European University Institute

DEPARTMENT OF ECONOMICS

Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity

Aleksei Netšunajev

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, March 2013

European University Institute
Department of Economics

Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity

Aleksei Netšunajev

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Helmut Lütkepohl, DIW Berlin and Freie Universität (External Supervisor)
Prof. Fabio Canova, European University Institute
Prof. Helmut Herwartz, Georg-August-Universität Göttingen
Prof. Markku Lanne, University of Helsinki

ACKNOWLEDGMENTS

It is a pleasure to express my appreciation to the people who have contributed to the completion of this work. I want to thank my parents, Natalja and Vladimir, and my wife Svetlana for their support and patience. I am grateful to my supervisor, Professor Helmut Lütkepohl, who encouraged and guided me through the preparation of the thesis. I am also grateful to Professors Arpad Abraham, Nicola Pavoni, Peter Hansen, and Fabio Canova. from the European University Institute for their valuable comments on parts of the thesis. I am also thankful for the comments I received during various seminars and conferences where I presented each of the papers that this thesis is comprised of.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

ABSTRACT

Structural vector autoregressions are of great importance in applied macroeconometric work. The main difficulty associated with structural analysis is to identify unique shocks of interest. In a conventional approach this is done via zero or sign restrictions. Heteroskedasticity is proposed for use in identification. Under certain assumptions when volatility of shocks changes over time, unique shocks can be obtained. Then formal testing of the restrictions and impulse response analysis can be performed. In this thesis I show how identification via heteroskedasticity can be used in different contexts. In the first chapter I analyze the dynamics of trade balances in response to macroeconomic shocks. I show that identifying restrictions, which are known in the literature, are rejected for two out of seven countries. Partially identified models fail to provide enough information to fully identify shocks. The second chapter, coauthored with my supervisor, demonstrates how one can benefit from identification via heteroskedasticity when sign restrictions are used. The approach is illustrated with a model of the crude oil market. It is shown that shocks identified via previously known sign restrictions are in line with the properties of the data. Use of tighter restrictions uncovers that the approach can be discriminative. The third chapter reconsiders the conflicting results in the debate on the effects of technology shocks on hours worked. Using six ways of identifying technology shocks, I find that not all of them are supported by the data. There is no clear-cut evidence in favor of positive reaction of hours to technology shocks. However, it is plausible for real wage and disentangled investment-specific and neutral technology shocks, even though conventional identification of the latter shocks is rejected.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Contents

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Introduction

Structural vector autoregressions (VAR) are of great importance in applied macroeconometric work. The main difficulty arising in the analysis is the identification of the structural shocks of interest. To obtain unique and economically interpretable shocks, one can usually use some theory describing the relationships of interest. Then exclusion of certain instantaneous or long-run effects on the variables, or in other words assuming zero effects helps to identify the shocks. Another way of identifying structural shocks of interest is based on sign restrictions, when one admits the whole range of shocks that are in line with the prespecified positive or negative effects.

These ways of identification have a common undesirable feature, namely there is no place for the data to speak up against the restrictions. The zero restrictions just identify shocks and are not testable. The sign restrictions exclude the reaction that is not in line with the assumed effects. However, recent proposals have been made to use heteroskedasticity to extract additional identifying information. When the volatility of shocks changes over time, the sample size can be partitioned according to the changes and under certain assumptions, unique shocks are obtained. These shocks are statistically identified, yet they may not be interpretable economically. Formal testing of the restrictions and impulse response analysis reveals if the economically meaningful shocks are in line with the properties of the data.

In the thesis I use the Markov switching (MS) mechanism to model and determine the changes in the volatility regimes of shocks endogenously from the data. The expectation maximization algorithm and maximum likelihood is used to estimate the models.

The thesis consists of three chapters that are separate research papers in structural time series econometrics. It is shown how identification via heteroskedasticity can be used in the context of zero and sign restrictions. The methodology is illustrated with several well known examples in the literature.

In the first chapter I analyze the dynamics of trade balances in response to macroeconomic shocks in the structural vector autoregressive framework. I use Markov switching vector autoregression models with volatility changes in the residuals, in combination with

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

conventional procedure to identify structural shocks. Quarterly data on relative output, real exchange rates and trade balances are analyzed. The identifying restrictions are rejected for two out of seven countries. The rejection comes from the scheme imposed to identify nominal shocks. Partially identified models fail to provide enough information to fully identify shocks. Impulse response analysis is conducted for the countries for which the structural shocks can be identified. I find nominal shocks to have permanent positive effects on trade balances. Volatility analysis shows that countries are different in terms of the volatility structure. Therefore, when one is dealing with a group of countries the identification scheme may not be suitable for all of them.

The second chapter was written in collaboration with Professor Helmut Lütkepohl, my thesis supervisor. My contribution to the paper is computational. I was in charge of the empirical part of the paper, namely estimation of the models, testing of the restrictions and impulse response analysis. The chapter focuses on the sign restricted structural VAR and shows how one can use identification via heteroskedasticity in the context of sign restrictions. Although in an ideal setting the sign restrictions specify shocks of interest, sign restrictions may be invalidated by measurement errors, data adjustments or omitted variables. So far there are no techniques for validating the shocks identified via such restrictions. The approach is illustrated by considering a small model for the market of crude oil. It is shown that shocks identified via sign restrictions are in line with the properties of the data. The restrictions used in the example are weak and it may be difficult to argue against them. Using similar but tighter restrictions, it is shown that the approach can be discriminative.

In the third chapter I reconsider the conflicting results in the debate connected to the effects of technology shocks on hours worked. Given the considerable dissatisfaction with the just-identifying long-run restrictions, I analyze whether the restrictions used in the literature are consistent with the properties of the data. In the chapter, changes in volatility of shocks, as well as non-linearities in the intercept term are modeled with the Markov switching mechanism. Using six ways of identifying technology shock, I find that not all of them are consistent with the data. Identification of permanent technology shocks, non-permanent technology shocks and permanent real wage shocks is supported by the data. Furthermore, permanent TFP shocks and permanent technology shocks after controlling for capital tax have less support from the data. Finally, disentangling investment specific and neutral technology shocks is not supported by the data, however, a neutral technology shock can be identified in the system. There is not strong evidence in favor of positive reaction of hours to different technology shocks. But it is plausible for real wage shocks and an investment-specific – neutral technology shocks tandem. The latter result is achieved even though conventional

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

identification of the shocks is rejected by the data.

The identification via heteroskedasticity is a useful tool that one should not neglect in a structural VAR analysis. Combining it with the conventional approach, one can either bring convincing arguments in favour of the identified shocks or conclude that the identification scheme is not supported by the data.

Chapter 1

Trade Balance Dynamics in Response to Macroeconomic Shocks: Conventional Identification versus Identification via Heteroskedasticity

Aleksei Netšunajev

1.1 Introduction

The world economy has become increasingly integrated in recent years. The rising volume of international trade in goods as a result of contemporary organization of production schemes, when the production of single elements of a final commodity is diffused among countries is a well known phenomenon. Trade in services has also become increasingly important during the last couple of decades. All this suggests that international trade flows are subject to a range of macroeconomic shocks. Hence, understanding the reaction to shocks is essential from a number of different perspectives, including domestic and international policy coordination. In this paper I focus on trade balance adjustment that could be directly interpreted as, for instance, reaction to monetary or fiscal policy changes. The theoretical and empirical model from Prasad (1999) is used as a benchmark in the current analysis. Prasad considers relative output, real exchange rate and trade balance of the G7 [1] countries in the structural vector

[1] The following ISO notation is used for the countries: Canada - CA, Italy - IT, United Kingdom - GB, Germany - DE, Unites States - US, Japan - JP, France - FR

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

autoregressive (SVAR) framework to analyze the dynamics of trade balance in response to macroeconomic shocks. The econometric analysis is conducted separately for each country allowing the author to provide countrywide quantitative estimates of the effects. In the conventional structural VAR analysis as used by Prasad (1999) the major issue is to identify structural shocks.

Prasad (1999) uses a strategy proposed by Blanchard and Quah (1989), where the empirical model is identified by imposing a set of long-run restrictions derived from a theoretical model. The restrictions are just-identifying and, hence, cannot be checked against the data by statistical tests. The restrictions proposed by Prasad (1999) may be subject to criticism on the following grounds. Firstly, they come from a stylized IS-LM model which may not be the best approximation of the data generation process. Hence, it is relevant to question if one should carry over the features of the models to the SVAR. Secondly, the economic implications of the restrictions might be controversial, as will be discussed later. In the recent literature statistical properties of the data have been used to extract additional identifying information.

In SVAR analysis identification via heteroskedasticity has been proposed and used by Rigobon (2003), Rigobon and Sack (2003), and Lanne and Lütkepohl (2008). The underlying assumption of those papers is that there are exogenously generated changes in the volatility of the shocks. Furthermore, one could extract the information about the volatility regimes from the data and partition the sample accordingly. Then the identification of the shocks is based on the assumption that the effects of shocks are the same regardless of the volatility regime in which they occur. In other words, the impulse responses are invariant throughout the sample period whereas the volatility of the shocks changes.

Lanne and Lütkepohl (2010), Lanne, Lütkepohl and Maciejowska (2010) and later Herwartz and Lütkepohl (2011) consider Markov switching vector error correction (VECM) models and show how residual heteroskedasticity could be used for identification purposes. As pointed out by Lanne et al. (2010) analysis based solely on statistical information could produce economically meaningless and unattractive results. That means, shocks and impulse responses may provide little information on the underlying economic processes. Hence, combining conventional and statistical identifying information could be useful and give more evidence in favor of a particular identification procedure which would shed more light on the nature of economic processes.

The present paper adapts the common approach of Lanne et al. (2010) and Herwartz and Lütkepohl (2011) and applies it to the trade balance analysis. I model the conditional heteroskedasticity of the residuals assuming Markov switching structure. Going for Markov

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

switching heteroskedasticity models has two purposes in the present study. Firstly, using specific covariance matrix parameterization I can formally test for the validity of possible restrictions and let the data speak, either in favor of or against the imposed identification structure. Secondly, it allows me to capture unobserved volatility regimes (states) of the considered empirical models and detect different behavior of the economic indicators conditional to the regimes.

Estimated state probabilities and variances of the variables usually have interesting economic interpretations. One would typically associate state probabilities with specific events in the economies and pin down the reaction of the variables to regime shifts via estimated variances. Thus the paper contributes to the literature in the following ways. Firstly, recent econometric and identification methodology is applied to the structural VAR case. The study of Prasad (1999) is reconsidered and the issue of identification of the shocks is addressed, giving the data the possibility to speak up against the imposed identifying restrictions.

Secondly, the results from the MS models show how G7 countries' trade balance and other aggregate variables behave in different volatility periods, allowing one to see which indicators are most vulnerable to economic fluctuations. Impulse responses shed light on the response of the trade balance to shocks and transmission mechanisms of the shocks.

The paper is structured as follows. In Section 1.2 the theoretical model and a SVAR with conventional identification procedure used by Prasad (1999) is shortly discussed. In Section 1.3 the MS extension of the econometric model that allows for statistical identification of the shocks is presented. Description of the data and detailed empirical analysis follows in Section 1.4. Conclusions are provided in the last section. All computations and estimations are done in Matlab.

1.2 Theoretical model and conventional SVAR Model

This section briefly presents the theoretical model that illustrates the main channels through which different types of macroeconomic shocks influence the dynamics of the trade balance. The model is adapted from Prasad (1999) and it is a stochastic version of the Mundell-Fleming model that incorporates sluggish price adjustment. With the exception of the interest rate and the trade balance, the variables in the model are in logarithms and, except for the trade balance, are expressed as deviations of domestic levels from foreign levels of the corresponding variables. The model reads:

$$y_t^d = d_t + \eta(s_t - p_t) - \sigma(i_t - E_t[p_{t+1} - p_t]) \tag{1.1}$$

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

$$p_t = (1-\theta)E_{t-1}p_t^e + \theta p_t^e, 0 < \theta \leq 1 \tag{1.2}$$

$$m_t^s - p_t = y_t - \lambda i_t \tag{1.3}$$

$$i_t = E_t[s_{t+1} - s_t] \tag{1.4}$$

In the model y_t^d denotes output demand, d_t is a demand shock, s_t is the nominal exchange rate, p_t is the aggregate price level, i_t is the nominal interest rate, p_t^e is equilibrium flexible price level, m_t^s is money supply. Equation (1.1) is an open economy IS equation, equation (1.2) captures sluggish adjustment of the price level to its flexible price equilibrium, where θ is the speed of adjustment, equation (1.3) is an LM equation and (1.4) is an interest rate parity condition. Remember that in the national income accounting identity, GDP is the sum of domestic demand and net export of goods and nonfactor services. It is therefore sufficient to specify the determinants of the trade balance since, given total output, this accounting identity then pins down total domestic demand. The two main determinants of the trade balance are assumed to be relative output and the real exchange rate. The equation for the home country's trade balance then reads:

$$tb_t = \xi q_t - \beta y_t \tag{1.5}$$

where ξ and β denote elasticities of the trade balance with respect to real exchange rate q_t and relative output y_t. This specification implies that, if business cycles were perfectly synchronized between the home and foreign countries, the composition of domestic output would depend solely on the level of the real exchange rate.

The stochastic processes that drive the relative supply of output, the relative real demand shock, and relative money supply are as follows (the associated shocks are supply, demand, and nominal shock). The first and the last stochastic processes are assumed to be random walks while the demand shock is allowed to have a permanent as well as a transitory component:

$$y_t^s = y_{t-1}^s + z_t$$

$$d_t = d_{t-1} - \gamma\delta_{t-1} + \delta_t$$

$$m_t = m_{t-1} + v_t$$

The innovations z_t, δ_t, v_t are assumed to be serially and mutually uncorrelated. The flexible-price rational expectations solution to the model is:

$$y_t^e = y_t^s \tag{1.6}$$

$$q_t^e = \frac{y_t^s - d_t}{\eta} + \frac{\sigma\gamma\delta_t}{\eta(\eta+\sigma)} \tag{1.7}$$

where y_t^e and q_t^e are the equilibrium output and real exchange rate. Trade balance then could be rewritten as:

$$tb_t^e = y_t^s\left(\frac{\xi}{\eta} - \beta\right) + \frac{\xi}{\eta}\left(\frac{\sigma\gamma\delta_t}{\eta+\sigma} - d_t\right) \tag{1.8}$$

Equations (1.6) - (1.8) represent the rational expectations solution to the model. These equations imply that the level of output is not affected by either nominal or demand shocks and nominal shock does not influence the level of the real exchange rate. These are the three restrictions that were used to identify the econometric model. The restriction on the level of output might be controversial, as it may well be that the demand shock shifts output to a new level. Recalling the fact that the data in Prasad (1999) was in first differences, the theoretical restrictions translate into the so-called long run restrictions as in Blanchard and Quah (1989). One could also see that in the model the level of trade balance is not influenced by nominal shocks. The last implication may be questionable, therefore it is left unused for the identification purposes.

The model indicates that the effects of supply shocks and nominal shocks on the trade balance in the short run are ambiguous and depend on the elasticities of the trade balance w.r.t. relative output and the real exchange rate. On the other hand, permanent demand shocks, which result in an appreciation of the real exchange rate simultaneously with a transitory increase in relative output, produce an unambiguous negative trade balance response.

Due to permanent (random walk) components of the shocks, the model implies that relative output, real exchange rate and trade balance are non-stationary in levels but stationary in first differences. Since the dynamics are determined by different shocks, the variables are not cointegrated. These implications will be further tested against the data.

The econometric methodology used to estimate the data is as follows. Consider a standard K dimensional reduced form VAR(p) model as in Lütkepohl (2005):

$$y_t = v + A_1 y_{t-1} + ... + A_p y_{t-p} + u_t \tag{1.9}$$

were $y_t = (y_{1t}, ..., y_{Kt})'$ is a $K \times 1$ dimensional vector of observations, v is the intercept term, A_i are $K \times K$ coefficient matrices for $i = 1, ..., p$ and u_t is a $K \times 1$ dimensional error term that has mean zero and positive definite covariance matrix $u_t \sim (0, \Sigma_u)$.

In the standard SVAR approach based on (1.9) a transformation of the reduced form residuals u_t is used to obtain the structural shocks denoted by ϵ_t. A transformation matrix B is chosen such that $\epsilon_t = B^{-1}u_t \sim (0, I_K)$, i.e. structural shocks are assumed to be

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

orthogonal with variances normalized to unity. Therefore $\Sigma_u = BB'$. To obtain identified, unique structural shocks, one should impose restrictions on the B matrix. Zero restrictions are usually used. A zero restriction on B implies that a certain shock does not have an instantaneous effect on one of the variables. The long-run restrictions could be imposed using methodology discussed by Blanchard and Quah (1989). They consider the accumulated effect of shocks to the system, which is summarized by the total impact matrix:

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

and they identify the structural innovations by placing zero restrictions on this matrix. For Ξ_∞ to exist, it must be that the $[I_K - A_1 - ... - A_p]$ matrix is invertible, that is, the VAR process is stationary. Prasad (1999) uses this approach for the 3-dimensional ($K = 3$) system. Given the dimensionality one needs to impose $K(K-1)/2 = 3$ restrictions on the total impact matrix. The restrictions on the long-run matrix can be transformed into restrictions on the B matrix in the way shown by Lütkepohl (2005). The restrictions are justified by the theoretical model discussed above, i.e. nominal and demand shocks have no effects on the level of output and that nominal shocks do not have a permanent effect on the level of the real exchange rate. Ordering the variables as $(RelativeOutput_t, RealExchangeRate_t, TB_t)'$ and associating supply, demand and nominal shocks with structural innovations $(\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t})'$ respectively, the total impact matrix has three zero restrictions: $\Xi_\infty = \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$. The unrestricted elements are denoted by asterisks. Notice, that these zero restrictions are just identifying and, hence, cannot be tested with statistical methods.

1.3 MS-VAR model with changing volatility

It is a well established fact that economic time series exhibit changes in their behavior throughout nearly any horizon of observations. These changes could represent such events in the economy as financial crises (Jeanne and Masson, 2000) or policy changes (Hamilton, 1989, 1994). Therefore, dynamics of economic indicators in different time periods have recently been of increased research interest. The Markov switching (MS) models for analyzing changes in the time series were first discussed by Hamilton (1989). Parameters of the considered autoregression were viewed as the outcome of a discrete-state Markov process. A general econometric framework for multiple time series when the data generation process is subject to

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

regime shifts has been constructed by Krolzig (1997) as a generalization and continuation of work by Hamilton (1988, 1989, 1994). A stable vector autoregression is analyzed, conditional on the regime, and the regime generating process is assumed to follow a Markov chain. The regime switching or Markov switching (MS) models were found to be applicable and outperforming traditional models in business cycle, fiscal and monetary analysis and in the analysis of financial data. Thus, regime switching models work well in many situations where multiple time series models have traditionally been used.

The way to proceed in allowing for MS in the multivariate structural VAR model is to consider the MS-VAR with heteroskedasticity, i.e different volatility regimes, a MSH-VAR in Krolzig (1997) notation. The main reasons for doing this in the current study is to account for potential heteroskedasticity and extract additional identifying information from the data. To allow for the regime change, one should think of u_t from (1.9) being dependent on the discrete Markov process s_t with states $1, ..., M$. The probabilistic model could be described as:

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

where p_{ij}s are the transition probabilities from state i to j. In addition to that the conditional distribution of u_t given s_t is assumed to be normal:

$$u_t|s_t \sim N(0, \Sigma_m)$$

The distributional assumption is used for setting up the likelihood function in the estimation procedure. The normality assumption is not crucial for the identification of shocks. The conditional likelihood function is:

$$f(y_t|s_t = m, Y_{t-1}) = (2\pi)^{-K/2} \det(\Sigma_m)^{-1/2} \exp(-0.5u_t' \Sigma_m^{-1} u_t), m = 1...M.$$

where matrix Y_{t-1} represents the history of observations up to time t. The main feature of the model is that Σ_m depends on the individual state. The identification of shocks is based on the assumption that they are orthogonal across states with variances changing across states. Remember that the B matrix is the matrix of instantaneous effects, hence state invariant instantaneous effects are assumed.

The assumption on the effects of structural shocks does not go beyond what is typically assumed in a conventional SVAR analysis. Orthogonality of the shocks across the sample is a standard assumption in the literature (Lütkepohl, 2005; Lanne et al., 2010).

If there are just two states with positive definite covariance matrices Σ_1 and Σ_2, a re-

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

sult of matrix algebra establishes that there exists a $K \times K$ matrix B such that $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2 B'$, where $\Lambda_2 = diag(\lambda_1, ..., \lambda_K)$ is a diagonal matrix with positive diagonal elements (Lanne et al., 2010).

It turns out that the decomposition of the state dependent covariance matrices is crucial for the identification procedure. Lanne et al. (2010) and Herwartz and Lütkepohl (2011) show how the decomposition of covariance matrices helps to test the identification restrictions on the B matrix. Consider the two state system. Decomposition $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2 B'$ gives a total of $K(K+1)$ equations which can be solved uniquely for the K^2 elements of B and the K diagonal elements of Λ_2, if the diagonal elements of Λ_2 are all distinct. Therefore B is uniquely identified (up to sign changes and interchange of columns). Thus restrictions imposed on B become over-identifying and testable. The diagonal elements of the matrix Λ_2 have a nice interpretation and could be thought of as relative variances of the structural shocks in State 2 versus State 1. Sign reversal of B is not a problem. As shown by Lanne et al. (2010) the matrix is still identified if all elements in any of the columns are multiplied by -1. This means to consider negative shocks instead of positive ones (or vice versa).

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

$$\Sigma_1 = BB', \Sigma_i = B\Lambda_i B', i = 2, ..., M$$

with diagonal Λ_i may have problems with uniqueness and existence. B is unique up to sign reversal in its columns if for each pair of equal diagonal elements in, say $\Lambda_2 = diag(\lambda_{21}, ..., \lambda_{2K})$ there is a corresponding pair of distinct diagonal elements in other $\Lambda_i = diag(\lambda_{i1}, ..., \lambda_{iK})$, $i = 3, ..., M$. For instance, if $\lambda_{2k} = \lambda_{2l}$ then it should be that $\lambda_{ik} \neq \lambda_{il}$ for some $i = 3, ..., M$. (Lanne et al., 2010). Assumption that the decomposition exists is restrictive. Therefore if there are more than two states, the invariance of the instantaneous effects of shocks across states could be tested by a likelihood ratio test where under the null $\Sigma_1, ..., \Sigma_M$ have the decomposition discussed above. As shown by Lanne et al. (2010) the corresponding likelihood ratio test has an asymptotic χ^2 distribution with

$$0.5MK(K+1) - K^2 - (M-1)K$$

degrees of freedom. If $M = 2$, the number of degrees of freedom is zero.

