Non-Gaussian Structural Vector Autoregression with Unknown Break Points

Keyan Liu*
Faculty of Social Sciences, University of Helsinki

December 26, 2024

Abstract

In this paper, I consider testing and estimating non-Gaussian Structural Vector Autoregressive models with unknown break points (SVAR-BP). This model extends traditional SVAR analysis by allowing for unknown breakpoints, capturing potential changes in both autoregressive coefficients and structural parameters. I employ the Partial Sample Generalized Method of Moments (PSGMM) to estimate the model and utilize the sup-Wald test to assess parameter stability. Additionally, I establish the asymptotic properties of the break point estimators and propose a sequential procedure for detecting and estimating multiple break points. My method is applied to a U.S. macroeconomic dataset from 1954 to 2023, where I identify significant structural breaks corresponding to key economic events. The results demonstrate the ability of my approach to detect and estimate multiple break points, modeling shifts in the dynamics of economic variables.

Keywords: Structural vector autoregression; Structural break; break point test.

*E-mail address: keyan.liu@helsinki.fi. I am grateful to Jani Luoto, Markku Lanne, Mika Meitz, and Leena Kalliovirta for their insightful comments. I thank the participants of the Econometrics Workshop at Helsinki GSE for their valuable feedback, which contributed to the development of this work. Financial support from the Research Council of Finland (grant 347986) is gratefully acknowledged.

1 Introduction

The Structural Vector Autoregression (SVAR) model is a widely used econometric tool for analyzing dynamic interactions among macroeconomic variables. Since Sims (1980) popularized SVARs, they have become a standard in empirical macroeconomics for studying the effects of structural shocks. SVAR models traditionally impose restrictions on contemporaneous relationships between variables to provide structural interpretations of economic systems. However, these models often assume constant relationships over time, which may overlook structural breaks—a limitation increasingly highlighted by empirical evidence that many economies undergo shifts due to policy changes, technological advancements, and external shocks.

Recent contributions, such as (Magnusson and Mavroeidis, 2014; Demetrescu and Salish, 2024; Lewis, 2021, 2024), emphasize the need for flexible models that can account for parameter instability within macroeconomic relationships. This concern aligns with research on time-varying SVARs, which examine how instability in structural parameters—particularly in the contemporaneous response matrix B—affects inference. While models like those of Carriero et al. (2018) and Angelini et al. (2019) allow for time-varying parameters, they generally assume a fixed structure in B, limiting their ability to capture full structural shifts. Bayesian approaches by Primiceri (2005) and Cogley and Sargent (2005) accommodate time-varying B but often impose strong prior assumptions, resulting in parameter inferences shaped more by the priors than by data-driven insights. Even frequentist models like those in Auerbach and Gorodnichenko (2012) that incorporate time-varying parameters in reduced-form models do not fully allow for variation in B.

Model instability has been widely documented, particularly during periods like the "Great Moderation," a decline in macroeconomic volatility since the mid-1980s, as noted by McConnell and Perez-Quiros (2000), Stock and Watson (2002), and Justiniano and Primiceri (2008). Some attribute this moderation to reduced shock volatility, or "good luck" (Sims and Zha, 2006; Stock and Watson, 2002), while others argue that shifts in monetary policy and structural changes in private-sector behavior contributed to this instability (Boivin and Giannoni, 2006; Inoue and Rossi, 2011). Recent studies also investigate the flattening relationship between inflation and the business cycle, as represented by the Phillips curve, documented by scholars like Del Negro et al. (2020) and Bergholt et al. (2023).

This paper addresses these challenges by introducing a Non-Gaussian Structural Vector Autoregression model with structural breaks (SVAR-BP), extending the SVAR framework to allow for unknown breakpoints in both autoregressive coefficients and structural parameters. Our approach addresses parameter instability and leverages non-Gaussianity in the error structure to achieve identification through higher-order moments, as demonstrated by Velasco (2023), Lanne and Luoto (2021), Lanne et al. (2023a) and Lanne et al. (2023b). By using skewness and kurtosis properties, the model captures asymmetries and heavy tails often observed in macroeconomic data, offering a robust solution to the problem of parameter instability.

Unlike models relying on exogenously specified break dates, a limitation in many empirical SVAR applications (Bacchiocchi and Fanelli, 2015; Rigobon, 2003; Bacchiocchi and Kitagawa, 2024), our model determines break dates endogenously. This flexibility allows for identifying structural shifts that align with major policy changes or financial crises, crucial in macroeconomic contexts where breaks often correspond to significant economic events.

Estimation of structural breaks and break points has been a significant focus in econometrics, particularly within time series models. Foundational methodologies by Bai (1994) and Bai and Perron (1998) introduced techniques for detecting and estimating multiple structural breaks in regression models, with later works like Bai (2000) and Ling (2016) extending these techniques to account for structural changes in autoregressive models. However, these models were largely developed under Gaussian assumptions, unable to capture the structural matrix B in SVAR contexts without additional restrictions. In contrast to these approaches, my paper introduces a novel break date estimation function specifically tailored to non-Gaussian SVAR models. This new method not only allows for endogenous break date estimation but also incorporates potential changes in the structural matrix B. Leveraging higher-order moment conditions such as skewness and kurtosis, the model identifies B through non-Gaussianity, thereby providing a robust framework for analyzing economic systems where structural dynamics may evolve over time.

The contributions of this paper are threefold: (1) the integration of non-Gaussianity into SVAR models to improve identification and enable testing for parameter instability, (2) a method to estimate potential break dates within the model, addressing a gap in the existing literature that typically assumes predetermined breakpoints, and (3) the development of a Partial Sample Generalized Method of Moments (PSGMM) estimator for the SVAR-BP model, facilitating robust parameter estimation and breakpoint detection in the presence of non-Gaussian errors and common volatility structures across the structural shocks. This sequential estimation approach builds on the structural break methodologies of Bai (1997) and Bai and Perron (2003).

To demonstrate the practical application of the model, I apply it to U.S. macroeconomic data from 1954 to 2024, encompassing several major economic events, including shifts in monetary policy regimes and the global financial crisis. The findings reveal significant structural breaks aligned with these events, highlighting the model's capacity to capture shifts in economic dynamics and improve understanding of how external shocks propagate through the economy.

By addressing the limitations of traditional SVAR models and incorporating structural breaks and non-Gaussian errors, this paper offers a flexible and robust framework for macroeconomic analysis. It contributes to the literature on structural breaks in time series models and the identification of SVARs with non-Gaussian shocks, offering insights into economic shock dynamics and enabling a method for endogenously testing possible break dates within the model.

The remainder of this paper is organized as follows. In Section 2, I develop the methodology for testing and estimating the SVAR model with one structural break. This includes a detailed discussion of how the break point is determined and how non-Gaussianity aids in model identification. Section 3 extends the analysis to the case of multiple structural breaks, introducing a sequential procedure for identifying and estimating multiple break points within the SVAR framework. In Section 4, I provide the statistical inference framework, constructing confidence sets for impulse responses hypothesis tests on the parameters and impulse responses across different regimes. Section 5 presents an empirical illustration using U.S. macroeconomic data, highlighting the effectiveness of the proposed methodology in detecting and interpreting structural breaks in economic relationships. Finally, Section 6 concludes the paper, summarizing the key findings and outlining directions for future research.

2 Model with one break point

I consider the following structural vector autoregressive model of order p with an unknown structural break (SVAR-BP):

$$\begin{aligned} y_t &= \nu_1 + A_{1,1} y_{t-1} + \cdots + A_{1,p} y_{t-p} + B_1 \varepsilon_t, \quad \text{for} \quad 1 \leq t \leq k, \\ y_t &= \nu_2 + A_{2,1} y_{t-1} + \cdots + A_{2,p} y_{t-p} + B_2 \varepsilon_t, \quad \text{for} \quad k+1 \leq t < T, \end{aligned} \tag{1}$$

where y_t is the n-dimensional time series of interest, ν_1 and ν_2 are $(n \times 1)$ intercept terms, and $\{A_{1,j}\}$ and $\{A_{2,j}\}$ are $(n \times n)$ parameter matrices. ε_t is an i.i.d. sequence satisfying $E(\varepsilon_t \varepsilon_t') = I_n$. I assume y_t to be partially stable, i.e.,

$$\det A_i(z) \stackrel{\text{def}}{=} \det\left(I_n - A_{i,1} z - \cdots - A_{i,p} z^p\right) \neq 0, \quad \text{for} \quad |z| \leq 1, \quad \text{for } i = 1, 2 \tag{2}$$

Model (1) extends the previous literature on Structural Change Vector AutoRegression by Bai et al. (1998) and Bai (2000) by considering changes not only in autoregressive coefficients but also in structural parameters B. This is a reasonable consideration, as it is recognized that structural breaks may have marked consequences on both the transmission and propagation mechanisms of shocks. However, considering changes in structural parameter B, we face the identification problem in the SVAR analysis. In this paper, I follow the literature on statistical identification through non-Gaussianity of the structural errors. Specifically, I use the error assumptions by Lanne et al. (2023a) and Mesters and Zwiernik (2022):

Assumption 1. *(third moments)*

(i) The component processes ε_{it}, $i = 1, \ldots, n$, have zero co-skewness.

(ii) At most one component of ε_t has zero skewness.

Assumption 2. *(fourth moments)*

(i) The component processes ε_{it}*,* $i = 1, \ldots, n$*, have zero co-kurtosis.*

(ii) At most one component of ε_t *has zero excess kurtosis.*

Assumption 3. *(reflection-invariant)*

(i) The component processes ε_{it}*,* $i = 1, \ldots, n$*, have zero co-kurtoses except that the* $n(n-1)/2$ *symmetric co-kurtosis* $E(\varepsilon_{it}^2 \varepsilon_{jt}^2) - 1, i < j$ *can be nonzero.*

(ii) $\sum_{i=1}^{n} E(\varepsilon_{it}^2 \varepsilon_{jt}^2) \neq \sum_{i=1}^{n} E(\varepsilon_{it}^2 \varepsilon_{kt}^2)$, *for all* $j \neq k$.

Mesters and Zwiernik (2022) demonstrate that each of these assumptions, combined with the i.i.d. condition for ε_t and an identity covariance matrix, is sufficient for locally identifying the matrix B in the linear process $u_t = B\varepsilon_t$. The assumptions leverage the non-Gaussian features of structural errors, using higher-order moments to distinguish structural components, which is essential for identifying B without relying on full independence of shocks.

Assumption 1 centers on third moments, specifically skewness. The zero co-skewness condition implies that interactions between different components of ε_t are symmetric, meaning that the cross-component third moments are zero. At most one component of ε_t may have zero skewness, ensuring that most components exhibit asymmetry. This asymmetry provides the necessary variability for identifying B, as the presence of skewness introduces additional dimensions for distinguishing structural shocks. Importantly, this assumption does not preclude common volatility processes, allowing the model to accommodate conditional heteroskedasticity without imposing strict independence among the components.

The second assumption, 2, focuses on fourth moments (kurtosis). This assumption requires zero co-kurtosis across the components, meaning no excess kurtosis in joint distributions. However, at most one component of ε_t may have zero excess kurtosis, ensuring that the majority of components display heavier tails or sharper peaks than Gaussian distributions. This non-Gaussianity enables the model to distinguish between structural shocks through the heavier-tailed characteristic of the data. Assumption 2 is stricter than

Assumption 1in that it does not accommodate common volatility processes among the components, thus limiting interaction effects in fourth moments. Nevertheless, this assumption remains weaker than full independence, allowing some interaction forms while maintaining flexibility in real-world applications.

Assumption 3 relaxes the zero co-kurtosis condition, permitting specific nonzero symmetric co-kurtosis terms, $E(\varepsilon_{it}^2\varepsilon_{jt}^2)-1$. This allows for potential common volatility processes among the components, facilitating situations where some components may exhibit correlated volatility. The condition that no two components share identical relationships in their fourth moments aids in identifying B by ensuring distinct variation patterns among components.

Together, these assumptions leverage non-Gaussian characteristics (skewness and kurtosis) to provide information for locally identifying the B matrix. The (ii) clauses in each assumption reflect "generic assumptions," which are typically met in real-world data where the errors are rarely purely Gaussian. Without these non-Gaussian features, specifically in cases where errors are Gaussian, B would remain unidentified, as Gaussian errors lack the asymmetry and heavy tails needed to separate structural shocks effectively.

2.1 Moment Conditions

The model can be consistently estimated using Generalized Method of Moments (GMM). In this section, I present the moment conditions utilized in the partial sample GMM estimation. From the assumption that the structural shocks ε_t are i.i.d., I derive the following $n + pn^2$ unconditional moment conditions:

$$E(\varepsilon_t \otimes Z_{t-1}) = 0_{n(np+1)\times 1} \tag{3}$$

where $Z_t = (1, y'_{t-1}, \ldots, y'_{t-p})'$ is a vector of lagged observations.

The assumption that ε_t has an identity variance-covariance matrix yields $n(n+1)/2$ additional orthogonality conditions of the form:

$$E[e'_i\varepsilon_t e'_j\varepsilon_t] - \delta_{ij} = 0, \quad \text{for all} \quad 0 \leq i, j \leq n \tag{4}$$

Here, e_i is the i-th column of the identity matrix I_n, and δ_{ij} is the Kronecker delta, defined as:

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

While the $2n + pn^2 + n(n-1)/2$ moment conditions in Equations (3) and (4) are necessary, they are insufficient for global identification of the parameters. To fully identify the model, higher-order moment conditions are employed. If Assumption 1 (concerning third moments) holds, the following co-skewness conditions are satisfied:

$$E[e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t] = 0, \quad \text{for} \quad 0 \leq i, j, k \leq n, \quad |\{i, j, k\}| > 1 \tag{5}$$

where $|\cdot|$ denotes the cardinality of the set.

