Structural Vector Autoregressions with Heteroskedasticity: A Comparison of Different Volatility Models

Helmut Lütkepohl*
Aleksei Netšunajev**

*DIW Berlin and Freie Universität Berlin, Germany
**Freie Universität Berlin, Germany

Structural Vector Autoregressions with Heteroskedasticity

A Comparison of Different Volatility Models[*]

Helmut Lütkepohl

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

Aleksei Netšunajev

Freie Universität Berlin, Fachbereich Wirtschaftswissenschaft
Boltzmannstr. 20, 14195 Berlin, Germany
email: Aleksei.Netsunajev@fu-berlin.de

March 30, 2015

Abstract. A growing literature uses changes in residual volatility for identifying structural shocks in vector autoregressive (VAR) analysis. A number of different models for heteroskedasticity or conditional heteroskedasticity are proposed and used in applications in this context. This study reviews the different volatility models and points out their advantages and drawbacks. It thereby enables researchers wishing to use identification of structural VAR models via heteroskedasticity to make a more informed choice of a suitable model for a specific empirical analysis. An application investigating the interaction between U.S. monetary policy and the stock market is used to illustrate the related issues.

Key Words: Structural vector autoregression, identification via heteroskedasticity, conditional heteroskedasticity, smooth transition, Markov switching, GARCH

JEL classification: C32

---

1 Introduction

Structural vector autoregressive (SVAR) models are typically identified by exclusion restrictions on the impact effects of the structural shocks (e.g., Sims (1980), Bernanke and Mihov (1998), Kilian (2009)), by restrictions on the long-run effects of the shocks (e.g, Blanchard and Quah (1989), King, Plosser, Stock and Watson (1991)) or by restrictions on the signs of the responses of specific variables to a shock (e.g., Faust (1998), Canova and De Nicoló (2002), Uhlig (2005)). Recently, a number of articles uses changes in the volatility of the variables for identification of SVAR models (e.g., Rigobon (2003)). A number of alternative approaches have been proposed for modelling the changes in volatility. For example, exogenous changes in the variances of the residuals are considered by Rigobon (2003), Rigobon and Sack (2003) and Lanne and Lütkepohl (2008). In contrast, Lanne, Lütkepohl and Maciejowska (2010) and Herwartz and Lütkepohl (2014), for example, model the changes in volatility by a Markov regime switching (MS) mechanism whereas Lütkepohl and Netšunajev (2014b) consider smooth transition of the residual volatility from one regime to another. Yet another approach is based on a generalized autoregressive conditional heteroskedastic (GARCH) error structure (e.g., Normandin and Phaneuf (2004), Bouakez and Normandin (2010), Weber (2010), Strohsal and Weber (2012)). Typically in this literature there is not much discussion why a particular volatility model is used.

The objective of this study is to compare the alternative approaches and discuss their advantages and drawbacks. Thereby we hope to provide a basis for making a rational decision which model to use for a particular application. Specifically we discuss the model setup, identification conditions for the shocks, estimation and inference methods related to the alternative models. As far as estimation and inference is concerned, we focus on frequentist procedures because Bayesian procedures have not been applied much so far in this context (see, however, Kulikov and Netšunajev (2013)). Hence, there is not much experience with Bayesian methods. For some of the models they are not yet fully developed.

We illustrate the different approaches in the SVAR analysis based on the study by Bjørnland and Leitemo (2009). These authors use conventional exclusion restrictions on the impact and long-run effects of the shocks in studying the interaction between monetary policy and the stock market in the U.S.. We demonstrate how the different volatility models can be used for generating additional identifying information and we discuss the advantages and drawbacks of the alternative models for the example system.

Earlier related reviews of some of the models discussed in the following are given by Lütkepohl (2013) and Lütkepohl and Velinov (2015). They

discuss a more limited set of volatility models and different, more limited examples. In particular, they do not compare the full range of volatility models in the context of a specific example. The latter of the two articles focuses specifically on combining restrictions on the long-run effects of the shocks with identifying information from changes in volatility. That topic is part of our general setup and, hence, it is included in the present study as a special case.

The remainder of this study is organized as follows. In the next section the basic SVAR model is introduced and the different volatility models are discussed in Section 3. The illustrative example is considered in Section 4 and concluding remarks are given in the final section.

2 The Baseline Model

The baseline model is a VAR of order p (VAR(p)) of the form

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

where $y_t = (y_{1t}, \ldots, y_{Kt})'$ is a vector of observable variables, the A_i are $(K \times K)$ coefficient matrices, ν is a $(K \times 1)$ constant term and the u_t are K-dimensional serially uncorrelated reduced form residuals with mean zero and nonsingular covariance matrix Σ_u.

The structural residuals are denoted by ε_t. They have zero mean and are serially uncorrelated. Typically they are also assumed to be instantaneously uncorrelated, that is, $\varepsilon_t \sim (0, \Sigma_\varepsilon)$, where Σ_ε is a diagonal matrix. Sometimes it is convenient to assume that the variances of the structural shocks are normalized to one so that Σ_ε is an identity matrix.

The structural residuals are typically obtained from the reduced form residuals by a linear transformation:

$$\varepsilon_t = B^{-1} u_t \quad \text{or} \quad u_t = B \varepsilon_t. \tag{2}$$

The matrix B contains the instantaneous effects of the structural shocks on the observed variables. Given the relation between the reduced form residuals and the structural residuals, the matrix B has to satisfy $\Sigma_u = B \Sigma_\varepsilon B'$. In other words, in principle B can be any matrix satisfying $\Sigma_u = B \Sigma_\varepsilon B'$. The relation between the reduced form and structural residuals does not uniquely determine the matrix B and, hence, the structural innovations are not uniquely determined without further assumptions.

The conventional approach is to impose further restrictions on B or its inverse directly to make it unique. These restrictions may be zero restrictions

2

indicating that a certain shock does not have an instantaneous impact on one of the variables (Sims (1980)) or it may be implied by a restriction on the long-run effects of a structural shock as in Blanchard and Quah (1989). In the latter approach the matrix of long-run effects of structural shocks is given by

$$\Xi_\infty = (I_K - A_1 - \cdots - A_p)^{-1} B,$$

assuming that the inverse exists. That condition is satisfied for stable, stationary processes without unit roots. For integrated and cointegrated processes the long-run effects matrix is related to the cointegration structure of the model (see, e.g., Lütkepohl (2005) or Lütkepohl and Velinov (2015)). For our purposes it is sufficient to know that the matrix of long-run effects can be computed from the reduced form and structural parameters. Imposing restrictions on the matrix of long-run effects implies restrictions on B for a given data generation process (DGP) and, hence, a given reduced form.

Typically the restrictions on B just-identify the structural model and, hence, the structural shocks. In other words, there are just enough restrictions for uniqueness of B and no more. If there are two competing sets of just-identifying assumptions or theories implying just-identifying restrictions, they lead to identical reduced forms and cannot be tested against the data. Hence, the conventional setup is often uninformative regarding the validity of specific economic theories. In the next section it is discussed how heteroskedasticity can be used to improve the situation in this case.

3 SVAR Models with Time-Varying Volatility

Suppose now that u_t is a heteroskedastic or conditionally heteroskedastic error term. This means that the variances or conditional variances of the reduced form and, hence, also the structural innovations change over time. In the macroeconomic literature such behaviour is well documented. Allowing for this feature means that there is more than one volatility regime present during the sample period and each regime is characterized by a covariance matrix $\Sigma_u(m)$, say. Suppose that there are M such regimes. If the corresponding covariance matrices can be decomposed as

$$\Sigma_u(1) = BB', \quad \Sigma_u(m) = B\Lambda_m B', \quad m = 2, \ldots, M, \tag{3}$$

with $\Lambda_m = \mathrm{diag}(\lambda_{m1}, \ldots, \lambda_{mK})$, then the structural shocks obtained as in (2) satisfy the basic condition that they are instantaneously uncorrelated in all

volatility regimes. The diagonal matrices Λ_m are the matrices of variances of structural shocks in regime m relative to regime 1. For convenience Λ_1 is chosen to be the identity matrix. In other words, the structural shocks are normalized to have unit variances in the first volatility regime.

In this setup the matrix of impact effects of the shocks, B, is (locally) uniquely determined by the decomposition in (3) if for any subscripts $k, l \in \{1, \ldots, K\}$, $k \neq l$, there is a $j \in \{2, \ldots, M\}$ such that $\lambda_{jk} \neq \lambda_{jl}$ (Lanne et al., 2010, Proposition 1). In other words, for any two subscripts k and l there must be at least one regime, where the corresponding relative variances are different. For example, for the case of just two different volatility states with covariance matrices

$$\Sigma_u(1) = BB', \quad \Sigma_u(2) = B\Lambda_2 B'$$

all diagonal elements of Λ_2 have to be distinct. For more than two states,

$$\lambda_{2k} = \lambda_{2l}, \ldots, \lambda_{Mk} = \lambda_{Ml} \tag{4}$$

must not hold for any combination k, l with $k \neq l$. In other words, local uniqueness is ensured if there is sufficient heterogeneity in the volatility changes. For instance, it is not enough that all variances change in proportion.