Under the stated assumptions maximum likelihood (ML) estimation is the most suitable method of estimation. If normality of the conditional distribution does not hold, the estimates are pseudo ML. The estimation procedure is based on the EM algorithm as in Hamilton (1994) and Krolzig (1997) and non-linear optimization of the log likelihood as in

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Lanne et al. (2010) and Herwartz and Lütkepohl (2011). Understandingly such a procedure is computationally demanding and the results may depend on the starting values. To reduce the possibility that the optimizer ends up in a local optimum the algorithm is repeated for a whole range of starting values and the estimates that correspond to the highest likelihood are chosen.

1.4 Application to International Trade

1.4.1 The data

The data collected for the current study is different from that used in Prasad (1999). The main reason for that is the willingness to broaden the time horizon and hence have more data to analyze rather than simply check and replicate the existing results. Nevertheless, the new datasets are constructed in a way that allows the capture of the same effects as in the benchmark paper. The data, apart from real exchange rates, originates from the IMF International Financial Statistics database. The source for the real exchange rate time series is the Bank for International Settlements (narrow indices). To construct the datasets I use quarterly data on GDP which is seasonally adjusted, exports of goods and services and imports of goods and services. Monthly real exchange rates were averaged to construct quarterly series.

For each G7 country I construct an index of relative demand by taking a trade-weighted average of GDP in the other G7 countries. In order to derive relative output the log of this index is subtracted from the log of domestic output. I use the same weights as in the benchmark paper, they can be seen in Table 1.1. They represent the proportion of the total trade for each country accounted for by the other G7 countries. These weights are further normalized so that for each G7 country the sum of the weights of the other six countries in the sample is 100. The trade balance is measured as the difference between exports of goods

Table 1.1: Trade weights for the G7 countries

Canada	France	Germany	Italy	Japan	U.K.	U.S.
83.8	66.9	56.7	68.9	57.6	64.6	63.5

and services and imports of goods and services. The trade balance is expressed as a ratio of total output in order to control for scale effects which is standard practice in the literature. All variables except for the trade balance are used in logs. The following notation for the variables is used throughout the paper Y - relative output (log), RER - real exchange rate (log), TB - trade balance.

The data was collected for each G7 country from 1964:Q1 through 2009:Q4. The two exceptions with respect to the time horizon are Germany and France. The French dataset starts from 1970:Q1 and the German one from 1978:Q3 due to the availability of import and export data in the International Financial Statistics database.

To see if the data preserves certain features previously reported in the literature I compute correlations of trade balance (EX-IM) and GDP of the countries (see Table 1.2). It is evident that trade balance is positively correlated with GDP for CA, IT, JP and FR. For the other countries I find negative correlations. Findings for all countries apart from GB and DE are consistent with Mendoza (1995). However it should be noted, firstly, that Mendoza (1995) applies a Hodrick-Prescott filter to the data which I don't do; secondly, I get a positive correlation for DE if I consider a time period only until 1990 as the author did.

1.4.2 Model selection and statistical analysis

In this section I deal with the model selection process and discuss such issues as stationarity of the series, lag order, number of states and statistical analysis of the restrictions. I consider first differenced VAR models for all countries as a suitable choice for the following reasons. First, there is strong evidence of unit roots in nearly all series with only some evidence of cointegration for one country - Japan. This will be discussed later. Second, it allows me to focus on a homogeneous sample. And third, this is in line with both the theoretical model described above and the benchmark paper. ADF tests for models with nonzero mean and trend were conducted to check for stationarity of the series. Test statistics are presented in Table 1.3. Presence of the unit root could not be rejected at 1% level for all series and all countries apart from the trade balance of Japan. To check for further evidence in favor of first differencing variables I conducted a series of cointegration tests. Neither Johansen's trace test (Johansen, 1991) nor Saikkonen and Lütkepohl's test for cointegration (Saikkonen and Lütkepohl, 2000) provide evidence for cointegration relations in the data for any of the countries except for Japan. In the latter case, one could not reject the cointegration rank $rk = 1$ at 10% level. These results are a slight problem, as the unit root test results for Japanese trade balances implies that the data exhibits different behavior than the theoretical model. Also existing cointegration relation requires error correction specification for the Japanese data. For those reasons I drop Japan from the subsequent analysis.

Furthermore in order to determine the lag order for each country I use Akaike Information Criterion (AIC) and Hannan-Quinn Criterion (HQ) values for VAR models without

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Markov switching. Based on AIC, lag order 1 is chosen for CA, IT, GB, DE; lag order 2 for US and lag order 4 for FR data. HQ criteria generally supports the selected lag order. It should be said that the benchmark paper uses likelihood ratio tests to determine the lag order. This results in the optimal lags being 8-12. In the current study however, I stick to the lags selected by AIC for two reasons. Firstly, there are few if any significant coefficients left in the subsequent lags and secondly, the optimization algorithm is sensitive to over-parametrization of the model. Equipped with the results I estimate VAR models for each country. Resulting residuals are shown in Figure 1.1. One can clearly see that the volatility of the residuals changes across time for each country. This speaks in favor of the necessity to model heteroskedasticity.

Choosing the number of volatility states is crucial for the analysis. Standard LR tests are not suitable for the purpose due to the fact that the usual regularity conditions are not fulfilled under the null hypothesis and some parameters are unidentified, hence the asymptotic null distribution of the likelihood ratio test statistic is not the χ^2 one. Although some special tests have been proposed in the literature they are computationally very demanding and complicated. An alternative way to choose the number of states is based on the multiple-regime representation of the MS-VAR(m, p) process and determines the state dimension m by means of complexity-penalized likelihood criteria, such as AIC, HQ or Swarz Criterion (SC). The criteria have a general form

$$C(\theta) = -2 \log L_T(\theta) + C_T \times dim(\theta)$$

where C_T is a constant that for AIC is equal to 2 and for SC is $\log T$. Psaradakis and Spagnolo (2003) examined the performance of the three procedures for selecting the number of regimes. The authors conclude that AIC is generally successful in choosing the correct state dimension, whereas HQ and SC have a tendency to underestimate the number of states.

In order to select the number of states and to test for the identification restrictions I estimate two and three state restricted, partially restricted and unrestricted models with state invariant B for each of the countries. I also estimate three state model without assumptions on B and use the results to test for the validity of state invariant instantaneous effects. AIC and SC values for two and three state models with time invariant B and with/without long run (BQ) restrictions are computed for each country and reported in Table 1.4.

Although there is no general agreement in the literature whether it is sufficient to use the information criteria to select number of states, it is instructive to compare models on those grounds as, for example, in Herwartz and Lütkepohl (2011). It is quite clear that AIC favors three state models for all countries with an exception of DE, whereas SC favors three states only for IT.

This should not be a surprise as, in general, AIC chooses models with a high number of estimated parameters. If one looks at the estimated state probabilities for three state models, then, unfortunately, for the majority of the countries one of the states would be associated with relatively few periods. Hence, the estimates could be unreliable due to the small number of observations associated with one of the states. Another argument for the two state models in the current analysis could be based on the standardized residuals. Those are computed as

$$\hat{\Sigma}_{t|t-1}^{-1/2}\hat{u}_t,$$

where $\hat{u}_t = y_t - \hat{\nu} - \hat{A}_1 y_{t-1} - \cdots - \hat{A}_p y_{t-p}$ and

$$\hat{\Sigma}_{t|t-1} = \sum_{m=1}^{M} \widehat{\Pr}(s_t = m|\mathbf{Y}_{t-1})\hat{\Sigma}_m.$$

Here $\hat{\Sigma}_{t|t-1}$ denotes the estimated residual covariance matrix conditional on the information up to time $t-1$, that is, conditional on $\mathbf{Y}_{t-1} = (Y_{t-1}, \ldots, Y_1)$. The residuals for the two state models are shown in Figure 1.2. If we compare those with the VAR residuals, then the standardized ones are much more homogeneous. Therefore the MS structure with two states captures the turbulences of volatility well enough. Consequently, in what follows I continue with the two state models.

The next question to address is whether the estimated B is locally unique. Even though the two state covariance matrix decomposition is unique (up to sign changes and rotation of columns) the estimation precision could be poor due to the features of the data. One should therefore take a closer look at the estimated λ_{ij} parameters. The estimates and standard errors are presented in Table 1.5. Note that for restricted models no ordering of the elements in Λ_2 is imposed, while for the unrestricted model the elements go from smallest to largest.

The estimates of λ_{ij}s are all different, but they may not be significantly different due to high standard errors (especially for CA, DE and US). Some words on the estimation of standard errors should be said. I have used the outer product estimate of the information matrix as in Hamilton (1994). Covariance matrix of the parameters is then the inverse of the information matrix, while standard errors are obtained by taking square roots of the diagonal elements of the covariance matrix.

Remembering that estimated parameters are asymptotically normally distributed, Wald and LR tests will suit for the purpose of testing the equality of the λ_{ij}s. The Wald test uses

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

estimates of the covariance matrices. The number of estimated parameters relative to the sample size and duration of the states is quite large and therefore, the covariance matrix could be poorly estimated. High standard errors constitute the evidence of that. Therefore Wald tests may only by of limited use whereas conducting LR tests could be beneficial. The LR test depends on the highly nonlinear likelihood optimization subject to constraints. Nevertheless, LR tests are preferable in the context. For uniqueness of the B in a two state model it suffices to test whether $\lambda_{2i} = \lambda_{2j}$ for all combinations of i and j.

The corresponding Wald and LR test statistics are shown in Table 1.6. On the one hand, Wald tests clearly fail to reject equality of λ_{ij}s for all countries apart from IT. For GB and FR two pairwise equalities are rejected. The results for the other countries reflect relatively high standard errors. On the other hand, looking at the LR statistics one could find strong evidence in favor of structural decomposition. For CA, IT and DE all null hypotheses of pairwise equality of λ_{ij}s are rejected at most at 7% significance level. This should be interpreted as evidence for the existence of decomposition for these datasets. For GB and FR two and one pairwise equalities could not be rejected respectively.

The rejection of equality comes from the estimation precision and should not be interpreted as evidence against the decomposition. As to US data, pairwise equalities could not be rejected at common significance level, the smallest p value is 0.261. This may be due to the features of the data, the Markov switching model may not be the best approximation of the data generation process for the dataset. The first state is represented only by a few observations (as it becomes clear from Figure 1.7), while the estimated number of parameters is high. Therefore the estimates of the B may be poor in that case.

Based on LR tests there is strong evidence for the identified and unique B for all countries apart from the US. This gives good grounds to look at the LR tests for different restrictions reported in Table 1.7. Considering two state model, the restrictions imposed by Prasad (1999) are not in line with the data for two of the analyzed countries.

The long run restrictions are rejected at 1% level for CA (note that the result is independent of the number of states). As to IT, GB, US and FR the restrictions for the two state case could not be rejected at least at 8% level, thus the data speaks quite strongly in favor of those. For DE restrictions are rejected at a low level for both two and three state models, but the data objects to the restrictions less in the three state model. The data for CA and DE did not support the restrictions and in the data for JP cointegration was present. Thus, the chosen structural model may not be the best one to apply to those countries' data. Looking at the information criteria, SC favors long run restrictions for GB, US and FR, whereas AIC always chooses models with neither assumptions on the B matrix nor long run restrictions.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Therefore SC values are consistent with the LR test results.

In order to analyze the reasons for the rejection of the restrictions, I estimate partially identified models for CA and DE. The estimated models are those with the identified demand shock $\xi_{1,2} = 0$, combination of the restrictions $\xi_{1,2} = 0, \xi_{2,3} = 0$ and $\xi_{1,2} = 0, \xi_{1,3} = 0$ and the identified nominal shock $\xi_{1,3} = 0, \xi_{2,3} = 0$. As becomes clear, the data does not object to the identification of the demand shock, but the identification of the nominal shock is problematic. For CA the model with maximum two restrictions $\xi_{1,2} = 0, \xi_{2,3} = 0$ is supported, while for DE only $\xi_{1,2} = 0$ could not be rejected at a common significance level. This means that the relations in the equations (1.6) and (1.7) may not be valid for all countries. In other words, the demand shocks are identified for CA and DE , but the other two may not have clear economic interpretation. This issue will be reconsidered in the impulse response analysis section.

Summing up, independently of the number of states the long run restrictions were rejected for CA and DE. However the model where only demand shocks are identified for those countries was supported by the data. On the other hand, sufficient support for the long run restrictions in the two state case was found for IT, GB, US and FR.

1.4.3 Volatility Structure and Impulse Response Analysis

Some joint features of the volatility structure could be seen among the countries. Estimated covariance matrices for the datasets are reported in Table 1.8, smoothed state probabilities are presented in Figures 1.3 - 1.8. Trade balance of the European countries and CA has the lowest variance in any of the volatility states, followed by RER. Relative output is the most volatile variable for these countries. Low volatility of trade balance could be an indicator of a very similar and interconnected trade structure among the countries. This would not be a surprise for the European Union countries who have been members of a customs union since 1958, but the data from Canada also speaks in favor of trade volatility structure similarity to EU countries.

On the other hand the US have an exactly reversed feature. Trade balance is the variable with the highest variance in the high volatility states. Although relative output of the US is also highly volatile in the state that was previously associated with economic recessions, it is not as volatile as the trade balance. It could be partly due to the fact that the US is the largest economy in the sample and reaction to changed economic and trade conditions is different if compared to relatively 'small' EU countries. The US economy is also known to be highly sensitive to the prices of raw material (mainly crude oil), that form a substantial part in traded goods. Another feature to note for the US is that RER is the variable with the lowest variance across volatility states. On the contrary in the EU + CA group of countries RER has higher variance than trade balance in high volatility states (DE is an exception). Thus the trade structure of US differs from that of the European countries. Estimated state probabilities also show that the reaction of the real economies to the recent economic downturn is different. In all of the countries Y experienced a period of high volatility. In CA and US it was also the RER that fluctuated considerably, while in the EU countries this was not the case. US is a unique economy where TB becomes volatile in periods of downturn.

In the previous subsections it became clear that restrictions for some of the G7 countries are rejected, thus full impulse response (IR) analysis is limited to the countries where the identification scheme is supported by the data. For CA and DE I present the impulse responses for the models with maximum identification supported by the data. For CA it is identified demand shock and $\xi_{2,3} = 0$ while for DE it is only identified demand shock. Equipped with the impulse responses I will try to identify the other two shocks for these countries. I also include US in the current analysis, but one should be cautious with it, as the B matrix may be poorly estimated.

Since the VARs were estimated in first differences, the impulse responses to the structural shocks are accumulated in order to show the level responses. The IRs we estimated in a classical way (Lütkepohl, 2005) using 1000 wild bootstrap replications to construct 68% confidence intervals. Confidence intervals we constructed, conditional on the estimated state probabilities and Λ_2 matrix as discussed in Herwartz and Lütkepohl (2011). Impulse responses are presented in Figures 1.9 - 1.12, where the reaction of TB to the identified supply, demand and nominal shocks is shown.

Over the long run, the trade balance response to a positive supply shock is positive for US. If I consider the reaction to a negative shock for US, the result would be in line with Prasad (1999). The response could be both positive and negative for the other countries, as confidence intervals around the responses are above and below zero, thus it is difficult to conclude what the exact reaction would be.

Demand shock (for instance, a domestic fiscal expansion) has negative effects for IT, US and FR but positive for GB. Unfortunately, with confidence intervals above and below zero for IT, GB and US it is difficult to pin down the exact effect. As to FR, the reaction of TB to positive demand shock is marginally negative. In the benchmark paper, Prasad (1999) finds minor negative impact of the demand shock on TB for FR, hence the results are consistent with previous findings. The shock would tend to increase output and cause an exchange rate appreciation, negatively affecting the trade balance through both the exchange rate and relative output channels.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

A nominal shock has two opposing effects on trade balance. It tends to both increase output and cause an exchange rate depreciation in the short run, thus only impulse responses reveal which of the effects dominate. It must be that nominal shock is transmitted to trade balance through depreciation of RER in all countries, as the shock leads to improvement of the TB. The theoretical model implies that nominal shocks (a relative monetary easing in the home country for instance) has no long run effect on trade balance. Looking at the impulse responses one can see that this is not the case. A nominal shock has a permanent positive effect on trade balance for all countries. This is consistent with the earlier findings by Prasad (1999) and Baldwin (1990) and is an indication of persistent exchange rate effects on trade dynamics. Therefore it would be improper to use this theoretical identifying information, as the data would speak up against it.

The IR for CA and DE are shown in Figures 1.13 and 1.14 respectively. As discussed by Prasad (1999) the reaction of TB to the shocks does not have clear expected signs, therefore labeling the first and third shocks may be difficult. In fact the reaction of the variables to shocks does not uncover the underlying economic mechanism in the current case. The reaction of Canadian TB to the 1st and 3d shocks is inconclusive due to large confidence intervals around the responses. However Y reacts positively to the first shock. For DE the reaction of TB to the first and the last shocks is significantly positive.

1.5 Conclusions

In the present paper I analyze the dynamics of the trade balance of G7 countries in response to macroeconomic shocks in the structural VAR framework. I use recent econometric methodology to exploit the volatility of the residuals of the structural VAR model to allow data to speak for or against identification restrictions implied by the theory. Volatility changes are modeled by a Markov process.

The quarterly data on relative output, real exchange rates and trade balances is obtained and analyzed for each G7 country. Due to the presence of cointegration I had to drop JP from the analysis. Similar data is used in the literature, where identification is achieved by theoretical restrictions. I argue that a theory-based approach may not be relevant for all of the countries. I find theoretical restrictions to be incompatible with the data for two out of seven countries. This result comes from the rejection of the identification scheme of the nominal shocks for Canada and Germany. The restrictions are supported for IT, GB, US and FR data, however estimation precision of the US model may be poor. Estimated Markov-switching VARs allow the association of high volatility periods with well-known economic

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

downturns (first and second oil crises, recessions in the 90s, recent financial crisis). Volatility analysis shows that during downturns trade balance is much more volatile in the US, than in the EU countries and CA. On the contrary real exchange rate fluctuates more in the EU and Canada, than in the US.

Impulse response analysis reveals that there is significant positive effect of nominal shocks on the trade balance for the countries where the identification is supported by the data. That finding is known in the literature but it contradicts the implications of the theoretical model discussed in this paper. The marginal negative effect of demand shock on TB in FR and positive effect of supply shock on TB in US are consistent with the previous findings. It would be improper to draw the conclusions on the other effects due to wide confidence intervals. The partially identified models for CA and DE fail to provide enough information to identify shocks fully.

Bibliography

Baldwin, R. (1990). Hysteresis in trade, *Empirical Economics* **15**: 127–142.

Blanchard, O. J. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances, *American Economic Review* **79**(4): 655–73.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* **57**(2): 357–84.

Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press.

Herwartz, H. and Lütkepohl, H. (2011). Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks, *Working Paper series*, EUI, Florence.

Johansen, S. (1991). Cointegration and hypothesis testing of cointegration vectors in gaussian vector autoregressive models, *Econometrica* **59**(6): 1551–1580.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions*, Springer.

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. and Lütkepohl, H. (2010). Structural vector autoregressions with nonnormal residuals, *Journal of Business & Economic Statistics* **28**(1): 159–168.

Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with markov switching, *Journal of Economic Dynamics and Control* **34**(2): 121–131.

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

Mendoza, E. G. (1995). The terms of trade, the real exchange rate, and economic fluctuations, *International Economic Review* **36**(1): 101–37.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Prasad, E. S. (1999). International trade and the business cycle, *The Economic Journal* **109**(458): 588–606.

Psaradakis, Z. and Spagnolo, N. (2003). On the determination of the number of regimes in markov-switching autoregressive models, *Journal of Time Series Analysis* **24**(2): 237–252.

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.

Saikkonen, P. and Lütkepohl, H. (2000). Testing for the cointegrating rank of a var process with an intercept, *Econometric Theory* **16**(3): pp. 373–406.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 1.2: Correlation of trade balance and GDP

Country	$Corr(TB, GDP)$
CA	0.4000
IT	0.5795
GB	-0.8323
DE	-0.1282
DE (until 1990)	0.7453
US	-0.8928
JP	0.6298
FR	0.3641

Table 1.3: ADF tests for stationarity

Country	Y	RER	TB
CA	-1.3307	-2.6235	-2.3775
IT	-2.9792	-2.2245	-3.0500
GB	-3.7903	-3.0141	-3.3554
DE	-2.4067	-2.7012	-2.1999
US	-2.0802	-2.2950	-2.3742
JP	-1.7913	-2.2361	-4.2054
FR	-2.7476	-2.9239	-1.7864
Critical values are 1% 5% 10% -3.96 -3.41 -3.13			

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 1.4: Comparison of MS models

Country	Model	$\log L_T$	AIC	SC
CA	2 states, unrestricted	1591.223	-3130.445	-3047.719
	2 states, identified demand shock ($\xi_{1,2} = 0$)	1590.862	-3131.725	**-3061.218**
	2 states, identified demand shock and $\xi_{2,3} = 0$	1590.116	-3132.233	-3055.871
	2 states, identified demand shock and $\xi_{1,3} = 0$	1586.811	-3125.622	-3049.259
	2 states, identified nominal shock ($\xi_{1,3} = 0$ and $\xi_{2,3} = 0$)	1584.484	-3120.969	-3053.282
	2 states, long run restrictions	1580.277	-3114.553	-3041.372
	3 states, unrestricted	1616.709	**-3161.418**	-3046.302
	3 states, state invariant B	1598.016	-3130.031	-3025.032
	3 states, long run restrictions, invariant B	1591.510	-3123.021	-3027.566
IT	2 states, unrestricted	1497.101	-2942.201	-2859.475
	2 states, long run restrictions	1495.019	-2944.038	-2870.857
	3 states, unrestricted	1520.835	**-2969.670**	-2854.554
	3 states, state invariant B	1514.757	-2963.516	-2858.515
	3 states, long run restrictions, invariant B	1514.555	-2969.111	**-2873.657**
GB	2 states, unrestricted	1394.782	-2737.565	-2654.837
	2 states, long run restrictions	1389.092	-2732.185	**-2659.003**
	3 states, unrestricted	1411.775	**-2751.551**	-2636.434
	3 states, state invariant B	1408.035	-2750.071	-2645.072
	3 states, long run restrictions, invariant B	1398.052	-2736.104	-2640.650
DE	2 states, unrestricted	1083.318	**-2114.636**	-2041.309
	2 states, identified demand shock ($\xi_{1,2} = 0$)	1082.256	-2114.513	**-2044.006**
	2 states, identified demand shock and $\xi_{2,3} = 0$	1067.811	-2087.621	-2019.934
	2 states, identified demand shock and $\xi_{1,3} = 0$	1064.476	-2080.952	-2013.265
	2 states, identified nominal shock ($\xi_{1,3} = 0$ and $\xi_{2,3} = 0$)	1066.886	-2085.772	-2018.085
	2 states, long run restrictions	1063.913	-2081.826	-2016.873
	3 states, unrestricted	1089.645	-2107.291	-2005.229
	3 states, state invariant B	1087.028	-2108.056	-2014.986
	3 states, long run restrictions, invariant B	1079.469	-2098.913	-2014.330
US	2 states, unrestricted	1678.181	-3286.362	-3175.197
	2 states, long run restrictions	1676.116	-3288.232	**-3186.595**
	3 states, unrestricted	1699.887	**-3309.775**	-3166.135
	3 states, state invariant B	1694.771	-3305.543	-3172.145
	3 states, long run restrictions, invariant B	1687.572	-3297.144	-3173.274
FR	2 states, unrestricted	1456.807	-2807.614	-2646.313
	2 states, long run restrictions	1453.489	-2806.979	**-2654.808**
	3 states, unrestricted	1484.611	**-2843.221**	-2650.513
	3 states, state invariant B	1476.734	-2833.469	-2650.864
	3 states, long run restrictions, invariant B	1469.689	-2825.378	-2651.904

Note: in **bold** the minimum value of the criterion for the particular country are denoted

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 1.5: Estimates of structural parameters, transition probabilities and their standard errors

Parameter	CA, unrestricted		IT, restricted		GB, restricted		DE, unrestricted		US, restricted		FR, restricted	
	estimate	st.error	estimate	st.error	estimate	st.error	estimate	st.error	estimate	st.error	estimate	st.error
λ_{21}	0.432	5.191	0.402	0.092	7.281	1.956	0.033	0.108	0.191	0.302	0.595	0.205
λ_{22}	2.223	23.710	0.053	0.016	5.397	1.589	0.097	0.712	0.113	0.156	0.032	0.010
λ_{23}	4.332	75.060	0.213	0.055	2.549	0.745	45049.106	162511.182	0.076	0.299	0.437	0.153
p_{11}	0.992	0.135	0.901	0.077	0.849	0.061	0.170	0.452	0.529	0.489	0.954	0.085
p_{12}	0.001	0.072	0.051	0.030	0.157	0.095	0.041	0.093	0.021	0.060	0.026	0.038

Note: By construction columns of $\hat{P}$ sum up to 1 and therefore only unrestricted elements are presented (last row of the matrix is calculated for given estimates).