Similarly, under Assumption 2 (concerning fourth moments), I obtain the following co-kurtosis conditions:

$$E[e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t e_l'\varepsilon_t] - \delta_{ijkl} = 0, \quad \text{for} \quad 0 \leq i, j, k, l \leq n, \quad |\{i, j, k, l\}| > 1 \tag{6}$$

Here, δ_{ijkl} is defined as:

$$\delta_{ijkl} = \begin{cases} 1 & \text{if } i, j, k, l \text{ form two distinct pairs}, \\ 0 & \text{otherwise}. \end{cases}$$

If Assumption 3 (reflection-invariance) is satisfied, I apply a modified version of the co-kurtosis condition in Equation (6), excluding the $n(n-1)/2$ symmetric co-kurtosis terms $E(\varepsilon_{it}^2\varepsilon_{jt}^2) - 1$ for $i < j$. Thus, I obtain:

$$E[e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t e_l'\varepsilon_t] = 0, \quad \text{for} \quad |\{i, j, k, l\}| > 1, \quad i, j, k, l \text{ do not form two distinct pairs} \tag{7}$$

These moment conditions are applicable over the full sample period. However, when a structural break is present, the structural shocks ε_t take different forms in the two subsamples:

$$B_1\varepsilon_t = y_t - \nu_1 - A_{1,1}y_{t-1} - \cdots - A_{1,p}y_{t-p}, \quad \text{for} \quad 1 \leq t \leq k \tag{8}$$

$$B_2\varepsilon_t = y_t - \nu_2 - A_{2,1}y_{t-1} - \cdots - A_{2,p}y_{t-p}, \quad \text{for} \quad k+1 \leq t \leq T \tag{9}$$

I define the following notation for the parameters of interest: $\pi_i = \text{vec}(\nu_i, A_{i,1}, \ldots, A_{i,p})$ for $i = 1, 2$, and $\vartheta_i = \text{vec}(B_i)$. The parameter vector θ_i is then written as $\theta_i = \text{vec}(\pi_i', \vartheta_i')'$. The true value of the parameters are denoted as π_{i0}, ϑ_{i0} and θ_{i0}. The moment conditions can be categorized as follows: - $f_s(y_t, \theta)$ includes moment conditions (3)-(5), - $f_k(y_t, \theta)$ includes moment conditions (3), (4), and (6), - $f_r(y_t, \theta)$ includes moment conditions (3), (4), and (7).

2.2 Identification Result

The following identification result applies to the model

Proposition 1. *In model 1, suppose ε_t has finite fourth moments,*

(i) If Assumption 1 is satisfied, then $E(f_s(y_t, \theta)) = 0 \Rightarrow \pi_i = \pi_{i0}, B_i = B_{i0}Q$.

(ii) If Assumption 2 is satisfied, then $E(f_k(y_t, \theta)) = 0 \Rightarrow \pi_i = \pi_{i0}, B_i = B_{i0}Q$.

(iii) If Assumption 3 is satisfied, then $E(f_r(y_t, \theta)) = 0 \Rightarrow \pi_i = \pi_{i0}, B_i = B_{i0}Q$.

Here, Q is a signed permutation matrix of size $n \times n$.

Proof. See Appendix A

□

Proposition 1 presents a crucial identification result for the SVAR model under the assumption of finite fourth moments, confirming that the structural parameters in our model are uniquely identified up to a signed permutation matrix. Specifically, the proposition states that under any of the third and fourth moment assumptions (Assumptions 1, 2, and 3), the vector parameter π is uniquely identified, and the structural impact matrix B_1 and B_2 is identified up to a matrix Q, which is a signed permutation matrix. This result ensures that each of these sets of moment conditions provides sufficient information to identify the model's structural parameters.

This identification result (i.e., identification up to sign and permutation) is well-established in the literature, particularly in the context of linear systems of the form $u_t = B\varepsilon_t$, where B is identified up to sign and permutation. For example, Keweloh (2021) and Mesters and Zwiernik (2022) provide similar identification results, demonstrating that the matrix B in a linear system is identifiable up to a signed permutation matrix when either skewness or kurtosis assumptions are applied. My contribution extends these results to the SVAR context, where both the structural impact matrix B and the VAR coefficients (contained in π) are shown to be globally identified. This extension is significant because it allows for the identification of not only the contemporaneous impact of shocks but also the autoregressive dynamics in a VAR system, which is essential for studying dynamic macroeconomic relationships.

Additionally, result (ii) in Proposition 1 is closely related to the identification framework in Lanne and Luoto (2021). They use a subset of co-kurtosis conditions to identify the matrix B in SVAR models. While they also discuss identification of the VAR coefficients, their result is local. In contrast, our proposition provides a global identification result for both B and π, ensuring robust and global identification of the entire model.

To acquire a uniquely globally identified model, I further establish a unique and sign configuration as proposed by Lanne et al. (2017). Define M_n as the set of all invertible n by n matrices.

Identification scheme. Let $B \in M_n$ and transform it to $\bar{B} = \Pi(B) = BPD$ as follows:

1. P is an $n \times n$ permutation matrix such that $G = BP$ satisfies $|g_{ii}| > |g_{ij}|$ for all $i < j$.

2. D is a diagonal matrix with ± 1 entries such that all diagonal elements of BPD are positive.

Define

$$\mathcal{B} = \{\bar{B} \in M_n : \exists B \in M_n : \Pi(B) = \bar{B}\}.$$

In this paper, I denote the parameter space $\pi_i \in \Pi$, where Π is a subset of $\mathcal{R}^{n+pn^2}$ and $B_i \in \mathcal{B}_1$, where $\mathcal{B}_1$ is a subset of $\mathcal{B}$.

2.3 Partial Sample GMM Estimator

The previous section outlined the moment conditions. In this section, I define the Partial Sample GMM Estimator (PSGMM) and derive its asymptotic properties in the absence of a structural break. This result is fundamental for constructing the structural break test, which will be discussed in the next section.

The theoretical results presented in the remainder of this paper hold regardless of which specific set of moment conditions, $f_s(y_t, \theta)$, $f_k(y_t, \theta)$, or $f_r(y_t, \theta)$, is used. For simplicity of notation, I will refer to the chosen set of moment functions generically as $f(y_t, \theta)$, representing any one of these three possible moment conditions. Under the null hypothesis of no structural break, the moment functions $E[f(y_t, \theta)] = 0$ are stable throughout the entire sample period. Under the alternative hypothesis, however, the moment functions $E[f(y_t, \theta_{10})] = 0$ hold for observations $1, \ldots, k$, and $E[f(y_t, \theta_{20})] = 0$ hold for the remaining $T - k$ observations.

If the breakpoint k were known, the GMM estimator for each subsample can be derived by minimizing the criterion function with respect to the respective parameters θ_1 and θ_2 in each segment. This allows us to estimate separate parameter values for the periods before and after the break, capturing potential structural shifts.

For the first subsample, I define the criterion function as follows:

$$\begin{aligned} Q_1(k, \theta_1) &= \bar{f}_1'(k, \theta_1)[W_1(k)]^{-1} \bar{f}_1(k, \theta_1), \\ \hat{\theta}_1(k) &= \arg\min_{\theta_1} Q_1(k, \theta_1), \end{aligned} \tag{10}$$

where

$$\bar{f}_1(k, \theta_1) = \frac{1}{k} \sum_{t=1}^{k} f(y_t, \theta_1).$$

Similarly, for the second subsample, the criterion function is defined as:

$$\begin{aligned} Q_2(k, \theta_2) &= \bar{f}_2'(k, \theta_2)[W_2(k)]^{-1} \bar{f}_2(k, \theta_2), \\ \hat{\theta}_2(k) &= \arg\min_{\theta_2} Q_2(k, \theta_2), \end{aligned} \tag{11}$$

where

$$\bar{f}_2(k, \theta_2) = \frac{1}{T-k} \sum_{t=k+1}^{T} f(y_t, \theta_2),$$

with $W_1(k)$ and $W_2(k)$ as weighting matrices. These matrices may depend on k, adjusting for differences in the number of observations within each subsample.

This two-step GMM approach allows the model to accommodate structural changes by tailoring parameter estimates to each regime. Estimating $\hat{\theta}_1(k)$ and $\hat{\theta}_2(k)$ separately enables the model to account for distinct dynamics across regimes, providing a more precise framework for analyzing potential breaks. In the subsequent section, this formulation forms the foundation for testing structural breaks by examining whether parameter values shift significantly at k.

I denote the true parameters under the null hypothesis of no structural break as $\theta_0' = (\theta_{10}', \theta_{10}')$. To facilitate the development of an asymptotic theory, the following assumptions are introduced:

Assumption 4. *The parameter spaces* Π *and* $\mathcal{B}_1$ *are compact.*

Assumption 5. *The weighting matrix* $W_1(k) = E(f(y_t, \theta_{10})f(y_t, \theta_{10})')$ *and* $W_2(k) = E(f(y_t, \theta_{20})f(y_t, \theta_{20})')$ *are both positive definite,*

and $E\|f(y_t, \theta_{10})\|^{2+\delta_1} < \infty$*,* $E\|f(y_t, \theta_{20})\|^{2+\delta_2} < \infty$ *for some* $\delta_1, \delta_2 > 0$*.*

Assumption 6. *The true value of break date* $k_0 = T\lambda_0$*, where* $\lambda_0 \in (0, 1)$*.*

The three assumptions outlined above provide the foundation for partial sample GMM estimation. Assumptions 4 and 5 are standard in the GMM literature. Assumption 4, which states that the parameter spaces Π and $\mathcal{B}_1$ are compact, ensures that the optimization procedure is well-behaved, as the parameters are confined within a bounded and closed set. Assumption 5 guarantees that the weighting matrix is optimal, ensuring that the moment functions are well-behaved, with finite second (and slightly higher) moments, and that their covariance matrix is positive definite. Together, these two assumptions ensure the existence of a global minimum during the estimation process. In practice, the weighting matrices are replaced by their consistent estimates:

$$W_1(k) = \frac{1}{k} \sum_{t=1}^{k} f(y_t, \hat{\theta}_1^0) \sum_{t=1}^{k} f(y_t, \hat{\theta}_1^0)', \quad W_2(k) = \frac{1}{T-k} \sum_{t=k+1}^{T} f(y_t, \hat{\theta}_2^0) \sum_{t=k+1}^{T} f(y_t, \hat{\theta}_2^0)'.$$

where $\hat{\theta}_1^0$ and $\hat{\theta}_2^0$ are the first-round OLS estimator from each subsample.

Assumption 6 is specific to structural break analysis. It assumes that the break point k_0 is proportional to the sample size T and falls within the interval $(0, 1)$. This ensures that the number of observations in each subsample increases as the total sample size increases. Such an assumption is essential for asymptotic analysis in structural break testing and estimation.

The following proposition presents the convergence result for the partial sample GMM estimator under the null hypothesis:

Proposition 2. *Suppose the structural errors satisfy one set of Assumptions 1, 2, or 3, corresponding to the moment function $f(y_t, \theta)$. Further, suppose Assumptions 4 through 6 hold, and $\theta_{10} = \theta_{20}$, Then,*

$$\sqrt{T}(\hat{\theta}(T\cdot) - \theta_0) \Rightarrow H(\cdot)^{-1}J(\cdot) \tag{12}$$

where

$$H(\lambda) = \begin{bmatrix} \lambda G'W^{-1}G & 0 \\ 0 & (1-\lambda)G'W^{-1}G \end{bmatrix}, \quad J(\lambda) = \begin{pmatrix} G'W^{-1/2}B(\lambda) \\ G'W^{-1/2}(B(1) - B(\lambda)) \end{pmatrix},$$

with $G = E\left(\frac{\partial f(y_t, \theta_{10})}{\partial \theta_1}\right)$, and $B(\lambda)$, $\lambda \in (0, 1)$, is a vector of independent Brownian motions on $(0, 1)$. The weak convergence holds in the Skorokhod space $D(0, 1)$.

For the proof of Proposition 2, see Appendix B.

The above proposition establishes the asymptotic properties of the partial sample GMM estimator under the null hypothesis of no structural break. It shows that the estimator, as a function of λ, converges at the standard $\sqrt{T}$-rate and follows a limiting Brownian motion. This result is crucial for constructing a structural break test, which I will discuss in the next section.

2.4 Structural Break Test

In this section, I formulate the null and alternative hypotheses of interest and propose a Wald test statistic to test for a structural break in the parameters of the model. my hypothesis is as follows:

$$H_0 : \theta_{10} = \theta_{20} \quad \text{versus} \quad H_1 : \theta_{10} \neq \theta_{20} \quad \text{for } 1 < k < T \tag{13}$$

When the break point k is known, classical approaches such as the Wald, LM (Lagrange Multiplier), or LR (Likelihood Ratio) tests can be used (see Andrews and Fair (1988) for examples of these tests within the GMM framework). Under the null hypothesis, these tests conform to a standard asymptotic χ^2 distribution.

In this paper, however, I consider the case where the break point k is unknown. I build upon the work of Andrews (1993), who developed a method for testing hypotheses with an unknown break point in the GMM framework. As shown in Proposition 2, the asymptotic variances for the estimators in the two subsamples are given by:

$$\text{Asy. Var}(\hat{\theta}_1) = \frac{1}{\lambda T} G' W^{-1} G, \tag{14}$$

$$\text{Asy. Var}(\hat{\theta}_2) = \frac{1}{(1-\lambda) T} G' W^{-1} G, \tag{15}$$

where $\lambda = k/T$.