Here local uniqueness refers to the fact that the matrix B is unique only up to changes in the signs of its columns and up to permutations of its columns. In other words, the matrix B is unique if the structural shocks are chosen as $\varepsilon_t = B^{-1} u_t$ and the signs and ordering of the shocks are fixed. Thus, allowing for multiple volatility regimes the assumption of regime invariant impact effects of the structural shocks, i.e., choosing $\varepsilon_t = B^{-1} u_t$ with a time-invariant transformation matrix implies uniqueness or identification without imposing additional economic restrictions. In the context of SVARs with heteroskedasticity the identification is shifted from imposing the just-identifying exclusion or long-run restrictions on B towards restrictions implied by the statistical model setup. These conditions have to be met in order to obtain exact (local) identification via changes in volatility. Once the statistical identification is achieved, additional economic restrictions can be imposed and can be tested formally as over-identifying restrictions.

There are several ways to model the change in volatility. Initially the change points were just specified exogenously by the analyst, possibly based on some prior statistical analysis (e.g., Rigobon (2003), Lanne and Lütkepohl (2008), Bacchiocchi and Fanelli (2012)). Other approaches let the data decide endogenously on the change points and possibly also the number of volatility regimes. This approach is used when modelling the changes in volatility

with a smooth transition specification (Lütkepohl and Netšunajev (2014b)), via a Markov switching process (e.g., Lanne et al. (2010)), or based on a generalized autoregressive conditional heteroskedasticity (GARCH) process (Normandin and Phaneuf (2004)). These approaches allow even for infinitly many volatility regimes. They are presented and discussed in more detail in the following sections.

3.1 Exogenously Specified Volatility Changes

A simple model for volatility changes is based on the assumption that the change points are exogenously given, that is,

$$\mathbb{E}(u_t u_t') = \Sigma_{u,t} = \Sigma_u(m) \quad \text{for} \quad t \in \mathcal{T}_m, \quad m = 1, \ldots, M, \tag{5}$$

where $\mathcal{T}_m = \{T_{m-1} + 1, \ldots, T_m\}$ ($m = 1, \ldots, M$) are M given volatility regimes usually consisting of consecutive time periods. The T_m, for $m = 1, \ldots, M - 1$, represent the time periods of volatility changes. The initial $T_0 = 0$ and $T_M = T$. The T_m are assumed to be known to the analyst and are, hence, taken as given in the analysis. In practice, they may be obtained with some prior statistical procedure, for instance, a testing procedure for changes in the residual covariance matrix.

Estimation of the model is straightforward. Under normality assumptions, the likelihood function can be set up and ML estimation can be used (see, e.g., Lütkepohl (2013)). If the data are not normally distributed, the estimators are quasi ML estimators with standard asymptotic properties. Since the likelihood function is nonlinear one may also want to consider a generalized least squares procedure instead of full ML.

To estimate the structural parameters, $\Sigma_u(1)$ has to be replaced by BB' and $B\Lambda_m B'$ replaces $\Sigma_u(m)$ ($m = 2, \ldots, M$) in the likelihood function. Note, however, that for $M > 2$ this imposes a restriction on the model that can be tested by a likelihood ratio (LR) test. Under Gaussian assumptions and the null hypothesis of the restriction being correct, the LR statistic has an asymptotic χ^2 distribution with

$$\frac{1}{2} MK(K + 1) - K^2 - (M - 1)K$$

degrees of freedom (see Lanne et al. (2010)). If the null hypothesis is rejected, the assumption of a time-invariant initial effects matrix B across all volatility regimes is not in line with the data. In that case, one could consider the possibility of a time-invariant initial effects matrix across some of the regimes.

If the decomposition of the covariance matrices is supported by the data, then the conditions for local uniqueness of B can be tested. In other words, using (4), null hypotheses

$$H_0 : \lambda_{2k} = \lambda_{2l}, \ldots, \lambda_{Mk} = \lambda_{Ml} \tag{6}$$

have to be tested for all pairs $k, l \in \{1, \ldots, K\}$ with $k \neq l$. These are conditions on the diagonal elements of the Λ_m matrices. They are identified if they are ordered uniquely even if B is not identified. For example, one may order the diagonal elements of Λ_2 from smallest to largest, if all elements are distinct. If some of the elements are equal, the ordering can be based on Λ_3 etc.. In any case, their asymptotic distribution can be used for testing the identifying conditions. Wald tests based on the asymptotic distribution of the λ_{mk} can be used for that purpose (see Herwartz and Lütkepohl (2014)). Alternatively, under Gaussian assumptions, LR tests can be considered (Lütkepohl and Netšunajev (2014a)).

Another issue is the determination of the volatility change points. As mentioned previously, in practice this is often done by prior statistical procedures (e.g., Ehrmann, Fratzscher and Rigobon (2011)). Again standard procedures, e.g., Chow tests can be used for this purpose. Thus, overall the statistical procedures related to the model are straightforward and computationally feasible.

Applications of the model are due to Rigobon (2003), Rigobon and Sack (2003, 2004), Lanne and Lütkepohl (2008) and Ehrmann et al. (2011), for example. While the straightforward estimation and inference procedures associated with the model make it attractive for use in practice, the fact that the volatility change points have to be prespecified is a disadvantage. Clearly, it is often plausible to assume that the volatility changes are endogenously generated. Therefore letting the data determine the volatility process including its possible change points is more appealing.

3.2 SVAR with Smooth Transition in Variances

One possible model with endogenously changing volatility assumes a smooth change in the residual covariance matrix. It was proposed in this context by Lütkepohl and Netšunajev (2014b). More precisely, the change in the covariance structure is modelled as a smooth transition from a volatility regime characterized by a positive definite covariance matrix $\Sigma_u(1)$ to a regime with a different positive definite covariance matrix $\Sigma_u(2)$. The transition is described by a smooth transition function $G(\gamma, c, s_t)$ that depends on parameters γ and c as well as a transition variable s_t. Lütkepohl and Netšunajev

(2014b) use a logistic function with transition variable $s_t = t$, that is,

$$G(\gamma, c, s_t) = (1 + \exp[-\exp(\gamma)(t - c)])^{-1}. \tag{7}$$

Because $\exp(\gamma) > 0$ for positive and negative values of γ, $G(\gamma, c, s_t) > 0$ is close to zero when t is much smaller than c and $G(\gamma, c, s_t) \to 1$ for $t \to \infty$. Formally, the reduced form residual covariance is described as

$$\mathbb{E}(u_t u_t') = \Sigma_{u,t} = (1 - G(\gamma, c, s_t))\Sigma_u(1) + G(\gamma, c, s_t)\Sigma_u(2). \tag{8}$$

This setup ensures a positive definite matrix $\Sigma_{u,t}$ because this matrix is a convex combination of two positive definite matrices. Notice that due to the covariance change during the transition period there are in fact infinitely many different covariance regimes.

The transition from the residual covariance matrix $\Sigma_u(1)$ to $\Sigma_u(2)$ can be used for identification purposes, using the decomposition

$$\Sigma_u(1) = BB' \text{ and } \Sigma_u(2) = B\Lambda_2 B', \tag{9}$$

where $\Lambda_2 = \mathrm{diag}(\lambda_{21}, \ldots, \lambda_{2K})$ is a diagonal matrix with positive diagonal elements as before. Apart from changing the signs of the columns of B this decomposition is unique for a given ordering of the diagonal elements of Λ_2 if the diagonal elements are all pairwise distinct (Lanne and Lütkepohl (2008)). Thus, if the B matrix from (9) is used to transform the reduced form residuals, the covariance matrices of the structural shocks are I_K and Λ_2 for the initial and final regimes, respectively. The diagonal elements of the Λ_2 matrix can thus be interpreted as variances of structural shocks in the final regime relative to the initial regime.

Under the assumption of normality the likelihood function of the model can be set up and ML estimation can be performed using standard numerical non-linear optimization algorithms. An iterative procedure for estimation is discussed in detail by Lütkepohl and Netšunajev (2014b). Since the range of the smoothness and threshold parameters $\{\gamma, c\}$ can be bounded, a grid search can be performed over the space of these two parameters. Thus, estimation is technically straightforward and the asymptotic properties of the parameter estimators are standard under usual assumptions for the DGP.

An important advantage of using this setup is that the identification condition based on the diagonal elements of Λ_2 becomes testable. Lütkepohl and Netšunajev (2014b) use Wald tests for this purpose. In a related context Lütkepohl and Netšunajev (2014a) use likelihood ratio tests. If the diagonal elements of Λ_2 are all distinct, any restrictions imposed on B in a conventional SVAR framework become over-identifying and can be tested against the data.