 DOI: 10.2870/73498

Table 1.6: Tests for equality of λ_{ij}-s for the models

Country	H_0	Wald	p-value	LR	p-value
CA	$\lambda_{21} = \lambda_{22}$	0.0093	0.9230	21.238	4.05×10^{-6}
	$\lambda_{22} = \lambda_{23}$	0.0017	0.9673	3.462	0.062
	$\lambda_{21} = \lambda_{23}$	0.0031	0.9555	21.236	4.06×10^{-6}
IT	$\lambda_{21} = \lambda_{22}$	11.859	5.7×10^{-4}	28.568	9.04×10^{-8}
	$\lambda_{22} = \lambda_{23}$	7.524	0.006	14.699	1.26×10^{-4}
	$\lambda_{21} = \lambda_{23}$	2.962	0.085	3.451	0.062
GB	$\lambda_{21} = \lambda_{22}$	0.508	0.475	0.591	0.442
	$\lambda_{22} = \lambda_{23}$	3.164	0.075	1.499	0.221
	$\lambda_{21} = \lambda_{23}$	4.356	0.037	6.635	0.010
DE	$\lambda_{21} = \lambda_{22}$	0.007	0.9292	36.662	1.40×10^{-9}
	$\lambda_{21} = \lambda_{23}$	0.076	0.7816	31.181	2.35×10^{-8}
	$\lambda_{22} = \lambda_{23}$	0.077	0.7816	36.664	1.4×10^{-9}
US	$\lambda_{21} = \lambda_{22}$	0.052	0.8186	0.384	0.535
	$\lambda_{21} = \lambda_{23}$	0.012	0.9109	0.232	0.630
	$\lambda_{22} = \lambda_{23}$	0.067	0.7949	1.261	0.261
FR	$\lambda_{21} = \lambda_{22}$	7.386	0.006	44.272	2.80×10^{-11}
	$\lambda_{22} = \lambda_{23}$	6.906	0.008	30.008	4.30×10^{-8}
	$\lambda_{21} = \lambda_{23}$	0.345	0.556	0.495	0.481

Table 1.7: Likelihood ratio tests for MS models

Country	MS states	H_0	H_1	LR	p-value
CA	2 states	BQ restrictions	state invariant B	21.892	6.8×10^{-5}
	2 states	$\xi_{1,2} = 0$	state invariant B	0.722	0.395
	2 states	$\xi_{1,2} = 0, \xi_{2,3} = 0$	state invariant B	2.214	0.331
	2 states	$\xi_{1,2} = 0, \xi_{1,3} = 0$	state invariant B	8.824	0.012
	2 states	$\xi_{1,3} = 0, \xi_{2,3} = 0$	state invariant B	13.478	0.001
	3 states	state invariant B	unrestricted	37.386	3.8×10^{-8}
	3 states	BQ restrictions	state invariant B	13.012	0.005
IT	2 states	BQ restrictions	state invariant B	4.164	0.244
	3 states	state invariant B	unrestricted	12.156	0.007
	3 states	BQ restrictions	state invariant B	0.404	0.939
GB	2 states	BQ restrictions	state invariant B	11.380	0.098
	3 states	state invariant B	unrestricted	8.254	0.041
	3 states	BQ restrictions	state invariant B	19.966	1.7×10^{-4}
DE	2 states	BQ restrictions	state invariant B	38.811	1.9×10^{-8}
	2 states	$\xi_{1,2} = 0$	state invariant B	2.124	0.145
	2 states	$\xi_{1,2} = 0, \xi_{2,3} = 0$	state invariant B	31.014	1.8×10^{-7}
	2 states	$\xi_{1,2} = 0, \xi_{1,3} = 0$	state invariant B	37.684	6.5×10^{-9}
	2 states	$\xi_{1,3} = 0, \xi_{2,3} = 0$	state invariant B	32.864	7.3×10^{-8}
	3 states	state invariant B	unrestricted	5.234	0.155
	3 states	BQ restrictions	state invariant B	15.118	0.002
US	2 states	BQ restrictions	state invariant B	4.130	0.247
	3 states	state invariant B	unrestricted	10.232	0.017
	3 states	BQ restrictions	state invariant B	14.398	0.002
FR	2 states	BQ restrictions	state invariant B	6.636	0.085
	3 states	state invariant B	unrestricted	15.754	0.003
	3 states	BQ restrictions	state invariant B	14.090	0.002

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 1.8: Estimated covariance matrices for the models

Country	$\Sigma_1 \times 10^{-4}$	$\Sigma_2 \times 10^{-4}$
CA	$\begin{bmatrix} 5.18 & & \\ 2.81 & 2.48 & \\ 0.12 & 0.02 & 0.47 \end{bmatrix}$	$\begin{bmatrix} 18.19 & & \\ 12.98 & 10.59 & \\ 0.95 & 0.01 & 0.8 \end{bmatrix}$
IT	$\begin{bmatrix} 26.05 & & \\ 6.41 & 8.72 & \\ 0.62 & -0.06 & 1.38 \end{bmatrix}$	$\begin{bmatrix} 10.43 & & \\ 2.21 & 0.85 & \\ 0.29 & 0.05 & 0.30 \end{bmatrix}$
GB	$\begin{bmatrix} 3.98 & & \\ 2.08 & 2.70 & \\ 0.18 & 0.11 & 0.38 \end{bmatrix}$	$\begin{bmatrix} 28.58 & & \\ 14.61 & 16.15 & \\ 1.00 & 0.68 & 0.99 \end{bmatrix}$
DE	$\begin{bmatrix} 46.43 & & \\ -5.41 & 0.84 & \\ -10.05 & 0.00 & 8.58 \end{bmatrix}$	$\begin{bmatrix} 17.36 & & \\ 5.50 & 2.36 & \\ 0.89 & 0.28 & 0.47 \end{bmatrix}$
US	$\begin{bmatrix} 25.44 & & \\ -0.16 & 0.67 & \\ 2.97 & 0.87 & 28.96 \end{bmatrix}$	$\begin{bmatrix} 4.82 & & \\ -0.02 & 0.07 & \\ 0.20 & 0.06 & 2.22 \end{bmatrix}$
FR	$\begin{bmatrix} 20.28 & & \\ 7.77 & 5.86 & \\ 0.32 & 0.36 & 0.32 \end{bmatrix}$	$\begin{bmatrix} 11.01 & & \\ 2.97 & 0.91 & \\ 0.04 & -0.01 & 0.12 \end{bmatrix}$

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 1.1: Residuals of the VAR models

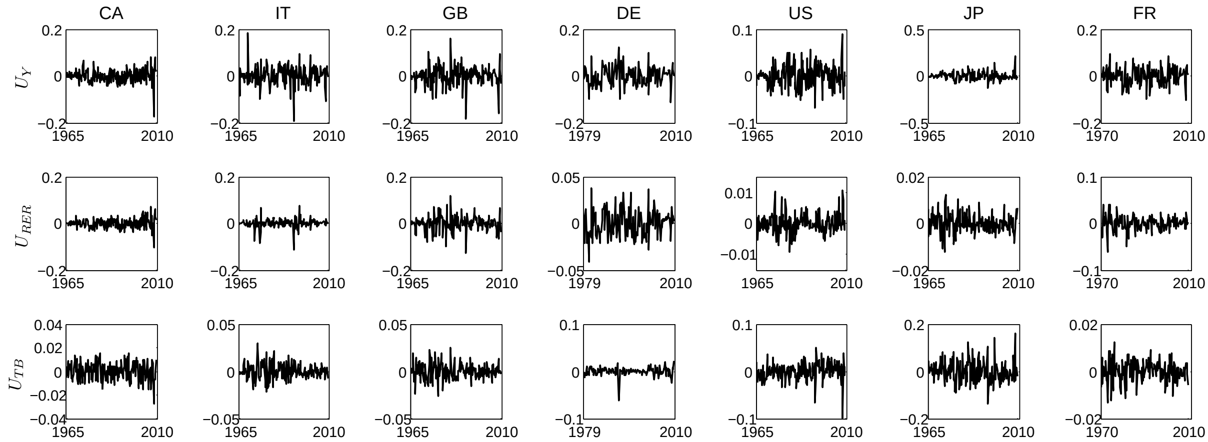
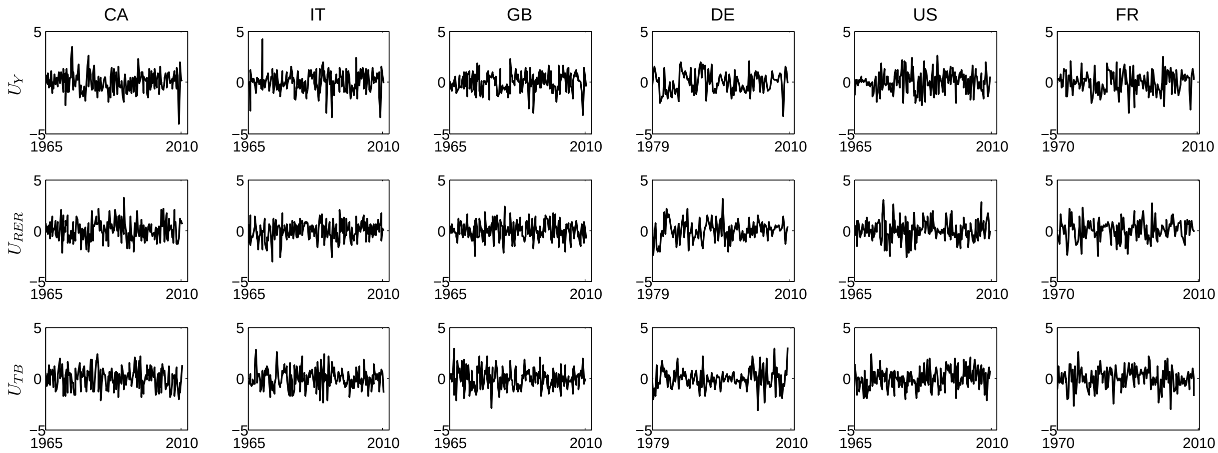


Figure 1.2: Standardized residuals of the 2 state MS-VAR models with state invariant B



Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 1.3: Smoothed state probabilities for CA

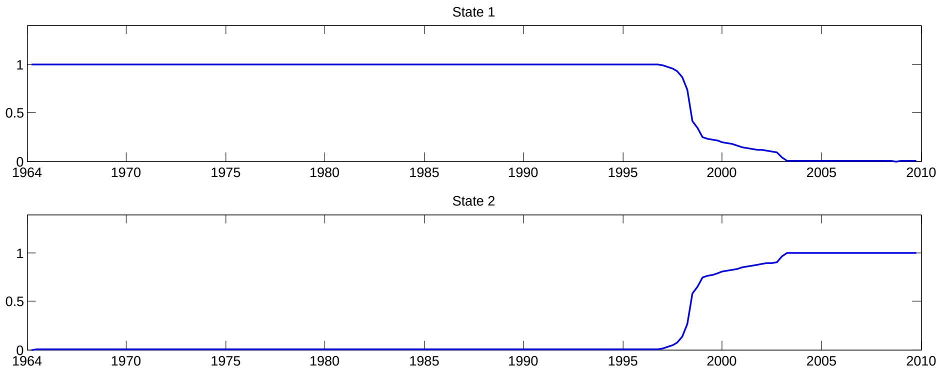


Figure 1.4: Smoothed state probabilities for IT

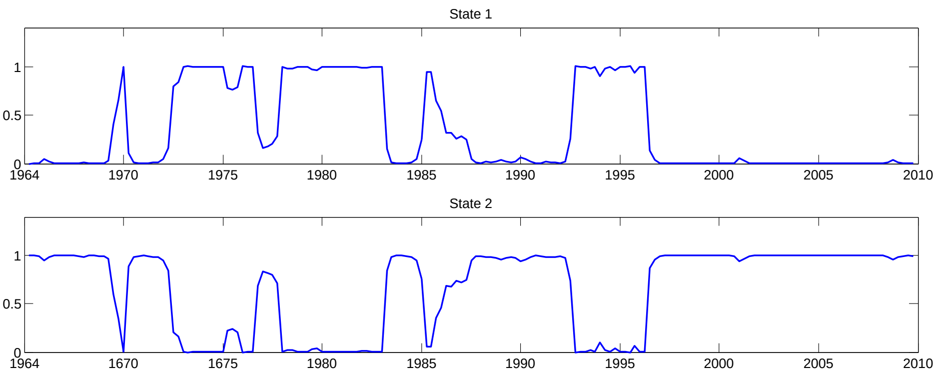
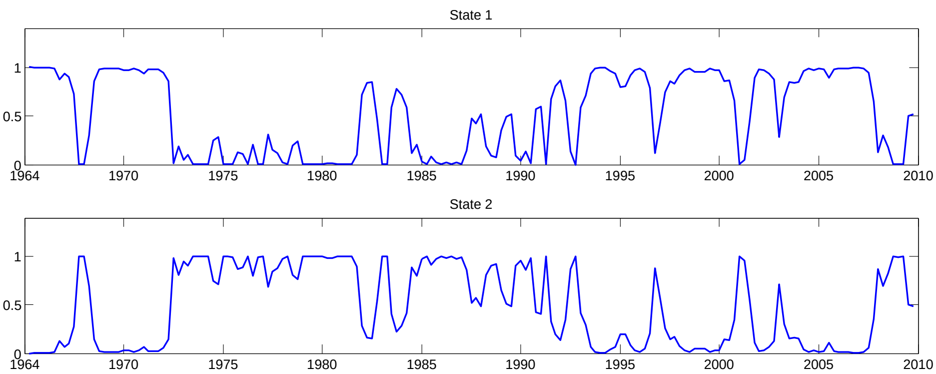


Figure 1.5: Smoothed state probabilities for GB



Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 1.6: Smoothed state probabilities for DE

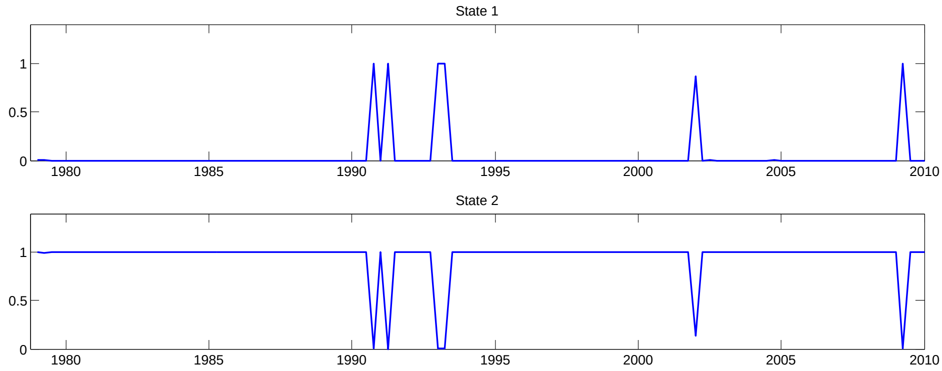


Figure 1.7: Smoothed state probabilities for US

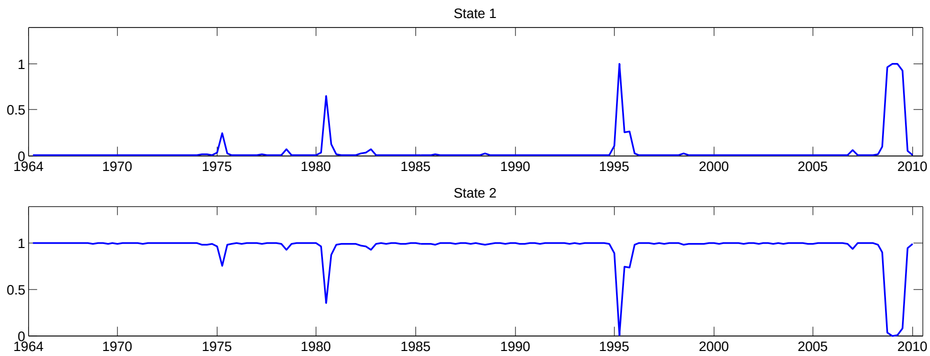
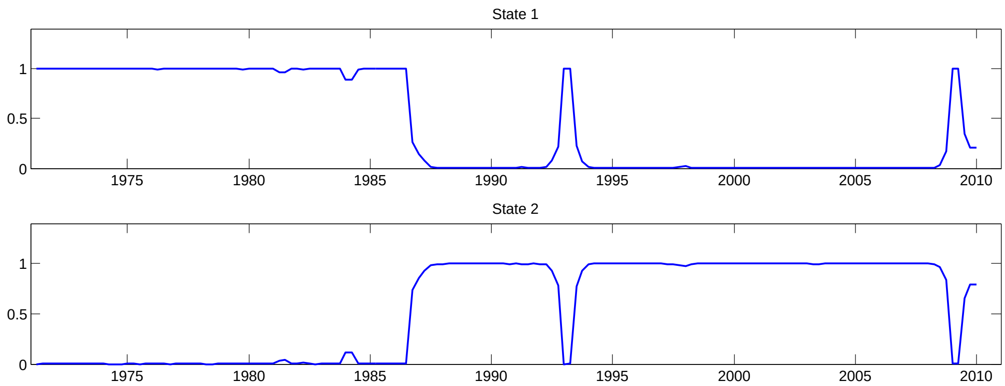


Figure 1.8: Smoothed state probabilities for FR



Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 1.9: Impulse responses of TB to shocks: IT

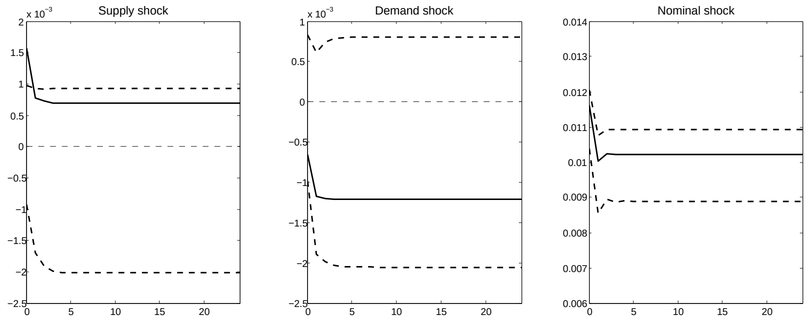


Figure 1.10: Impulse responses of TB to shocks: GB

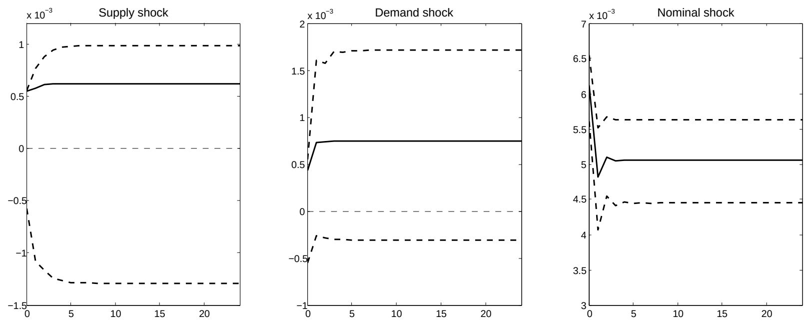
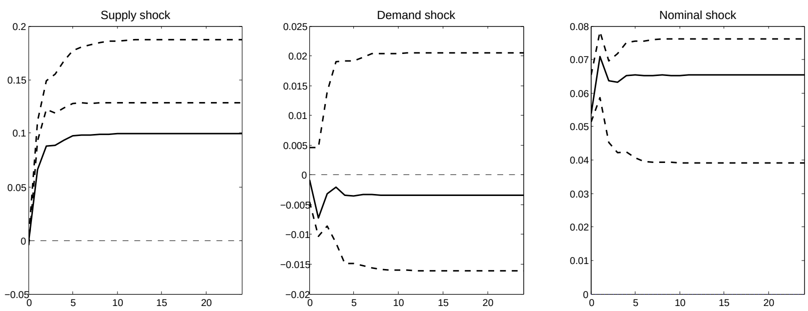


Figure 1.11: Impulse responses of TB to shocks: US



Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 1.12: Impulse responses of TB to shocks: FR

