DEPARTMENT OF ECONOMICS

EUI Working Papers

ECO 2012/13
DEPARTMENT OF ECONOMICS

REACTION TO TECHNOLOGY SHOCKS IN MARKOV-SWITCHINGS STRUCTURAL VARs: IDENTIFICATION VIA HETEROSKEDASTIOCITY

Aleksei Netšunajev

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
DEPARTMENT OF ECONOMICS

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

ALEKSEI NETŠUNAJEV

EUI Working Paper **ECO** 2012/13

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

March 26, 2012

Reaction to Technology Shocks in Markov-switchings Structural VARs: Identification via Heteroskedasticity [1]

by

Aleksei Netšunajev
Department of Economics, European University Institute
Via della Piazzola 43, I-50133 Firenze, ITALY
E-mail: aleksei.netsunajev@eui.eu

Abstract. The paper reconsiders the conflicting results in the debate connected to the effects of technology shocks on hours worked in the bivariate system. Given major dissatisfaction with the just-identifying long-run restrictions, I analyze whether the restrictions used in the literature are consistent with the data. Modeling volatility of shocks using Markov switching structure allows to obtain additional identifying information and perform tests of the restrictions that were just-identifying in classical structural vector autoregression analysis. Using four datasets where hours worked are modeled differently, I find that the standard restriction, identifying the technology shocks as the only sources of variation in labor productivity, has major support by the data. Taking into account important low frequency movements in the hours worked series yields a result consistent with the recent findings: hours decline in response to a positive technology shock. I also show that the use of a standard Hodrick-Prescott filter may be problematic in the context.

Key Words: Technology shocks, Markov switching model, heteroskedasticity

JEL classification: C32

[1] I thank participants of the research seminar at the Bank of Estonia for useful comments on an earlier version of the paper. The research for this paper was carried out in part while the author was a vising researcher at DIW Berlin

1 Introduction

A standard real business cycle model implies that per capita hours worked rise after a permanent shock to technology. This prediction is at the center of the literature that assessing if 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 the 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 the question of interest, but the most common is based on the structural vector autoregressive (SVAR) models. In a seminal paper, Gali (1999) identifies the technology shocks using long-run restrictions and find 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. 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 identified, as in Gali (1999), can be classified as technology shocks. Using different identifying assumptions they find that all but one specification produced the result similar to Gali (1999). 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 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 Hodrick-Prescott filter to the data (Francis and Ramey, 2009), the response of hours worked to a technology shock becomes negative. Francis and Ramey (2009) argue that the low frequency movements are connected to sectoral shifts in hours and changes of the age composition of the work force.

It should be noted, that the studies listed above may share some common shortcomings. First, the underlying assumptions just-identify the technology shocks and leave no place for the data to speak up against the restrictions. The problem of just-identified shocks is discussed, among others, by Lanne and Lütkepohl (2008), Lanne, Lütkepohl and Maciejowska (2010), Herwartz and Lütkepohl (2011). Second, studies of technology shocks (for example, Gali (1999), Francis and Ramey (2005), Christiano et al. (2003)) 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 take into account heteroskedasticity as it allows to extract additional identifying information from the data (Rigobon, 2003). Therefore modeling heteroskedasticity can be used as a tool of validating the restrictions that are just identifying in a conventional SVAR analysis.

Thus the aim of the current paper is to reconsider the reaction to technology shocks and relax some of the assumptions common in this literature. For that purpose I estimate a series of Markov-switching (MS) models that allow to capture the changes in volatility and intercept, provide a framework to test for the validity of the identifying restrictions and assess 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. First, to provide additional motivation to the paper, different identification schemes of technology shocks are discussed in Section 2. Then the structural MS-VAR model deployed in the current analysis is described in Section 3. The data are discussed in Section 4. Section 5 provides the empirical analysis. The last section concludes.

2 Identification of shocks

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

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

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 get unique structural shocks one needs to place $K(K-1)/2$ restrictions. For that reason the B matrix is often assumed to be lower triangular. Then the B is the matrix of instantaneous effects of the unique structural shocks.

In the related technology shock literature one usually considers a bivariate system in the spirit of Gali (1999). Using long run restrictions one identifies 2 shocks: technology shocks and non-technology shocks. The shocks are identified in the following system, which is a moving average representation of a VAR:

$$\begin{bmatrix} \Delta x_t \\ \Delta n_t \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_t^z \\ \varepsilon_t^m \end{bmatrix} \tag{2}$$

where x_t denotes the log of labor productivity, n_t denotes the log of labor 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 (2). It should be noted, that the restritions on Ξ_∞ can be transformed to the restrictions on the B as shown in Lütkepohl (2005).

Most common identifying assumption restricts $C_{12}(1) = 0$ implying that only technology shocks have long run effect on the labor productivity (Gali, 1999). The non-technology shocks could be interpreted then as demand shocks (Gali, 1999). Francis and Ramey (2005) discuss that a similar reaction of variables could be expected due to permanent changes in capital income taxation. Expanding the bivariate system and controlling for capital income tax, they conclude that it does not resolve the problem of the negative reaction of hours. Based on that finding I do not include the capital income in the models I consider.

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 put differently, 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.

3 The Setup of the Econometric Model

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. The identification is based on the assumptions that the system is stable over time (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 not more restrictive then 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 rather is a tool to test if economic restrictions that are just-identifying in the conventional SVAR are consistent with the data. In the current paper I consider a modified version of the model by Lanne et al. (2010).

Consider the VAR(p):

$$Y_t = \nu_{s_t} + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + U_t, \tag{3}$$

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,

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

In addition to the state dependent covariance matrices I allow also 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 the data similar to the data I use, tests for structural breaks in the productivity growth series and finds them to be likely. Therefore the model deployed in the subsequent analysis has to capture potential non-regularities also 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.

I do not allow the autoregressive parameters to be state dependent. Understandingly time-varying coefficient models have been used, for example, in monetary policy analysis (Primiceri, 2005; Sims and Zha, 2006) or oil market analysis (Baumeister and Peersman, 2010). The drifting coefficients capture possible time variation in the lag structure of the model and are economically interpreted as time dependent monetary rule (Sims and Zha, 2006) or speed of oil price adjustment and time varying elasticities (Baumeister and Peersman, 2010). In the current study it would be more common to allow for the time dependent covariance matrices to capture changes in the volatility of shocks and variations in the intercept to capture the business cycle component in the spirit of Krolzig (1997). It should be noted that allowing for a time dependent lag structure would make impulse response analysis much more involved and hardly comparable with the results in the existing literature. Leaving the VAR coefficients to be time invariant can be to some extent supported by the findings of Sims and Zha (2006), who have shown that the time varying coefficient model is less favored by the data. Admittedly, their results were obtained using a different system of variables.