Apart from the straightforward estimation and inference procedures of the model it has also the advantage of transparency of the mechanism of volatility change. This, however, may turn out to be a disadvantage because it may not be very realistic in many macroeconomic applications. A simple move from one volatility regime to another during the sample period is clearly a very special situation. Of course, the model can be extended by including further transition terms so that the volatility can move between more states. Also one could consider other transition variables or transition functions that allow for more movement between two volatility states. If such volatility changes are suspected it is a question, however, whether alternative specifications may be more suitable for capturing them. One such alternative specification is a Markov switching mechanism that drives the volatility changes. It is presented next.

3.3 SVARs with Markov Switching in Variances

Lanne et al. (2010) propose to use a Markov regime switching mechanisms in the VAR residuals for modelling the volatility changes and identifying structural shocks. The related statistical methodology for using it in the SVAR framework is partly developed by Herwartz and Lütkepohl (2014). In this approach the distribution of the error term u_t is assumed to depend on a discrete Markov process s_t $(t = 0, \pm 1, \pm 2, \dots)$ with states $1, \dots, M$, and transition probabilities

$$p_{ij} = \Pr(s_t = j | s_{t-1} = i), \quad i, j = 1, \dots, M.$$

The conditional distribution of u_t given s_t is assumed to have a covariance matrix that is different across states,

$$u_t | s_t \sim (0, \Sigma_u(s_t)). \tag{10}$$

The model allows for Markov switching in the residual covariances only and not in the VAR parameters. It captures conditional heteroskedasticity. Notice that there are finitely many Markov states only. Since the model assigns probabilities to these states in each period, there are actually infinitely many possible conditional covariance matrices. Thus, the model can also capture smooth transitions from one state to another because a particular state does not necessarily come up with probability one but the system may be in between states. In other words, the actual volatility may be described by a mixture of different states, each state being weighted by a certain probability in each period.

Structural identification is obtained in this setting by using the covariance decomposition

$$\Sigma_u(1) = BB', \quad \Sigma_u(m) = B\Lambda_m B', \quad m = 2,\ldots,M, \tag{11}$$

as in the model with exogenously determined changes in volatility. The conditions are the same as those discussed in Section 3.1.

Estimation of the model can be done by ML if distributional assumptions are made. Lanne et al. (2010) and Herwartz and Lütkepohl (2014) assume a normal conditional distribution of u_t, i.e., $u_t|s_t \sim \mathcal{N}(0, \Sigma_u(s_t))$, set up the likelihood function accordingly and present a suitable optimization algorithm. The actual optimization task is difficult for larger models. There are several obstacles that make estimation difficult. First, the likelihood function is actually unbounded so that strictly speaking there is no maximum. The problem can be overcome by ensuring that all covariance matrices are positive definite and their eigenvalues are bounded away from zero. Second, the likelihood function is highly nonlinear and has local optima. This problem can be alleviated by using a large number of different starting values for the optimization algorithm. Third, there is the well-known label-switching problem which arises if the ordering of the regimes changes in the course of the optimization. These obstacles make it difficult to compute ML estimates for models with many parameters, that is, for models with many volatility regimes, many variables or large lag orders. It may be noteworthy that overcoming these problems is also not straightforward in a Bayesian approach.

If ML estimates can be obtained they have standard asymptotic properties under suitable assumptions. In fact, the normality assumption may be relaxed. Asymptotic properties can also be derived if the true distribution is different so that ML based on the conditionally Gaussian distribution is actually a quasi ML procedure. Note however, that assuming a Gaussian conditional distribution implies a nonnormal unconditional distribution of the data. Hence, the assumption may make sense even for financial data, for example.

In any case, the difficult numerical computation of the ML estimates is a drawback of the model when systems with larger numbers of variables are of interest. It also makes it difficult to use bootstrap techniques that require estimation of a large number of models. Herwartz and Lütkepohl (2014) propose a fixed design wild bootstrap method that conditions on some of the estimated parameters. Little is known about the actual properties of the method.

Applications of the MS-SVAR model are available in a number of studies. Examples are Lanne et al. (2010), Herwartz and Lütkepohl (2014), Lütkepohl

and Netšunajev (2014a), Netšunajev (2013), Lütkepohl and Velinov (2015), Velinov (2013).

The model has appeal because it lets the data assign the observations to the different volatility regimes. The regimes obtained in this way often have a clear interpretation. For example, they may be linked to the business cycle or other economically relevant features or periods of particular interest from the point of view of economics. Identification conditions are available that can be checked easily. If the model is small enough to compute Gaussian ML estimates easily, then standard procedures for inference can be used straightforwardly. The drawback of the model is, however, that for larger models with long lag orders, a larger set of variables or many different volatility regimes estimation is difficult and potentially unreliable. Moreover, even for small models that can be estimated easily, little is known about the properties of the proposed methods for constructing confidence bands around impulse responses, for example.

3.4 SVARs with GARCH

Suppose now that the reduced form errors u_t follow a multivariate GARCH process. The information available up to and including period t is contained in the set $\mathcal{F}_t$ and the conditional covariance matrix of u_t given $\mathcal{F}_{t-1}$ is assumed to be of the form

$$\mathbb{E}(u_t u_t' | \mathcal{F}_{t-1}) = B\Sigma_{\varepsilon,t|t-1}B', \tag{12}$$

where $\Sigma_{\varepsilon,t|t-1} = \mathrm{diag}(\sigma^2_{1,t|t-1}, \ldots, \sigma^2_{K,t|t-1})$ is a diagonal matrix. The individual variances of structural shocks are assumed to have a GARCH(1,1) structure:

$$\sigma^2_{k,t|t-1} = (1 - \gamma_k - g_k) + \gamma_k \varepsilon^2_{k,t-j} + g_k \sigma^2_{k,t-j|t-j-1}, \quad k = 1, \ldots, K. \tag{13}$$

In other words, the structural shocks are assumed to be orthogonal and their variances are modelled by individual GARCH$(1, 1)$ processes. Of course, in principle one could consider higher order GARCH processes. Because this is not done in practice, we also focus on GARCH(1,1) processes to avoid unnecessarily clumsy notation. Notice also that the individual GARCH processes imply an unconditional variance of 1. Hence, the unconditional covariance matrix of the structural errors is the identity and

$$\mathbb{E}(u_t u_t') = \Sigma_u = BB'.$$

The setup of our model very much resembles the GO-GARCH specification proposed by van der Weide (2002) and the model considered by Normandin and Phaneuf (2004).

Sentana and Fiorentini (2001) and Milunovich and Yang (2013) provide general results on the identification of B and, hence, the structural shocks. Specifically, B is (locally) identified if at least $K - 1$ of the GARCH(1,1) processes are nontrivial, that is, $\gamma_k \neq 0$ for at least $K - 1$ of the K processes in (13). In essence, there has to be enough heterogeneity in the conditional variances to identify the shocks. At most one shock may be homoskedastic.

In the empirical literature using GARCH for identifying structural shocks in an SVAR analysis, formal tests of the identification condition are typically not applied. Such tests are desirable in particular in macroeconometric studies because they often involve variables for which one would not expect conditional heteroskedasticity a priori, especially when data of relatively low frequency, such as quarterly data are used. Lütkepohl and Milunovich (2015) propose and explore the properties of tests for identification in SVAR-GARCH models that were originally suggested by Lanne and Saikkonen (2007) in a related context. It tests the null hypothesis that there are $K - 2$ nontrivial GARCH components against the alternative that there are more such components. Rejecting the null hypothesis implies identification. The test proceeds as follows.

To emphasize the fact that there may only be $h < K$ nontrivial GARCH(1,1) components, the process generating the reduced-form errors is written as

$$u_t = B \begin{bmatrix} \Lambda_{t|t-1}^{1/2} & 0 \\ 0 & I_{K-h} \end{bmatrix} \varepsilon_t, \tag{14}$$

where $\varepsilon_t \sim \text{iid}(0, I_K)$ and

$$\Lambda_{t|t-1} = \begin{bmatrix} \sigma_{1,t|t-1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{h,t|t-1}^2 \end{bmatrix} \tag{15}$$

is an $(h \times h)$ diagonal matrix with univariate GARCH(1,1) processes on the diagonal. In other words,

$$u_{t|t-1} \sim \left(0, \Sigma_{t|t-1} = B \begin{bmatrix} \Lambda_{t|t-1} & 0 \\ 0 & I_{K-h} \end{bmatrix} B' \right).$$

