

1871

Discussion  
Papers

# Structural Vector Autoregressive Models with more Shocks than Variables Identified via Heteroskedasticity

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

#### IMPRESSUM

© DIW Berlin, 2020

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

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

ISSN electronic edition 1619-4535

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

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

# Structural Vector Autoregressive Models with more Shocks than Variables Identified via Heteroskedasticity

Helmut Lütkepohl<sup>1</sup>

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

May 15, 2020

**Abstract.** In conventional structural vector autoregressive (VAR) models it is assumed that there are at most as many structural shocks as there are variables in the model. It is pointed out that heteroskedasticity can be used to identify more shocks than variables. However, even if there is heteroskedasticity, the number of shocks that can be identified is limited. A number of results are provided that allow a researcher to assess how many shocks can be identified from specific forms of heteroskedasticity.

*Key Words:* Structural vector autoregression, identification through heteroskedasticity, structural shocks

*JEL classification:* C32

---

<sup>1</sup>I thank Max Diegel for inspiring discussions of the subject and Adrian Pagan for comments on an earlier version of the paper.

# 1 Introduction

In traditional structural vector autoregressive (VAR) analysis, the number of shocks is equal to the number of variables (see, e.g., Sims (1980) or a number of models reviewed in Kilian and Lütkepohl (2017)). Even if only a subset of shocks is of interest for some economic analysis, additional shocks are typically specified and identified in some way such that all shocks are contemporaneously uncorrelated. Also in some dynamic stochastic general equilibrium (DSGE) models, the number of shocks is equal to the number of variables (e.g., Smets and Wouters (2007)).

However, in recent years, economic models have been constructed, where the number of shocks exceeds the number of observable variables. For example, Christiano, Motto and Rostagno (2014) specify such a model, where news shocks play an important role. Another example is Ireland (2011), who constructs a small New-Keynesian DSGE model with more shocks than observed variables. Plagborg-Møller and Wolf (2017) allow for more shocks than variables in a nonparametric setting which extends the standard VAR approach. In addition, Pagan and Robinson (2020) list a number of situations, where more shocks than observed variables appear for technical reasons in the model. For example, such features may be implied by measurement errors in some variables or they may be due to using unobserved component models or time-varying parameter models.

Pagan and Robinson (2020) also point out that ignoring the fact that the number of shocks exceeds the dimension of the VAR process can lead to distorted impulse responses in a standard structural VAR analysis and to severe problems in the interpretation of the evidence from these models.

In the present study, it is shown that heteroskedasticity can be used to identify more shocks than there are observed variables in a VAR model. I consider a VAR model where the volatility of the residuals changes during the sample period and there are actually  $M$  different volatility regimes, where  $M > 1$  is some positive integer. An upper bound for the number of identifiable shocks is given and conditions are presented that have to be satisfied for identification. Moreover, a rank condition for identification is derived.

The following notation is used. The column vectorizing operator for matrices is denoted by  $\text{vec}$  and  $\text{vech}$  signifies the operator that stacks the columns of a  $(K \times K)$  square matrix from the main diagonal downward in a  $\frac{1}{2}K(K+1)$ -dimensional column vector. The symbol  $\mathbf{D}_K$  denotes a  $(K^2 \times \frac{1}{2}K(K+1))$  duplication matrix,  $\mathbf{D}_K^+$  is its Moore-Penrose generalized inverse, and  $\mathbf{K}_{KK}$  denotes a  $(K^2 \times K^2)$  commutation matrix (see Lütkepohl (1996) for precise definitions of these matrices and rules for working with them). Moreover,  $\text{diag}(a_1, \dots, a_N)$  denotes an  $(N \times N)$  diagonal matrix with

scalars  $a_1, \dots, a_N$  on the main diagonal and  $\mathbf{S}_N$  signifies an  $(N^2 \times N)$  dimensional matrix with elements 0 and 1 such that  $\text{vec}(\text{diag}(a_1, \dots, a_N)) = \mathbf{S}_N(a_1, \dots, a_N)'$ , e.g., for  $N = 2$ ,

$$\mathbf{S}_N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}'.$$

The structure of the remaining part of the paper is as follows. The model setup is presented in the next section. The main results on identification through heteroskedasticity are given in Section 3 and conclusions are drawn in Section 4.

