EESTI PANK

EUROSÜSTEEM

Working Paper Series
2/2018

STRUCTURAL VECTOR AUTOREGRESSION WITH TIME VARYING TRANSITION PROBABILITIES:

WENJUAN CHEN
ALEKSEI NETŠUNAJEV

The Working Paper is available on the Eesti Pank web site at:
http://www.eestipank.ee/en/publications/series/working-papers

For information about subscription call: +372 668 0998; Fax: +372 668 0954
e-mail: publications@eestipank.ee

DOI: 10.23656/25045520/022018/0153

ISBN 978-9949-606-27-6 (hard copy)
ISBN I978-9949-606-28-3 (pdf)

Eesti Pank. Working Paper Series, ISSN 1406-7161; 2/2018 (hard copy)
Eesti Pank. Working Paper Series, ISSN 2504-5520; 2/2018 (pdf)

Structural vector autoregression with time varying transition probabilities: identifying uncertainty shocks via changes in volatility

Wenjuan Chen and Aleksei Netšunajev*

Abstract

Structural vector autoregressive models with regime-switching variances have been used to test structural identification strategies. In these models the transition probabilities are assumed to be constant over time. In reality these probabilities may depend on certain economic fundamentals that help predicting turning points. This paper is the first to introduce time varying probabilities into structural VAR model that is identified via volatility. A generalized Expectation-Maximization algorithm is developed for estimation of the model. For empirical illustration the model is applied to test two sets of assumptions used for identification of uncertainty shocks. A formal test rejects the hypothesis that uncertainty shocks do not influence macroeconomics variables on impact but support the alternative of non-negligible contemporaneous effects.

JEL Codes: C32, D80, E24
Keywords: structural vector autoregression; Markov switching; time varying transition probabilities; identification via heteroscedasticity; uncertainty shocks; unemployment dynamics

*Wenjuan Chen: Freie Universitat Berlin. E-mail: wenjuan.chen@fu-berlin.de. Aleksei Netšunajev (corresponding author): Eesti Pank. E-mail: aleksei.netsunajev@eestipank.ee. We thank participants in the 2017 Nordic Econometric Meeting, and participants in the Research Seminar at Free University Economics Department for their helpful comments.

Non-technical summary

In structural vector autoregressive (SVAR) models, it is critical for structural shocks to be identified convincingly, since impulse response analysis could be sensitive to various restrictions that are assumed for identification purposes. For this reason a large number of papers in recent years have heteroscedasticity for identification. SVAR models with Markov regime switching in variances are particularly widely applied. The transition probabilities are assumed to be constant over time in that strand of literature.

This paper is the first to introduce time-varying transition probabilities into Markov-switching structural VAR models that are identified through volatility. We let the transition probabilities depend on various economic fundamentals through a logistic function so that information that help predict turning points can be used. Using the regime switching variances lets us adopt statistical tests to discriminate between competing conventional identification schemes. We estimate the model using maximum likelihood and a flexible EM algorithm.

In the empirical application we use US data and investigate two different types of identification strategy for uncertainty shocks. The first strategy is based on the hypothesis that an uncertainty shock may have contemporaneous effects on macroeconomic variables. The alternative strategy assumes that uncertainty shocks have no contemporaneous impact on the macroeconomic variables.

Our estimation results lead to several new insights. The information criteria of the model show that the Markov-switching model with time varying transition probabilities outperforms the standard model with constant probabilities. It turns out that the seven-quarter moving average of GDP growth is the most preferred transition variable if compared to lagged inflation, federal funds rate and GDP by the means of information criteria. The likelihood ratio test rejects the identification scheme that forces the uncertainty shocks to have no impact on macro variables, but the alternative allowing the uncertainty shock to have contemporaneous effects on macroeconomic variables is supported by the data. This finding demonstrates the power of our method for differentiating between the economic assumptions that are used for identification purposes.

Contents

1 Introduction

In structural vector autoregressive (SVAR) models, it is critical for structural shocks to be identified convincingly, since impulse response analysis could be sensitive to various restrictions that are assumed for identification purposes. Moreover, it is often hard to find economic theories that are able to justify the identifying restrictions. For this reason a large number of papers in recent years have used statistical properties of the data such as heteroscedasticity for identification (Rigobon and Sack (2003), Lanne and Lütkepohl (2008), and Lütkepohl and Netšunajev (2017b)). SVAR models with Markov regime switching in variances are particularly widely applied, as in Lanne, Lütkepohl and Maciejowska (2010) and Herwartz and Lütkepohl (2014), and can be used to test different types of structural identification schemes.

However, it is assumed in the strand of literature that uses Markov-switching in variances for identification that the transition probabilities are constant over time. These probabilities may actually vary over time in reality though, and can depend on some underlying economic fundamentals (see Diebold, Lee and Weinbach (1994), Filardo (1994), and Bazzi, Blasques, Koopman and Lucas (2017)). This paper is the first to address this issue by introducing time-varying transition probabilities into Markov-switching structural VAR models that are identified through volatility. We let the transition probabilities depend on various economic fundamentals through a logistic function so that information from economic fundamentals that help predict turning points can be used. We develop a generalised expectation maximisation (EM) algorithm to estimate the model. Using the regime switching variances lets us adopt statistical tests to discriminate between competing conventional identification schemes, which is in the spirit of Lanne and Lütkepohl (2008) and Herwartz and Lütkepohl (2014).

As an empirical illustration, we investigate two different types of identification strategy for uncertainty shocks in a system in a similar way to Caggiano, Castelnuovo and Groshenny (2014). Starting from the work by Bloom (2009), a growing number of papers have studied the role of uncertainty in the economy (see Alexopoulos and Cohen (2009), Bachmann, Elstner and Sims (2013), Colombo (2013), Nodari (2014), and Baker, Bloom and Davis (2016)). Linear structural VAR models are particularly used in many empirical papers for identifying uncertainty shocks. To the best of our knowledge no studies provide a formal test for differentiating between various assumptions for identifying uncertainty shocks. Our framework provides over-identifying information through changes in variances so as to test different types of identification scheme formally.

Our estimation results shed new light on several issues. Most importantly, the information criteria of the model show that the Markov-switching model with time varying transition probabilities outperforms the standard model with constant probabilities. The choice of the economic fundamental that governs the transition probabilities plays the key role in our analysis and so we estimate models with many alternative candidates, such as lagged unemployment, lagged federal funds rates, and lagged GDP growth rates. Following Auerbach and Gorodnichenko (2012), Bachmann and Sims (2012), Berger and Vavra (2014) and Caggiano et al. (2014), we also consider the moving averages of seven quarter-on-quarter GDP growth rates as a candidate. It turns out that the seven-quarter moving average of GDP growth is the most preferred transition variable according to the information criteria.

Further, our model allows us to test identifying restrictions formally, while in the conventional SVAR setup it is impossible to discriminate between different structural assumptions. The VAR studies on uncertainty shocks, including Caggiano et al. (2014), Alexopoulos and Cohen (2009), and Nodari (2014), typically assume recursive zero restrictions. Two different types of identification strategy are considered in Caggiano et al. (2014). The first specification is based on the hypothesis that an uncertainty shock may have contemporaneous effects on macroeconomic variables. The alternative specification assumes that uncertainty shocks have no contemporaneous impact on the macroeconomic variables. The likelihood ratio test rejects the identification scheme that forces the uncertainty shocks to have no impact on macro variables, but the alternative allowing the uncertainty shock to have contemporaneous effects on macroeconomic variables is supported by the data.

The remainder of the paper is organised as follows. Section 2 sets up the SVAR model with time varying transition probabilities and discusses how it can be estimated and used for identification purposes. The empirical example analysing the relation between economic policy uncertainty and US unemployment is discussed in Section 3. The last section summarises the conclusions from our study.