Partitioning the transformation matrix B such that $B = [B_1 : B_2]$, where B_1 is $(K \times h)$ and B_2 is $(K \times (K - h))$, the first part, B_1 is identified and, hence, can be estimated consistently. Lanne and Saikkonen (2007) propose to choose B_2 as an orthogonal complement of the estimated B_1. Denoting by $\hat{B}$ the estimated B matrix obtained in this way, they propose to test the

last $K - h$ components of $\hat{B}^{-1}\hat{u}_t$, say $\hat{e}_{2t}$, for remaining GARCH components using GARCH tests based on portmanteau type statistics (see Lütkepohl and Milunovich (2015) for details). More precisely, they consider test statistics based on autocorrelations of the quantities

$$\xi_t = \hat{e}'_{2t}\hat{e}_{2t} - T^{-1}\sum_{t=1}^{T}\hat{e}'_{2t}\hat{e}_{2t}$$

and

$$\vartheta_t = \text{vech}(\hat{e}_{2t}\hat{e}'_{2t}) - T^{-1}\sum_{t=1}^{T}\text{vech}(\hat{e}_{2t}\hat{e}'_{2t}),$$

where vech denotes the half-vectorization operator. The test statistics are

$$Q_1(H) = T\sum_{j=1}^{H}[\tilde{\gamma}(h)/\tilde{\gamma}(0)]^2, \tag{16}$$

where

$$\tilde{\gamma}(j) = T^{-1}\sum_{t=j+1}^{T}\xi_t\xi_{t-j},$$

and

$$Q_2(H) = T\sum_{j=1}^{H}\text{tr}[\tilde{\Gamma}(j)'\tilde{\Gamma}(0)^{-1}\tilde{\Gamma}(j)\tilde{\Gamma}(0)^{-1}], \tag{17}$$

where

$$\tilde{\Gamma}(j) = T^{-1}\sum_{t=j+1}^{T}\vartheta_t\vartheta'_{t-j} \quad \text{for} \quad j = 0, 1, \ldots,$$

and $\tilde{\Gamma}(j) = \tilde{\Gamma}(-j)'$ for $j < 0$. Under the null hypothesis of only h GARCH components, the statistics $Q_1(H)$ and $Q_2(H)$ have asymptotic χ^2 distributions with H and $H(K-h)^2(K-h+1)^2/4$ degrees of freedom, respectively. The number H of autocorrelations to be included has to be fixed by the analyst. A value $H = 1$ is considered by Lanne and Saikkonen (2007) and Lütkepohl and Milunovich (2015).

For a third possible test, Lütkepohl and Milunovich (2015) consider the regression

$$\eta_t = \delta_0 + D_1\eta_{t-1} + \cdots + D_H\eta_{t-H} + w_t, \tag{18}$$

where $\eta_t = \text{vech}(\hat{e}_{2t}\hat{e}'_{2t})$, δ_0 is a $\frac{1}{2}(K-h)(K-h+1)$-dimensional fixed vector, the D_i, $i = 1,\ldots,H$ are $\frac{1}{2}(K-h)(K-h+1) \times \frac{1}{2}(K-h)(K-h+1)$ parameter matrices and w_t is an error term. The LM test statistic for testing the null

$$H_0 : D_1 = \cdots = D_H$$

is

$$Q_3(H) = \tfrac{1}{2}T(K-h)(K-h+1) - T\text{tr}[\tilde{\Sigma}_w\tilde{\Gamma}(0)^{-1}] \tag{19}$$

where $\tilde{\Sigma}_w$ is the estimated residual covariance matrix of (18). The critical values are obtained from the $\chi^2(\frac{1}{2}H(K-h)^2(K-h+1)^2)$ distribution. Lütkepohl and Milunovich (2015) also choose $H = 1$ for implementing this test.

The estimation of the parameters of the SVAR-GARCH model may be done by ML under Gaussian assumptions, that is, ε_t is assumed to be normally distributed. Lanne and Saikkonen (2007) derive a form of the likelihood that does not depend on unidentified parameters for given h. Thus, estimation is feasible. The actual optimization of the log-likelihood can still be numerically challenging for large models. Therefore designing useful bootstrap methods, for example, for estimating impulse responses leaves room for further research.

In some studies using the SVAR-GARCH approach, other GARCH models are considered (e.g., Weber (2010), Strohsal and Weber (2012)). In fact, given the additional identifying information in the GARCH structure, it is not even necessary to insist on instantaneously uncorrelated shocks and this assumption is given up in some related literature. The shocks can be identified by some other assumption such as constant conditional correlations (see Weber (2010)). In any case, this does not make the treatment of the model simpler. A drawback is the challenging likelihood function that is difficult to maximize for larger models with many variables although the procedure described in the foregoing may alleviate the task. Another drawback is that the periods of different volatility, such as low volatility and high volatility periods, may not be obvious from the model. On the other hand, the conditional heteroskedastic structure has been applied extensively for financial data and, hence, it has some appeal for such data.

4 The Relation Between U.S. Monetary Policy and the Stock Market

4.1 The Data and the Structural Model

In this section the different heteroskedastic and conditionally heteroskedastic SVARs are illustrated in the context of an investigation of the interaction between U.S. monetary policy and the stock market. We use all three models that allow for endogenously changing volatility and reconsider a study reported by Bjørnland and Leitemo (2009). These authors use short-run and long-run restrictions in a conventional SVAR setting to identify the shocks of interest. The study by Bjørnland and Leitemo (2009) is chosen as point of departure for several reasons. First, these authors propose a novel identification scheme that is not testable in a conventional setting while the SVARs with heteroskedasticity allow to test the restrictions. Second, the model of Bjørnland and Leitemo (2009) is at monthly frequency and we use the extended sample period 1970M1 - 2007M6 and, hence, we have 450 observations for each variable making, for instance, GARCH model estimation feasible. The end of the sample is determined to avoid distortions from the crisis that started around the middle of 2007. Typical macro-VARs on quarterly frequency would not contain more than 250 observations. The data is collected as close as possible to Bjørnland and Leitemo (2009). We consider a five dimensional VAR with the vector of variables being $y_t = (q_t, \pi_t, c_t, \Delta sp_t, r_t)'$, where:

- q_t is the linearly detrended log of an industrial production index;

- π_t is the scaled ($\times 100$) annual change in the log of consumer prices (CPI);

- c_t represents the scaled ($\times 100$) annual change in the log of the World Bank (non energy) commodity price index;

- sp_t is the log of the S&P 500 stock price index deflated by the CPI to measure the real stock prices. The series is first differenced to represent monthly returns (Δsp_t);

- r_t denotes the Federal Funds rate.

The data (with the exception of the commodity price index obtained from the World Bank) is downloaded from the Federal Reserve Bank of St. Louis database FRED.

14

From the point of view of our analysis it is of particular interest to test the restrictions specified by Bjørnland and Leitemo (2009) using different models and, hence, to check whether they are in line with the data. Bjørnland and Leitemo (2009) criticize the recursive identification of monetary policy and stock price shocks. They suggest that monetary policy shocks may have an impact on the stock market within the month. To overcome this problem combining short-run and long-run restrictions is proposed. The identification scheme can be visualized as follows:

$$B = \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \quad \text{and} \quad \Xi_\infty = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & 0 \\ * & * & * & * & * \end{bmatrix}, \qquad (20)$$

where the asterisks indicate unrestricted elements and zeros denote elements restricted to zero and B and Ξ_∞ denote the matrices of impact and long-run effects, respectively, of the shocks, as before. The shocks of particular interest are the shocks ordered fourth and fifth, corresponding to the last two columns of the matrices. This identification suggests that the last shock is the monetary policy shock and it has no immediate impact on industrial production, inflation and commodity prices as well as no long-run effect on stock prices. The shock ordered fourth is viewed as the stock price shock and it has no contemporaneous effect on the real side of the economy. Note that the first three shocks are not of interest for the current analysis. They can be identified arbitrarily in a conventional SVAR analysis.

In the following we consider alternative sets of restrictions for the initial effects matrix B and the long-run effects matrix Ξ_∞ and explore their compatibility with the data.

R1: B and Ξ_∞ are restricted as in (20) (Bjørnland-Leitemo identification).

R2: Only the two last columns of B and Ξ_∞ are restricted as in (20).

R3: Only B is restricted as in (20) and no restrictions are imposed on Ξ_∞.

The first set of restrictions constrains the impact effects and long-run responses just as in the Bjørnland-Leitemo setup. The second set imposes only the restrictions used for the shocks of interest, i.e., the stock market shock and the monetary policy shock. Finally, R3 represents the full set of restrictions from Bjørnland-Leitemo except for the long-run restriction. Thus, imposing only those restrictions does not fully identify the shocks of interest. All these sets of restrictions are either under-identifying or just-identifying

in a conventional framework. Hence, testing them requires additional identifying restrictions that we get from the changes in volatility in the variables.

4.2 The Volatility Model

We estimate the standard VAR and three (conditional) heteroskedastic VARs with different assumptions on the evolution of volatility. The VAR order is chosen to be three as suggested by the Akaike information criterion (AIC) in a conventional model that does not allow for heteroskedasticity. Bjørnland and Leitemo (2009) use four lags based on lag reduction tests for their shorter data range.

