Journal of Econometrics 183 (2014) 104–116

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks

Helmut Herwartz [a], Helmut Lütkepohl [b,*]

[a] Department of Economics, Georg-August-University, Platz der Göttinger Sieben 5, D-37073 Göttingen, Germany

[b] Department of Economics, Freie Universität Berlin and DIW Berlin, Mohrenstr. 58, D-10117 Berlin, Germany

ARTICLE INFO

Article history:
Available online 26 June 2014

JEL classification:
C32

Keywords:
Vector autoregressive model
Markov process
EM algorithm
Impulse responses

ABSTRACT

In structural vector autoregressive (SVAR) analysis a Markov regime switching (MS) property can be exploited to identify shocks if the reduced form error covariance matrix varies across regimes. Unfortunately, these shocks may not have a meaningful structural economic interpretation. It is discussed how statistical and conventional identifying information can be combined. The discussion is based on a VAR model for the US containing oil prices, output, consumer prices and a short-term interest rate. The system has been used for studying the causes of the early millennium economic slowdown based on traditional identification with zero and long-run restrictions and using sign restrictions. We find that previously drawn conclusions are questionable in our framework.

1. Introduction

Identifying structural shocks is a major issue in structural vector autoregressive (SVAR) analysis. A range of alternative proposals have been made for this purpose. For example, recursive models which impose a triangular structure on the instantaneous effects of the shocks have been popular in the earlier SVAR literature (e.g., Sims, 1980, Amisano and Giannini, 1997, Lütkepohl, 2005, Chapter 9). Later restrictions on the long-run effects of shocks became popular (Blanchard and Quah, 1989; King et al., 1991; Pagan and Pesaran, 2008) as well as sign or shape restrictions for the shocks (Uhlig, 2005; Canova and De Nicoló, 2002; Faust, 1998). Typically these restrictions rely on potentially controversial economic or institutional believes about the system of interest. Given that the restrictions are often just-identifying, it is not possible to test them against the data in a conventional SVAR analysis. The problem is also present when identification relies on sign restrictions. In that case only those impulse responses are retained which satisfy the prior assumptions of the investigator. Thus, the assumptions are satisfied by construction.

There are two main problems related to these kinds of identification restrictions which both result from the fact that the data are not informative on the validity of the restrictions. First, controversial views on the underlying economic structures cannot be resolved by statistical tests. Second, assuming that the reduced form is a valid description of the data generation process, the data have no opportunity to reflect a general incompatibility of the identifying restrictions and the model. For example, restrictions may be valid within a larger model with additional variables but impose a structure on the actual model under investigation which results in unrealistic impulse responses. In this context it may be worth remembering that a number of models produced a 'price puzzle', that is, a price level increase in response to a contractionary monetary policy shock, which disappeared in a larger model with forward-looking variables capturing expectations.

For these reasons it is of interest that sometimes statistical properties of the data may contain further information that is usually not accounted for in the identification of shocks in a conventional SVAR analysis, as pointed out by Sentana and Fiorentini (2001), Rigobon (2003), Normandin and Phaneuf (2004), Lanne and Lütkepohl (2010) and others. In particular, in these articles residual heteroskedasticity or conditional heteroskedasticity is used for extracting additional identifying information from the data. This approach was also used by Lanne et al. (2010) who consider a Markov regime switching (MS) mechanism for modeling changes in the volatility of the residuals.

We will build on the latter approach and consider the question how this statistical information can be combined with conventional identifying information in a meaningful way. An identification procedure which draws exclusively on statistical data

* Corresponding author.
E-mail addresses: hherwartz@uni-goettingen.de (H. Herwartz), hluetkepohl@diw.de (H. Lütkepohl).

http://dx.doi.org/10.1016/j.jeconom.2014.06.012

properties may end up with structural shocks which are not meaningful economically. That is, the shocks and corresponding impulse responses may not be informative about the underlying economic mechanisms. Moreover, the shocks obtained from our setup are unique only up to permutation. Hence, even if economically meaningful shocks are found, economic properties of the shocks or the associated impulse responses have to be used for labeling them. In other words, economic information has to be used in addition to the statistical properties. On the other hand, it is clear that the economic assumptions have to be in line with the sample information for using them in this context. This feature can be checked given the statistical properties of the data. We will discuss how the two types of identifying information can be combined beneficially. To that end we will also discuss some technical extensions of the basic approach set out in Lanne et al. (2010). More precisely, we discuss how to overcome problems related to the optimization of the log-likelihood function and bootstrap methods for impulse responses. Note that the latter methods were not proposed by Lanne et al. (2010) because of the computational complexities involved. We propose a bootstrap method which is feasible in practice.

The main issues will be illustrated with an empirical model from Peersman (2005). He uses SVAR technology to investigate the causes of the recession in major economies at the beginning of the new millennium and attributes the economic slowdown to a combination of shocks in oil prices, monetary policy, aggregate supply and aggregate demand. The actual contribution of these shocks depends on the identification strategy used for the shocks. In particular, he compares a conventional identification scheme using zero restrictions on the instantaneous and long-run effects of shocks and a strategy based on sign restrictions.

As mentioned earlier, both of these identification approaches have the drawback in the present context that they do not leave room for the data to speak up against the restrictions. Therefore, in this study we use an identification strategy which avoids this shortcoming. It is assumed that there are changes in the volatility which are driven by an MS mechanism. Distinct states of volatility provide an additional source of identifying information that is utilized to check restrictions which are just-identifying in a conventional SVAR analysis.

The structure of the paper is as follows. In the next section we present the basic underlying SVAR model with conventional identification based on instantaneous and long-run effects of the shocks. In Section 3 the MS extension and some related technical problems are discussed. From the outset we discuss the models with the US example system in mind for which the detailed empirical analysis is presented in Section 4. Conclusions follow in Section 5. The Appendix contains details on the estimation algorithm.

The following abbreviations are used throughout: VAR for vector autoregressive or autoregression, SVAR for structural VAR, VECM for vector error correction model, MS for Markov regime switching, ML for maximum likelihood, LR for likelihood ratio, AIC for Akaike information criterion, SC for Schwarz information criterion and IR for impulse response.

2. The conventional SVAR model

We consider a vector error correction model (VECM) for a K-dimensional vector y_t,

$$\Delta y_t = \nu_0 + \nu_1 t + \alpha\beta' y_{t-1} + \sum_{i=1}^{p} \Gamma_i \Delta y_{t-i} + B\varepsilon_t, \tag{1}$$

where Δ signifies the differencing operator, $\alpha\beta' y_{t-1}$ is the error correction term containing the cointegration relations $\beta' y_{t-1}$, $\nu_0 + \nu_1 t$ is a linear trend term and ε_t is the vector of K structural residuals which is assumed to have a diagonal covariance matrix. The quantity B is a $(K \times K)$ matrix of instantaneous effects of the shocks.

In the framework of this model restrictions for the instantaneous effects of the shocks are placed on B, whereas long-run restrictions are placed on the matrix of long-run effects,

$$\Xi = \beta_\perp \left[\alpha'_\perp \left(I_K - \sum_{i=1}^{p} \Gamma_i \right) \beta_\perp \right]^{-1} \alpha'_\perp B, \tag{2}$$

where $\beta_\perp$ and $\alpha_\perp$ denote $(K \times (K - r))$ dimensional orthogonal complements of the $(K \times r)$ dimensional matrices β and α, respectively. Here r is the cointegrating rank (see, e.g., Lütkepohl, 2005, Chapter 9 for details).

In the empirical section we consider a four-dimensional US system $y_t = (oil_t, q_t, p_t, s_t)'$, where oil_t is the price of oil, q_t is output, p_t is a consumer price index and s_t is a short-term interest rate. The first three variables are treated as integrated of order one and not cointegrated whereas the interest rate is assumed to be stationary on theoretical grounds although for the actual variable used in the empirical study there is also some evidence for a unit root. Thus, the only 'cointegration vector' in (1) is $\beta = (0, 0, 0, 1)'$ and, hence, $\beta_\perp = [I_3\ 0]'$, where I_3 denotes a (3×3) identity matrix. Accordingly, $\mathrm{rk}(\Xi) = 3$ and the last row of Ξ consists of zeros. Moreover, B is a (4×4) matrix of instantaneous effects of the shocks $\varepsilon_t = (\varepsilon_t^{oil}, \varepsilon_t^{s}, \varepsilon_t^{d}, \varepsilon_t^{m})'$, where the components represent oil price shocks, aggregate supply shocks, demand or spending shocks and monetary policy shocks, respectively.

In his conventional identification scheme Peersman (2005) assumes that aggregate supply, demand and monetary policy shocks have no instantaneous impact on oil prices and monetary policy shocks also have no immediate impact on output. Moreover, he assumes that demand and monetary shocks are neutral in the long-run and, thus, have only transitory effects on output. These assumptions translate into the following restrictions on the contemporaneous and long-run effects matrices:

$$B = \begin{bmatrix} * & 0 & 0 & 0 \\ * & * & * & 0 \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \quad \text{and} \quad \Xi = \begin{bmatrix} * & * & * & * \\ * & * & 0 & 0 \\ * & * & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3}$$

Here unrestricted elements are denoted by asterisks. The zero restrictions imposed on B and Ξ are just-identifying, and, hence, they cannot be tested against the data in the conventional framework.

Peersman (2005) points out that these restrictions are not uncontroversial and therefore he also performs an analysis which relies on sign restrictions only. For example, oil prices may react to demand or supply shocks within the same quarter when they occur. Moreover, there are economic models which allow for instantaneous effects of monetary policy shocks on output. The long-run restrictions may be problematic in this context because demand and monetary policy shocks may affect the steady-state level of capital (see, e.g., Gali, 1992). Other restrictions may be more appropriate instead. For example, the Fed may not respond instantaneously to oil price shocks (e.g., Kilian and Lewis, 2011, Nakov and Pescatori, 2010). Hence, it is useful to check these assumptions carefully. In the next section we discuss the formal framework which will be used for this purpose.

3. A model with different volatility regimes

3.1. The model setup

Following Lanne et al. (2010) we assume that the distribution of the reduced form error term $u_t = B\varepsilon_t$ depends on a Markov process s_t such that

$$u_t | s_t \sim \mathcal{N}(0, \Sigma_{s_t}). \tag{4}$$

Here s_t $(t = 0, \pm 1, \pm 2, \ldots)$ is a discrete Markov process with states $1, \ldots, M$ and transition probabilities

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

The conditional normality assumption in (4) is not critical for our analysis. It is just made for convenience to set up the likelihood function for ML estimation. If conditional normality does not hold, our estimators will just be pseudo ML estimators. Note, however, that conditional normality of the residuals for each state implies an unconditional nonnormal distribution in general. In fact, our assumptions cover a rich distribution class for the residuals.