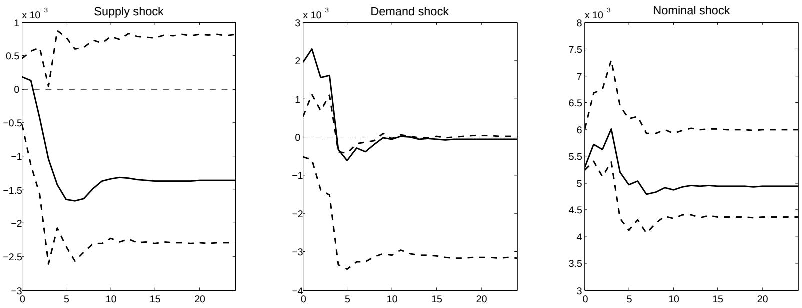


Figure 1.13: Impulse responses to shocks: CA

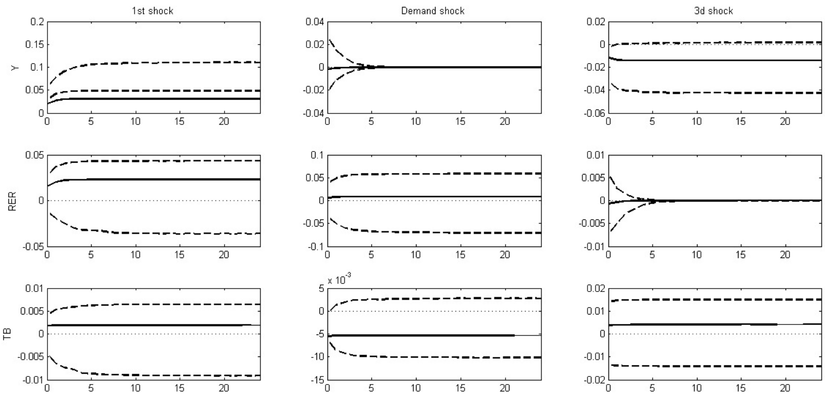
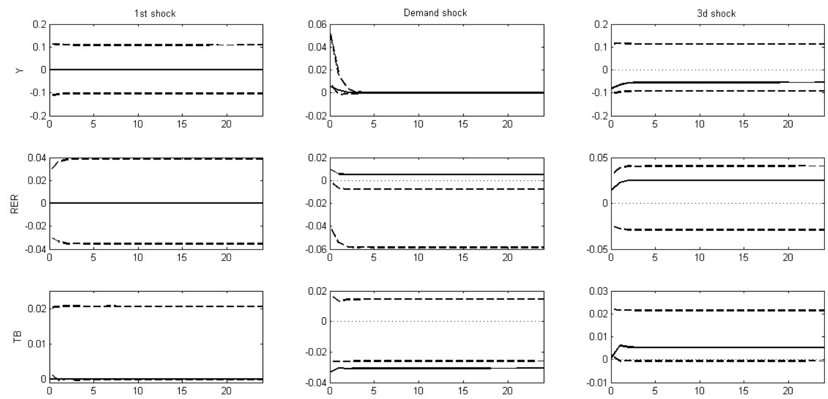


Figure 1.14: Impulse responses to shocks: DE



Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Chapter 2

Disentangling Demand and Supply Shocks in the Crude Oil Market: How to Check Sign Restrictions in Structural VARs

Helmut Lütkepohl[1], **Aleksei Netšunajev**

2.1 Introduction

Identifying the shocks of interest is a major problem in structural vector autoregressive (SVAR) analysis. Often just-identifying restrictions are imposed, for example on the instantaneous effects of the shocks as in Sims (1980) and Amisano and Giannini (1997) or on the long-run effects as in Blanchard and Quah (1989) and King, Plosser, Stock and Watson (1991) (see Lütkepohl, 2005, for a textbook exposition). Typically the restrictions imposed are thereby just-identifying and, hence, cannot be checked against the data by statistical tests. There is some dissatisfaction with these types of equality restrictions because there is often no agreement on them. Reasons may be that different underlying economic models imply alternative sets of restrictions or that they may not imply sufficiently precise restrictions to uniquely identify all shocks of interest.

[1]We thank Lutz Kilian, Helmut Herwartz and participants of the EC^2 Conference in Florence in December 2011 as well as seminars at the University of Venice, the University of Zurich and the Deutsche Bundesbank for comments on an earlier version of this paper. The research for this paper was carried out in part while the first author was a Bundesbank Professor.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Therefore, identification by inequality or sign restrictions for the effects of shocks on certain variables is proposed by a number of authors (Faust, 1998; Canova and De Nicoló, 2002; Uhlig, 2005). In the latter article shocks are identified by specifying the sign of their effect on certain variables on impact and possibly in some of the following periods. All shocks that are in line with the sign restrictions are considered admissible. Since it appears to be easier to agree on such softer restrictions, this approach is seen as an attractive alternative to identification via more classical exclusion restrictions and consequently it has been applied by a number of researchers (Mountford and Uhlig, 2009; Lippi and Nobili, 2011; Peersman and Straub, 2009; Canova and Pappa, 2007).

In some literature, dynamic stochastic general equilibrium (DSGE) models are used to check the signs of responses to shocks of interest and the corresponding sign restrictions are used to characterize the shocks. These then imposed on a SVAR model for the identification of the shocks (Canova, 2002; Dedola and Neri, 2007; Pappa, 2009). DSGE models are highly stylized models of an economy and as such not necessarily a good description of the data. Therefore it makes sense to assume that they cannot give a precise picture of the reactions to shocks in actual economic systems but may be able to suggest at least a direction of the responses of the variables to shocks. Using these directions or signs of the responses in a VAR model which is specified so as to fit the data well, it is hoped to provide a better understanding of the effects of shocks in practice.

Unfortunately, using sign restrictions for identifying structural shocks has some drawbacks as well. First of all, they typically do not identify the shocks uniquely but allow for a range of admissible shocks that are all in line with the sign restrictions. Therefore the range of possible responses may also be large and hence, sign restrictions may deliver only a rather diffuse picture of the reactions of variables to a shock of interest. Secondly, identification via sign restrictions does not provide a possibility to check the validity of the restrictions because only those shocks that satisfy the prespecified sign restrictions are considered. Thus, any analysis based upon them is conditional on the sign restrictions being correct. In other words, the data have no possibility to speak up against the restrictions. This criticism has been raised against conventional just-identifying restrictions as well, of course, but it is no less valid in the context of sign restrictions. One may argue that the problem is less severe in the context of sign restrictions because they are much weaker restrictions than equality restrictions. Even if the sign restrictions are derived from generally accepted economic models, there is a potential gap between the empirical and the theoretical models that may invalidate the sign restrictions in the empirical model. Reasons may be that the variables used in the empirical model do not correspond exactly to those considered in the theoretical model,

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

for instance, due to measurement errors, trend and/or seasonal adjustment or using data with an observation frequency that does not correspond to that of the theoretical model. In addition, the economic model describes the relations within a set of variables and may lack other effects that may, however, be important in practice. For example, the theoretical model may be one for a closed economy whereas foreign effects may not be negligible in the actual system and the available data. In other words, there may be an omitted variables problem that could affect the empirical model.

In this study a proposal is made as to how to let the data speak about the validity of identifying restrictions in general and sign restrictions in particular. The idea is to use changes in volatility to support the identification of shocks. In SVAR analysis identification via heteroskedasticity is proposed and used by Rigobon (2003), Rigobon and Sack (2003), and Lanne and Lütkepohl (2008) among others. These authors essentially assume that there are exogenously generated changes in the volatility of the shocks and partition the sample period accordingly. They then base the identification of the shocks on the assumption that the effects of shocks are the same regardless of the volatility regime in which they occur. In other words, they assume that the impulse responses are invariant throughout the sample period whereas the volatility of the shocks may change. Additional, statistical identifying restrictions thereby become available, which can be used to check restrictions that are just-identifying in the conventional approach. We will use the approach proposed by Lanne, Lütkepohl and Maciejowska (2010) and model the changes in volatility endogenously by means of a Markov regime switching (MS) mechanism. A related proposal, based on different models for conditional heteroskedasticity, is used by Normandin and Phaneuf (2004), Bouakez and Normandin (2010) and others for identifying shocks.

In the present study it will be shown how changes in volatility can provide additional identifying information that can be used for checking the validity of sign restrictions. Of course, this approach requires that there are changes in volatility during the sample period and cannot be used if volatility is time-invariant. If volatility varies sufficiently this device can generate a full set of unique shocks that are in line with the data. These shocks may not have an economic interpretation. If none satisfy the sign restrictions it means that there is no shock that is acceptable to the data and satisfies the sign restrictions at the same time. For example, if there are sign restrictions characterising a shock as a technology shock and none of the shocks acceptable to the data satisfies those restrictions, one may conclude that a technology shock cannot be isolated in the system under consideration. Using the shocks identified by changes in volatility may result in unique shocks and not in a range of feasible shocks and, hence, it may also provide more precise impulse responses.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

In any case, the additional identification information obtained from heteroskedasticity or conditional heteroskedasticity can be helpful for checking the validity of the sign restrictions in the empirical model. Consequently, the two drawbacks of sign restrictions mentioned earlier, that is, lack of precise impulse responses and the inability to reject the restrictions, can be addressed if there are changes in volatility.

One may argue that the typical approach for sign restricted SVARs uses Bayesian methods for estimation that are based on independently, identically distributed (i.i.d.) residuals. Hence, changes in volatility are excluded and our model framework does not nest the typical model assumed under sign restrictions. Although such arguments are valid, it may be worth noting that sign restrictions can be used in conjunction with classical estimation methods (see Moon and Schorfheide (2009) and Moon, Schorfheide, Granziera and Lee (2009) for a discussion of the frequentist approach to inference in such models). In that framework they can be justified under assumptions more general than i.i.d. for the errors. Moreover, even though Bayesian methods require stringent assumptions regarding the error distribution of the VAR model, in applied work they will typically be approximately satisfied at best. Our approach also provides a statistical check of some of these assumptions.

In this study we use a classical approach for inference. In other words, we use likelihood based methods to avoid any distortions by imposing priors on the parameters, although Bayesian methods are more common in conjunction with sign restrictions. However, our main arguments do not depend on the method for estimation. They could be equally well placed in a Bayesian framework.

To illustrate the approach, a study by Kilian and Murphy (2011) will be reconsidered. These authors analyze a system with the following three variables:

- $\Delta prod_t$ - percent change in global crude oil production,
- q_t - log detrended index of real economic activity,
- p_t - log of real price of oil.

They specify an oil supply shock, an aggregate demand shock and an oil-market specific demand shock purely by sign restrictions. They point out that sign restrictions are not enough to get a precise picture of the effects of such shocks and they propose to add further information to pin down the effects more precisely. The example they use is particularly suitable for our purposes because changes in volatility of oil production and the price of oil are diagnosed in the related literature (see Baumeister and Peersman (2010)). In our framework, these volatility changes can be used to address the question whether the system

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

is suitable for isolating the three shocks of interest and also to obtain more information which can help to narrow down the effects of the shocks.

The study is organised as follows. The model setup and some technical details of our inference procedures are given in Section 2.2. The empirical study illustrating the method for checking the sign restrictions is presented in Section 2.3 and conclusions are provided in Section 2.4.

2.2 The Model Setup and Inference

The reduced form of our model is a K-dimensional VAR(p),

$$Y_t = \nu + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + U_t, \tag{2.1}$$

where ν is a constant term, the A_js ($j = 1, \ldots, p$) are ($K \times K$) coefficient matrices and U_t is a zero-mean error term.

In a conventional SVAR model the structural shocks are obtained from the reduced form residuals by a linear transformation, $\varepsilon_t = B^{-1}U_t$ or $B\varepsilon_t = U_t$, where B is such that ε_t has identity covariance matrix, that is, $\varepsilon_t \sim (0, I_K)$, and the reduced form residual covariance matrix is decomposed as $E(U_t U_t') = \Sigma_U = BB'$.

Following Lanne et al. (2010), in our setup the distribution of the reduced form error term U_t is assumed to depend on a discrete Markov process s_t ($t = 0, \pm1, \pm2, \ldots$) 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 distribution of U_t given s_t is assumed to be normal,

$$U_t | s_t \sim \mathcal{N}(0, \Sigma_{s_t}). \tag{2.2}$$

Our model allows for Markov switching in the residual covariances only and not in other parameters of the model. Although there are finitely many states and in practice the number of different states will be small, the model can capture smooth transitions from one state to another. This is because a particular state does not necessarily come up with probability one but the system may be in-between states, that is, in a given period the actual volatility may be described by a mixture of different states, each state being weighted by a certain probability. Thus, in this sense our model is similar to a multivariate GARCH or stochastic

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

volatility model and can capture similar changes in volatility. We do not allow other parameters than the residual covariance to be state-dependent. We thereby impose more regularity on our models than in the MS-SVAR models considered by Rubio-Ramirez, Waggoner and Zha (2005), Sims and Zha (2006) and Sims, Waggoner and Zha (2008) or in SVAR models with time-varying coefficients (Cogley and Sargent, 2005; Primiceri, 2005; Baumeister and Peersman, 2010). In the latter paper, a system similar to our example system analyzed in Section 2.3 is considered in the framework of a time-varying coefficient SVAR model. We will return to their results in the empirical section.

In the following sections we will use the abbreviation MS(M)-VAR(p) model for a VAR model with p lags and M Markov states. We emphasise that we use this notation for simplicity although all VAR coefficients apart from the residual covariances are state-invariant. This fact is not reflected in our notation. For example, Krolzig (1997) uses the notation MSH-VAR for our model type.

The fact that the covariances Σ_{s_t} can vary across states is used in our framework for identifying structural shocks that are consistent with the statistical properties of the data. For example, if there are just two states ($M = 2$), then there exists a decomposition $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2 B'$, where $\Lambda_2 = \text{diag}(\lambda_{21}, \dots, \lambda_{2K})$ is a diagonal matrix with positive diagonal elements. If the λ_{2i}s are all distinct, this decomposition is unique apart from changes in the sign and permutations of the columns of B and corresponding changes in the ordering of the λ_{2i}s (Lanne et al. (2010)). Thus, for a given ordering of the λ_{2k}, if the structural shocks are orthogonal, have the same instantaneous effects across states and are normalized such that they have unit conditional variance in the first state, then they are uniquely determined by the transformation $\varepsilon_t = B^{-1}U_t$. Therefore, any additional restrictions from economics become over-identifying and can be checked against the data. For instance, exclusion restrictions can be checked with formal statistical tests. Moreover, if a set of sign restrictions is postulated for the system to be suitable for the desired analysis, the impulse responses corresponding to the shocks identified via the MS structure must satisfy the prespecified sign restrictions that characterize the economic shocks of interest. Alternatively, if the shocks do not satisfy the sign restrictions, these restrictions are not compatible with the statistical properties of the data. The reasons could be those mentioned in the introduction, that is, omitted variables, measurement errors, aggregation problems or distortions due to data transformations. If, however, the impulse responses satisfy the sign restrictions, labels may be attached to the shocks accordingly.

It may be worth emphasising that the requirement of having distinct λ_{2i}s is crucial for exact identification of all shocks. The λ_{2i}s can be interpreted as variances in State 2 relative

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

to those in State 1. Thus, distinct λ_{2i}s imply that the volatility changes are not homogenous across all variables. Moreover, one of the λ_{2i}s may, of course, be 1, that is, one of the shocks may have the same variance in both states, as long as there is enough heterogeneity in the volatility of the other shocks. Furthermore, even if some of the λ_{2i}s are identical, it may still be possible to identify those shocks with distinct relative variances in State 2. Of course, in that situation it may not be possible to arrive at clear conclusions regarding the validity of the sign restrictions. An important advantage of our approach is that the crucial identification restrictions (distinct λ_{2i}s) can be checked with statistical tests rather than having to assume it.

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

$$\Sigma_1 = BB', \quad \Sigma_i = B\Lambda_i B', \quad i = 2, \ldots, M, \tag{2.3}$$

with diagonal Λ_i matrices is restrictive. Lanne et al. (2010) discuss a likelihood ratio (LR) test for these restrictions. Denoting the diagonal elements of Λ_j by $\lambda_{j1}, \ldots, \lambda_{jK}$, uniqueness of B up to sign and permutation of the shocks is ensured for 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). Although this condition for exact (local) identification is apparently more complicated than in the two state case, it is also potentially more likely to be satisfied if there are more than two states and, hence, more heterogeneity in the volatility. Also in this case all shocks may be identified even if one of them has homogenous volatility across all states. Again, the identification condition can be checked by statistical tests. If it is satisfied, the resulting shocks are unique and need to satisfy any valid sign restrictions. In turn, any restrictions that the shocks do not satisfy, are not compatible with the data.

Our identification techniques are the same as those used in the literature on identification via heteroskedasticity (see, in particular, Lanne and Lütkepohl (2008)). While in that literature heteroskedasticity is used, we consider conditional heteroskedasticity. In our approach changes in volatility are determined endogenously from the data. In principle, the periods associated with a particular volatility state can be spread irregularly throughout the sample. Thus, the volatility states are more flexible than in the standard approach based on heteroskedasticity.

Since we assume normality of the residuals conditional on the states, the likelihood function can be set up and the model can be estimated by maximum likelihood (ML). The likelihood function is given in Lanne et al. (2010) and a detailed discussion of the related estimation problems can be found in Herwartz and Lütkepohl (2011). They also present the details of an EM algorithm for optimizing the likelihood and point out that there are many local maxima. In fact, the numerical problems are challenging, in particular when a large number of different Markov states are allowed for. The procedure is a quasi ML procedure if the normality assumption in (2.2) does not hold. The normality assumption is not essential for the asymptotic properties of the estimates but is used for setting up the likelihood function.

Herwartz and Lütkepohl (2011) discuss a fixed design wild bootstrap procedure for constructing confidence intervals for impulse responses in the presently considered model class. They propose to construct bootstrap samples conditional on the ML estimates so that for our model setup they are constructed as

$$Y_t^* = \hat{\nu} + \hat{A}_1 Y_{t-1} + \cdots + \hat{A}_p Y_{t-p} + U_t^*, \tag{2.4}$$

where $U_t^* = \eta_t \hat{U}_t$ and η_t is a random variable with values 1 and -1, each with probability 0.5. Thereby potential heteroskedasticity and the pattern of contemporaneous dependence of the data is preserved. We bootstrap parameter estimates θ^* of $\theta = \text{vec}[\nu, A_1, \ldots, A_p]$ and B^* of B, conditionally on the initially estimated transition probabilities and Λ_m, $m = 2, \ldots, M$, to alleviate the computational burden as in Herwartz and Lütkepohl (2011). Notice that computing the bootstrap impulse responses still requires nonlinear optimization of the log-likelihood and, hence, is computationally demanding. We use the ML estimates as starting values in the bootstrap replications. In our empirical analysis we consider 68% standard percentile confidence intervals based on 1000 replications. The confidence level is in line with much of the related literature.

2.3 Empirical Analysis

2.3.1 Previous identification restrictions

Kilian and Murphy (2011) are primarily interested in the effects of demand and supply shocks in the crude oil market on the real price of oil. They argue that the exclusion restrictions used by Kilian (2009) for identifying the shocks could be questioned. In the latter article Kilian identifies the shocks by assuming that oil-market specific demand shocks ($\varepsilon_t^{\text{oil-d}}$) do not have an instantaneous effect on oil production and real activity and aggregate demand shocks ($\varepsilon_t^{\text{aggr-d}}$) have no immediate impact on oil production. In other words, B is lower-triangular

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

such that

$$\begin{bmatrix} U_t^{\Delta prod} \\ U_t^q \\ U_t^p \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{\text{oil-s}} \\ \varepsilon_t^{\text{aggr-d}} \\ \varepsilon_t^{\text{oil-d}} \end{bmatrix}. \tag{2.5}$$

Kilian and Murphy (2011) point out that a zero short-run oil supply elasticity of aggregate demand may not be realistic and the same is true for the short-run effect of an oil-market specific demand shock on real activity. Therefore Kilian and Murphy (2011) propose the use of sign restrictions that do not fix these effects at zero. More precisely, they impose the sign restrictions given in Table 2.1 for the impact effects. They also point out that these restrictions do not identify the shocks uniquely but result in a rather large set of admissible shocks and correspondingly imprecise impulse responses. They narrow down the range of admissible shocks by imposing bounds on the oil supply elasticity. In our approach we can get additional information from the data, as demonstrated in the following pages.

Concerns regarding the exclusion restrictions imposed by Kilian (2009) are also expressed by Baumeister and Peersman (2010) who consider a similar system of variables. They use sign restrictions in the framework of a time-varying coefficient VAR model that allows for time-varying coefficients in addition to heteroskedasticity in the residuals. They find that during our sample period there was a rise in oil price volatility while at the same time the volatility of oil production declined. Hence, it seems plausible to use our approach and utilize the changes in the volatility of the shocks for identification.

2.3.2 The data and the VAR model

We use monthly data from 1973m2 - 2006m12, as in Kilian (2009). An updated dataset is used by Kilian and Murphy (2011). These authors use a VAR(24) model. However, model selection criteria favour smaller VAR orders. We base our analysis on a VAR(3) model that is selected by the more generous AIC. Notice that we have to use a nonlinear optimization algorithm for maximizing the likelihood function. Such algorithms run into problems for heavily parameterized models. Admittedly, standard residual tests indicate some remaining autocorrelation in the residuals of the VAR(3) model. The same is true for the VAR(24) model, however, and we have checked that qualitatively similar impulse responses result from both models in a conventional analysis.

In Figure 2.1 the residuals of both the VAR(24) and the VAR(3) models are plotted. While the two sets of residuals naturally look a bit differently, they both show that the residuals of oil production are much more volatile in the first part of the sample while the

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

oil price residuals are substantially more volatile in the later sample period from the middle of the 1980s onwards. The figure thus confirms the observation made by Baumeister and Peersman (2010).

2.3.3 The number of MS states

In Table 2.2 some statistics are presented for a range of different models. In particular, models with different numbers of states and restrictions on the impact effects matrix B are compared. Considering only unrestricted VAR(3) models, the MS(3)-VAR(3) is favoured by SC while AIC prefers a 4-state model. It is worth pointing out that models with MS are much preferred to the VAR(3) model without allowance for changes in volatility. Of course, this result is not surprising given the residual graphs in Figure 2.1. For the MS(3) models the SC and AIC values are further reduced by imposing the restriction of a state-invariant impact effects matrix B and lower-triangularity of this matrix. The SC value also declines when these restrictions are imposed and only MS(4) models are compared while the same is not true for AIC which is minimized for the least restricted model. Thus, based on the model selection criteria models with state invariant and recursive impact effects have some support although they are generally not favored.

