’Structural vector autoregressions with Markov switching’ M.Lanne, H.Lütkepohl & K.Maciejowska (2010)

Pierre Rouillard

1. Summary of the paper

A large part of the Structural VAR framework analysis has to do with (orthogonal) structural shocks identification. Several approachs have been developed and discussed throughout the years, such as recursive identification (Sims (1980) [1] i.e imposing zero restrictions so that variables do not depend contemporaneously on the shocks ordered after), short and long-run restrictions (respectively zero restrictions: on a subset of shocks for specific variable(s) and on some coefficients of the long-run matrix). These identification schemes yield *exact* identification in the sense that a shock is uniquely identified through precise estimation of the matrix B_0 (see R1). However sign restrictions have also been discussed and in this case, we have a pool of plausible models and thus only *partial* identification. All these identification procedures usually rely on economic theory to justify restriction choices.

The framework that Lanne et al. [2] leverages, known as identification via heteroskedasticity, makes use of regimes with solely different variances for the structural shocks, which increases the number of estimated moments and is able to yield exact identification for B_0. Thus, changes in the volatility of the VAR errors (and hence the observed variables) can be used to assist in the precise identification of structural shocks. In other words, it extracts further information from the data to provide additional conditions needed for identification. Still, this result relies on the assumption that between regimes it is solely the variance of estimated residuals that changes so that the coefficient matrices are not time/regime dependent (in particular impulse response functions do not depend on the state, see R2). Thus comes the most important feature of this framework: any additional restriction on B_0 only over-identifies the model and can then be tested using standard tools[1]. The framework also allows for a comparion between *unrestricted* (Markov-switching based) and *restricted* (standard SVAR with restrictions) impulse response functions '*IRFs*' results. This makes for a strong contender to compare work already done in the litterature, especially papers that use recusrive identification assumptions in standard SVAR, as long as changes in volatility can be argued for.

In detail, an homogenous Markov structure is assumed on the residuals such that we have, in general, $\text{Var}(u_t|s_t) = \Sigma_{s_t}$ with $s_t \in \{1, \dots, M\}$ (M regimes) and transition matrix P. The authors further argue that the normality assumption $u_t|s_t \sim \mathcal{N}(0, \Sigma_{s_t})$ is,

[1] Wald test for equality, likelihood ratio '*LR*' test ...

in theory, not mandatory but comes in very handy to define the likelihood used in the estimation procedure. Moreover, they tackle the restrictiveness issue of such (conditional) normality hypothesis by illustrating that with constant rows for the transition matrix[2] P the model reduces to a system of usual SVAR with mixed normal errors and the probabilities that u_t follows each law are given by the first column of P. They argue such class of distributions is already much more diverse, complete and can accomodate more situations compared to the standard simple normality assumption.

Using the *simultaneous diagolanization of a positive semi definite matrix and symmetric matrices* from Lütkepohl[3](1996) we get that for a 2-regime system, the estimated variance matrices can be such that : $\Sigma_1 = B.B^T$ and $\Sigma_2 = B.\Gamma_2.B^T$ with B positive definite and Γ_2 a real diagonal matrix with nonnegative diagonal coefficients. Conditional on all coefficients of Γ_2 being distinct (which can be tested via usual Wald test), estimates of Σ_1, Σ_2 effectively yield $2.\frac{n(n+1)}{2} = n(n+1)$ moments which uniquely identifies the n^2 coefficients of B and the n diagonal coefficients of Γ_2. Otherwise if $\Sigma_1 \neq \Sigma_2$, but some coefficients are equal, we still identify Γ_2 but B is no longer unique! A similar decomposition for $M > 2$ regimes (see [4]) does not necessarily exist: it potentially needs to be assumed and tested via LR-test. Such test will then assess to what extent a constant matrix B can be used to diagonalize covariance matrices associated with all states. Still, the authors provide a result that ensures the uniqueness of B (up-to-sign): for any pair of diagonal coefficients, at least one of the matrices of $\{\Gamma_i\}$ is such that these coefficients are distinct. Regardless of the number of regimes, the resulting shocks $u_t = B.e_t$ are orthogonal in all states. In this framework the Markov process only affects the covariance matrix, and it is precisely these differences in error covariance matrices that provide sufficient conditions (from the data alone) to uniquely identify shocks without diverging much from the standard SVAR framework.