The crucial feature in (4) is that the covariances Σ_{s_t} can vary across states. This fact is used by Lanne et al. (2010) to identify structural shocks which are consistent with the statistical data properties and to test restrictions which are just-identifying in the conventional setup. To see how this can be done suppose first that there are just two states ($M = 2$). Then there exists a decomposition $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2 B'$, where $\Lambda_2 = \operatorname{diag}(\lambda_{21}, \ldots, \lambda_{2K})$ is a diagonal matrix with positive diagonal elements. If the λ_{2i}'s are all distinct, this decomposition is in fact unique apart from changes in sign and permutations in the λ_{2i}'s and the corresponding columns of B. Thus, if we assume that the structural shocks are orthogonal across states, have the same instantaneous effects in each state and are normalized such that they have unit variance in the first state, then they are uniquely determined by the transformation $\varepsilon_t = B^{-1} u_t$. Hence, any restrictions imposed on B and Ξ are over-identifying and can be tested against the data.

Notice that the assumptions for the effects of the structural shocks do not go beyond what is typically assumed in a classical framework. In particular, orthogonality across the sample is a standard assumption in structural VAR analysis. Also, if no distinction between volatility states is made, a classical analysis has no reason to allow for changes in the instantaneous effects during the sample period. Hence, making the assumption in our framework as well is plausible and not more restrictive than in a standard SVAR analysis. Moreover, the standardization of the variances of the structural residuals is common in the classical framework. It could be replaced by imposing a unit diagonal on B and a diagonal covariance matrix of the structural shocks in State 1. Notice that in our setup the diagonal elements of the matrix Λ_2 can be interpreted as relative variances of the structural shocks in State 2 versus State 1.

It is important to note, however, that the ordering of the diagonal elements of Λ_2 can be changed freely. This will also change the order of the shocks and the columns of B. Thus, although the shocks and their instantaneous responses in B are unique, they can be permuted without changing the products BB' and $B\Lambda_2 B'$. This property is important to remember when it comes to labeling or interpreting the shocks economically. The sequence in which the shocks appear in the vector ε_t is arbitrary if an arbitrary ordering of the λ_{2j} is used. Attaching meaningful labels to the shocks usually requires taking into account the underlying economic mechanisms. For example, in a system with a demand and a supply shock only, one may label the shock which is neutral in the long-run as demand shock and the one with long-run effects as supply shock.

The critical assumption for uniqueness of the shocks is that the diagonal elements of Λ_2 all have to be distinct. This, however, is a property which can be checked with statistical tests. If there are equal elements on the diagonal of Λ_2, B will no longer be (locally) unique. The elements of Λ_2 are still identified if $\Sigma_1 \neq \Sigma_2$. Thus, we can test equality of the diagonal elements of Λ_2. In other words, for identification purposes we can go much further with statistical analysis than in a conventional framework which does not take advantage of potential volatility changes during the sample period.

If there are more than two volatility states, the corresponding covariance matrix decomposition

$$\Sigma_1 = BB', \qquad \Sigma_i = B\Lambda_i B', \quad i = 2, \ldots, M, \tag{5}$$

with diagonal Λ_i's becomes restrictive. In fact, in that case it can be tested and thereby the assumption of invariant instantaneous effects of the structural shocks across states can be checked. The corresponding likelihood ratio (LR) test has an asymptotic χ^2-distribution with

$$\tfrac{1}{2}MK(K+1) - K^2 - (M-1)K \tag{6}$$

degrees of freedom (Lanne et al., 2010).

Denoting the diagonal elements of Λ_j by $\lambda_{j1}, \ldots, \lambda_{jK}$, uniqueness of B up to sign is ensured for models with more than two states 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). Again this condition can be checked by statistical tests.

The possible sign changes of the elements of B are another source of nonuniqueness. The precise condition is that each column of B can be multiplied by -1 without affecting the decomposition in (5). Hence, B is only locally unique which is sufficient for asymptotic inference. From the point of view of interpreting the results this nonuniqueness is also no problem because changing the signs of all elements in a column of B just means to consider negative instead of positive shocks or vice versa. Hence, the economist interpreting the impulse responses just needs to decide whether s/he is interested in positive or negative shocks.

This discussion suggests that statistically identified shocks may not have much meaning for economic analysis. In fact, the shocks identified by the statistical properties of the model may be mixtures of economically relevant shocks. However, there are two basic devices which may be helpful for associating statistically identified shocks with economic shocks of interest.

First, if the statistically identified shocks coincide with the economic shocks, their interpretation is straightforward. To find out whether we are in this lucky situation, we may test the identifying restrictions of a conventional identification scheme by means of statistical tests. If the restrictions are not rejected, we may impose them and then attach the usual economic interpretation to them. In case the economic identification is controversial, it is obviously an advantage to be able to test it against the data. Rejecting the restrictions may be seen as a signal of a problem. For instance, the underlying theory may simply be false. Of course, it may also be a deficiency of the statistical model which leads to a rejection of the restrictions. For example, there could be omitted variables, time-varying parameters, nonlinearities or errors-in-variables problems that do not allow certain shocks to be identified in the way assumed by the analyst. We will return especially to this issue in the empirical section. In any case, being able to test the economic identifying restrictions is an advantage because it can signal problems related to the interpretation of the shocks.

The second device that may be helpful in associating statistically identified shocks with economics derives from the changes in volatility in different periods during the sample. In some cases economic background knowledge may suggest different volatility of the shocks in different periods, which may be used for labeling the shocks. Again, this issue will be illustrated in the empirical section.

3.2. Estimation

We use classical ML estimation based on a log-likelihood derived from the conditional normality assumed in (4). The likelihood function is highly nonlinear which requires numerical optimization techniques. The objective function has several local optima in addition to those which follow from the identification issues discussed in the previous subsection. Moreover, the variances have to be bounded away from zero. In fact, the covariance matrices in the states must be nonsingular with determinants bounded away from zero. We impose restrictions on the eigenvalues of the state covariance matrices to ensure nonsingularity. Furthermore, the diagonal elements of the Λ_i matrices are bounded away from zero. An EM algorithm as described in Krolzig (1997) is used for the actual likelihood maximization task. Details are given in the Appendix.

Given the difficulties associated with the optimization of the likelihood function, classical residual based bootstrap methods are problematic for generating confidence intervals for the impulse responses (IRs). It has to be ensured that only bootstrap replications are considered in an area of the parameter space corresponding to the same parameterization as in the original estimation step. In particular, the same sign and ordering of the shocks has to be ensured. Sign changes of the shocks can be prevented by enforcing a particular instantaneous response of one of the variables. For example, a monetary policy shock increases the interest rate on impact. Finally, given that the MS model exploits patterns of vector heteroskedasticity, any potential resampling scheme must preserve second order features of the data. To account for these issues, resampling of IRs is performed throughout in the spirit of a fixed design wild bootstrap (Goncalves and Kilian, 2004). Conditionally on ML parameter estimates, bootstrap samples are determined as

$$\Delta y_t^* = \hat{\nu}_0 + \hat{\nu}_1 t + \hat{\alpha}\beta' y_{t-1} + \sum_{i=1}^{p} \hat{\Gamma}_i \Delta y_{t-i} + u_t^*. \tag{7}$$

In (7) $u_t^* = \psi_t \hat{u}_t$, where ψ_t is a random variable which is independent of the y_t and has a Rademacher distribution, that is, it has values 1 and -1, each with probability 0.5. Apart from preserving potential heteroskedasticity, multiplying the residual vectors $\hat{u}_t$ from the original estimation by a scalar quantity with mean zero and unit variance imitates the pattern of contemporaneous dependence of the data. Throughout, bootstrap parameter estimates θ^* of $\theta = \mathrm{vec}[\nu_0, \nu_1, \alpha, \Gamma_1, \ldots, \Gamma_p]$ and B^* of B are determined conditionally on the initially estimated diagonal elements in $\hat{\Lambda}_i$, $i = 2, \ldots, M$, and transition probabilities $\hat{p}_{ij}$, $i, j = 1, \ldots, K$, i.e., the relative variance parameters and transition probabilities are not subjected to resampling. Bootstrap IRs are obtained by nonlinear optimization of the log-likelihood with starting value being the vector of ML estimates. Apart from these modifications, the bootstrap confidence intervals are standard percentile intervals based on 1000 replications and using the 16th and 84th quantiles of the bootstrap distribution. Hence, we consider 68% confidence intervals in line with Peersman (2005).

3.3. Model selection

Choosing the number of volatility states is critical for this type of analysis. Standard tests are problematic for this purpose because some parameters are not identified under a null hypothesis of a smaller number of states. Although tests for the number of states have been proposed for this situation (e.g., Hansen, 1992, Garcia, 1998), we will rely on model selection criteria for choosing the number of states. They were found to work reasonably well for MS models in performance comparisons by Psaradakis and Spagnolo (2003, 2006).

Model selection criteria are also useful for comparing models with various types of restrictions even if some of the parameters may not be identified. For under-identified SVAR models the likelihood is the same as for the corresponding reduced form model. If two models with the same reduced form are compared, model selection criteria choose the more restricted model due to their penalty term for the number of parameters. This issue will be important in comparing different MS-SVAR models because at the time of model selection the identification properties may not be fully resolved (see Section 4.2 for further discussion).

4. Empirical analysis of a US system

4.1. The data

As mentioned earlier, we use the variables and quarterly US data from Peersman (2005). The variables are an oil price index (oil_t), a GDP index multiplied by 100 (q_t), a consumer expenditure index multiplied by 100 (p_t) and a 3-months interest rate (s_t).[1] We use the variables considered by Peersman except that we have multiplied output and prices by 100 to ensure a balanced scaling of the residual covariance matrices. This scaling is helpful for the nonlinear optimization of the log-likelihood. We also consider Peersman's sample period from 1980Q1–2002Q2 to ensure comparability of the results although longer series are available.[2]

There has been some discussion in the literature about changes in volatility of shocks during our sample period. In particular, it is a well established empirical fact that the volatility was reduced during the *Great Moderation* which started in the middle of the 1980s in the US (e.g., McConnell and Perez-Quiros, 2000, Mills and Wang, 2003, Stock and Watson, 2005). Thus, one may be able to use this and possibly other changes in volatility for identifying shocks.

Peersman uses his models to examine the causes of the economic downturn at the beginning of the new millennium. With this objective in mind one may consider other variables as well. For example, monetary aggregates such as M1 or Divisia variables (see Barnett, 1980) and other quantities related to the market for crude oil such as oil inventories (Kilian and Murphy, 2014) come to mind. Moreover, there has been some discussion of possible nonlinearity of the effects of oil price shocks (Hamilton, 2003; Kilian and Vigfusson, 2011). We do not consider such extensions because the main objective of the present study is to illustrate problems related to the interpretation of the shocks and to propose solutions to these problems.

