Michał Bernardelli[‡]

# The procedure of business cycle turning points identification based on hidden Markov models

## Abstract

In the paper the procedure, based on hidden Markov chains with conditional normal distributions and uses algorithms such as time series decompositions (STL), Baum-Welch algorithm, Viterbi algorithm and Monte Carlo simulations, is proposed to analyze data out of the business tendency survey conducted by the Research Institute for Economic Development, Warsaw School of Economics. There are considered three types of models, namely, with two-state, three-state and four-state Markov chains. Results of the procedure could be treated as an approximation of business cycle turning points.

The performed analysis speaks in favor of multistate models. Due to, an increasing with the number of states, numerical instability, it is not obvious which model should be considered as the best one. For this purpose various optimization criteria are taken into consideration: information criteria (AIC, BIC) and the maximum-likelihood, but also frequency of obtaining a given set of parameters in the Monte Carlo simulations. The results are confronted with the turning points dated by OECD. The tested models were compared in terms of their effectiveness in detecting of turning points.

The procedure is a step into automation of business cycle analysis based on results of business tendency surveys. Though this automation covers only some models from millions of possibilities, the procedure turns out to be extremely accurate in business cycle turning points identification, and the approach seems to be an excellent alternative for classical methods.

Keywords: hidden Markov model, Viterbi algorithm, Baum-Welch algorithm, business tendency surveys, business cycle turning points

JEL classification: C63, C83, E37

---

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

# 1. Introduction

The analysis of business cycles is one of the primary sources of assessment of current and future economic situation. Certainly, the future level of economic development depends on many factors such as the gross domestic product, exports, rate of employment, level of production or other, often self-constructed indicators. Many different econometric methods are used to identify turning points. These are mainly ARIMA-based methods (Cleveland, 1972; Bell, 1984; Wildi & Schips, 2005) that are often used with the filters such as Hodrick-Prescott (1997), or Christiano-Fitzgerald (2003). Another class of econometric methods widely used in business cycles analysis is a logistic regression (Lamy, 1997; Birchenhall *et al*., 1999; Chin *et al*., 2000; Sensier *et al*., 2004). There is also a group of spectral methods based on the Fourier transform (see Addo *et al*., 2012). A construction of any econometric or spectral method, however, is problematic due to the bulk of various data as well as due to the potential presence of unspecified variables in developed models or simply restrictive assumptions about the model and input data. Thus, even unequivocal identification of turning points in an economy is not an easy task. As an alternative to these approaches, Markov models could be used (see Hamilton, 1994; Bhar & Hamori, 2004; Koskinen & Oeller, 2004; Mamon *et al*., 2007). Based on their non-deterministic character and weak assumptions, in many fields one can get at least comparable or often better result.

The paper describes a procedure to analyze data out of the business tendency survey in the manufacturing industry in Poland. The survey is conducted by the Research Institute for Economic Development, Warsaw School of Economics. The procedure is based on hidden Markov chains with conditional normal distributions and uses algorithms such as time series decompositions (STL), Baum-Welch algorithm, Viterbi algorithm and Monte Carlo simulations. There were considered models with two-state, three-state and four-state Markov chains. As an input not only answers to individual questions from the survey were analyzed, but also panel data were included, namely time series that consist of answers to a pair of survey questions. The more states, the better fit, but also the more numerical instability and longer time of computations. As an optimization criteria in the procedure, information criteria (AIC, BIC), maximum-likelihood and frequency of obtaining a given set of parameters in the Monte Carlo simulations are considered. After finding sets of parameters of suitable models the Viterbi path is calculated. It is the path of a state with the highest probability (due to the model parameters). Results of the procedure could be treated as an approximation of business cycle turning points. Obviously, it is sometimes necessary to consider a time delay between answers of survey respondents and changes in business activity, but for many of the examined input data it is a clear pattern and a strong premise to use it as a leading indicator. The results were confronted with the dating of business cycle turning points identified by OECD. The tested models were compared in terms of their effectiveness in detecting of coming changes in business situation. The study speaks in favor of multistate models. Furthermore, the use of panel data is justified, and in many cases recommended, due to the higher quality of the fitted model.

Although hidden Markov models are well-established in theory and practice of business cycles analysis (see Abberger & Nierhaus, 2010), their usage is mainly limited to two states. Generalization to multistate chains gives opportunity to deal with more flexible and efficient models. The second issue is the exploration of the Viterbi paths. The Viterbi algorithm is often used in, for example, pattern recognition and DNA sequencing but rather rarely exploited in macroeconomic applications. Such a merger between multistate hidden Markov chains and the Viterbi paths is innovative in the area of business tendency surveys analysis.

