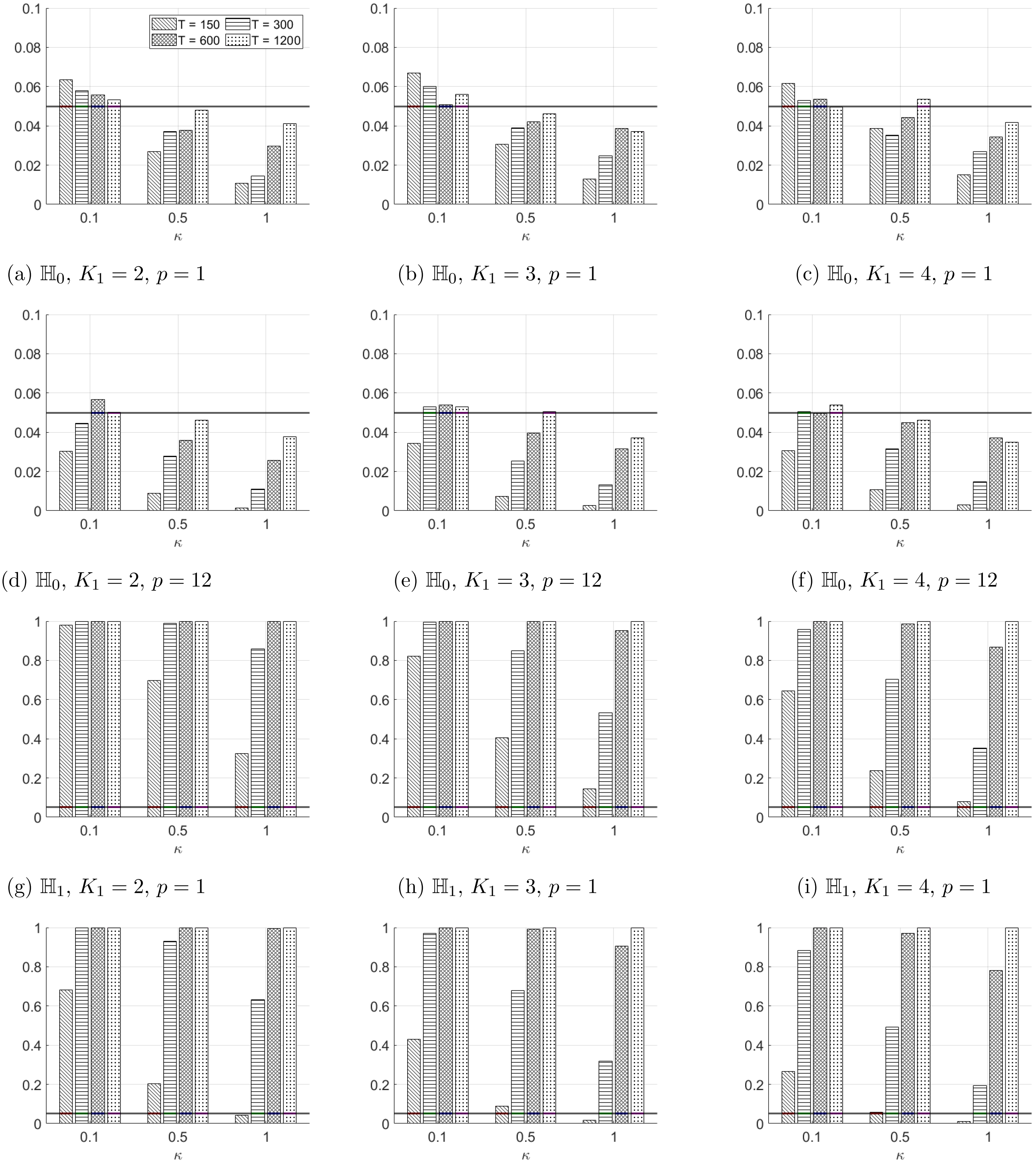


(k) $\mathbb{H}_0 : \beta(1) = \beta(3)$,
$\rho = 0.5$, $p = 12$



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

Figure 2: Relative rejection frequencies for DGP1 under $\mathbb{H}_1$. Nominal level 5%.



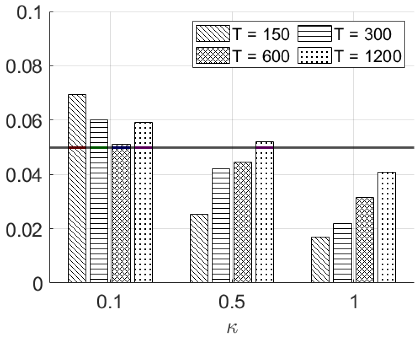
(a) $\mathbb{H}_0$, $K_1 = 2$, $p = 1$ (b) $\mathbb{H}_0$, $K_1 = 3$, $p = 1$ (c) $\mathbb{H}_0$, $K_1 = 4$, $p = 1$

(d) $\mathbb{H}_0$, $K_1 = 2$, $p = 12$ (e) $\mathbb{H}_0$, $K_1 = 3$, $p = 12$ (f) $\mathbb{H}_0$, $K_1 = 4$, $p = 12$

