Michał Bernardelli, Monika Dędys[‡]

The Viterbi Path of hidden Markov models in an analysis of business tendency surveys

Abstract

The aim of the paper is to show that turning points detection can be treated as a problem of pattern recognition. In the paper there are presented the results of applying normal hidden Markov models to a number of survey balances. Beyond a classical two-scale assessment of business activity a slightly more fuzzy classification of states is considered. To determine periods of unclear or difficult to evaluate situation unobservable Markov chains with three and four states are introduced. The outputs of the Viterbi algorithm, i.e. the most likely paths of unobservable states of Markov chains, are a basis of the proposed classification. The comparison of these paths with the business cycle turning points dated by OECD is described. The results obtained for three- and four-state Markov chains are close to those established in the references time series and seem to improve the speed with which, especially downshifts, are signaled. Furthermore, these results are more favorable than outcomes provided by conventional two-state models.

The method proposed in this paper seems to be a very effective tool to analyze results of business tendency surveys, in particular, when multistate Markov chains are considered. Moreover, proposed decompositions allow an easy comparison of two time series as far as turning point are concerned. In the paper survey balances are compared with ‘hard’ economic data such as sold manufacturing production. The results confirm the accuracy of assessment provided by survey respondents.

Keywords: hidden Markov model, Viterbi algorithm, business tendency surveys, business cycle turning point detection

JEL classification: C63, C83, E37

[‡] Warsaw School of Economics, Collegium of Economic Analysis, michal.bernardelli@sgh.waw.pl, monika.dedys@sgh.waw.pl, tel. +48 605 106 485.

1. Introduction

Markov-switching time-series models have been played a prominent role in the analysis of business cycle for decades. The idea of two unobservable regimes changing parameters of observed time series have been explored and generalized by many researchers. We claim that beyond a bivalent assessment of business activity, it is worth to consider a little more fuzzy classification. We would like to distinguish periods of unclear or difficult to evaluate situation, something between poor and good states of an economy, a signal of a change to come. One approach to this problem was presented in Abberger & Nierhaus (2010). A two-state underlying Markov chain and estimates of smoothed probability were proposed there as the basis for classification. In this paper we suggest a little bit different solution. Namely, we treat turning point detection as a problem of pattern recognition. We apply so called a normal hidden Markov model, i.e. the simplest type of a Markov-switching model with observable components being conditionally independent Gaussian variables, to business tendency survey data. Using the Viterbi algorithm we determine the most likely path of unobservable Markov chains with two, three and four states. We compare results of such decomposition with the business cycle turning points dated by OECD. It turns out that the described transformation of survey balances may provide crucial signals of coming changes in an economy, especially when downturns are considered.

It is worth to emphasize that the idea of involving multistate underlying Markov chain to analyze business cycles is not new (see Artis *et al.*, 1998; Çakmaklı *et al.*, 2013). However, in our research the meaning of states is slightly different and some additional conditions are imposed on transition probabilities. Moreover, we apply the Viterbi algorithm which seems to be very rarely used in an analysis of macroeconomic data. Furthermore, its application appears to be limited to two-state models only (Boldin, 1994).

The purpose of the paper is to show that the Viterbi algorithm provides a very effective tool for an analysis of business tendency surveys data, in particular when multistate Markov chains are considered. The most likely path of the Markov chain gives crucial information about considered economic time series and could be helpful in dating of business cycle turning points.

Besides Markov-switching models (Hamilton, 1994; Koskinen & Oeller, 2004) other methods of determining business cycle turning points are known. Obviously, all the methods use a sort of a (hopefully) leading indicator. It provides information about current and future states of a business cycle. A wide range of econometric methods should be mentioned. Most of all model-based methods (see Cleveland, 1972; Bell, 1984; Wildi & Schips, 2005) rely on ARIMA or state-space model-representations of the data generating process (DGP) often used with filters such as the Hodrick-Prescott (1997) or Christiano-Fitzgerald (2003) ones. Also econometric models based on logistic regression are used (Lamy, 1997; Birchenhall *et al*., 1999; Chin *et al*., 2000; Sensier *et al*., 2004). Finally, there is a group of spectral methods based on frequency filtering that use the Fourier (or other) transform (Addo *et al*., 2012).

The paper is composed of four sections. The basic terminology and methodology are presented in section 2. The third section is devoted to description of conducted experiments and their results. The paper ends with conclusions in section 4.

2. Models and the method of estimation

We focus on hidden Markov model (HMM), i.e. on the bivariate discrete stochastic process $\{X_k, Y_k\}_{k\geq 0}$ satisfying the following conditions:

- the process $\{X_t\}_{t\geq 0}$ is the homogenous Markov chain (MC);
- conditionally on the process $\{X_t\}_{t\geq 0}$ the observations $\{Y_t\}_{t\geq 0}$ are independent, and for each t the conditional distribution of Y_t depends on X_t only.

The Markov chain $\{X_t\}_{t\geq 0}$ is not observable. Its state space is denoted by S. When Y_t has univariate or multivariate Gaussian distribution, which is a common case in macroeconomic application, we say about normal HMM.

The problem, which in natural way arises when one applies HMM models to analyze business cycles, is as follows. Having information about the realization of observable variables in some period of time (say from $t = 1$ to $t = T$), one could try to estimate the state of unobservable MC at a fixed time $n \leq T$. The most common approach is to use the smoothed probability:

$$P(X_n = i | Y_1 = y_1, \dots, Y_T = y_T), \tag{1}$$

or the filtered probability:

$$P(X_n = i | Y_1 = y_1, \dots, Y_n = y_n) \tag{2}$$

to deal with this problem.

To estimate the path of MC some kind of ‘step by step decoding’ is processed. The state with the highest smoothed probability (respectively, filtered probability) is assigned to the particular time point n for $n \leq T$.

In contrast to the method mentioned above, certain kind of ‘global decoding’ is possible. Instead of a single point of time one could consider the whole period covered by the analysis and look for the most likely path of MC. To be more precise, the sequence $(\tilde{x}_1, \tilde{x}_2,, \dots, \tilde{x}_T) \in S^T$ satisfying the following condition:

$$\begin{aligned} P(X_1 = \tilde{x}_1, X_2 = \tilde{x}_2, \dots, X_T = \tilde{x}_T | Y_1 = y_1, \dots, Y_T = y_T) \\ = \max_{(x_1, x_2, \dots, x_T) \in S^T} P(X_1 = x_1, X_2 = x_2, \dots, X_T = x_T | Y_1 \\ = y_1, \dots, Y_T = y_T) \end{aligned} \tag{1}$$

is the object of interest.

To estimate the parameters the well-known Baum-Welch algorithm is used (Baum *et al.*, 1970). The results of this deterministic method strongly depend on initial values of the parameters. Therefore, they may be far from optimal. In order to increase the chances of finding the optimal solution, the calculation can be repeated several times for the same set of data and different initial values. For *k*-state HMM model preselecting of the following values is required (see Bernardelli, 2013):

- initial distribution of an unobserved Markov chain (*k* parameters),
- transition probabilities of an unobserved Markov chain k^2 parameters,
- means and variances of conditional distribution of an observed variable (*2k* parameters).

In our research the initial values are chosen randomly using independent and identically distributed draws from a univariate distribution. The number of draws used for parameters estimation of the time series being under study varies between 1.000 and 10.000. The number of trial's repetitions depends on the number of HMM's states and numerical stability of computations.

The best estimates of parameters of models are chosen with selection criteria taking into account the following indicators:

- Akaike's information criterion (AIC),
- Bayesian information criterion (BIC),
- the log likelihood value,
- frequency of obtaining certain solution of the Baum-Welch algorithm (with an accuracy of one decimal place).