In practice, if changes in volatility are suspected, it is useful to think about the most suitable volatility model for a particular application. For the system of interest here, volatility changes are likely, for example, because a financial variable such as Δsp_t is included. It is not easy to link the volatility changes to specific exogenous events and, hence, to specific parts of the sample period. Therefore a model that assigns the volatility states endogenously makes more sense than a model with exogenous volatility changes. Given the stock returns in the system one may think of a GARCH model. However, also the MS model can capture similar conditional heteroskedastic structures. Therefore it is a priori equally suitable if enough volatility states are allowed for. If the volatility change is sufficiently smooth, also the ST model is a possibility. It is used, in fact, by Lütkepohl and Netšunajev (2014b) for the presently considered system and we draw on the results from that study for the ST-SVAR model in the following. In deciding on the volatility model it may be useful to inspect the residuals from a standard VAR analysis. They are presented in Figure 5 and show that there are indeed considerable changes in volatility throughout the sample and capturing them with a model with exogenously assigned volatility regimes may not be easy. Therefore we consider all three models with endogenous assignment of volatility states in the following.

Clearly, a visual inspection of the residuals is not enough for making a good decision on the most suitable volatility model. Such visual tools have to be complemented with statistical criteria such as information criteria. Therefore we show values of the Akaike information criterion and the Schwarz criterion (SC) for different models in Table 1. Notice that in all the volatility models a state-invariant initial effects matrix B is assumed and the VAR order is 3 for all models. The GARCH model is estimated with 5 univariate GARCH(1,1) components. It turns out that all (conditionally) heteroskedastic SVARs perform much better than the standard VAR according to AIC and SC. In fact, both criteria are minimized for the MS-SVAR

model with 3 volatility states. We did not try to estimate MS models with more volatility states because for a 5-dimensional model that would have been a major computational challenge and the estimates cannot be expected to be reliable.

The ultimate objective is, of course, to capture the volatility well with our models. Therefore it is instructive to look at the standardized residuals

$$\hat{u}_{kt}/\hat{\sigma}_{kk,t}$$

obtained for the different models. Here $\hat{\sigma}^2_{kk,t}$ is the kth diagonal element of the estimated residual covariance matrix or conditional covariance matrix for period t, depending on the model considered. These quantities are plotted in Figure 2 for the models with time-varying volatility. Notice that these are not the structural residuals obtained from these models but reduced-form residuals. The graphs are just meant to give a visual impression of possible remaining volatility changes or the ability of the different models to actually capture the volatility changes in the reduced-form residuals.

Comparing to the residuals in Figure 5, all models appear to improve on a model without allowance for volatility changes. However, the volatility models differ in their ability to capture the volatility in the data well. The visual impression is that the ST-SVAR model leaves quite some unexplained volatility in the residuals. Also, the MS(2) model appears to be inferior to the MS(3) model. For example, it leaves more heterogeneity in the commodity price residual series than the MS(3) model. Moreover, the GARCH residuals are somewhat disappointing and clearly inferior to the MS(3) residuals. Note, for example, large outliers in the u_t^r series that are not captured by the GARCH model. Also some of the other residual series leave room for improvement relative to the corresponding quantities from the MS(3) model. For example, the volatility changes in the residuals of the output equation, u_t^q, are not well captured by the GARCH model. Thus, it is no accident that the AIC and SC criteria prefer the MS(3) model as well.

These results are perhaps better understood when looking at the transition function of the ST-SVAR model and the smoothed state probabilities of the MS(2)-SVAR and MS(3)-SVAR models in Figure 3. The MS models indicate that the change in volatility is more complex than what is captured by the ST-SVAR model. There does not appear to be a smooth transition from one volatility regime to another if the data are allowed to assign different regimes in a different way as in the MS models. Also allowing for a third state leads to a quite different separation of the sample period than a two state model. This explains why the MS(3) model is preferred over the ST and the MS(2) models. Despite the fact that we favor the MS(3)-SVAR for the

17

present dataset, we continue looking also at the other models for illustrative purposes.

4.3 Identification Analysis

The first question we have to address is the identification issue. For all three models statistical identification conditions are linked to the changes in volatility. For the ST and MS models they are linked to the relative variances. Estimates of these quantities are shown in Table 2. For the ST and MS(2) models the relative variances in the second state are all less than one. Hence, the ST model moves the volatility from a high volatility state to a lower volatility state. This corresponds well to the fact that the MS(2) model assigns the second, hence, lower volatility state to the latter part of the sample (see Figure 3). In contrast, the first state dominates the second half of the sample in the MS(3) model. Therefore it is plausible that the relative variances of the second and third states assigned by the MS(3) model are all larger than one, that is, the variances are larger than in the first state.

Recall that we have a fully identified ST-SVAR model if all the relative variances are distinct. Taking into account also the standard errors in Table 2, the estimates suggest that at least the first four relative variances are distinct. Formal tests that they are all distinct are presented in Table 3. Note that we have to test $K(K-1)/2 = 10$ pairs for a system of dimension $K = 5$. For the smooth transition model, the null hypotheses of pairwise equality is rejected at a 10% significance level for all pairs but the last. Thus, we conclude that we have some statistically identified shocks but perhaps not a fully identified ST-SVAR model. In particular, the first three shocks are identified via heteroskedasticity. For a fully identified set of structural shocks we need additional restrictions such as those or some of those used by Bjørnland and Leitemo (2009).

The standard errors of the estimated relative variances for the MS models in Table 2 tend to be larger than for the ST model and also the estimated relative variances are not clearly distinct in a number of cases. Therefore it is not surprising that the corresponding tests for full identification shown in Tables 3 and 4 do not reject some of the relevant hypotheses. However, they do indicate that the volatility models contain at least some additional identifying information because some of the relevant hypotheses are clearly rejected with p-values much smaller than 5% or even 1%. Interestingly, in the MS(2)-SVAR model the tests indicate that the first relative variance is clearly distinct from the others and, hence, the first shock is identified while in the MS(3)-SVAR model the last shock is identified. Note, however, that we are using Wald tests for checking the identification restrictions and these

tests may not be very reliable (see Lütkepohl and Netšunajev (2014a)). Thus, there may be more identified shocks. Our current statistical procedures do not provide much evidence for them, however.

The identification issue in the context of the SVAR-GARCH model is a somewhat different one. We have to check the number of univariate GARCH components underlying the GARCH structure. Ignoring the problem and estimating a model with five GARCH components as we have done in the foregoing, some of the GARCH parameters may actually not be identified. Thus, the analysis of the SVAR-GARCH model should ideally start with a thorough investigation of the number of underlying univariate GARCH components. This is of particular importance in analyses such as the present one where it is not clear whether the GARCH structure captures changes in volatility well enough in all variables. Notably variables such as industrial production are not necessarily well modelled by GARCH. Therefore we use the tests presented in Section 3.4 to investigate the issue. The results are shown in Table 5. Since the tests reject fewer than five GARCH components ($r < 5$), it makes sense to use the GARCH model with five components which also fully identifies all five shocks. As discussed earlier, this does not necessarily mean that the GARCH-SVAR model is the best one. In fact, we have seen earlier that it may be inferior to the MS(3)-SVAR model in some respects. What our tests show, however, is that the GARCH specification has a likelihood with curvature against restrictions on the B and Ξ_∞ matrices if these restrictions are not in line with the data in the framework of the SVAR-GARCH model.

The overall conclusion of our identification analysis is that all models may present some power against restrictions on the short-run and long-run effects matrices B and Ξ_∞. Therefore it makes sense to use all four models to test the restrictions discussed earlier. The test results are presented in Table 7. Obviously, all the null hypotheses can be rejected against models with fewer restrictions in all the four volatility setups. Thus there is strong evidence against any set of the restrictions. Rejecting R1 means, of course, that the Bjørnland-Leitemo restrictions are rejected. Such a rejection may be due to only one or a small subset of restrictions being false. To make sure that they are not rejected because the restrictions used to identify the shocks of no interest for an analysis of the stock market and monetary shocks, we also test R2 against a model without zero restrictions on B and Ξ_∞. Again, these restrictions are clearly rejected regardless of the volatility model used. Moreover, the impact restrictions R3 are rejected in the absence of the long-run restriction. Similarly, the zero restriction on Ξ_∞ is individually rejected in a test of R1 against R3. These results support findings in Lütkepohl and Netšunajev (2014b) where also other ST-SVAR models are considered.

19

Of course, one could argue that it is not clear whether the Bjørnland-Leitemo restrictions or the volatility models are rejected by the tests. However, given that all the volatility models agree in their rejections, this sheds some doubt on the restrictions. After all, in the conventional approach the changes in volatility are ignored and the restrictions are just identifying so that the data have no opportunity to speak up against them.

4.4 Impulse Response Analysis

Having rejected the Bjørnland-Leitemo restrictions one wonders what can be said about their main issue of interest namely the interaction between monetary policy and the stock market. The volatility analysis so far has not provided a full set of economic shocks. Even in the GARCH model where full identification is found, the shocks delivered by the SVAR-GARCH approach are just statistically identified and may not have economic meaning. For example, they may be mixtures of economic shocks. For the other models there is not even a full set of properly identified shocks. So the question arises whether we can learn something from our analysis regarding the economic questions of interest. To answer that question we can take a look at the identified shocks and see whether they can be interpreted as economically meaningful shocks. This can best be assessed by looking at the corresponding impulse responses.