The estimation procedure consists of computing the maximum or pseudo-maximum likelihood estimators for the parameters (depending on assuming conditional normality or not). Regarding the optimization solution in itself of such likelihood, they simply mention using an extension of the EM algorithm of Dempster, Laird and Rubin (1977) for the specific case of hidden-Markov chain parametrized covariance matrices but do not provide much detail. In addition, the usual hypothesis underlying the SVAR framework relies on stationarity of the processes but in case some variables are only

[2] Constant rows for the transition matrix = same probability to switch between all states and yields independancy of current state relative to previous date state

[3] *Handbook of Matrices* H. Lütkepohl (page 86) - Humboldt-Universität zu Berlin Germany

cointegrated, and especially if the cointegration relationship is unknown, the authors provide a sketch of the stance to adopt. Still they mention such problems remain to be tackled and demonstrated by further research.

Finally, they make use of the strong over-identification test feature of the MS-SVAR framework to assess the identification restrictions used in two papers on monetary policy analysis. They also compare their unrestricted IRFs to the articles' results. Primiceri (2005) introduces a three-variable model (inflation rate, unemployment rate and three-month treasuries yield) for the US and procedes to identify shocks via standard recursive identification. The justification of such identification scheme hinges on the assumption that monetary policy shocks do not contemporaneously affect economic variables: the effect is at least one-period delayed. Using a 2-regime MS-VAR model, Lanne et al. show that structural shocks can be fully identified as diagonal coefficients of the estimated Γ_2 matrix are sufficiently distinct (assessment via Wald test). The recursive identification -now an over-indentifying assumption- with underlying hypothesis that B_0 is a lower triangular matrix can then be tested via LR test and it appears this null hypothesis is rejected. Precisely, after identifying the monetary shock amidst the three estimated shocks, it appears B_0 has a negative and significant coefficient on the contemporaneous reaction of the unemployment rate to a monetary policy shock, contrary to the inital assumption made by Primiceri. However, rejecting the lower-triangularity assumption in itself is not a necessary result that invalidates Primiceri's conclusions on the variables reactions to shocks. This is rather assessed by comparing the restricted and unrestricted IRFs, and it appears responses to a monetary policy shock derived from the unrestricted MS-VAR model are stronger (in absolute value) in the short-run. Robustness checks were performed and do not change this result.

The Sims et al. (2008) three-variable (log GDP, inflation and short-term rate) model is also assessed. Despite initial estimation done via MS-VAR, shocks are not identified via heteroskedasticity changes but rather via again recursive identification. Following similar assessment to Primiceri's paper Lanne et al. show strong evidence for a uniquely identified matrix B_0. As the model is this time composed of three regimes, the uniqueness result is derived from the extended result [4]. A likelihood ratio test on the lower-triangularity hypothesis can then be ran on the estimated B_0 and it is rejected at 10% significance level but not at 5%. The *simultaneous diagolanization* decomposition, which does not necessarily exist for three or more regimes, is also accounted for and tested via LR test. The test amounts to checking that all matrices $B_0^{-1}.\Sigma_i.B^T$ are diagonal[4]. The null hypothesis is not rejected and combined with little evidence to reject the lower-triangularity hypothesis, Lanne et al. results from the MS-SVAR framework, in this case, strengthens the confidence in the recursive identification scheme. Overall the MS-VAR framework appears to be the perfect fit to test restrictions which are used as *'necessary-identification-conditions'* in standard SVAR.