2 The regime switching model with time varying transition probabilities

2.1 The model setup

Consider the standard VAR model of order p:

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

where y_t is the $K \times 1$ vector of variables of interest, v is the $K \times 1$ intercept terms, $A_i s$ are the $K \times K$ coefficient matrices, and u_t is the vector of reduced form residuals which has zero mean and covariance matrix Σ_u. In order to obtain economically meaningful structural residuals ε_t with zero mean and identity covariance matrix, a linear transformation is commonly used: $u_t = B\varepsilon_t$ or $Au_t = \varepsilon_t$. In the conventional case the identifying restrictions are usually imposed on the matrix B or on its inverse $A = B^{-1}$.

Now let the distribution of u_t depend on a Markov process s_t with M discrete states, $s_t \in 1 \ldots M$. The transition probabilities are usually assumed to be constant over time: $p_{ij} = Pr(s_t = j | s_{t-1} = i)$, but here we allow them to be time varying. Specifically, we follow Diebold et al. (1994) and let the transition probabilities depend on a vector of economic fundamentals x_t and assume that they evolve according to a logistic function. In a simple two-regime case the matrix of transition probabilities P_t is:

$$P_t = \begin{pmatrix} p_t^{11} = \frac{e^{x'_{t-1}\beta_{11}}}{1+e^{x'_{t-1}\beta_{11}}} & p_t^{21} = 1 - p_t^{22} \\ p_t^{12} = 1 - p_t^{11} & p_t^{22} = \frac{e^{x'_{t-1}\beta_{22}}}{1+e^{x'_{t-1}\beta_{22}}} \end{pmatrix}.$$

The superscripts in p_t^{ij} indicate that a switch from regime i to regime j takes place and β_{ij} is a vector of the parameters to be estimated. For the case with three regimes, the transition probability matrix is:

$$\begin{pmatrix} p_t^{11} = \frac{e^{x'_{t-1}\beta_{11}}}{1+e^{x'_{t-1}\beta_{11}}+e^{x'_{t-1}\beta_{12}}} & p_t^{21} = \frac{e^{x'_{t-1}\beta_{21}}}{1+e^{x'_{t-1}\beta_{21}}+e^{x'_{t-1}\beta_{22}}} & p_t^{31} = 1 - p_t^{32} - p_t^{33} \\ p_t^{12} = \frac{e^{x'_{t-1}\beta_{12}}}{1+e^{x'_{t-1}\beta_{11}}+e^{x'_{t-1}\beta_{12}}} & p_t^{22} = \frac{e^{x'_{t-1}\beta_{22}}}{1+e^{x'_{t-1}\beta_{21}}+e^{x'_{t-1}\beta_{22}}} & p_t^{32} = \frac{e^{x'_{t-1}\beta_{32}}}{1+e^{x'_{t-1}\beta_{32}}+e^{x'_{t-1}\beta_{33}}} \\ p_t^{13} = 1 - p_t^{11} - p_t^{12} & p_t^{23} = 1 - p_t^{21} - p_t^{22} & p_t^{33} = \frac{e^{x'_{t-1}\beta_{33}}}{1+e^{x'_{t-1}\beta_{32}}+e^{x'_{t-1}\beta_{33}}} \end{pmatrix}$$

Structural shocks in the model can be identified by the assumption that only the variances of the shocks change across states while impulse responses are not affected, meaning that the instantaneous effects are the same across the states. If there are just two regimes with positive definite covariance matrices Σ_1 Σ_2, it is known that a matrix B exists that satisfies $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2 B'$ where Λ_2 is a diagonal matrix with positive diagonal elements $\Lambda_2 = diag(\lambda_{21}, ...\lambda_{2K})$. Lanne et al. (2010) prove that the matrix B is unique up to changes in sign, given that the diagonal elements of Λ_2 are distinct and ordered in a certain way. Therefore, any restrictions set upon B in a conventional VAR model become over-identifying in our framework.

For the case with more than two regimes, the covariance matrices are decomposed in the same way: $\Sigma_1 = BB'$, $\Sigma_i = B\Lambda_i B'$, $i = 2, ..., M$, where Λ_i are diagonal matrices. The condition for the B to be unique is that if one pair of diagonal elements from Λ_2 are the same, there must be another pair of distinct diagonal elements from some other Λ_i. For example, if $\lambda_{2k} = \lambda_{2l}$, then there must be a pair $\lambda_{ik}, \lambda_{il}$ so that $\lambda_{ik} \neq \lambda_{il}$ from $i = 3, ..., M$. Unfortunately there are no formal tests to see whether the pairwise inequality $\lambda_{ik} \neq \lambda_{il}$ holds in the estimated model. Testing a null hypothesis of no identification $H_0 : \lambda_{21} = \lambda_{22}$ implies that some parameters are not identified under H_0 and standard χ^2 asymptotic properties are not valid (Lütkepohl and Netšunajev, 2017a).

2.2 The estimation

We use maximum likelihood estimation based on a log-likelihood function derived from conditional normality: given the state, the distribution of u_t is assumed to be normal, so $u_t|s_t \sim N(0, \Sigma_{s_t})$. The log likelihood function is highly nonlinear, so numerical optimisation techniques are required. Therefore we adopt the expectation maximisation (EM) algorithm of Herwartz and Lütkepohl (2014), which builds on Diebold et al. (1994) for the actual likelihood optimisation task. The iterative algorithm consists of an expectation step where the estimates of the unobserved regime probabilities are obtained, and a maximisation step where the transition parameters, structural parameters and VAR parameters are estimated.

The expectation step of the algorithm closely follows Kim (1994), Krolzig (1997) and Herwartz and Lütkepohl (2014). In the smoothing part of the expectation step we introduce the filter from Kim (1994), which is not part of the algorithm of Diebold et al. (1994). By doing this we economise on the iterations needed to compute the smoothed regime probabilities that incorporate the information from the full sample.

In the maximisation step the transition parameters, the structural parameters and the VAR parameters are estimated. We add an additional step to the maximisation part of the algorithm of Herwartz and Lütkepohl (2014) to estimate the transition parameters β_{ij}. As the first order conditions of the likelihood function are nonlinear in β_{ij}, we use linear approximation of p_t^{ij} around β_{ij}^{n-1}, which comes from the previous iteration. Consider β_{11} as an example:

$$p_t^{11}(\beta_{11}^n) \approx p_t^{11}(\beta_{11}^{n-1}) + \frac{\partial p_t^{11}(\beta_{11})}{\partial \beta_{11}}\bigg|_{\beta_{11}=\beta_{11}^{n-1}} (\beta_{11} - \beta_{11}^{n-1}).$$

When the linear approximations are further substituted for the probabilities into the first order conditions, the conditions become linear and may be rearranged to give the closed form solution for β_{ij}.

Even though we obtain closed form solutions to the estimate for the transition probabilities, the structural parameters B and $\Lambda_m, m = 2, ..., M$ still have to be estimated by numerical methods. The objective function is nonlinear and can have several local optimums, so we run the estimation over various initial values. With those estimates in hand the VAR parameters of the model are obtained by generalised least squares as in Herwartz and Lütkepohl (2014). The detailed procedure of our algorithm is given in the Appendix.

3 Macroeconomic impact of uncertainty shocks

3.1 The data

We apply our method to a four dimensional VAR and study the effects of uncertainty shocks on the real economy, following Caggiano et al. (2014) closely. Since the seminal paper by Bloom (2009), a growing number of research papers have studied the impact of uncertainty shocks on macroeconomic variables. One strand of the literature has studied the role of uncertainty shocks in dynamic stochastic general equilibrium models. Another strand used VAR models to identify uncertainty shocks and study their effects. This paper contributes to the second strand of literature by extracting information from heteroscedasticity for the identification of uncertainty shocks.