4.2. Statistical analysis

We start from a similar model as Peersman (2005), the VECM in (1) with three lags. The model corresponds to Peersman's who considers a VAR model in first differences of the first three variables and the interest rate in levels. He also uses three lags. Regarding the choice of the lag order, we have used the same as Peersman although it is not difficult to find arguments for lower or higher lag orders. Using Peersman's choice seems reasonable because his analysis serves as a benchmark.

Several questions have to be addressed at the model specification stage in our setup. First, we have to decide on the number of volatility states. Then we have to check whether a statistical identification of shocks is possible and whether all or some of the economic identifying restrictions from Peersman are consistent with the data. In other words, we have to check whether the identified shocks can be given an economic interpretation.

In Table 1 the log-likelihood maxima and associated model selection criteria (AIC and SC) are given for a range of models. Comparing only unrestricted models, AIC favors a 3-state MS model while SC selects a 2-state MS model. None of the criteria goes for the VECM without MS. Within the group of 3-state models both AIC and SC reach their minima for a model with Peersman's zero restrictions imposed on the matrix of instantaneous effects, B. In contrast, considering only 2-state models, the version with the four

[1] The data are from the archive of the *Journal of Applied Econometrics* associated with Peersman (2005).

[2] We have also performed a similar analysis for an extended sample period with data from 1970Q1–2002Q2. It turns out that the model selection results and our main conclusions are remarkably robust.

Table 1
Comparison of MS-SVAR models for $y_t = (oil_t, q_t, p_t, s_t)'$ with lag order $p = 3$, intercept and linear trend term (sample period: 1980Q1–2002Q2).

Model	$\log L_T$	AIC	SC
VECM without MS	−62.41	264.8	436.6
2-state MS, unrestricted	−11.16	186.3	387.6
2-state MS, four zero restrictions on B	−13.39	182.8	374.2
2-state MS, zero restr. on B and long-run demand shock restr.	−14.03	182.1	371.0
2-state MS, zero restr. on B and long-run monetary shock restr.	−23.28	200.5	389.5
2-state MS, all restrictions	−27.39	206.8	393.3
3-state MS, unrestricted	7.46	177.1	412.7
3-state MS, state-invariant B	4.07	171.9	392.8
3-state MS, four zero restrictions on B	2.58	166.8	377.9
3-state MS, zero restr. on B and long-run demand shock restr.	0.06	169.9	378.5
3-state MS, zero restr. on B and long-run monetary shock restr.	−9.21	188.4	397.0
3-state MS, all restrictions	−12.13	192.3	398.4

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. Likelihood maximization subject to a bound of 0.01 for the λ_{ij}'s and 0.001 for the eigenvalues of the state covariance matrices.

Table 2
Estimates of structural parameters of 2-state MS-SVAR model for $y_t = (oil_t, q_t, p_t, s_t)'$ with lag order $p = 3$, intercept and linear trend term (sample period: 1980Q1–2002Q2).

Parameter	Unrestricted model		Demand neutrality + short-run restrictions		Fully restricted model	
	Estimate	Std. dev.	Estimate	Std. dev.	Estimate	Std. dev.
λ_{21}	0.012	0.006	0.095	0.037	0.013	0.010
λ_{22}	0.102	0.075	0.131	0.071	0.419	0.205
λ_{23}	0.843	0.489	1.062	0.526	1.194	0.586
λ_{24}	16.52	6.990	19.60	7.522	10.10	3.917
p_{11}	1	na	0.972	0.083	1	na
p_{22}	0.941	1.747	0.871	0.508	0.941	4.967

Note: Standard errors are obtained from the inverse of the outer product of numerical first order derivatives (*gradp*, Gauss 9.0). na stands for 'not available' due to an estimate at the boundary of the parameter space.

zero restrictions on B and the long-run neutrality of a demand shock (i.e., no permanent effect on output) is favored by both criteria. Both AIC and SC prefer some of the 3-state models over an unrestricted 2-state model. On the other hand, the overall minimum SC value is obtained for a 2-state model with zero restrictions on B and long-run neutrality of the demand shock. The less parsimonious AIC favors 3-state models, however. Thus, there is evidence for both 2-state and 3-state models. Given the relatively small sample size, estimation of these models was a challenge and in that sense a 3-state model may be over-parameterized. On balance we decided that it may still be worthwhile to continue with both types of MS models and compare the results.

Notice also that the evidence against a model without MS is quite strong, that is, the likelihood improves substantially when MS in volatility is allowed for. Moreover, it is reassuring that the 3-state model with unrestricted state covariance matrices does not have a much better likelihood than a model which imposes state invariant instantaneous effects. Neither AIC nor SC favor the fully unrestricted model over one with a state-invariant B. Further support for a state-invariant B is obtained from the LR test reported in Table 6. The p-value is 0.746. Hence, the null hypothesis of a state-invariant B cannot be rejected at common significance levels. Thus, allowing for changing volatility during the sample period and state-invariant instantaneous effects are both supported by the data.

A more difficult question is, however, whether the fact that AIC and SC select models with restrictions on B is evidence in favor of the restrictions or a reflection of a lack of identification. As mentioned in Section 3.3, in an under-identified model, the model selection criteria theoretically favor the one with fewer parameters. To investigate the identification issue, it is necessary to look at the λ_{ij} variance parameters. The estimates and their standard errors for some 2- and 3-state models are presented in Tables 2 and 3, respectively. Note that the order of the λ_{ij}'s for the unrestricted models is in principle arbitrary while for the restricted models the order is the one that is optimal for accommodating the restrictions. Because the λ_{2i}'s turn out to be ordered from smallest to largest in the fully restricted model, we use the same ordering for the unrestricted model.

Apparently the estimated λ_{2i}'s for the unrestricted 2-state model in Table 2 are all different. Whether they are significantly different is not clear, given the relatively large standard errors. The question is further explored in Table 4, where Wald tests are presented for null hypotheses of equality of the λ_{2i}'s. Notice that the λ_{ij}'s are identified even if they are identical. The estimated λ_{2i}'s have asymptotic normal distributions under standard assumptions. Hence, we use Wald tests based on that distribution. Since the number of parameters in our models is quite large relative to the number of sample observations, the estimate of the covariance matrix may be poor, however, and Wald tests may have poor small sample properties. Considering their p-values, the Wald tests leave open the possibility that the smallest λ_{2i}'s are identical. This outcome is not surprising given the relatively large standard deviations of the estimated λ_{2i}'s in Table 2. It reflects the limited sample information on the one hand and the complexity of the model on the other hand. In any case, overall there is at least weak evidence that all λ_{2i}'s may be distinct and, hence, the shocks are identified by purely statistical means. However, it may be worth keeping the problems related to these tests in mind.

If the model is fully identified, any restrictions on B reduce the dimensionality of the parameter space. Hence, using the log-likelihood maxima reported in Table 1, we can perform LR tests of the different sets of restrictions. They are shown in Table 6 and deliver the outcome suggested also by AIC and SC, namely that the long-run restriction for the monetary policy shock is clearly rejected and, hence, also the set of all restrictions jointly is not supported. One may argue that the test results hinge on the assumption of distinct λ_{2i}'s. While this is true, it may be worth remembering that, if some of the λ_{2i}'s were not distinct, the degrees of freedom for the LR tests would be reduced so that the actual p-values might actually be even smaller than stated in Table 6. Thus,

Table 3
Estimates of structural parameters of 3-state MS-SVAR models for $y_t = (oil_t, q_t, p_t, s_t)'$ with lag order $p = 3$, intercept and linear trend term (sample period: 1980Q1–2002Q2).

Parameter	Unrestricted, state-invariant B		Model with short-run restrictions		Demand neutrality + short-run restrictions		Fully restricted model	
	Estimate	Std. dev.	Estimate	Std. dev.	Estimate	Std. dev.	Estimate	Std. dev.
λ_{21}	0.048	0.037	0.049	0.026	0.063	0.038	26.99	27.42
λ_{22}	0.332	0.173	0.405	0.224	0.458	0.236	0.620	0.959
λ_{23}	1.025	0.582	1.719	0.846	1.824	0.873	0.991	0.967
λ_{24}	3.250	1.822	2.920	1.483	2.946	1.382	0.336	0.533
λ_{31}	0.220	0.162	0.275	0.302	0.331	0.433	2.672	1.865
λ_{32}	0.044	0.068	0.019	0.017	0.020	0.018	0.698	0.574
λ_{33}	1.768	1.733	1.391	1.279	1.381	1.270	2.912	3.013
λ_{34}	64.99	51.00	35.78	23.38	42.68	30.61	32.63	32.60
p_{11}	0.920	0.256	0.933	0.344	0.933	0.382	0.928	0.095
p_{21}	0.080	0.255	0.067	0.332	0.067	0.377	0.045	0.081
p_{12}	0.102	0.258	0.085	0.383	0.084	0.469	0.387	0.569
p_{22}	0.845	0.222	0.885	0.311	0.885	0.373	0.613	0.548
p_{23}	0.239	0.759	0.192	0.863	0.197	0.894	0.000	0.126
p_{33}	0.761	0.760	0.808	0.843	0.803	0.848	0.750	0.465

Note: By construction the columns of $\hat{P}$ add to unity, hence, only six of the nine elements are reported. Standard errors are obtained from the inverse of the outer product of numerical first order derivatives (*gradp*, Gauss 9.0).

Table 4
Wald tests for equality of λ_{2i}'s for the unrestricted 2-state MS-SVAR model from Table 2.

H_0	Test statistic	p-value
$\lambda_{21} = \lambda_{22}$	1.450	0.229
$\lambda_{21} = \lambda_{23}$	2.284	0.131
$\lambda_{21} = \lambda_{24}$	5.543	0.019
$\lambda_{22} = \lambda_{23}$	2.893	0.089
$\lambda_{22} = \lambda_{24}$	5.577	0.018
$\lambda_{23} = \lambda_{24}$	5.228	0.022

Table 5
Wald tests for equality of λ_{ij}'s for the 3-state MS-SVAR model with state-invariant, unrestricted B from Table 3.