The paper is composed of six sections. The short description of hidden Markov models is given in Section 2. Section 3 presents the description of data, whereas Section 4 specifies the procedure. Section 5 presents the results from numerical experiments exploring the usefulness of the procedure. The paper sums up with conclusions in Section 6.

## 2. Hidden Markov models

Hidden Markov models (HMM) are widely used in analysis of processes and patterns in many fields. They are an excellent tool when one can distinguish two layers: one visible which is used to uncover the second, a hidden layer. Therefore it is common in pattern recognition: the first layer is an observed sequence of emissions, whereas the second layer is a sequence of states (symbols) which we do not know but need to discover. Application in speech, handwriting or gesture recognition (Jelinek, 1997) are well known examples. Markov models are also one of basic tools in analysis of data in bioinformatics (Durbin *et al*., 1998). In econometrics HMM are mostly used to analyze financial and macroeconomic time series (Cappé *et al*., 2005).

A hidden Markov model could be defined as a stochastic process (see Cappé *et al*., 2005). It could be also considered as the simplest dynamic Bayesian network (Ghahramani, 2001). It is possible to give an equivalent definition that uses the terminology from the field of finite-state probabilistic machine (or finite-state probabilistic automaton) (Rabin, 1963).

Let $S_X$ be a finite $k$-element set, so called the set of states, with the specified state $S_1$ treated as an initial state. We assume therefore that $k$ is greater than zero. In other words, we assume that the set of states is non-empty. Furthermore, let:

$$P = \left[p_{i,j}\right]_{i,j=1}^{k} \tag{1}$$

be a matrix of probabilities of transitions, where $p_{i,j}$ is the probability of transition from the state $i$ to the state $j$. It is assumed that the transition matrix is stochastic, that is for every $i$:

$$\sum_{j=1}^{k} p_{i,j}. \tag{2}$$

The Markov chain is an ordered triple ($S_X$, $S_1$, $P$). The characteristic feature of the Markov chain is so called lack of memory, which means that the next state depends only on current state but not on the whole history of getting to this state.

Hidden Markov models are known in mathematics and computer science as the probabilistic automaton. They are an extension of the Markov chain for an additional alphabet Σ, symbols of which are emitted in the specific state with the given probability distribution. We assume that in every state some symbol is emitted. For the finite alphabet the HMM in the state $i \in S_X$ is emitting the symbol $x \in \Sigma$ with the probability $e_i(x)$, and, next, it changes the state to $j$ with the probability $p_{i,j}$. In the case of continuous probabilities by $e_i(x)$ a probability distribution is meant, e.g. Gaussian. In both cases observable are only the symbols emitted by the model, but the current state of the hidden Markov chain remains unobservable (see Figure 1).

Hidden Markov chains with a $k$-element set of states are simply called $k$-state HMM. In the paper two-, three- and four-state models are considered. Also, an assumption on the probability distribution of emitting the symbol is taken. For every state symbols are emitting with normal distribution probability.

Each hidden Markov model thus is defined by these parameters:

- $k$ – number of states,
- set of symbols (alphabet) Σ, where $n$ is a number of symbols,
- initial probabilities for every state ($k$ parameters),

- transition matrix $P$, that is a matrix of probabilities of transitions between two states ($k^2$ parameters),
- parameters of normal distribution defining probability of emission of symbol in each state ($2kn$ parameters).

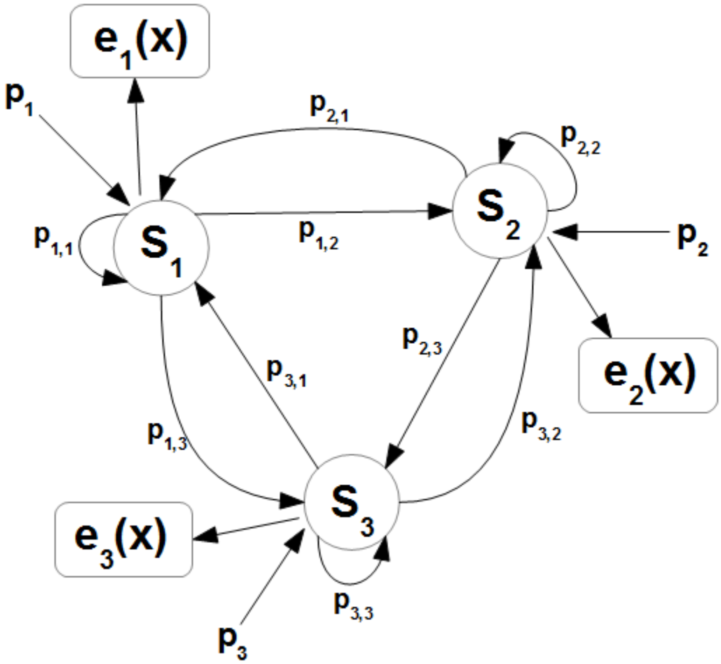


Figure 1. Scheme of a three-state hidden Markov model with a pair of normal probability distributions of emitting symbols.