There are various measures of uncertainty, with the most widely used being the CBOE's Volatility Index (VIX), which is an index of 30-day option-implied volatility in the S&P 500 stock index, and the economic policy uncertainty index (EPU), which is based on newspaper coverage frequency. Baker et al. (2016) show that the two measures move closely together, but the EPU index shows stronger responses to political events such as the election of a new president, the September 11 attacks, or political debates over taxes and government spending, while the VIX has a stronger connection to events in financial markets such as the Asian financial crisis. Furthermore, the VIX only covers publicly traded firms, which account for around a third of private employment (see Davis, Haltiwanger, Jarmin and Miranda (2007)), but the EPU index reflects not only stock market volatility but also major political events that affect employment on a national level. Although Caggiano et al. (2014) use the VIX index in their paper, we prefer the EPU measure, given the availability of the data and their relevance for unemployment analysis.

The VAR contains the following vector of variables $y_t = (EPU_t, \pi_t, U_t, FFR_t)'$. The variables are defined in the following way:

- EPU_t stands for the economic policy uncertainty index developed by Baker et al. (2016), which is a proxy for uncertainty;
- the inflation rate π_t is calculated as the quarter-on-quarter percentage growth rate of the implicit GDP deflator;
- U_t is the civilian unemployment rate;
- FFR_t is the federal funds rate.

Quarterly observations of monthly data are constructed by quarterly averaging. The sample runs from 1962Q3 to 2012Q3, as in Caggiano et al. (2014). The source of the EPU index is the website http://www.policyuncertainty.com/, while the other time series are obtained from the FRED database provided by the Federal Reserve Bank of St. Louis.

3.2 Model comparison

The summary statistics for the estimated models are shown in Table 1. We report the log likelihood and Akaike information criterion (AIC) for different types of model, among them a linear VAR model, the Markov-switching model with constant transition probabilities, and the Markov-switching model with time-varying transition probabilities (TVTP). Results are reported for the Markov-switching models for both two-regime and three-regime specifications. The lag order is selected to be three for all models as derived from the AIC of the linear VAR. The choice of the information criterion is based on the simulation study by Luetkepohl and Schlaak (2017), who report the AIC to be slightly advantageous for model selection purposes in heteroscedastic SVARs.

Table 1: Comparison of VAR(3) Models for the period 1962Q3–2012Q3

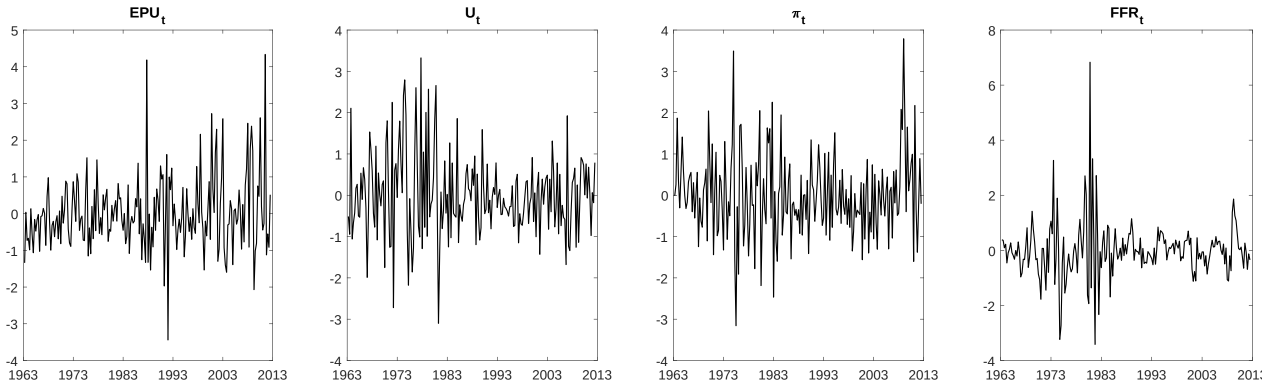
Model	Fundamentals	$\log L_T$	AIC
VAR(3), linear	none	–1408.72	2941.44
MS-SVAR with constant transition probability			
MS, 2 regimes	none	–1273.59	2695.18
MS, 3 regimes	none	–1247.74	2659.49
MS-SVAR with time varying transition probability			
MS-TVTP, 2 regimes	$(1; \overline{\Delta GDP}_{t-1})$	–1268.57	2689.14
MS-TVTP, 3 regimes	$(1; \overline{\Delta GDP}_{t-1})$	–1234.14	2656.29
MS-TVTP, 3 regimes, state invariant B	$(1; \overline{\Delta GDP}_{t-1})$	–1236.67	2649.34
MS-TVTP*, 3 regimes, state invariant B	$(1; \overline{\Delta GDP}_{t-1})$	–1239.79	2651.58

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. $\overline{\Delta GDP}_{t-1}$ is the seven-quarter moving average of GDP growth rates. MS-TVTP* is the model with two extra zero restrictions on the transition parameters, which improves the estimation efficiency and remains very close to the MS-TVTP model in terms of the AIC. Since the MS-TVTP model produces several large standard errors, we use the MS-TVTP* model as the benchmark in the following analysis.

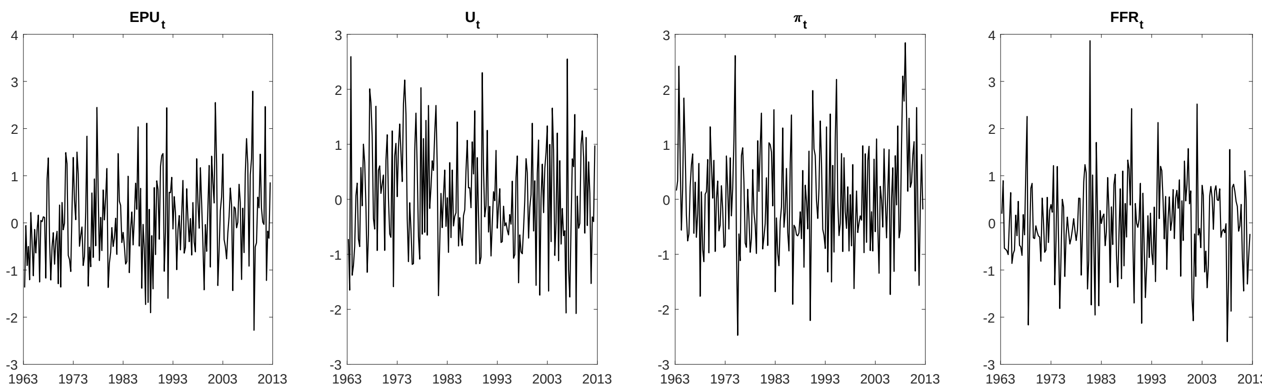
The transition variable plays a key role in our analysis, so we estimate and compare models with various transition variables, including lags of inflation, lags of unemployment, lags of the federal funds rate, and lags of GDP growth rates. We also consider the moving average involving seven realisations of quarter-on-quarter GDP growth rates, following Auerbach and Gorodnichenko (2012), Bachmann and Sims (2012), Berger and Vavra (2014) and Caggiano et al. (2014). Of these transition variables, the moving average of GDP growth rates, denoted as $\overline{\Delta GDP}_{t-1}$, performs best according to the AIC. Though we omit GDP from the VAR system to stay close to Caggiano et al. (2014), it is worth noting that Auerbach and Gorodnichenko (2012) and Bachmann and Sims (2012) include GDP in the VAR part of the model and still find the lagged moving average outperforms the alternative transition variables. We focus on the models with this transition variable in what follows.

Moreover, the choice of the number of regimes is also critical for our analysis. Following Psaradakis and Spagnolo (2006) and Herwartz and Lütkepohl (2014), we use the information criteria as the tool for selecting the number of regimes. Judged by this criterion, the model without any regime switching performs the worst, while the model with time varying transition probabilities with three regimes performs best. This finding is also supported by the plot of the standardised residuals of the linear

VAR model and of the MS-TVTP 3-regime model in Figure 1. The residuals of the model that takes the changing volatility into account are much more regular than those of the standard VAR model.