H_0	Test statistic	p-value
$\lambda_{21} = \lambda_{22}, \lambda_{31} = \lambda_{32}$	3.19	0.202
$\lambda_{21} = \lambda_{23}, \lambda_{31} = \lambda_{33}$	3.60	0.166
$\lambda_{21} = \lambda_{24}, \lambda_{31} = \lambda_{34}$	4.43	0.109
$\lambda_{22} = \lambda_{23}, \lambda_{32} = \lambda_{33}$	2.18	0.336
$\lambda_{22} = \lambda_{24}, \lambda_{32} = \lambda_{34}$	1.81	0.405
$\lambda_{23} = \lambda_{24}, \lambda_{33} = \lambda_{34}$	3.11	0.211
$\lambda_{21} = \lambda_{22} = \lambda_{23} = \lambda_{24}$	10.4	0.016
$\lambda_{31} = \lambda_{32} = \lambda_{33} = \lambda_{34}$	2.93	0.403

null hypotheses that are rejected under the present assumptions would also be rejected if some λ_{2i}'s were actually equal. On the other hand, the short-run restrictions (zero restrictions for B) and the long-run neutrality restriction for the demand shock are not rejected at conventional levels within the 2-state MS model class. This conclusion may be affected by equal λ_{2i}'s. It is, however, supported by the fact that the estimated λ_{2i}'s in Table 2 do not change much when the zero restrictions and the neutrality restriction of a demand shock are imposed whereas a considerable change in the estimated λ_{2i}'s is observed when neutrality of the monetary shock is imposed in addition. These results are in contrast with Peersman (2005) who concludes from a comparison of conventional and sign restricted SVARs that the zero restrictions on the instantaneous effects for the oil price may be too stringent while he finds evidence for long-run neutrality of a monetary shock.

Turning now to the models with three states, the estimated λ_{ij}'s for different versions can be found in Table 3 and tests for pairwise equality of the λ_{ij}'s of the model with state-invariant B are shown in Table 5. In the 3-state MS model we need for uniqueness of B that for each pair $i, j \in \{1, \dots, K\}$, $i \neq j$, either $\lambda_{2i} \neq \lambda_{2j}$ or $\lambda_{3i} \neq \lambda_{3j}$. Hence, we test joint null hypotheses $H_0 : \lambda_{2i} = \lambda_{2j}$ and $\lambda_{3i} = \lambda_{3j}$, as shown in Table 5. The Wald tests do not reject these null hypotheses at conventional significance levels but they do reject that all diagonal elements of Λ_2 are identical. The test is also presented in Table 5 and has a p-value of 0.016. Obviously, such a complex model is difficult to estimate from our limited sample which is likely to undermine the power of our tests. Taking into account the results of all the tests, there is evidence that at least some λ_{ij}'s are distinct. Hence, tests of restrictions on B can be performed but have to be interpreted cautiously.

In the class of 3-state models in Table 1 AIC and SC both favor a model with the four zero restrictions on B specified in (3). Some LR tests of restrictions on B in the 3-state model are also presented in Table 6. They support a model with the four short-run zero restrictions on B but also do not reject the long-run neutrality of the demand shock. They reject the long-run money neutrality restriction, however.

These results are based on the degrees of freedom parameters obtained for a fully identified model. Given that the results in Table 5 provide weak support for such an assumption at best, the p-values in Table 6 are better thought of as upper bounds for the actual asymptotic p-values, as explained earlier. Hence, the fully restricted model with all of Peersman's restrictions is clearly rejected and so is the model with the four short-run restrictions and the long-run money neutrality restriction. The situation is less clear for the model with all but the long-run money neutrality restriction. The p-value obtained under the assumption of a χ^2-distribution with five degrees of freedom does not give rise to rejecting at conventional significance levels, the p-value being 0.155. However, this result may just reflect the lack of sufficient sample information against the null hypothesis.

Overall we conclude from our statistical analysis that a model without MS in the residual covariance is clearly inferior to models with MS. Both a 2-state and a 3-state MS model have some support from the data. Within the class of 2-state models the one with the four conventional zero restrictions from (3) on B and the long-run restriction associated with the demand shock is the favorite model. This model also has some support from the data in the 3-state class but here the situation is more ambiguous. Still, we will pay special attention to these models in the following. None of these models would be fully identified in a conventional setting. Hence, the interpretation of the resulting shocks is not obvious. In the next subsection we will discuss whether and how the volatility of the shocks can help in labeling them.

4.3. Analysis of states

In order to discuss the question how the MS structure can help in labeling the shocks, it is useful to consider the estimated residual

Table 6
LR tests of restrictions for MS-SVAR models for $y_t = (oil_t, q_t, p_t, s_t)'$ with lag order $p = 3$, intercept and linear trend term (sample period: 1980Q1–2002Q2).

Model	H_0	H_1	LR	df	p-value
2-state MS	$b_{12} = b_{13} = b_{14} = b_{24} = 0$	Unrestricted	4.46	4	0.347
	$b_{12} = b_{13} = b_{14} = b_{24} = 0, \xi_{23} = 0$	Unrestricted	5.74	5	0.332
	$b_{12} = b_{13} = b_{14} = b_{24} = 0, \xi_{24} = 0$	Unrestricted	24.24	5	0.000
	$b_{12} = b_{13} = b_{14} = b_{24} = 0, \xi_{23} = \xi_{24} = 0$	Unrestricted	32.46	6	0.000
3-state MS	State-invariant B	Unrestricted	6.78	10	0.746
	$b_{12} = b_{13} = b_{14} = b_{24} = 0$	State-inv. B	2.98	4	0.561
	$b_{12} = b_{13} = b_{14} = b_{24} = 0, \xi_{23} = 0$	State-inv. B	8.02	5	0.155
	$b_{12} = b_{13} = b_{14} = b_{24} = 0, \xi_{24} = 0$	State-inv. B	26.56	5	0.000
	$b_{12} = b_{13} = b_{14} = b_{24} = 0, \xi_{23} = \xi_{24} = 0$	State-inv. B	32.40	6	0.000

Note: LR $= 2(\log L_T - \log L_T^r)$, where L_T^r denotes the maximum likelihood under H_0 and L_T denotes the maximum likelihood under H_1 from Table 1.

Table 7
Estimated state covariance matrices ($\times 100$) of MS-SVAR models for $y_t = (oil_t, q_t, p_t, s_t)'$ with lag order $p = 3$, intercept and linear trend term (sample period: 1980Q1–2002Q2).

	Unrestricted 2-state MS-VAR	3-state MS-VAR with state invariant, unrestricted B
Σ_1	$\begin{bmatrix} 6.973 & -0.157 & 3.100 & 3.547 \\ & 19.375 & -1.294 & 8.916 \\ & & 3.911 & 1.680 \\ & & & 16.039 \end{bmatrix}$	$\begin{bmatrix} 8.928 & 1.621 & 3.811 & 5.440 \\ & 22.739 & -2.264 & 10.351 \\ & & 4.426 & 0.244 \\ & & & 12.494 \end{bmatrix}$
Σ_2	$\begin{bmatrix} 0.186 & 0.378 & -0.016 & -1.325 \\ & 7.357 & 2.184 & 1.954 \\ & & 1.085 & 3.901 \\ & & & 163.250 \end{bmatrix}$	$\begin{bmatrix} 0.461 & 0.202 & 0.091 & -0.224 \\ & 11.860 & 0.866 & 5.046 \\ & & 1.653 & 0.283 \\ & & & 17.352 \end{bmatrix}$
Σ_3		$\begin{bmatrix} 2.406 & -0.832 & 0.302 & -10.664 \\ & 13.999 & 5.492 & 29.569 \\ & & 2.899 & 15.869 \\ & & & 313.366 \end{bmatrix}$

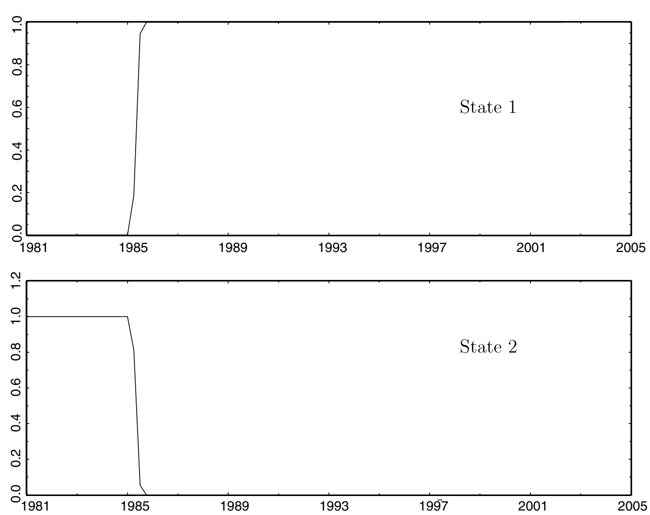

Fig. 1. Smoothed state probabilities of the unrestricted 2-state MS-SVAR model.

covariance matrices of 2- and 3-state MS models with unrestricted B matrix presented in Table 7. In an unrestricted 2-state model a substantially larger variance is observed for the interest rate equation. Thus, the second state is associated with high volatility in the interest rate. The state probabilities are plotted in Fig. 1 where it can be seen that State 2 is associated with the first half of the 1980s while State 1 corresponds to the Great Moderation period afterwards. Once reached, State 1 is never left during our sample period. This explains the estimate $\hat{p}_{11} = 1$ for the unrestricted model in Table 2. In other words, the estimated State 1 is an absorbing state.

Considering the estimated λ_{2i}'s, i.e., the relative variances in State 2 in Table 2, it is apparent that the last shock is the one with relatively high volatility in State 2 across all models, that is, irrespective of the restrictions which are imposed on B. Thus, the volatility analysis suggests that the last shock may be the monetary policy shock even in models which are not fully identified by conventional restrictions imposed on B. In particular, in the preferred model where some of the conventional identifying restrictions are imposed, one may suspect, taking into account the volatility of the shocks, that the last shock is a monetary policy shock. Looking at the associated state probabilities of the preferred model with all but the monetary policy shock neutrality restriction in Fig. 2, they are a bit different from those in Fig. 1. Again much of the first half of the 1980s is associated with State 2, but also a short period in 1988/89 and in 1992/93. The last period follows the Persian Gulf War (late 1990) which was associated with turbulence in the oil market and associated reactions of monetary policy (Kilian, 2008a,b). Hence, the high volatility shocks in State 2 in this model may be a mixture of monetary policy and oil price shocks. In other words, it is not obvious that the last shock can really be classified as a monetary policy shock in our preferred model. We will return to the issue of classifying the shocks when we discuss the IRs in the next subsection.

For the 3-state MS model the situation is slightly different. The state covariance matrices for the model with unrestricted, state-invariant B are given in Table 7. The last state is again one with a high volatility in the interest rate equation while in the first state the oil price equation exhibits much higher volatility than in the other states. The corresponding smoothed state probabilities are depicted in Fig. 3. They show that the third state is confined to the high volatility period in the first half of the 1980s but now the remaining period is subdivided in two states. In other words, the Great Moderation period during the 1980s and 1990s is divided up among the first and second states. It is not obvious to relate the periods assigned to a particular state with specific events associated with higher or lower volatility. In fact, in the covariance matrices associated with the first two states there is no uniform ordering of the variances. More precisely, the first three variances are larger in State 1 while the fourth variance is larger in State 2.