Source: own compilation.

## 3. Description of the input data

The input data is balances taken from business tendency surveys in the manufacturing industry conducted monthly by the Research Institute for Economic Development, Warsaw School of Economics. Each month the survey consists of eight questions:

Q1 – volume of production
Q2 – volume of orders
Q3 – volume of export orders
Q4 – finished goods inventories
Q5 – selling prices of products

Q6 – level of employment
Q7 – financial standing
Q8 – general economic situation in Poland,

each one in two versions: retrospective, concerning what happened in the last 3-4 months ('AS-IS'), and prospective, concerning what is expected to happen in the next 3-4 months ('TO-BE'). For the calculations data from March 1997 to February 2014 were taken. Having analyzed results of numerical experiments, it has been found that models based on respondents' expectations are less accurate and worse fit than models including AS-IS balances. The same experiments (Bernardelli & Dędys, 2012) suggest that seasonal and random components should be filtered out of the input time series. Therefore the data were pre-processed. In order to decompose the raw time series the procedure STL from the R package was used. STL procedure is an implementation of an algorithm based on local weighted regression method called "loess" (see Cleveland, 1990). Figure 2 presents the decomposition of Q1 balance, where:

*data = seasonal + trend + irregular component (remainder),*

and Figure 3 illustrates the decomposition of Q7 balance.

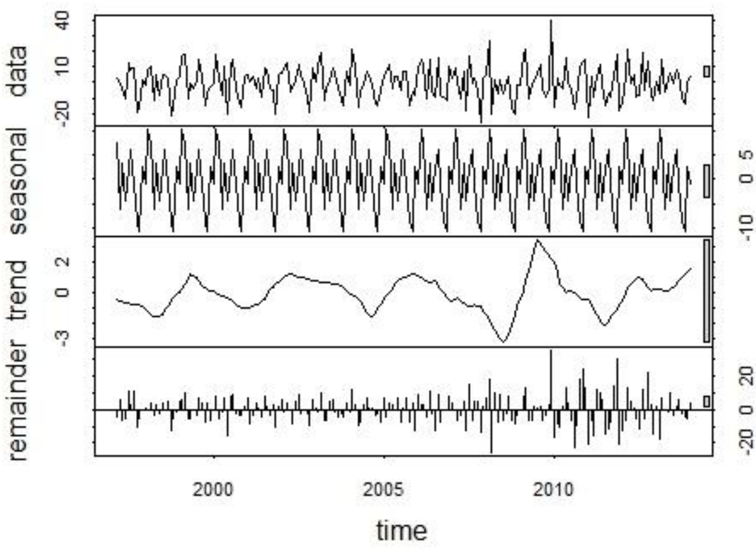


Figure 2. Time series decomposition with the STL procedure for Q1.

Source: own computation.

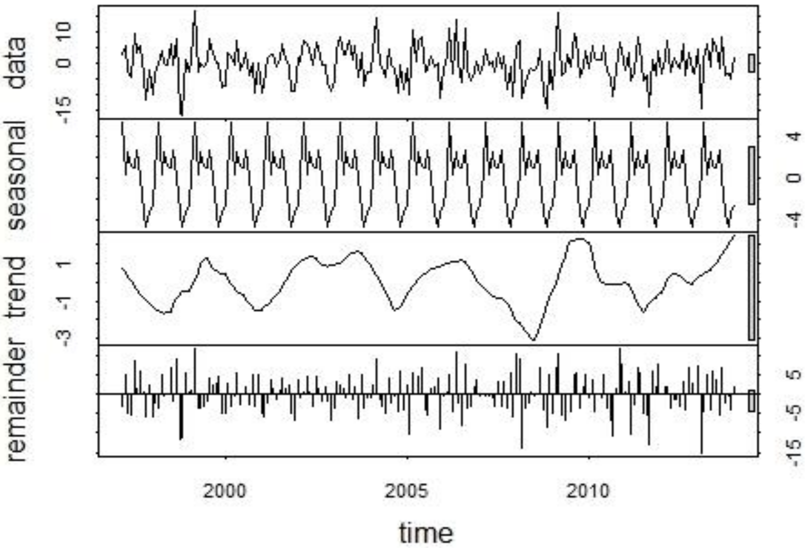


Figure 3. Time series decomposition with the STL procedure for Q7.

Source: own computation.

Basic descriptive statistics for these balances, before and after decomposing them, are given in Table 1.

## 4. Description of the procedure

The procedure takes on decomposed time series (only trend) and returns the path of states that has the highest probability in the whole considered period. For the sake of numerical stability (Bernardelli, 2012) and ease of interpretation computation was restricted to models with two, three and four states.