(a) Residuals of the linear VAR model



(b) Residuals of the Markov switching model with time varying probabilities

Figure 1: Residuals of various models

We have noticed that the MS-TVTP 3-regime model produces very large standard errors in part of the transition parameters, and therefore we estimated different specifications that restrict some of the transition parameters to zero. We found that the MS-TVTP* model with two extra zero restrictions on the transition parameters in the second regime is very close to the MS-TVTP model in its maximum log likelihood and AIC, but improves the estimation efficiency substantially. The p-value of the likelihood ratio test is 0.044, which only marginally indicates support for the restrictions. Nevertheless we take this model as the benchmark model in the analysis and discussion of the robustness of our findings. The estimated parameters of the transition function for the MS-TVTP* 3-regime specification are reported in Table 2. They are estimated quite precisely with the standard errors being mostly smaller than the estimates.

Figure 2 compares the estimated smoothed regime probabilities from two different models. The subfigure on the top displays the regime probabilities estimated for the model that lets the transition probabilities depend on the moving averages of GDP growth rates. The first regime, which is the most volatile one, covers a period in the beginning of the 1970s, the beginning of the 1980s, and

Table 2: Estimated transition parameters and their standard errors

	β_{11}	β_{12}	β_{21}	β_{22}	β_{32}	β_{33}
Estimate(intercept)	5.28	–1.94	–0.53	–3.64	28.83	–18.09
Estimate(slope)	4.68	–3.91	0	0	13.73	–5.77
Std.err.(intercept)	4.56	1.88	3.64	6.50	12.45	7.79
Std.err.(slope)	5.91	9.17	0	0	7.08	3.15

Note: This table reports the estimated transition parameters from the MS-TVTP* 3-regime model.

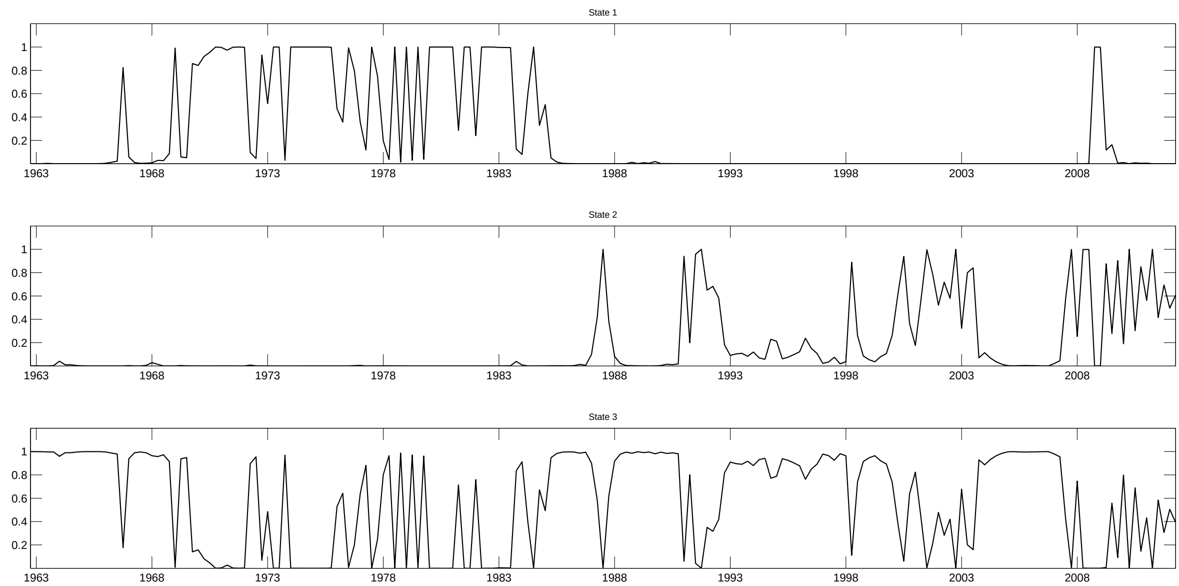
the recent financial crisis. The second regime covers most of the first half of the 2000s and the recovery period following the financial crisis. The third regime, which represents the least volatile regime, covers almost the whole period from 1984 to 2008 with a few exceptions. The timing of this low volatility regime corresponds to the well-known phenomenon named the Great Moderation. The graph at the bottom shows the estimated regime probabilities of the model that assumes constant transition probabilities.

There are noticeable differences when the assumption of constant transition probabilities is imposed. The MS-TVPT* model estimates for example that a short period consequent to the 2008 financial crisis is in the most volatile state. Given the depth of the recession in 2008 it is reasonable to think that it should be in the period of high volatility. In contrast, MS(3) estimates the 2008 financial crisis to be in a relatively calm state. Meanwhile, the whole period from the end of the 1960s to the beginning of the 1980s is estimated by MS(3) to be constantly in the most volatile regime. MS-TVPT* finds though that certain intervals in this period are also relatively calm. This may be intuitive economically as the time between the oil crises in the 1970s may be considered a period of lower volatility.

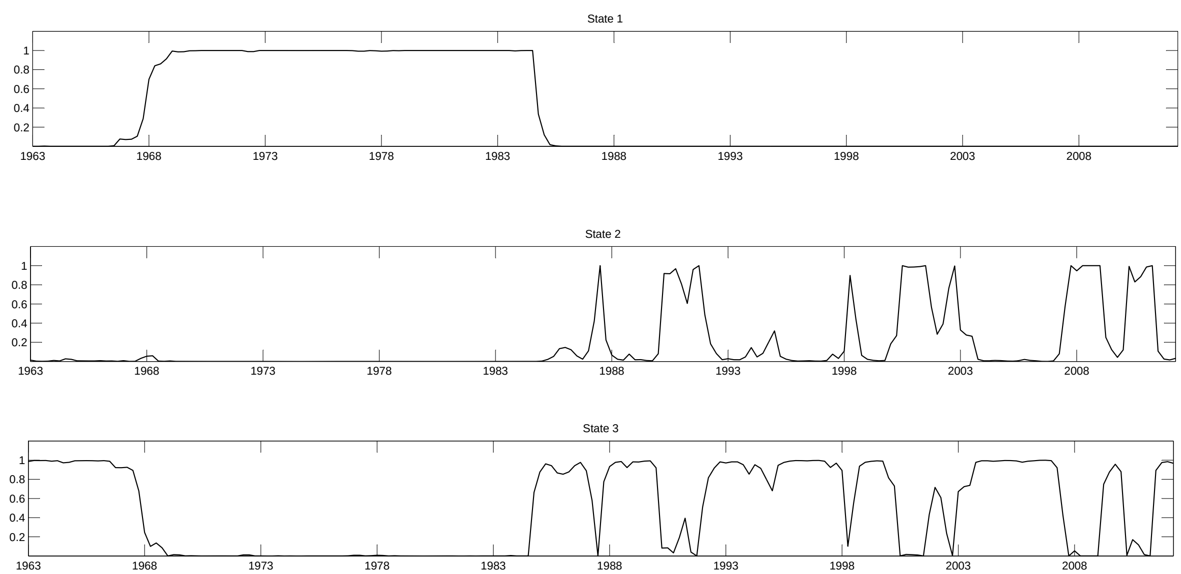
The estimated time varying transition probabilities in Figure 3 provide more evidence for the importance of relaxing the constant transition probability assumption. An example is p_t^{21}, which represents the probability of switching from the medium regime to the most volatile regime. This probability p_t^{21} takes the value of zero for most of the sample, but towards the end of the sample for the period around the 2008 financial crisis it rises to a half, which suggests that it is highly likely that the underlying state will switch from the medium regime to the most volatile regime. If constant transition probabilities are assumed, they would look like a straight horizontal line, and the information contained in the moving averages of the GDP growth rates for the transition probabilities would have been lost.