Let $V = G'W^{-1}G$ represent the asymptotic variance. A consistent estimator for V is given by:

$$\hat{V}_i = \hat{G}_i{}' \hat{W}_i{}^{-1} \hat{G}_i, \tag{16}$$

where $\hat{G}_1$ and $\hat{G}_2$ are estimated as:

$$\hat{G}_1 = \frac{1}{T\lambda} \sum_{t=1}^{T\lambda} \frac{\partial f(y_t, \hat{\theta}_1)}{\partial \theta_1}, \tag{17}$$

$$\hat{G}_2 = \frac{1}{T(1-\lambda)} \sum_{t=T\lambda+1}^{T} \frac{\partial f(y_t, \hat{\theta}_2)}{\partial \theta_2}, \tag{18}$$

and $\hat{W}_1$ and $\hat{W}_2$ are estimated as:

$$\hat{W}_1 = \frac{1}{T\lambda} \sum_{t=1}^{T\lambda} f(y_t, \hat{\theta}_1) f(y_t, \hat{\theta}_1)', \tag{19}$$

$$\hat{W}_2 = \frac{1}{T(1-\lambda)} \sum_{t=T\lambda+1}^{T} f(y_t, \hat{\theta}_2) f(y_t, \hat{\theta}_2)'. \tag{20}$$

The Wald statistic for testing the null hypothesis is then defined as:

$$Wald_T(\lambda) = T(\hat{\theta}_1(\lambda) - \hat{\theta}_2(\lambda))' \left(\frac{\hat{V}_1}{\lambda} + \frac{\hat{V}_2}{1-\lambda} \right)^{-1} (\hat{\theta}_1(\lambda) - \hat{\theta}_2(\lambda)). \tag{21}$$

To test the hypothesis (13), the following test statistic is used:

$$\sup_{\lambda \in \Lambda} Wald_T(\lambda),$$

where Λ is a closed subset of $(0, 1)$. The asymptotic null distribution of the test statistic is given in the following proposition.

Proposition 3. *Under the assumptions of Proposition 2,*

$$Wald_T(\cdot) \Rightarrow Q_P(\cdot) \quad \textit{and} \quad \sup_{\lambda \in \Lambda} Wald_T(\lambda) \xrightarrow{d} \sup_{\lambda \in \Lambda} Q_P(\lambda),$$

where

$$Q_P(\lambda) = \frac{(B(\lambda) - \lambda B(1))'(B(\lambda) - \lambda B(1))}{\lambda(1-\lambda)},$$

and $B(\lambda)$ is a vector of independent Brownian motions on the interval $[0, 1]$.

Proof. Since in the proof of proposition 2 I have shown that my model satisfies assumptions 1-3 in Andrews (1993), proposition 3 is a direct result of Therorem 3 in Andrews (1993). □

Proposition 3 builds upon Andrews (1993) foundational work on structural break testing in econometrics, specifically leveraging the limiting distribution of the sup-Wald statistic to allow for non-parametric structural break detection. The convergence of the sup-Wald statistic to a limiting distribution involving Brownian motions aligns with methods developed in Andrews (1993), where asymptotic results provide a robust basis for inference in structural change contexts without knowing break points a priori. By extending these approaches to the SVAR framework, my model is equipped to handle potentially unknown breaks, which are frequent in economic data subject to shifts in policy or structural dynamics. Furthermore, the link between Proposition 3 and Andrews (1993) reinforces the application of critical values for sup-Wald statistics, which are readily available in Table 1 of Andrews (1993) and provide reference points to maintain statistical rigor in finite samples.

This proposition has particular value for the SVAR model in identifying structural shifts, which could represent changes in macroeconomic policy regimes or shifts in economic structure. Unlike standard break testing approaches, the use of the sup-Wald statistic enables a more flexible framework that detects breaks without pre-specifying breakpoints, fitting well within the non-Gaussian SVAR framework developed here. By validating the use of sup-Wald critical values, this approach ensures that the test maintains both statistical size and power in large samples, an essential consideration in applied macroeconomic analysis.

Given the finite sample limitations of real-world data, my empirical application restricts the interval to $\Lambda = (0.20, 0.80)$. This decision follows the practice of excluding extreme ends of the sample to prevent distortions in the Wald statistic, as suggested by empirical structural break literature such as Andrews (1993). By drawing on this interval restriction, the analysis achieves a balance between capturing meaningful breaks and avoiding edge effects that could compromise test power. In this context, my application of the sup-Wald statistic with finite sample adjustments further supports accurate inference on structural changes within SVAR systems.

2.5 Estimating the Break Point

In the previous section, I discussed how to test for parameter instability in the Non-Gaussian SVAR model using the sup-Wald statistic. In this section, I focus on the problem of estimating the break point when a structural break is detected.

The procedure for estimating the break date begins with testing for the existence of an unknown structural break over the full sample period $t = 1, \ldots, T$. If the null hypothesis is rejected by the sup-Wald statistic, indicating the presence of a structural break, the break date is then estimated by minimizing the sum of the objective functions across the two subsamples. Specifically, the break point $\hat{k}$ is estimated as:

$$\hat{k} = \arg\min_{k} \left[Q_1(k, \hat{\theta}_1(k)) + Q_2(k, \hat{\theta}_2(k)) \right] \tag{22}$$

where $Q_i(k, \theta_i)$ and the estimators $\hat{\theta}_1(k), \hat{\theta}_2(k)$ are defined as in equations (10) and (11). From Assumption 6, I also obtain an estimate of λ as $\hat{\lambda} = \hat{k}/T$.

I next show that the estimated date obtained by criterion 22 is consistent. First, notice that the identification result in Proposition 1 does not secure global identification of the break date. To that end, I impose the following identification assumption:

Assumption 7 (Unique identification of the break date)**.** *Suppose there exists a constant $M > 0$, and denote $D_M = \{k : |k - k_0| > M\}$. Then,*

$$\min_{k \in D_M} \min_{\theta_1, \theta_2 \in (\Pi \times \mathcal{B}_1)^2} |\frac{1}{k} \sum_{t=1}^{k} E(f(y_t, \theta_1))| + |\frac{1}{T-k} \sum_{t=k+1}^{T} E(f(y_t, \theta_2))| > 0$$

Assumption 7 is essential to ensure the uniqueness of the break point estimation in the model, as defined by Equation (22). This assumption specifies that the true break point k_0 is uniquely identifiable. This is a stronger condition than simply assuming parameter distinctness across regimes ($\theta_{10} \neq \theta_{20}$). In the literature on structural break estimation, assumptions of this nature are critical for consistency in break point estimation, particularly in econometric models where structural parameters are allowed to change at unknown points in time. For instance, Bai and Perron (1998) and Bai and Perron (2003) utilize similar uniqueness conditions in their analysis of multiple structural changes, ensuring that each break point can be consistently estimated based on distinct parameter shifts in the model. Likewise, Qu and Perron (2007) extend this approach to multivariate settings, showing that such conditions are necessary to accurately pinpoint structural changes even when multiple parameters exhibit shifts concurrently.

The assumption made here contributes to the growing literature by reinforcing the stability of the estimated break point $\hat{k}$ against potential local minima in the objective function. Without this condition, the model might yield multiple candidate break points, complicating inference on the true location of the structural change.

The following proposition presents the asymptotic properties of the estimated break point and the parameter estimates in each regime.

Proposition 4. *Suppose the assumptions of Proposition 2 hold, except that $\theta_{10} \neq \theta_{20}$. In addition, Assumption 7 holds. Then:*

1. *The estimators $\hat{\theta}_1(\hat{k})$ and $\hat{\theta}_2(\hat{k})$ converge to their true values at the standard $\sqrt{T}$ rate:*

$$\sqrt{T}\left(\hat{\theta}_1(\hat{k}) - \theta_{10}\right) \xrightarrow{d} N(0, G'_{\theta_{10}} W^{-1}_{\theta_{10}} G_{\theta_{10}}), \tag{23}$$

$$\sqrt{T}\left(\hat{\theta}_2(\hat{k}) - \theta_{20}\right) \xrightarrow{d} N(0, G'_{\theta_{20}} W^{-1}_{\theta_{20}} G_{\theta_{20}}), \tag{24}$$

where $G_{\theta_{10}}, G_{\theta_{20}}$ are the gradients of the moment conditions and $W_{\theta_{10}}, W_{\theta_{20}}$ are the corresponding weighting matrices.

2.

$$\hat{k} = k_0 + O_p(1),$$

where $O_p(1)$ denotes the stochastic boundness. Equivlently, the estimated break point fraction $\hat{\lambda}$ is T-consistent for the true break point fraction λ_0.

Proof. See Appendix C. □

The above proposition outlines two essential results on the behavior of parameter estimates and the estimated break point in the presence of structural breaks within a Non-Gaussian SVAR model framework. First, the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ for each subsample converge at the standard $\sqrt{T}$ rate, even when a structural break exists. This consistency and asymptotic normality imply that structural breaks do not compromise the asymptotic properties of the parameter estimates, preserving the reliability of inference. These properties enable confidence interval construction for impulse responses via the delta method, facilitating inference across different regimes.

Second, the proposition establishes the consistency of the estimated break point $\hat{k}$, where $\hat{k} = k_0 + O_p(1)$, indicating that the estimation error remains bounded. Equivalently, the break fraction $\hat{\lambda} = \hat{k}/T$ converges to the true fraction λ_0 as the sample size increases, allowing precise break date estimation in larger samples.

These results align with prior foundational work on break point estimation and asymptotic properties, particularly by Bai and Perron (1998) and Bai (2000) in the structural break context. Specifically, Bai and Perron (1998) confirmed convergence rates and asymptotic distributions for break points in multivariate regression models, while Bai (2000) extended these results to vector autoregressive models under Gaussian assumptions. This proposition advances their findings by adapting these asymptotic properties within the Non-Gaussian SVAR framework, incorporating the dynamics of both autoregressive coefficients and the structural matrix B, where B is identified through non-Gaussian characteristics.

This extension is significant, offering a robust, consistent estimator for endogenous break dates that applies to complex economic systems in which structural shifts often correspond with macroeconomic events.

3 Model with multiple Break Points

In this section, I extend the structural vector autoregressive (SVAR) model framework to account for multiple structural break points and establish the corresponding theoretical results. Specifically, I consider a structural vector autoregressive model of order p with m break points. The model is formulated as follows:

$$y_t = \begin{cases} \nu_1 + A_{1,1}y_{t-1} + \cdots + A_{1,p}y_{t-p} + B_1\varepsilon_t, & \text{for } 1 \leq t \leq k_1, \\ \nu_2 + A_{2,1}y_{t-1} + \cdots + A_{2,p}y_{t-p} + B_2\varepsilon_t, & \text{for } k_1 + 1 \leq t \leq k_2, \\ \vdots & \\ \nu_m + A_{m,1}y_{t-1} + \cdots + A_{m,p}y_{t-p} + B_m\varepsilon_t, & \text{for } k_{m-1} + 1 \leq t \leq T, \end{cases} \tag{25}$$

where ν_i are $n \times 1$ intercept terms, and $A_{i,j}$ are $n \times n$ parameter matrices for $i = 1, \ldots, m$ and $j = 1, \ldots, p$, representing the time-varying dynamics across different regimes. The structural errors ε_t are assumed to be an i.i.d. sequence with $E(\varepsilon_t \varepsilon_t') = I_n$. The process y_t is stable within each regime, as guaranteed by the following stationarity condition:

$$\det A_i(z) \stackrel{\text{def}}{=} \det\left(I_n - A_{i,1}z - \cdots - A_{i,p}z^p\right) \neq 0 \quad \text{for } |z| \leq 1, \quad i = 1, \ldots, m, \tag{26}$$

3.1 Sequential procedure

In practical applications, the number of break points, m, is typically unknown. To determine the number of break points and estimate their locations, I employ the sequential break-point estimation procedure originally proposed by Bai (1997). This procedure is particularly useful in macroeconomic time series analysis where structural changes in the economy often occur at unknown points in time. The procedure follows the following steps: First, the initial break point is estimated using the formula (22). Once the first break is identified, the sample is split into two subsamples: the first subsample containing observations up to the estimated break point k_1, and the second subsample containing the remaining observations. The parameter constancy in each subsample is tested using the sup-Wald statistic. If the test rejects the null hypothesis of constancy in any subsample, a break point is estimated using the formula (22) within that subsample. This recursive process continues, splitting the subsamples further whenever parameter constancy is rejected. The process terminates when the constancy test is not rejected for all subsamples.

By following this hierarchical approach, the number of break points is determined, and their locations are estimated. The final number of estimated break points is denoted as $\hat{m}$, where $\hat{m} = \text{number of subsamples} - 1$, and this should ideally converge to the true number of break points m_0.

Next, I discuss the theoretical foundation for the consistency and asymptotic properties of this sequential procedure. Before presenting the key results, I redefine Assumptions (5) to (6) to account for the presence of multiple break points.

Assumption 8. *In each subsample, the weighting matrix $W_i(k_i) = E(f(y_t, \theta_{i0})f(y_t, \theta_{i0})')$ is positive definite, and $E\|f(y_t, \theta_{i0})\|^{2+\delta_i} < \infty$ for some $\delta_i > 0$.*

Assumption 9. *Each break point is denoted as $k_i = T\lambda_i$, where $\lambda_i \in (0, 1)$, representing the normalized location of the break in terms of sample size T.*

Given these assumptions, I now demonstrate the consistency of the estimated break points. Even in the presence of multiple breaks, the estimation procedure is consistent, as shown in the following proposition:

Proposition 5. *For the model in equation (25), suppose the structural errors satisfy one set of Assumptions 1, 2, or 3, corresponding to the moment function $f(y_t, \theta)$. Further, suppose Assumptions 4, 8, and 9 hold, and k_{i0} satisfies the identification assumption 7. Then, the estimated break points $\hat{k}_i$, obtained using equation (22), are asymptotically bounded, i.e.,*

$$\hat{k}_i = k_{i0} + O_p(1),$$

Equivalently, the estimated break fractions $\hat{\lambda}_i$ are T-consistent.

The proof follows directly from the proof of proposition 4 in Appendix C.