In the case of a two-state hidden Markov chain it is assumed that the zero state is associated with periods determined by the respondents as worse, while the state denoted by one is related to the situation assessed as better. In the case of three-state chains there is an additional state ½ symbolizing the transient situation between states 0 and 1. It is the state designed for situations uncertain and difficult to unambiguous classification. The space of states of four-state hidden Markov chains has the form {0, ⅓, ⅔, 1}. State 0 indicates strong economic downturn, state 1 indisputable economic recovery, while states ⅓ and ⅔ are transients. The state ⅓ should be interpreted as indicating the uncertain status of worse economic situation in the country, whereas the state ⅔ suggests rather better economic conditions.

Table 1. Descriptive statistics for the questions about Level of production and Financial standing.

| Statistics | Question 1 Level of production | | | |
|---|---|---|---|---|
| | data | seasonal | trend | remainder |
| minimum | -25.30 | -10.83 | -3.24 | -25.67 |
| 1-quantile | -7.95 | -5.28 | -0.79 | -5.52 |
| median | 0.30 | -0.94 | -0.01 | -0.07 |
| mean | -0.09 | -0.05 | -0.03 | -0.01 |
| 3-quantile | 6.90 | 3.10 | 0.76 | 5.16 |
| maximum | 40.10 | 10.07 | 3.47 | 35.14 |

| Statistics | Question 7 Financial standing | | | |
|---|---|---|---|---|
| | data | seasonal | trend | remainder |
| minimum | -16.20 | -4.74 | -3.09 | -15.21 |
| 1-quantile | -3.05 | -3.01 | -0.80 | -3.51 |
| median | 0.00 | 0.41 | 0.05 | -0.05 |
| mean | 0.053 | -0.003 | 0.033 | -0.015 |
| 3-quantile | 2.90 | 1.90 | 0.90 | 3.76 |
| maximum | 14.70 | 5.35 | 2.48 | 11.59 |

Source: own calculations.

The important assumption was made about probabilities of the transition matrix. Non-zero probabilities are permitted only between the adjacent states. This is the reflection of economically justified situation of gradual changes in the economy. Of course, this assumption is meaningless in the case of two states and makes sense only when the number of states is greater than two.

The procedure of business cycle turning points identification based on hidden Markov models can be described in the following steps:

(1) pre-processing the input data: choose the time series (a single or a pair of questions from the survey), and decompose them (using STL procedure);

(2) choose $M$ initial approximations of parameters of conditional normal distributions (parameters for every state and each of the input time series); initial points could be chosen randomly;

(3) for each point from the step (2) use the Baum-Welch algorithm to estimate parameters of the hidden Markov model; based on the expected values of conditional distributions find the correct order of states;
(4) group the parameters of all calculated models (at most *M* results) on the basis of – rounded to one decimal place – expected values of conditional distributions; for each group define a representative model with parameters being averages of the respective parameters of models from this particular group;
(5) for representative models from each group calculate the most probable path of a hidden Markov chain using the Viterbi algorithm;
(6) based on various optimization criteria or/and comparison with the reference time series choose the best HMM model.

Now let's present each step of the procedure in more detail way. The first step was described in the previous section. It is worth to emphasize that it is possible to take any combination of time series as an input – even answers to all eight questions from the survey. Although adding more data could improve model fitting, it is not a rule (Bernardelli, 2013b).

The second step of the procedure concentrates on choosing the right initial parameters to the model. For the *k*-state hidden Markov chain the following parameters need to be defined:

- initial probabilities for each of *k* states – in the procedure all equal to 1/*k;*
- the transition matrix *P* – with zero probabilities of transition between non-adjacent states only 4 + 3(*k* - 2) non-zero elements of the matrix need to be specify. By default probabilities are set as follows:

$$p_{i,j} = \begin{cases} \frac{1}{2} & i = 1, j = 1,2; \quad i = k, j = k-1, k \\ \frac{1}{3} & i = 2,3,\dots,k-1, j = i-1, i, i+1 \end{cases}. \tag{3}$$

In the procedure for *k* = 4 matrix has the form:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}; \tag{4}$$

- parameters $(\mu, \sigma)$ of independent normally distributed[1] $n$ random variables defining probability of emission of a symbol in each state, where $\mu = [\mu_1, \mu_2, \dots, \mu_n]^T$ is a vector of expected values and $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_n]^T$ – vector of standard deviations. There are $2kn$ parameters that determine the most, final values of model parameters. In the procedure these initial values are chosen from the following intervals: an expected value $\mu_{i,j} \in \left[\overline{\mu_{i,j}} - 3\overline{\sigma_{i,j}}, \overline{\mu_{i,j}} + 3\overline{\sigma_{i,j}}\right]$ and a standard deviation $\sigma_{i,j} \in \left[0.5\overline{\sigma_{i,j}}, 3\overline{\sigma_{i,j}}\right]$, where $\overline{\mu_{i,j}}$ and $\overline{\sigma_{i,j}}$ are empirical parameters calculated for every state ($i$=1,2,…,$k$) and each of input time series ($j$=1,2,…,$n$). Of course, intervals could be wider, but according to three sigma rule for the normal distribution in high probability they cover the vast majority of possible values.