3.3 Analysis of identification strategies

We next proceed by analysing the structural shocks identified using the model proposed. It is important to check whether the estimated model is identified by at least comparing the pairs of relative variances. The estimates of these parameters along with their standard errors in our preferred model are shown in Table 3. Where no formal tests for identification exist, the standard errors of the variances have to be examined (Lütkepohl and Netšunajev, 2017a). For the preferred three state model with time varying transition probabilities, the estimates of the Λ_2 and Λ_3 matrices are quite precise and heterogeneous with standard errors much lower than the corresponding point estimates. Thus



(a) MS-TVTP* 3-regime model with time varying transition probabilities



(b) MS 3-regime model with constant transition probabilities

Figure 2: A comparison of estimated smoothed regime probabilities

Note: State 1 is the state with the highest volatility. State 2 is the state with medium volatility and State 3 is the one with the lowest volatility.

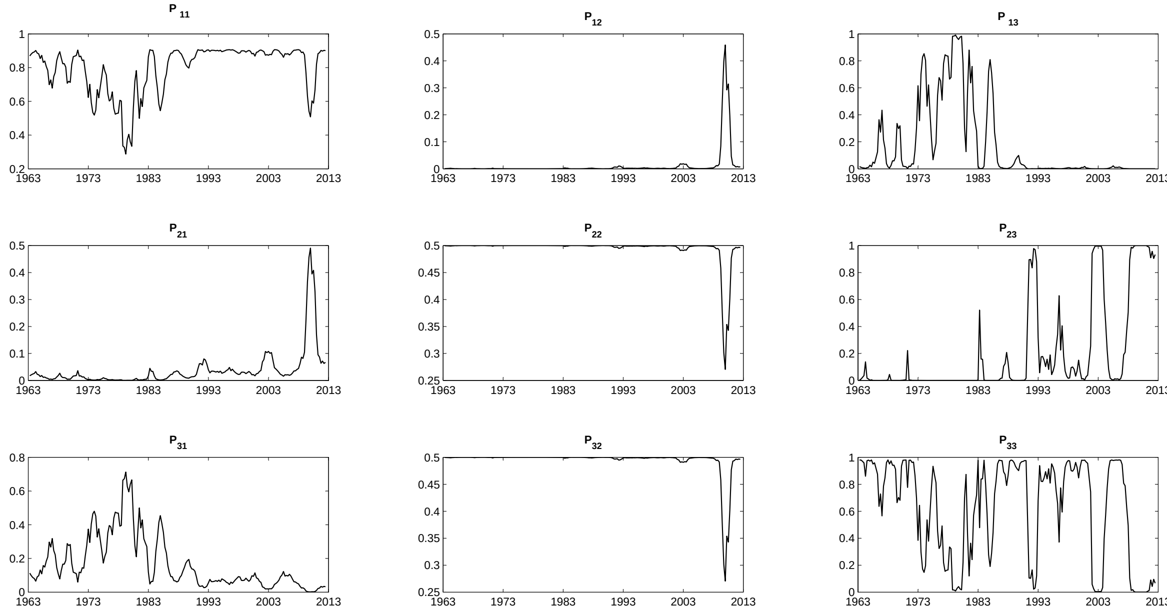


Figure 3: Time varying transition probabilities

there are good reasons to believe that we have additional information on volatility and that the structural matrix B is well identified and the tests for restrictions have power.

The two types of Cholesky ordering are to be tested next, as in Caggiano et al. (2014). The first ordering, $y_t = (EPU_t, \pi_t, U_t, FFR_t)'$, which assumes that the uncertainty shocks contemporaneously affect macroeconomic variables, fits the economic intuition well. However, even though the second ordering counter-intuitively assumes that uncertainty shocks cannot have impact effects on macroeconomic indicators on impact, it is shown by Caggiano et al. (2014) that the responses of unemployment to an uncertainty shock under the second ordering look very similar to those under the first ordering.

Table 3: Estimated relative variances of the MS(3) model with state invariant B

Parameter	Estimate	Standard error
λ_{21}	0.09	0.10
λ_{22}	8.75	5.75
λ_{23}	0.30	0.19
λ_{24}	0.03	0.03
λ_{31}	0.20	0.10
λ_{32}	0.94	0.54
λ_{33}	0.16	0.10
λ_{34}	0.03	0.01

Now that we have additional information from the changes in volatility, we could test whether either of the two Cholesky orderings can be confirmed by the data. The results of testing are presented in Table 4. Under the assumption of two regimes, neither Cholesky ordering can be rejected, but this is not the case for the three regime specifications. With three regimes, we first have to test the structural covariance matrix decomposition where H_0: $\Sigma_1 = BB'$, $\Sigma_2 = B\Lambda_2 B'$, $\Sigma_3 = B\Lambda_3 B'$, and an alternative where the covariance matrices are fully unrestricted. We do not reject the decomposition for MS-TVPT* received from the LR test with six degrees of freedom and a value of 5.06 and $p = 0.53^*$. Thus we proceed by testing the Cholesky orderings of interest.

Given three regimes, the identification scheme B_1 in Table 4 assuming that the uncertainty shocks have contemporaneous effects on macroeconomic variables cannot be rejected with $p = 0.57$. On the contrary, the Cholesky ordering B_2 assuming that uncertainty shocks cannot have instantaneous impact effects on macroeconomic variables is rejected with $p = 0.04^\dagger$. These results indicate that the model with only two regimes may lack information that is needed for identification. The model with three regimes captures the pattern of volatility better, and thus the estimated variances of structural shocks are much more distinct. The LR test is more powerful for the case with three regimes.

Table 4: Tests for different identification schemes

H_0	H_1	df	LR statistic	p-value
B_1	MS-TVTP, 2 regime	6	6.34	0.39
B_2	MS-TVTP, 2 regime	6	10.02	0.12
B_1	MS-TVTP*, 3 regime, state invariant B	6	4.82	0.57
B_2	MS-TVTP*, 3 regime, state invariant B	6	13.04	0.04

Note: The identification strategy B_1 stands for the six zero restrictions on the upper triangular part of the B matrix in the spirit of the Cholesky decomposition with the variables ordered as $y_t = (EPU_t, \pi_t, U_t, FFR_t)'$. The identification strategy B_2 represents the six zero restrictions on B when the EPU index is ordered as the last variable in the VAR system $y_t = (\pi_t, U_t, FFR_t, EPU_t)'$. The models under H_1 impose no identifying restrictions, while the models under H_0 are restricted with the identification scheme B_1 or B_2.

The existing literature did not take a clear stand on whether uncertainty shocks should be thought of as affecting the macroeconomic variables on impact or not. While many researchers assumed the uncertainty shock affected the economy on impact and based their analysis on such assumptions, no formal argument was proposed. Our analysis exploiting information from changes in volatility is able to distinguish between the two hypotheses. We find support for the identification scheme that uses the non-negligible contemporaneous effects of uncertainty shocks on macroeconomic variables.

[*] The LR test leads to the same conclusion for the MS-TVPT model.

[†] For the MS-TVPT model, the Cholesky ordering B_2 is also rejected at the 5% significance level.

4 Conclusions

In the paper we propose a structural vector autoregressive model where the changes in volatility are governed by a Markov process with time varying transition probabilities. Time varying transition probabilities are assumed to depend on fundamental economic variables. The structural parameters of the model are identified with changes in the volatility of shocks. Additional information that comes from the time variation in the variances of structural shocks allows conventional identifying restrictions to be tested. We estimate the model using maximum likelihood and a flexible EM algorithm.

In the empirical illustration of our model, we apply this method to identify uncertainty shocks following the study by Caggiano et al. (2014). Based on the information criteria our model with time varying transition probabilities fits the example data better than a standard Markov-switching model like that in Lanne et al. (2010), which assumes constant transition probabilities. This is most likely due to the useful information contained in the transition variable, which in our case is the moving average of seven quarter-on-quarter GDP growth rates.