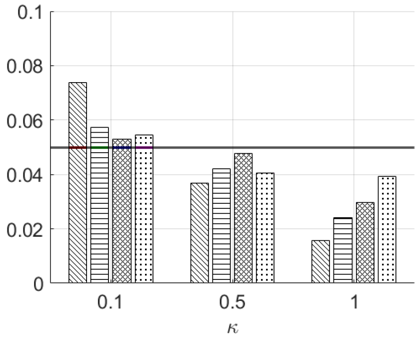
(g) $\mathbb{H}_1$, $K_1 = 2$, $p = 1$ (h) $\mathbb{H}_1$, $K_1 = 3$, $p = 1$ (i) $\mathbb{H}_1$, $K_1 = 4$, $p = 1$

(j) $\mathbb{H}_1$, $K_1 = 2$, $p = 12$ (k) $\mathbb{H}_1$, $K_1 = 3$, $p = 12$ (l) $\mathbb{H}_1$, $K_1 = 4$, $p = 12$

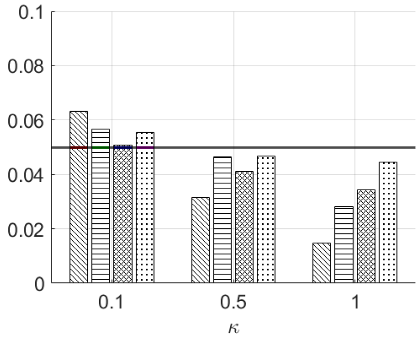
Figure 3: Relative rejection frequencies for DGP2. Nominal significance level 5%. True and assumed volatility change point at $T_1 = \frac{1}{2}T$.