The HMM model, considered as the best for the particular input data set, is used to compute the most likely path which consists of a sequence of states of MC (throughout the whole period under consideration). This path is an output of the Viterbi algorithm. It is worth noting that despite of deterministic nature of the algorithm, the method of 'decoding' states of unobserved MC as a whole has a non-deterministic character. The block diagram of the proposed algorithm is presented in Figure 1.

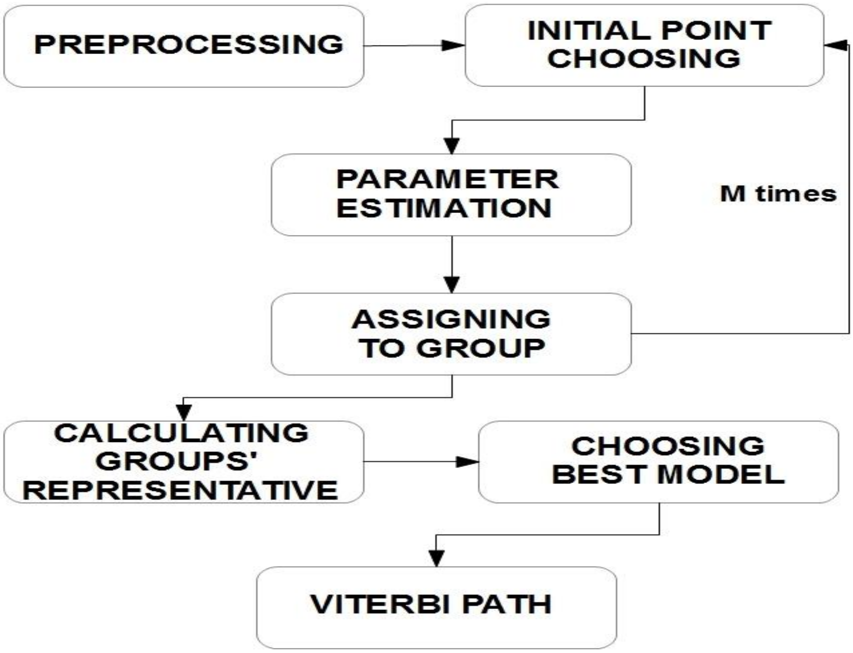

Figure 7. Schema of the algorithm of business cycle turning points detection.

Source: own compilation.

The primary step in the analysis is the defragmentation of time series into two types of periods: those associated with relatively good conditions and those which are rather connected with worse situation. To conduct such classification we consider normal HMM with state space of the form $S = \{0, 1\}$. An observable component Y_t corresponds to economic time series being under the study and the following condition must be satisfied:

$$Y_n|_{X_n=0} \sim N(\mu_0, \sigma_0) \text{ and } Y_n|_{X_n=1} \sim N(\mu_1, \sigma_1), \tag{4}$$

where $\mu_0 < \mu_1$. The state 0 corresponds to those points of time in which survey respondents report deterioration of situation, while the state 1 is associated with improvement. The most likely path of MC reflects changes in business activity in the two-scale strict separation of the states. We refer to this model using notation HMM(2). An underlying Markov chain is denoted by MC(2).

As mentioned, we would like to distinguish periods of unclear or difficult to evaluate situation. For this purpose we introduce a Markov chain with an extended state space $S = \left\{0, \frac{1}{2}, 1\right\}$. The state $\frac{1}{2}$ should therefore correspond to such uncertain, transient period. The meaning of the states 0 and 1 is the same as in the standard two-state model. An extended three-state model is defined as follows:

$$Y_n|_{X_n=i} \sim N(\mu_i, \sigma_i), \tag{5}$$

for $i = 0, \frac{1}{2}, 1$, where $\mu_0 < \mu_{\frac{1}{2}} < \mu_1$. Additionally we assume that $p(0,1) = p(1,0) = 0$ to reflect smoothing of changes. This model is denoted by HMM(3), while an unobservable MC by MC(3).

To carry out the more precise classification, the third model, denoted by HMM(4), was constructed. To distinguish definitely good periods, worse but still positive, definitely bad and moderately bad ones, the four-level scale should be taken into consideration. Above assessments are associated respectively with states 1, $\frac{2}{3}$, 0 and $\frac{1}{3}$ of MC. Therefore HMM model is introduced as follows:

$$Y_n|_{X_n=i} \sim N(\mu_i, \sigma_i), \tag{6}$$

for $i = 0, \frac{1}{3}, \frac{2}{3}, 1$, where $\mu_0 < \mu_{\frac{1}{3}} < \mu_{\frac{2}{3}} < \mu_1$. As in the case of the second, three-state model, we assume that only transitions between adjacent states are possible, so:

$$p(0,1) = p(1,0) = p\left(0, \frac{2}{3}\right) = p\left(\frac{2}{3}, 0\right) = p\left(\frac{1}{3}, 1\right) = p\left(1, \frac{1}{3}\right) = 0, \tag{7}$$

We want to emphasize that we do not claim the HMM model is the data generating process. In our research we simply treat turning point detection in terms of pattern recognition.

3. Results of empirical analysis

3.1. Input time series

The paper applies models and techniques described in the previous section to data of the business tendency survey in the Polish manufacturing industry, which is conducted by the Research Institute for Economic Development, Warsaw School of Economic, on the monthly basis. In this survey the respondents evaluate changes in certain areas of business activity. They give answers to eight questions. For every question there are three possible options to choose from: increase/improve, decrease/worsen or no change. A balance is calculated as a difference between percentages of positive (increase/improve) and negative (decrease/worsen) answers. For each category of activity the respondents assess current and future changes. We show that by simply transforming the balance time series it may signal a turning point in business activity. In our research we mainly focus on the balance of production (Q1, see Figure 2). However, we pay also some attention to the balance of finished goods inventories (Q4, see Figure 3). Basic descriptive statistics for those balances are given in Table 1.

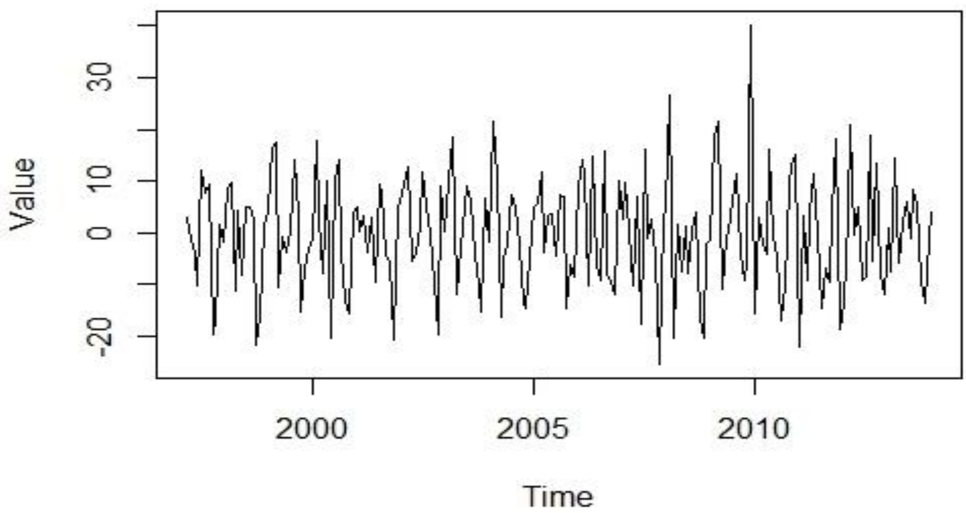

Figure 8. Time series with balances of positive and negative answers for the question about Level of production.

Source: own computation.

Moreover, we analyze the time series of sold production in constant prices (data from the Central Statistical Office of Poland) in order to check whether our results correspond to the 'real world' data. Depending on the context, the data sample covers March 1997 to February 2014, or January 2006 to February 2014.