In deciding on the number of MS states, it may be worth looking at the smoothed state probabilities for the sample period. They are shown in Figure 2.2 where a state-invariant B matrix is imposed for the MS(3) and MS(4) models. The corresponding state covariance matrices are given in Table 2.3. Looking at the state probabilities of the MS(2)-VAR(3) model first, it can be seen that the first part of the sample is mainly associated with State 2 while State 1 dominates the second part of the sample. In Table 2.3 it can be seen that the volatility of oil production is relatively high in State 2 whereas the volatility of the oil price is much higher in State 1 than in State 2. This "puzzle" is further investigated by Baumeister and Peersman (2010), as mentioned earlier. For the purposes of our study the important point is that volatility changes occur during the sample period.

Looking at the smoothed probabilities of the MS(3) model in Figure 2.2 it becomes clear that the first two states are similar to those of the MS(2) model while the third state apparently captures some special periods. From Table 2.3 it becomes clear that the third state is characterized by very high volatility both in oil production and the real price of oil. Thus, it captures special states of particularly high volatility in the crude oil market. Many of the State 3 periods can be associated with specific events which were important for the crude oil market as listed in Barsky and Kilian (2004). For instance, there are peaks associated with the October War and oil embargo from October 1973 through early 1974, the

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Iranian Revolution between October 1978 and February 1979, the outbreak of the Iran-Iraq war in September 1980, the invasion of Kuwait in August 1990, and the important OPEC meeting in March 1999. Thus, the three-state model is not only favoured by SC but also makes good sense when looking at the developments in the crude oil market.

The situation is somewhat different for the four state model. The first and second states are associated with similar periods as in the MS(3) model, as seen in Figure 2.2 and also the associated covariance matrices in Table 2.3 are similar. Thus, the first state is once again associated with periods of low volatility in oil production while the second state captures periods with higher volatility in the price of oil. In the third and fourth states, the volatility of oil production is again very high while the volatility of the price of oil is very large in State 3 only. Unfortunately, State 3 is only associated with relatively few periods. Hence, the model is difficult to estimate and the estimates are unreliable. Overall, from a statistical point of view as well as considerations of events in the crude oil market, the model is more problematic than a MS(3) model and, hence, we favor the MS(3)-VAR(3) model in the following.

In Figure 2.3 the standardized residuals of the MS(3)-VAR(3) model are presented. They are determined as

$$\hat{\Sigma}_{t|t-1}^{-1/2}\hat{u}_t,$$

where $\hat{u}_t = y_t - \hat{\nu} - \hat{A}_1 y_{t-1} - \cdots - \hat{A}_p y_{t-p}$ and

$$\hat{\Sigma}_{t|t-1} = \sum_{m=1}^{M} \widehat{\Pr}(s_t = m|\mathbf{Y}_{t-1})\hat{\Sigma}_m.$$

Here $\hat{\Sigma}_{t|t-1}$ denotes the estimated residual covariance matrix conditional on information up to time $t-1$, that is, conditional on $\mathbf{Y}_{t-1} = (Y_{t-1}, \ldots, Y_1)$. Clearly, the volatility of the standardized residuals in Figure 2.3 is more regular than in Figure 2.1. In other words, the MS structure captures the changes in volatility well. Hence, we will focus the analysis on the MS(3)-VAR(3) model in the following, which is also plausible from a subject matter point of view.

2.3.4 Statistical analysis of MS(3)-VAR(3) model

Because we are interested in using the MS structure for identification purposes, a main question of interest is whether a state-invariant initial effects matrix B is compatible with the data and whether the associated relative variances are sufficiently different to obtain a statistical identification of the shocks. In Table 2.2 we already see that both AIC and SC support a state-invariant B matrix for the MS(3) model. In Table 2.4 the likelihood ratio test from Lanne et al. (2010) for the null hypothesis of a state-invariant B is presented. Its p-value is 0.63 and hence exceeds conventional significance levels substantially. Thus, the LR test also supports a state-invariant initial effects matrix.

The estimated λ_{ij}s of the MS(3)-VAR(3) model with state-invariant B are shown in Table 2.5 together with estimated standard errors. The standard errors indicate that estimation precision is quite reasonable and hence, we can hope for sufficient heterogeneity to get identification. Statistical tests of the relevant hypotheses are presented in Table 2.6. Recall that B is locally identified (identified apart from changes in sign and permutation of its columns) if for each pair of subscripts (i, j) there is a Λ_m matrix for which the corresponding diagonal elements are different, that is, $\lambda_{mi} \neq \lambda_{mj}$ for some $m \in \{2, \ldots, M\}$. For our three-dimensional system we thus have to check three pairs of subscripts. The corresponding Wald and LR test results are shown in Table 2.6. The Wald tests are attractive in the present context because they are very easy to compute from the estimates of the model with state-invariant B. On the other hand, they are known to be unreliable for highly nonlinear null hypotheses. Two of the three null hypotheses are clearly rejected by the Wald tests at a 5% significance level while the third one is rejected at a 10% level.

We also computed LR tests which were computationally more demanding. They require an additional likelihood optimization under the null hypothesis, which is a challenging task for the present model class being considered with the risk of ending up in some local optimum. Note that the LR tests are computed for models where the diagonal elements of Λ_2 are sorted in increasing order. The LR test values result in very small p-values and strongly support identification of B. Thus, we conclude that the data is in line with identified shocks. These shocks are the only ones in our framework that are time-invariant with the same impact effects across states. Note that time-invariant shocks and impulse responses are also assumed by Kilian (2009) and Kilian and Murphy (2011), but not by Baumeister and Peersman (2010), who allow for time-varying coefficients. We emphasise that our identifying restrictions so far are based on data driven procedures and the data are not objecting to them in the framework of the MS(3)-VAR(3) model.

Of course, our identification is a statistical one and there is still a question of whether it is consistent with economically meaningful shocks, in particular with the shocks considered by Kilian (2009) and Kilian and Murphy (2011). The question of whether our statistically identified shocks are in line with the sign identified shocks considered by Kilian and Murphy (2011) will be considered when we look at impulse responses in the next subsection. In the

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

present context however, we can perform a statistical test of the exclusion restrictions used by Kilian (2009), which are given in Equation (2.5). Having identifying information for the shocks from other sources, allows us to apply simple LR tests because Kilian's recursive identification scheme becomes over-identifying in the context of our model, as explained in Section 2.2. The AIC and SC values for models with state-invariant, lower-triangular B matrix in Table 2.2 and the corresponding LR tests for a MS(3)-VAR(3) model presented in Table 2.4 all support a lower-triangular B. In particular, the LR tests in Table 2.4 have very large p-values and hence, do not reject lower-triangularity of B. This holds for both a test against an unrestricted MS(3) model and against a MS(3) model with state-invariant B.

Further support for a lower-triangular B is presented in Table 2.5, where the estimated λ_{ij}s of the model with a lower-triangular B matrix are given. They are very close to those of the corresponding model with unrestricted B. Note that the columns of B can no longer be permuted when the matrix is triangular. Thus, the ordering of the λ_{ij}s is the one corresponding to the lower-triangular B matrix. It results in λ_{2j}s in decreasing order and hence, we order the λ_{2j}s for the model with unrestricted, state-invariant B accordingly. Apparently, from a statistical point of view, the assumption of a lower-triangular initial effects matrix is not restrictive. In the next section we will explore whether the sign restrictions used in the related literature are also in line with our statistically identified shocks. An impulse response analysis will be performed for that purpose.

2.3.5 Impulse response analysis

Given that we have identified shocks, we can also compute impulse responses from our MS(3)-VAR(3) model with state-invariant initial effects matrix B. These are displayed in Figure 2.4 together with 68% bootstrap confidence intervals. The scaling of the impulse responses is determined by B, which in turn is scaled such that the residuals in State 1 have identity covariance matrix. To some extent this is arbitrary because the numbering of the states is arbitrary. The scaling is also quite different from that used by Kilian (2009). We emphasise that our normalization of the shocks in the first state affects only the scaling but not the shape and the sign of the impulse responses.

A priori we do not know which economically interpretable shocks are represented by the statistically identified shocks. However, the fact that our ordering of the shocks is in line with a lower-triangular, recursive identification scheme indicates that they may be labeled as in (2.5), that is, the first shock may be viewed as an oil supply shock, the second as an aggregate demand shock and the third as an oil-market specific demand shock. With this labeling in mind, they can be seen to be in line with the sign restrictions given in Table 2.1. Only

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

the oil-market specific demand shock could be viewed as potentially problematic because its initial effect on output tends to be positive. Zero is at the lower end of the 68% confidence interval. Because the confidence level could be regarded as pretty low, our analysis provides little basis for questioning the sign restrictions. Using a slightly larger confidence level would result in intervals with negative values for the initial response of output. In other words, the data does not object to the sign restrictions, that is, the system as set up here and the data used are consistent with the sign restrictions adopted by Kilian and Murphy (2011).

Notice that the impulse responses of oil production are cumulated responses of $\Delta prod_t$. They fall outside the respective 68% bootstrap confidence bands. That feature is also observed in other studies and is not uncommon in the SVAR literature. It may be due to the skewness and bias of the distribution of the impulse response estimates.

The fact that our method does not exclude any of the restrictions imposed in the previous studies may be seen as a weakness, indicating a lack of power of our procedure. In defence of our method we note that the restrictions in Table 2.1 are indeed not very tight. In the literature using sign restrictions it is uncommon to restrict the impact effects only. Assuming that the sign restrictions hold for some quarters after the shock has occurred would be a more common set of restrictions in the related literature. If we consider such restrictions for illustrative purposes, it turns out that our shocks do not satisfy them. For example, if we stick with the 68% confidence intervals around the impulse responses and consider this as the possible range of impulse responses supported by the data, there is no shock that raises oil production, reduces real activity and increases the price of oil for four quarters after hitting the system. Thus, in our framework the data does not support the existence of an oil-market specific demand shock satisfying much more stringent conditions. This result illustrates that our procedure does have some discriminatory power. The restrictions used by Kilian and Murphy (2011) are apparently very weak and hence, difficult to reject by the data, although more stringent restrictions can be rejected. Of course, imposing weak restrictions and still being able to draw conclusions from the analysis is an advantage of the Kilian-Murphy study and our analysis gives further support to their results.

Kilian and Murphy (2011) show that if only the sign restrictions from Table 2.1 are used, some elasticities become unreasonably large. Therefore, they search for other types of information that can be used to obtain more reasonable responses. As mentioned earlier, they also impose restrictions on the oil supply elasticity. In the context of our approach we can use the statistical identification restrictions and obtain the impulse responses with 68% confidence intervals in Figure 2.4. In some cases they are not very precise but can be further restricted by imposing statistically valid restrictions such as lower-triangularity of B. The

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

corresponding impulse responses are depicted in Figure 2.5. Overall they are not much more precise than those from the unrestricted model in Figure 2.4, apart from the impact effects of aggregate demand and oil-market specific demand shocks that are now restricted to zero.

One may of course argue that the main gains in precision relative to the impulse responses shown in Kilian and Murphy (2011) and Kilian (2009) come from the smaller VAR lag order. Recall that these authors use lag order $p = 24$ while we use $p = 3$. Indeed our smaller order results in smoother impulse responses. They are qualitatively similar to those from Kilian (2009), however. To show this, we present the impulse responses of a conventional, recursively identified VAR(24) in Figure 2.6 and the corresponding quantities from a VAR(3) in Figure 2.7. The confidence intervals shown in these figures are obtained from a standard percentile procedure. The ones in Figure 2.6 look very similar to the corresponding ones in Figure 3 of Kilian (2009). Notice that Kilian reports one- and two-standard error bounds that are naturally symmetric around the impulse responses while we use (potentially asymmetric) 68% bounds for ease of comparison with our previous results.

The overall conclusion from our study is that using the statistical information more fully allows for more precise inference on the effects of structural shocks in a SVAR analysis. It provides additional identifying information that can complement the information from economic theory and allows to check restrictions that cannot be checked against the data in a conventional analysis with exclusion or sign restrictions. It also has the advantage of providing information on the overall compatibility of the model and the data used with the theoretical considerations of economists. In other words, it can provide information on the compatibility of the data despite potential deviations from an ideal model world due to omitted variables, measurement errors or other data problems.

2.4 Conclusions

We argue that restrictions derived from some economic model may not be valid in a corresponding empirical model. Reasons may be that the economic model is a simplification of reality that explicitly ignores important effects in the real economy. For example, there may be an omitted variables problem if only those variables that are described in the theoretical model are included in the empirical model. Moreover, the empirical variables may not correspond precisely to the variables underlying the economic model due to measurement errors, data adjustments such as trend or seasonal adjustments or because the theoretical variables are not measurable. In a conventional SVAR analysis one would still impose the restrictions suggested by economic reasoning. If the restrictions just-identify the shocks it may go unnoticed that the empirical model does not reflect the actual relationships between the variables involved. Similarly, if only sign restrictions are used to identify the shocks, in a conventional analysis the data have no basis to speak up against the model. Therefore, there are proposals to use statistical properties such as heteroskedasticity to complement economic restrictions and check restrictions that are just-identifying in a conventional SVAR framework. In this study we propose to use similar devices for checking sign restrictions.

More precisely, we propose to model changes in the volatility of the shocks by a MS mechanism and we show how this device can be used to check sign restrictions in addition to exclusion restrictions on the impact and long-run effects of the shocks. If the volatility structure is rich enough it may provide identified shocks which should correspond to the economic shocks if the empirical model is in line with the economic model. In this case, the statistically identified shocks have to satisfy, for example, the sign restrictions or other restrictions derived from economic considerations.

For illustrative purposes we use a small, three-dimensional model of the market for crude oil consisting of the change in oil production, the price of oil and an index of economic activity. For this model, oil supply and demand shocks, as well as an aggregate demand shock, are specified with sign restrictions for the corresponding impulse responses, on the one hand, and changes in the volatility of the residuals are observed on the other. The changes in volatility are used to obtain identified shocks. It turns out that these shocks satisfy a previously used weak set of sign restrictions but not a slightly stricter set of such restrictions. Our analysis allows us not only to check previously used sign restrictions but it also enables us to impose further exclusion restrictions that may help obtain more precise estimated impulse responses. More generally, the study demonstrates how identifying restrictions in general and sign restrictions in particular can be checked in our framework.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Bibliography

Amisano, G. and Giannini, C. (1997). *Topics in Structural VAR Econometrics*, 2nd edn, Springer, Berlin.

Barsky, R. B. and Kilian, L. (2004). Oil and the macroeconomy since the 1970s, *Journal of Economic Perspectives* **18**: 115–134.

Baumeister, C. and Peersman, G. (2010). Sources of the volatility puzzle in the crude oil market, *Working Paper 2010/634*, Universiteit Gent.

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**: 139153.

Canova, F. (2002). Validating monetary DSGE models through VARs, *Discussion Paper 3442*, CEPR.

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

Canova, F. and Pappa, E. (2007). Price differentials in monetary unions: The role of fiscal shocks, *Economic Journal* **117**: 713–734.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII US, *Review of Economic Dynamics* **8**: 262–302.

Dedola, L. and Neri, S. (2007). What does a technology shock do? A VAR analysis with model-based sign restrictions, *Journal of Monetary Economics* **54**: 512–549.

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

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Herwartz, H. and Lütkepohl, H. (2011). Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks, *Working Paper series*, EUI, Florence.

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.

Kilian, L. and Murphy, D. (2011). Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market VAR models, *Journal of the European Economic Association* .

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.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*, Springer-Verlag, Berlin.

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.

Lippi, F. and Nobili, A. (2011). Oil and the macroeconomy: A quantitative structural analysis, *Journal of the European Economic Association* .

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

Moon, H. R. and Schorfheide, F. (2009). Bayesian and frequentist inference in partially identified models, *Working Paper 14882*, NBER.

Moon, H. R., Schorfheide, F., Granziera, E. and Lee, M. (2009). Inference for VARs identified with sign restrictions, *manuscript*, USC, Los Angeles.

Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks?, *Journal of Applied Econometrics* **24**: 960–992.

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

Pappa, E. (2009). The effects of fiscal shocks on employment and the real wage, *International Economic Review* **50**: 217–244.

Peersman, G. and Straub, R. (2009). Technology shocks and robust sign restrictions in a euro area SVAR, *International Economic Review* **50**: 727–750.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

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

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.

Rubio-Ramirez, J. F., Waggoner, D. and Zha, T. (2005). Markov-switching structural vector autoregressions: Theory and applications, *Discussion Paper*, Federal Reserve Bank of Atlanta.

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

Sims, C. A., Waggoner, D. F. and Zha, T. (2008). Methods for inference in large multiple-equation Markov-switching models, *Journal of Econometrics* **146**: 255–274.

Sims, C. A. and Zha, T. (2006). Were there regime switches in U.S. monetary policy?, *American Economic Review* **96**: 54–81.

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.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 2.1: Sign Restrictions for Impact Responses from Kilian and Murphy (2009)

	oil supply shock	aggregate demand shock	oil-market specific demand shock
oil production	$-$	$+$	$+$
real activity	$-$	$+$	$-$
oil price	$+$	$+$	$+$

Table 2.2: Comparison of MS-VAR(3) Models for $Y_t = (\Delta prod_t, q_t, p_t)'$

Model	$\log L_T$	AIC	SC
VAR(3) without MS	-4271.47	8614.94	8758.99
MS(2), unrestricted	-4106.82	8301.64	8477.71
MS(2), B lower-triangular	-4108.85	8299.70	8463.76
MS(3), unrestricted	-4048.84	8205.68	8421.76
MS(3), state-invariant B	-4049.70	8201.40	8405.47
MS(3), state-invariant, lower-triangular B	-4050.13	8196.27	8388.34
MS(4), unrestricted	-4013.32	8158.64	8422.73
MS(4), state-invariant B	-4030.42	8180.85	8420.93
MS(4), state-invariant, lower-triangular B	-4038.40	8190.81	8418.89

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.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 2.3: Estimated State Covariance Matrices of MS(m)-VAR(3) Models with State-Invariant B, $m = 2, 3, 4$, for $Y_t = (\Delta prod_t, q_t, p_t)'$

	$m = 2$	$m = 3$	$m = 4$
Σ_1	$\begin{bmatrix} 155.59 & & \\ 2.02 & 12.38 & \\ -9.43 & 2.36 & 58.29 \end{bmatrix}$	$\begin{bmatrix} 99.38 & & \\ 0.96 & 10.51 & \\ -0.86 & 1.09 & 42.84 \end{bmatrix}$	$\begin{bmatrix} 90.92 & & \\ 1.33 & 9.89 & \\ -0.98 & 0.49 & 40.27 \end{bmatrix}$
Σ_2	$\begin{bmatrix} 912.86 & & \\ 18.49 & 35.38 & \\ -0.56 & 0.56 & 2.94 \end{bmatrix}$	$\begin{bmatrix} 533.95 & & \\ 9.39 & 33.86 & \\ 5.07 & 0.64 & 2.45 \end{bmatrix}$	$\begin{bmatrix} 368.56 & & \\ 6.42 & 23.89 & \\ -0.60 & 1.06 & 1.52 \end{bmatrix}$
Σ_3		$\begin{bmatrix} 2499.03 & & \\ 65.38 & 37.42 & \\ 16.46 & 5.31 & 190.49 \end{bmatrix}$	$\begin{bmatrix} 947.64 & & \\ 17.89 & 39.83 & \\ -5.97 & 2.03 & 222.26 \end{bmatrix}$
Σ_4			$\begin{bmatrix} 2365.54 & & \\ 47.36 & 57.70 & \\ -3.55 & 2.49 & 9.59 \end{bmatrix}$

Table 2.4: LR Tests of Restrictions for MS(3)-VAR(3) Models for $Y_t = (\Delta prod_t, q_t, p_t)'$

H_0	H_1	LR	df	p-value
state-invariant B	unrestricted MS(3)-VAR(3)	1.72	3	0.63
state-invariant, lower-triangular B	state-invariant B	0.86	3	0.84
state-invariant, lower-triangular B	unrestricted MS(3)-VAR(3)	2.58	6	0.86

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 for the model under H_1.

Table 2.5: Estimates of Structural Parameters of MS(3)-VAR(3) Models for $Y_t = (\Delta prod_t, q_t, p_t)'$ with State-invariant B

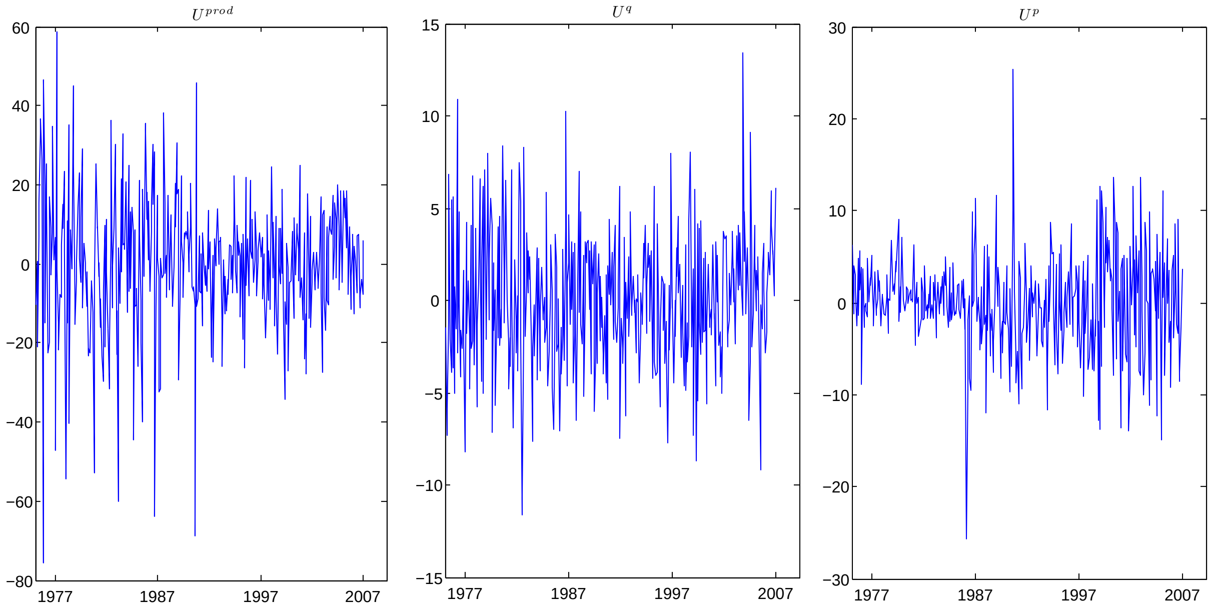
	unrestricted B		lower-triangular B	
parameter	estimate	std.dev.	estimate	std.dev.
λ_{21}	5.384	0.933	5.429	0.944
λ_{22}	3.210	0.485	3.220	0.489
λ_{23}	0.056	0.009	0.057	0.009
λ_{31}	25.235	6.938	25.097	7.107
λ_{32}	3.387	1.242	3.481	1.258
λ_{33}	4.441	1.018	4.382	1.011

Note: Standard errors are obtained from the inverse of the outer product of numerical first order derivatives.