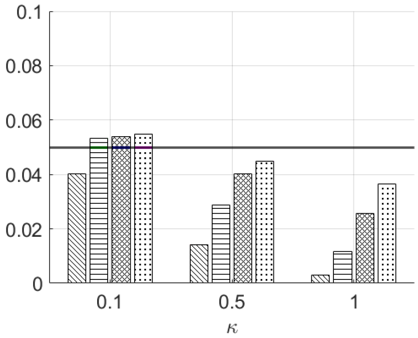
(a) $\mathbb{H}_0$, $K_1 = 2$, $p = 1$



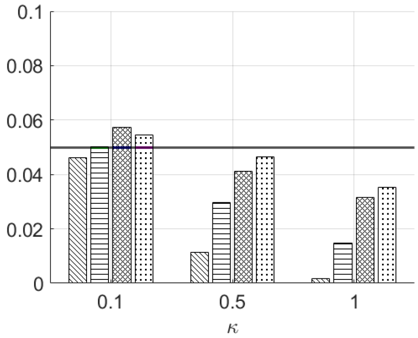
(b) $\mathbb{H}_0$, $K_1 = 3$, $p = 1$



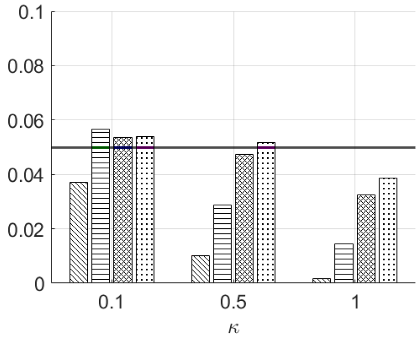
(c) $\mathbb{H}_0$, $K_1 = 4$, $p = 1$



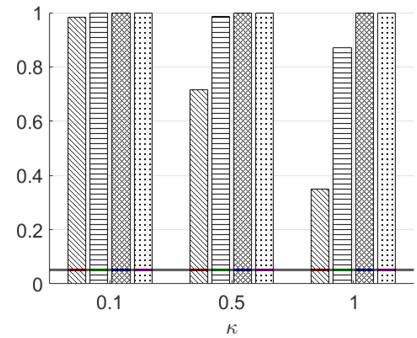
(d) $\mathbb{H}_0$, $K_1 = 2$, $p = 12$



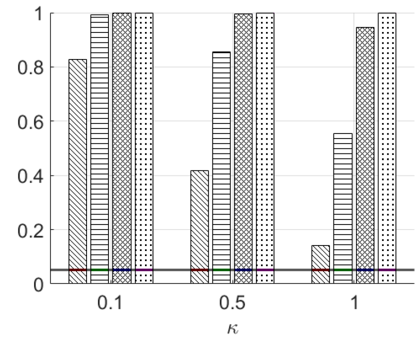
(e) $\mathbb{H}_0$, $K_1 = 3$, $p = 12$



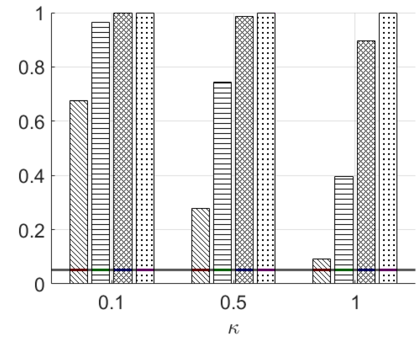
(f) $\mathbb{H}_0$, $K_1 = 4$, $p = 12$



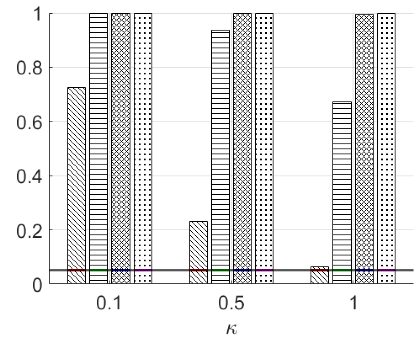
(g) $\mathbb{H}_1$, $K_1 = 2$, $p = 1$



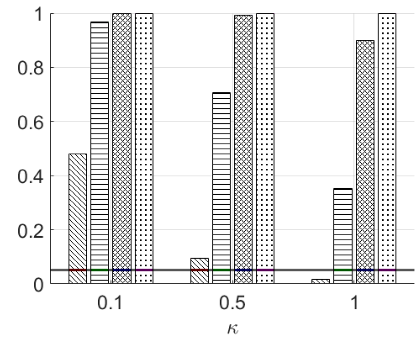
(h) $\mathbb{H}_1$, $K_1 = 3$, $p = 1$



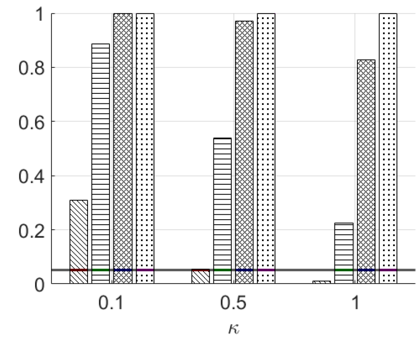
(i) $\mathbb{H}_1$, $K_1 = 4$, $p = 1$



(j) $\mathbb{H}_1$, $K_1 = 2$, $p = 12$

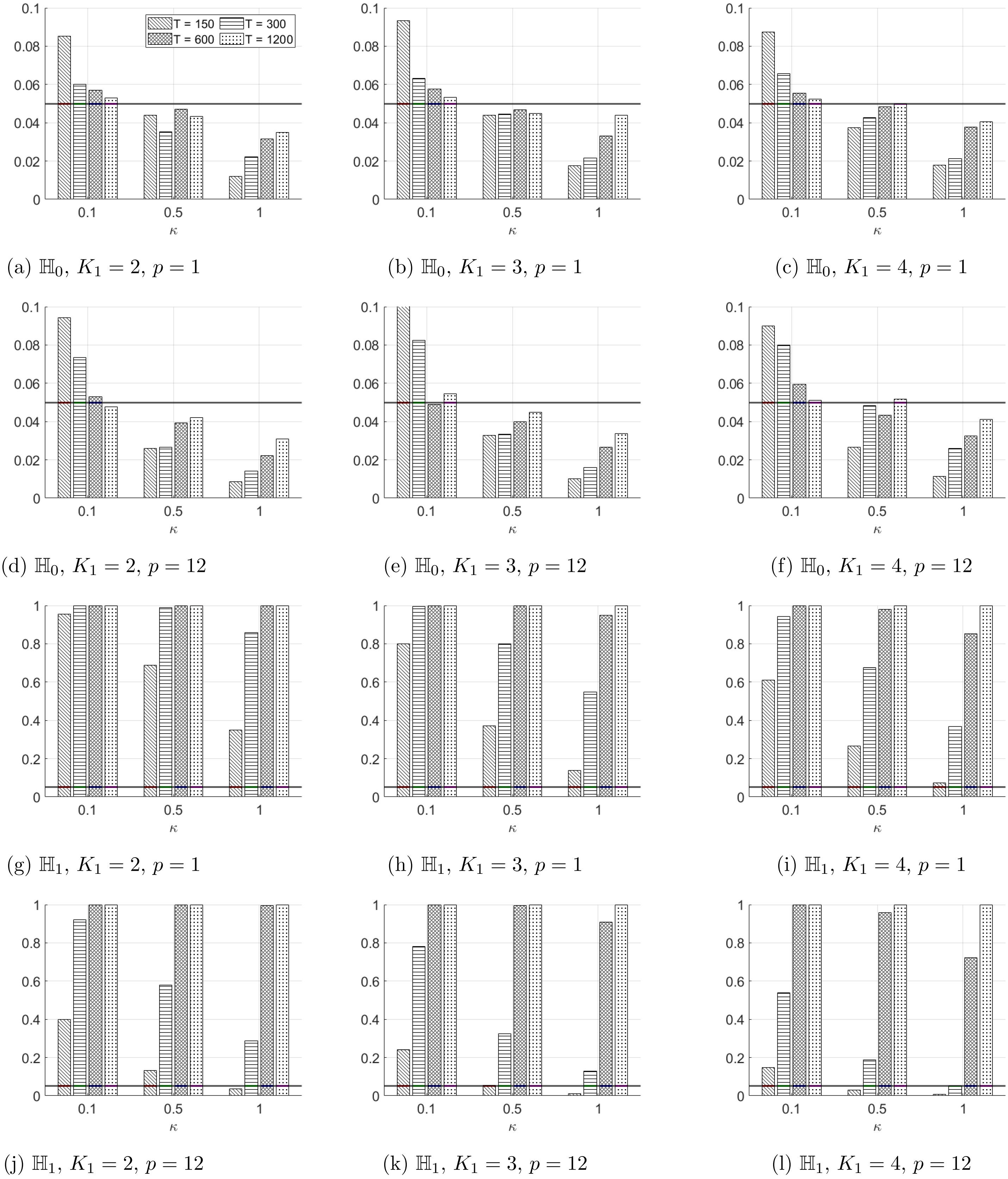


(k) $\mathbb{H}_1$, $K_1 = 3$, $p = 12$



(l) $\mathbb{H}_1$, $K_1 = 4$, $p = 12$