The changes in the volatility of the residuals are used in this framework to test if the identified shocks are in line with the properties of the data. For instance, if there are two volatility states ($M = 2$), then there exists a decomposition of the covariance matrices $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2 B'$, where $\Lambda_2 = \text{diag}(\lambda_{21}, \ldots, \lambda_{2K})$ is a diagonal matrix with the positive diagonal entries. The Λ_2 matrix is then 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 $\varepsilon_t = B^{-1}U_t$. Then any further restrictions induced by theoretical models become over-identifying and testable.

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

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

with diagonal Λ_i matrices is restrictive. Lanne et al. (2010) discuss conditions and ways to test for the exact (local) identification in the cases where the number of states is greater then 2. If these are satisfied, the resulting shocks are unique and imposed restrictions are over-identifying and testable.

It is worth pointing out that the requirement of having distinct relative variances is necessary for exact identification of all shocks. An important advantage of the approach adopted in the current paper is that the equality of λ_{mi}s can be checked with standard statistical tests. An extensive discussion of tests for 2 and 3 state MS models can be found in Lanne et al. (2010) and Herwartz and Lütkepohl (2011).

Since I assume 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). The authors emphasize that the normality assumption is not essential for the asymptotic properties of the estimates but is just used for setting up the likelihood function. In the current paper the expectation maximization (EM) algorithm of Herwartz and Lütkepohl (2011) is adopted and updated to allow for changes in the intercept.

3.2 Bootstrapping confidence bands

In the MS models bootstrapping confidence bands for impulse responses may be problematic, therefore a discussion of the procedure deployed in the present paper is useful. Herwartz and Lütkepohl (2011) discuss a fixed design wild bootstrap for constructing confidence intervals for impulse responses. They propose to construct bootstrap samples conditional on estimated state probabilities and the ML estimates. For the current model, I take into account the changes in the intercepts when constructing the bootstrapped series. One of the ways to do it is to use a weighted average of the intercept for each t, with the weights being the estimated state probabilities. Then for the current model the bootstrapped series can be represented as:

$$Y_t^* = \mu_t + \hat{A}_1 Y_{t-1} + \cdots + \hat{A}_p Y_{t-p} + U_t^*, \tag{6}$$

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 and $\hat{\nu}_{s_t}$ 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, conditioning on the initially estimated transition probabilities. The procedure allows the preservation of potential heteroskedasticity and changes in the intercept of the data. Therefore the weights for the intercept in the bootstrap loop do not change.

Note that Herwartz and Lütkepohl (2011) condition also on the estimated Λ_i, $i = 2, \ldots, M$ matrices. I relax this assumption and estimate the weighting matrices, also in the bootstrap step. In order to eliminate potential interchanges of columns of the B matrix one has 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 that, 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 $\hat{B}$. That is done to avoid interchanges in signs of the B and reduce confidence bands as discussed in Lütkepohl (2011). 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. In the current example I fix the elements on the main diagonal of the B to be positive. In the end of 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 has nothing to do with the validation of sign restricted impulse responses and is just a device for reducing confidence bands for impulse responses.

It should be emphasized that computing the bootstrapped impulse responses in that way requires 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% percentile confidence intervals based on 1000 replications.

4 The Data

I use quarterly data from 1947:Q1 through 2010:Q4 to estimate the MS models. The data is relatively standard to that literature and is similar to that used by both Francis and Ramey (2005) and Francis and Ramey (2009). For the series on labor productivity and labor input, I use the Bureau of Labor Statistics series 'Index of output per hour' and 'Index of hours' in the business sector, respectively. Labor input is put on a per capita basis by dividing by the population age 16 and above. The variables are used in logarithms.

Standard ADF tests for both productivity and hours indicate the presence of a unit root. However Christiano et al. (2003) argue that hours per capita cannot logically have a unit root as it is a bounded process. They enter hours in levels and find that a positive technology shock leads to an increase in hours worked.

Fernald (2007) shows that the low frequency movements in the data, namely a U-shaped trend in both productivity growth and hours per capita, is driving the results of Christiano et al. (2003). When he removes the low frequency movements from either productivity growth or hours per capita, the response of hours worked to a positive technology shock is negative. Francis and Ramey (2009) further exploit the low frequency properties of the data, and show that the U-shapes in trend in hours per capita are not coincidental, but are due to entry of the baby boom generation into the labor market. They use the Hodrick-Prescott (HP) filter as well as demographic controls to remove a potentially distorting trend from hours and reach similar results as Fernald (2007).

Understandably, the trend present in the hours series should be taken into account. Following Francis and Ramey (2009), I remove low-frequency movements in hours per capita using a conservative standard HP filter with a parameter $\lambda = 16000$. The resulting trend is U-shaped similar to the Figure 3 of Francis and Ramey (2009) and is not shown to conserve space. Francis and Ramey (2009) enter filtered data both in levels and in first differences. Both specifications produce a result similar to Gali (1999): a positive technology shock leads to a decline in hours. I use only a levels specification as it seems to be the most problematic one.

However there could be a problem with the HP filtered data. Recall that the standard HP filter is a two sided filter that uses information at time $t-1$ as well as $t+1$ to process data at time t. That could possibly change the information set compared to the unfiltered data when it comes to the estimation of state probabilities and computing impulse responses.[2] Therefore I foresee that there may be potential complications with this dataset, as the relevant outcomes of the estimation would be computed using different information sets. In order to minimize possible alteration of the information set, I construct another dataset where I use a one-sided HP filter as in Stock and Watson (1999). The trend is extracted using a Kalman filter and it is similarly U shaped.

After the above mentioned manipulations, I end up with 4 datasets:

1. Both productivity and per capita hours enter in first differences;

2. Productivity enters in first differences and per capita hours in levels;

[2] In fact, a similar problem could be present in the data due to revisions that are conducted by the issuing authority. However potential influence on the results that arises from the revisions should not be as dramatic as explicit manipulation of the data with any sort of filtering. Even if some revision distortions are present in the productivity series, they remain the same across datasets as the productivity data is not manipulated.

3. Productivity enters in first differences and per capita hours in levels is HP filtered with two-sided filter.

4. Productivity enters in first differences and per capita hours in levels is HP filtered with one-sided filter.

The data is plotted in Figure 1.

For the results to be comparable to the ones known in the related literature, I use models with 4 lags. It is not far from the number of lags that the more generous Akaike Information Criterion (AIC) indicates (2 for the dataset $y_t = [\Delta x_t, \Delta n_t]'$ and 3 for the other datasets). Four lags should be also enough to capture the dynamics in the model with the quarterly data.

It should be emphasized that the methodology discussed in Section 3 does not allow to discriminate among the four datasets and determine which would be the most appropriate one to use. It is rather possible to discuss the results induced by the data and assess the validity of the identifying restrictions for each dataset. That is presented in the next section.

5 Empirical analysis

5.1 Analysis of states

In Table 1 the range of estimated models together with the corresponding values of the log-likelihood, Akaike Information Criterion (AIC) and Schwarz Criterion (SC) are presented. In the current study models with 2 and 3 Markov states, with different restrictions for each of the 4 datasets, are compared. First, according to the information criteria, the models with MS are much preferred to the standard VAR(4) models. This result is consistent with the residual graphs in Figure 2, where reduced volatility of the residuals in the second half of the samples can be seen.