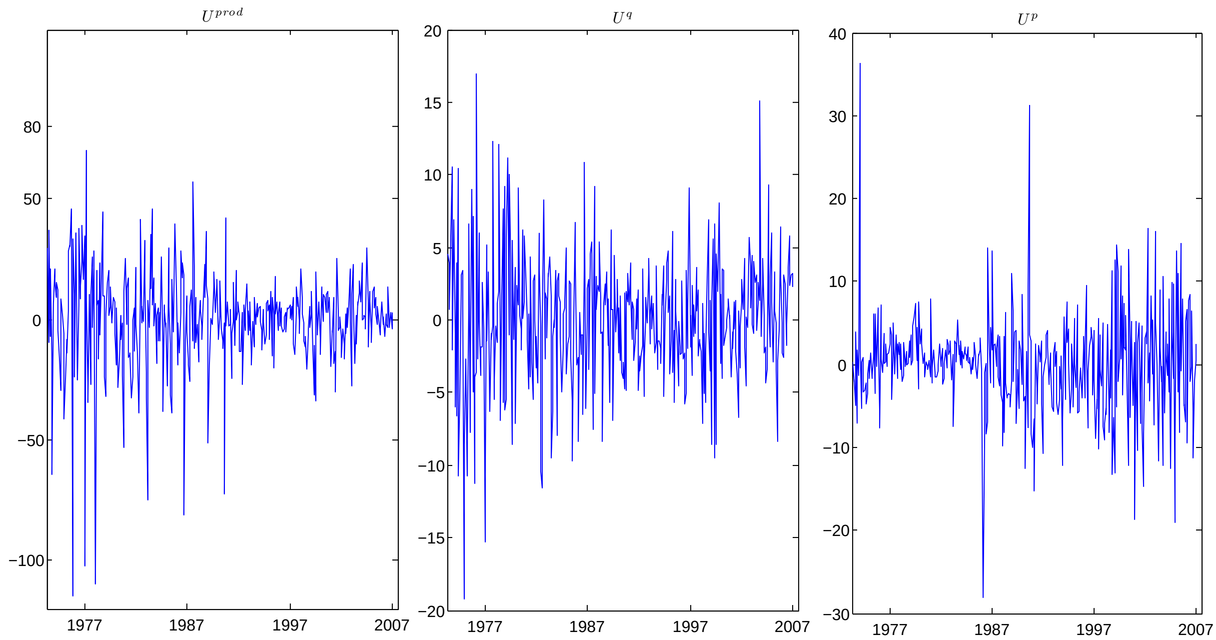
Table 2.6: Tests for Equality of λ_{ij}s for MS(3)-VAR(3) Model with State-invariant B

H_0	Wald statistic	p-value	LR statistic	p-value
$\lambda_{21} = \lambda_{22}$, $\lambda_{31} = \lambda_{32}$	7.99	0.02	20.39	3.7×10^{-5}
$\lambda_{21} = \lambda_{23}$, $\lambda_{31} = \lambda_{33}$	7.87	0.02	21.04	2.7×10^{-5}
$\lambda_{22} = \lambda_{23}$, $\lambda_{32} = \lambda_{33}$	5.16	0.07	27.15	1.3×10^{-6}

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498



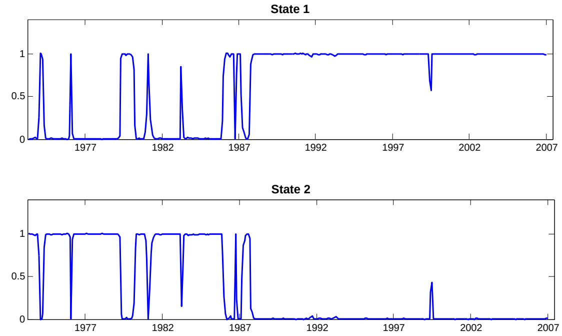
(a) Residuals of VAR(24) model



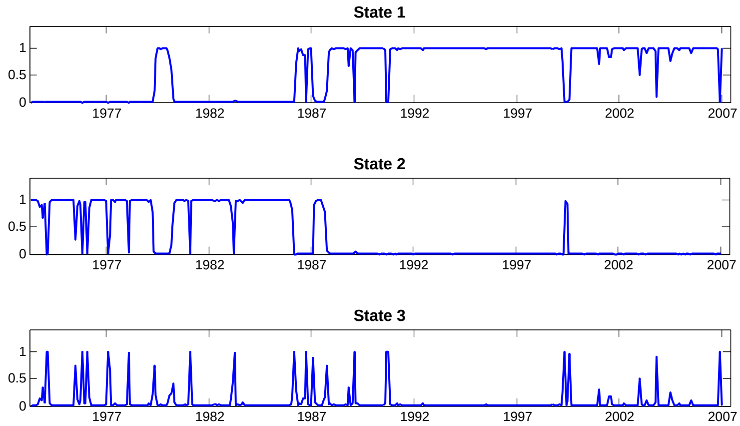
(b) Residuals of VAR(3) model

Figure 2.1: Residuals of VAR(24) and VAR(3) models.

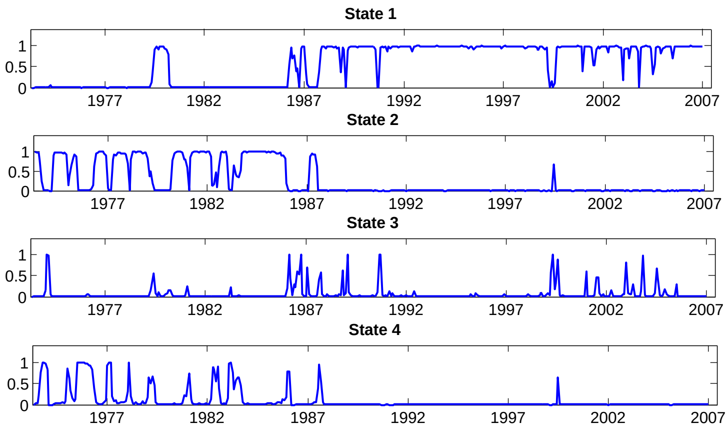
Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute
DOI: 10.2870/73498



(a) Smoothed state probabilities for MS(2)-VAR(3)



(b) Smoothed state probabilities for MS(3)-VAR(3)



(c) Smoothed state probabilities for MS(4)-VAR(3)

Figure 2.2: Smoothed state probabilities of unrestricted MS(m)-VAR(3) models for $m = 2, 3, 4$.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

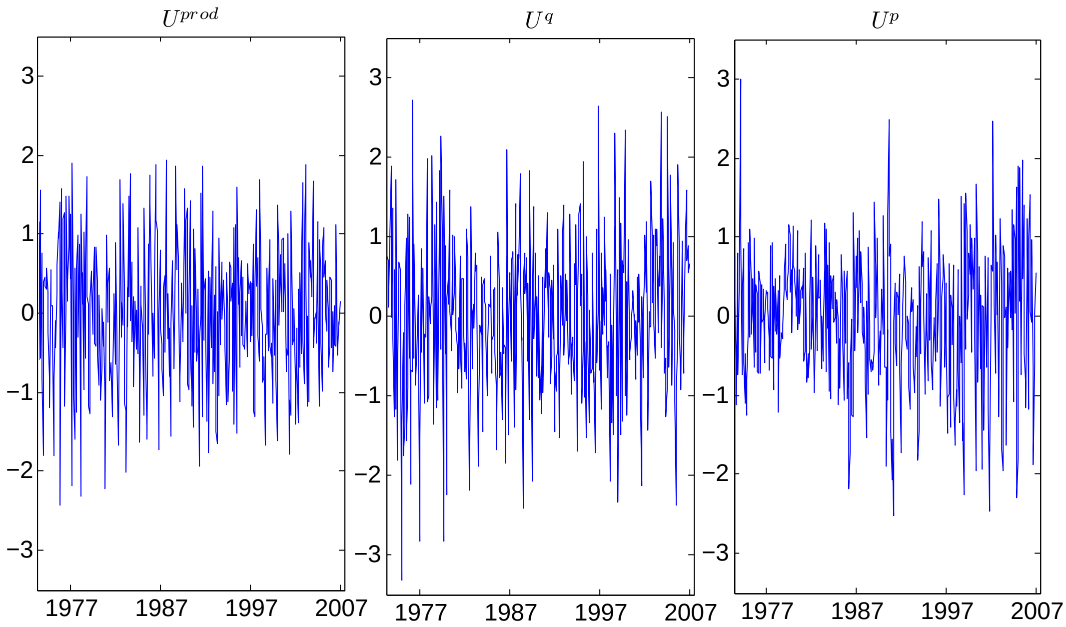


Figure 2.3: Standardized residuals of MS(3)-VAR(3) model with state-invariant B.

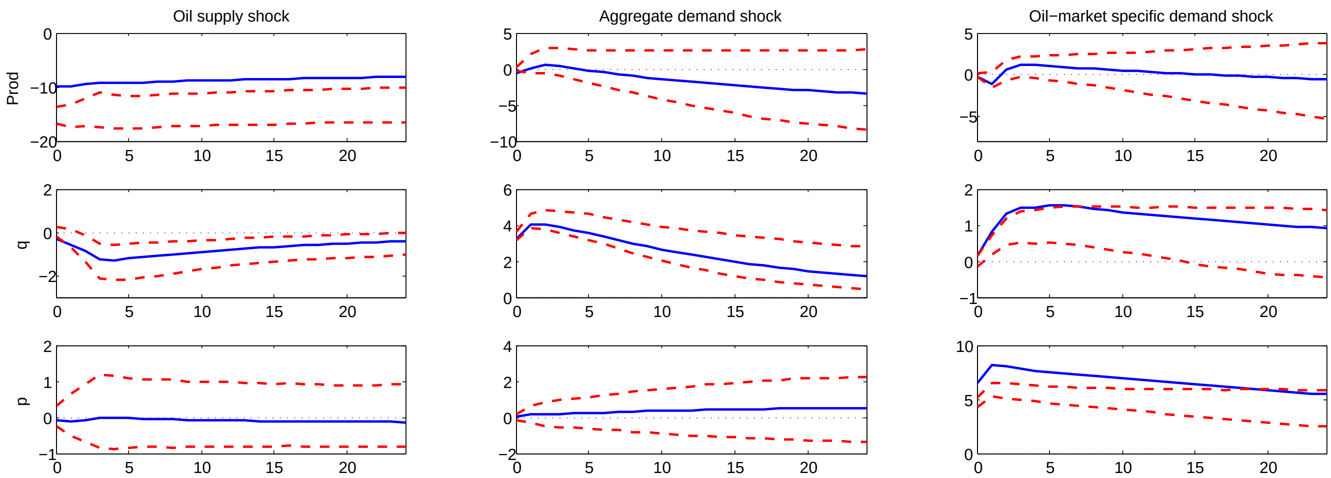


Figure 2.4: Impulse responses with 68% confidence bounds of the MS(3)-VAR(3) model with state-invariant B.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute
DOI: 10.2870/73498

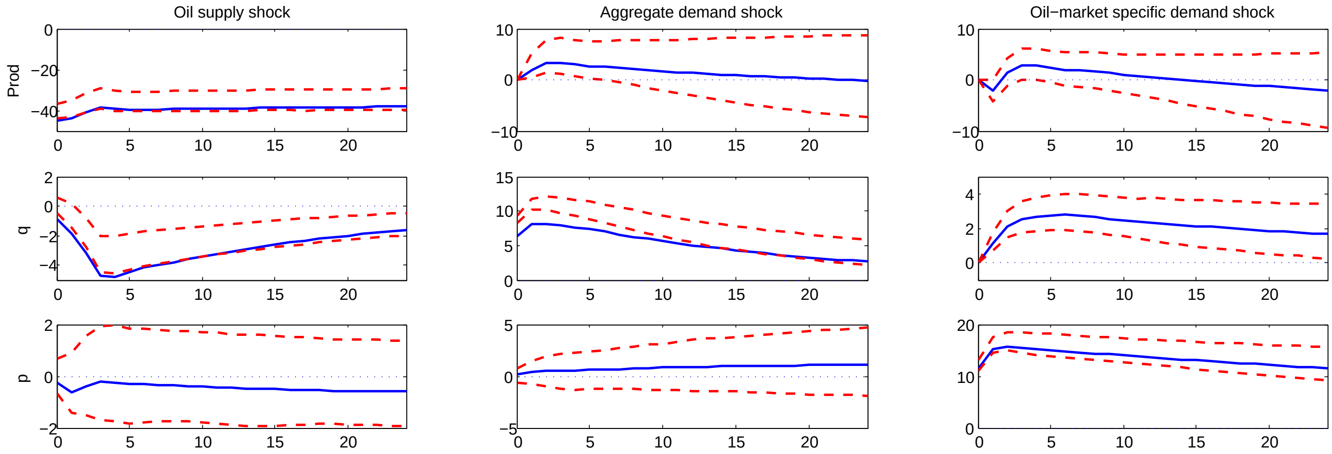


Figure 2.5: Impulse responses with 68% confidence bounds of the MS(3)-VAR(3) model with state-invariant, lower-triangular B.

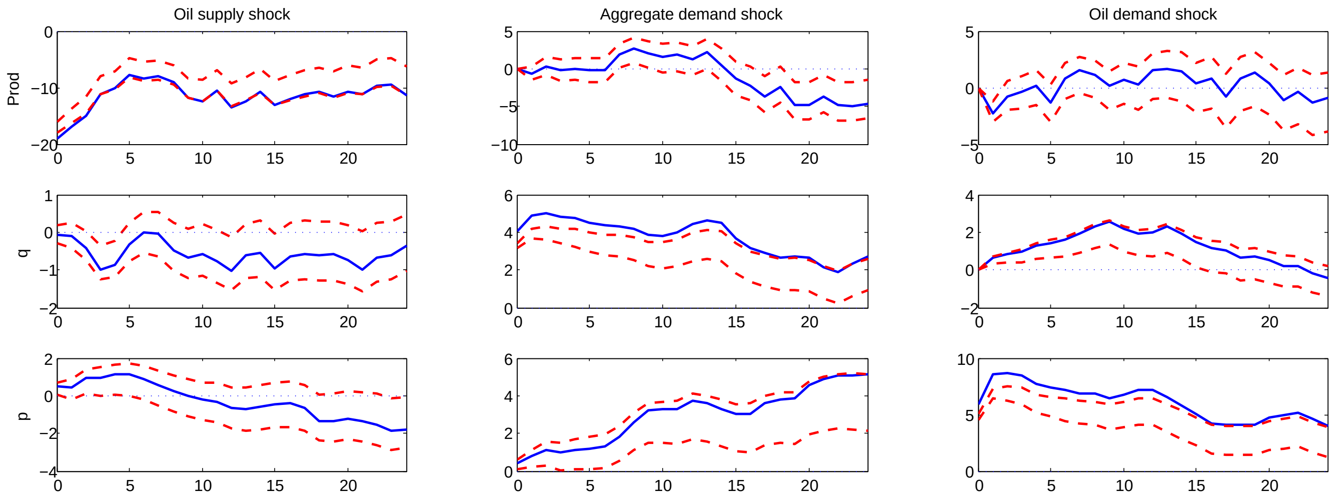


Figure 2.6: Impulse responses with 68% confidence bounds of a recursively identified VAR(24) model.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

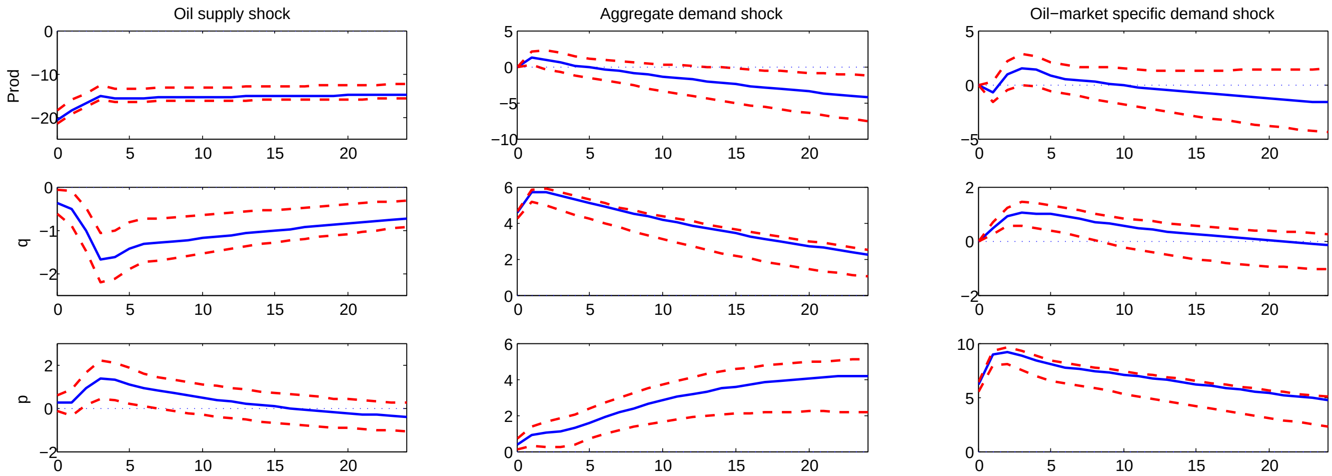


Figure 2.7: Impulse responses with 68% confidence bounds of a recursively identified VAR(3) model.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute
DOI: 10.2870/73498

Chapter 3

Reaction to Technology Shocks in Markov-switching Structural VARs: Identification via Heteroskedasticity

Aleksei Netšunajev [1]

3.1 Introduction

A standard real business cycle model implies that hours worked per capita rise after a permanent shock to technology. This prediction is at the centre of the literature that assesses whether it is consistent with the data. The general conclusion reached is that it is not. Not surprisingly, the result has attracted a lot of attention as technology shocks are a significant source of fluctuations in productivity and employment.

In the literature, one can find a variety of methods used to study reaction of hours worked to technology shocks, but the most common is based on structural vector autoregressive (SVAR) models. In a seminal paper, Gali (1999) identifies the technology shocks or, to put it differently, permanent productivity shocks, using long-run restrictions and he finds that hours worked fall after a positive technology shock. Several papers consider similar systems as in Gali (1999) and try to assess the validity of the identifying restrictions. A similar identification is used in Gali, Lopez-Salido and Valles (2003), Christiano, Eichenbaum and Vigfusson (2003), Francis and Ramey (2005), and Francis and Ramey (2009). The study by Francis and Ramey (2005) questions whether the shocks that are identified as in Gali (1999) can be classified as technology shocks. Using different

[1] My thanks to Professor Helmut Lütkepohl, participants in the seminar at the Bank of Estonia and the European University Institute, as well as an anonymous referee for useful comments on an earlier version of the paper and Professor Valery Ramey for providing the data.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

identifying assumptions, they find that all but one specification produced the result similar to Gali (1999). In other words, Francis and Ramey (2005) show that permanent real wage and permanent productivity shocks, after controlling for capital tax rate, produce a negative reaction of hours worked.

Christiano et al. (2003) find that treating per capita hours worked as a difference stationary process yields the result that hours worked fall after the technology shock; if, on the contrary, hours worked are assumed to be a stationary process, the result is the opposite: hours worked rise after the technology shock. Fernald (2007) and Francis and Ramey (2009) argue that there are low frequency movements in hours per capita that may distort the results of the SVAR in Christiano et al. (2003). After either detrending the data (Fernald, 2007) or applying a filter to the data (Francis and Ramey, 2009), the response of hours worked to a neutral technology shock becomes negative.

Fisher (2002) proposes to disentangle investment-specific and neutral technology shocks. Similarly, Canova, Lopez-Salido and Michelacci (2010) consider the effects of neutral and investment-specific technology shocks on hours. Both studies show that hours worked fall in response to neutral shocks and increase in response to investment-specific shocks. Chang and Hong (2006) propose to identify the permanent total factor productivity (TFP) shocks in a way that is similar to Gali (1999). They show that the reaction of hours worked to a permanent TFP shock is positive.

It should be noted that the studies listed above may share some common shortcomings. Firstly, the underlying assumptions just-identify the macroeconomic shocks and leave no place for the data to speak out against restrictions. The problem of just-identified shocks is discussed, among others, by Lanne and Lütkepohl (2008), Lanne, Lütkepohl and Maciejowska (2010), and Herwartz and Lütkepohl (2011). Second, studies of technology shocks (Gali, 1999; Francis and Ramey, 2005; Christiano et al., 2003; Canova et al., 2010; Chang and Hong, 2006) ignore relevant features of the data, namely heteroskedasticity. The presence of time-varying volatility is extensively discussed and documented by Kim and Nelson (1999), McConnell and Perez-Quiros (2000), Blanchard and Simon (2001), Stock and Watson (2003), so it should be taken into account.

It is useful to consider heteroskedasticity as it allows additional identifying information to be extracted from the data (Rigobon (2003)). In the present context this is important given the mixed evidence on the reaction of hours on technology shocks. Modeling heteroskedasticity can be used as a way of validating the restrictions that are just-identifying in a conventional SVAR analysis and for checking how different identification methods comply with the properties of the data.

Thus, the aim of the current paper is to reconsider the reaction of hours worked to technology shocks and to relax some of the assumptions common in this literature. For this purpose, I estimate a series of Markov-switching (MS) models that allow the changes in volatility and intercept to be captured, provide a framework to test for the validity of the identifying restrictions, and assess the labeling of identified shocks as technology shocks. The model used in the paper is a modified version of the model used by Lanne et al. (2010) and Herwartz and Lütkepohl (2011).

The rest of the paper is organized as follows. I provide additional motivations for the paper, while different identification schemes of technology shocks and the data are discussed in Section 3.2. In Section 3.3 the structural MS-VAR model deployed in the current analysis is described. Section 3.4 provides the empirical analysis. The last section concludes the analysis.

3.2 Identification of shocks

Consider a standard K-dimensional reduced form VAR with p lags:

$$Y_t = \nu + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + U_t, \tag{3.1}$$

where ν is a constant intercept term, the A_js ($j = 1, \ldots, p$) are $(K \times K)$ coefficient matrices and U_t is a zero-mean error term.