Figure 4: Relative rejection frequencies for DGP2. Nominal significance level 5%. True volatility change point at $T_1 = \frac{1}{2}T$. Assumed volatility change point at $T_1 = \frac{2}{5}T$.



(a) $\mathbb{H}_0$, $K_1 = 2$, $p = 1$ (b) $\mathbb{H}_0$, $K_1 = 3$, $p = 1$ (c) $\mathbb{H}_0$, $K_1 = 4$, $p = 1$

(d) $\mathbb{H}_0$, $K_1 = 2$, $p = 12$ (e) $\mathbb{H}_0$, $K_1 = 3$, $p = 12$ (f) $\mathbb{H}_0$, $K_1 = 4$, $p = 12$

(g) $\mathbb{H}_1$, $K_1 = 2$, $p = 1$ (h) $\mathbb{H}_1$, $K_1 = 3$, $p = 1$ (i) $\mathbb{H}_1$, $K_1 = 4$, $p = 1$

(j) $\mathbb{H}_1$, $K_1 = 2$, $p = 12$ (k) $\mathbb{H}_1$, $K_1 = 3$, $p = 12$ (l) $\mathbb{H}_1$, $K_1 = 4$, $p = 12$

Figure 5: Relative rejection frequencies for DGP2. Nominal significance level 5%. True volatility change point at $T_1 = \frac{1}{2}T$. Volatility change point obtained via minimizing the likelihood criterion $\psi(T_1)$ given in equation (15) over a rough grid of sample points.