## 2 Model Setup

Suppose the  $K \geq 2$  observable variables  $y_t = (y_{1t}, \dots, y_{Kt})'$  are generated by the reduced-form VAR( $p$ ) process

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

where  $u_t$  is a zero mean white noise process with nonsingular covariance matrix  $\Sigma_u$ , i.e.,  $u_t \sim (0, \Sigma_u)$ . There are no deterministic terms because they are not important in the following discussion. In practice, they may be added without changing the essential results.

Suppose that there are  $N \geq K$  structural errors,  $w_t = (w_{1t}, \dots, w_{Nt})'$ , which are linearly related to the reduced form errors by

$$u_t = B w_t. \tag{1}$$

In other words, the structural matrix  $B$  is  $(K \times N)$  and has rank  $K$ . Without loss of generality, it is also assumed that none of the columns of  $B$  is zero because if the  $n^{\text{th}}$  column were zero, this would imply that the  $n^{\text{th}}$  structural shock has no impact on the reduced-form errors and, hence, on the observed variables. Thus, it would be redundant.

Without loss of generality, it is also assumed that  $w_t$  has identity covariance matrix, i.e.,  $w_t \sim (0, I_K)$ . Thus,

$$\Sigma_u = B B'.$$

Clearly, these  $K(K + 1)/2$  linearly independent relations between reduced form parameters  $\Sigma_u$  and structural parameters  $B$  cannot identify all  $K \cdot N$  structural parameters. However, they can be identified under suitable assumptions if the reduced form residuals are heteroskedastic, a case which will be considered next.

Because  $\text{rk}(B) = K$ , the impact effects of the shocks must be linearly dependent if  $N > K$ . This feature also has implications for the impulse responses of the structural shocks. As usual, the impulse responses are obtained as  $\Phi_i B$  for  $i = 1, \dots, H$ , where the  $(K \times K)$  matrix  $\Phi_i = \sum_{j=1}^i \Phi_{i-j} A_j$  can be obtained recursively from the VAR slope coefficients for  $i = 1, 2, \dots$ , using  $\Phi_0 = I_K$  (Lütkepohl, 2005, Section 2.1.2).

### 3 Identification via Heteroskedasticity

Suppose that a sample  $y_1, \dots, y_T$  of size  $T$  is available and there are  $M$  different volatility regimes such that covariance changes occur at prespecified dates throughout the sample,

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

where  $\mathcal{T}_m = \{T_{m-1} + 1, \dots, T_m\}$ ,  $m = 1, \dots, M$ , are  $M$  given volatility regimes, for simplicity associated with consecutive time periods. The volatility changes occur at time periods  $T_m$ , for  $m = 1, \dots, M - 1$ , with  $T_0 = 0$  and  $T_M = T$ .

It is assumed that the structural shocks are contemporaneously uncorrelated and that a time-invariant transformation  $u_t = B w_t$  exists such that the regime dependent covariance matrices are of the form

$$\Sigma_u(m) = B \Lambda_m B', \quad m = 1, \dots, M, \quad (3)$$

where  $\Lambda_m = \text{diag}(\boldsymbol{\lambda}_m)$  with  $\boldsymbol{\lambda}_m = (\lambda_{1,m}, \dots, \lambda_{N,m})'$  ( $m = 1, \dots, M$ ) are diagonal matrices with strictly positive diagonal elements. Without loss of generality, I assume that the variances of the structural errors  $w_t$  are normalized to unity in the first volatility state, i.e.,  $\Lambda_1 = I_K$ . This normalization implies that the  $\boldsymbol{\lambda}_m$ ,  $m = 2, \dots, M$ , represent the variances of the structural shocks relative to the first regime. I collect them in the  $(N \times (M - 1))$  matrix  $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_2, \dots, \boldsymbol{\lambda}_M]$ .

Under suitable conditions, the structural parameters  $B$  and  $\boldsymbol{\Lambda}$  can be recovered uniquely from the reduced-form parameters  $\Sigma_u(m)$ ,  $m = 1, \dots, M$ . The following proposition spells out necessary conditions for the structural parameters to be identified.