With extra information extracted from changes in variances we test the two different types of identification strategy used in Caggiano et al. (2014). Using the preferred three regime MS-TVTP model, we reject the identification strategy that restricts the uncertainty shocks to have no contemporaneous effects on macroeconomic variables. However, we do not reject the alternative identification strategy that allows for these contemporaneous effects to be present. This finding demonstrates the power of our method for differentiating between the economic assumptions that are used for identification purposes.

A Appendix. Estimation of the MS-SVAR model with time-varying transition probabilities

The section describes in detail the expectation maximization (EM) algorithm based on Krolzig (1997), Herwartz and Lütkepohl (2014) and Diebold et al. (1994), and presents the estimation procedure for structural VAR model with changes in volatility of shocks where the transition probability matrix is also allowed to vary over time.

Definitions

The baseline model is the VAR(p) of the form:

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

Denote:
M - number of states, assumed to be three in this appendix,
K - number of variables in the vector y,
p - number of lags.

Let the matrix $X = [x_0, x_1, ..., x_T]$ contain transition variables up to T with the entries for specific t given by a $(J+1) \times 1$ vector x_t of J economic variables that affect the transition probabilities and a leading one for a constant.

Define $\xi_t = \begin{bmatrix} I(s_t = 1) \\ \vdots \\ I(s_t = M) \end{bmatrix}$, then $E(\xi_t) = \begin{bmatrix} Pr(s_t = 1) \\ \vdots \\ Pr(s_t = M) \end{bmatrix}$, where $I()$ is an indicator function which takes value 1 if statement in the argument is true and 0 otherwise.

Further define

$$\xi_{t|s} = E(\xi_t | Y_s, X_s) = \begin{bmatrix} Pr(s_t = 1 | Y_s, X_s) \\ \vdots \\ Pr(s_t = M | Y_s, X_s) \end{bmatrix}$$, where $Y_s = (y_1, ..., y_s), X_s = (x_0, ..., x_s)$

Next:

$$\xi^{(2)}_{t|T} = \begin{bmatrix} Pr(s_t = 1 | s_{t-1} = 1, Y_T, X) \\ \vdots \\ Pr(s_t = 1 | s_{t-1} = M, Y_T, X) \\ \vdots \\ Pr(s_t = M | s_{t-1} = 1, Y_T, X) \\ \vdots \\ Pr(s_t = M | s_{t-1} = M, Y_T, X) \end{bmatrix}.$$

We let the transition probability matrix to be time varying for M state Markov process. Define P_t as the time varying transition matrix, which yields $\xi_{t+1|t} = P_t \xi_{t|t}$, for $t = 0, 1, ..., T-1$. Advancing on Diebold et al. (1994) we shows the closed form solutions for estimating models with $M = 2$ and $M = 3$ as these appear to be the most important in practice. Expressions for models with $M > 3$ may be derived analogously. The individual elements of the P_t matrix evolve as logistic functions of $x'_{t-1}\beta_{ij}$. Then β_{ij} is the $(J + 1) \times 1$ vector of parameters. It is convenient to collect the individual β_{ij} vectors into a matrix $\beta = [\beta_{11}\ \beta_{22}]$ for 2 regimes and $\beta = \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{32} \\ \beta_{12} & \beta_{22} & \beta_{33} \end{bmatrix}$ for 3 regimes. The matrix β_0 denotes the initial values for the transition parameters. The matrix P_t is defined as:

$$P_t = \begin{pmatrix} Pr(s_{t+1} = 1|s_t = 1, \beta, x_{t-1}) & \dots & Pr(s_{t+1} = 1|s_t = M, \beta, x_{t-1}) \\ \vdots & \ddots & \vdots \\ Pr(s_{t+1} = M|s_t = 1, \beta, x_{t-1}) & \dots & Pr(s_{t+1} = M|s_t = M, \beta, x_{t-1}) \end{pmatrix}$$

The following matrices illustrate the details. Note that the subscripts for β_{ij} and superscripts for p_t^{ij} denote transition from state i to state j. Transition probability matrix for $M = 2$:

$$\begin{pmatrix} p_t^{11} = \frac{e^{x'_{t-1}\beta_{11}}}{1+e^{x'_{t-1}\beta_{11}}} & p_t^{21} = 1 - p_t^{22} \\ p_t^{12} = 1 - p_t^{11} & p_t^{22} = \frac{e^{x'_{t-1}\beta_{22}}}{1+e^{x'_{t-1}\beta_{22}}} \end{pmatrix}$$

Transition probability matrix for $M = 3$:

$$\begin{pmatrix} p_t^{11} = \frac{e^{x'_{t-1}\beta_{11}}}{1+e^{x'_{t-1}\beta_{11}}+e^{x'_{t-1}\beta_{12}}} & p_t^{21} = \frac{e^{x'_{t-1}\beta_{21}}}{1+e^{x'_{t-1}\beta_{21}}+e^{x'_{t-1}\beta_{22}}} & p_t^{31} = 1 - p_t^{32} - p_t^{33} \\ p_t^{12} = \frac{e^{x'_{t-1}\beta_{21}}}{1+e^{x'_{t-1}\beta_{11}}+e^{x'_{t-1}\beta_{12}}} & p_t^{22} = \frac{e^{x'_{t-1}\beta_{22}}}{1+e^{x'_{t-1}\beta_{21}}+e^{x'_{t-1}\beta_{22}}} & p_t^{32} = \frac{e^{x'_{t-1}\beta_{32}}}{1+e^{x'_{t-1}\beta_{32}}+e^{x'_{t-1}\beta_{33}}} \\ p_t^{13} = 1 - p_t^{11} - p_t^{12} & p_t^{23} = 1 - p_t^{21} - p_t^{22} & p_t^{33} = \frac{e^{x'_{t-1}\beta_{33}}}{1+e^{x'_{t-1}\beta_{32}}+e^{x'_{t-1}\beta_{33}}} \end{pmatrix}$$

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

where $f()$ is conditional distribution function:

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

Covariance matrices have decomposition as previously described: $\Sigma_1 = BB', \Sigma_m = B\Lambda_m B'$ for $m = 2, ..., M$

Further the following notation is used:
$\odot$ elementwise multiplication,
$\oslash$ elementwise devision,
$\otimes$ Kronecker product,
I_K is a $K \times K$ dimensional identity matrix,
$1_M = (1, ..., 1)'$ is a $M \times 1$ dimensional vector of ones,
$\theta = \text{vec}(v, A_1, A_2, ..., A_P)$ is the vector of VAR coefficients
$Z'_{t-1} = (1, y'_{t-1}, y'_{t-2}, ..., y'_{t-p})$ is the matrix of ones and lagged observations.

Initial values

The following starting values are used for the iterations:

P_t is calculated for given X and β_0 for $1, ..., T$.

$\hat{\theta} = \text{vec}(\hat{v}, \hat{A}_1, ..., \hat{A}_p) = \left[\sum_{t=1}^{T} Z_{t-1}Z'_{t-1} \otimes I_K\right]^{-1} \sum_{t=1}^{T} (Z_{t-1} \otimes I_K)y_t$

$B = T^{-1}(\sum_{t=1}^{T} \hat{u}_t\hat{u}'_t)^{1/2} + B_0$, where $\hat{u}_t = y_t - (Z'_{t-1} \otimes I_K)\hat{\theta}$

and B_0 is a matrix of random numbers coming form standard normal distribution and scaled by a factor of 10^{-5}.

$\Lambda_1 = I_K, \Lambda_m = c_m I_K, m = 2, ..., M$

$\xi_{0|0} = M^{-1}1_M$

Expectation step

For given $P_t, \theta, \Sigma_m, m = 1, 2, ..., M$ and $\xi_0 = \xi_{0|0}$ the following parameters are computed:

η_t for $t = 1, 2, ..., T$,

$\xi_{t|t} = \frac{\eta_t \odot P_t \xi_{t-1|t-1}}{1'_M(\eta_t \odot P_t \xi_{t-1|t-1})}$, for $t = 1, 2, ..., T$.