For the computational purpose all intervals which contain values of possible parameters must be discretized. Let us consider in more detail discretization of parameters $\mu$ and $\sigma$. Let $m_{i,j}^{\mu}$ and $m_{i,j}^{\sigma}$ be numbers of nodes in the interval for respectively an expected value and a standard deviation of the $j$-th input time series of the $i$-th state. The number of all nodes $M$ in the discretization grid is defined by the formula:

$$m = \prod_{i=1}^{k} \prod_{j=1}^{n} m_{i,j}^{\mu} m_{i,j}^{\sigma}. \tag{5}$$

Mesh nodes may be distributed uniformly, but it is not always the best possible choice. Assuming that $n = 1$ and for all $i$ = 1,2,…,$k$ values $m = m_{i,1}^{\mu} = m_{i,1}^{\sigma}$ number of nodes in the grid for different numbers of states are given in Table 2. For a pair of questions ($n$ = 2), numbers of nodes are squares of numbers given in this table.

[1] One may consider other probability distributions than a multivariate normal distribution.

Table 2. Number of nodes for different sizes of discretization grid and different numbers of states in hidden Markov chain for a single question.

| $M$ | $k = 2$ | $k = 3$ | $k = 4$ |
|---|---|---|---|
| 2 | $1.60\times10^{1}$ | $6.40\times10^{1}$ | $2.56\times10^{2}$ |
| 5 | $6.25\times10^{2}$ | $15.6\times10^{3}$ | $3.91\times10^{5}$ |
| 10 | $10.0\times10^{3}$ | $10.0\times10^{5}$ | $10.0\times10^{7}$ |
| 20 | $1.60\times10^{5}$ | $6.40\times10^{7}$ | $25.6\times10^{9}$ |
| 50 | $6.25\times10^{6}$ | $15.6\times10^{9}$ | $39.1\times10^{12}$ |
| 100 | $10.0\times10^{7}$ | $10.0\times10^{11}$ | $10.0\times10^{15}$ |

Source: own calculations.

Each of the nodes is an initial point for calculations performed in the third step, in which the Baum-Welch algorithm is used to estimate HMM parameters. In the implementation the procedure *fit* from the depmixS4 library of the package R was used. The Baum-Welch algorithm is an iterative method that maximizes the expected value (Baum *et al.*, 1970). More precisely it is a representative of an Expectation-Maximization class of methods that calculates maximum likelihood. Due to the way of finding the maximum, the Baum-Welch algorithm should be classified as a greedy algorithm. Thus, obtained solutions may be far from optimal. There is no guarantee that the result is really a global maximum. Depending on the initial parameters the solution found by the algorithm may be only the local maximum. This is why the algorithm is used repeatedly for the same input data, but different initial parameters. Due to the high dimension of the grid, a number of nodes is increasing exponentially with an increasing number of states as well as with an increasing size of the panel data (see Table 2). The computation time is proportional to the number of nodes $M$. Therefore to get the result in a reasonable time the mesh used in the procedure must by rather thick. This is the reason why some random initial points are chosen. In this way the probability of finding local minimum that is not the global one is significantly decreased. This added randomness is in fact equivalent to the Monte Carlo approach and each use of the Baum-Welch algorithm to the Monte Carlo simulation.

Calculated parameters of a model, due to numerical rounding, are almost always unique. However the differences between parameters of two models could be really small, for example they can differ on the eighth decimal place. All models are assigning to groups on the basis of rounded to one decimal place expected values of conditional distributions. All respective parameters of models in a particular group are average. Parameters obtained in this way define the representative model in each group.

In the fifth step of the procedure for representative models from every group the most probable path of a hidden Markov chain using the Viterbi algorithm (Viterbi, 1967) is calculated. The implementation from the *posterior* procedure included in the package R was used. The Viterbi algorithm is an example of dynamic programming algorithm. The output data is the most likely sequence of hidden states which are commonly called the Viterbi path.

The purpose of the last step of the procedure is to choose the best model and, connected with it, the Viterbi path from representative models of all considered group. The choice could be made based on various optimization criteria as well as on comparison with the reference time series. The criteria used in the procedure are:

- Akaike information criterion (AIC),
- Bayesian information criterion (BIC),
- value of likelihood function,
- frequency of obtaining a given set of parameters in the Baum-Welch algorithm (size of each group).