**Proposition 1. (Counting Rule)**

The following conditions have to be satisfied for the structural parameters  $B$  and  $\boldsymbol{\Lambda}$  to be solved uniquely from the relations (3):

- (1)  $N < \frac{1}{2}K(K + 1)$

Table 1: Minimal Number of Necessary Volatility States  $M$  for Selected Numbers of Variables  $K$  and Shocks  $N$

No. of variables $K$	No. of shocks $N$													
	2	3	4	5	6	7	8	9	10	11	12	13	14	
2	2	-	-	-	-	-	-	-	-	-	-	-	-	-
3		2	4	10	-	-	-	-	-	-	-	-	-	-
4			2	3	5	7	12	27	-	-	-	-	-	-
5				2	3	4	5	6	8	11	16	26	56	

Note: A dash indicates that condition (1) of Proposition 1 is not satisfied.

$$(2) \quad M \geq 2(K - 1)N / (K^2 + K - 2N) \quad \square$$

**Proof:** Given the symmetry of  $\Sigma_u(m)$ , the sets of relations in (3) each provide at most  $\frac{1}{2}K(K + 1)$  independent equations. Thus, to solve for all  $KN + (M - 1)N$  unknown structural parameters in  $B$  and  $\mathbf{\Lambda}$ , we need  $M$  to be such that

$$\frac{1}{2}MK(K + 1) \geq KN + (M - 1)N. \quad (4)$$

This inequality can only be satisfied if  $N < \frac{1}{2}K(K + 1)$  which proves (1) and rewriting the inequality (4) gives condition (2) of the proposition.  $\square$

Table 1 shows the minimal number of volatility states,  $M$ , for selected values of  $K$  and  $N$ . For example, in a bivariate model ( $K = 2$ ), the impact effects of no more than two shocks can be recovered by heteroskedasticity. For larger values of  $N$ , the conditions of Proposition 1 cannot hold. For a 3-dimensional VAR process, the impact effects of up to  $N = 5$  shocks can be identified through heteroskedasticity and estimated consistently. To estimate unique impact effects of 5 shocks, at least 10 volatility regimes are needed, however. If there are 5 observable variables in the VAR, it is even possible to identify the impact effects of up to 14 shocks. That requires 56 volatility states, however, which is not a typical situation in macroeconomic models.

The next proposition spells out another condition that has to be satisfied for  $B$  to be (locally) identified.

**Proposition 2. (Necessary Condition for Identification)**

The matrix  $B$  is unique up to column sign and column permutation only if

$\forall l, n \in \{1, \dots, N\}$  with  $l \neq n$ ,  $\exists m \in \{2, \dots, M\}$  such that  $\lambda_{lm} \neq \lambda_{nm}$ .  $\square$

**Proof:** We show that the  $l^{\text{th}}$  and  $n^{\text{th}}$  column of  $B$  are not identified if  $\lambda_{lm} = \lambda_{nm}$  for  $m = 2, \dots, M$ . Since the columns of  $B$  can be permuted arbitrarily, we assume that, without loss of generality,  $\lambda_{1m} = \lambda_{2m} = \lambda_0$  for  $m = 2, \dots, M$ , and we define the orthogonal ( $N \times N$ ) matrix

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & I_{N-2} \end{bmatrix},$$

where

$$Q_1 = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

is an orthogonal ( $2 \times 2$ ) matrix for  $\theta \in [0, 2\pi]$ . Clearly, if  $B$  satisfies the relations in (3) and  $\lambda_{1m} = \lambda_{2m} = \lambda_0$ , for  $m = 2, \dots, M$ , then

$$\Sigma_u(m) = BQ\Lambda_m Q' B', \quad m = 1, \dots, M,$$

because  $BQ Q' B' = B B' = \Sigma_u(1)$  due to the orthogonality of  $Q$  and

$$\begin{aligned} BQ \begin{bmatrix} \lambda_0 & 0 & 0 & \dots & 0 \\ 0 & \lambda_0 & 0 & \dots & 0 \\ 0 & 0 & \lambda_{3m} & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_{Nm} \end{bmatrix} Q' B' &= B \begin{bmatrix} \lambda_0 Q_1 Q_1' & 0 & \dots & 0 \\ 0 & \lambda_{3m} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_{Nm} \end{bmatrix} B' \\ &= B\Lambda_m B' = \Sigma_u(m) \end{aligned}$$