$\xi_{t|T} = (P'_t(\xi_{t+1|T} \oslash P_t\xi_{t|t})) \odot \xi_{t|t}$, for $t = T-1, ..., 0$.

$\xi^{(2)}_{t|T} = \text{vec}(P'_t) \odot ((\xi_{t+1|T} \oslash P_t\xi_{t|t}) \otimes \xi_{t|t})$, for $t = 1, ..., T-1$.

Maximization step

Estimation of transition parameters β

Given the smoothed probabilities, the expected complete-data log likelihood are non-linear in the β transition parameters. Taking into account the logistic transition function, the first order conditions for β are given as follows:

M = 2:

$\sum_{t=2}^{T} x_{t-1} \{Pr(s_t = 1|s_{t-1} = 1, Y_T, X) - p_t^{11} Pr(s_{t-1} = 1|Y_T, X)\} = 0,$

$\sum_{t=2}^{T} x_{t-1} \{Pr(s_t = 2|s_{t-1} = 2, Y_T, X) - p_t^{22} Pr(s_{t-1} = 2|Y_T, X)\} = 0,$

M = 3:

$\sum_{t=2}^{T} x_{t-1} \{Pr(s_t = 1|s_{t-1} = 1, Y_T, X) - p_t^{11} Pr(s_{t-1} = 1|Y_T, X)\} = 0,$

$\sum_{t=2}^{T} x_{t-1} \{Pr(s_t = 2|s_{t-1} = 1, Y_T, X) - p_t^{12} Pr(s_{t-1} = 1|Y_T, X)\} = 0,$

$\sum_{t=2}^{T} x_{t-1}\left\{Pr(s_t = 1|s_{t-1} = 2, Y_T, X) - p_t^{21} Pr(s_{t-1} = 2|Y_T, X)\right\} = 0,$
$\sum_{t=2}^{T} x_{t-1}\left\{Pr(s_t = 2|s_{t-1} = 2, Y_T, X) - p_t^{22} Pr(s_{t-1} = 2|Y_T, X)\right\} = 0,$
$\sum_{t=2}^{T} x_{t-1}\left\{Pr(s_t = 3|s_{t-1} = 2, Y_T, X) - p_t^{32} Pr(s_{t-1} = 3|Y_T, X)\right\} = 0,$
$\sum_{t=2}^{T} x_{t-1}\left\{Pr(s_t = 3|s_{t-1} = 3, Y_T, X) - p_t^{33} Pr(s_{t-1} = 3|Y_T, X)\right\} = 0.$

Using the following Taylor approximation of the elements of P_t matrix, we find the closed-form solution for all β vectors. Consider β_{11} as an example:
$p_t^{11}(\beta_{11}^n) \approx p_t^{11}(\beta_{11}^{n-1}) + \frac{\partial p_t^{11}(\beta_{11})}{\partial \beta_{11}}\Big|_{\beta_{11}=\beta_{11}^{n-1}} (\beta_{11} - \beta_{11}^{n-1})$

where β_{11}^{n-1} is the β_{11} coming from the previous iteration of the algorithm. The closed-form solutions for β are given as follows.

$\underline{M = 2}$:

$$\beta_{11} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{1t}^{11} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{Jt}^{11} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{1t}^{11} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{Jt}^{11} \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t = 1|s_{t-1} = 1, Y_T, X) - Pr(s_{t-1} = 1|Y_T, X)[p_t^{11} - \beta_{11}^{j-1} \frac{\partial P_t^{11}}{\partial \beta_{11}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t = 1|s_{t-1} = 1, Y_T, X) - Pr(s_{t-1} = 1|Y_T, X)[p_t^{11} - \beta_{11}^{j-1} \frac{\partial P_t^{11}}{\partial \beta_{11}}] \right\} \end{pmatrix}$$

$$\beta_{22} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1} = 2|Y_T, X) p_{1t}^{22} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1} = 2|Y_T, X) p_{Jt}^{22} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1} = 2|Y_T, X) p_{1t}^{22} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1} = 2|Y_T, X) p_{Jt}^{22} \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t = 2|s_{t-1} = 2, Y_T, X) - Pr(s_{t-1} = 2|Y_T, X)[p_t^{22} - \beta_{22}^{j-1} \frac{\partial P_t^{22}}{\partial \beta_{22}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t = 2|s_{t-1} = 2, Y_T, X) - Pr(s_{t-1} = 2|Y_T, X)[p_t^{22} - \beta_{22}^{j-1} \frac{\partial P_t^{22}}{\partial \beta_{22}}] \right\} \end{pmatrix}$$

$\underline{M = 3}$:

$$\beta_{11} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{1t}^{11} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{Jt}^{11} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{1t}^{11} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1} = 1|Y_T, X) p_{Jt}^{11} \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t = 1|s_{t-1} = 1, Y_T, X) - Pr(s_{t-1} = 1|Y_T, X)[p_t^{11} - \beta_{11}^{j-1} \frac{\partial P_t^{11}}{\partial \beta_{11}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t = 1|s_{t-1} = 1, Y_T, X) - Pr(s_{t-1} = 1|Y_T, X)[p_t^{11} - \beta_{11}^{j-1} \frac{\partial P_t^{11}}{\partial \beta_{11}}] \right\} \end{pmatrix}$$

$$\beta_{12} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=1|Y_T,X) p_{1t}^{12} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=1|Y_T,X) p_{Jt}^{12} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=1|Y_T,X) p_{1t}^{12} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=1|Y_T,X) p_{Jt}^{12} \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t=2|s_{t-1}=1,Y_T,X) - Pr(s_{t-1}=1|Y_T,X)[p_t^{12} - \beta_{12}^{j-1} \frac{\partial P_t^{12}}{\partial \beta_{12}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t=2|s_{t-1}=1,Y_T,X) - Pr(s_{t-1}=1|Y_T,X)[p_t^{12} - \beta_{12}^{j-1} \frac{\partial P_t^{12}}{\partial \beta_{12}}] \right\} \end{pmatrix}$$

$$\beta_{21} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=2|Y_T,X) p_{1t}^{21} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=2|Y_T,X) p_{Jt}^{21} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=2|Y_T,X) p_{1t}^{21} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=2|Y_T,X) p_{Jt}^{21} \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t=1|s_{t-1}=2,Y_T,X) - Pr(s_{t-1}=2|Y_T,X)[p_t^{21} - \beta_{21}^{j-1} \frac{\partial P_t^{21}}{\partial \beta_{21}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t=1|s_{t-1}=2,Y_T,X) - Pr(s_{t-1}=2|Y_T,X)[p_t^{21} - \beta_{21}^{j-1} \frac{\partial P_t^{21}}{\partial \beta_{21}}] \right\} \end{pmatrix}$$

$$\beta_{22} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=2|Y_T,X) p_{1t}^{22} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=2|Y_T,X) p_{Jt}^{22} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=2|Y_T,X) p_{1t}^{22} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=2|Y_T,X) p_{Jt}^{22} \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t=2|s_{t-1}=2,Y_T,X) - Pr(s_{t-1}=2|Y_T,X)[p_t^{22} - \beta_{22}^{j-1} \frac{\partial P_t^{22}}{\partial \beta_{22}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t=2|s_{t-1}=2,Y_T,X) - Pr(s_{t-1}=2|Y_T,X)[p_t^{22} - \beta_{22}^{j-1} \frac{\partial P_t^{22}}{\partial \beta_{22}}] \right\} \end{pmatrix}$$

$$\beta_{32} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=3|Y_T,X) p_{1t}^{32} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=3|Y_T,X) p_{Jt}^{32} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=3|Y_T,X) p_{1t}^{32} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=3|Y_T,X) p_{Jt}^{32} \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t=2|s_{t-1}=3,Y_T,X) - Pr(s_{t-1}=3|Y_T,X)[p_t^{32} - \beta_{32}^{j-1} \frac{\partial P_t^{32}}{\partial \beta_{32}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t=2|s_{t-1}=3,Y_T,X) - Pr(s_{t-1}=3|Y_T,X)[p_t^{32} - \beta_{32}^{j-1} \frac{\partial P_t^{32}}{\partial \beta_{32}}] \right\} \end{pmatrix}$$