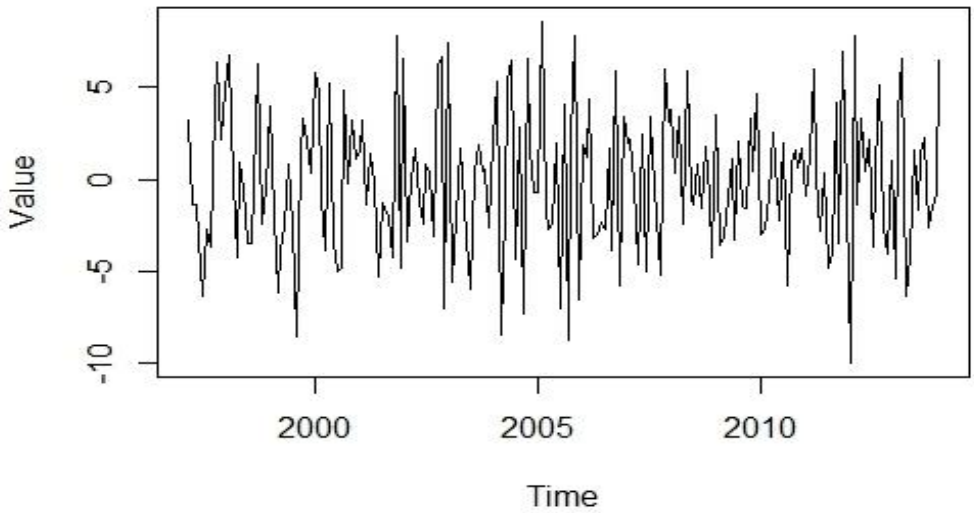

Figure 9. Time series with balances of positive and negative answers for the question about Stocks of finished goods.

Source: own computation.

Table 3. Descriptive statistics for the questions about Level of production and Stocks of finished goods.

Statistics	Q1	Q2
Minimum	-25.30000	-10.00000
1-quantile	-7.95000	-2.75000
Median	0.30000	-0.10000
Mean	-0.08670	-0.03005
3-quantile	6.90000	2.55000
Maximum	40.10000	8.60000

Source: own calculations.

The idea of the extended state space of an unobservable MC has already been examined in Bernardelli & Dędys (2012). We have transformed there the balance of production into the most likely path of MC(3). We have also constructed some kind of a compound three-state reference time series. In order to achieve this here, besides having reference turning points dated by OECD, we need their dating from another source. We decided to use one made by Drozdowicz-Bieć (2008). It seems that periods for which there exist discrepancies between researchers could be paired with the 'uncertain' state ½ of MC. The states 0 and 1 are assigned to consistent indications. That is, if in both component time series contraction is found in the same period, then the constructed compound time series takes on the 0 state. And if there is an agreement in each of the two time series about expansion in a particular period, then the state 1 is assumed to appear in the compound reference time series. Figure 4 presents the results of comparison of the Viterbi path of the model for Q1 with the constructed compound reference time series.

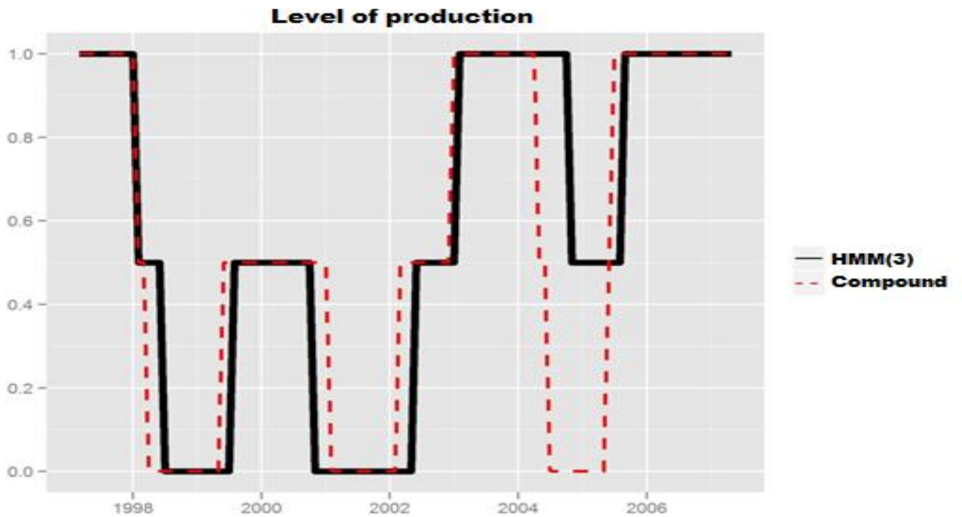

Figure 10. Comparison of the compound reference time series with the Viterbi path for 3-state HMM for Q1.

Source: own computation.

3.2. Period March 1997 – February 2014

The first part of research covers the period from March 1997 to February 2014. For reference we use turning points dated by OECD. Figure 6 presents the results of applying HMM(2) to the Q1 balance. The unobservable MC(2) signals almost all the turning points but with some delay of one to six months. Unfortunately, one phase of contraction has been missed. Two troughs are signaled in advance. Figure 7 shows the most likely path of MC(3) for the same balance. It seems that introducing the third state may improve the detection of turning points. MC(3) does signal the missing contraction phase. It sends, however, some false signals about expansions. It is worth noting that MC(3) leads almost all peaks. The application of HMM(4) seems to give even better results (see Figure 8). The signal of missed contraction is strengthened, while false information about expansion is weakened. All noted downturns of the reference series are led. The comparison of all: two-, three- and four-state HMMs is illustrated by Figure 5.

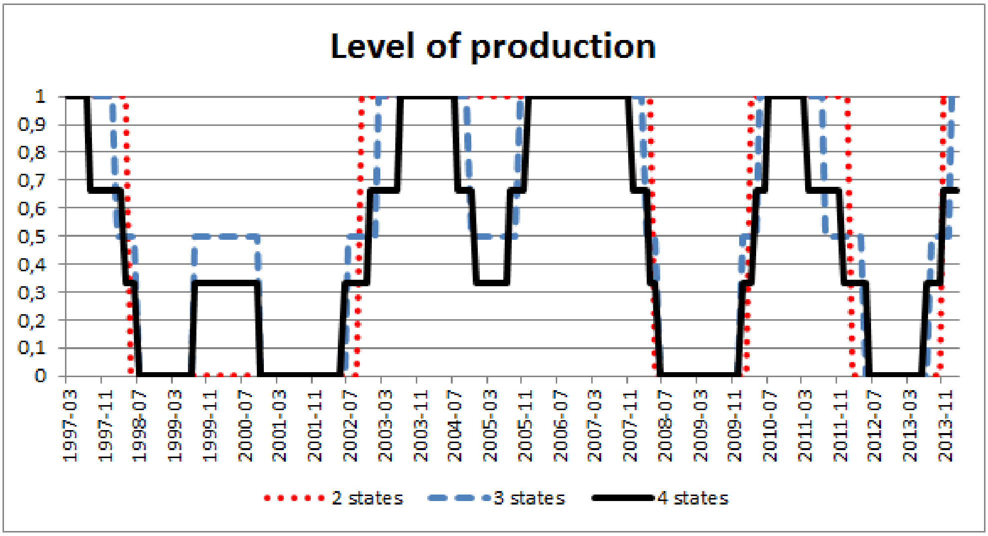

Figure 11. Comparison of the Viterbi paths for HMM(2), HMM(3) and HMM(4) for Q1.

Source: own computation.