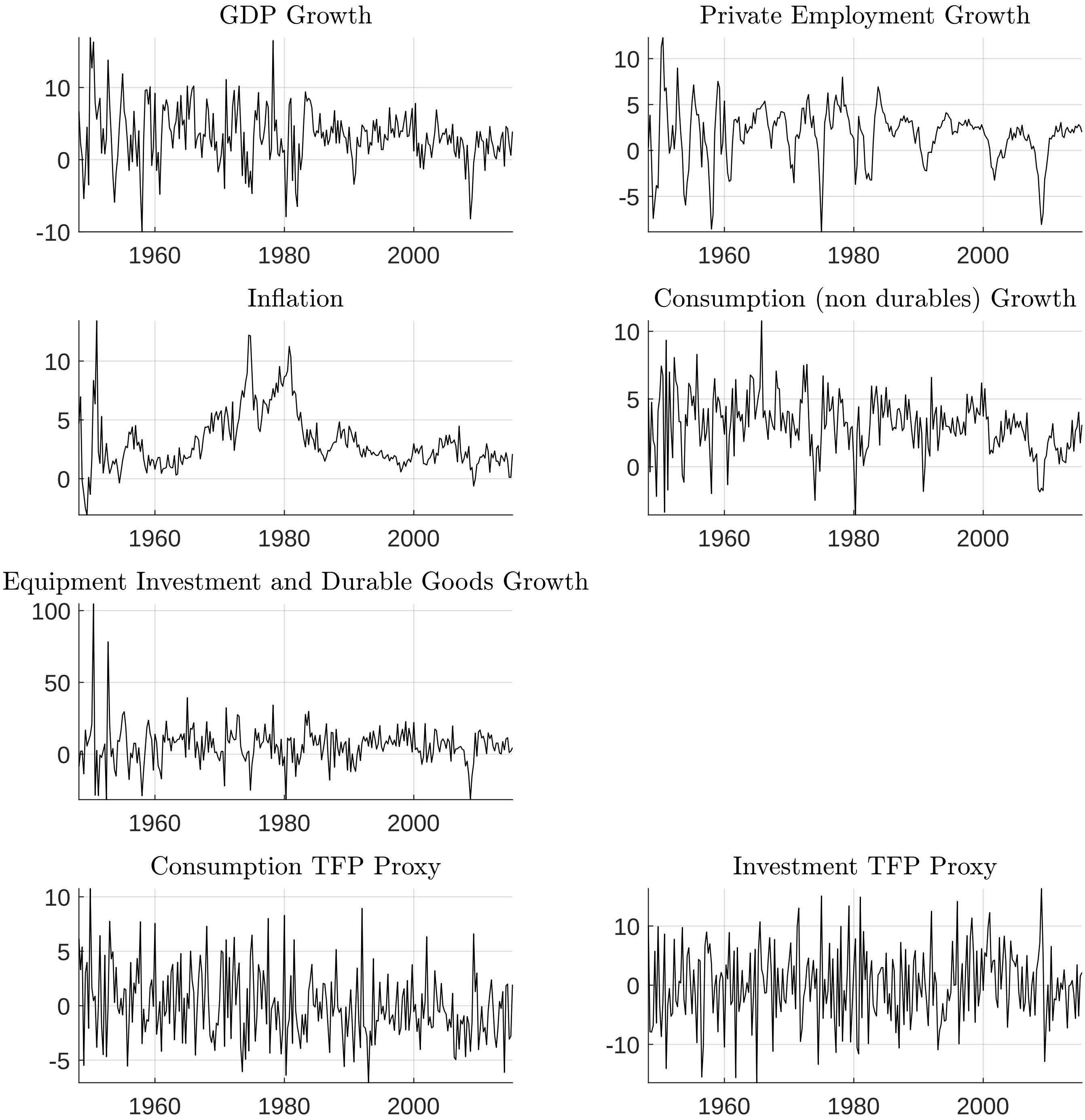


Figure 6: Variables and proxies of the Lunsford (2015) dataset.

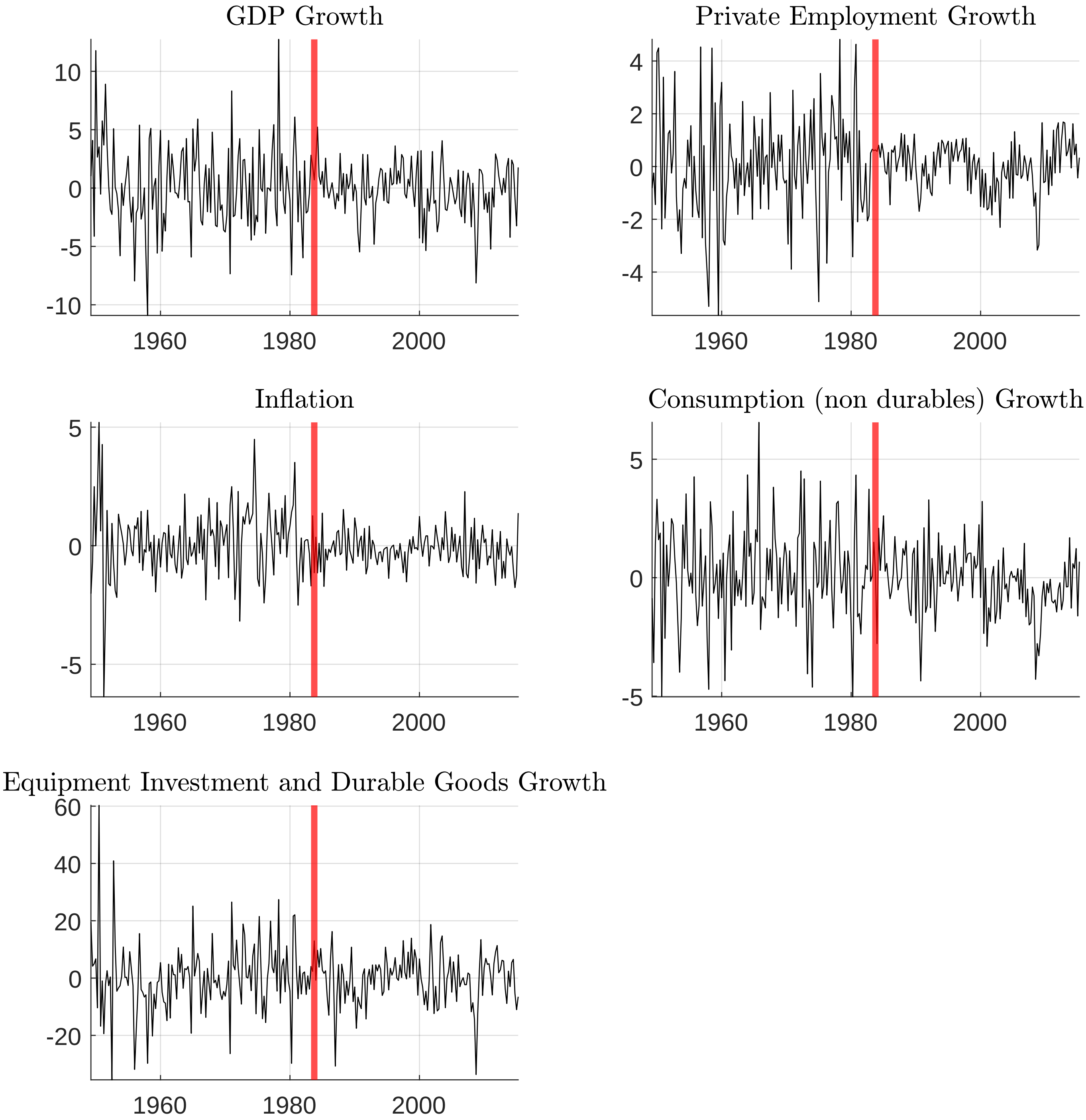


Figure 7: OLS residuals from 1949Q2 to 2015Q2 for the VAR(4) U.S. macro model. The solid red line indicates the possible variance change point in 1983Q4.