Comparing only unrestricted MSIH-VAR(4) models, the MSIH(2)-VAR(4) is favored by SC while AIC prefers a 3-state model. Among 2 state models, SC values are further reduced by imposing identification restrictions, whereas AIC is reduced for the lower-triangular Ξ_∞ model for datasets with first differenced and stationary hours as well as for the one sided filtered data.

If only 3 state models are considered, one would note that AIC favors a model with state invariant B for the first differenced and one sided filtered data, whereas models without assumption of state-invariant B are preferred for the other datasets. That is not a surprise as AIC would select, in general, the least restricted models. SC favors a lower triangular Ξ_∞ for all the datasets, apart from two sided filtered data.

Thus, based on the model selection criteria, 2 state models are favored by SC and 3 state models by AIC. Models with identification restrictions and state invariant B have support mainly by SC, although they are not favored generally. The information criteria are used to select the number of states by, among others, Psaradakis and Spagnolo (2003) and Lanne et al. (2010). Notwithstanding, apart from AIC and SC, one would need additional evidence in favor of the number of states for the particular cases I have.

In addition to the model selection criteria, it is useful to look at the smoothed state probabilities in deciding on the number of states. They are shown in Figure 3 for the 2 state models and in Figure 4 for the 3 state models. The corresponding state covariance matrices for 2 state models are given in Table 2. The figures show that volatility changes are present during the sample period; both 2 and 3 state models capture it. The smoothed state probabilities for models $y_t = [\Delta x_t, \Delta n_t]'$ and $y_t = [\Delta x_t, n_t]'$ are nearly identical for 2 state case. The state-invariant B matrix is imposed for the MSIH(3) models. Again state probabilities for models $y_t = [\Delta x_t, \Delta n_t]'$ and $y_t = [\Delta x_t, n_t]'$ look similar. On the other hand, the state probabilities for the filtered data look quite different. From Table 2 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 variance of productivity and hours is 3-4 times 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. Apart from that, the estimated state probabilities reveal the great moderation phenomena that started in the beginning of the 80s and lasted until the late 90s.

It should be emphasized that the estimated state probabilities of the 2 state models for the two-sided HP filtered data look different if compared to the unfiltered data. The low volatility period is estimated to have more pronounced peaks in the period of late 1940-s and beginning of 50s as well as around 1960 and 1970. The high volatility state has more pronounced peaks around 1990 and 1999-2000. The period associated with the great moderation is similar for the both filtered datasets. The changes in the estimated state probabilities are due to the filtering and may be an indication of the changed information structure of the data. On the contrary, the estimated state probabilities for the one sided HP filtered data resemble the unfiltered data much more. The main difference is in a pronounced peak in the high volatility State 1 around year 2000, which is not seen in the unfiltered data.

Looking at Figure 4 it becomes clear that the high volatility State 1 of MSIH(2) models is now split into 2 states for the majority of the datasets. The smoothed probabilities of the split states are estimated as peaks that interchange often. That means that the MSIH(3) model might hardly distinguish between these states and the states do not have a clear interpretation. On the contrary, the estimated state probabilities of the one-sided HP filtered data (Figure 4, (d)) can be interpreted. State 1 represent the low volatility state, State 2 is the high volatility state and State 3 captures the periods of notably low volatility of the hours series. The third state, which is representing some rare events in the economy, is only associated with relatively few periods. Hence, the model may be difficult to estimate and the estimates are potentially unreliable.

Summing up, from a statistical point of view as well as considering the estimated state probabilities, the 3 state models seem to be more problematic for the current datasets. On the other hand, the 2 state models do the job well enough in capturing the changes in volatility and potential non-linearities in the intercept. However one could think that there are not enough convincing arguments to exclude either 2 state or 3 state models from the procedure of testing the identifying restrictions.

Therefore the testing of the restrictions in the next section will be discussed in the context of both 2 and 3 state models. A little ahead it will be shown, that the main conclusions regarding the validity of the identification schemes will not depend on the number of states. Further the impulse response analysis will be conducted for the datasets that support the restrictions.

5.2 Statistical analysis of MSIH models

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 the Section 3, that to obtain a statistical identification of the shocks for a 2 state model, it is enough to check if 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. The standard errors indicate that estimation precision is quite good for the 2 state models and, hence, I anticipate that the estimates are statistically different.

Recall that B is locally identified in the 2 state model (apart from changes in 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}. In the related literature Wald and likelihood ratio (LR) tests are used in the context (see for example Lanne et al. (2010) and Herwartz and Lütkepohl (2011)). In both papers, if the Wald test rejects equality, then the LR tests do so as well. Therefore in the current paper I check the results of the Wald tests first and if I get enough evidence in favor of the distinct λ_{ij}s , I skip the LR tests. The results are presented in Table 4. Two of the four null hypotheses are rejected by the tests at a 5% significance level, while the other two are not far away from 5% and are clearly rejected at a 10% level. Hence, based on the Wald tests, there is enough evidence in favor of the unique B. In other words, the shocks are uniquely identified by the data.

That means that I have achieved a statistical identification of the 2 state models. The obtained shocks are unique but they are not labeled economically. With that identification in hand the economic restrictions on Ξ_∞ become overidentifying. The main question is whether the data supports the economically meaningful shocks identified by Gali (1999) and Francis and Ramey (2005). The usual LR tests are applicable to perform the testing of the restrictions. The outcomes of the LR tests are shown in Table 5. The SC values and LR test for the datasets $y_t = [\Delta x_t, \Delta n_t]'$ and $y_t = [\Delta x_t, n_t]'$ support the lower-triangular Ξ_∞ matrix. For the dataset $y_t = [\Delta x_t, n_t]'$ the data does not object also to the upper triangularity of the Ξ_∞. Therefore the original identification scheme of Gali (1999) seems to be consistent with the properties of the unfiltered data.

The results for the two sided filtered data are quite interesting. Both identification schemes tend to be rejected by the data. The p values of the LR tests are around 0.07. That is potentially an indication of the problem with the changed information set as discussed before. The results of the tests for the one sided filter support this conjecture. These are similar to the unfiltered data, the long run restriction as in Gali (1999) is supported with $p = 0.2146$. The SC value also favors the model restricted in that way. The other tested restriction $C_{21}(1) = 0$ has less support with p value being slightly below 10%. These results are similar to the ones obtained with the dataset where $y_t = [\Delta x_t, \Delta n_t]'$.

The results of the tests are consistent with the estimated relative variances, which are shown in Table 3. It should be noted that the λ_{2i}s for restricted and unrestricted models are quite similar for the cases where LR tests favor the identification. One should note, however, that the ordering of λ_{2i}s is different for the restricted model. The columns of B can no longer be permuted when the Ξ_∞ is triangular. Thus, the ordering of the λ_{2i}s corresponds to the lower-triangular Ξ_∞ matrix for the restricted cases. For the unrestricted cases, the λ_{2i}s are forced to be in increasing order to prevent permutations in B for the reasons discussed in Section 3.