Because we favor the MS(3)-SVAR model on the basis of our statistical analysis and because in this model we found only one clearly identified shock we present impulse responses for this shock in Figure 5 together with 68% confidence intervals. On impact the shock moves up the interest rate and, hence, qualifies as a contractionary monetary policy shock. On the other hand, there is a pronounced initial increase in output and inflation which would not be expected for responses to a monetary policy shock. Therefore we follow Lütkepohl and Netšunajev (2014b) and interpret the shock identified via heteroskedasticity as a demand side monetary shock rather than a supply side shock of the money market. This is also in line with the money demand shock studied in Kulikov and Netšunajev (2013) and Favara and Giordani (2009). Hence, in our system and in the system we do not find a shock that qualifies as a supply side shock of the money market. Hence, the restrictions may be rejected because they do not identify the desired shock.

Turning now to the response of stock prices (sp_t), it is seen in Figure 5 that there is a significant decline for a few months after which the stock prices return to insignificant territory. Such a response is in fact quite plausible as one would expect the stock market to absorb monetary shocks quickly. Note that the impulse responses in Figure 5 are the accumulated responses of Δsp_t

and, thus, they are the responses of sp_t rather than the returns.

For the other models we have also plotted impulse responses to the clearly identified shocks and checked whether any one of them increases the interest rate significantly on impact. It turns out that for each of the models only one shock really qualifies as a monetary shock. We present the impulse responses of these shocks in Figure 5. As can be seen in the figure, the responses look very similar in all the models. In particular, they qualitatively tell the same story regarding the stock market response. The stock returns decline initially but move back up after a few months. Thus, we can conclude that there is some evidence of a short-term stock market response to activities in the money market. It is striking how little the volatility model affects the results. The important point seems to be that volatility changes are taken into account at all. The specific model for describing them is of secondary importance even if it is not an ideal description of the volatility changes. Of course, this latter conclusion may be specific to this example and cannot be generalized based on the available evidence.

5   Conclusions

In this study we have reviewed the volatility models in current use for identifying structural shocks in VAR analysis. The great advantage of this tool is that it offers potentially additional identifying information that can be used to investigate assumptions that are just-identifying in a conventional SVAR framework and thereby opens up the possibility to confront the data with such assumptions. We have pointed out the advantages and drawbacks of exogenously specified changes in volatility, smoothly changing volatility and volatility changes modelled by Markov switching and GARCH processes. Understanding the specific features of these models can help in making more informed decisions which model to use in a particular application.

Given that little about the actual volatility changes is typically known in practice for a specific set of variables of interest, exogenously specified volatility changes usually require considerable pretesting or preliminary analysis of possible changes. On one hand, they are very rigid in representing the break in volatility and assume that there is no transition for an extended period after some change has occurred. On the other hand, estimation is relatively easy.

Models that allow for endogenously determined volatility changes such as ST-SVAR, MS-SVAR and SVAR-GARCH models remove the burden of specifying the change points. Also the changes as such can be very flexible. The greatest flexibility in this respect is offered by MS-SVAR and SVAR-GARCH

models. On the other hand, these models are also the ones most difficult to estimate. In fact, estimation still poses considerable computational problems for models of larger size, that is, for models with many variables, large lag orders or, in the case of MS models, large numbers of different volatility states. For these models not only estimation of the model parameters is difficult but they also pose considerable problems for constructing bootstrap confidence intervals for impulse responses, for example.

On the positive side, for all model types precise conditions are known for full identification of all shocks. Moreover, statistical tests are available for checking the identification conditions formally.

We have illustrated the use of the models and some of the statistical issues involved by a detailed analysis of a monthly U.S. model for investigating the relation between monetary policy and the stock market. We find that previously used restrictions for identifying the shocks that are just-identifying in a conventional framework can be formally tested against the data and are in fact rejected by all volatility models. We find a Markov switching model with three regimes to be the most suitable model. Unfortunately, we do not find a full set of economically interpretable shocks using our preferred volatility model. Only a potential monetary shock and its responses can be clearly identified. It indicates that money market activities have some impact on stock returns. The effects are absorbed by the stock market quickly, however, as one might expect. This result is quite robust with respect to the volatility model used in this application. Whether or not this is a more general result is pure speculation. It shows, however, the importance of allowing for volatility changes in a system of variables where such changes actually occur.

References

Bacchiocchi, E. and Fanelli, L. (2012). Identification in structural vector autoregressive models with structural changes, *Departmental Working Papers 2012-16*, Department of Economics, Management and Quantitative Methods at Università degli Studi di Milano.

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

Bjørnland, H. C. and Leitemo, K. (2009). Identifying the interdependence between US monetary policy and the stock market, *Journal of Monetary Economics* **56**: 275 – 282.

Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances, *American Economic Review* **79**: 655–673.

Bouakez, H. and Normandin, M. (2010). Fluctuations in the foreign exchange market: How important are monetary policy shocks?, *Journal of International Economics* **81**: 139–153.

Canova, F. and De Nicoló, G. (2002). Monetary disturbances matter for business fluctuations in the G-7, *Journal of Monetary Economics* **49**: 1131–1159.

Ehrmann, M., Fratzscher, M. and Rigobon, R. (2011). Stocks, bonds, money markets and exchange rates: Measuring international financial transmission, *Journal of Applied Econometrics* **26**: 948–974.

Faust, J. (1998). The robustness of identified VAR conclusions about money, *Carnegie-Rochester Conference Series in Public Policy* **49**: 207–244.

Favara, G. and Giordani, P. (2009). Reconsidering the role of money for output, prices and interest rates, *Journal of Monetary Economics* **56**: 419–430.

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

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market, *American Economic Review* **99**: 1053–1069.

King, R. G., Plosser, C. I., Stock, J. H. and Watson, M. W. (1991). Stochastic trends and economic fluctuations, *American Economic Review* **81**: 819–840.

Kulikov, D. and Netšunajev, A. (2013). Identifying monetary policy shocks via heteroskedasticity: A Bayesian approach, *Bank of Estonia Working Papers WP2013-9*, Bank of Estonia.

Lanne, M. and Lütkepohl, H. (2008). Identifying monetary policy shocks via changes in volatility, *Journal of Money, Credit and Banking* **40**: 1131–1149.

Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching, *Journal of Economic Dynamics and Control* **34**: 121–131.

Lanne, M. and Saikkonen, P. (2007). A multivariate generalized orthogonal factor GARCH model, *Journal of Business & Economic Statistics* **25**: 61–75.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. (2013). Identifying structural vector autoregressions via changes in volatility, *Advances in Econometrics* **32**: 169–203.

Lütkepohl, H. and Milunovich, G. (2015). Testing for identification in SVAR-GARCH models - reconsidering the impact of monetary shocks on exchange rates, *Discussion Paper 1455*, DIW Berlin.

Lütkepohl, H. and Netšunajev, A. (2014a). Disentangling demand and supply shocks in the crude oil market: How to check sign restrictions in structural VARs, *Journal of Applied Econometrics* **29**: 479–496.

Lütkepohl, H. and Netšunajev, A. (2014b). Structural vector autoregressions with smooth transition in variances: The interaction between u.s. monetary policy and the stock market, *Discussion Papers of DIW Berlin 1388*, DIW Berlin, German Institute for Economic Research.

Lütkepohl, H. and Velinov, A. (2015). Structural vector autoregressions: Checking identifying long-run restrictions via heteroskedasticity, *Journal of Economic Surveys* .

Milunovich, G. and Yang, M. (2013). On identifying structural VAR models via ARCH effects, *Journal of Time Series Econometrics* **5**: 117–131.

Netšunajev, A. (2013). Reaction to technology shocks in Markov-switching structural VARs: Identification via heteroskedasticity, *Journal of Macroeconomics* **36**: 51–62.

Normandin, M. and Phaneuf, L. (2004). Monetary policy shocks: Testing identification conditions under time-varying conditional volatility, *Journal of Monetary Economics* **51**: 1217–1243.

Rigobon, R. (2003). Identification through heteroskedasticity, *Review of Economics and Statistics* **85**: 777–792.

Rigobon, R. and Sack, B. (2003). Measuring the reaction of monetary policy to the stock market, *Quarterly Journal of Economics* **118**: 639–669.

Rigobon, R. and Sack, B. (2004). The impact of monetary policy on asset prices, *Journal of Monetary Economics* **51**: 1553–1575.

Sentana, E. and Fiorentini, G. (2001). Identification, estimation and testing of conditionally heteroskedastic factor models, *Journal of Econometrics* **102**: 143–164.