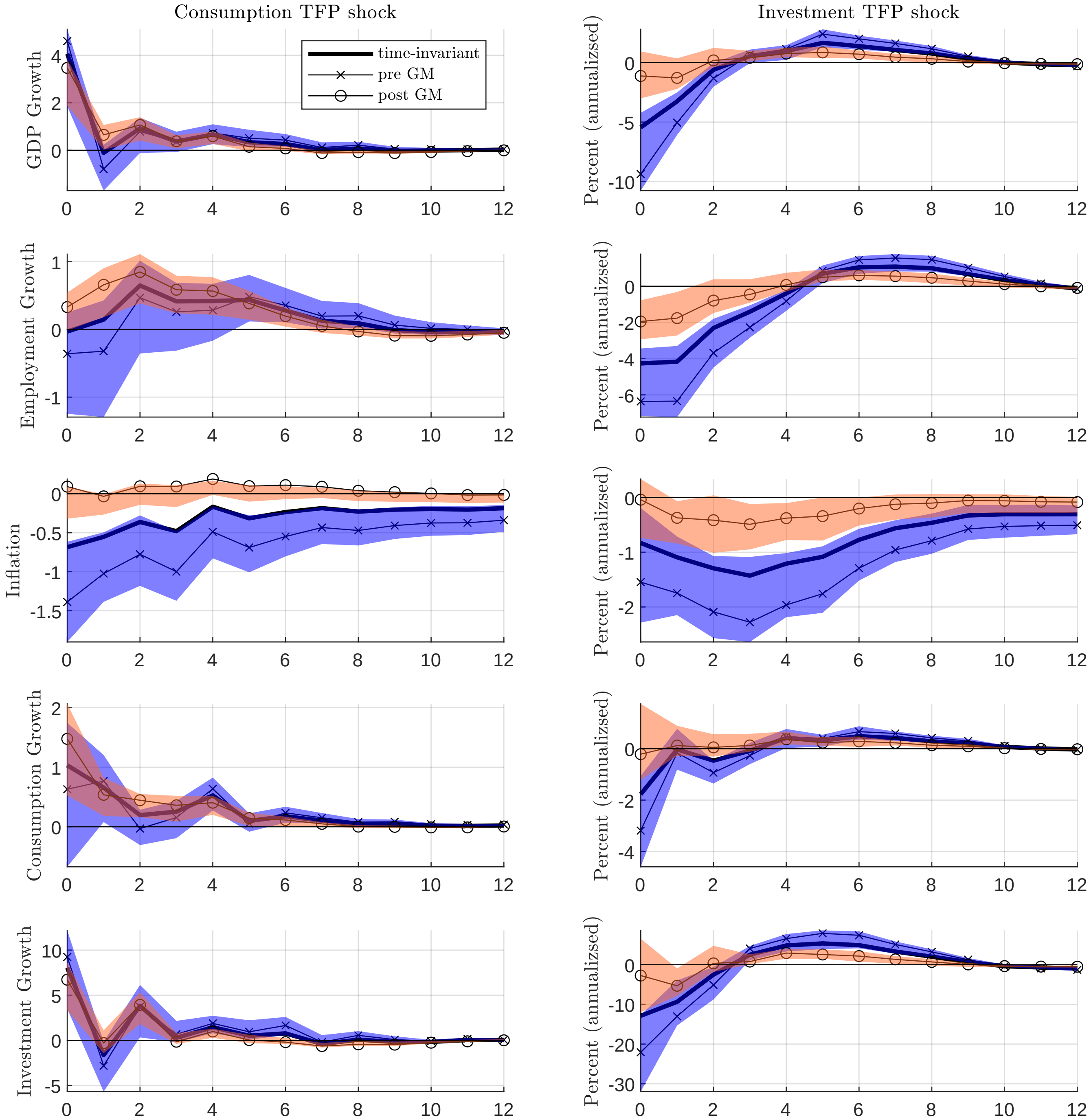


Figure 8: Comparison of responses to TFP shocks assuming time-varying or, alternatively, time-invariant dynamic responses with 90% pointwise confidence intervals based on a MBB. (Volatility regime 1: 1949Q2 - 1983Q4 (blue); Volatility regime 2: 1984Q1 - 2015Q2 (red); Time-invariant impulse responses: solid black line.)