In a conventional SVAR model, the structural shocks are usually obtained from the reduced form residuals by a linear transformation, $\varepsilon_t = B^{-1} U_t$ or $B\varepsilon_t = U_t$, where B is such that ε_t has identity covariance matrix, that is, $\varepsilon_t \sim (0, I_K)$, and the reduced form residual covariance matrix is decomposed as $E(U_t U_t') = \Sigma_U = BB'$. To obtain unique structural shocks, one needs to place $K(K-1)/2$ restrictions. For this reason the B matrix is often assumed to be lower triangular. Thus the B is the matrix of instantaneous effects of the unique structural shocks.

In the related technology shock literature, a bivariate system is usually considered in the spirit of Gali (1999). Using long-run restrictions, one identifies two kinds of shocks: technology shocks and non-technology shocks. The shocks are identified in the following system, which is a moving average representation of a VAR:

$$\begin{bmatrix} \Delta x_t \\ \Delta n_t \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_t^z \\ \varepsilon_t^m \end{bmatrix} \tag{3.2}$$

where x_t denotes the log of labour productivity, n_t denotes the log of labour input, ε_t^z is the technology shock and ε_t^m is the non-technology shock, $C_{ij}(L)$ is a polynomial in the lag operator and Δ is the difference operator.

In the present paper I follow the strategy proposed by Blanchard and Quah (1989) and place the restrictions on the total impact matrix $\Xi_\infty = (I_K - A_1 - ... - A_p)^{-1} B$, which is identical to restricting the system in (3.2). It should be noted that the restrictions on Ξ_∞ can be transformed to the restrictions on B as shown in Lütkepohl (2005).

The most common identifying assumption restricts $C_{12}(1) = 0$, implying that only technology shocks have long-run effects on labour productivity (Gali, 1999). The non-technology shocks could thus be interpreted as demand shocks (Gali, 1999).

Another way of identifying technology shocks in the bivariate system is proposed by Francis and Ramey (2005). They argue that technology shocks should not have a long-run effect on hours or, in other words, they exclude permanent technology shocks. This restriction is implemented by constraining $C_{21}(1) = 0$ above. Francis and Ramey (2005) argue that the resulting residuals in the productivity equation may contain other shocks in addition to the productivity shock. For instance, these could be monetary shocks that have no long-run effect on hours. Therefore, this identification is different from the original one in Gali (1999) and may be problematic.

Francis and Ramey (2005) consider an alternative long-run restriction involving real wages using a theoretical result, i.e. that only a technology shock should have a permanent effect on real wages. Thus, an alternative way to identify the technology shock is to substitute real wages for productivity and to impose $C_{12}(1) = 0$.

Francis and Ramey (2005) discuss the notion that permanent changes in capital income taxation can also have permanent effects on productivity. To control for this, they include current and four lags of the level of capital tax rates as exogenous variables in the VAR. On the contrary, I add the capital income tax series to the system above and untangle the technology shocks and capital income tax shocks using the long-run restrictions provided in Francis and Ramey (2005). Both technology shocks and capital income tax shocks can affect labour productivity in the long run, while a non-technology shock cannot. Furthermore, permanent shifts in technology should not affect the long-run labour supply, while a capital income tax shock can have permanent effects on labour. Note that the described system would not be identified in the conventional SVAR, while the restrictions are testable in the MS-VAR framework.

Following Chang and Hong (2006), one can identify the permanent TFP shocks for the aggregate economy. This is done by substituting a TFP measure for productivity and imposing $C_{12}(1) = 0$.

Furthermore, augmenting the bivariate system with the real price of investment, one can disentangle investment-specific technology shocks and neutral technology shocks. Solely investment-specific technology shocks affect the price of investment in the long run, while both investment-specific and neutral shocks affect labour productivity in the long run. The identification corresponds to a lower-triangular Ξ_∞ matrix for the ordering of variables price of investment, productivity and hours.

Table 3.1 summarizes the variations of technology shocks used in the subsequent analysis. Notation for the variables is as follows: x_t log of labour productivity, n_t log of per capita hours worked, w_t log of real wage, TFP_t measure of total factor productivity, τ_t measure of capital tax, i_t log of real price of investment.

I use quarterly data from 1947:Q1 through 2010:Q4 to estimate the models for permanent technology and non-permanent technology shocks. The data is obtained from the Bureau of Labor Statistics. Standard ADF tests for both productivity and hours indicate the presence of a unit root. Christiano et al. (2003) argue that hours per capita cannot logically have a unit root as it is

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 3.1: Models used to study technology shocks

	Used by	Data	Restrictions
Model 1	Gali (1999)	$y_t = [\Delta x_t, \Delta n_t]'$	$C_{12}(1) = 0$
Model 2	Francis and Ramey (2005)	$y_t = [\Delta x_t, \Delta n_t]'$	$C_{21}(1) = 0$
Model 3	Francis and Ramey (2005)	$y_t = [\Delta w_t, \Delta n_t]'$	$C_{12}(1) = 0$
Model 4	Chang and Hong (2006)	$y_t = [\Delta TFP_t, \Delta n_t]'$	$C_{12}(1) = 0$
Model 5	Francis and Ramey (2005)	$y_t = [\Delta x_t, \Delta n_t, \tau_t]'$	$C_{12}(1) = C_{21}(1) = 0$
Model 6	Canova et al. (2010)	$y_t = [\Delta i_t, \Delta x_t, \Delta n_t]'$	$C(1)$ lower triangular

a bounded process. However, hours series have low frequency movements that should be taken into account (Fernald, 2007; Francis and Ramey, 2009). In the present paper I use hours series in first differences, as it is consistent with the technology shock literature. I use the data kindly provided by Professor Valery Ramey to estimate models for permanent real wage shocks and technology shocks after controlling for capital income tax. The data runs from 1947:Q1 through 2003:Q1 for the real wage model and through 1997:Q4 for capital income tax model. Furthermore, I use publicly available data from Chang and Hong (2006) and Canova et al. (2010) for permanent TFP (yearly data) and investment-specific technology shock models. All variables except the tax rate are entered in logarithms. Lag order four is selected for all the datasets to remain consistent with the previous studies. One exception is the permanent TFP data where lag order two, as suggested by the Akaike information criterion (AIC), has been chosen.

3.3 The Model

3.3.1 Markov Switching SVAR

Identification via heteroskedasticity initially appeared with Rigobon (2003). In SVAR analysis, it is proposed and used by Rigobon and Sack (2003) and Lanne and Lütkepohl (2008), among others. These authors show that if there are exogenously generated changes in the volatility of the shocks, the structural parameters could be effectively recovered from the reduced form model. This identification is based on the assumptions that the system is stable over time (the effects of shocks are the same regardless of the volatility regime) and that the structural shocks are orthogonal. These assumptions are usually implicit in the conventional structural VAR analysis, and hence are no more restrictive than usual. In particular, they are also common to the technology shock literature.

In the present paper I consider conditional heteroskedasticity, which allows for changes in the volatility to be determined from the data. I use the approach proposed by Lanne et al. (2010) and model the changes in volatility and intercept by a Markov regime-switching (MS) mechanism. It should be noted that the approach does not label shocks economically but is rather a tool to test

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

whether economic restrictions that are just-identifying in the conventional SVAR are consistent with the data. Specifically, I consider a modified version of the model by Lanne et al. (2010).

Consider the VAR(p):

$$Y_t = \nu_{s_t} + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + U_t, \tag{3.3}$$

I assume that the time dependent intercept ν_{s_t} as well as the distribution of the reduced form error term U_t 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 be normal,

$$U_t | s_t \sim N(0, \Sigma_{s_t}). \tag{3.4}$$

In addition to the state dependent covariance matrices, I also allow the intercept term ν_{s_t} to be dependent on the Markov process. Models with similar features, changes in covariances and intercept, are used in the empirical business cycle literature as, for example, in Hamilton (1989) and Krolzig (1997). Fernald (2007), using data similar to the data I use, tests for structural breaks in the productivity growth series and finds them to be likely. Potential breaks in the intercept of the hours series are also discussed by Canova et al. (2010). Therefore, the model deployed in the subsequent analysis must also capture potential non-regularities in the intercept. In the following I will stick to the notation similar to Krolzig (1997). MSIH(M)-VAR(p) will denote models with changes in the intercept and volatility where M denotes the number of Markov states and p the lag length.

The changes in the volatility of the residuals are used in this framework to test whether the identified shocks are in line with the properties of the data. For instance, if there are two volatility states ($M = 2$), then a decomposition of the covariance matrices $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2 B'$ exists, where $\Lambda_2 = \text{diag}(\lambda_{21}, \dots, \lambda_{2K})$ is a diagonal matrix with positive diagonal entries. The Λ_2 matrix is thus the matrix of relative variances. Suppose λ_{2i}s are all distinct. Then the decomposition is unique up to changes in the sign and permutations of the columns of B and corresponding changes in the ordering of the weighting matrix Λ_2 (Lanne et al., 2010).

Thus, under the assumptions of orthogonality and state invariant instantaneous effects, the structural shocks are uniquely determined by the transformation $\varepsilon_t = B^{-1} U_t$. Consequently, any further restrictions induced by theoretical models become over-identifying and testable. Under the normality assumption, the likelihood ratio test is suitable for the purpose. Degrees of freedom of the asymptotic χ^2 distribution of the test statistic coincide with the number or restrictions being tested.

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

$$\Sigma_1 = BB', \quad \Sigma_i = B\Lambda_i B', \quad i = 2, \ldots, M, \tag{3.5}$$

with diagonal Λ_i matrices is restrictive. An assumption that the decomposition exists imposes restrictions on the covariance matrices which can be tested. Hence, if there are three or more states with covariance matrices $\Sigma_1, ..., \Sigma_M$, the invariance of the initial effects of the shocks across states can be checked by a likelihood ratio test. According to the null hypothesis being tested, the covariance matrices have representations as in (3.5). The degrees of freedom for the asymptotic χ^2 is $0.5MK(K+1) - K^2 - K(M-1)$. In other words, the number of elements in B and diagonal elements of $M-1$ $\Lambda_i, i = 2...M$ matrices are substracted from the number of distinct elements in the M covariance matrices (Lanne et al., 2010).

It is worth pointing out that the requirement of having distinct relative variances is necessary for an exact identification of all shocks. The B is (locally) unique, if for each pair of equal diagonal elements, say, in $\Lambda_2 = diag(\lambda_{21}, ..., \lambda_{2K})$, there is a corresponding pair of distinct diagonal elements in one of the other $\Lambda_i = diag(\lambda_{i1}, ..., \lambda_{iK})$ (Lanne et al., 2010). For instance, if $\lambda_{2k} = \lambda_{2l}$ then $\lambda_{ik} \neq \lambda_{il}$ for $i = 3, ..., M$ must exist. An important advantage of the approach adopted in this paper is that the equality of λ_{mi}s can be checked with both Wald and likelihood ratio tests. Wald tests do not require the full optimization of the model under the alternative that is advantageous in the current setup. However, it may happen that the standard errors of the $\Lambda_m s$ are poorly estimated. In this situation, LR tests may be useful. Further discussion of tests for two and three state MS models can be found in Lanne et al. (2010) and Herwartz and Lütkepohl (2011).

Since I assume the normality of the residuals conditional on the states, the likelihood function can be set up and the model is estimated by maximum likelihood (ML). The concentrated likelihood function and detailed discussion of the related estimation problems can be found in Herwartz and Lütkepohl (2011). In the present paper the expectation maximization (EM) algorithm of Herwartz and Lütkepohl (2011) is adopted and updated to allow for changes in the intercept. The likelihood function is nonlinear, therefore numerical optimization is used. For estimation purposes, I bound diagonal elements of $\Lambda_i, i = 2, ..., M$ matrices away from zero. The optimization runs for a set of starting values to reduce the possibility of getting stuck in a local optimum.

3.3.2 Bootstrapping confidence bands

In the MS models, bootstrapping confidence bands for impulse responses may be problematic; therefore, discussion of the procedure deployed in the present paper is useful. Herwartz and Lütkepohl (2011) propose a fixed design wild bootstrap for constructing confidence intervals for impulse responses. They suggest constructing bootstrap samples conditional on estimated state probabilities and the ML estimates. For the current model, I take into account the changes in the intercepts

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

when constructing the bootstrapped series. One of the ways to do this is to use a weighted average of the intercept for each t, with the weights being the estimated state probabilities. Thus for the current model, the bootstrapped series can be represented as:

$$Y_t^* = \mu_t + \hat{A}_1 Y_{t-1} + \cdots + \hat{A}_p Y_{t-p} + U_t^*, \tag{3.6}$$

where $\mu_t = (\hat{\xi}_t \hat{\nu}_{s_t})'$ and $\hat{\xi}_t = [\hat{\xi}_{1t}, ..., \hat{\xi}_{Mt}]$ is a $1 \times M$ vector of estimated state probabilities for period t, $\hat{\nu}_{s_t}$ is a $M \times K$ matrix of estimated state dependent intercepts, $U_t^* = \eta_t \hat{U}_t$ and η_t is a random variable that has Rademacher distribution (takes values 1 and -1 with probability 0.5).

Note that I do not bootstrap a history of the hidden regimes but rather take it as given following Herwartz and Lütkepohl (2011). I bootstrap parameter estimates θ^* of $\theta = \text{vec}[\nu_{s_t}, A_1, \ldots, A_p]$ and B^* of B, conditional on the initially estimated transition probabilities. Therefore the weights for the intercept in the bootstrap loop do not change.

It should be noted that Herwartz and Lütkepohl (2011) also condition on the estimated Λ_i, $i = 2, \ldots, M$, matrices. I relax this assumption and estimate the weighting matrices in the bootstrap step. In order to eliminate any potential interchanges of columns of the B matrix one needs to impose an ordering of the diagonal elements of $\Lambda_i, i = 2, ..., M$, for unrestricted models. However, no additional ordering of the relative variances (diagonal elements of Λ_i) is required if just-identifying restrictions on the B or Ξ_∞ are imposed.

Apart from this, in each iteration of the bootstrap, I check if signs of the diagonal elements of the B^* are consistent with the signs of the diagonal elements of the initial estimate $\hat{B}$. This is done to avoid interchanges in signs of the B and to reduce confidence bands as discussed by Lütkepohl (2012). In general, to fix the sign, one should choose elements in the $\hat{B}$ with the lowest standard errors and carry over the signs to the bootstrap loop. For instance, I fix the elements on the main diagonal of the B to be positive for productivity-hours data. Then at each bootstrap step, if an element on the main diagonal of the B^* is negative, the relevant column of the B^* is multiplied by -1. Note that this procedure is simply a device for reducing confidence bands for impulse responses.

It should be emphasized that computing the bootstrapped impulse responses in this way requires a nonlinear optimization of the log likelihood as in the maximization step of the EM algorithm and is computationally demanding. I use ML estimates of $\hat{\theta}$ as starting values in each bootstrap replication. In the empirical analysis, I consider 90% confidence intervals based on 1000 replications.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

3.4 Empirical analysis

3.4.1 Statistical Analysis

For the purpose of validating restrictions, MS models with two and three states are estimated. In the current section I will focus primarily on the analysis of two state models. As will become clear further down, the main arguments regarding the identification will be valid independent of the number of states. Where any differences are detected, they will be discussed. Detailed results for three state models are available as an appendix upon request.

In Table 3.2, the range of estimated two state models together with the corresponding values of the log likelihood, the Akaike Information Criterion (AIC) and the Schwarz Criterion (SC), are presented. In the current study, models with different restrictions for each of the datasets are compared. According to the information criteria, the models with MS are preferred to the standard VAR models. As can be seen the values of the AIC and SC are reduced when the identifying restrictions are imposed. For the investmen-specific technology shock model, the SC favors the most restrictive model, whereas the AIC supports a set of restrictions that identify non-technology shock, leaving investment-specific and neutral technology shocks unidentified.

Table 3.2: Comparison of MSIH(2) Models

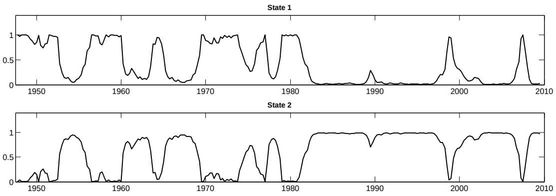
Model	Restrictions	LogL_t	AIC	SC
Model 1	VAR without MS	1696.55	-3351.10	-3277.06
	Unrestricted	1729.88	-3403.76	-3305.49
	$C_{12}(1) = 0$	1729.42	**-3404.85**	**-3310.09**
Model 2	$C_{21}(1) = 0$	1728.31	-3402.61	-3307.86
Model 3	VAR without MS	1442.10	-2842.20	-2770.55
	Unrestricted	1489.28	-2922.55	-2827.53
	$C_{12}(1) = 0$	1488.94	**-2923.88**	**-2832.25**
Model 4	VAR without MS	239.17	-452.34	-428.29
	Unrestricted	258.84	-477.69	-440.68
	$C_{12}(1) = 0$	258.32	**-478.64**	**-443.48**
Model 5	VAR without MS	1969.01	-3848.02	-3698.70
	Unrestricted	2026.71	-3941.42	-3756.96
	$C_{12}(1) = C_{21}(1) = 0$	2025.08	**-3942.16**	**-3762.98**
Model 6	VAR without MS	-523.23	1136.46	1280.89
	Unrestricted	-492.99	1097.98	1276.16
	$C_{12}(1) = 0$	-494.30	1098.60	1273.91
	$C_{13}(1) = C_{23}(1) = 0$	-493.85	**1095.72**	1268.83
	$C(1)$ lower tr.	-497.04	1100.08	**1268.71**

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.

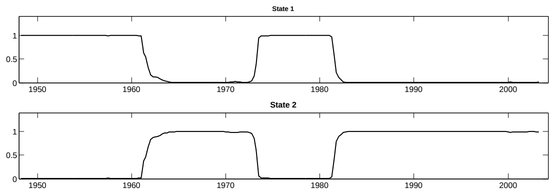
In addition to the model selection criteria, it is useful to look at the smoothed state probabilities.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

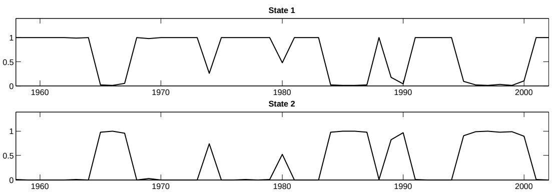
Figure 3.1: Smoothed state probabilities of MSIH(2) models



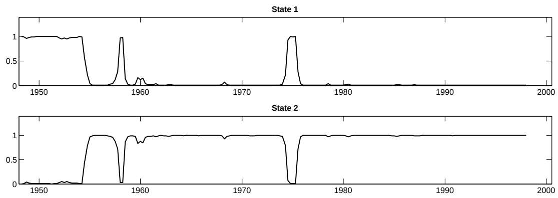
(a) Models 1 and 2



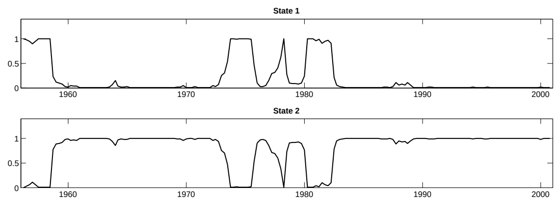
(b) Model 3



(c) Model 4



(d) Model 5



(e) Model 6

These are shown in Figure 3.1 for the two state models. The corresponding state covariance matrices for two state models are given in Table 3.3. The figures show that volatility changes are present during the sample period. From Table 3.3 it becomes clear that the States 1 and 2 of the MSIH(2) models can be interpreted as high and low volatility states, respectively. The variances of all of the variables are significantly lower in State 2 relative to State 1.

Periods of high volatility can be associated with the periods of economic downturns in the sample period for Models 1 and 2 (Figure 3.1 (a)). The estimated state probabilities reveal the great moderation phenomena that started at the beginning of the 80s and lasted until the late 90s. A similar picture can be seen in Figure 3.1 (c), where a measure of TFP is substituted for productivity. Estimated smoothed probabilities of the other datasets may not have a clear economic interpretation. For instance, both 3 variable models have a long duration of the low volatility state with a high volatility state being in place at the beginning and in the middle of the samples.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 3.3: Estimated State Covariance Matrices of MSIH(2) Models

Model	$\Sigma_1 \times 10^{-3}$	$\Sigma_2 \times 10^{-3}$
Models 1 and 2	$\begin{bmatrix} 0.124 & \\ 0.017 & 0.093 \end{bmatrix}$	$\begin{bmatrix} 0.033 & \\ -0.007 & 0.026 \end{bmatrix}$
Model 3	$\begin{bmatrix} 0.264 & \\ -0.063 & 0.107 \end{bmatrix}$	$\begin{bmatrix} 0.045 & \\ -0.008 & 0.030 \end{bmatrix}$
Model 4	$\begin{bmatrix} 0.212 & \\ 0.085 & 0.501 \end{bmatrix}$	$\begin{bmatrix} 0.016 & \\ -0.001 & 0.023 \end{bmatrix}$
Model 5	$\begin{bmatrix} 0.114 & & \\ 0.006 & 0.122 & \\ 0.029 & 0.128 & 0.539 \end{bmatrix}$	$\begin{bmatrix} 0.067 & & \\ 0.001 & 0.042 & \\ 0.001 & 0.011 & 0.049 \end{bmatrix}$
Model 6	$\begin{bmatrix} 996.3 & & \\ -202.3 & 827.1 & \\ 522.3 & 287.1 & 1323.1 \end{bmatrix}$	$\begin{bmatrix} 573.0 & & \\ -36.8 & 158.4 & \\ 7.3 & -7.5 & 247.5 \end{bmatrix}$

I intend to use the MS structure for identification purposes; therefore, the main question of interest is whether assumptions needed for local identification are satisfied. Recall from Section 3.3 that to obtain a statistical identification of the shocks for a two state model, it is enough to check whether the associated relative variances of unrestricted models are sufficiently different from each other. The estimates of λ_{2i}s together with the estimated standard errors for a range of MSIH(2) models are shown in Table 3.4. The standard errors indicate that the estimation precision is quite good for the two variable models and reasonable for the three variable models. Hence, I anticipate that the estimates are statistically different.

Table 3.4: Estimates of structural parameters of unrestricted MSIH(2) Models

Data	$\hat{\lambda}_{21}$	std.dev	$\hat{\lambda}_{22}$	std.dev	$\hat{\lambda}_{23}$	std.dev
Models 1 and 2	0.181	0.060	0.396	0.099		
Model 3	0.169	0.038	0.309	0.081		
Model 4	0.025	0.032	0.122	0.213		
Model 5	0.090	0.074	0.437	0.373	0.589	0.943
Model 6	0.138	0.068	0.244	0.162	1.021	0.573

Note: Standard errors are obtained from the inverse of the outer product of numerical first order derivatives.