The consistency of the estimated break points $\hat{k}_i$ is a critical result, as it ensures that the estimated locations of the break points will converge to the true break points k_{i0} as the sample size grows. Moreover, this result holds even when the estimation procedure is applied within subsamples, meaning that the sequential break estimation procedure remains valid in the presence of multiple structural breaks.

I now turn to the consistency of the number of break points $\hat{m}$ determined by the sequential procedure. Specifically, I show that the estimated number of breaks converges to the true number of breaks m_0, as formalized in the following proposition:

Proposition 6. *Suppose that the size of the test α_T converges to zero at the rate T^{-1}. Then, under the assumptions of Proposition 5,*

$$P(\hat{m} = m_0) \to 1 \quad \text{as } T \to \infty.$$

Proof. See Appendix D. □

This proposition ensures that the sequential procedure correctly identifies the true number of breaks with probability approaching one as the sample size increases. This result is crucial for the reliability of the procedure in practical applications, where the number of structural breaks is often unknown.

Finally, I state the following corollary, which guarantees that the limiting distribution of the parameter estimates is not affected by the break-point estimation process:

Corollary 9.1. *Under the assumptions of Proposition 6, the limiting distribution of the estimated parameters $\hat{\theta}_i$ is the same as the distribution that would be obtained if the true break points $k_{10}, \ldots, k_{m0}$ were known.*

Proof. The proof follows directly from proof of proposition 4 in Appendix C. □

This corollary highlights the robustness of the sequential procedure, ensuring that the estimation of break points does not distort the asymptotic properties of the parameter estimates and allows for valid inference even in the presence of multiple structural breaks.

4 Statistical Inference

In this section, I present the inference methods for the SVAR model with multiple structural breaks. Specifically, I focus on testing hypotheses related to the estimated parameters in each regime and analyzing impulse responses across different break periods. Given the presence of multiple structural breaks and non-Gaussian errors, the inference procedure needs to account for the variations in the model's dynamics across different subsamples.

4.1 Parameter Tests

To perform valid statistical inference in the presence of structural breaks, I extend the GMM framework discussed in the previous subsection to allow for regime-specific parameter estimation. Let $\hat{\theta}_i$ denote the GMM estimator of the parameter vector θ_{i0} in the i-th regime, $i = 1, \ldots, m$.

From the asymptotic results established in Corollary 9.1 and Proposition 5, we know that the GMM estimator $\hat{\theta}_i$ is consistent for θ_{i0}, the true parameter vector in the i-th regime. This consistency holds in the presence of multiple breaks, and the estimator converges at rate T under standard regularity conditions, as discussed earlier. Furthermore, $\hat{\theta}_i$ is asymptotically normally distributed with a regime-specific covariance matrix:

$$G'_{i0} W_{\theta_{i0}}^{-1} G_{i0},$$

where $G_{i0} = E\left(\frac{\partial f(y_t, \theta_{i0})}{\partial \theta}\right)$ is the Jacobian matrix of the moment conditions evaluated at θ_{i0}, and $W_{\theta_{i0}} = E[f(y_t, \theta_{i0}) f(y_t, \theta_{i0})']$ is the covariance of the moment conditions.

Given the results from Corollary 9.1, hypotheses about the parameters in each regime can be tested using the Wald statistic. Suppose I want to test a set of restrictions on the parameters of the form $H_0 : r(\theta_{i0}) = 0$, where $r(\cdot)$ is a differentiable vector-valued function of the parameters. The Wald test statistic for testing these restrictions is given by:

$$Wald_T^{(i)} = Tr(\hat{\theta}_i)' \left[R\left(\hat{\theta}_i\right) \left(G_{iT} W_{\theta_{iT}} G'_{iT}\right)^{-1} R\left(\hat{\theta}_i\right)' \right]^{-1} r(\hat{\theta}_i),$$

where $R(\theta) = \frac{\partial r(\theta)}{\partial \theta}$ is the Jacobian of the restrictions, and G_{iT} and $W_{\theta_{iT}}$ are consistent estimators of the matrices G_{i0} and $W_{\theta_{i0}}$, respectively. Under the null hypothesis, the Wald statistic follows an asymptotic chi-squared distribution with degrees of freedom equal to the number of restrictions s.

The asymptotic properties of these estimators and tests are directly inherited from the consistency and limiting normality results of Corollary 9.1 and Proposition 5, which establish that the estimators $\hat{k}_i$ and $\hat{\theta}_i$ are T-consistent and normally distributed under the true model with multiple break points.

4.2 Impulse Response Analysis

Impulse response functions (IRFs) are widely used to study the dynamic effects of structural shocks on macroeconomic variables. In the presence of multiple structural breaks, as modeled in this paper, the IRFs must be computed separately for each regime, reflecting the different dynamics in each period.

For each regime i, the reduced-form representation of the SVAR model yields a moving average (MA) process for the endogenous variables y_t:

$$y_t = \mu_i + \sum_{k=0}^{\infty} C_{i,k} B_i \varepsilon_{t-k}, \tag{27}$$

where μ_i is the unconditional expectation of y_t in regime i, and $C_{i,k}$ are the moving average coefficients in regime i, obtained recursively as

$$C_{i,k} = \sum_{l=1}^{k} C_{i,k-l} A_{i,l},$$

where $A_{i,l} = 0$ for $l > p$. The matrix B_i represents the contemporaneous impact matrix of structural shocks in regime i, and ε_t is the vector of structural shocks.

The j-th column of $C_{i,k} B_i$ contains the impulse responses of the system to the structural shock $\varepsilon_{j,t}$, where the (l, j)-th element of this matrix represents the response of the l-th variable in the system to a unit shock in $\varepsilon_{j,t}$ at horizon k. Specifically, the impulse response of variable $y_{l,t+k}$ to shock $\varepsilon_{j,t}$ is given by:

$$\frac{\partial y_{l,t+k}}{\partial \varepsilon_{j,t}} = \iota_l' C_{i,k} B_i \iota_j,$$

where ι_l and ι_j are unit vectors. I denote this impulse response coefficient by $\psi_{i,k,l,j}(\pi_i, \vartheta_i)$, where π_i is the vector of VAR parameters, and ϑ_i contains the elements of the impact matrix B_i in regime i.

To estimate the impulse response coefficients, I replace the true parameter vectors π_i and ϑ_i with their consistent estimators $\hat{\pi}_i$ and $\hat{\vartheta}_i$, obtained from the SVAR model estimated in each regime. The estimated IRF is denoted by $\hat{\psi}_{i,k,l,j}(\hat{\pi}_i, \hat{\vartheta}_i)$.

4.2.1 Asymptotic Properties of the IRFs

To perform inference on the IRFs, I derive their asymptotic distribution. Following the results of Corollary 9.1 and Proposition 5, the estimators $\hat{\pi}_i$ and $\hat{\vartheta}_i$ are asymptotically normally distributed. Using the delta method, I obtain the asymptotic distribution of the estimated impulse responses. Specifically:

$$T^{1/2}\left[\hat{\psi}_{i,k,l,j}(\hat{\pi}_i, \hat{\vartheta}_i) - \psi_{i,k,l,j}(\pi_i, \vartheta_i)\right] \xrightarrow{d} N(0, \sigma^2_{i,k,l,j}),$$

where the asymptotic variance $\sigma^2_{i,k,l,j}$ is given by:

$$\sigma^2_{i,k,l,j} = \frac{\partial \psi_{i,k,l,j}(\pi_i, \vartheta_i)}{\partial(\pi_i', \vartheta_i')} G_{i0}' W_{\theta_{i0}}^{-1} G_{i0} \frac{\partial \psi_{i,k,l,j}(\pi_i, \vartheta_i)}{\partial(\pi_i', \vartheta_i')'},$$

The asymptotic confidence intervals for the IRFs are computed using this variance-covariance structure. Alternatively, a bootstrap procedure can be employed to construct confidence intervals for the IRFs, as it may provide more accurate coverage in small samples.

4.3 Testing Impulse Responses Across Regimes

To formally assess whether the impulse responses differ significantly across regimes, I test the null hypothesis that the impulse response functions are identical across break periods. Specifically, I test:

$$H_0 : \psi_{i,k,l,j} = \psi_{i',k,l,j} \quad \text{for all } i, i' \in \{1, \dots, m\},\ l, j \in \{1, \dots, n\},\ k \geq 0.$$

This hypothesis can be tested using Wald-type tests or through bootstrapped confidence intervals that compare the impulse responses across different regimes. A rejection of the null hypothesis would indicate that the structural breaks significantly affect the transmission mechanism of shocks, suggesting that the responses of variables to the shocks vary across regimes.

The tests are conducted within the GMM framework, ensuring that the asymptotic properties of the estimators, as established in Proposition 5 and Corollary 9.1, hold. This approach provides a robust framework for understanding the impact of structural breaks on impulse response dynamics.

5 Empirical Illustration

I estimate a SVAR model with break points on monthly U.S. macroeconomic data spanning the period from 1954:VII to 2024:I, comprising 835 observations. The model consists of three variables $y_t = (\pi_t, u_t, r_t)'$, where π_t denotes inflation, u_t the unemployment gap, and r_t the federal funds rate.

During the period from 2009 to 2023, when the federal funds rate was constrained by the zero lower bound, it is replaced by the "shadow rate" of Wu and Xia (2023) to account for unconventional monetary policy. The data for inflation and the unemployment gap are extracted from the Federal Reserve Economic Database (FRED). Inflation (π_t) is computed as the logarithmic difference of the Consumer Price Index for All Urban Consumers (mnemonic CPIAUCSL), multiplied by 1200 to express it in annualized percentage terms. The unemployment gap (u_t) is calculated as the difference between the observed unemployment rate (mnemonic UNRATE) and the natural rate of unemployment (mnemonic NROU). The natural rate of unemployment is originally reported as quarterly data, which is linearly interpolated to match the monthly frequency of the other variables in the dataset.

To determine the appropriate lag order for the SVAR model, I conduct a lag selection based on the reduced-form VAR using the full sample. The Akaike Information Criterion (AIC) suggests a model with 5 lags, while the more parsimonious Schwarz Information Criterion (SIC) recommends 2 lags. Given that the 5-lag model includes a substantial number of parameters, there is a risk of overfitting, particularly when structural breaks reduce the number of observations in each subsample. Consequently, I opt for a SVAR(2) model, as it balances parsimony with statistical adequacy, providing a more reliable fit without overfitting.

I estimate the SVAR-BP model using three sets of moment functions, corresponding to $f_s(y_t, \theta)$, $f_k(y_t, \theta)$, and $f_r(y_t, \theta)$. These estimation methods are referred to as GMM1, GMM2, and GMM3, respectively. For each set of moment functions, I apply the sequential procedure outlined in Section 3.1 to identify and estimate both the break dates and the model parameters. To determine the optimal set of moment conditions, I evaluate whether the generic assumptions hold by computing the sample skewness and kurtosis from the estimated structural errors. Let $\hat{\epsilon}_s$, $\hat{\epsilon}_k$, and $\hat{\epsilon}_r$ denote the structural errors estimated using GMM1, GMM2, and GMM3, respectively. The QQ-plots in Figure 1 suggest that all three sets of structural errors exhibit non-Gaussianity, regardless of the estimation method used. I then proceed to check the generic conditions for each set of moment functions. GMM1 is not prefered due to a failure to meet its generic assumption that at most one structural error has zero skewness. The estimated skewness values for the structural errors under GMM1 are 0.04, 3.38, and -0.70. These values indicate that one component has nearly zero skewness (0.04) and another is only weakly skewed (-0.70). Such low skewness levels undermine the robustness of the third-moment conditions required for GMM1, as they suggest a potential weak identification issue where the structural parameters may not be reliably estimated due to insufficient asymmetry in the error distribution (Stock and Wright, 2000). GMM2 is also not prefered, despite satisfying the generic condition, because the value of the J-statistic—a measure of overidentification, was significantly higher compared to GMM3, indicating a poorer fit of the model to the data. GMM3, which uses the reflection-invariant moment conditions, is ultimately prefered. This set of moment functions satisfies the generic conditions, and the J-statistic indicates a better overall fit, making GMM3 the optimal choice for estimating the SVAR-BP model in this context.

Using the GMM3 estimation method and the sequential procedure described in Section 3, I identify two significant break points in the model: March 1981 and November 2008. These dates correspond to crucial events in U.S. economic history and align closely with structural shifts noted in previous studies, underscoring the importance of accounting for these transitions within macroeconomic models.

The first break point, March 1981, aligns with a major shift in monetary policy under Federal Reserve Chairman Paul Volcker, as the Fed aggressively raised interest rates to control the high inflation that characterized the 1970s. This transition to a restrictive monetary policy regime had substantial economic effects, contributing to a significant recession and, ultimately, to a longer-term stabilization of inflation. This finding is consistent with earlier studies such as Bernanke and Mihov (1998), McConnell and Perez-Quiros (2000),Stock and

Watson (2002), Sims and Zha (2006), Smets and Wouters (2007), Justiniano and Primiceri (2008), Canova et al. (2008), and Canova (2009).

The second break point, November 2008, aligns with the peak of the global financial crisis following the collapse of Lehman Brothers in September 2008. This period was marked by severe disruptions in financial markets, leading the Federal Reserve to implement unconventional monetary policies, including near-zero interest rates and quantitative easing, to stabilize the economy. This break date is consistent with studies such as Baker et al. (2016) and Stock and Watson (2012).

Both identified break points underscore pivotal shifts in U.S. economic policy, highlighting the necessity of modeling structural breaks to capture the economy's evolving dynamics accurately. The consistency of these findings with established literature further validates the model's capacity to endogenously identify significant economic regime changes, enhancing our understanding of how structural breaks impact macroeconomic relationships over time.