C Additional Simulation Results

Table C.1: Relative Rejection Frequencies for DGP1 with Lag Length p, $\rho = 0$, and Nominal Significance Level $\alpha = 5\%$

		$\mathbb{H}_0$					
		$p = 1$			$p = 12$		
κ	T	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$
0.2346	150	0.078	0.077	0.076	0.074	0.069	0.072
	300	0.067	0.066	0.063	0.064	0.062	0.059
	600	0.057	0.057	0.058	0.061	0.060	0.057
	1200	0.057	0.054	0.054	0.051	0.049	0.054
1	150	0.068	0.072	0.074	0.062	0.050	0.059
	300	0.060	0.057	0.063	0.053	0.053	0.061
	600	0.052	0.051	0.049	0.056	0.059	0.056
	1200	0.054	0.051	0.055	0.054	0.051	0.053
3	150	0.047	0.044	0.050	0.034	0.026	0.040
	300	0.050	0.056	0.054	0.050	0.037	0.045
	600	0.055	0.047	0.050	0.047	0.045	0.051
	1200	0.054	0.049	0.050	0.051	0.052	0.050
		$\mathbb{H}_1$					
		$p = 1$			$p = 12$		
κ	T	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$
0.2346	150	0.716	0.096	0.500	0.402	0.013	0.258
	300	0.928	0.146	0.793	0.864	0.061	0.683
	600	0.997	0.212	0.971	0.994	0.171	0.953
	1200	1.000	0.328	0.999	1.000	0.300	0.998
1	150	0.651	0.068	0.446	0.307	0.010	0.212
	300	0.897	0.106	0.725	0.807	0.025	0.592
	600	0.991	0.160	0.938	0.988	0.113	0.921
	1200	1.000	0.238	0.996	1.000	0.205	0.996
3	150	0.511	0.041	0.335	0.199	0.008	0.159
	300	0.821	0.071	0.595	0.671	0.013	0.464
	600	0.976	0.114	0.844	0.955	0.048	0.818
	1200	1.000	0.166	0.972	0.999	0.133	0.968

Table C.2: Relative Rejection Frequencies for DGP1 with Lag Length p, $\rho = 0.5$, and Nominal Significance Level $\alpha = 5\%$

		$\mathbb{H}_0$					
		$p = 1$			$p = 12$		
κ	T	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$
0.2346	150	0.078	0.077	0.079	0.069	0.065	0.069
	300	0.059	0.067	0.065	0.061	0.060	0.066
	600	0.060	0.057	0.059	0.060	0.058	0.052
	1200	0.052	0.053	0.055	0.057	0.051	0.056
1	150	0.070	0.066	0.071	0.066	0.059	0.063
	300	0.059	0.057	0.061	0.054	0.060	0.061
	600	0.050	0.053	0.056	0.055	0.058	0.057
	1200	0.055	0.055	0.053	0.059	0.051	0.052
3	150	0.051	0.045	0.058	0.034	0.024	0.036
	300	0.050	0.048	0.053	0.051	0.041	0.046
	600	0.052	0.052	0.056	0.054	0.049	0.044
	1200	0.049	0.047	0.050	0.053	0.047	0.049
		$\mathbb{H}_1$					
		$p = 1$			$p = 12$		
κ	T	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$	$\beta(1) = \beta(2)$	$\beta(1) = \beta(3)$	$\beta(2) = \beta(3)$
0.2346	150	0.712	0.092	0.484	0.419	0.019	0.267
	300	0.934	0.155	0.800	0.861	0.073	0.669
	600	0.997	0.237	0.968	0.994	0.193	0.958
	1200	1.000	0.340	0.999	1.000	0.308	0.998
1	150	0.678	0.079	0.441	0.333	0.015	0.224
	300	0.923	0.127	0.731	0.828	0.042	0.587
	600	0.994	0.194	0.929	0.989	0.148	0.912
	1200	1.000	0.280	0.991	1.000	0.243	0.993
3	150	0.560	0.067	0.346	0.211	0.009	0.176
	300	0.860	0.103	0.591	0.723	0.023	0.479
	600	0.986	0.156	0.841	0.971	0.080	0.793
	1200	1.000	0.209	0.960	1.000	0.189	0.950

Table C.3: Relative Rejection Frequencies for DGP2 with Lag Length p and Nominal Significance Level 5%, Correctly Specified Volatility Change Point

		κ	K_1	$T = 150$	$T = 300$	$T = 600$	$T = 1200$
Under $\mathbb{H}_0$		0.1	2	0.063	0.058	0.056	0.053
			3	0.067	0.060	0.051	0.056
			4	0.062	0.053	0.054	0.050
		0.5	2	0.027	0.037	0.038	0.048
	$p = 1$		3	0.031	0.039	0.042	0.046
			4	0.039	0.035	0.044	0.054
		1	2	0.011	0.014	0.030	0.041
			3	0.013	0.025	0.039	0.037
			4	0.015	0.027	0.034	0.042
		0.1	2	0.030	0.045	0.057	0.050
			3	0.034	0.053	0.054	0.053
			4	0.031	0.051	0.049	0.054
		0.5	2	0.009	0.028	0.036	0.046
	$p = 12$		3	0.007	0.025	0.040	0.050
			4	0.011	0.031	0.045	0.046
		1	2	0.001	0.011	0.026	0.038
			3	0.003	0.013	0.031	0.037
			4	0.003	0.015	0.037	0.035
Under $\mathbb{H}_1$		0.1	2	0.980	1.000	1.000	1.000
			3	0.822	0.995	1.000	1.000
			4	0.645	0.959	1.000	1.000
		0.5	2	0.697	0.990	1.000	1.000
	$p = 1$		3	0.405	0.848	0.998	1.000
			4	0.236	0.704	0.985	1.000
		1	2	0.326	0.858	0.999	1.000
			3	0.144	0.533	0.950	1.000
			4	0.080	0.351	0.867	0.999
		0.1	2	0.681	1.000	1.000	1.000
			3	0.429	0.970	1.000	1.000
			4	0.266	0.885	1.000	1.000
		0.5	2	0.202	0.931	1.000	1.000
	$p = 12$		3	0.090	0.679	0.992	1.000
			4	0.058	0.491	0.972	1.000
		1	2	0.042	0.631	0.996	1.000
			3	0.017	0.317	0.904	1.000
			4	0.011	0.194	0.780	0.998