Recall that B is locally identified in the two state model (apart from changes in the sign and permutation of its columns) if each pair of the diagonal elements of the Λ_2 matrix is distinct. For the two-dimensional system I thus have to check the equality of one pair of the diagonal elements λ_{21} and λ_{22}. For the three variable systems, three pairwise equalities must be checked. In the related literature Wald and likelihood ratio (LR) tests are used in the context (Lanne et al., 2010; Herwartz and Lütkepohl, 2011). Given that some of the standard errors of the Λ_2 elements shown in Table 3.4 are relatively high, the Wald test may perform poorly. Therefore I use computationally more demanding LR tests. The results are presented in Table 3.5. For the two variable models, the null hypotheses are rejected by the tests at a 5% significance level. For the three variable models the situation is somewhat different. For the capital tax augmented data, one pairwise equality cannot be rejected at a high level with $p = 0.418$. The remaining tests produce p values at around 0.2. Recall that there are only two restrictions to test using this data, and therefore the very high p value is not a big problem. Test results for the last specification are better, with the highest p value being 0.189 and the others below 10%. The results for the three variable models may be caused by an imprecisely estimated Λ_2 rather than by the true equality of its diagonal elements, as the required covariance matrix decomposition always exists for two states. The relatively short duration of the second state may have influenced the estimation precision of Λ_2.

Hence, based on the LR tests, there is strong evidence in favour of a unique B for the two variable models, as well as enough evidence for three variable models. This means that I have achieved a statistical identification of the two state models. The shocks obtained are unique but they are not labeled economically. With this identification in hand the economic restrictions on Ξ_∞ become overidentifying. The main question is whether the data supports the economically meaningful technology shocks identified by Gali (1999), Francis and Ramey (2005), Chang and Hong (2006) and Canova et al. (2010).

Table 3.5: Test for Equality of $\lambda_{ij}s$ for unrestricted MSIH(2) Models

Data	$H_0 : \lambda_{21} = \lambda_{22}$		$H_0 : \lambda_{22} = \lambda_{23}$		$H_0 : \lambda_{21} = \lambda_{23}$	
	LR	p	LR	p	LR	p
Models 1 and 2	6.275	0.012				
Model 3	4.221	0.039				
Model 4	5.955	0.014				
Model 5	1.625	0.202	1.623	0.203	0.656	0.418
Model 6	1.720	0.189	2.810	0.093	12.267	0.001

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 for the model under H_1. Here under H_1 are unrestricted MSIH(2) Models

The usual LR tests are applicable to perform testing of the restrictions. A small Monte Carlo experiment shows, that the probability to reject a true null hypothesis is 7%, showing the test has reasonable power. The outcomes of the LR tests are shown in Table 3.6. The LR test for the datasets support the lower-triangular Ξ_∞ matrix at a 5% level for all the models, with the exception of investment-specific technology data. One can reject the identification scheme of investment-specific and neutral technology shocks at a 5% level. To understand the sources of the rejection, additional models with identified neutral and non-technology shocks are estimated. The separate identification of these two shocks is supported by the data. Therefore imposing both restrictions simultaneously leads to the rejection of the lower triangular Ξ_∞ matrix. The non-permanent technology shock identified in Francis and Ramey (2005) also seems to have less support from the

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

data with $p = 0.07$. Clearly, the remaining ways of identifying technology shocks are consistent with the properties of the data.

Table 3.6: LR Tests of Restrictions for MSIH(2) Models

Model	Restriction under H_0	DF	LR	p-value
Model 1	$C_{12}(1) = 0$	1	0.92	0.337
Model 2	$C_{21}(1) = 0$	1	3.14	0.07
Model 3	$C_{12}(1) = 0$	1	0.67	0.41
Model 4	$C_{12}(1) = 0$	1	1.05	0.30
Model 5	$C_{12}(1) = C_{21}(1) = 0$	2	3.26	0.19
Model 6	$C_{12}(1) = 0$	1	2.26	0.13
	$C_{13}(1) = C_{23}(1) = 0$	2	1.74	0.42
	$C(1)$ lower tr.	3	8.10	0.04

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 for the model under H_1. Here under H_1 are unrestricted MSIH(2) Models

The following differences in the testing outcomes should be mentioned for the three state models. State invariant instantaneous responses are not supported for Models 4, 5, and 6. For Model 4, the restriction is rejected at a 5% level in favour of the fully unrestricted three state model. The identifying restrictions are strongly rejected for both three variable models. However, this result is driven by the rejection of the state invariant B for Model 5 and Model 6. For Model 5, the identifying restriction itself cannot be rejected at a 5% level for the alternative of state invariant B. The identification of investment-specific technology shocks for Model 6 is rejected at a high level, while the identification of neutral technology shocks is not.

Table 3.7: Estimates of structural parameters of restricted MSIH(2) Models

Data	$\hat{\lambda}_{21}$	std.dev	$\hat{\lambda}_{22}$	std.dev	$\hat{\lambda}_{23}$	std.dev
Model 1	0.381	0.098	0.189	0.044		
Model 2	0.357	0.088	0.208	0.045		
Model 3	0.179	0.038	0.301	0.085		
Model 4	9.726	17.14	64.81	63.86		
Model 5	3.043	1.601	11.59	4.580	1.562	0.911

Note: Standard errors are obtained from the inverse of the outer product of numerical first order derivatives.

In Table 3.7 the relative variances of the structural shocks for the restricted models are presented. Model 6 is omitted as the joint identification of investment-specific and neutral technology shocks is not supported by the data. Note that for restricted Models 1, 2, and 3 State 1 is the high volatility state ($\hat{\lambda}_{2j}s < 1$), but for Models 4 and 5 State 1 is the low volatility state ($\hat{\lambda}_{2j}s > 1$). The relative variances associated with technology shocks are $\hat{\lambda}_{21}$s for all models. If $\lambda_{21} > \lambda_{22}$, then the volatility of the shock associated with λ_{21} is higher than the volatility of the shock associated

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

with λ_{22}. The technology shocks as in Gali (1999) and Francis and Ramey (2005) are more volatile in the low volatility states than the non-technology shocks. As the low volatility state can be associated with good times in the economy, it is reasonable that non-technology shocks (say, demand shocks) exhibit low volatility. There is more variation in the technology shocks, as technological development is an ongoing process. Plausibly, the technological innovations are implemented in good times when more resources are available, hence giving rise to adjustments in productivity rather than employment. Moreover, productivity is more volatile than hours in both states (see Table 3.3).

The TFP shocks exhibit a lower increase in variance in the high volatility state than non-TFP shocks. Hence, when the economy is in turbulent times, the work force - rather than TFP - reacts. In the high volatility state the technology shocks, after controlling for capital tax, have higher volatility than capital tax shocks but lower then non-technology shocks. The work force exhibits more pressure in bad times relative to technology, supporting the result for the TFP shocks.

The outcome of the testing can be briefly summarized as follows: (1) the identification of permanent technology shocks as in Gali (1999), non-permanent technology shocks, and permanent real wage shocks as in Francis and Ramey (2005) is supported for the models with two and three Markov states; (2) permanent TFP shocks as in Chang and Hong (2006) and permanent technology shocks after controlling for capital tax (Francis and Ramey, 2005) are supported in two state models; (3) disentangling investment-specific and neutral technology shocks as in Fisher (2002) and Canova et al. (2010) is not supported by the data independent of the number of states. However, a neutral technology shock can be identified in the system. With these results, impulse response analysis is performed.

3.4.2 Impulse Response Analysis

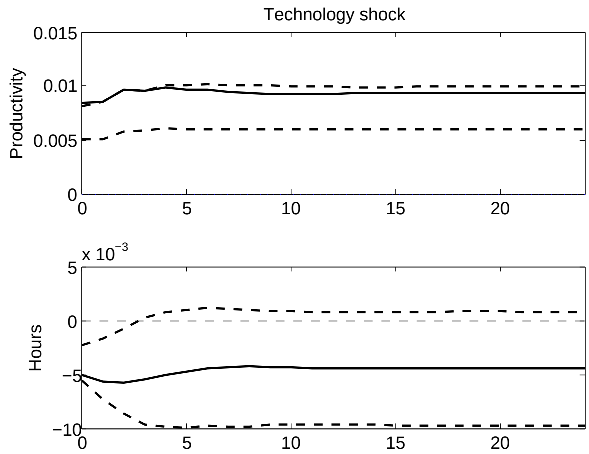
Given that the majority of the identification schemes were supported by the data, the impulse response (IR) analysis may be performed for the supported identification. The impulse responses for the variables that enter in first differences are accumulated. Some of the impulse responses fall outside the respective 90% bootstrap confidence bands. This feature has also been observed in some other studies and is not uncommon in the VAR literature. In the current study it might be due to a complex optimization step in the bootstrap cycle.

In Figure 3.2, the responses to the technology shock identified as in Gali (1999) are shown. The responses are consistent with the previous findings in the literature (Gali, 1999; Christiano et al., 2003; Francis and Ramey, 2005): productivity improves significantly, while hours are negative on impact; they then rise but remain negative. It should be noted that the upper confidence band starting from around 4-th quarter is above 0. This feature is also common to the results in the related literature.

In Figure 3.3, the responses to non-permanent technology and real wage shock identified in

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 3.2: Responses to a positive technology shock, Model 1



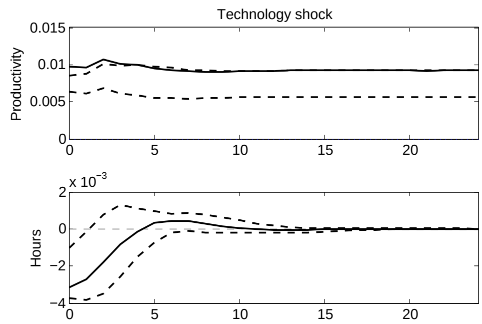
Francis and Ramey (2005) are shown. The dynamics is similar to the permanent technology shock: productivity (real wages) increase while hours worked drop on impact. The lowest bound of the 90% confidence interval for the response of hours to wage shock is actually above zero at horizon five and later. In other words, the response of hours is positive at that particular confidence level. However, the lower bound is so close to zero that for a wider confidence interval the response would become insignificant. Figure 3.3(c) shows the reaction of productivity and hours to technology shocks after controlling for capital tax. The responses do not change much with respect to Figure 3.3(a). The response of productivity is positive on impact and of hours worked remains insignificant. Hence, there is no full proof evidence in favour of positive and significant reaction of hours to these types of technology shocks.

Figure 3.4 shows the impulse responses to the TFP shock as in Chang and Hong (2006). On impact, the TFP is positive and hours negative. Then both start rising. Hours do not become positive in the horizon of six years; moreover, the reaction is insignificant on nearly all the response horizon. This contrasts findings in Chang and Hong (2006). This should not be surprising, given that the lower confidence band is only slightly above zero in this study (Chang and Hong (2006), Figure 1).

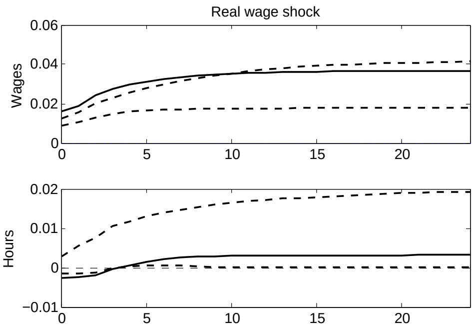
The last specification considered disentangled neutral and investment-specific technology shocks. Recall that full identification was not supported by the data, but an identified neutral technology shock was supported for two and three state models. Therefore it may be useful to study the impulse responses of the system with an identified neutral technology shock and try to identify the investment-specific technology shocks. The responses are shown in Figure 3.5. The investment-specific technology shock is identified as the only shock that positively and significantly increases

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

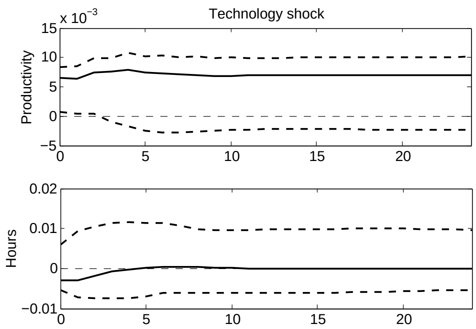
Figure 3.3: Responses to shocks identified as in Francis and Ramey (2005)



(a) Response to non-permanent technology shock, Model 2



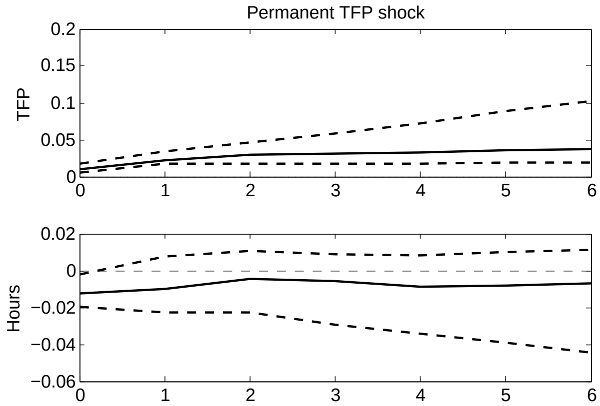
(b) Response to real wage shock, Model 3



(c) Responses to a positive technology shock after controlling for capital tax, Model 5

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 3.4: Responses to a positive TFP shock, Model 4



the real price of investment on impact. Unfortunately, the confidence bands for the impulse responses are quite wide for the model. One can see that within the limit the response of hours worked is positive, but that it is insignificant for both types of technology shocks.

Given the variety of studied impulse responses, there is no strong clear-cut evidence in favor of a positive and significant reaction of hours to different technology shocks, although it is plausible for some models.

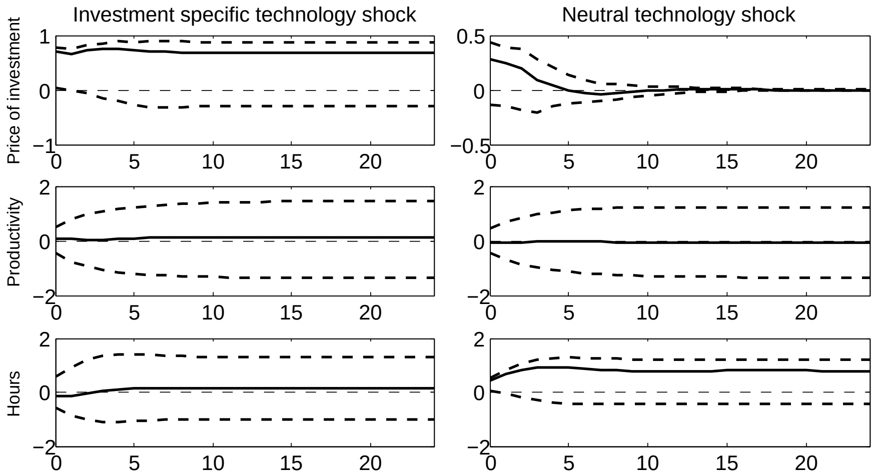
3.5 Conclusions

In the present paper I reconsider the effect of technology shocks on productivity and hours worked. I use Markov switching VAR instead of a standard VAR and assume that the intercept and variance-covariance matrices change over time. The reason for doing this is that proposals have been made to use heteroskedasticity in order to complement and test just-identifying economic restrictions. Identification via heteroskedasticity is particularly useful in the current analysis as there are several ways to identify technology shocks discussed in the literature.

Different identification schemes with long-run restrictions are used by Gali (1999), Francis and Ramey (2005), Chang and Hong (2006), Fisher (2002) and Canova et al. (2010). The studies listed above propose studying permanent and non-permanent technology shocks, permanent real wage shocks, permanent TFP shocks as well as to disentangle investment-specific and neutral technology shocks. In the conventional framework potentially competing restrictions are just-identified and hence not testable. In contrast, the present setup of the econometric model allows for the extraction

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Figure 3.5: Responses to an investment-specific and neutral technology shocks, Model 6



of additional information from the data and to test just-identifying long-run restrictions.

The results of the testing procedure show that the identification of permanent technology shocks as in Gali (1999), non-permanent technology shocks and permanent real wage shocks as in Francis and Ramey (2005) is supported by the data. Furthermore, permanent TFP shocks as in Chang and Hong (2006) and permanent technology shocks after controlling for capital tax (Francis and Ramey (2005)) are supported in two state models. Finally, disentangling investment-specific and neutral technology shocks as in Fisher (2002) and Canova et al. (2010) is not supported by the data independent of the number of states. However, a neutral technology shock can be identified in the system.

In conclusion, given the variety of impulse responses studied, there is no strong evidence in favor of a positive reaction of hours to technology shocks. A positive and significant reaction is plausible only for a real wage shocks and investment-specific – neutral technology shocks tandem. The latter result is achieved even though the original identification by Fisher (2002) and Canova et al. (2010) is rejected by the data. However, a better way of computing impulse responses would be useful for more precise inference.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Bibliography

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

Blanchard, O. and Simon, J. (2001). The long and large decline in u.s. output volatility, *Brookings Papers on Economic Activity* **32**(1): 135–174.

Canova, F., Lopez-Salido, D. and Michelacci, C. (2010). The effects of technology shocks on hours and output: a robustness analysis, *Journal of Applied Econometrics* **25**(5): 755–773.

Chang, Y. and Hong, J. H. (2006). Do technological improvements in the manufacturing sector raise or lower employment?, *American Economic Review* **96**(1): 352–368.

Christiano, L. J., Eichenbaum, M. and Vigfusson, R. (2003). What happens after a technology shock?, *NBER Working Papers 9819*, National Bureau of Economic Research, Inc.

Fernald, J. G. (2007). Trend breaks, long-run restrictions, and contractionary technology improvements, *Journal of Monetary Economics* **54**(8): 2467 – 2485.

Fisher, J. D. M. (2002). Technology shocks matter, *Working Paper Series WP-02-14*, Federal Reserve Bank of Chicago.

Francis, N. and Ramey, V. A. (2005). Is the technology-driven real business cycle hypothesis dead? shocks and aggregate fluctuations revisited, *Journal of Monetary Economics* **52**(8): 1379–1399.

Francis, N. and Ramey, V. A. (2009). Measures of per capita hours and their implications for the technology-hours debate, *Journal of Money, Credit and Banking* **41**(6): 1071–1097.

Gali, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?, *American Economic Review* **89**(1): 249–271.

Gali, J., Lopez-Salido, J. D. and Valles, J. (2003). Technology shocks and monetary policy: assessing the fed's performance, *Journal of Monetary Economics* **50**(4): 723–743.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* **57**(2): 357–84.

Herwartz, H. and Lütkepohl, H. (2011). Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks, *Working Paper series*, EUI, Florence.

Kim, C.-J. and Nelson, C. R. (1999). Has the u.s. economy become more stable? a bayesian approach based on a markov-switching model of the business cycle, *The Review of Economics and Statistics* **81**(4): 608–616.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*, Springer-Verlag, Berlin.

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.

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

Lütkepohl, H. (2012). Reducing confidence bands for simulated impulse responses, *Discussion Papers of DIW Berlin 1235*, DIW Berlin, German Institute for Economic Research.

McConnell, M. M. and Perez-Quiros, G. (2000). Output fluctuations in the united states: What has changed since the early 1980's?, *American Economic Review* **90**(5): 1464–1476.

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.

Stock, J. H. and Watson, M. W. (2003). Has the business cycle changed and why?, *NBER Macroeconomics Annual 2002, Volume 17*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 159–230.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute
DOI: 10.2870/73498

Appendix

Table 3.8: Comparison of MSIH(3) Models

Data	MSIH(3) Model	LogL_t	AIC	SC
Model 1	Unrestricted	1747.46	-3420.92	-3290.48
	State inv.B	1746.87	**-3421.75**	-3295.41
	$C_{12}(1) = 0$	1745.06	-3420.12	**-3297.29**
Model 2	$C_{21}(1) = 0$	1744.81	-3419.63	-3296.79
Model 3	Unrestricted	1497.49	-2920.98	-2794.75
	State inv.B	1497.43	-2922.87	-2800.69
	$C_{12}(1) = 0$	1497.25	**-2924.50**	**-2805.72**
Model 4	Unrestricted	268.39	-478.78	-425.13
	State inv.B	265.54	-475.09	-425.13
	$C_{12}(1) = 0$	264.61	**-475.23**	**-427.06**
Model 5	Unrestricted	2057.41	-3976.82	-3747.87
	State inv.B	2050.01	**-3968.01**	-3750.65
	$C_{12}(1) = C_{21}(1) = 0$	2047.19	-3966.38	**-3754.02**
Model 6	Unrestricted	-443.27	**1024.54**	**1245.99**
	State inv.B	-455.69	1043.39	1253.38
	$C_{12}(1) = 0$	-456.27	1042.54	1249.72
	$C_{13}(1) = C_{23}(1) = 0$	-465.95	1059.91	1263.90
	$C(1)$ lower tr.	-482.12	1090.24	1290.70

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.

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498

Table 3.9: LR Tests of Restrictions for MSIH(3) Models

Data	H_0	H_1	DF	LR	p-value
Model 1	St.inv. B	Unrestr.	1	1.18	0.27
	$C_{12}(1) = 0$	St.inv. B	1	3.62	0.05
	$C_{12}(1) = 0$	Unrestr.	2	4.80	0.09
	$C_{21}(1) = 0$	St.inv. B	1	4.12	0.04
Model 2	$C_{21}(1) = 0$	Unrestr.	2	5.30	0.07
Model 3	St.inv. B	Unrestr.	1	0.10	0.74
	$C_{12}(1) = 0$	St.inv. B	1	0.36	0.54
	$C_{12}(1) = 0$	Unrestr.	2	0.47	0.79
Model 4	St.inv. B	Unrestr.	1	5.68	0.01
	$C_{12}(1) = 0$	St.inv. B	1	1.83	0.17
	$C_{12}(1) = 0$	Unrestr.	2	7.53	0.02
Model 5	St.inv. B	Unrestr.	3	14.82	0.008
	$C_{12}(1) = C_{21}(1) = 0$	St.inv. B	2	5.60	0.06
	$C_{12}(1) = C_{21}(1) = 0$	Unrestr.	5	20.42	0.001
Model 6	St.inv. B	Unrestr.	3	24.84	0.00
	$C_{12}(1) = 0$	St.inv. B	1	1.16	0.28
	$C_{13}(1) = C_{23}(1) = 0$	St.inv. B	2	20.52	0.00
	$C(1)$ lower tr.	St.inv. B	3	52.87	0.00
	$C(1)$ lower tr.	Unrestr.	6	77.71	0.00

Netšunajev, Aleksei (2013), Structural Vector Autoregressions with Markov Switching: Identification via Heteroskedasticity
European University Institute DOI: 10.2870/73498