2. Literature review

Herwartz et Lütkepohl (2014) [3] directly build on the work of Lanne et al. to address the issue of how standard identifying assumption can complete statistical information. Indeed, shocks identified purely through the prism of statistical properties (as it is done with MS-SVAR) may not end up deing economically relevant. In such case we effectively extract more information from the data but we cannot economically interpret the results. These issues are discussed through a quarterly model for the US for oil prices, output, price level and short-term interest rate. They provide more detailed statistical guidelines regarding model selection and interpretation compared to Lanne et al. Information criterion are used to compare different unrestricted models to the restricted and their interpretation provide support from the data for the final choice. Regarding the labeling of shocks it is mainly done through volatility analysis in the different regimes: state probabilities are interpreted in light of historical economic knowledge. Thus, shocks are identified by regime-specific volatilities in light of economic events/periods. Noteworthy additions of Herwartz et Lütkepohl are also on issues with standard residual based bootstrap methods for the optimization of the log-likelihood and especially a thorough description of the adapted EM algorithm when covariance matrices are parametrized by a Markov switching process (which was not present in Lanne et al.).

On the IRF litterature in the context of MS-VAR models, earlier work focused on computing IRFs for specific simple specifications (i.e with only the intercept being regime-dependent). In 2003 Ehrmann et al. [4] proposed a procedure to compute a *regime-dependent impulse response function* in the general framework of MS-VAR i.e with regime-dependent coefficients and covariance matrices. Parameters of the model and the markov chain are estimated by standard EM algorithm procedure. With estimated parameters, the identification problem needs to be dealt with and they use restrictions assumptions from the standard SVAR literature. However, the derived IRFs are

[4]The alternative hypothesis is then: at least one is not diagonal, and in this case the *simultaneous diagolanization* decomposition hypothesis is rejected.

only modelisation of the system response *conditional* on the regime state at which the disturbance occurs. The modeled reaction thus corresponds to the *within-regime* system response (assumes the system stays indefinitely in this regime), but not the unconditional (general) reponse. To that end Karamé (2010) [5] makes a step further and proposes more general IRFs corresponding to the system general reaction following a shock, encompassing the unconditional response. Cavicchioli (2023) [6] proposes a continuation of such work by unifying results under a matrix representation framework. A complete breakdown of assumptions and theorems underlying the IRFs derived from general MS-VAR model, extending initial results from standard VAR of Hamilton, is also provided.

Empirical applications on financial and commodity market topics dwell further by using general MS-VAR models. For instance, Sharestani et Rafei (2020) [7] focus on the case of oil price shocks on the Iranian stock exchange. The (broader) MS-SVAR framework, is mainly considered to assess varying persistence effect of oil price shocks on Iranian financial markets between 2 regimes. The number of states choice was adressed through information criteria. Hou et Nguyen (2018) [8] attempt to provide more insights regarding the structure of the US natural gas market using the MS-VAR framework. Specifically, the markov chain structure allows for regime recurrence which could be more attractive to study commodity markets and understand links between the different type of shocks compared to standard structural break models. Note that they use the general MS-VAR with state-specific coefficients and covariance matrices but compute the IRFs *within* a regime. Infering results from such IRFs would assume the system indeed stays for a long time in each regime, and their results show is it most likely the case as almost all regimes span a considerable length of time.