The discussion until now was based on the more parsimonious 2 state models. I would like to emphasize that the same analysis, with roughly same conclusions, can be done for the 3 state models. The corresponding results of the estimations are shown in the Appendix.

The estimated structural parameters and their standard errors for the 3 state models are shown Table A.1. The tests for the (local) identification of the B in the 3 state mode context are more complicated. It requires to test whether the diagonal elements of Λ_2 and Λ_3 are jointly equal. The result of the testing is shown in Table A.2. Due to relatively high standard errors, Wald tests fail to reject pairwise equality of the relative variances. Therefore computationally more demanding LR tests are conducted. These indicate that the relative variances are distinct in the 3 state models and hence, the shocks with the assumed state invariant B models are statistically identified.

The fact that the testing of the restrictions for the 3 state models yields similar results to the 2 state models can be seen in Table A.3. Recall that in the 3 state models the assumption on the state invariant instantaneous effects can be tested. The state invariant B is supported for the datasets where hours enter in first differences and for the one sided filter data where the p values for the relevant tests are above 15% level.

For the dataset $y_t = [\Delta x_t, \Delta n_t]$, the restriction $C_{12}(1) = 0$ cannot be rejected at the 10% level for the alternative of unrestricted MSIH(3) model with the $p = 0.0907$. Admittedly, the p value is lower, than in the 2 state model case. The test of the restriction $C_{21}(1) = 0$ for the alternative of unrestricted MSIH(3) model has $p = 0.0707$, which is close to the result for the 2 state model. As in the 2 state model case, the original identification of Gali (1999) has more support from the data than the restriction of Francis and Ramey (2005), which excludes permanent effect of technology shocks on hours worked.

For the one sided filtered data $y_t = [\Delta x_t, n_t]$ the restriction of Gali (1999) is supported with the $p = 0.1262$. That finding is similar to the 2 state model case. The alternative identification of technology shocks restricting $C_{21}(1) = 0$ has substantially less support from the data and can be rejected at a 5% level.

Now I will discuss the restrictions for the datasets where the state invariant B was rejected. The dataset $y_t = [\Delta x_t, n_t]$ shows very little support for the identification condition (state invariant B). Therefore all of the restrictions are rejected here. That was not the case in the 2 state model. Apparently the hours series in the dataset has the distorting trend and that could be driving the controversial results. The dataset will be further discussed in the impulse response analysis section in the context of the 2 state model.

For the two sided filtered data, the p value of the test for the state invariant B is just below 5% level. If one believes in the assumption of the state invariant B both possible restrictions are rejected, as in the 2 state model case.

The outcome of the testing can be briefly summarized as follows: (1) the identification of shocks as in Gali (1999) is supported by 3 out of 4 datasets for the models with 2 Markov states and by 2 out of 4 datasets for the models with 3 Markov states; (2) the dataset where the two sided HP was used appear to be problematic for the impulse response analysis independent of the number of states; and (3) excluding permanent technology shocks, as in Francis and Ramey (2005), is strongly supported only by the unfiltered data where $y_t = [\Delta x_t, n_t]'$ in the 2 state model case.

As a robustness check of the outcomes of the testing, I performed the same analysis for the models with a constant intercept. The 2 and 3 state models were estimated and the testing of the restrictions was performed. The results for the 2 state models confirm the results induced by the MSIH(2) models: the identification by Gali (1999) is supported by the data, whereas restriction $C_{21}(1) = 0$ is not favored by the dataset $y_t = [\Delta x_t, \Delta n_t]'$ and by the one-sided HP filtered data. The 3 state models seem to be more problematic in the context of the constant intercept models. The state invariant B is rejected for the dataset $y_t = [\Delta x_t, \Delta n_t]'$ and the one sided filtered data. Assuming the state invariant B for these datasets, the identification by Gali (1999) can be confirmed. When hours are assumed to be stationary, none of the identifying restrictions can be rejected. On the contrary, for the 2 sided filtered data all of the restrictions can be rejected at a high significance level. Admittedly, there are some discrepancies in the results compared to MSIH model that are mainly due to more complex 3 state models. Notwithstanding the main message remains the same: Gali (1999) identification has more support from the data than Francis and Ramey (2005) identification that excludes permanent technology shocks. The detailed results are available upon request.

5.3 Impulse Response Analysis

Given that some of the identification schemes were supported by the data, the impulse response (IR) analysis may be performed for these schemes and datasets that support those. The dataset where hours were subject to a two sided HP filter is excluded from the analysis, as too little evidence in favor of the restrictions was found. In the current section both 2 and 3 state models are considered.

The impulse responses for the productivity were accumulated for all datasets, the responses for hours were accumulated if hours entered in first differences. Some of the impulse responses fall outside the respective 90% bootstrap confidence bands. That feature has been observed in some other studies as well and is not uncommon in the SVAR literature. In the current study it might be due to a complex optimization step in the bootstrap cycle.

In Figure 5 the responses to the technology shock in MSIH(2) models identified as in Gali (1999) are shown for the datasets that support the restriction. For the first dataset where hours enter in first differences, the responses are consistent with the previous findings in the literature (Gali (1999), Christiano et al. (2003), Francis and Ramey (2005) and others): the productivity improves significantly, while hours are negative on impact, then rise but remain negative. It should be noted, that the upper confidence band starting from around 4-th quarter is above 0. That feature is also common to the results in the related literature.

The results for the data where hours enter in levels is also consistent with Christiano et al. (2003): both productivity and hours improve. But as discussed before, this result is shown to be driven solely by the low frequency fluctuations in the data. Removing these by one sided HP filter I obtain results similar to Francis and Ramey (2009): as in the difference specification, hours fall on impact of the positive technology shock, and then start rising. In Francis and Ramey (2009) the response is positive during quarters 4-13, however that is not the case in the current setup. As one can see in Figure 5 (c), the response is negative in the beginning of the response horizon, then it becomes only slightly positive during quarters 8-13 and remains close to 0 from then on. The confidence bands are relatively wide and are above and below 0, therefore it can not be clearly stated whether the reaction of hours remains negative or becomes positive. It should be emphasized, that the results for the one sided HP filtered data rather confirm the results obtained in Francis and Ramey (2009), but as shown above, the standard HP filtered data may not suite for the impulse response analysis.

The IRs to the technology shock of the MSIH(3) models identified, as in Gali (1999), are shown in Figure 7. The reaction of variables is quite similar to the MSIH(2) models. For the dataset $y_t = [\Delta x_t, \Delta n_t]'$ the reaction of hours is negative on impact, but after 3 quarters it becomes positive. However the confidence bands of that horizon are above and below zero, therefore it remains unclear if the reaction to the technology shock is positive or negative in the limit. The reaction of hours in Figure 7 (b) is negative on impact and the dynamics is similar to the IR of the MSIH(2) (Figure 5 (c)).