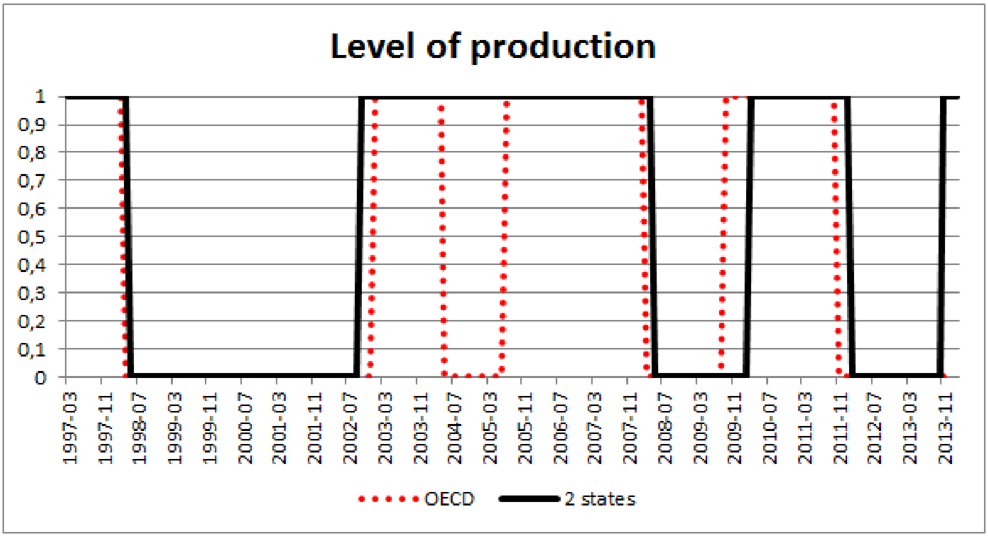

Figure 12. Comparison of the OECD reference time series with the Viterbi path for HMM(2) for Q1.

Source: own computation.

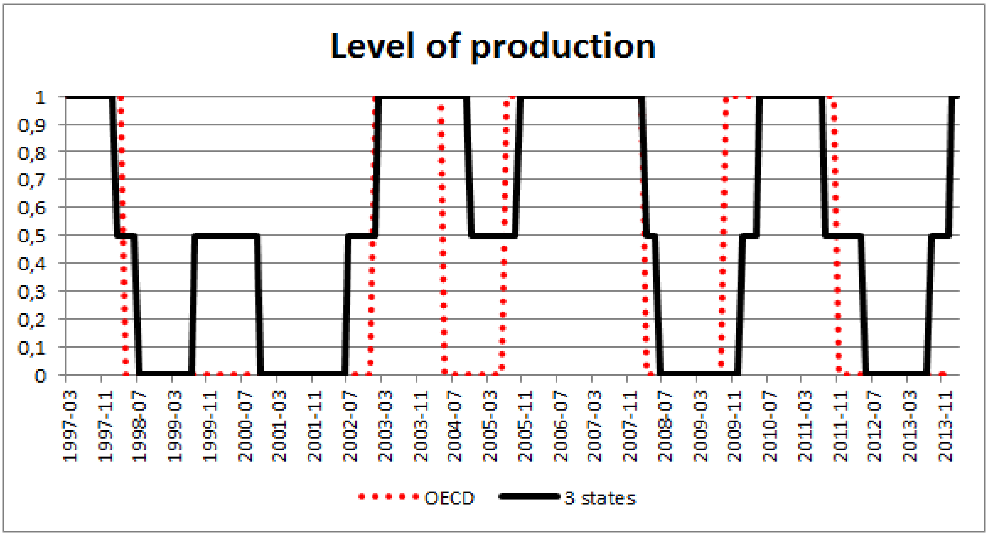

Figure 13. Comparison of the OECD reference time series with the Viterbi path for HMM(3) for Q1.

Source: own computation.

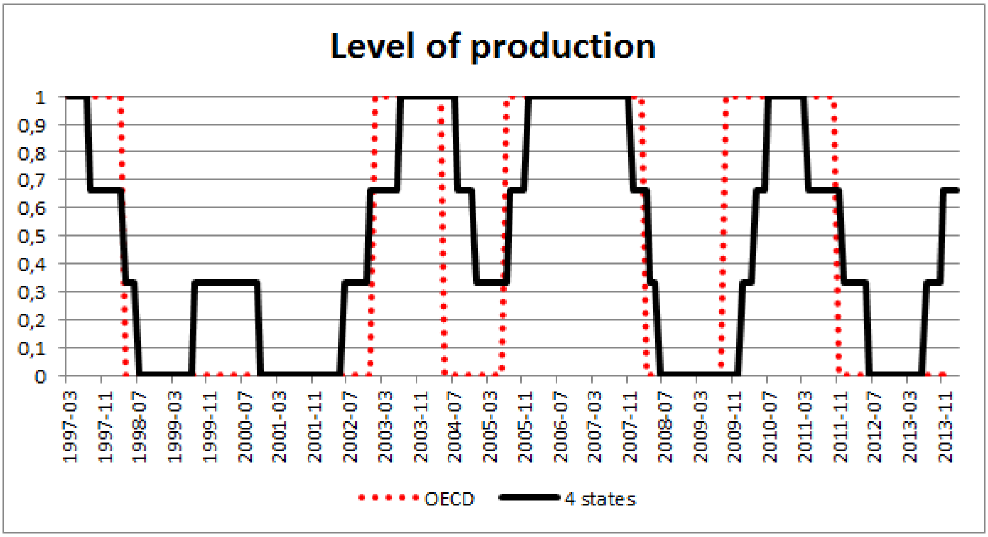

Figure 14. Comparison of the OECD reference time series with the Viterbi path for HMM(4) for Q1.

Source: own computation.

The results of the decomposition of the Q4 balance are presented in Figure 9. In fact the balance is multiplied by minus one before processing to match the reference time series. While comparing the most probable path of MC(2) with the reference time series (see Figure 10), one can distinguish two periods with different accuracy in detecting the turning points. In the first period, up to the end of 2005, the turning points are signaled with a lag of 4-6 months, while after 2005 all the turning points are led by 3-4 months or coincident. As in the case of Q1, HMM(3) and, specifically, HMM(4) perform better (see Figures 11 and 12).

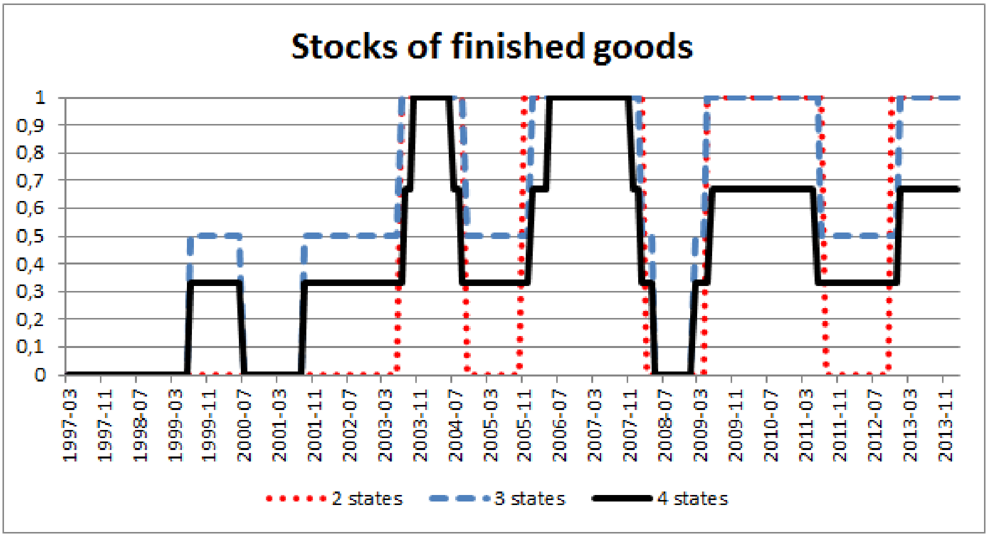

Figure 15. Comparison of the Viterbi paths for HMM(2), HMM(3) and HMM(4) for Q4.

Source: own computation.