for  $m = 2, \dots, M$ , because  $Q_1$  is an orthogonal matrix. Thus, if  $B$  satisfies (3), then  $BQ$  also satisfies the relations in expression (3). In other words,  $B$  is not identified.  $\square$

The proposition states that some heterogeneity in the variance changes is required to solve uniquely for  $B$  and, hence, the shocks. If the variances of two or more shocks change but they change in the same proportion so that the relative changes are identical, then some columns of  $B$  will not be identified even if there are sufficiently many volatility changes to satisfy the counting rules. In principle, the data are informative about the variance heterogeneity. A formal statistical test for a special case was proposed by Lütkepohl, Meitz, Netšunajev and Saikkonen (2020).



The following proposition gives a sufficient condition for  $B$  to be (locally) identified.

**Proposition 3. (Rank Condition)**

The matrix  $B$  is unique up to column sign and column permutation and the matrix  $\mathbf{\Lambda}$  is unique up to row permutation if the following matrix has rank  $(K + M - 1)N$ :

$$\begin{bmatrix} 2\mathbf{D}_K^+(B \otimes I_K) & 0 & \dots & 0 \\ 2\mathbf{D}_K^+(B\Lambda_2 \otimes I_K) & \mathbf{D}_K^+(B \otimes B)\mathbf{S}_N & & 0 \\ \vdots & & \ddots & \vdots \\ 2\mathbf{D}_K^+(B\Lambda_M \otimes I_K) & 0 & \dots & \mathbf{D}_K^+(B \otimes B)\mathbf{S}_N \end{bmatrix}, \quad (5)$$

$(\frac{1}{2}K(K+1)M \times (K+M-1)N)$

where  $\mathbf{D}_K$  and  $\mathbf{S}_N$  are the matrices specified in Section 1. □

**Proof:** Let

$$\varphi(B, \mathbf{\Lambda}) = \begin{bmatrix} \text{vech}(BB' - \Sigma_u(1)) \\ \text{vech}(B\Lambda_2 B' - \Sigma_u(2)) \\ \vdots \\ \text{vech}(B\Lambda_M B' - \Sigma_u(M)) \end{bmatrix} \quad (\frac{1}{2}K(K+1)M \times 1)$$

be a function from the  $(K + M - 1)N$ -dimensional Euclidean space to the  $\frac{1}{2}K(K + 1)M$ -dimensional space. According to Theorem 6 of Rothenberg (1971),  $\varphi(B, \mathbf{\Lambda}) = 0$  can be solved uniquely in a neighbourhood of a point  $(B_0, \mathbf{\Lambda}_0)$  if and only if the matrix of first partial derivatives at the point  $(B_0, \mathbf{\Lambda}_0)$  has rank  $(K + M - 1)N$  (see Lütkepohl (2005, Proof of Proposition 9.1) for a related argument). Thus, the proposition follows by noting that

$$\begin{aligned} \frac{\partial \text{vech}(B\Lambda_m B' - \Sigma_u(m))}{\text{vec}(B)'} &= \mathbf{D}_K^+(I_{K^2} + \mathbf{K}_{KK})(B\Lambda_m \otimes I_K) \\ &= 2\mathbf{D}_K^+(B\Lambda_m \otimes I_K) \end{aligned}$$

and

$$\frac{\partial \text{vech}(B\Lambda_m B' - \Sigma_u(m))}{\boldsymbol{\lambda}'_m} = \mathbf{D}_K^+(B \otimes B)\mathbf{S}_N$$