$$\beta_{33} = \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=3|Y_T,X) p_{1t}^{33} & \dots & \sum_{t=2}^{T} x_{1,t-1} Pr(s_{t-1}=3|Y_T,X) p_{Jt}^{33} \\ \vdots & \ddots & \vdots \\ \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=3|Y_T,X) p_{1t}^{33} & \dots & \sum_{t=2}^{T} x_{J,t-1} Pr(s_{t-1}=3|Y_T,X) p_{Jt}^{33} \end{pmatrix}^{-1}$$

$$\times \begin{pmatrix} \sum_{t=2}^{T} x_{1,t-1} \left\{ Pr(s_t=3|s_{t-1}=3,Y_T,X) - Pr(s_{t-1}=3|Y_T,X)[p_t^{33} - \beta_{33}^{j-1} \frac{\partial P_t^{33}}{\partial \beta_{33}}] \right\} \\ \vdots \\ \sum_{t=2}^{T} x_{J,t-1} \left\{ Pr(s_t=3|s_{t-1}=3,Y_T,X) - Pr(s_{t-1}=3|Y_T,X)[p_t^{33} - \beta_{33}^{j-1} \frac{\partial P_t^{33}}{\partial \beta_{33}}] \right\} \end{pmatrix}$$

where $p_{1t}^{ii}, \cdots, p_{Jt}^{ii}$ are denoting the elements in the vector of partial derivatives as used in the Taylor approximation.

Estimation of structural parameters B and Λ_m :

Define $T_m = \sum_{t=1}^{T} \xi_{mt|T}$, where $\xi_{mt|T}$ denotes the m-th element of the vector $\xi_{t|T}$. Estimation of B and Λ_m is done by minimizing the likelihood function:

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

Then compute:
$\hat{\Sigma}_1 = \hat{B}\hat{B}', \hat{\Sigma}_m = \hat{B}\hat{\Lambda}_m\hat{B}'$ for $m = 2, ..., M$

Estimation of VAR parameters:

Estimates of the parameter vector θ are given by:
$$\hat{\theta} = \left[\sum_{m=1}^{M}\left(\sum_{t=1}^{T}\xi_{mt|T}Z_{t-1}Z_{t-1}'\right)\otimes\hat{\Sigma}_m^{-1}\right]^{-1}\sum_{t=1}^{T}\left(\sum_{m=1}^{M}\xi_{mt|T}Z_{t-1}\otimes\hat{\Sigma}_t^{-1}\right)y_t$$

Initial regime probabilities are updated according to:
$\xi_{0|0} = \xi_{0|T}$

Convergence Criteria

Relative change in the value of the log-likelihood function is used as convergence criteria. The log-likelihood is evaluated for given $P_t, \theta, \Sigma_m, m = 1, 2, ..., M$ and $\xi_{0|0}$ in the end of the expectation step. Given:

η_t for $t = 1, 2, ..., T$,

$\xi_{t|t-1} = P_t \xi_{t-1|t-1}$, for $t = 1, 2, ..., T$,

$\xi_{t|t} = \frac{\eta_t \odot P_t \xi_{t|t-1}}{1'_M (\eta_t \odot P_t \xi_{t|t-1})}$, for $t = 1, 2, ..., T$.

The log likelihood is:

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

$f(y_t | Y_{t-1}) = \xi'_{t|t-1} \eta_t$.

Estimation of β, B, Λ_m and θ are iterated until convergence, i.e. relative change Δ in the log-likelihood is negligibly small (does not exceed tolerance value $\alpha = 10^{-9}$) for k-th and $(k-1)$-th rounds of iterations:

$\Delta = \frac{\log L_T(k) - \log L_T(k-1)}{\log L_T(k-1)} < \alpha$.

References

Alexopoulos, M. and Cohen, J. (2009). Uncertainty times, uncertainty measures, *Technical report*, University of Toronto, Department of Economics, Working Paper No. 325.

Auerbach, A. and Gorodnichenko, Y. (2012). Measuring the output responses of fiscal policy, *American Econmic Journal: Economic Policy* **4**(2): 1–27.

Bachmann, R., Elstner, S. and Sims, E. (2013). Uncertainty and economic activity: evidence from business survey data, *American Economic Journal: Economic Policy* **4**(2): 1–27.

Bachmann, R. and Sims, E. (2012). Confidence and the transmission of government spending shocks, *Journal of Monetary Economics* **59**: 235–249.

Baker, S., Bloom, N. and Davis, S. (2016). Measuring economic policy uncertainty, *Technical report*, Northwestern University, Standford University, and University of Chicago.

Bazzi, M., Blasques, F., Koopman, S. J. and Lucas, A. (2017). Time varying transition probabilities for Markov regime switching models, *Journal of Time Series Analysis* **38**: 458–478.

Berger, D. and Vavra, J. (2014). Measuring how fiscal shocks affect durable spending in recessions and expansions, *American Economic Review Papers and Proceedings* **104**(5): 112–115.

Bloom, N. (2009). The impact of uncertainty shocks, *Econometrica* **77**(3): 623–685.

Caggiano, G., Castelnuovo, E. and Groshenny, N. (2014). Uncertainty shocks and unemployment dynamics in US recessions, *Journal of Monetary Economics* **67**: 78–92.

Colombo, V. (2013). Economic policy uncertainty in the US: does it matter for the euro area?, *Economics Letters* **121**(1): 39–42.

Davis, S., Haltiwanger, J., Jarmin, R. and Miranda, J. (2007). Volatility and dispersion in business growth rates: Publicly traded versus privately held firms, *NBER Macroeconomics Annual* **21**: 107–180.

Diebold, F., Lee, J.-H. and Weinbach, G. (1994). Regime switching with time varying transition probabilities, *Business Cycles: Durations, Dynamics, and Forecasting* pp. 144–165.

Filardo, A. (1994). Business-cycle phases and their transitional dynamics, *Journal of Business and Economic Statistics* **12**(3): 299–308.

Herwartz, H. and Lütkepohl, H. (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks, *Journal of Econometrics* **183**(1): 104–116.

Kim, C.-J. (1994). Dynamic linear models with Markov-switching, *Journal of Econometrics* **60**(1): 1–22.

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**(6): 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.

Luetkepohl, H. and Schlaak, T. (2017). Choosing between Different Time-Varying Volatility Models for Structural Vector Autoregressive Analysis, *Discussion Papers of DIW Berlin 1672*, DIW Berlin, German Institute for Economic Research.

Lütkepohl, H. and Netšunajev, A. (2017a). Structural vector autoregressions with heteroskedasticity: A review of different volatility models, *Econometrics and Statistics* **1**: 2 – 18.

Lütkepohl, H. and Netšunajev, A. (2017b). Structural Vector Autoregressions with Smooth Transition in Variances - The Interaction Between US Monetary Policy and the Stock Market, *Journal of Economic Dynamics and Control* **84**: 43–57.

Nodari, G. (2014). Financial regulation policy uncertainty and credit spreads in the US, *Journal of Macroeconomics.* **41**: 122–132.

Psaradakis, Z. and Spagnolo, N. (2006). Joint determination of the state dimension and autoregressive order for models with Markov regime switching, *Journal of Time Series Analysis* **27**(5): 753–766.

Rigobon, R. and Sack, B. (2003). Measuring the reaction of monetary policy to the stock market., *Quarterly Journal of Economics* **118**(2).

Working Papers of Eesti Pank 2018

No 1
Sang-Wook (Stanley), Cho Julián P. Daz. Skill premium divergence: the roles of trade, capital and demographics