Figures 6 and 8 show the reaction to the technology shock identified as a shock that excludes permanent effects on hours (Francis and Ramey, 2005) in MSIH(2) and MSIH(3) models respectively. One can see that productivity rises and hours fall, which is consistent with the findings in Francis and Ramey (2005). However the dataset that strongly supports the restriction in MSIH(2) model is the non filtered levels specification and it is plagued with the problem of the low frequency fluctuations discussed by Fernald

(2007). Therefore it can be that the results are driven by these features of the raw data. Apparently the results of the MSIH(3) model for the dataset $y_t = [\Delta x_t, \Delta n_t]'$ speak against this conjecture: the IRs in Figure 7 are very similar to Figure 5: productivity increases while hours significantly fall on impact.

The variety of the IRs discussed in this section have a common feature: the reaction of hours is negative on impact and it does not become significantly positive within the response horizon. The only exception is the dataset where the low frequency movements are not taken into account. Hence the results known in the literature is not objected by the data: hours worked fall after a positive technology shock.

6 Conclusions

In the present paper I reconsider the effect of technology shocks on productivity and hours worked in the well known bivariate system. I used 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 there are proposals 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.

It is shown that conflicting results in the debate concerning the effects of technology shocks on hours stems from the low-frequency movements in the measure of hours. Therefore I construct 4 datasets that have been used in the literature and treat hours differently: non-stationary, stationary, hours filtered using standard HP filter, hours filtered using the one sided HP filter.

Different identification schemes with long run restrictions are used by Gali (1999) and Francis and Ramey (2005). The latter propose excluding permanent technology shocks, while Gali (1999) assumes that the technology shocks are the only sources of variation in the productivity. In the conventional framework potentially competing restrictions are just identifying and hence not testable. The setup of the econometric model allows for the extraction of additional information out of the data and for the testing of validity of just identifying long run restrictions in the bivariate system.

The result of the testing procedure show that the classical identification scheme proposed by Gali (1999) is supported by the majority of the datasets. The other just-identifying restriction, proposed by Francis and Ramey (2005), has less support in the data. That conclusion is robust to the number of Markov states assumed in the system. Interesting implications for the identification schemes has the standard two sided HP filter. I argue that it is rather difficult to justify any of the restrictions for the data filtered with the standard HP filter. The problem could be due to the changed information structure of the data. A good argument for that reveals the data that was filtered using the one sided HP filter.

Having enough evidence in favor of the restrictions, I perform impulse response analysis. The bottomline result suggests that the reaction of hours worked to that technology shock identified as the only source of variation in productivity is negative on impact. That is the result that has strong support in the data. However the reaction over a longer response horizon is not clearly negative as the confidence bands of the responses are above and below zero. That is true when the low frequency movements are accounted for in the data. The findings are in line with the literature and are robust to a range of model specification issues.

References

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.

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.

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.

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.

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. (2011). Reducing confidence bands for simulated impulse responses, *Unpublished manuscript.*

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.

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

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.

Sims, C. A. and Zha, T. (2006). Were there regime switches in U.S. monetary policy?, *American Economic Review* **96**(1): 54–81.

Stock, J. H. and Watson, M. W. (1999). Forecasting inflation, *Journal of Monetary Economics* **44**(2): 293–335.

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.

Table 1: Comparison of MS-VAR(4) Models

	$y_t = [\Delta x_t, \Delta n_t]'$			$y_t = [\Delta x_t, n_t]'$			$y_t = [\Delta x_t, n_t]'$, two sided filtered			$y_t = [\Delta x_t, n_t]'$, one sided filter		
Model	$\log L_T$	AIC	SC	$\log L_T$	AIC	SC	$\log L_T$	AIC	SC	$\log L_T$	AIC	SC
VAR without MS	1696.55	-3351.10	-3275.41	1702.68	-3363.36	-3289.66	1712.95	-3383.90	-3310.21	1726.97	-3379.94	-3250.09
MSIH(2), unrestricted	1729.88	-3403.76	-3305.49	1733.02	-3410.01	-3311.78	1746.30	-3436.60	-3338.34	1758.82	-3461.65	-3363.39
MSIH(2), $C_{12}(1) = 0$	1729.42	-3404.85	**-3310.09**	1732.48	-3410.96	**-3316.21**	1744.69	-3435.38	-3340.63	1758.05	-3462.10	**-3367.35**
MSIH(2), $C_{21}(1) = 0$	1728.31	-3402.61	-3307.86	1732.09	-3410.18	-3315.43	1744.75	-3435.50	**-3340.75**	1757.45	-3460.9	-3366.15
MSIH(3), unrestricted	1747.46	-3420.92	-3290.48	1755.24	**-3436.48**	-3306.4	1767.55	**-3461.11**	-3330.67	1774.60	-3475.21	-3344.77
MSIH(3), state inv. B	1746.87	**-3421.75**	-3295.41	1746.49	-3420.99	-3294.66	1765.61	-3459.23	-3332.89	1773.67	**-3475.35**	-3349.01
MSIH(3), $C_{12}(1) = 0$	1745.06	-3420.12	-3297.29	1745.32	-3420.65	-3297.82	1760.35	-3450.69	-3327.87	1772.53	-3475.07	-3352.24
MSIH(3), $C_{21}(1) = 0$	1744.81	-3419.63	-3296.79	1743.73	-3417.47	-3294.64	1758.89	-3447.77	-3324.95	1770.22	-3470.45	-3347.62

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

Table 2: Estimated State Covariance Matrices of MSIH(2)-VAR(4) Models

	$y_t = [\Delta x_t, \Delta n_t]'$	$y_t = [\Delta x_t, n_t]'$	$y_t = [\Delta x_t, n_t]'$, two sided filter	$y_t = [\Delta x_t, n_t]'$, one sided filter
$\Sigma_1 \times 10^{-3}$	$\begin{bmatrix} 0.124 & \\ 0.017 & 0.093 \end{bmatrix}$	$\begin{bmatrix} 0.118 & \\ 0.019 & 0.090 \end{bmatrix}$	$\begin{bmatrix} 0.127 & \\ 0.012 & 0.082 \end{bmatrix}$	$\begin{bmatrix} 0.109 & \\ 0.012 & 0.059 \end{bmatrix}$
$\Sigma_2 \times 10^{-3}$	$\begin{bmatrix} 0.033 & \\ -0.007 & 0.026 \end{bmatrix}$	$\begin{bmatrix} 0.034 & \\ -0.006 & 0.025 \end{bmatrix}$	$\begin{bmatrix} 0.030 & \\ -0.007 & 0.022 \end{bmatrix}$	$\begin{bmatrix} 0.030 & \\ -0.005 & 0.018 \end{bmatrix}$