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

Strohsal, T. and Weber, E. (2012). The signal of volatility, *Discussion Paper 2012-043*, Humboldt University Berlin, SFB 649.

Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure, *Journal of Monetary Economics* **52**: 381–419.

van der Weide, R. (2002). GO-GARCH: A multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics* **17**: 549–564.

Velinov, A. S. (2013). Can stock price fundamentals really be captured? Using Markov switching in heteroskedasticity models to test identification restrictions, *Discussion Paper 1350*, DIW Berlin.

Weber, E. (2010). Structural conditional correlation, *Journal of Financial Econometrics* **8**: 392–407.

Table 1: Comparison of SVAR(3) Models with State Invariant B

Model	$\log L_T$	AIC	SC
SVAR(3)	-3159.344	6508.689	6899.067
ST-SVAR(3)	-2878.255	5976.510	6428.527
MS(2)-SVAR(3)	-2826.742	5877.484	6337.719
MS(3)-SVAR(3)	-2774.614	5791.230	6288.448
SVAR(3)-GARCH(1,1)	-2891.971	6013.942	6486.505

Note: L_T – likelihood function, AIC $= -2\log L_T + 2 \times$no of free parameters, SC $= -2\log L_T + \log T \times$no of free parameters.

Table 2: Estimates of Relative Variances for Smooth Transition and Markov Switching Models

parameter	ST-SVAR estimate	ST-SVAR std.dev.	MS(2)-SVAR estimate	MS(2)-SVAR std.dev.	MS(3)-SVAR	MS(3)-SVAR
λ_{21}	0.019	0.002	0.019	0.003	3.724	0.849
λ_{22}	0.315	0.057	0.271	0.097	2.741	0.626
λ_{23}	0.548	0.088	0.371	0.125	5.243	1.309
λ_{24}	0.867	0.154	0.428	0.135	3.544	0.761
λ_{25}	0.927	0.172	0.682	0.356	2.568	0.608
λ_{31}					1.859	0.593
λ_{32}					2.554	0.747
λ_{33}					2.693	0.863
λ_{34}					3.981	1.153
λ_{35}					82.265	16.065

26

Table 3: Tests for Equality of λ_{2k} for Unrestricted ST-SVAR and MS(2) Model

H_0	ST-SVAR		MS(2)-SVAR	
	Wald statistic	p-value	Wald statistic	p-value
$\lambda_{21} = \lambda_{22}$	26.463	2.686×10^{-7}	6.6233	0.0101
$\lambda_{21} = \lambda_{23}$	35.720	2.277×10^{-9}	7.9320	0.0049
$\lambda_{21} = \lambda_{24}$	29.806	4.673×10^{-8}	9.0093	0.0027
$\lambda_{21} = \lambda_{25}$	27.498	1.572×10^{-7}	3.4660	0.0626
$\lambda_{22} = \lambda_{23}$	4.729	0.029	0.7049	0.4012
$\lambda_{22} = \lambda_{24}$	10.731	0.001	1.3050	0.2533
$\lambda_{22} = \lambda_{25}$	10.583	0.001	0.9135	0.3392
$\lambda_{23} = \lambda_{24}$	2.814	0.093	0.1255	0.7232
$\lambda_{23} = \lambda_{25}$	3.687	0.054	0.4901	0.4839
$\lambda_{24} = \lambda_{25}$	0.061	0.805	0.3201	0.5715

Table 4: Tests for Pairwise Equality of λ_{2k} and λ_{3k} for MS(3)-SVAR Model

H_0	MS(3)-SVAR	
	Wald statistic	p-value
$\lambda_{21} = \lambda_{22}, \lambda_{31} = \lambda_{32}$	2.011	0.356
$\lambda_{21} = \lambda_{23}, \lambda_{31} = \lambda_{33}$	1.138	0.565
$\lambda_{21} = \lambda_{24}, \lambda_{31} = \lambda_{34}$	2.921	0.232
$\lambda_{21} = \lambda_{25}, \lambda_{31} = \lambda_{35}$	29.753	0.000
$\lambda_{22} = \lambda_{23}, \lambda_{32} = \lambda_{33}$	2.993	0.223
$\lambda_{22} = \lambda_{24}, \lambda_{32} = \lambda_{34}$	1.193	0.551
$\lambda_{22} = \lambda_{25}, \lambda_{32} = \lambda_{35}$	26.424	0.000
$\lambda_{23} = \lambda_{24}, \lambda_{33} = \lambda_{34}$	2.758	0.251
$\lambda_{23} = \lambda_{25}, \lambda_{33} = \lambda_{35}$	34.043	0.000
$\lambda_{24} = \lambda_{25}, \lambda_{34} = \lambda_{35}$	28.324	0.000

Table 5: Tests for Identification in GARCH-SVAR Model

r under H_0	$Q_1(1)$	df	p-value	$Q_2(1)$	df	p-value	$Q_3(1)$	df	p-value
1	24.026	1	9.5×10^{-7}	251.520	100	4.8×10^{-15}	4036.911	100	0
2	20.446	1	6.1×10^{-6}	148.644	36	1.2×10^{-15}	2252.564	36	0
3	12.712	1	3.6×10^{-4}	31.256	9	2.6×10^{-4}	899.776	9	0
4	15.127	1	1.1×10^{-4}	15.127	1	1.1×10^{-4}	12.421	1	4.2×10^{-4}

Table 6: Estimates of GARCH Parameters for Model with Unrestricted B and Ξ_∞

	γ_k		g_k	
k	estimate	std.dev.	estimate	std.dev.
1	0.349	0.078	0.089	0.182
2	0.342	0.021	0.614	0.039
3	0.245	0.233	0.234	0.223
4	0.125	0.005	0.819	0.037
5	0.095	0.004	0.832	0.036