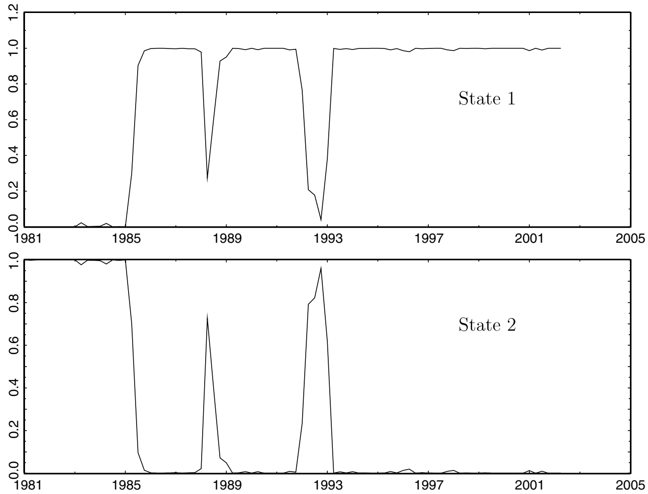

Fig. 2. Smoothed state probabilities of the preferred 2-state MS-SVAR model with four zero restrictions on B and a long-run demand shock neutrality restriction.

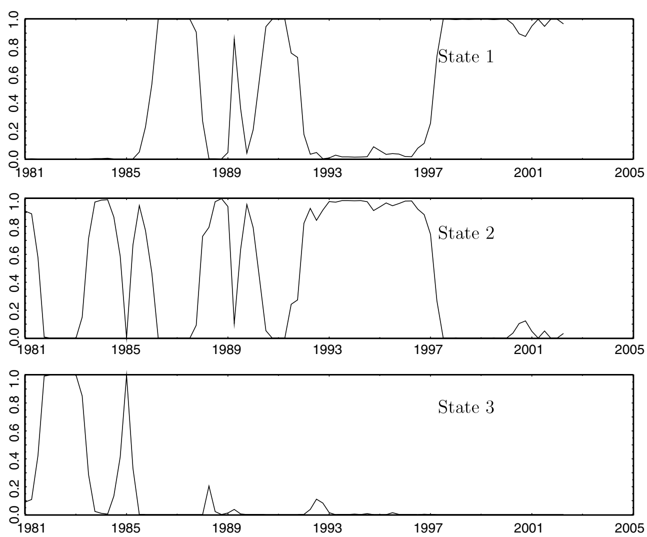

Fig. 3. Smoothed state probabilities of the 3-state MS-SVAR model with state-invariant, unrestricted B.

The only markedly higher variance is the first one in State 1 which is more than ten times the corresponding quantity in State 2.

Considering also the estimated λ_{ij}'s of the rejected, fully restricted 3-state MS model in Table 3 it can be seen that they are quite different from those of models which are not rejected by the data. Hence, the shocks in the latter models are likely to represent different shocks from those based on Peersman's restrictions. In other words, the shocks in the models not rejected by the data may not represent the same ones as in Peersman's model. It has to be seen whether the IRs may suggest appropriate labels for the shocks. IR analysis is considered next.

4.4. Impulse response analysis

It may be instructive to start the IR analysis by looking at the IRs obtained with the fully restricted 2-state MS model in Fig. 4. These IRs are quite similar to Peersman's when he uses conventional restrictions although our model has an MS structure and is hence different from his. An exception is the oil price reaction to a demand shock which becomes negative after a couple of quarters in Fig. 4 while it is significantly positive in Peersman's Fig. 1 even after a few years. Another difference to Peersman is the oil price response to a monetary policy shock. After an initial positive response it becomes quickly insignificant in Fig. 4 whereas it is significantly negative in Peersman's study.

Generally, a main difference to Peersman's results is that our confidence intervals for the IRs appear to be partly wider and less symmetric around the estimated IRs. There are a couple of factors that contribute to this outcome. First, Peersman uses a Bayesian approach to estimate IRs and construct confidence bounds, ignoring changes in volatility. In contrast, our approach is purely classical. Given that his IRs are median responses drawn from some posterior, their similarity to our classical IRs is a signal for the robustness of the results. Second, including the MS structure in the models increases the dimension of the parameter space and, hence, the estimation uncertainty. Moreover, ignoring volatility changes may lead to biased confidence intervals for IRs. Therefore, somewhat wider confidence intervals for our IRs are not surprising. The other difference to Peersman's IRs is that some of our confidence bands are much more asymmetric around the IRs. The median of the posterior distribution used by Peersman for estimating the IRs is within the confidence intervals by construction whereas the actual IRs of the system may in fact reach outside the confidence bands. Fry and Pagan (2011) mention this problem in the context of sign restricted IRs. Our confidence intervals may just reflect this feature. They should just be interpreted as an indication of estimation uncertainty in the IRs.

We emphasize that a contractionary monetary policy shock brings down output and the price level after some time and, hence, delivers plausible responses in our study as well as in Peersman's. While the effect on the price level is long lasting, the effect on output tapers off after some years due to the neutrality restriction imposed on the long-run effect.

In Fig. 5 we present the IRs from the unrestricted 2-state MS model. They are largely similar to those in Fig. 4. Actually the responses to the oil price shock (the first column in Fig. 5) are very similar to those in the fully restricted model in Fig. 4. We associate the second and third shocks with supply and demand on the basis of the IRs. The IRs show some differences to the fully restricted ones in Fig. 4 but they are qualitatively similar. Considering, for instance, the response of prices, the second shock has characteristics of a supply shock while the third one is recognized as a demand shock. Of course, such an interpretation assumes that our statistically identified shocks are actually supply and demand shocks.

Finally, the last shock in the unrestricted system, which was identified as a candidate for a monetary policy shock on the basis of the volatility analysis, has similar effects as in the restricted model except that now we have a 'price puzzle'. In other words, in the unrestricted model an interest rate increase goes together with a lasting increase in the price level which is in sharp contrast to the corresponding IRs in Fig. 4. There are a number of alternative explanations for this counter intuitive result. First, the fourth shock is not truly a monetary policy shock in the unrestricted model but perhaps a mixture of different economic shocks, as suggested by the analysis of the states in the previous subsection. Second, there may be important variables missing in the model so that the IRs do not properly reflect the actual responses to the shocks. The latter explanation has prompted earlier researchers to include forward-looking variables such as commodity prices in the model and there is no strong reason why the problem should not be present in the current model. Third, there may be other reasons such as model misspecification, errors-in-variables and the like. Such problems, if they exist, are apparently covered up in Peersman's models with conventional and sign restrictions.

To explore the problem further we show the IRs of the preferred 2-state MS-SVAR model in Fig. 6 which is not rejected by the data. The underlying model incorporates the short-run (zero) restrictions on B and the demand shock long-run neutrality restriction for output. Again all IRs are qualitatively similar to those of the

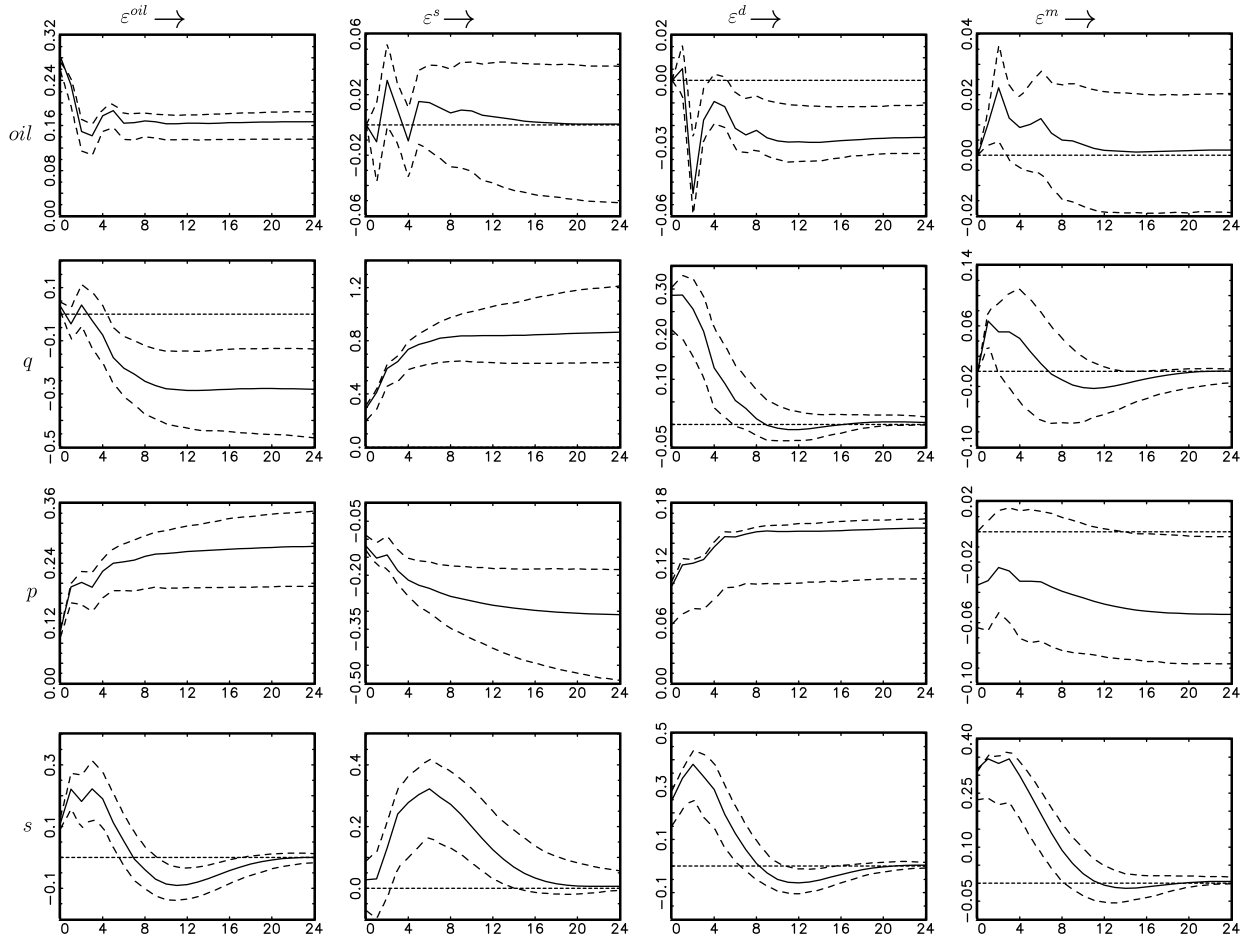

Fig. 4. Impulse responses with 68% confidence bounds of the fully restricted 2-state MS-SVAR model, with *oil*, *q*, *p*, *s* (ε^{oil}, ε^{s}, ε^{d}, ε^{m}) referring to oil prices, output, price level and short term interest rates (oil price, supply, demand and monetary shocks).

fully restricted model in Fig. 4 with one major exception. The price response to a contractionary monetary policy shock is grossly different from that in the fully restricted model. In other words, the 'price puzzle' persists. Given that the long-run restriction of the output response to a monetary policy shock is strongly rejected in our framework, we conclude that the model may not be a good one for studying the causes of the early millennium recession. Note that monetary policy shocks were regarded as potentially important for the slowdown. Hence, it is of particular concern that their impact is not captured properly by the model.