Table 3: Estimates of Structural Parameters of MSIH(2)-VAR(4) Models

		$y_t = [\Delta x_t, \Delta n_t]'$		$y_t = [\Delta x_t, n_t]'$		$y_t = [\Delta x_t, n_t]'$, two sided filter		$y_t = [\Delta x_t, n_t]'$, one sided filter	
Model	Parameter	estimate	std.dev.	estimate	std.dev.	estimate	std.dev.	estimate	std.dev.
Unrestricted MSIH(2)	λ_{21}	0.181	0.060	0.189	0.058	0.166	0.050	0.189	0.041
	λ_{22}	0.396	0.099	0.428	0.113	0.367	0.099	0.413	0.104
MSIH(2), $C_{12}(1) = 0$	λ_{21}	0.381	0.098	0.188	0.205	0.332	0.110	0.432	0.107
	λ_{22}	0.189	0.044	0.416	0.100	0.207	0.055	0.186	0.039
MSIH(2), $C_{21}(1) = 0$	λ_{21}	0.357	0.088	0.473	0.232	0.337	0.087	0.371	0.468
	λ_{22}	0.208	0.045	0.225	0.066	0.207	0.046	0.220	0.045

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

Table 4: Tests for Equality of λ_{ij}'s for unrestricted MSIH(2)-VAR(4) models

$H_0 : \lambda_{21} = \lambda_{22}$	$y_t = [\Delta x_t, \Delta n_t]'$	$y_t = [\Delta x_t, n_t]'$	$y_t = [\Delta x_t, n_t]'$, two sided filter	$y_t = [\Delta x_t, n_t]'$, one sided filter
Wald statistic	3.459	4.206	3.851	3.720
p-value	0.0629	0.0403	0.0497	0.0538

Table 5: LR Tests of Restrictions for MSIH(2)-VAR(4) Models

		$y_t = [\Delta x_t, \Delta n_t]'$		$y_t = [\Delta x_t, n_t]'$		$y_t = [\Delta x_t, n_t]'$,two sided filter		$y_t = [\Delta x_t, n_t]'$,one sided filter	
H_0	H_1	LR	p-value	LR	p-value	LR	p-value	LR	p-value
MSIH(2), $C_{12}(1) = 0$	unrestricted MSIH(2)	0.92	0.3375	1.08	0.2987	3.22	0.0727	1.54	0.2146
MSIH(2), $C_{21}(1) = 0$	unrestricted MSIH(2)	3.14	0.0764	1.86	0.1726	3.10	0.0783	2.74	0.0979

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.

Figure 1: The data

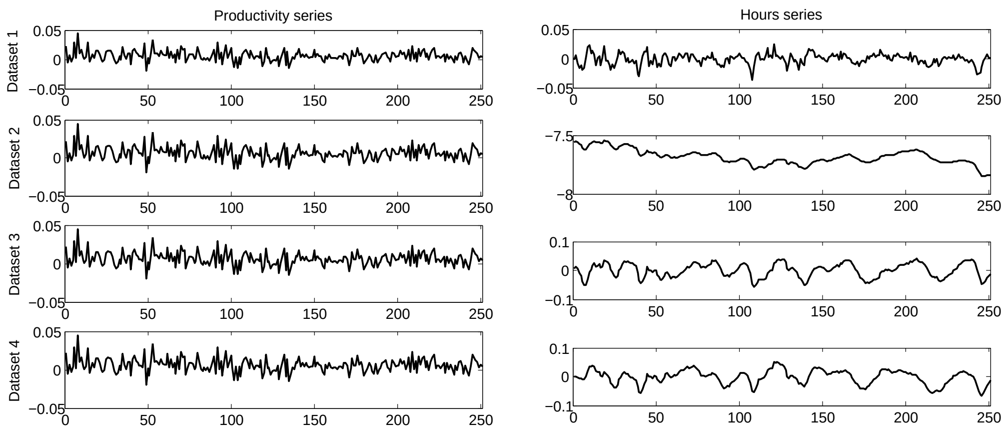


Figure 2: Residuals of VAR(4) models

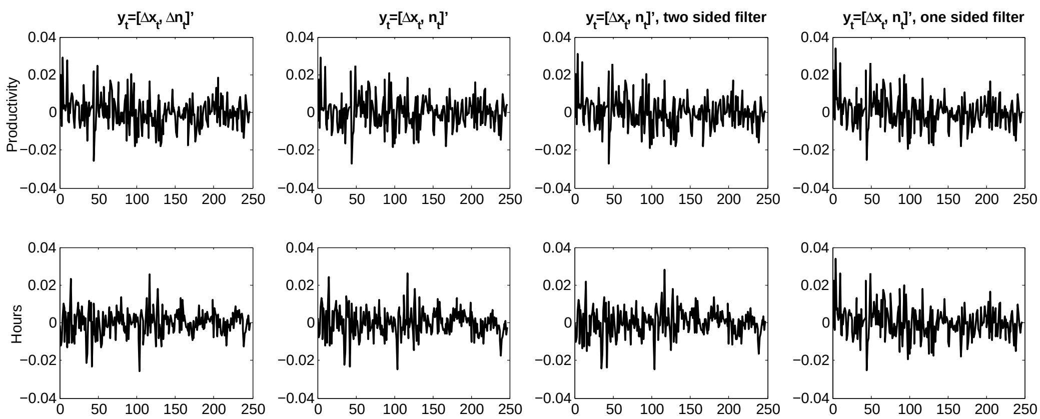
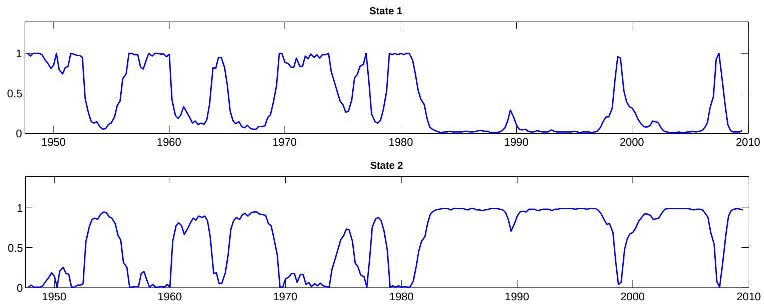
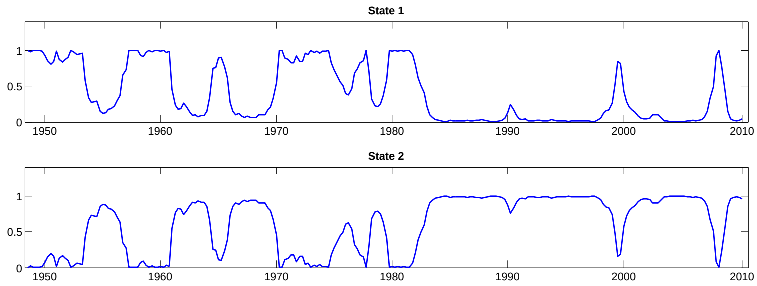


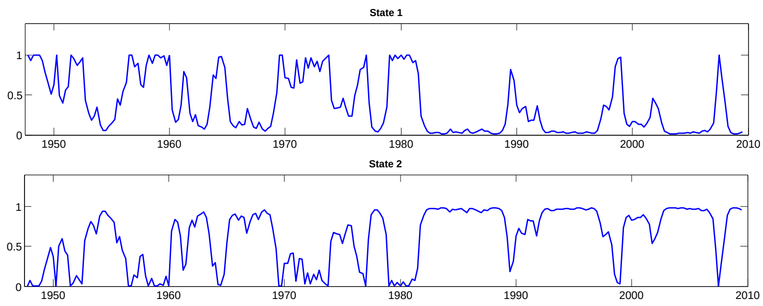
Figure 3: Smoothed state probabilities of MSIH(2)-VAR(4) models