The question occurs: can the balances of future situation improve the detection of turning points? The answer seems to be negative. In Figure 13 there are presented results of applying HMM(2) to the balances of predicted production. Only the results obtained for the period after year 2005 seem to be satisfactory. Before 2005 all downturns are missed, and afterwards the upper turning points are signaled with a lag. Figure 14 shows the most likely path of MC(2) fitted to the balance of expected finished goods inventories. In this case the estimates obtained for the period up to the year 2006 are closer to the reference time series. All the turning points after that year are not, however, captured by the algorithm, and the turning points before 2006 are lagged.

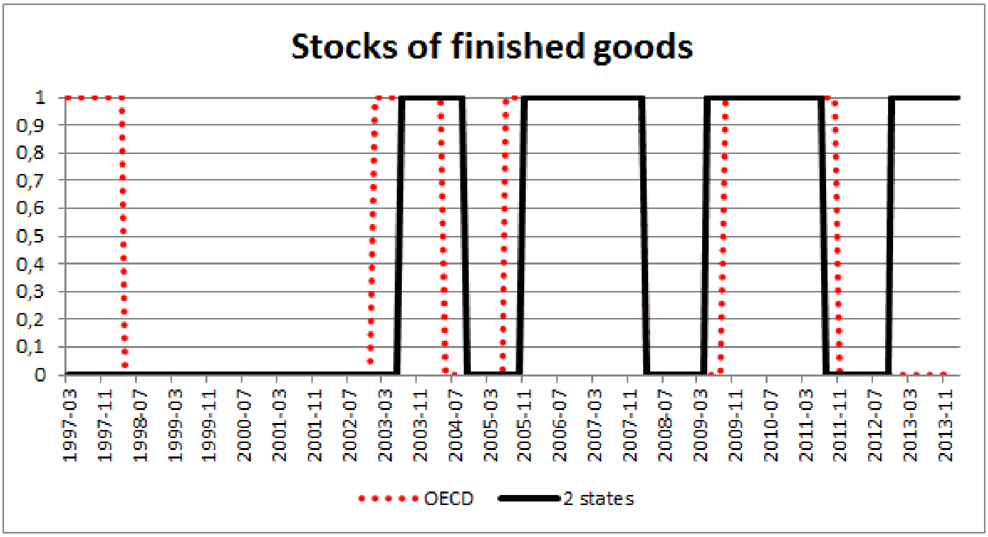

Figure 16. Comparison of the OECD reference time series with the Viterbi path for HMM(2) for Q4.

Source: own computation.

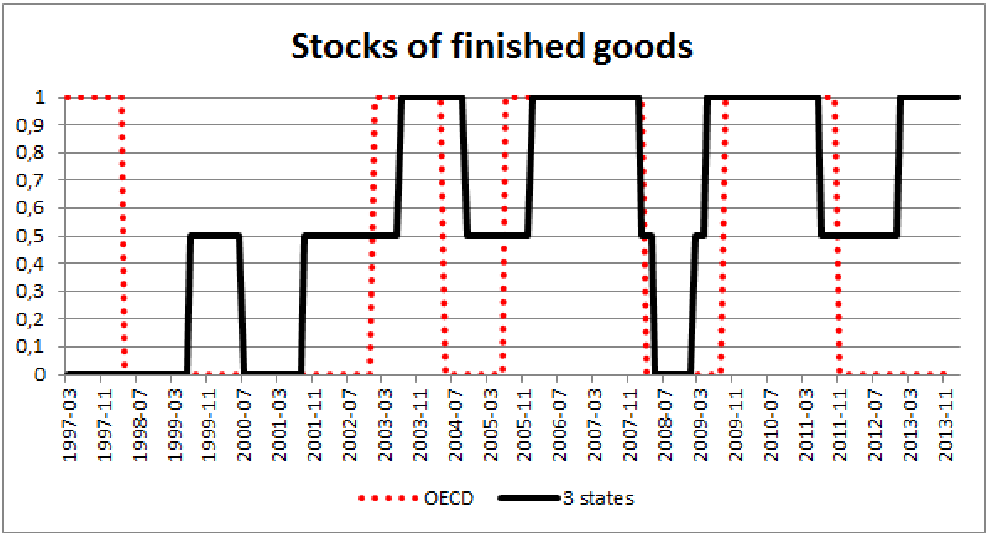

Figure 17. Comparison of the OECD reference time series with the Viterbi path for HMM(3) for Q4.

Source: own computation.

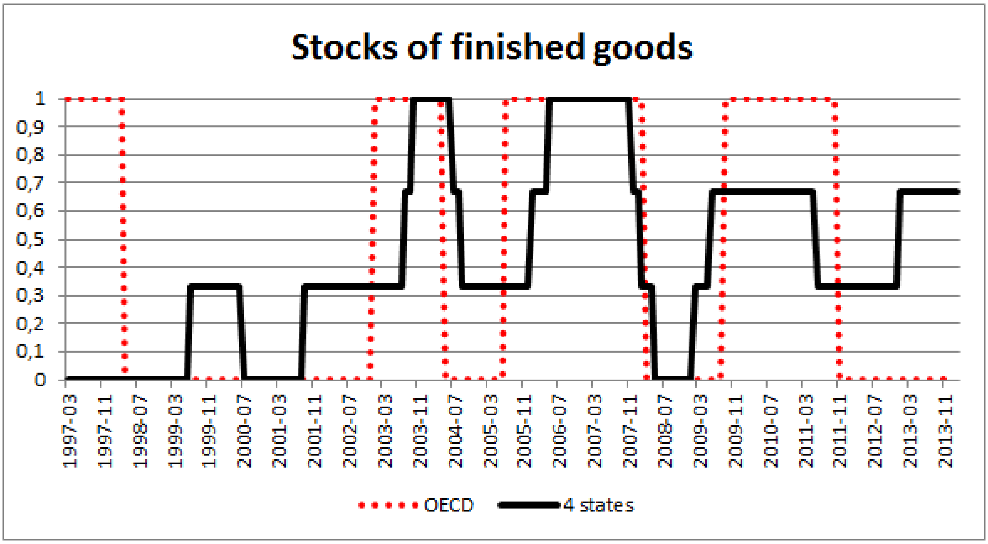

Figure 18. Comparison of the OECD reference time series with the Viterbi path for HMM(3) for Q4.

Source: own computation.

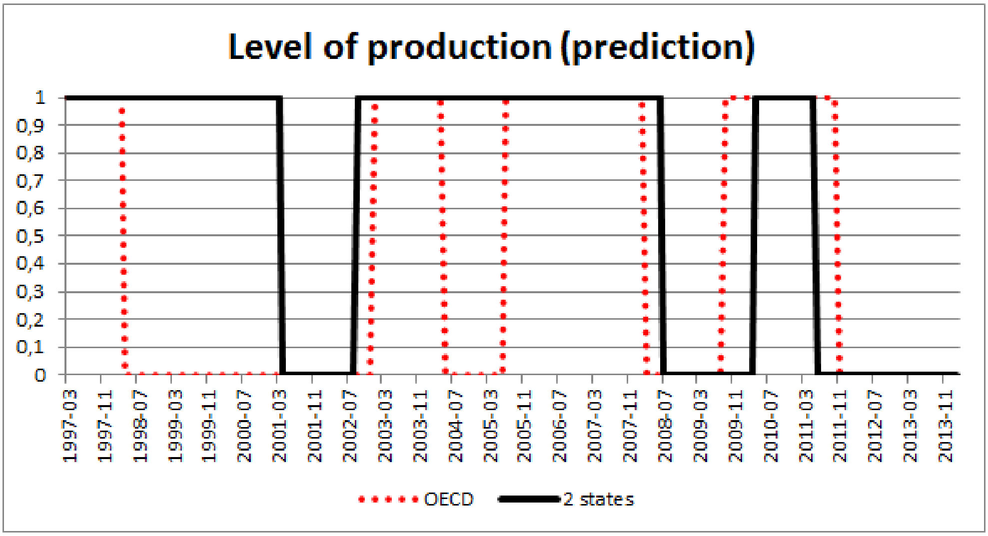

Figure 19. Comparison of the OECD reference time series with the Viterbi path for HMM(2) for Q1 (predicted).