In Table 8 we show the forecast error variance decompositions, conditional on the states, associated with 'monetary policy shocks' obtained from the fully restricted and the preferred 2-state MS-SVAR models. Not surprisingly, in both models the forecast error variance components due to these shocks are quite different across states. For instance, the shocks contribute a much larger share to the forecast error variance of output and prices in State 2 than in State 1. From the point of view of this study the more interesting observation is, however, that the components differ substantially across the two models. For example, in State 2 the 'monetary policy shocks' contribute less than 10% to the forecast error variance of output for horizons of two or more years in the fully restricted model while their contribution in the preferred model is more than 50%. Clearly, if the last shock in the preferred model was viewed as a monetary policy shock and the associated forecast error variance components were interpreted accordingly, this could lead to substantially different conclusions than those drawn from a model with Peersman's restrictions.

To get further support for the result that misleading conclusions may be drawn from a fully restricted model we take a look at the preferred 3-state MS-SVAR model next. The IRs of the 3-state model with the four zero restrictions on *B* and the long-run demand shock neutrality, that is, our preferred 3-state MS-SVAR model, are depicted in Fig. 7. They are largely in line with Peersman's IRs except that there is again a positive response of the price index to a monetary policy shock, that is, the 'price puzzle' persists. We just mention that also in the 3-state MS model the 'price puzzle' disappears when we impose all of Peersman's conventional restrictions. Thus, these results are overall quite robust even across rather different models.

Given that the long-run neutrality restriction of the monetary policy shock for the output responses is strongly rejected by the data, we conclude that a model with Peersman's restrictions is a questionable tool for IR analysis more generally. Our analysis suggests that it may be necessary to include further variables in the model or modify the model in some other way to obtain reliable predictions of the reactions of the variables to the shocks of interest. Such a conclusion is difficult to draw in a conventional framework where the data cannot object to the just-identifying restrictions or in a setup using sign restrictions. Hence, the analysis demonstrates the virtues of our setup.

5. Conclusions

In this paper we consider the possibility of using changes in the volatility of the residuals of a VAR model to get identifying information for structural shocks. Volatility changes are modeled by means of an MS process. It is shown how this feature can be used for evaluating the validity of conventional restrictions. It is argued,

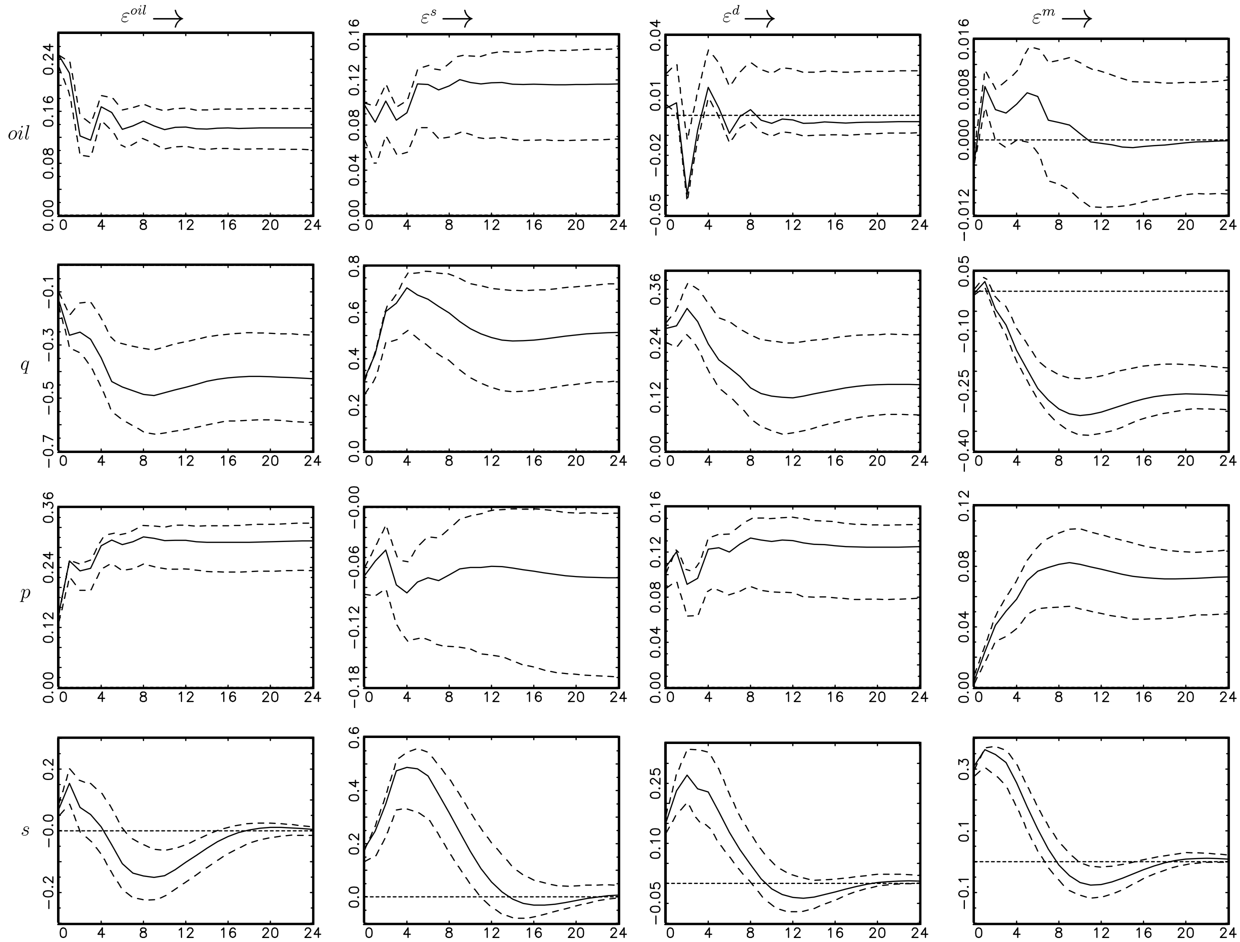

Fig. 5. Impulse responses with 68% confidence bounds of the unrestricted 2-state MS-SVAR model, with *oil*, *q*, *p*, *s* (ε^{oil}, ε^{s}, ε^{d}, ε^{m}) referring to oil prices, output, price level and short term interest rates (oil price, supply, demand and monetary shocks).

Table 8
Conditional forecast error variance components due to 'Monetary Policy Shocks' for 2-state MS-SVAR models for $y_t = (oil_t, q_t, p_t, s_t)'$ with lag order $p = 3$, intercept and linear trend term (sample period: 1980Q1–2002Q2).

Model	State	Variable	Forecast horizon					
			4	8	12	16	20	24
Fully restricted	1	*oil*	0.004	0.004	0.003	0.002	0.002	0.002
		q	0.006	0.003	0.002	0.001	0.001	0.001
		p	0.012	0.010	0.011	0.012	0.013	0.013
		s	0.389	0.342	0.314	0.308	0.308	0.308
	2	*oil*	0.557	0.528	0.441	0.375	0.329	0.295
		q	0.106	0.057	0.043	0.033	0.027	0.022
		p	0.240	0.215	0.230	0.249	0.262	0.270
		s	0.889	0.882	0.874	0.872	0.872	0.872
Preferred	1	*oil*	0.000	0.000	0.000	0.000	0.000	0.000
		q	0.005	0.016	0.021	0.023	0.024	0.024
		p	0.012	0.015	0.016	0.015	0.015	0.015
		s	0.220	0.144	0.133	0.131	0.131	0.131
	2	*oil*	0.044	0.037	0.033	0.031	0.030	0.029
		q	0.355	0.682	0.750	0.768	0.775	0.780
		p	0.409	0.482	0.495	0.496	0.496	0.496
		s	0.959	0.944	0.940	0.939	0.939	0.939

however, that shocks identified purely with statistical means may not be meaningful for economic analysis and it is discussed how identifying statistical information can be combined with economic restrictions for a meaningful interpretation of the shocks.

The issues involved have been discussed in the framework of a quarterly model for the US for oil prices, output, price level and a short-term interest rate. The system has been used previously for analyzing the causes of the early millennium slowdown of the US economy using alternatively conventional just-identifying and sign restrictions for the identification of the shocks. We have argued that these approaches have the drawback of leaving insufficient room for the data to object to the crucial assumptions underlying the analysis. In contrast, taking into account the statistical identifying information can disclose incompatibility of the data with conventional identifying or sign restrictions. It is shown that the US system is a questionable tool for analyzing the economic issues of interest in the present context because the data do not support the economic identifying assumptions.