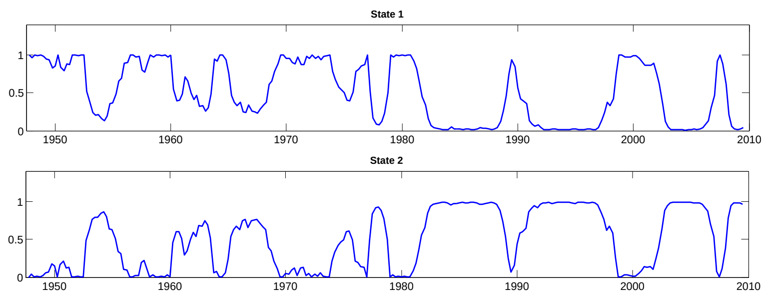
(a) Dataset 1: $y_t = [\Delta x_t, \Delta n_t]'$



(b) Dataset 1: $y_t = [\Delta x_t, n_t]'$

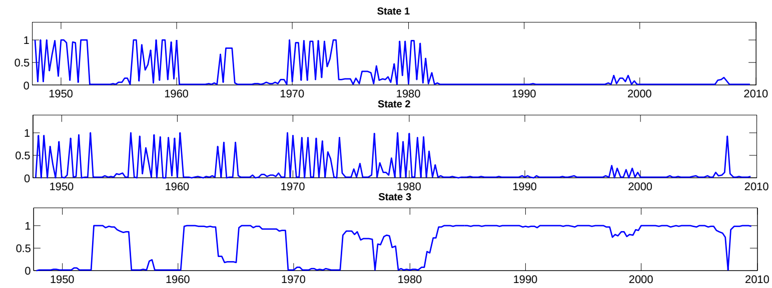


(c) Dataset 3: $y_t = [\Delta x_t, n_t]'$, two sided filter

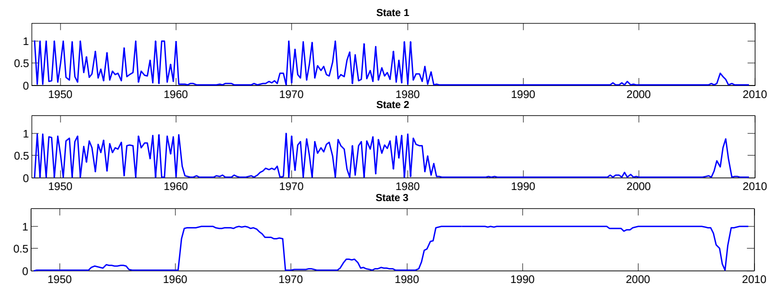


(d) Dataset 4: $y_t = [\Delta x_t, n_t]'$, one sided filter

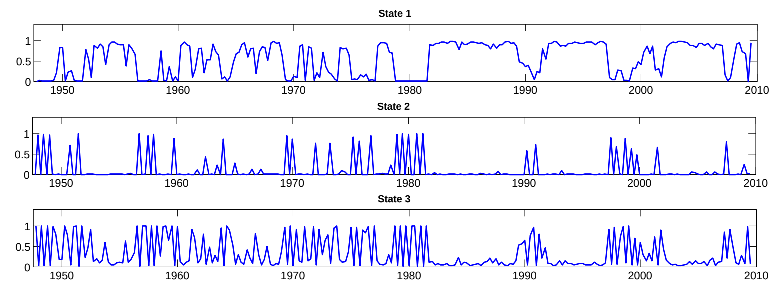
Figure 4: Smoothed state probabilities of MSIH(3)-VAR(4) models



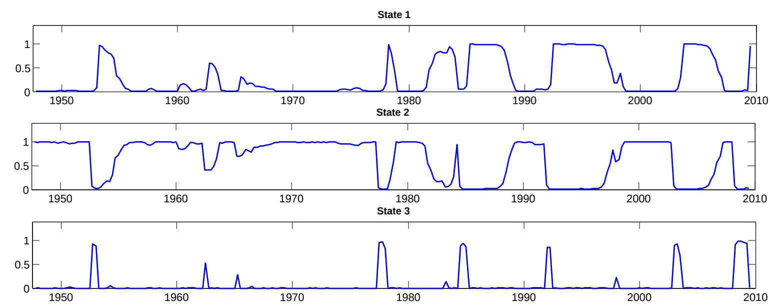
(a) Dataset 1: $y_t = [\Delta x_t, \Delta n_t]'$



(b) Dataset 2: $y = [\Delta x_t, n_t]'$

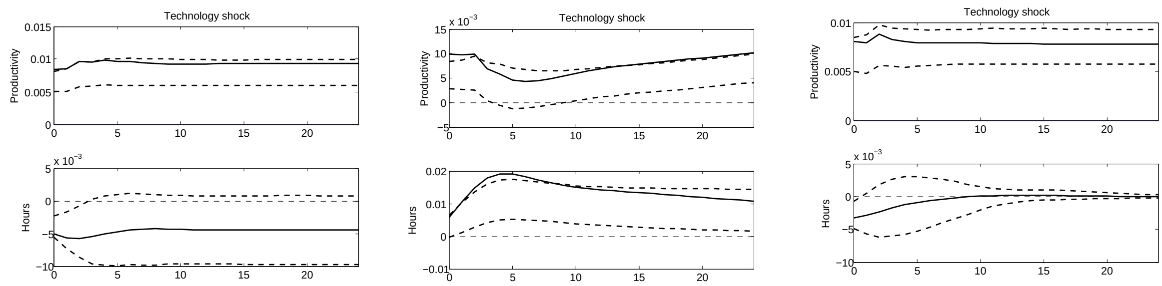


(c) Dataset 3: $y_t = [\Delta x_t, n_t]'$, two sided filter



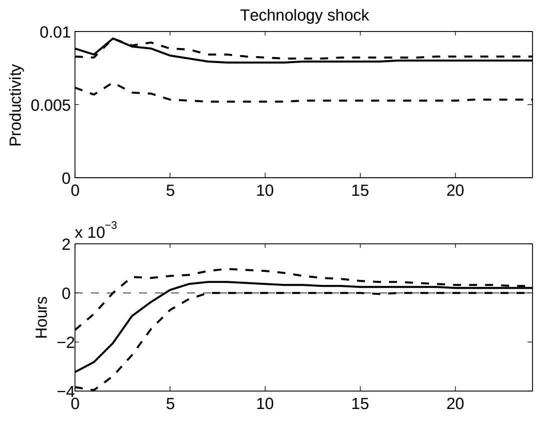
(d) Dataset 4: $y_t = [\Delta x_t, n_t]'$, one sided filter