Source: own computation.

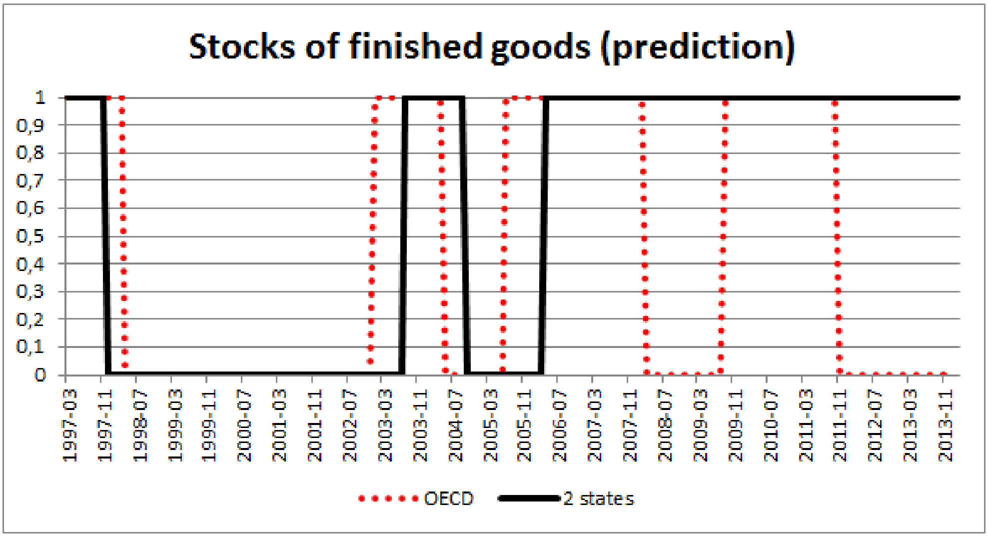

Figure 20. Comparison of the OECD reference time series with the Viterbi path for HMM(2) for Q4 (predicted).

Source: own computation.

3.3. Period January 2006 – February 2014

In the second part of the study we focus on the period from January 2006 to February 2014. In this period all turning points are reflected by the most likely paths of MC. The procedure of estimation is repeated for the Q1 and Q4 balances. In addition, they are confronted with 'hard' economic data – the index of sold manufacturing production (in constant prices).

The most likely path of Markov chain fitted to the Q1 balance are drawn in Figures 15-17. There is no surprise that the results slightly differ from the previous ones. Almost all of the previously observed properties of decomposition are present in new output. As before, downturns are signaled at the same time or in advance. And multistate MCs, especially MC(3), help to detect the reference turning points more accurately and with a lead.

The Viterbi paths obtained for the sold production index are presented in Figures 18-20. All the turning points are detected properly. As in previous cases, one can observe that the most likely paths of MC(3) and MC(4) are leading. Furthermore, the peaks are signaled earlier.

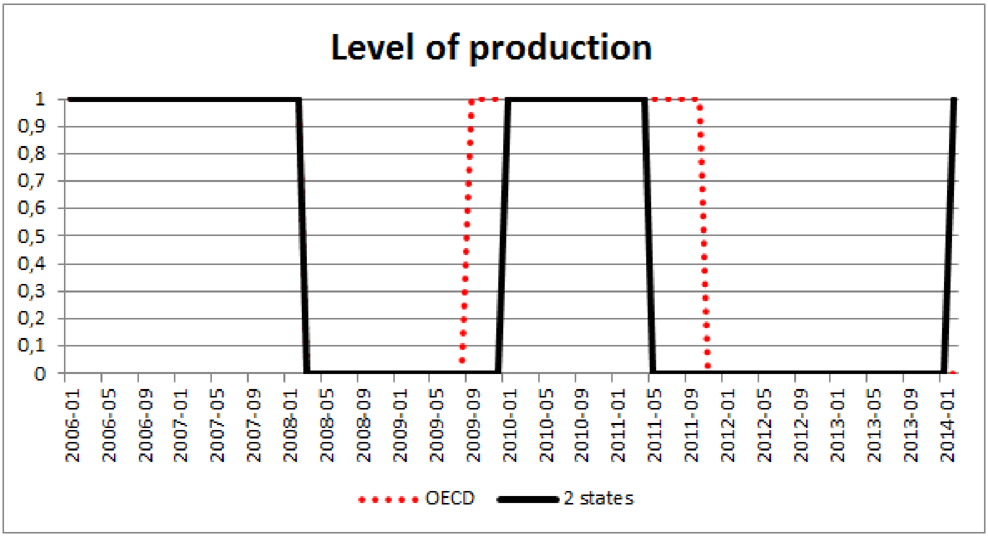

Figure 21. Comparison of the OECD reference time series with the Viterbi path for HMM(2) for Q1.

Source: own computation.

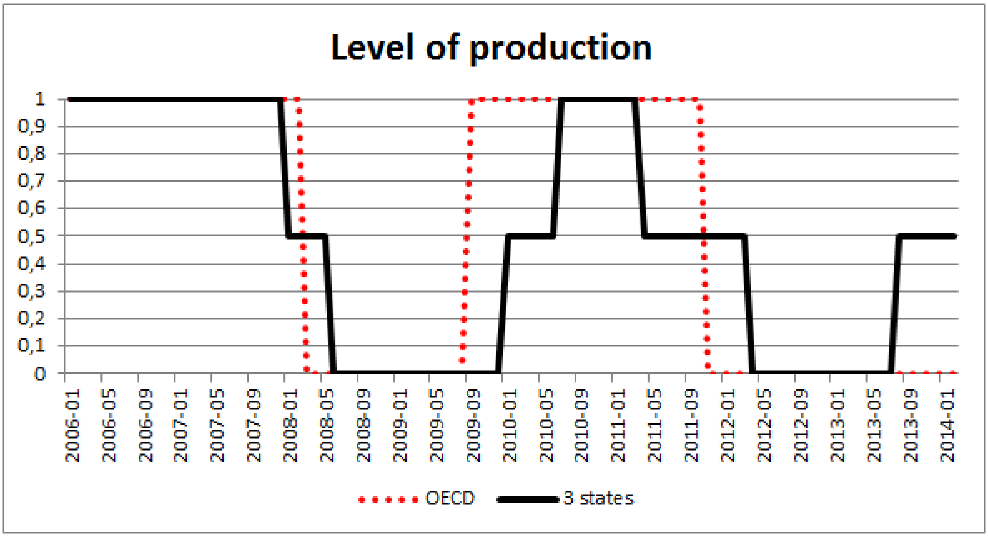

Figure 22. Comparison of the OECD reference time series with the Viterbi path for HMM(3) for Q1.

Source: own computation.