The GMM3 estimate of the matrix of impact effects in each period is (asymptotic standard deviation in parentheses):

$$\hat{B}_1 = \begin{bmatrix} \underset{(0.162)}{2.470} & \underset{(0.141)}{-0.084} & \underset{(0.128)}{0.591} \\ \underset{(0.008)}{0.011} & \underset{(0.013)}{0.235} & \underset{(0.013)}{-0.068} \\ \underset{(0.035)}{-0.033} & \underset{(0.037)}{-0.009} & \underset{(0.116)}{0.988} \end{bmatrix} \quad \hat{B}_2 = \begin{bmatrix} \underset{(0.219)}{2.710} & \underset{(0.177)}{0.292} & \underset{(0.046)}{-0.155} \\ \underset{(0.012)}{-0.039} & \underset{(0.007)}{0.153} & \underset{(0.011)}{0.017} \\ \underset{(0.006)}{0.030} & \underset{(0.015)}{-0.065} & \underset{(0.028)}{0.278} \end{bmatrix} \tag{28}$$

$$\hat{B}_3 = \begin{bmatrix} \underset{(0.157)}{2.695} & \underset{(0.047)}{-0.647} & \underset{(0.200)}{-0.131} \\ \underset{(0.019)}{-0.036} & \underset{(0.091)}{0.783} & \underset{(0.009)}{0.006} \\ \underset{(0.012)}{0.018} & \underset{(0.004)}{0.008} & \underset{(0.015)}{0.186} \end{bmatrix} \tag{29}$$

The GMM3 estimation yields distinct matrices of impact effects, $\hat{B}_1$, $\hat{B}_2$, and $\hat{B}_3$, each corresponding to different subsample periods based on the identified structural breaks.

5.1 Interpretation of Impulse Response Functions

The impulse responses of inflation, the unemployment gap, and the federal funds rate(or shadow rate) to the different shocks are displayed from figure 2 to 4 for each period. The analysis below focuses on the third shock in each period, which shows a statistically significant impact on the federal funds rate and can be labeled as the monetary policy shock. This labeling is consistent with studies such as Bernanke and Blinder (1992) and Christiano et al. (1999), who identify monetary policy shocks through significant responses in interest rates.

5.1.1 1954-07-01 to 1981-03-01

In the first period (1954-1981), the impulse response of inflation to a monetary policy shock is initially positive, exhibiting the so-called "price puzzle," where inflation temporarily rises following a contractionary policy shock. This phenomenon has been noted in previous literature, including Sims (1992) and Eichenbaum (1995), suggesting that the price puzzle might be due to model simplicity or the delayed effect of monetary policy transmission. Over subsequent periods, inflation begins to decline, aligning with the expected outcome of a contractionary policy aimed at reducing inflation in the medium term. The unemployment gap initially decreases, indicating a short-term drop in unemployment, but turns positive after a few periods, suggesting that the restrictive policy ultimately leads to an economic slowdown and rising unemployment. The significant positive response of the federal funds rate on impact corroborates the labeling of this third shock as a monetary policy shock, consistent with the response patterns seen in historical monetary tightening cycles, such as those analyzed by Bernanke and Blinder (1992).

5.1.2 1981-04-01 to 2008-11-01

During the second period (1981-2008), the impact of the monetary policy shock is more subdued but still consistent with a controlled inflationary environment, likely due to the disinflationary policies implemented post-Volcker. The initial response of inflation to a monetary policy shock is positive but weaker than in the earlier period, with inflation declining more quickly, reflecting the efficacy of monetary policy in this era, which reduced the price puzzle effect. Unemployment initially decreases slightly in response to the shock, but similar to the first period, it ultimately increases, reflecting the medium-term contractionary effects of monetary policy on economic activity. The federal funds rate exhibits a strong, significant response on impact, further supporting the identification of this third shock as a monetary policy shock. This aligns with findings from Christiano et al. (1999) and Smets and Wouters (2007), who observed effective monetary policy transmission in controlling inflation during this period.

5.1.3 2008-11-01 to 2024-01-01

In the final period (2008-2024), which covers the post-financial crisis era, the monetary policy shock exhibits unique characteristics due to unconventional monetary policies at the zero lower bound. Following a contractionary policy shock, the shadow rate—a proxy used to capture the effective stance of monetary policy when the federal funds rate is near zero—responds positively, indicating a tightening of monetary conditions. This response aligns with the expected behavior of interest rates under contractionary policy, as also observed by Wu and Xia (2016), who highlight the effectiveness of the shadow rate in representing policy shifts under these constrained conditions. The inflation response is initially close to zero, reflecting the subdued inflationary environment of the period, even amid monetary tightening. Over time, inflation shows a delayed decline, consistent with the extended period of low inflation following the financial crisis. The unemployment gap initially decreases, suggesting an initial stimulative effect on the labor market, but turns positive in the medium term, indicating that the contractionary effects begin to take hold as the economy adjusts.

5.1.4 Overall Implications

Across all periods, the monetary policy shock has generally led to a medium-term negative impact on inflation and economic activity, with varying intensity and response patterns due to changing monetary frameworks. In the early period, more volatile responses are observed, consistent with the less structured monetary environment pre-Volcker. Post-1981, responses become more controlled, aligning with disinflationary policies. The final period shows a distinct response pattern due to unconventional policy measures, reflecting the challenges of monetary policy in the post-crisis era.

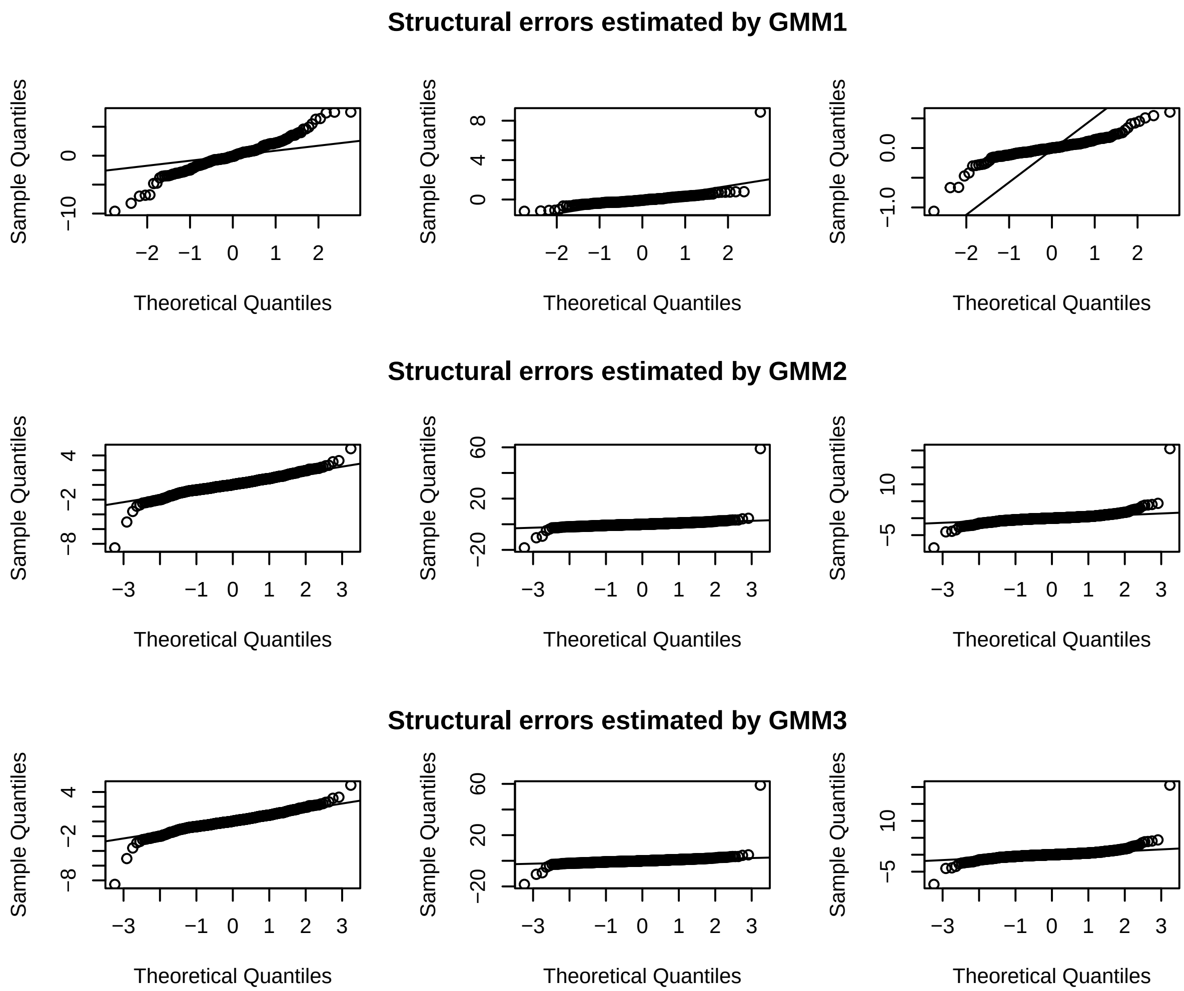

Figure 1: Quantile-quantile plots of the structural errors of the equations of inflation, unemployment gap and the federal funds rate (from left to right), estimated by 3 sets of moment functions

Figure 2: Impulse responses from 1954-07-01 to 1981-03-01. Each row contains the impulse responses of one shock on all variables. The shaded areas are the pointwise 95% confidence bands obtained by delta method.

Figure 3: Impulse responses from 1981-04-01 to 2008-11-01. Each row contains the impulse responses of one shock on all variables. The shaded areas are the pointwise 95% confidence bands obtained by delta method.

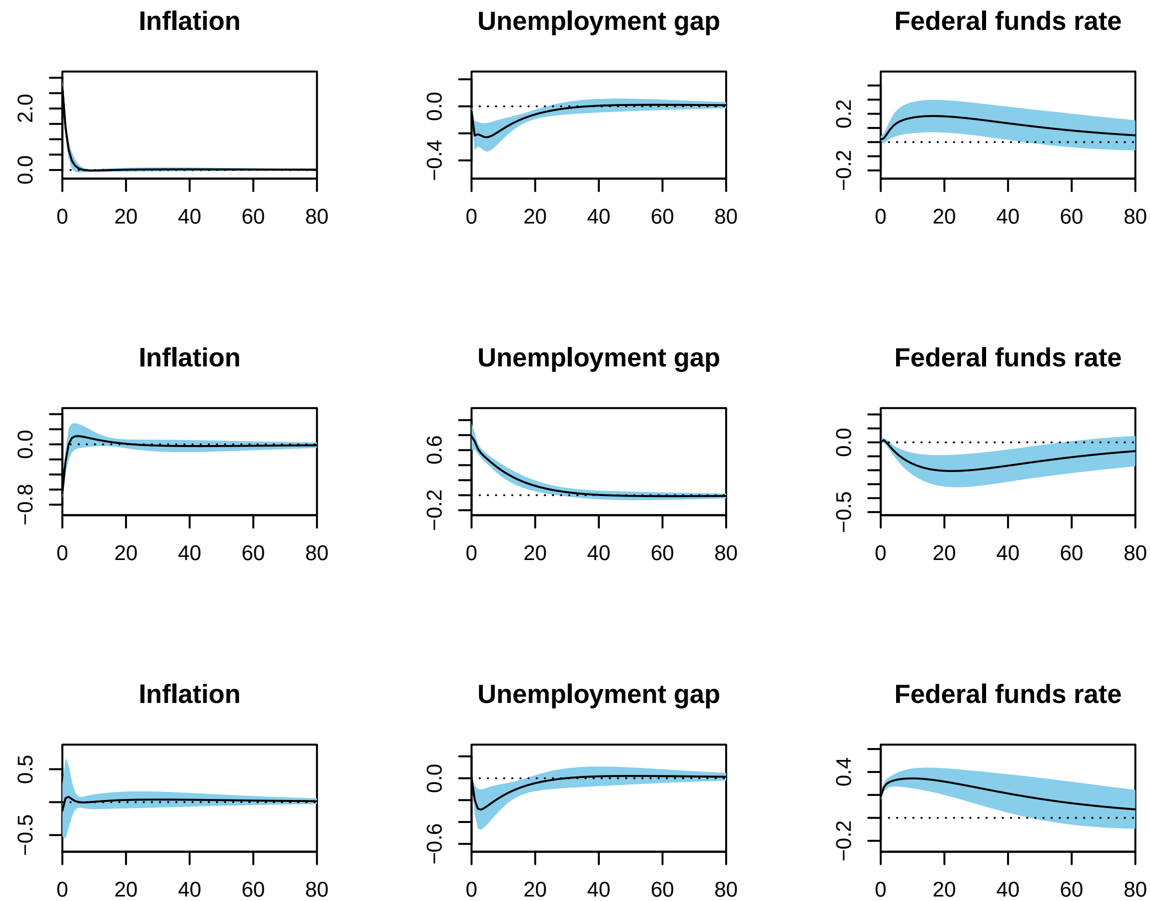

Figure 4: Impulse responses from 2008-11-01 to 2024-01-01. Each row contains the impulse responses of one shock on all variables. The shaded areas are the pointwise 95% confidence bands obtained by delta method.

6 Conclusion

This paper introduces a novel approach to estimating structural breaks within the framework of a non-Gaussian Structural Vector Autoregression (SVAR) model, allowing for endogenous identification of break points in both autoregressive coefficients and structural parameters. Using the Partial Sample Generalized Method of Moments (PSGMM) estimator and leveraging non-Gaussian error structures, the model identifies structural shifts through higher-order moments. This approach addresses the limitations of traditional SVAR models by capturing the potential instability of macroeconomic relationships over time, especially in response to significant economic events.