The MS-VAR framework has been extended to account for mixed-frequency data, a topic generally discussed in the context of Dynamic Factor Models (see Stock & Watson 2002). One extension of the standard VAR framework to mixed frequency data (MF-VAR) was proposed by Mariano et Murasawa (2010) [9]. The main idea is to use a latent monthly (unobserved) underlying variable for the the quarterly series (usually a geometric mean of the latent variable). Assuming a standard VAR model on the vector composed of the latent variable and the other observed monthly series, and using it to define a state vector[5] a state-space representation of the VAR can be derived. The original vector, in the measurement equation, composed of quarterly and monthly data is still only observed every third period. After subtule transformations (namely handling missing values as random draws whose realisations were all zeros) and variable definitions, the measurement equation can be rewritten and a monthly time-varying space state model *without* missing values obtained. It is estimated using the EM algorithm and the Kalman smoother is then used with the filtered value considered as a new coincident index (with appropriate transformation) to be compared with the Stock-Watson Experimental Coincident Index and the NBER recession dates. Camacho (2013) [10] builds on this work and considers the extension where the assumed VAR process for the state vector is now a Markov-Switching VAR. The computed monthly estimates of GDP growth for the US by Mariano et Murasawa is then augmented by the Markov-Switching version. Foroni et al. (2015) [11] also consider the Markov-Switching extension of Mariano et Murasawa (2010) but, amongst other things, compare it with the regime switching version of the Mixed-Frequency VAR of Ghysels (2012) [12]. The latter differs from the work of Mariano et Murasawa as it does not hinge on latent variables, rather it leverages a specific representation of the data: a *Mixed and Periodic Stacked VAR* representation. The main takeaway from this representation is that the model can be estimated using regular tools, namely with a maximum likelihood estimation which is much less computational intensive than Kalman filtering.

3. Other comments

Overall I believe the paper is really informative and clear regarding shocks identification using heteroskedasticity changes, but it does require previous knowledge about structural VARs in general to fully apprehend the subtility of identification without just-identifying assumptions (recursive identification, sign restrictions etc...).

For my own understanding there are a few points that I would have prefered to be more clear. Stating the stationarity assumption and addressing cointegration issues from the start (in the general setup definition for instance) would, in my opinion, allow the authors to remove any possible ambiguity rather than only mentioning it at the end of the estimation part. Properties of the markov process are not really addressed, whereas papers such as Foroni et al [11] do state that the Markov chain is aperiodic and irreducible, standard assumptions that we saw during the course for Markov-Switching types of models. I understand that the computational and estimation part was not the main addressing point of the paper but as I still lack familiarity with the EM algorithm I felt Herwartz et Lütkepohl [3] gave more insights about what is really happening.

[5] The exact definition depends on the VAR order

The empirical part is really informative and well detailed to assess the uniqueness result for the covariance matrix decomposition. The testing procedures and explanations for their interpretation make the model much more illustrated. And with these examples we can fully grasp to what extent the possibility to test the just-identifying assumptions used in the literature make such model innovative and valuable. I believe it was to remain as close as possible to the two original papers used for illustration, but the discussions from both Herwartz et Lütkepohl [3] and Sharestani et Rafei [7] on the lag order selection for the VAR and the use of information criteria to select the number of regimes provided more insights on how to really build a MS-VAR from scratch. In the same vein, the issue of the economic interpretation of the exactly indentified structural shocks is quickly looked over in Lanne et al but more addressed again in Herwartz et Lütkepohl [3]. Still, it does not reduce the quality and interest of the paper as I believe it serves it purpose pretty well: the MS-VAR framework is nicely summarized, a sufficient condition ensuring exact shock identification is provided and empirical illustrations on two models from the associated literature clearly highlight the strenghs of such model.

I was already interested in Structural VAR models after taking the course of Ricco Giovanni on the topic, but this project allowed me dwell more in the literature and further understand underlying constrains of just-indentifying assumptions.

4. Notes

I use usual notations and equations of standard Structural VAR, mainly derived from the 2023/24 ENSAE course '*Advanced Time-Series Analysis*' of Ricco Giovanni :

$$\begin{aligned}
\text{SVMA}(\infty): &\quad Y_t = \mu + F_0 u_t + F_1 u_{t-1} + F_2 u_{t-2} + (\dots) \quad u_t \sim WN(0, I_n)^6 \\
\text{SVAR}(\infty)^7: &\quad B_0 Y_t = K + B_1 Y_{t-1} + B_2 Y_{t-2} + (\dots) + u_t \\
\text{SVAR(p)}: &\quad B_0 Y_t = K + B_1 Y_{t-1} + B_2 Y_{t-2} + (\dots) + B_p Y_{t-p} + u_t \quad u_t \sim WN(0, I_n) \\
\text{Estimated VAR(p)}: &\quad Y_t = C + A_1 Y_{t-1} + A_2 Y_{t-2} + (\dots) + A_p Y_{t-p} + e_t \quad e_t \sim WN(0, \Sigma) \\
\text{(R1)} \qquad &\Rightarrow u_t = B_0 . e_t
\end{aligned}$$