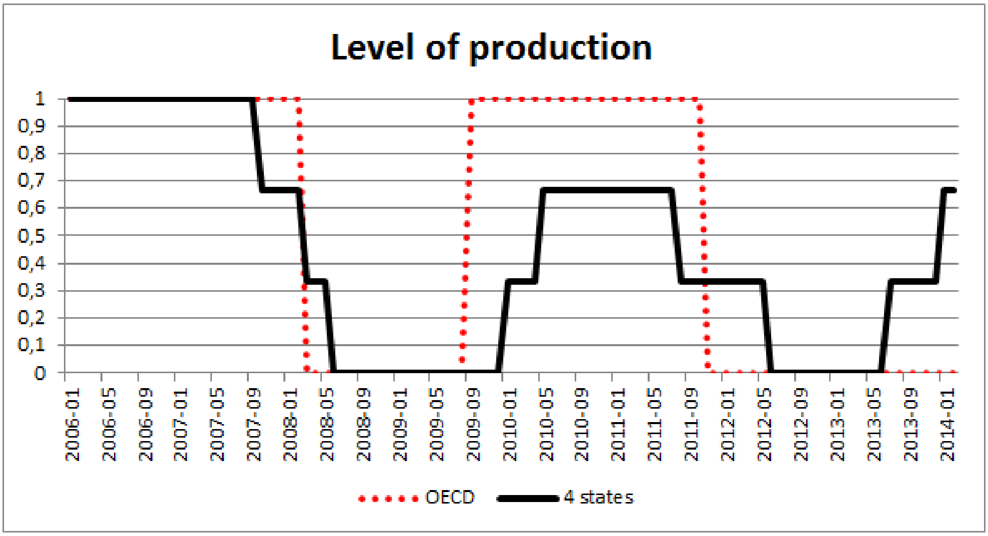

Figure 23. Comparison of the OECD reference time series with the Viterbi path for HMM(4) for Q1.

Source: own computation.

Figure 24. Comparison of the OECD reference time series with the Viterbi path for HMM(2) for the index of sold production.

Source: own computation.

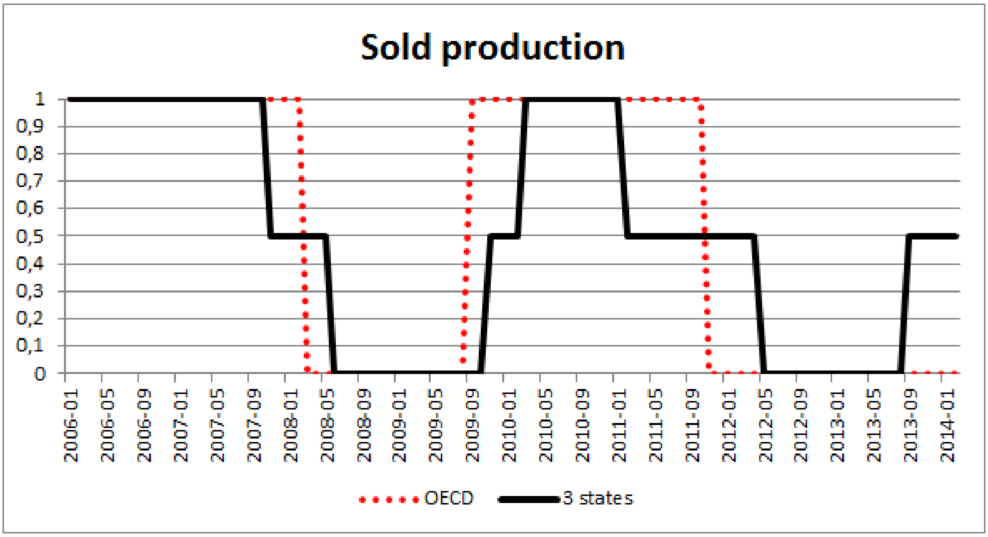

Figure 25. Comparison of the OECD reference time series with the Viterbi path for HMM(3) for the index of sold production.

Source: own computation.

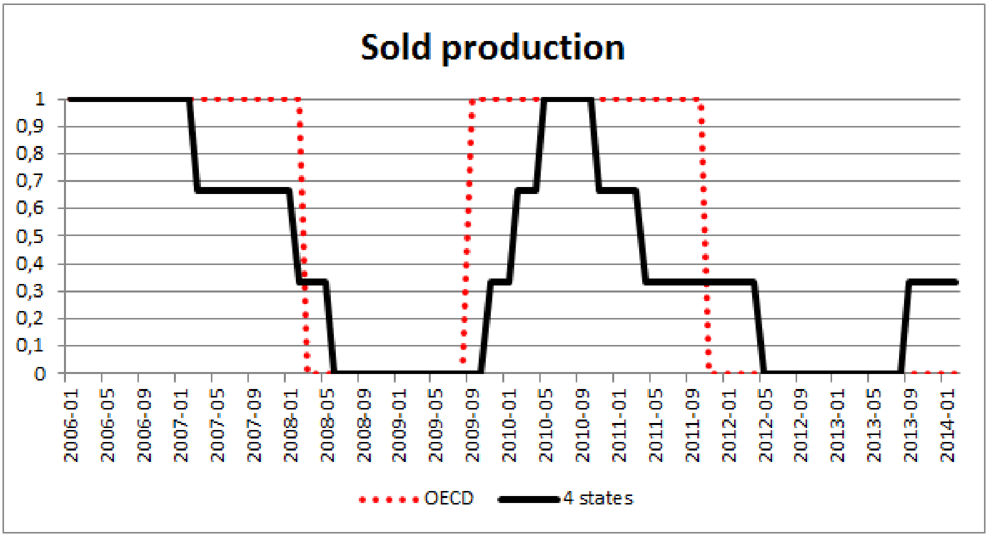

Figure 26. Comparison of the OECD reference time series with the Viterbi path for HMM(4) for the index of sold production.

Source: own computation.

The comparison of the most likely paths of MC with the index of sold manufacturing production and the Q1 balance (current) leads to the conclusion of high accuracy of the survey respondents' assessments. Figure 21 illustrates the Viterbi paths for the two-state MC. One can observe almost ideal fit. The paths differ only at the last point of time. The most likely paths obtained for HMM(3) seem to be very close as well (see Figure 22). In the case of the four-state MC the fit is not good alike, but in our opinion the accuracy is still satisfactory (see Figure 23). It is difficult to unambiguously decide on which pattern fits the best to the reference turning points. On the one hand, the survey balances are generally slightly lagging behind the index of sold manufacturing production. On the other hand, they signal the reference turning points with higher accuracy.

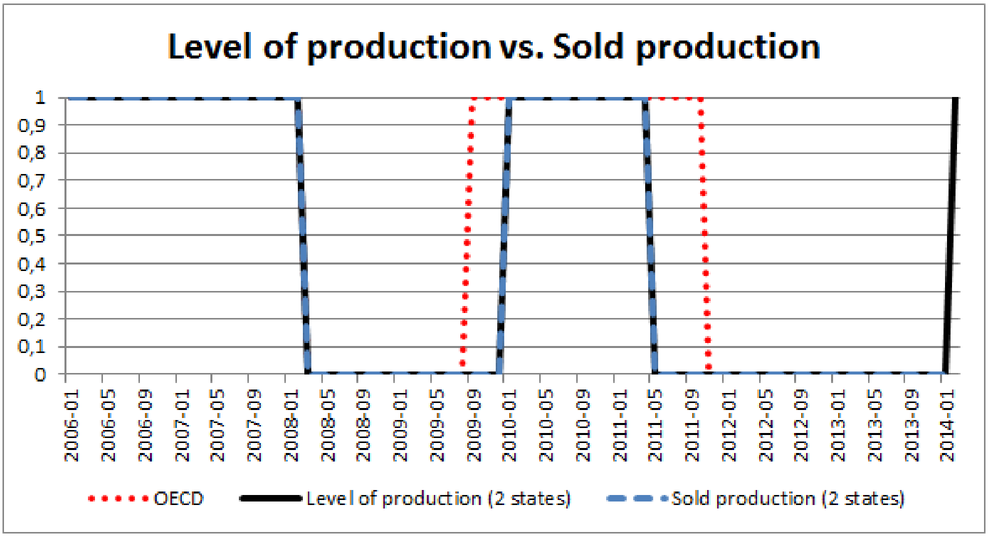

Figure 27. Comparison of the OECD reference time series with the Viterbi paths for HMM(2) for Q1 and the index of sold production.

Source: own computation.