Figure 5: Responses to a positive technology shock identified as in Gali (1999), MSIH(2) models



(a) Dataset 1: $y_t = [\Delta x_t, \Delta n_t]'$ (b) Dataset 2: $y_t = [\Delta x_t, n_t]'$ (c) Dataset 4: $y_t = [\Delta x_t, n_t]'$, one sided filter

Figure 6: Responses to a positive non-permanent technology shock, MSIH(2) models



(a) Dataset 2: $y_t = [\Delta x_t, n_t]'$

A Appendix

Table A.1: Estimates of Structural Parameters of MSIH(3)-VAR(4) Models

		$y_t = [\Delta x_t, \Delta n_t]'$		$y_t = [\Delta x_t, n_t]'$		$y_t = [\Delta x_t, n_t]'$, two sided filter		$y_t = [\Delta x_t, n_t]'$, one sided filter	
Model	Parameter	estimate	std.dev.	estimate	std.dev.	estimate	std.dev.	estimate	std.dev.
MSIH(3), state inv.B	λ_{21}	0.023	0.018	0.122	0.048	0.015	0.010	2.659	0.706
	λ_{22}	1.718	1.466	1.224	0.515	8.476	3.928	5.745	1.562
	λ_{31}	0.270	0.130	0.315	0.103	3.524	0.881	0.867	0.689
	λ_{32}	0.256	0.129	0.250	0.080	4.656	1.163	0.010	0.008
MSIH(3), $C_{12}(1) = 0$	λ_{21}	0.774	0.415	4.016	1.057	3.084	0.818	0.030	0.018
	λ_{22}	0.010	0.007	3.294	0.861	5.432	1.112	1.488	0.764
	λ_{31}	3.073	0.695	0.001	0.006	4.578	2.078	0.289	0.091
	λ_{32}	5.524	1.199	0.243	0.167	0.010	0.006	0.237	0.078
MSIH(3), $C_{21}(1) = 0$	λ_{21}	0.010	0.005	2.764	0.759	3.492	0.894	2.778	0.913
	λ_{22}	1.227	0.637	5.661	1.403	6.611	1.481	3.366	1.082
	λ_{31}	0.296	0.070	1.017	0.503	3.097	1.422	6.526	2.147
	λ_{32}	0.229	0.055	0.192	0.162	0.010	0.008	5.703	1.613

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

Table A.2: Tests for Equality of λ_{ij}'s for unrestricted MSIH(3)-VAR(4) models

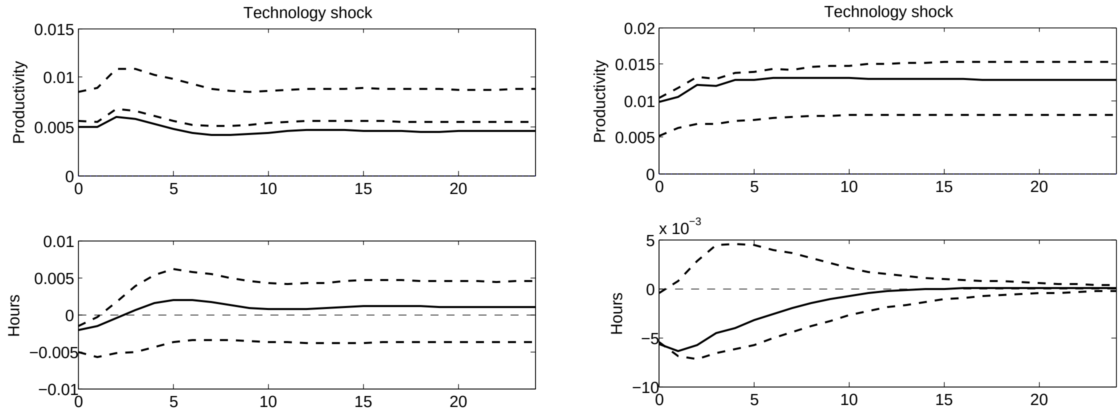
$H_0: \lambda_{21} = \lambda_{22}, \lambda_{31} = \lambda_{32}$	$y_t = [\Delta x_t, \Delta n_t]'$	$y_t = [\Delta x_t, n_t]'$	$y_t = [\Delta x_t, n_t]'$, two sided filter	$y_t = [\Delta x_t, n_t]'$, one sided filter
Wald statistic	1.43	6.18	3.87	4.93
p-value	0.48	0.04	0.14	0.08
LR statistic	26.23	9.27	23.78	7.48
p-value	2.01×10^{-6}	0.0097	6.8×10^{-6}	0.023

Table A.3: LR Tests of Restrictions for MSIH(3)-VAR(4) Models

		$y_t = [\Delta x_t, \Delta n_t]'$		$y_t = [\Delta x_t, n_t]'$		$y_t = [\Delta x_t, n_t]'$,two sided filter		$y_t = [\Delta x_t, n_t]'$,one sided filter	
H_0	H_1	LR	p-value	LR	p-value	LR	p-value	LR	p-value
State inv. B	Unrestricted MSIH(3)	1.18	0.2774	17.50	2.8×10^{-5}	3.88	0.0489	1.86	0.1726
State inv. B, $C_{12}(1) = 0$	State invariant B	3.62	0.0571	2.34	0.1261	10.52	0.0012	2.28	0.1311
State inv. B, $C_{12}(1) = 0$	Unrestricted MSIH(3)	4.80	0.0907	19.84	4.9×10^{-5}	14.40	0.0007	4.14	0.1262
State inv. B, $C_{21}(1) = 0$	State invariant B	4.12	0.0424	5.52	0.0188	13.44	0.002	6.90	0.0086
State inv. B, $C_{21}(1) = 0$	Unrestricted MSIH(3)	5.30	0.0707	23.02	1.0×10^{-5}	17.32	0.0002	8.76	0.0125

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.

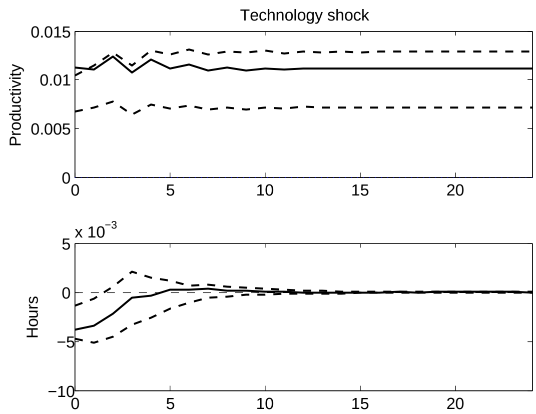
Figure 7: Responses to a positive technology shock identified as in Gali (1999), MSIH(3) models



(a) Dataset 1: $y_t = [\Delta x_t, \Delta n_t]'$

(b) Dataset 4: $y_t = [\Delta x_t, n_t]'$, one sided filter

Figure 8: Responses to a positive non-permanent technology shock, $y_t = [\Delta x_t, \Delta n_t]'$, MSIH(3) model



European
University
Institute
European University Press