The findings also contribute to the growing literature on structural break analysis within time series models, validating the model's ability to identify shifts endogenously and underscoring the relevance of non-Gaussian features for identifying structural matrices. This research expands the methodological toolkit for analyzing evolving macroeconomic relationships and provides a robust framework for future studies exploring the implications of structural changes in economic systems.

Empirical analysis on U.S. macroeconomic data from 1954 to 2024 highlights the efficacy of this method, identifying two key structural breaks in March 1981 and November 2008. These dates correspond closely with historical shifts in U.S. monetary policy: the 1981 transition under Chairman Volcker's aggressive anti-inflationary measures, and the 2008 peak of the global financial crisis that triggered unconventional monetary policies. These breaks underscore the critical role of structural shifts in macroeconomic modeling and demonstrate how the proposed SVAR-BP model adapts to capture these pivotal changes in economic dynamics.

A notable limitation of the current framework is the potential for weak identification when non-Gaussian characteristics in structural errors do not sufficiently distinguish between economic shocks. While higher-order moment conditions can provide valuable identifying information, weak skewness or kurtosis may limit the robustness of inference in some applications. This issue aligns with findings by Hoesch et al. (2024) and Lee and Mesters (2024), who propose robust inference techniques within standard SVAR settings but have yet to extend these approaches to frameworks involving structural breaks. Developing robust methods for SVAR models with break points remains an important area for future research, particularly in light of the complex economic contexts where structural shifts are present.

Another promising direction for addressing weak identification lies in integrating multiple sources of identifying information to strengthen the model's inferential power. Recent studies have explored the combined use of sign restrictions and non-Gaussianity to create more precise identification frameworks. For instance, Drautzburg and Wright (2023) propose an intersection-based approach, combining sign restrictions with non-Gaussian characteristics to construct an identified set that reflects both the desired economic sign restrictions and the independence properties of the shocks. Future research might also explore methodologies like those proposed by Keweloh et al. (2023) and Keweloh (2024), which incorporate proxy variables or short-run restrictions within a non-Gaussian context. Additionally, approaches that leverage external instruments to improve shock labeling, such as Crucil et al. (2023), may offer valuable insights for advancing robust identification in models with break points.

Appendix A Proof of proposition 1

In this appendix, I provide a formal proof of Proposition 1, demonstrating the identification of the parameters in the structural VAR (SVAR) model under the stated assumptions.

Let $u_t = B_i \varepsilon_t$. The moment conditions are given by:

$$E(u_t \otimes Z_{t-1}) = E(B\varepsilon_t \otimes Z_{t-1}) = (B \otimes I_{np+1})E(\varepsilon_t \otimes Z_{t-1}) = 0_{n(np+1)\times 1} \tag{A.1}$$

The last equality holds due to the condition in equation 3, which is satisfied by the i.i.d properties of the errors ε_t.

Step 1: Identification of π_i

I first show that $\pi_i = \pi_{i0}$. Define $C_i = (\nu_i, A_{i,1}, \ldots, A_{i,p})$ and express the residual u_t as:

$$u_t = y_t - C_i Z_{t-1}$$

Substituting this into equation (A.1), I have:

$$E(u_t Z'_{t-1}) = E((y_t - CZ_{t-1})Z'_{t-1}) = 0$$

Rearranging terms, I obtain:

$$C_i E(Z_{t-1}Z'_{t-1}) = E(y_t Z'_{t-1})$$

Now, suppose both C_i and $C_{i,0}$ satisfy the equation above. This leads to:

$$(C_i - C_{i,0})E(Z_{t-1}Z'_{t-1}) = 0$$

By the theorem of Mann and Wald (1943), $E(Z_{t-1}Z'_{t-1})$ is invertible if y_t follows a stable, stationary VAR(p) process, and ε_t are i.i.d with finite fmyth moments. Therefore, $C_i = C_{i,0}$, and since $\pi_i = \text{vec}(C_i)$, it follows that $\pi_i = \pi_{i,0}$.

Step 2: Identification of B_i

Having shown that $\pi_i = \text{vec}(C_i)$, I now turn to the identification of the matrix B_i. For the process $u_t = B_i \varepsilon_t$, I need to demonstrate that $B_i = B_{i0} Q$, where Q is a signed permutation matrix.

Using the conditions of the proposition, I apply the results of Mesters and Zwiernik (2022):

- Under Assumptions 1 or 2, Theorem 5.3 in Mesters and Zwiernik (2022) shows that $B_i = B_{i0} Q$.
- Under Assumption 3, Theorem 5.10 in Mesters and Zwiernik (2022) similarly ensures that $B_i = B_{i0} Q$.

Thus, under the given assumptions, the matrix B_i is identified up to a signed permutation matrix Q, completing the proof.

Appendix B Proof of Proposition 2

In this appendix, I provide the proof of Proposition 2. The proof is based on Theorems 1 and 3 in Andrews (1993). I verify the assumptions of Theorem 1 one by one. First, note that my PSGMM estimator is a special case of the estimator defined in Andrews (1993) with the optimal choice of the asymptotic weighting matrix. Therefore, Assumptions 2 and 3 from Andrews (1993) are satisfied. For the remaining proof, I demonstrate that the series y_t and moment conditions are near epoch dependent (NED), that the moment functions are continuous and differentiable, and that the gradient of the moment conditions is bounded and uniformly consistent for $\hat{\theta}$.

NED of y_t

I begin by considering the moving average (MA) representation of y_t:

$$y_t = \mu + \sum_{k=0}^{\infty} \Phi_k B \varepsilon_{t-k}, \tag{B.1}$$

where μ is the unconditional expectation of y_t, Φ_0 is the identity matrix, and $\Phi_k, k = 1, 2, \ldots,$ are obtained recursively as $\Phi_k = \sum_{l=1}^{k} \Phi_{k-l} A_l$, where $A_k = 0$ for $k > p$. Set $\mathcal{F}_{t-m}^{t+m} = \sigma(\varepsilon_{t-m}, \ldots, \varepsilon_{t+m})$ and define $\hat{y}_{m,t} = E(y_t|\mathcal{F}_{t-m}^{t+m})$. For all t, I have:

$$E\|y_t - \hat{y}_{m,t}\| = E\| \sum_{j=m+1}^{\infty} \Phi_j B(\varepsilon_{t-j} - E(\varepsilon_{t-j}|\mathcal{F}_{t-m}^{t+m}))\| \tag{B.2}$$

where $\|\cdot\|$ is the standard L^2 Euclidean norm. For the second term on the right-hand side of (B.2), I have:

$$E\|E(\varepsilon_t|\mathcal{F}_{t-m}^{t+m})\| \leq E\|E(\varepsilon_t|\mathcal{F}_{t-m}^{t+m})\|_1 = \sum_{i=1}^{n} E\left|E(\varepsilon_{t,i}|\mathcal{F}_{t-m}^{t+m})\right| \leq nE\|\varepsilon_t\|. \tag{B.3}$$

Combining (B.2) and (B.3), I obtain:

$$E\|y_t - \hat{y}_{m,t}\| \leq (n+1)BE\|\varepsilon_t\| \sum_{j=m+1}^{\infty} \|\Phi_j\|, \tag{B.4}$$

where, because ε_t has mean zero and the identity covariance matrix, $E\|\varepsilon_t\| = n$. The right-hand side of Equation (B.4) decays exponentially fast to zero as $m \to \infty$, since the elements in the coefficient matrices Φ_j are bounded by an exponentially declining series. Thus, I have shown that y_t is an L^2-NED sequence on $\{\varepsilon_t\}$. Since convergence in mean square implies convergence in probability by Markov's inequality, y_t is also an L^2-NED sequence on $\{\varepsilon_t\}$.

NED of Moment Functions

Next, I demonstrate that the moment functions $f(y_t, \theta)$ are also L^2-NED sequences on $\{\varepsilon_t\}$. I denote the vector of moment conditions as:

$$f(y_t, \theta) = \begin{bmatrix} \varepsilon_t \otimes Z_{t-1} \\ h_2(y_t, \theta) \\ h_i(y_t, \theta) \end{bmatrix},$$

where $h_2(y_t, \theta)$ represents the moments in (4), and $h_i(y_t, \theta)$ is one of $h_3(y_t, \theta)$, $h_4(y_t, \theta)$, or $h_r(y_t, \theta)$ from (5), (6), and (7), respectively.

Define the truncated version $\hat{f}_m(y_t, \theta) = E\left(f(y_t, \theta) \middle| \mathcal{F}_{t-m}^{t+m}\right)$. Then I have:

$$\begin{aligned} E\|f(y_t, \theta) - \hat{f}(y_t, \theta)\| = {} & E\|\varepsilon_t \otimes Z_{t-1} - E(\varepsilon_t \otimes Z_{t-1} \mid \mathcal{F}_{t-m}^{t+m})\| \\ & + E\|h_2(y_t, \theta) - E(h_2(y_t, \theta) \mid \mathcal{F}_{t-m}^{t+m})\| \\ & + E\|h_i(y_t, \theta) - E(h_i(y_t, \theta) \mid \mathcal{F}_{t-m}^{t+m})\|. \end{aligned} \tag{B.5}$$

Since ε_t is i.i.d., the second and third terms on the right-hand side of (B.5) are zero for all m. Therefore, I focus on the first term:

$$E\|\varepsilon_t \otimes Z_{t-1} - E(\varepsilon_t \otimes Z_{t-1} \mid \mathcal{F}_{t-m}^{t+m})\| = \sum_{i=1}^{p} E\|\varepsilon_t \otimes y_{t-i} - E(\varepsilon_t \otimes y_{t-i} \mid \mathcal{F}_{t-m}^{t+m})\|. \tag{B.6}$$

Since ε_t is independent, and I have already shown that y_t is NED on $\{\varepsilon_t\}$, by Corollary 4.3 in Gallant and White (1988), $\varepsilon_t \otimes y_{t-i}$ is also NED on $\{\varepsilon_t\}$. Therefore, the right-hand side of (B.6) converges to zero as $m \to \infty$. Thus, I have demonstrated that $f(y_t, \theta)$ is an L^2-NED sequence on $\{\varepsilon_t\}$.

Properties of the moment functions

The moment function $f(y_t, \theta)$ is a composition of one linear map and two polynomial maps in θ. Therefore, the moment function is continuous and differentiable. To compute the derivative of all the moment condition with respect to the parameters of the model (i.e., the coefficients $\theta = \text{vec}(\nu, A_1, \ldots, A_p, B))$, I differentiate the residual ε_t the structural shocks are defined as:

$$\varepsilon_t = B^{-1}(y_t - \nu - A_1 y_{t-1} - \cdots - A_p y_{t-p})$$

The derivative of ε_t with respect to ν is:

$$\frac{\partial \varepsilon_t}{\partial \nu} = -B^{-1}$$

Using the property $\text{vec}(Ay) = (I_n \otimes y')\text{vec}(A)$, I have:

$$\frac{\partial \varepsilon_t}{\partial \text{vec}(A_i)} = -B^{-1}(I_n \otimes y'_{t-i})$$

Using $\varepsilon_t = B^{-1}u_t$ and properties of matrix derivatives:

$$\frac{\partial \varepsilon_t}{\partial \text{vec}(B)} = -(\varepsilon'_t \otimes B^{-1})$$

The full derivative is:

$$\frac{\partial \varepsilon_t}{\partial \theta} = \left[-B^{-1},\ -B^{-1}(I_n \otimes y'_{t-1}),\ \ldots,\ -B^{-1}(I_n \otimes y'_{t-p}),\ -(\varepsilon'_t \otimes B^{-1})\right] \tag{B.7}$$

Next, I compute the derivative of the moment conditions with respect to θ.

1. Moment Condition 1 (Eq. 3):

$$E[\varepsilon_t \otimes Z_{t-1}] = 0,$$

where $Z_t = (1, y'_{t-1}, \ldots, y'_{t-p})'$.

The derivative of this moment condition with respect to θ is:

$$E\left[\frac{\partial(\varepsilon_t \otimes Z_{t-1})}{\partial \theta}\right] = E\left[\frac{\partial \varepsilon_t}{\partial \theta} \otimes Z_{t-1}\right].$$

Substituting $\frac{\partial \varepsilon_t}{\partial \theta}$, I get:

$$E\left[\frac{\partial(\varepsilon_t \otimes Z_{t-1})}{\partial \theta}\right] = E\left[\Big(-B^{-1},\ \ -B^{-1}(I_n \otimes y'_{t-1}),\ \ \ldots,\ \ -B^{-1}(I_n \otimes y'_{t-p}),\ \ -(\varepsilon'_t \otimes B^{-1})\Big) \otimes Z_{t-1}\right].$$

2. Moment Condition 2 (Eq. 4):

$$E[e_i'\varepsilon_t e_j'\varepsilon_t] - \delta_{ij} = 0$$

for all $0 \leq i, j \leq n$.

The derivative of this moment condition with respect to θ is:

$$E\left[\frac{\partial(e_i'\varepsilon_t e_j'\varepsilon_t)}{\partial\theta}\right] = E\left[e_i'\frac{\partial\varepsilon_t}{\partial\theta}e_j'\varepsilon_t + e_i'\varepsilon_t e_j'\frac{\partial\varepsilon_t}{\partial\theta}\right].$$

Substituting $\frac{\partial\varepsilon_t}{\partial\theta}$, I have:

$$E\left[\frac{\partial(e_i'\varepsilon_t e_j'\varepsilon_t)}{\partial\theta}\right] = E\left[e_i'\Big(-B^{-1}, \quad -B^{-1}(I_n \otimes y_{t-1}'), \quad \ldots, \quad -B^{-1}(I_n \otimes y_{t-p}'), \quad -(\varepsilon_t' \otimes B^{-1})\Big)\, e_j'\varepsilon_t\right].$$

3. Moment Condition 3 (Eq. 5):

$$E[e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t] = 0$$

for $0 \leq i, j, k \leq n$ and $|\{i, j, k\}| > 1$.