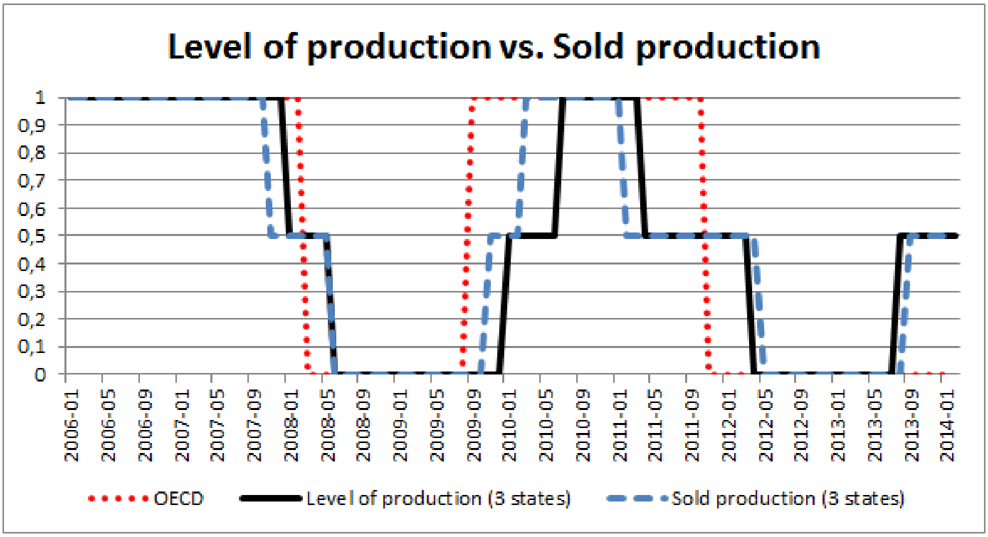

Figure 28. Comparison of the OECD reference time series with the Viterbi paths for HMM(3) for Q1 and the index of sold production.

Source: own computation.

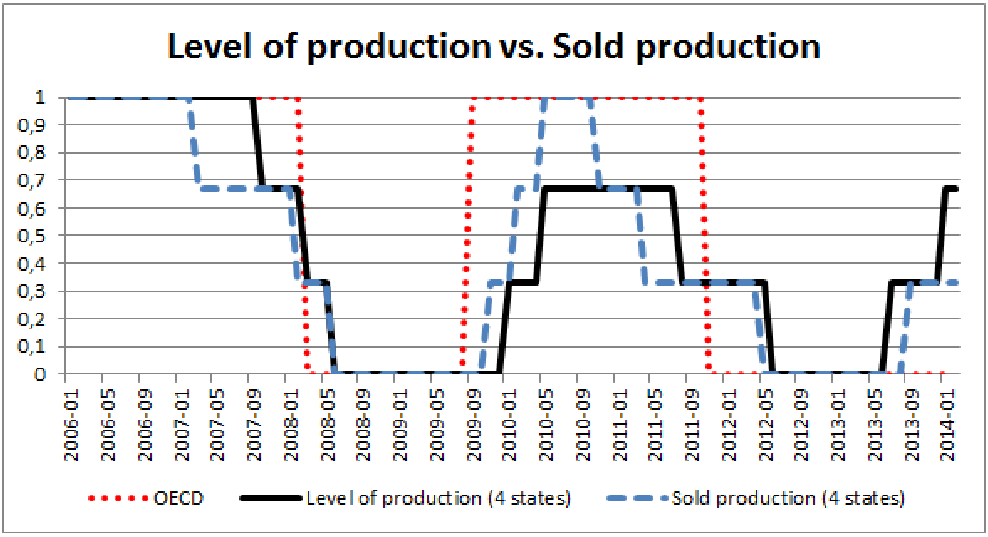

Figure 29. Comparison of the OECD reference time series with the Viterbi paths for HMM(4) for Q1 and the index of sold production.

Source: own computation.

4. Conclusions

The analysis shows that the presented method of time series decomposition can be really useful in detecting turning points, especially when multistate underlying Markov chains are considered. The study also found the business tendency survey data, gathered by the Research Institute for Economic Development, Warsaw School of Economics, to be satisfactory indicators of changes in real business activity. Specifically, it is worth to emphasize that downshifts of the reference time series were, in general, signaled with a lead. The preliminary results also suggest that in detecting turning points the balances of current situation play a more important role than the balances of predictions. The problem of the optimal sample size calls for further examination. The results encourage to real-time analysis.

References

Abberger, K., Nierhaus, W. (2010). Markov-switching and the Ifo business climate: the Ifo business cycle traffic lights, *Journal of Business Cycle Measurement and Analysis*, 2, 1-13.

Addo, P. M., Billio, M., Guegan, D. (2012). Alternative methodology for turning point detection in business cycle: A wavelet approach, *Documents de travail du Centre d'Economie de la Sorbonne*, 23.

Artis, M. J., Krolzig, H-M., Taro, J. (1998). *The European business cycle*, CEPR Discussion Paper No. 2242.

Baum, L. E., Petrie, T., Soules, G., Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, *The Annals of Mathematical Statistics*, 41 (1), 164-171.

Bell, W. (1984). Signal extraction for nonstationary time series, *Annals of Statistics*, 12 (2), 646-664.

Bernardelli, M. (2013). Non-classical Markov models in the analysis of business cycles in Poland, *Roczniki Kolegium Analiz Ekonomicznych SGH*, 30, 59-74.

Bernardelli, M., Dędys, M. (2012). Hidden Markov models in analysis of results of business tendency surveys, *Prace i Materiały Instytutu Rozwoju Gospodarczego SGH*, 90, 159-181.

Birchenhal, C. R., Jessen, H., Osborn, D. R., Simpson, P. (1999). Predicting US business-cycle regimes, *Journal of Business and Economic Statistics*, 17 (3), 313-323.

Boldin, M. (1994). Dating turning points in the business cycle, *Journal of Business*, 67 (1), 97-131.

Çakmaklı, C., Paap, R., van Dijk, D. (2013). Measuring and predicting heterogeneous recessions, *Journal of Economic Dynamics & Control*, 37 (11), 2195-2216

Cappé, O, Moulines, E., Rydén, T. (2005). *Inference in hidden Markov models*, New York: Springer.

Chin, D., Geweke, J., Miller, P. (2000). *Predicting turning points*, Technical Report 267, Reserve Bank of Minneapolis.

Christiano, L. J., Fitzgerald, T. J. (2003). The band pass filter, *International Economic Review*, 44 (2), 435-465.

Cleveland, W. S. (1972). *Analysis and forecasting of seasonal time series*, PhD thesis, University of Wisconsin-Madison.

Drozdowicz-Bieć, M. (2008), Od recesji do boomu. Wahania cykliczne polskiej 1990-2007, Prace i Materiały Instytutu Rozwoju Gospodarczego SGH, 80, 25-40.

Hamilton, J. D. (1994). *Time series analysis*, Princeton.

Hodrick, R. J., Prescott, E. C. (1997), Postwar US business cycles: An empirical investigation, *Journal of Money Credit and Banking*, 29 (1), 1-16.

Koskinen, L., Oeller, L. E. (2004). A classifying procedure for signaling turning points, *Journal of Forecasting*, 23 (3), 197-214.

Lamy, R. (1997). *Forecasting US recessions: Some further results from probit models*, Technical report, Finance Canada.

Mitra, S., Date, P. (2010). Regime switching volatility calibration by the Baum-Welch method, *Journal of Computational and Applied Mathematics*, 234, 3243-3260.

Sensier, M., Artis, M., Osborn, D. R., Birchenhall, C. R. (2004). Domestic and international influences on business cycle regimes in Europe, *International Journal of Forecasting*, 20 (2), 343-357.

Wildi, M., Schips, B. (2005). *Signal extraction: How (in)efficient are model-based approaches? An empirical study based on TRAMO-SEATS and Census X12-ARIMA*, Technical Report 96, Zurich: KOF Swiss Economic Institute.