Table 7: Tests for Identifying Restrictions in Heteroskedastic SVAR Models

H_0	H_1	LR statistic	df	p-value
R1	ST-SVAR with unrestricted B, Ξ_∞	35.845	10	8.9×10^{-5}
R2	ST-SVAR with unrestricted B, Ξ_∞	30.909	7	6.4×10^{-5}
R3	ST-SVAR with unrestricted B, Ξ_∞	22.491	9	0.0074
R1	R3	13.354	1	2.5×10^{-4}
R1	MS(2)-SVAR with unrestricted B, Ξ_∞	52.505	10	9.1×10^{-8}
R2	MS(2)-SVAR with unrestricted B, Ξ_∞	54.586	7	1.4×10^{-13}
R3	MS(2)-SVAR with unrestricted B, Ξ_∞	19.097	9	0.024
R1	R3	33.408	1	7.4×10^{-9}
R1	MS(3)-SVAR with unrestricted B, Ξ_∞	54.134	10	4.5×10^{-8}
R2	MS(3)-SVAR with unrestricted B, Ξ_∞	49.350	7	1.93×10^{-8}
R3	MS(3)-SVAR with unrestricted B, Ξ_∞	27.921	9	9.8×10^{-4}
R1	R3	26.212	1	3.5×10^{-7}
R2	SVAR-GARCH with unrestricted B, Ξ_∞	90.047	10	3.7×10^{-15}
R2	SVAR-GARCH with unrestricted B, Ξ_∞	92.229	7	0.000
R3	SVAR-GARCH with unrestricted B, Ξ_∞	56.052	9	6.7×10^{-9}
R1	R3	33.995	1	4.5×10^{-9}
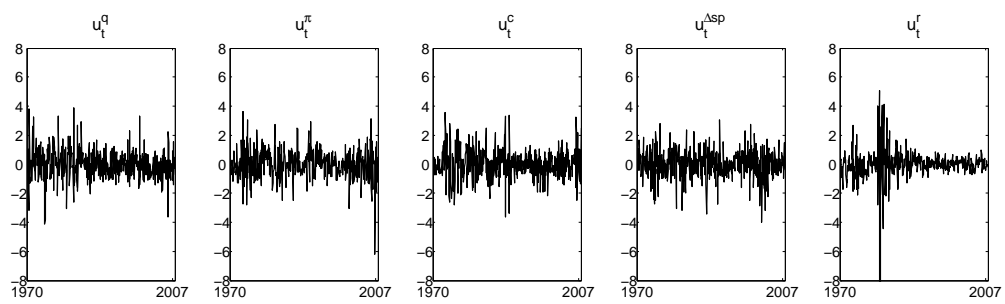
Figure 1: Residuals obtained from VAR(3) model without allowance for heteroskedasticity.
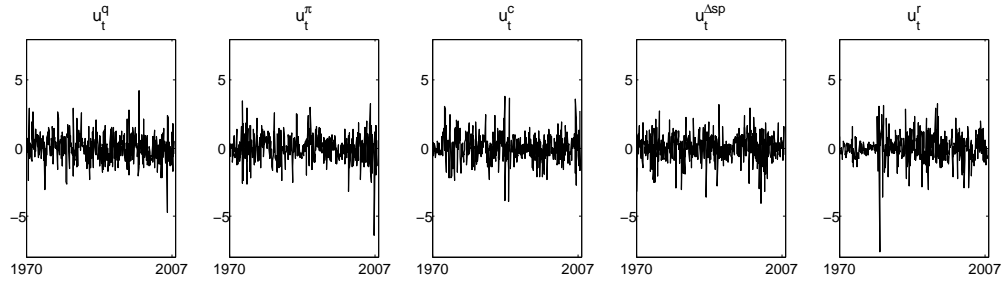
(a) Standardized residuals obtained from ST-SVAR model
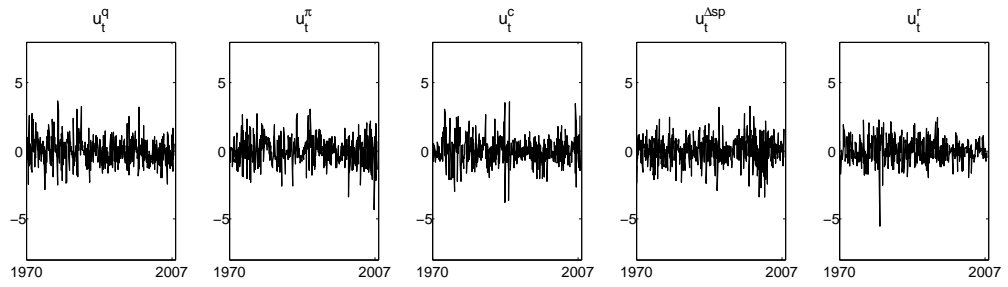
(b) Standardized residuals obtained from MS(2)-SVAR model
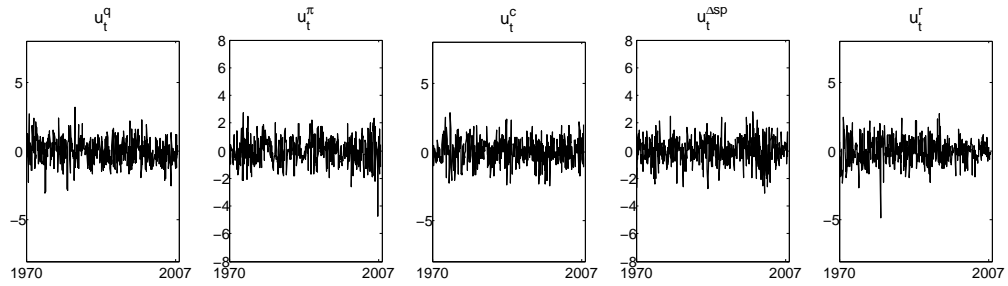
(c) Standardized residuals obtained from MS(3)-SVAR model
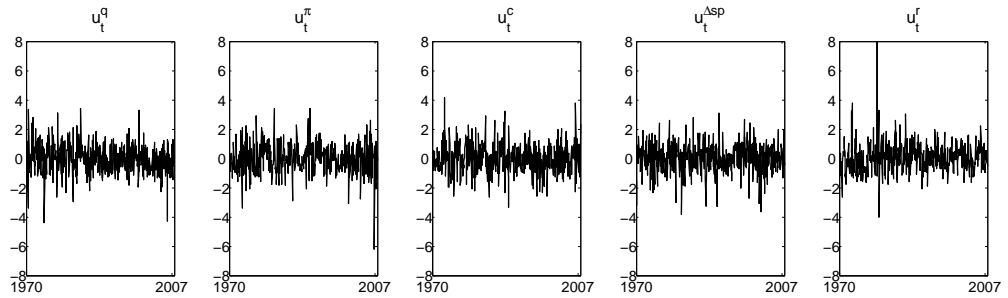
(d) Standardized residuals obtained from SVAR-GARCH model

Figure 2: Standardized residuals obtained from different models allowing for time-varying volatility.

30
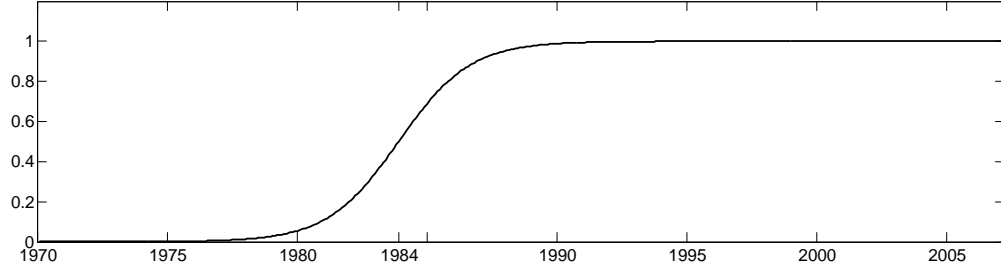
(a) Transition function of ST-SVAR model
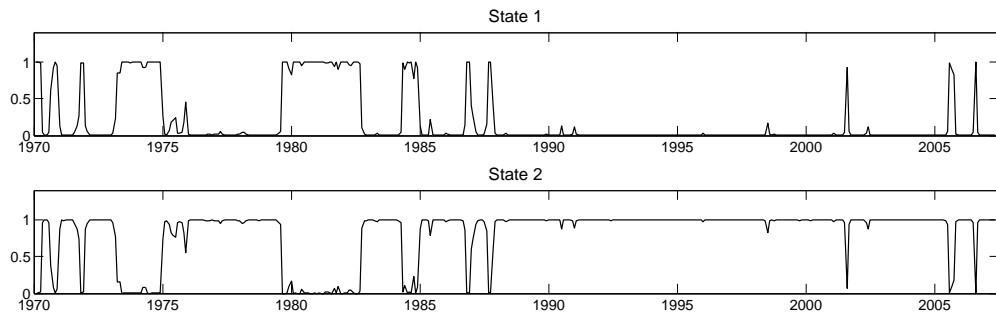
(b) Smoothed regime probabilities of MS(2)-SVAR model
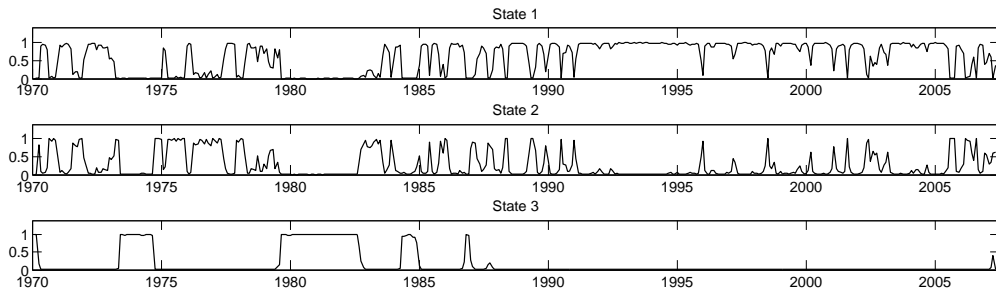
(c) Smoothed regime probabilities of MS(3)-SVAR model

Figure 3: Volatility regimes obtained from different models.

31
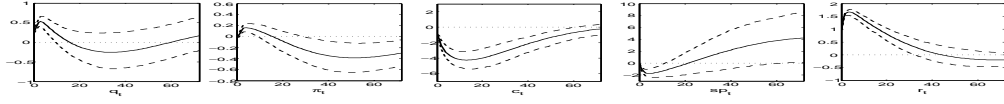
Figure 4: Impulse responses to monetary shock obtained from MS(3)-SVAR model.
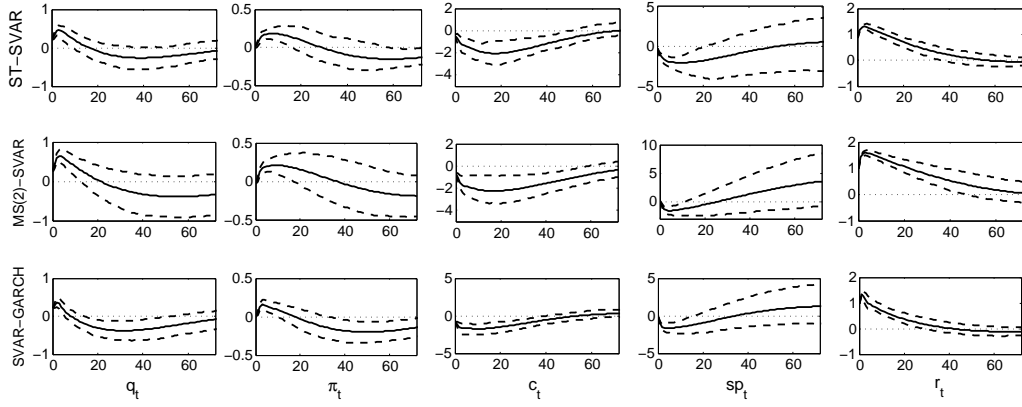
Figure 5: Impulse responses to monetary shock obtained from models.

SFB 649 Discussion Paper Series 2015

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.