The Viterbi paths were also compared with the reference time series, that is with dating of turning points in Poland evaluated by OECD. In order to verify the usefulness of the procedure numerical experiments were performed. Specification of the experiments and their results are described in the next section.

## 5. Numerical experiments

Many numerical experiments were conducted using the procedure described in the previous section. They were designed to answer research questions such as:

1) usefulness of respondents' expectations (Bernardelli & Dędys, 2012) – it turns out that TO-BE balances do not rather increase the accuracy of the detection of turning points;
2) impact of the number of states on the quality of the fit (Bernardelli & Dędys, 2012) – adding one or two states to the model seems to enrich business cycle analysis and the accuracy of the dating of turning points;

3) size of the panel data on input (Bernardell, 2013a) – increasing the number of input time series could improve, and in many cases it does, the quality of the business cycle approximation;
4) effect of optimization criteria for the quality of the fit (Bernardelli, 2013b) – procedure of the turning points identification should be treated as a multi-criteria optimization. Using information criteria doesn't always lead to an optimum hidden Markov model and the Viterbi path;
5) comparison of a non-deterministic version with the deterministic one (Bernardelli, 2014) – because of the so called 'the curse of dimension' Monte Carlo simulations are the only achievable way of getting the reliable fit in the reasonable time;
6) stability of the computations (Bernardelli, 2012) – numerical stability of the procedure worsens with the increasing number of states in the model, but – in comparison with the other Monte Carlo algorithms – should be considered as highly acceptable;
7) usefulness in other fields like transport analysis (Dorosiewicz, 2013) – the idea of the turning points identification procedure and the implementation of the whole algorithm was found very promising for other than business tendency surveys input data.

In the numerical experiments all the balances as well as all their pairs were examined. The analysis was focused on AS-IS balances. There were considered hidden Markov chains with two, three and four states. The size of the mesh was chosen such that the number of nodes was equal to 10.000. In addition 1.000 initial points was randomly chosen. It means, that the Baum-Welch algorithm (the third step) for every input data was executed $M$ = 11.000 times. In order to compare HMM path $x_t$ with the reference time series $r_t$ the following measure was used:

$$\rho(x,r) = \sum_{t=1}^{T} |r_t - x_t|, \qquad (6)$$

where $T$ is the length of the considered period. Lower values of the indicator $\rho$ mean closer similarity between the Viterbi path and the reference time series. Obviously, it is reasonable to take into consideration a time delay between the answers of the survey respondents and changes in the economy. The measure that takes into account the possibility of a shift between time series could have the following form:

$$\tilde{\rho}(x,r,s) = \frac{1}{T-|s|}\sum_{t=n_1}^{n_2}|r_t - x_{t+s}|, \quad (7)$$

where $n_1 = \max\{0,s\}$, $n_2 = \min\{T, T+s\}$ and $s \in \{-3,-2,-1,0,1,2,3\}$. The maximum shift is therefore assumed to be one quarter. The measure $\tilde{\rho}$ should be considered as an average equivalent of the measure $\rho$. The results of the numerical experiments are gathered in Table 3[2].

Table 3. Results of numerical experiments: values of optimization criteria and comparison measures.

| Questions | Number of states | AIC | BIC | logLik | Frequency [%] | $\rho$ | Shift | $\tilde{\rho}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 states | 1261 | 1284 | -623 | 99 | 29.0 | 1 | 0.13 |
| | 3 states | 1168 | 1214 | -570 | 96 | 39.5 | 2 | 0.18 |
| | 4 states | 1119 | 1195 | -536 | 73 | 40.3 | 1 | 0.19 |
| 2 | 2 states | 1390 | 1413 | -688 | 98 | 29.0 | 0 | 0.14 |
| | 3 states | 1296 | 1343 | -634 | 92 | 44.0 | 2 | 0.20 |
| | 4 states | 1244 | 1320 | -599 | 84 | 49.7 | 3 | 0.22 |
| 4 | 2 states | 1392 | 1415 | -689 | 100 | 45 | 1 | 0.21 |
| | 3 states | 1259 | 1305 | -615 | 72 | 51.5 | 2 | 0.24 |
| | 4 states | 1158 | 1235 | -556 | 19 | 69.0 | 2 | 0.32 |
| 7 | 2 states | 1317 | 1340 | -652 | 99 | 37.0 | 3 | 0.15 |
| | 3 states | 1212 | 1259 | -592 | 55 | 52.0 | 3 | 0.22 |
| | 4 states | 1120 | 1196 | -537 | 91 | 57.7 | 3 | 0.26 |
| 1 + 2 | 2 states | 2611 | 2647 | -1294 | 99 | 29.0 | 0 | 0.14 |
| | 3 states | 2384 | 2451 | -1172 | 66 | 42.5 | 3 | 0.19 |
| | 4 states | 2237 | 2340 | -1088 | 11 | 48.3 | 2 | 0.22 |
| 1 + 6 | 2 states | 2669 | 2705 | -1323 | 100 | 35.0 | 0 | 0.17 |
| | 3 states | 2513 | 2580 | -1237 | 72 | 50.0 | 2 | 0.23 |
| | 4 states | 2409 | 2512 | -1173 | 52 | 60.0 | 0 | 0.31 |
| 1 + 7 | 2 states | 2556 | 2592 | -1267 | 100 | 30.0 | 1 | 0.14 |
| | 3 states | 2412 | 2479 | -1186 | 38 | 43.5 | 3 | 0.20 |
| | 4 states | 2266 | 2369 | -1102 | 63 | 57.3 | 3 | 0.26 |