The derivative of this moment condition with respect to θ is:

$$E\left[\frac{\partial(e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t)}{\partial\theta}\right] = E\left[e_i'\frac{\partial\varepsilon_t}{\partial\theta}e_j'\varepsilon_t e_k'\varepsilon_t + e_i'\varepsilon_t e_j'\frac{\partial\varepsilon_t}{\partial\theta}e_k'\varepsilon_t + e_i'\varepsilon_t e_j'\varepsilon_t e_k'\frac{\partial\varepsilon_t}{\partial\theta}\right].$$

Substituting $\frac{\partial\varepsilon_t}{\partial\theta}$, I get:

$$E\left[\frac{\partial(e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t)}{\partial\theta}\right] = E\left[e_i'\Big(-B^{-1}, \quad -B^{-1}(I_n \otimes y_{t-1}'), \quad \ldots, \quad -B^{-1}(I_n \otimes y_{t-p}'), \quad -(\varepsilon_t' \otimes B^{-1})\Big)\, e_j'\varepsilon_t e_k'\varepsilon_t\right].$$

4. Moment Condition 4 (Eq. 6):

$$E[e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t e_l'\varepsilon_t] - \delta_{ijkl} = 0$$

for $0 \leq i, j, k, l \leq n$ and $|\{i, j, k, l\}| > 1$.

The derivative of this moment condition with respect to θ is:

$$E\left[\frac{\partial(e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t e_l'\varepsilon_t)}{\partial\theta}\right] = E\left[\sum_{\text{all permutations of } e_i', e_j', e_k', e_l'} \frac{\partial\varepsilon_t}{\partial\theta}\right].$$

Substituting $\frac{\partial\varepsilon_t}{\partial\theta}$, this results in a summation over all permutations of e_i', e_j', e_k', e_l'.

5. Moment Condition 5 (Eq. 7):

$$E[e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t e_l'\varepsilon_t] = 0$$

for $|\{i,j,k,l\}| > 1$ and i,j,k,l do not form two distinct pairs.

The derivative of this moment condition with respect to θ is similar to the previous one:

$$E\left[\frac{\partial(e_i'\varepsilon_t e_j'\varepsilon_t e_k'\varepsilon_t e_l'\varepsilon_t)}{\partial\theta}\right] = E\left[\sum_{\text{all permutations of } e_i',e_j',e_k',e_l'} \frac{\partial\varepsilon_t}{\partial\theta}\right].$$

From the above calculation, I can easily see that the first order derivative $\frac{\partial f(y_t,\theta)}{\theta}$ is countinous and bounded. Next, I evaluate the first-order derivative of $f(y_t,\theta)$ at the true value θ_0, where $B = B_0$ and the structural shocks satisfy the moment conditions imposed by Assumptions 1, 2, and 3. The results are simplified due to $E(\varepsilon_t) = 0$, $E(\varepsilon_t\varepsilon_t') = I_n$, and the co-skewness and co-kurtosis assumptions. The detailed derivation of the first-order derivative evaluated at θ_0 shows that it is of full column rank. The proof is similar to Proposition 1 in Lanne and Luoto (2021), so I omit it here.

Verify the assumptions of theorem 1 in Andrews (1993)

I am now ready to prove Proposition 2 by verifying the assumptions of Theorem 1 in Andrews (1993).

Assumptions 1(a) and 1(b) are satisfied because $f(y_t,\theta)$ is NED, and by Assumption 5, which further guarantees that the moment functions have finite second (and higher) moments.

Assumption 1(c) is satisfied since y_t is stationary, and by Assumption 5, which further guarantees that the process has finite second (and higher) moments. which ensures that the covariance matrix of the moment condition is positive definite.

Assumptions 1(e), 2, and 3 are satisfied because I use an optimal weighting matrix in my model, ensuring efficiency and consistency. This choice also adheres to the criteria for the asymptotic properties of the GMM estimator described by Andrews (1993).

Assumptions 1(f) and 1(g) are verified by my previous computation of the derivatives, and by Assumption 5, which ensures that the moment functions are continuously differentiable, and their gradients are bounded.

Assumption 1(h) holds because I have shown that $\frac{\partial f(y_t,\theta)}{\partial\theta}$ is of full column rank, ensuring local identification of the parameters, and again by Assumption 5.

Assumption 1(4) addresses the uniform consistency of $\hat{\theta}$. Andrews (1993) provides a set of conditions, known as Assumption A, under which the PS-GMM estimator $\hat{\theta}(\cdot)$ is consistent for θ_0, uniformly over $\pi \in \Pi$ under the null hypothesis.

Next, I verify that Assumption A holds in my model: - **Assumption A(a)** and **A(b)** are satisfied as shown in previous sections, ensuring the moment conditions are NED and well-behaved. - **Assumption A(c)** holds by Assumption 4, which ensures that the parameter space is compact and bounded. - **Assumptions A(d), A(e), and A(f)** concern the properties of the moment functions, which I have verified in the earlier subsections by demonstrating that the moment conditions are continuous and bounded. - **Assumption A(g)** is satisfied by Proposition 1 and Assumption 4, which guarantees global identification of the model parameters under my error assumptions.

Therefore, I have shown that my model and the PSGMM estimator satisfy all the assumptions in Theorem 1 of Andrews (1993). Thus, Proposition 2 is a direct result of applying Theorem 1 in Andrews (1993).

Appendix C Proof of Proposition 4

In this appendix, I provide the proof of the asymptotic properties of the break date estimator for the SVAR model with an unknown break date. First,Recall from equation (22) :

$$\hat{k} = \arg\min_k \left[Q_1(k, \hat{\theta}_1(k)) + Q_2(k, \hat{\theta}_2(k))\right] \tag{C.1}$$

where

$$Q_1(k, \hat{\theta}_1(k)) = \left(\frac{1}{k}\sum_{t=1}^{k} f(y_t, \hat{\theta}_1(k))\right)' W_1^{-1} \left(\frac{1}{Tk}\sum_{t=1}^{k} f(y_t, \hat{\theta}_1(k))\right),$$
$$Q_2(k, \hat{\theta}_2(k)) = \left(\frac{1}{T-k}\sum_{t=k+1}^{T} f(y_t, \hat{\theta}_2(k))\right)' W_2^{-1} \left(\frac{1}{T-Tk}\sum_{t=k+1}^{T} f(y_t, \hat{\theta}_2(k))\right).$$

I first prove part (2) of the proposition, which shows the consistency of $\hat{k}$. For some constant $M > 0$, consider the probability $P(|\hat{k} - k_0| > M)$, and I aim to show that it converges to zero in probability as $T \to \infty$.

To simplify the notation, define the combined objective function $Q(k, \hat{\theta}_1(k), \hat{\theta}_2(k))$ as:

$$Q(k, \hat{\theta}_1(k), \hat{\theta}_2(k)) = Q_1(k, \hat{\theta}_1(k)) + Q_2(k, \hat{\theta}_2(k)).$$

Denote $D_M = \{k : |k - k_0| > M\}$. By definition of $\hat{k}$ in (C.1), I obtain:

$$\begin{aligned} P\left(|\hat{k} - k_0| > M\right) &\le P\left(\min_{k \in D_M} Q(k, \hat{\theta}_1(k), \hat{\theta}_2(k)) - Q(k_0, \hat{\theta}_1(k_0), \hat{\theta}_2(k_0)) \ge 0\right) \\ &\le P\left(\min_{k \in D_M} \min_{\theta_1, \theta_2 \in (\Pi \times \mathcal{B}_1)^2} Q(k, \theta_1, \theta_2) - Q(k_0, \hat{\theta}_1(k_0), \hat{\theta}_2(k_0)) \ge 0\right). \end{aligned} \tag{C.2}$$

Next, I examine the limiting behavior of the function $Q(k, \theta_1, \theta_2)$ for any hypothetical θ_1 and θ_2.

Denote

$$Q_{10}(k, \theta_1) = \left(\frac{1}{k}\sum_{t=1}^{k} E[f(y_t, \theta_1)]\right)' W_1^{-1} \left(\frac{1}{k}\sum_{t=1}^{k} E[f(y_t, \theta_1)]\right).$$

From lemma A3 in Andrews(1994), I obtain the uniform convergence:

$$Q_1(k,\theta_1) \xrightarrow{p} Q_{10}(k,\theta_1).uniformly\ in\ k\ and\ \theta_1$$

and :

$$Q_2(k,\theta_2) \xrightarrow{p} Q_{20}(k,\theta_2),\ uniformly\ in\ k\ and\ \theta_2$$

where

$$Q_{20}(k,\theta_2) = \left(\frac{1}{T-k}\sum_{t=k+1}^{T} E[f(y_t,\theta_2)]\right)' W_2^{-1} \left(\frac{1}{T-k}\sum_{t=k+1}^{T} E[f(y_t,\theta_2)]\right).$$

Let $Q_0(k,\theta_1,\theta_2) = Q_{10}(k,\theta_1)+Q_{20}(k,\theta_2)$. Add and subtract $Q_0(k,\theta_1,\theta_2)$ and $Q_0(k_0,\theta_{10},\theta_{20})$ in equation (C.2):

$$\begin{aligned}\min_{k\in D_M}\min_{\theta_1,\theta_2\in(\Pi\times\mathcal{B}_1)^2} & Q(k,\theta_1,\theta_2) - Q_0(k,\theta_1,\theta_2) \\ & + Q_0(k_0,\theta_{10},\theta_{20}) - Q(k_0,\hat\theta_1(k_0),\hat\theta_2(k_0)) \\ & + Q_0(k,\theta_1,\theta_2) - Q_0(k_0,\theta_{10},\theta_{20}).\end{aligned} \tag{C.3}$$

The first and second terms converge to zero following the previous discussion. Furthermore, since $\hat\theta_1(k_0) \xrightarrow{p} \theta_{10}$ and $\hat\theta_2(k_0) \xrightarrow{p} \theta_{20}$ due to the consistency of the standard GMM estimator, the third term satisfies:

$$\min_{k\in D_M}\min_{\theta_1,\theta_2\in(\Pi\times\mathcal{B}_1)^2} Q_0(k,\theta_1,\theta_2) - Q_0(k_0,\theta_{10},\theta_{20}) > 0,$$

by Assumption 7. This implies that $P(|\hat k - k_0| > M)$ converges to zero in probability. Hence, the asymptotic bounded of $\hat k$ is established, i.e,

$$\hat k = k_0 + O_p(1) \tag{C.4}$$

Given that $\hat\lambda = \hat k/T$ and $\lambda_0 = k_0/T$, I directly have that The estimated break point fraction $\hat\lambda$ is T-consistent for the true break point fraction λ_0.

I now prove part (1) of the proposition. From equation (10), I have:

$$\hat\theta_1(\hat k) = \arg\min_{\theta_1} \left(\frac{1}{\hat k}\sum_{t=1}^{\hat k} f(y_t,\theta_1)\right)' W_1^{-1} \left(\frac{1}{\hat k}\sum_{t=1}^{\hat k} f(y_t,\theta_1)\right).$$

To analyze the convergence of the right-hand side, I apply the uniform law of large numbers. Consider the quantity $sup_{\theta_1}|\frac{1}{\hat{k}}\sum_{t=1}^{\hat{k}} f(y_t,\theta_1) - E[f(y_t,\theta_1)|$.

$$\sup_{\theta_1}|\frac{1}{\hat{k}}\sum_{t=1}^{\hat{k}} f(y_t,\theta_1) - E[f(y_t,\theta_1)]| \le \sup_{\theta_1}|\frac{1}{\hat{k}}\sum_{t=1}^{\hat{k}} f(y_t,\theta_1) - \frac{1}{\hat{k}}\sum_{t=1}^{\hat{k}} E[f(y_t,\theta_1)]| \tag{C.5}$$

$$+\sup_{\theta_1}|\frac{1}{\hat{k}}\sum_{t=1}^{\hat{k}} E[f(y_t,\theta_1)] - E[f(y_t,\theta_1)]| \tag{C.6}$$

The first term on the RHS converges to zero in probability because of uniform law of numbers. Now consider the second term. Suppose, without loss of generality, that $\hat{k} > k_0$. Then,

$$\sup_{\theta_1}|\frac{1}{\hat{k}}\sum_{t=1}^{\hat{k}} E[f(y_t,\theta_1)] - E[f(y_t,\theta_1)]| = \sup_{\theta_1}|\frac{1}{\hat{k}}\sum_{t=1}^{k_0} E[f(y_t,\theta_1)] + \frac{1}{\hat{k}}\sum_{t=k_0+1}^{\hat{k}} E[f(y_t,\theta_1)] - E[f(y_t,\theta_1)]| \tag{C.7}$$

$$\le \frac{\hat{k}-k_0}{\hat{k}}\sup_{\theta_1} E[f(y_t,\theta_1)] + \frac{\hat{k}-k_0}{\hat{k}}\sup_{t}\sup_{\theta_1} E[f(y_t,\theta_1)] \tag{C.8}$$

$$\xrightarrow{p} 0 \tag{C.9}$$

where the inequality holds because $f(y_t,\theta)$ is stationary and bounded over $\theta \in \Pi \times \mathcal{B}_1$, and the limit holds due to the stochastic bounded of $\hat{k} - k_0$.

Therefore, under regular conditions, I have $\sqrt{T}\left(\hat{\theta}_1(\hat{k}) - \theta_{10}\right) \xrightarrow{d} N(0, G'_{\theta_{10}} W^{-1}_{\theta_{10}} G_{\theta_{10}})$. Similarly, $\sqrt{T}\left(\hat{\theta}_2(\hat{k}) - \theta_{20}\right) \xrightarrow{d} N(0, G'_{\theta_{20}} W^{-1}_{\theta_{20}} G_{\theta_{20}})$. This completes the proof.