The last equation highlights the need to identify B_0 (n^2 elements matrix) to get the structual shocks u_t from the reduced-form (estimated) residuals e_t in the usual framework. As $\Sigma = B_0^{-1}.(B_0^{-1})^T$ and Σ is symmetric, this yields $\frac{n(n+1)}{2}$ moments. We then need $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$ other restrictions if we want to fully identify B_0 in the standard SVAR framework.

Usual Impulse Response Function for Structural VARs are obtained through the approximation of the Wold decomposition by inverting the estimated finite-order VAR(p) and with $\{C_j\}_j$ the coefficients of the Wold representation we have :

$$Y_t = \mu + C_0 u_t + C_1.B_0^{-1} u_{t-1} + C_1.B_0^{-1} u_{t-2} + (\dots)$$

$$\text{(R2)} \qquad \Rightarrow \text{IRF}_{t+j} = \frac{\partial Y_{t+j}}{\partial u_t} = C_j . B_0^{-1}$$

In case the coefficient matrices from the estimated VAR are not regime-dependent the $\{C_j\}_j$ from the Wold decomposition won't be either. [R2] also shows that different estimations of B_0 will be reflected in different IRFs, and thus different reponse profiles of the system following shocks.

[4] *Simultaneous diagolanization of a positive semi definite matrix and symmetric matrices* for $M > 2$ regimes: $\Sigma_1 = B.B^T$ and $\Sigma_i = B.\Gamma_i.B^T$ with B positive definite and Γ_i a diagonal matrix.

[7] '*WN*' white noise

[7] Obtained by inverting the infinite moving average form SVMA(∞)

References

[1] Cristopher A. Sims. Macroeconomics and reality. Econometrica, 48(1):1–48, 1980.

[2] Markku Lanne, Helmut Lütkepohl, and Katarzyna Maciejowska. Structural vector autoregressions with markov switching. Journal of Economic Dynamics and Control, 34(2):121–131, 2010.

[3] Helmut Herwartz and Helmut Lütkepohl. Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks. Journal of Econometrics, 183(1):104–116, 2014.

[4] Natacha Valla Michael Ehrmann, Martin Ellison. Regime-dependent impulse response functions in a markov-switching vector autoregressive model. Economics Letters, 78:295–299, 2003.

[5] Frederic Karamé. Impulse-response functions in markov-switching structural vector autoregressions: A step further. Economics Letters, 106(3):162–165, 2010.

[6] Maddalena Cavicchioli. Impulse response function analysis for markov switching var models. Economics Letters, 232:111357, 2023.

[7] Parnia Shahrestani and Meysam Rafei. The impact of oil price shocks on tehran stock exchange returns: Application of the markov switching vector autoregressive models. Resources Policy, 65:101579, 2020.

[8] Bao H. Nguyen Chenghan Hou. Understanding the us natural gas market : A markov switching var approach. Energy Economics, 75:42–53, 2018.

[9] Yasutomo Murasawa Roberto S. Mariano. A coincident index, common factors, and monthly real gdp. Oxford Bulletin of Economics and Statistics, 72(1):27–46, 2010.

[10] Maximo Camacho. Mixed-frequency var models with markov-switching dynamics. Economics Letters, 121(3):369–373, 2013.

[11] Claudia Foroni, Pierre Guérin, and Massimiliano Marcellino. Markov-switching mixed-frequency var models. International Journal of Forecasting, 31(3):692–711, 2015.

[12] Eric Ghysels. Mixed frequency vector autoregressive models. University of North Carolina, 2012.