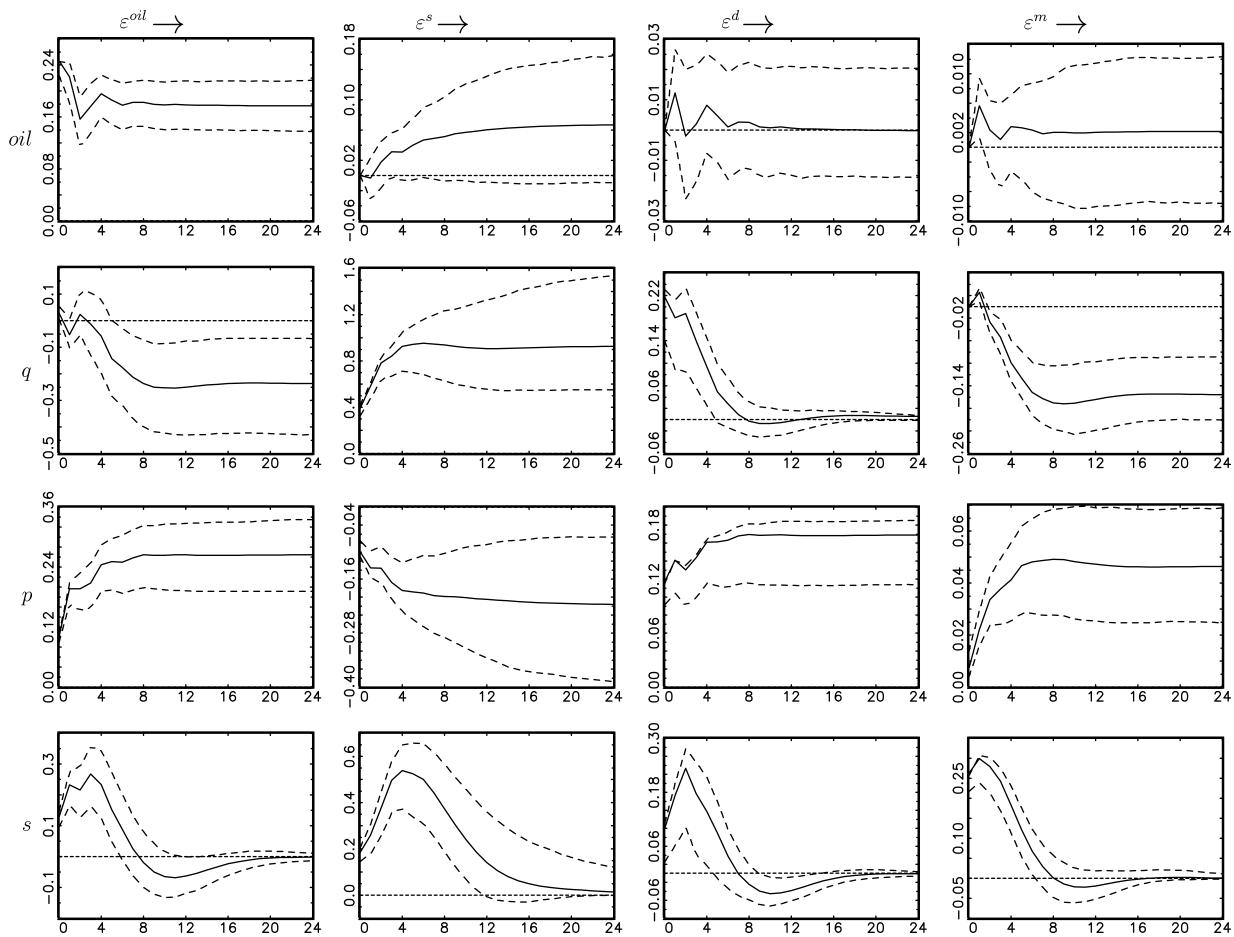

Fig. 6. Impulse responses with 68% confidence bounds of the 2-state MS-SVAR model with short-run restrictions and a long-run demand shock neutrality restriction, with *oil*, *q*, *p*, *s* (ε^{oil}, ε^{s}, ε^{d}, ε^{m}) referring to oil prices, output, price level and short term interest rates (oil price, supply, demand and monetary shocks).

Omitted variables may be a potential reason for the incompatibility of the conventional identifying restrictions and the data. Hence, future research of business cycle fluctuations may want to consider systems with additional or other variables which capture the transmission of monetary policy or may be of importance as explanatory factors. For example, forward-looking variables such as commodity prices or monetary aggregates may be included. With a view on the early millennium slowdown one may also want to consider variables related to financial markets or the wealth effects associated with the financial market contraction. Alternatively, one may consider adding a further shock which takes care of effects due to omitted variables as in Rigobon and Sack (2003).

Of course, there could be other reasons for rejecting the previously used identification assumptions. For instance, there may be errors-in-variables or model deficiencies such as nonlinearities or varying parameters not captured by the present setup. Such features may require using a different model class altogether.

Acknowledgments

We thank Jurgen Doornik, Lutz Kilian and participants of a *Workshop on "Modern Economic Forecasting: Alternative and Complimentary Approaches"*, Oslo, 31 May–1 June 2010, as well as two anonymous referees for helpful comments on an earlier version of this paper. Part of the research for this study was conducted while the first author was a Fernand Braudel Fellow in the Department of Economics at the European University Institute, Florence. The first author also gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (HE 2188-3/1).

Appendix. The EM algorithm

The EM algorithm presented in Krolzig (1997) is used. We adapt it to the case where the state covariances are parameterized as in (5) and provide information on specific implementations. Computations are performed with Gauss 9.0.

Notation and definitions

Define

$$\xi_t = \begin{bmatrix} \mathbf{I}(s_t = 1) \\ \vdots \\ \mathbf{I}(s_t = M) \end{bmatrix}, \quad \text{thus}, E(\xi_t) = \begin{bmatrix} \Pr(s_t = 1) \\ \vdots \\ \Pr(s_t = M) \end{bmatrix},$$

where $\mathbf{I}(\cdot)$ is an indicator function which is one if the condition in the argument holds and zero otherwise. Define

$$\xi_{t|s} = E(\xi_t | Y_s) = \begin{bmatrix} \Pr(s_t = 1|Y_s) \\ \vdots \\ \Pr(s_t = M|Y_s) \end{bmatrix},$$

where $Y_s = (y_1, \ldots, y_s)$. Note that

$$\xi_{t+1|t} = P\xi_{t|t}, \quad t = 0, \ldots, T-1,$$

where

$$P = \begin{bmatrix} \Pr(s_{t+1} = 1|s_t = 1) & \cdots & \Pr(s_{t+1} = 1|s_t = M) \\ \vdots & \ddots & \vdots \\ \Pr(s_{t+1} = M|s_t = 1) & \cdots & \Pr(s_{t+1} = M|s_t = M) \end{bmatrix}$$

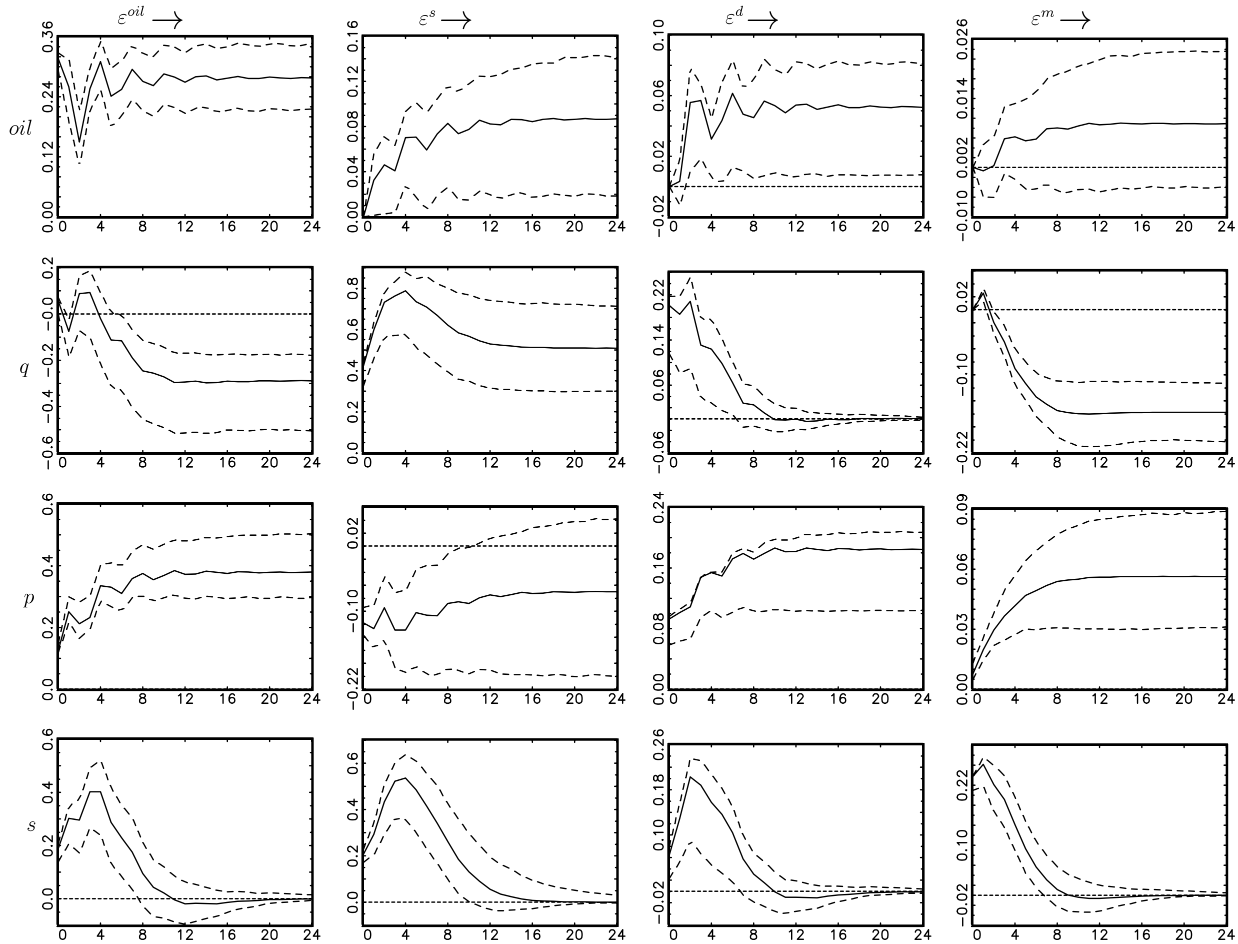

Fig. 7. Impulse responses and 68% confidence bounds for the 3-state MS-SVAR model with short-run restrictions and long-run demand shock neutrality restriction, with *oil*, *q*, *p*, *s* (ε^{oil}, ε^{s}, ε^{d}, ε^{m}) referring to oil prices, output, price level and short term interest rates (oil price, supply, demand and monetary shocks).

is the transition matrix. Moreover, define

$$\eta_t = \begin{bmatrix} f(y_t|s_t = 1, Y_{t-1}) \\ \vdots \\ f(y_t|s_t = M, Y_{t-1}) \end{bmatrix},$$

where

$$f(y_t|s_t = m, Y_{t-1}) = (2\pi)^{-K/2} \det(\Sigma_m)^{-1/2} \exp\left\{-\frac{1}{2}u_t'\Sigma_m^{-1}u_t\right\},$$

with

$$\Sigma_1 = BB', \qquad \Sigma_m = B\Lambda_m B', \quad m = 2, \ldots, M.$$

Furthermore, the following notation is used:

$\odot$ elementwise multiplication,
$\oslash$ elementwise division,
$\mathbf{1}_M = (1, \ldots, 1)'$ is an $(M \times 1)$ vector of ones,
$\theta = \text{vec}[\nu_0, \nu_1, \alpha, \Gamma_1, \ldots, \Gamma_p]$,
$Z_{t-1}' = [1, t, (\beta' y_{t-1})', \Delta y_{t-1}', \ldots, \Delta y_{t-p}']$.

EM algorithm

Starting values

$$P = M^{-1}\mathbf{1}_M\mathbf{1}_M',$$

$$\begin{aligned}\hat{\theta} &= \text{vec}[\hat{\nu}, \hat{\alpha}, \hat{\Gamma}_1, \ldots, \hat{\Gamma}_p] \\ &= \left[\sum_{t=1}^{T} Z_{t-1}Z_{t-1}' \otimes I_K\right]^{-1} \sum_{t=1}^{T}(Z_{t-1} \otimes I_K)\Delta y_t,\end{aligned}$$

$$B = \left(T^{-1}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t'\right)^{1/2} + B^0 \quad \text{with } \hat{u}_t = \Delta y_t - (Z_{t-1}' \otimes I_K)\hat{\theta}$$

and B^0 a matrix of small random numbers,

$\Lambda_m = I_K, \quad m = 2, \ldots, M,$

$\xi_{0|0} = \mathbf{1}_M/M.$

To ensure the detection of some 'global' maximum of the log-likelihood we use at least 10 000 distinct initial parameter choices for the elements in B.