Source: own calculations.

[2] For the clarity of the table, there has been presented the results only of selected questions and their combinations.

All calculated values of $\tilde{\rho}$ are close to the ideal zero value and the calculated Viterbi paths seem to indicate turning points at the same time or in advance comparing to the reference time series. The Viterbi paths for exemplary input data with references to OECD turning points time series for Poland are shown in Figures 4-6 (Q1 and Q7).

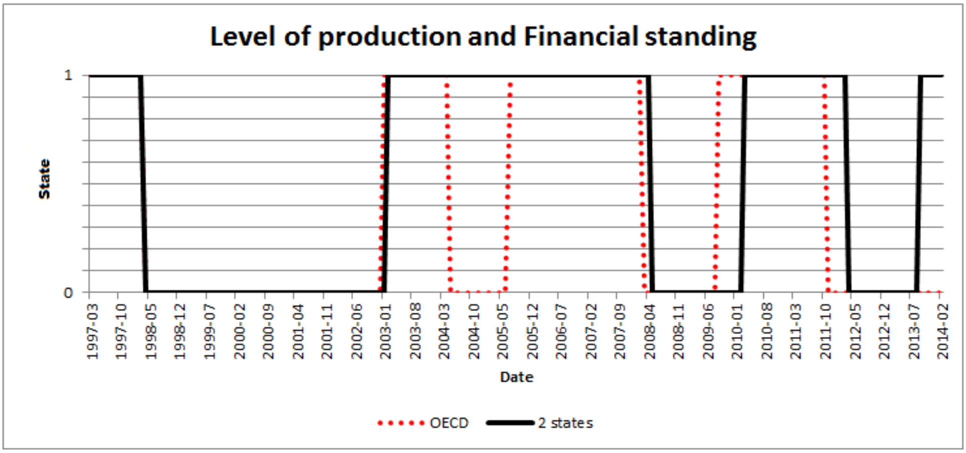


Figure 4. Comparison OECD reference time series with the Viterbi path for 2-state HMM for Q1 and Q7.

Source: own computation.

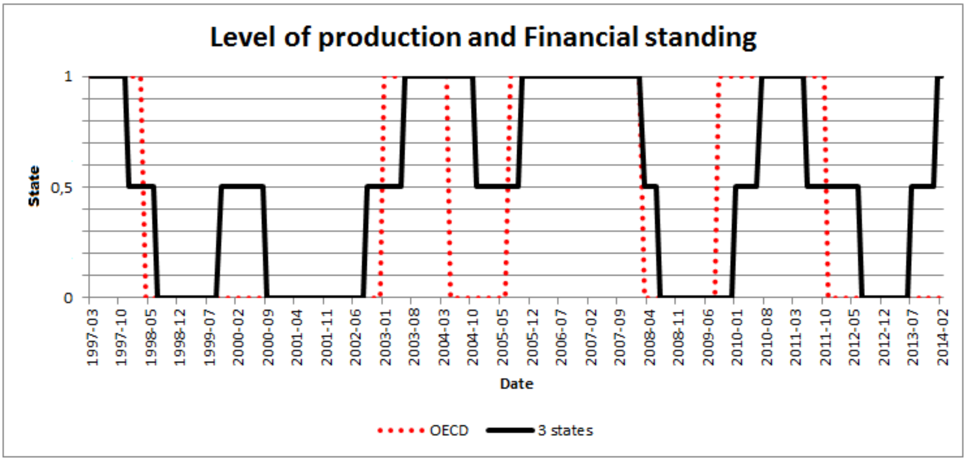


Figure 5. Comparison OECD reference time series with the Viterbi path for 3-state HMM for Q1 and Q7.

Source: own computation.