Appendix D Proof of Proposition 6

Proof. I adapt the logic from Bai (1997) to prove my proposition 6. The goal is to show that:

$$P(\hat{m} = m_0) \to 1, \quad \text{as } T \to \infty.$$

First, consider the event that the estimated number of breaks is less than the true number, i.e., $\hat{m} < m_0$. When the estimator underestimates the number of breaks, there must exist a segment $[\hat{k}_p, \hat{k}_q]$ that contains at least one true break point that has not been detected by the estimation procedure. Let $k_{r0} \in (\hat{k}_p, \hat{k}_q)$ be a true break point in this segment. Since k_{r0} is not detected by $[\hat{k}_p$ and $\hat{k}_q]$, by proposition 5 neither $\hat{k}_p/T$ and $\hat{k}_q/T$ converges to k_{r0}/T in probability. That is:

$$k_{r0} - \hat{k}_p > T\epsilon_0 \quad \text{and} \quad \hat{k}_q - k_{r0} > T\epsilon_0,$$

for some constant $\epsilon_0 > 0$.

I now consider the SupWald test statistic based on the subsample $[\hat{k}_p, \hat{k}_q]$. I first show that this statistic is of order T. To see this, recall from the definition of Wald statistic in equation (21), $Wald_T(\lambda) = T(\hat{\theta}_1(\lambda) - \hat{\theta}_2(\lambda))' \left(\frac{\hat{V}_1}{\lambda} + \frac{\hat{V}_2}{1-\lambda}\right)^{-1} (\hat{\theta}_1(\lambda) - \hat{\theta}_2(\lambda))$. Consider the statistic at the true break point k_{r0}: $Wald_T(\lambda_{r0})$, where $\lambda_{r0} = k_{r0}/T$.

$$Wald_T(\lambda_{r0}) = T(\hat{\theta}_1(\lambda_{r0}) - \hat{\theta}_2(\lambda_{r0}))' \left(\frac{\hat{V}_1}{\lambda_{r0}} + \frac{\hat{V}_2}{1-\lambda_{r0}}\right)^{-1} (\hat{\theta}_1(\lambda_{r0}) - \hat{\theta}_2(\lambda_{r0})) \tag{D.1}$$

$$\xrightarrow{p} T(\theta_{1,r0} - \theta_{2,r0})' \left(\frac{V_1}{\lambda_{r0}} + \frac{V_2}{1-\lambda_{r0}}\right)^{-1} (\theta_{1,r0} - \theta_{2,r0}) \tag{D.2}$$

$$= O_p(T) \tag{D.3}$$

where $\theta_{1,r0}$ and $\theta_{2,r0}$ denotes the true parameter in this subsample. The converge in (D.2) holds because of the proof in Appendix C. The last equality holds because k_{r0} is the true break point so that $\theta_{1,r0} \neq \theta_{2,r0}$ and the weighting matrix is positive definite from assumption 8. Next since $\sup Wald_T(\lambda) \geq Wald_T(\lambda_{r0}) = O_p(T)$, this implies that the $\sup Wald_T(\lambda)$ is $O_p(T)$.

Then, there exists a constant $\pi > 0$ such that, for every $\epsilon > 0$:

$$P(\sup F \geq \pi T) \geq 1 - \epsilon,$$

for sufficiently large T.

Let $\alpha_T = KT^{-1}$ and set the threshold $C_T = 3 \log T$, following the discussion in Bai (1997) after Lemma 10. Under this choice of α_T, I have:

$$P(\sup F \geq C_T) \geq 1 - \epsilon,$$

for sufficiently large T, or equivalently:

$$P(\sup F \geq C_T) \to 1 \quad \text{as } T \to \infty.$$

Thus, the null hypothesis of parameter constancy will be rejected with probability approaching 1 as the sample size increases. This implies that $P(\hat{m} < m_0) \to 0$ as $T \to \infty$.

Next, consider the event where $\hat{m} > m_0$. For $\hat{m} > m_0$ to be true, it must be the case that for some i, at a certain stage in the sequential estimation process, one rejects the null hypothesis of parameter constancy for an interval $[\hat{k}_i, \hat{k}_{i+1}]$, where $\hat{k}_i = k_{i0} + O_p(1)$ and $\hat{k}_{i+1} = k_{(i+1)0} + O_p(1)$. In other words, the interval contains no true nontrivial break point, but the null hypothesis is rejected. Therefore, I have:

$$P(\hat{m} > m_0) \leq P\left(\exists i, \text{ reject parameter constancy for } [\hat{k}_i, \hat{k}_{i+1}]\right)$$

Given the sequential nature of the testing procedure, this probability can be bounded by:

$$P(\hat{m} > m_0) \leq \sum_{i=0}^{m_0} P\left(\text{reject parameter constancy for } [\hat{k}_i, \hat{k}_{i+1}]\right),$$

where $\hat{k}_0 = 1$ and $\hat{k}_{m_0+1} = T$, the endpoints of the sample.

Now, since $\hat{k}_i - \hat{k}_{i-1} = O_p(T)$, the length of the segment, and the test statistic computed for the interval $[\hat{k}_i, \hat{k}_{i+1}]$, denoted as $\sup Wald_i$, converges in distribution to some quantity. By proposition 3, for large T (and hence large $n = \hat{k}_i - \hat{k}_{i-1}$), the supremum test statistic $\sup Wald$ converges to a limiting distribution, and:

$$P(\sup Wald_i > C_T) \to 0 \quad \text{as } T \to \infty,$$

where $C_T = 3\log T$, as defined earlier.

Thus, for sufficiently large T, I have:

$$P(\hat{m} > m_0) \leq (m_0 + 1) \max_{0 \leq i \leq m_0} P(\sup F_i > C_T) \to 0 \quad \text{as } T \to \infty.$$

This establishes that the probability of overestimating the number of breaks converges to zero as the sample size increases, provided that $\alpha_T \to 0$, i.e, $C_T \to \infty$.

Thus, the proof is complete. □

References

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica* **61**(4): 821–856.

Angelini, G., Bacchiocchi, E., Caggiano, G. and Fanelli, L. (2019). Uncertainty across volatility regimes, *Journal of Applied Econometrics* **34**(3): 437–455.

Auerbach, A. J. and Gorodnichenko, Y. (2012). Measuring the output responses to fiscal policy, *American Economic Journal: Economic Policy* **4**(2): 1–27.

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**(6): 761–779.

Bacchiocchi, E. and Kitagawa, T. (2024). SVARs with Breaks: Identification and Inference, *arXiv preprint arXiv:2405.04973* .

Bai, J. (1994). Least Squares Estimation of a Shift in Linear Processes, *Journal of Time Series Analysis* **15**(5): 453–472.

Bai, J. (1997). Estimating Multiple Breaks One at a Time, *Econometric Theory* **13**(3): 315–352.

Bai, J. (2000). Vector Autoregressive Models with Structural Changes in Regression Coefficients and in Variance-Covariance Matrices, *Annals of Economics and Finance* **1**(2): 303–339.

Bai, J., Lumsdaine, R. L. and Stock, J. H. (1998). Testing for and Dating Common Breaks in Multivariate Time Series, *The Review of Economic Studies* **65**(3): 395–432.

Bai, J. and Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica* **66**(1): 47–78.

Bai, J. and Perron, P. (2003). Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics* **18**(1): 1–22.

Baker, S. R., Bloom, N. and Davis, S. J. (2016). Measuring economic policy uncertainty, *The quarterly journal of economics* **131**(4): 1593–1636.

Bergholt, D., Furlanetto, F. and Vaccaro-Grange, G. (2023). The Flattening of the Phillips Curve, *Journal of Economic Perspectives* **37**(2): 165–194.

Bernanke, B. S. and Blinder, A. S. (1992). The federal funds rate and the channels of monetary transmission, *American Economic Review* **82**(4): 901–921.

Bernanke, B. S. and Mihov, I. (1998). Measuring monetary policy, *The quarterly journal of economics* **113**(3): 869–902.

Boivin, J. and Giannoni, M. P. (2006). Has Monetary Policy Become More Effective?, *The Review of Economics and Statistics* **88**(3): 445–462.

Canova, F. (2009). What explains the great moderation in the u.s.? a structural analysis, *Journal of the European Economic Association* **7**(4): 697–721.

Canova, F., Gambetti, L. and Pappa, E. (2008). The structural dynamics of output growth and inflation: Some international evidence, *Economic Journal* **118**(533): 1282–1312.

Carriero, A., Clark, T. E. and Marcellino, M. (2018). Measuring uncertainty and its impact on the economy, *Review of Economics and Statistics* **100**(5): 799–815.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end?, *Handbook of Macroeconomics* **1**: 65–148.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post wwii us, *Review of Economic dynamics* **8**(2): 262–302.

Crucil, R., Hambuckers, J. and Maxand, S. (2023). Do monetary policy shocks affect financial uncertainty? a non-gaussian proxy svar approach, *A Non-gaussian Proxy SVAR Approach (June 5, 2023)* .

Del Negro, M., Lenza, M., Primiceri, G. and Tambalotti, A. (2020). What's Up with the Phillips Curve?, *Brookings Papers on Economic Activity* **2020**(Spring): 301–373.

Demetrescu, M. and Salish, N. (2024). (structural) var models with ignored changes in mean and volatility, *International Journal of Forecasting* **40**(2): 840–854.

Drautzburg, T. and Wright, J. H. (2023). Refining set-identification in vars through independence, *Journal of Econometrics* **235**(2): 1827–1847.

Eichenbaum, M. (1995). Some empirical evidence on the effects of shocks to monetary policy on exchange rates, *Quarterly Journal of Economics* **110**(4): 975–1000.

Hoesch, L., Lee, A. and Mesters, G. (2024). Locally robust inference for non-gaussian svar models, *Quantitative Economics* **15**(2): 523–570.

Inoue, A. and Rossi, B. (2011). Identifying the Sources of Instabilities in Macroeconomic Fluctuations, *The Review of Economics and Statistics* **93**(4): 1186–1204.

Justiniano, A. and Primiceri, G. E. (2008). The Time-Varying Volatility of Macroeconomic Fluctuations, *American Economic Review* **98**(3): 604–641.

Keweloh, S. A. (2021). A generalized method of moments estimator for structural vector autoregressions based on higher moments, *Journal of Business & Economic Statistics* **39**(3): 772–782.

Keweloh, S. A. (2024). Uncertain short-run restrictions and statistically identified structural vector autoregressions, *arXiv preprint arXiv:2303.13281* .

Keweloh, S. A., Klein, M. and Prüser, J. (2023). Estimating fiscal multipliers by combining statistical identification with potentially endogenous proxies, *arXiv preprint arXiv:2302.13066* .

Lanne, M., Liu, K. and Luoto, J. (2023a). Identifying Structural Vector Autoregression via Leptokurtic Economic Shocks, *Journal of Business & Economic Statistics* **41**(4): 1341–1351.

Lanne, M., Liu, K. and Luoto, J. (2023b). Identifying Structural Vector Autoregressions via Non-Gaussianity of Potentially Dependent Structural Shocks, *Available at SSRN 4564713* .

Lanne, M. and Luoto, J. (2021). GMM Estimation of Non-Gaussian Structural Vector Autoregression, *Journal of Business & Economic Statistics* **39**: 69–81.

Lanne, M., Meitz, M. and Saikkonen, P. (2017). Identification and Estimation of Non-Gaussian Structural Vector Autoregressions, *Journal of Econometrics* **196**(2): 288–304.

Lee, A. and Mesters, G. (2024). Locally robust inference for non-gaussian linear simultaneous equations models, *Journal of Econometrics* **240**(1): 105647.

Lewis, D. J. (2021). Identifying shocks via time-varying volatility, *The Review of Economic Studies* **88**(6): 3086–3124.

Lewis, D. J. (2024). Identification based on higher moments.

Ling, S. (2016). Estimation of change-points in linear and nonlinear time series models, *Econometric Theory* **32**(2): 402–430.

Magnusson, L. and Mavroeidis, S. (2014). Identification-Robust Inference in SVARs, *Journal of Monetary Economics* **66**: 47–60.

Mann, H. B. and Wald, A. (1943). On Stochastic Limit and Order Relationships, *The Annals of Mathematical Statistics* **14**(3): 217–226.

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**(5): 1464–1476.

Mesters, G. and Zwiernik, P. (2022). Non-Independent Components Analysis, *arXiv preprint arXiv:2206.13668* .

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy, *The Review of Economic Studies* **72**(3): 821–852.

Qu, Z. and Perron, P. (2007). Estimating and Testing Structural Changes in Multivariate Regressions, *Econometrica* **75**(2): 459–502.

Rigobon, R. (2003). Identification Through Heteroskedasticity, *The Review of Economics and Statistics* **85**(4): 777–792.

Sims, C. A. (1980). Macroeconomics and Reality, *Econometrica* **48**(1): 1–48.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy, *European Economic Review* **36**(5): 975–1000.

Sims, C. A. and Zha, T. (2006). Were There Regime Switches in US Monetary Policy?, *American Economic Review* **96**(1): 54–81.

Smets, F. and Wouters, R. (2007). Shocks and frictions in u.s. business cycles: A bayesian dsge approach, *American Economic Review* **97**(3): 586–606.

Stock, J. H. and Watson, M. W. (2002). Has the Business Cycle Changed and Why?, *NBER Macroeconomics Annual* .

Stock, J. H. and Watson, M. W. (2012). Disentangling the channels of the 2007-2009 recession, *Technical report*, National Bureau of Economic Research.

Stock, J. H. and Wright, J. H. (2000). GMM with Weak Identification, *Econometrica* **68**(5): 1055–1096.

Velasco, C. (2023). Identification and estimation of structural varma models using higher order dynamics, *Journal of Business & Economic Statistics* **41**(3): 819–832.

Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound, *Journal of Money, Credit and Banking* **48**(2-3): 253–291.