Table C.4: Relative Rejection Frequencies for DGP2 with Lag Length p and Nominal Significance Level 5%, Misspecified Volatility Change Point at $0.4T$

		κ	K_1	$T = 150$	$T = 300$	$T = 600$	$T = 1200$
Under $\mathbb{H}_0$	$p = 1$	0.1	2	0.069	0.060	0.051	0.059
			3	0.074	0.057	0.053	0.054
			4	0.063	0.057	0.051	0.055
		0.5	2	0.025	0.042	0.045	0.052
			3	0.037	0.042	0.048	0.041
			4	0.032	0.046	0.041	0.047
		1	2	0.017	0.022	0.031	0.041
			3	0.016	0.024	0.030	0.039
			4	0.015	0.028	0.034	0.045
	$p = 12$	0.1	2	0.040	0.053	0.054	0.055
			3	0.046	0.050	0.057	0.055
			4	0.037	0.057	0.054	0.054
		0.5	2	0.014	0.029	0.040	0.045
			3	0.011	0.030	0.041	0.046
			4	0.010	0.029	0.047	0.052
		1	2	0.003	0.012	0.026	0.036
			3	0.002	0.015	0.031	0.035
			4	0.002	0.014	0.032	0.039
Under $\mathbb{H}_1$	$p = 1$	0.1	2	0.983	1.000	1.000	1.000
			3	0.827	0.993	1.000	1.000
			4	0.675	0.964	1.000	1.000
		0.5	2	0.716	0.987	1.000	1.000
			3	0.416	0.855	0.996	1.000
			4	0.279	0.744	0.986	1.000
		1	2	0.350	0.871	0.998	1.000
			3	0.142	0.556	0.945	1.000
			4	0.092	0.395	0.897	0.999
	$p = 12$	0.1	2	0.726	0.999	1.000	1.000
			3	0.480	0.968	1.000	1.000
			4	0.309	0.888	0.999	1.000
		0.5	2	0.232	0.937	1.000	1.000
			3	0.095	0.706	0.993	1.000
			4	0.054	0.539	0.970	1.000
		1	2	0.063	0.672	0.995	1.000
			3	0.019	0.352	0.898	1.000
			4	0.012	0.224	0.827	0.999

Table C.5: Relative Rejection Frequencies for DGP2 with Lag Length p and Nominal Significance Level 5%, Volatility Change Point Search via Likelihood Criterion

		κ	K_1	$T = 150$	$T = 300$	$T = 600$	$T = 1200$
Under $\mathbb{H}_0$	$p = 1$	0.1	2	0.085	0.060	0.057	0.053
			3	0.093	0.063	0.058	0.053
			4	0.087	0.066	0.055	0.052
		0.5	2	0.044	0.035	0.047	0.043
			3	0.044	0.045	0.047	0.045
			4	0.037	0.043	0.048	0.050
		1	2	0.012	0.022	0.032	0.035
			3	0.017	0.022	0.033	0.044
			4	0.018	0.021	0.038	0.040
	$p = 12$	0.1	2	0.094	0.074	0.053	0.048
			3	0.111	0.083	0.049	0.055
			4	0.090	0.080	0.059	0.051
		0.5	2	0.026	0.027	0.039	0.042
			3	0.033	0.033	0.040	0.045
			4	0.026	0.048	0.043	0.052
		1	2	0.008	0.014	0.022	0.031
			3	0.010	0.016	0.027	0.034
			4	0.011	0.026	0.032	0.041
Under $\mathbb{H}_1$	$p = 1$	0.1	2	0.956	1.000	1.000	1.000
			3	0.800	0.997	1.000	1.000
			4	0.610	0.942	1.000	1.000
		0.5	2	0.688	0.989	1.000	1.000
			3	0.372	0.801	0.998	1.000
			4	0.266	0.675	0.980	1.000
		1	2	0.349	0.858	0.999	1.000
			3	0.138	0.549	0.948	1.000
			4	0.072	0.368	0.854	1.000
	$p = 12$	0.1	2	0.400	0.922	1.000	1.000
			3	0.240	0.781	1.000	1.000
			4	0.146	0.538	0.999	1.000
		0.5	2	0.133	0.578	1.000	1.000
			3	0.054	0.324	0.994	1.000
			4	0.028	0.189	0.957	1.000
		1	2	0.036	0.287	0.995	1.000
			3	0.012	0.129	0.909	1.000
			4	0.008	0.050	0.721	0.997