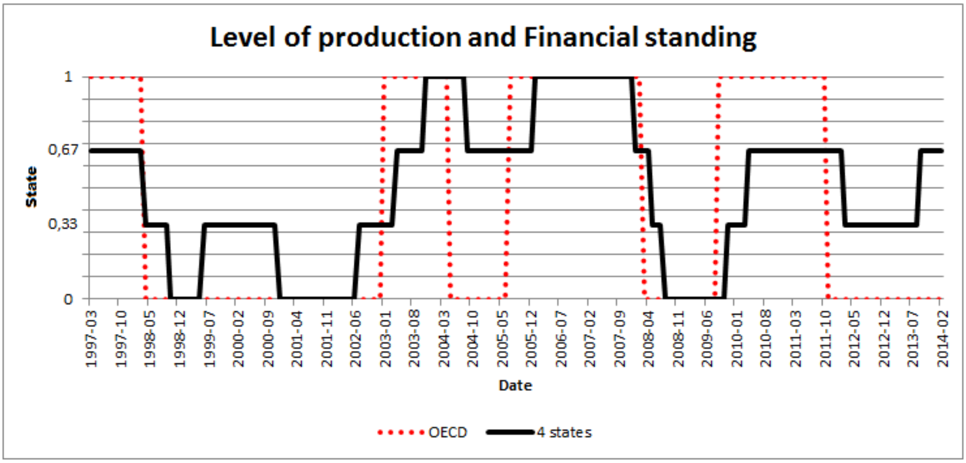


Figure 6. Comparison OECD reference time series with the Viterbi path for 4-state HMM for Q1 and Q7.

Source: own computation.

The Viterbi path for the two-state hidden Markov model declares almost all turning points with sometimes a significant delay of few months. One can also see that one phase of contraction has been missed. Adding the third state to the HMM seems to improve the detection of turning points. First of all, the missing contraction is captured. Although the false, but weak possibility of expansion is also signaled, and the rest of up- and downturns are announced in advance or with small delays in almost all cases. Situation looks even better for four states. False signals are weakened and the dating of the turning points is even more ahead of time than in case of the three-state path.

## 6. Summary

Based on the results of numerical experiments it is justified to draw the following conclusions about the usefulness of the described procedure. Definitely, models computed by the procedure provide satisfactory approximation of business cycle turning points. It is also a flexible and efficient way of an analysis of business tendency surveys balances. Main advantages are ease of generalization, minimal assumptions and high accuracy of the fit.

The procedure exploits not only multistate hidden Markov chains and panel data on input, but also an effective Viterbi algorithm. The procedure is not meant to be a tool for a complete automation of an analysis of changes in business activity based on business tendency surveys in the manufacturing industry in Poland. It was, however, developed to choose from millions of possible models the ones to further analyze. Such an approach turns out to be very useful and the results of the procedure extremely accurate in business cycle turning points identification. Thus, it is an excellent alternative for classical methods and definitely it is worth to continue work in this subject. The next major step would be to add TO-BE balances to the procedure.

## 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.

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. (2012). *Non-classical Markov models - numerical problems*, manuscript.

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

Bernardelli, M. (2013b). *Optimization criteria in the algorithm using hidden Markov models in the analysis of the economic data*, in *Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne*, vol. 2, Zieliński Z. (ed.), Kielce: Wydawnictwo Wyższej Szkoły Handlowej, 43-53.

Bernardelli, M. (2014). Parallel deterministic procedure based on hidden Markov models for the analysis of economic cycles in Poland, *Roczniki Kolegium Analiz Ekonomicznych SGH*, 34, 75-87.

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.

Bhar, R., Hamori, S. (2004). *Hidden Markov Models. Applications to Financial Economics*, New York: Springer.

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.

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, R., Cleveland, W. S., McRae, J. E., Terpenning, I. (1990). *STL:* A seasonal-trend decomposition procedure based on 'loess', *Journal of Official Statistics*, 6, 3-73.

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

Dorosiewicz, S. (2013). *Koniunktura w transporcie. Metodyka badań, wyniki, modele,* Warszawa: Instytut Transportu Samochodowego.

Durbin, R., Eddy, S., Krogh, A., Mitchison, G. (1998). *Biological sequence analysis: Probabilistic models of proteins and nucleic acids*, Cambridge: Cambridge University Press.

Ghahramani, Z. (2001). An introduction to hidden Markov models and Bayesian networks, *International Journal of Pattern Recognition and Artificial Intelligence*, 15 (1), 9-42.

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.

Jelinek, F. (1997). *Statistical methods for speech recognition*, Cambridge: MIT Press.

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.

Mamon, R. S., Elliott, R. J. (eds.) (2007). *Hidden Markov models in finance*, New York: Springer.

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

Rabin, M. O. (1963). Probabilistic automata, *Information and Control*, 6 (3), 230-245.

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.

Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, *IEEE Transactions on Information Theory*, 13 (2), 260-269.

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.