(see the rules in Chapter 10 of Lütkepohl (1996) and the derivation on p. 361 of Lütkepohl (2005)). In other words, the matrix in expression (5) is just

the matrix of partial derivatives of  $\varphi(B, \mathbf{\Lambda})$ . □

Note that the rank condition in Proposition 3 globally identifies  $B$  and  $\mathbf{\Lambda}$  if the ordering of the rows of  $\mathbf{\Lambda}$  and the column signs of  $B$  are fixed uniquely. In some situations suitable standardisations are easy to establish. For example, if there are just two volatility states ( $M = 2$ ) and, hence,  $\mathbf{\Lambda}$  consists only of one column, then the elements in the column could be ordered in some way, e.g., from largest to smallest. Moreover, the column signs of  $B$  can be chosen such that the first nonzero element in each column is positive. Recall that none of the columns of  $B$  is zero under our assumptions.

However, the proposition may not be very operational because it involves the unknown structural parameters. Of course, in some cases an inspection of the matrix in (5) may provide a clear indication of the rank condition to be satisfied. In other cases that may not be obvious, however. A possible way out in such a situation may be to replace the unknown parameters by random choices from the feasible parameter space and check the resulting rank. Such a check of identification makes sense because the matrix is rank deficient everywhere in the parameter space or it has rank  $(K + M - 1)N$  almost everywhere. Thus, if a rank of  $(K + M - 1)N$  is found for a random set of parameter values, this indicates identification (see also the discussion on p. 361 of Lütkepohl (2005)).

The volatility model assumed so far is restrictive in some respects and can be extended easily. The discussion assumes that there are  $M$  volatility states. That structure is also available if the volatility changes are generated endogenously by a Markov-switching process, for example, as assumed in some of the related literature (see, e.g., Lanne, Lütkepohl and Maciejowska (2010), Herwartz and Lütkepohl (2014)). The previous discussions and results also cover such models.

Having identified the impact effects matrix  $B$  through heteroskedasticity provides unique shocks. These shocks are not necessarily shocks of economic interest, as pointed out in the related literature (e.g., Kilian and Lütkepohl (2017, Chapter 14)). Further features of the shocks of interest must be taken into consideration to locate the shocks with economic interpretations in the set of identified shocks. Of course, in some cases one may not even expect all identified shocks to have economic interpretations. For example, Pagan and Robinson (2020) consider shocks that are due to measurement errors.

## 4 Conclusions

In some economic models the variables may be driven by more shocks than there are observable variables. In traditional structural VAR models such a situation is not foreseen. However, if there is heteroskedasticity, this feature can potentially be used to identify more shocks than there are variables. For specific types of heteroskedasticity, this study provides conditions for identification and it also presents upper bounds for the number of shocks that can be identified through heteroskedasticity. The conditions can be useful in applied work to assess, at an early stage, whether sufficient indentifying information can be obtained from heteroskedasticity to identify all potential shocks underlying a specific model setup.

## References

- Christiano, L. R., Motto, R. and Rostagno, M. (2014). Risk shocks, *American Economic Review* **104**: 27–65.
- Herwartz, H. and Lütkepohl, H. (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks, *Journal of Econometrics* **183**: 104–116.
- Ireland, P. N. (2011). A New Keynesian perspective on the great recession, *Journal of Money, Credit and Banking* **43**: 31–54.
- Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*, Cambridge University Press, Cambridge.
- Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching, *Journal of Economic Dynamics and Control* **34**: 121–131.
- Lütkepohl, H. (1996). *Handbook of Matrices*, John Wiley & Sons, Chichester.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.
- Lütkepohl, H., Meitz, M., Netšunajev, A. and Saikkonen, P. (2020). Testing identification via heteroskedasticity in structural vector autoregressive models, *Econometrics Journal* **23**.
- Pagan, A. and Robinson, T. (2020). Too many shocks spoil the interpretation, *Technical report*, University of Melbourne.

- Plagborg-Møller, M. and Wolf, C. K. (2017). Instrumental variable identification of dynamic variance decompositions, *Technical report*, Princeton University.
- Rothenberg, T. J. (1971). Identification in parametric models, *Econometrica* **39**: 577–791.
- Sims, C. A. (1980). Macroeconomics and reality, *Econometrica* **48**: 1–48.
- Smets, F. and Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach, *American Economic Review* **97**: 586–606.