Expectation step

For given $P, \theta, \Sigma_m,\ m = 1, \ldots, M$, and $\xi_0 = \xi_{0|0}$ compute

$$\eta_t, \qquad \xi_{t|t-1} = P\xi_{t-1|t-1}, \qquad \xi_{t|t} = \frac{\eta_t \odot \xi_{t|t-1}}{\mathbf{1}_M'(\eta_t \odot \xi_{t|t-1})},$$

$$t = 1, \ldots, T,$$

(choose $\xi_{iT|T} \leq \xi_{jT|T}$ for $i < j$ to avoid label switching, that is, the problem of iterating between likelihoods which correspond to different orderings or labeling of the states)

$$\xi_{t|T} = \left(P'(\xi_{t+1|T} \oslash P\xi_{t|t})\right) \odot \xi_{t|t}, \quad t = T-1, \ldots, 0,$$

and

$$\xi^{(2)}_{t|T} = \mathrm{vec}(P') \odot \left[(\xi_{t+1|T} \oslash P\xi_{t|t}) \otimes \xi_{t|t}\right], \quad t = 0, 1, \ldots, T-1.$$

Maximization step

Estimate P

$$\mathrm{vec}(\hat{P}') = \left(\sum_{t=0}^{T-1} \xi^{(2)}_{t|T}\right) \oslash \left(\mathbf{1}_M \otimes \sum_{t=1}^{T} \xi_{t|T}\right).$$

Estimate B and Λ_m

Define $\hat{u}_t = \Delta y_t - (Z'_{t-1} \otimes I_K)\hat{\theta}$ and $T_m = \sum_{t=1}^{T} \xi_{mt|T}$ and estimate B and Λ_m, $m = 2, \ldots, M$, by minimizing

$$l(B, \Lambda_2, \ldots, \Lambda_M) = T \log|\det(B)| + \frac{1}{2}\mathrm{tr}\left(B'^{-1}B^{-1}\sum_{t=1}^{T}\xi_{1t|T}\hat{u}_t\hat{u}'_t\right) + \sum_{m=2}^{M}\left[\frac{T_m}{2}\log\det(\Lambda_m) + \frac{1}{2}\mathrm{tr}\left(B'^{-1}\Lambda_m^{-1}B^{-1}\sum_{t=1}^{T}\xi_{mt|T}\hat{u}_t\hat{u}'_t\right)\right]$$

possibly subject to restrictions on B from (3) and impose a lower bound of 0.01 for the diagonal elements of Λ_m, $m = 2, \ldots, M$, to avoid singularity of the covariance matrix. Then define

$$\widehat{\Sigma}_1 = \widehat{B}\widehat{B}', \qquad \widehat{\Sigma}_m = \widehat{B}\hat{\Lambda}_m\widehat{B}' \quad m = 2, \ldots, M.$$

Estimate θ

$$\hat{\theta} = \left[\sum_{m=1}^{M}\left(\sum_{t=1}^{T}\xi_{mt|T}Z_{t-1}Z'_{t-1}\right) \otimes \widehat{\Sigma}_m^{-1}\right]^{-1} \times \sum_{t=1}^{T}\left(\sum_{m=1}^{M}\xi_{mt|T}Z_{t-1} \otimes \widehat{\Sigma}_m^{-1}\right)\Delta y_t.$$

Iterate estimation of B, Λ_m and θ until convergence.

Estimate ξ_0

$\xi_{0|0} = \xi_{0|T}$.

The expectation and maximization steps are iterated until convergence. We only consider models where all eigenvalues of $\hat{\Sigma}_1, \ldots, \hat{\Sigma}_M$ are greater than 0.001.

Likelihood function and convergence criteria

We use relative changes in the value of the log-likelihood function and the parameter values as convergence criteria.

The log-likelihood is evaluated as follows. For given P, θ, Σ_m, $m = 1, \ldots, M$, and $\xi_{0|0}$ compute for $t = 1, \ldots, T$,

$$\eta_t, \qquad \xi_{t|t-1} = P\xi_{t-1|t-1}, \qquad \xi_{t|t} = \frac{\eta_t \odot \xi_{t|t-1}}{\mathbf{1}'_M(\eta_t \odot \xi_{t|t-1})}.$$

Then

$$\log L_T = \sum_{t=1}^{T}\log f(y_t|Y_{t-1}),$$

where

$$f(y_t|Y_{t-1}) = \sum_{j=1}^{M}\Pr(s_t = j|Y_{t-1})f(y_t|s_t = j, Y_{t-1}) = \xi'_{t|t-1}\eta_t.$$

Note that the $\xi_{t|t-1}$ are not the smoothed transition probabilities but are obtained from the ones based on the given parameter values, that is, based on the parameter values obtained in a particular step of the estimation algorithm with $\xi_{0|0} = \xi_{0|T}$ being typically the smoothed estimate of the initial state.

Estimation of standard errors

Let γ_1 be the vector of all parameters in θ. Moreover, γ_2 consists of vec(B), the diagonal elements of Λ_m, $m = 2, \ldots, M$, and all $M(M-1)$unrestricted parameters in P (recalling that the columns of P sum to one). Let $\gamma = (\gamma'_1, \gamma'_2)'$. We use the outer product of numerical first order derivatives (&gradp, Gauss 9.0)

$$S = \sum_{t=1}^{T}\frac{\partial l_t(\gamma)}{\partial\gamma}\frac{\partial l_t(\gamma)}{\partial\gamma'}.$$

Standard errors for parameter estimates are determined as square roots of the diagonal elements of the inverse of this matrix under the presumption that the matrix is blockdiagonal with respect to γ_1 and γ_2.

References

Amisano, G., Giannini, C., 1997. Topics in Structural VAR Econometrics, second ed. Springer, Berlin.
Barnett, W.A., 1980. Economic monetary aggregates: an application of index number and aggregation theory. J. Econometrics 14, 11–48.
Blanchard, O., Quah, D., 1989. The dynamic effects of aggregate demand and supply disturbances. Amer. Econ. Rev. 79, 655–673.
Canova, F., De Nicoló, G., 2002. Monetary disturbances matter for business fluctuations in the G-7. J. Monetary Econ. 49, 1131–1159.
Faust, J., 1998. The robustness of identified VAR conclusions about money. Carnegie–Rochester Conf. Ser. Public Policy 49, 207–244.
Fry, R., Pagan, A., 2011. Sign restrictions in structural vector autoregressions: a critical review. J. Econom. Lit. 49, 938–960.
Gali, J., 1992. How well does the IS-LM moldel fit postwar US data. Quart. J. Econ. 107, 709–738.
Garcia, R., 1998. Asymptotic null distribution of the likelihood ratio test in Markov switching models. Internat. Econom. Rev. 39, 763–788.
Goncalves, S., Kilian, L., 2004. Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. J. Econometrics 123, 89–120.
Hamilton, J.D., 2003. What is an oil shock? J. Econometrics 113, 363–398.
Hansen, B.E., 1992. The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP. J. Appl. Econometrics 7, S61–S82. (erratum: 11, 195–198).
Kilian, L., 2008a. The economic effects of energy price shocks. J. Econom. Lit. 46, 871–909.
Kilian, L., 2008b. Exogenous oil supply shocks: how big are they and how much do they matter for the US economy? Rev. Econ. Stat. 90, 216–240.
Kilian, L., Lewis, L.T., 2011. Does the Fed respond to oil price shocks? Econ. J. 121, 1047–1072.
Kilian, L., Murphy, D.P., 2014. The role of inventories and speculative trading in the global market for crude oil. J. Appl. Econ. 29, 454–478.
Kilian, L., Vigfusson, R.J., 2011. Are the responses of the US economy asymmetric in energy price increases and decreases? Quant. Econ. 2, 419–453.
King, R.G., Plosser, C.I., Stock, J.H., Watson, M.W., 1991. Stochastic trends and economic fluctuations. Amer. Econ. Rev. 81, 819–840.
Krolzig, H.-M., 1997. Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis. Springer-Verlag, Berlin.
Lanne, M., Lütkepohl, H., 2010. Structural vector autoregressions with nonnormal residuals. J. Bus. Econom. Statist. 28, 159–168.
Lanne, M., Lütkepohl, H., Maciejowska, K., 2010. Structural vector autoregressions with Markov switching. J. Econom. Dynam. Control 34, 121–131.
Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin.
McConnell, M.M., Perez-Quiros, G., 2000. Output fluctuations in the United States: what has changed since the early 1980's? Amer. Econ. Rev. 90, 1464–1476.
Mills, T.C., Wang, P., 2003. Have output growth rates stabilised? Evidence from the G-7 economies. Scott. J. Polit. Econ. 50, 232–246.
Nakov, A., Pescatori, A., 2010. Oil and the great moderation. Econ. J. 120, 131–156.
Normandin, M., Phaneuf, L., 2004. Monetary policy shocks: testing identification conditions under time-varying conditional volatility. J. Monetary Econ. 51, 1217–1243.
Pagan, A.R., Pesaran, M.H., 2008. Econometric analysis of structural systems with permanent and transitory shocks. J. Econom. Dynam. Control 32, 3376–3395.
Peersman, G., 2005. What caused the early millennium slowdown? Evidence based on vector autoregressions. J. Appl. Econometrics 20, 185–207.
Psaradakis, Z., Spagnolo, N., 2003. On the determination of the number of regimes in Markov-switching autoregressive models. J. Time Ser. Anal. 24, 237–252.
Psaradakis, Z., Spagnolo, N., 2006. Joint determination of the state dimension and autoregressive order for models with Markov regime switching. J. Time Ser. Anal. 27, 753–766.
Rigobon, R., 2003. Identification through heteroskedasticity. Rev. Econ. Stat. 85, 777–792.
Rigobon, R., Sack, B., 2003. Measuring the reaction of monetary policy to the stock market. Quart. J. Econ. 118, 639–669.
Sentana, E., Fiorentini, G., 2001. Identification, estimation and testing of conditionally heteroskedastic factor models. J. Econometrics 102, 143–164.
Sims, C.A., 1980. Macroeconomics and reality. Econometrica 48, 1–48.
Stock, J.H., Watson, M.W., 2005. Understanding changes in international business cycle dynamics. J. Eur. Econom. Assoc. 3, 968–1006.
Uhlig, H., 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. J. Monetary Econ. 52, 381–419.