Dynamic Factor Analysis as a Methodology of Business Cycle Research

Konstantin A. Kholodilin

Contents

Chapter 1

Introduction

1.1 The Object of Research

In this thesis we consider the application of dynamic factor analysis (with possibly nonlinear dynamics) to the study of the cyclical fluctuations of a modern market economy. Thus, the object of our research is the business cycle, while the principal tool is dynamic factor model.

However, before trying to study business cycle properties, we have to define the notion of a business cycle and its phases — expansion and recession — in particular, since this would allow us to delineate the object of our research and to employ the most appropriate model which permits grasping these properties. Therefore we start with discussing several definitions of the business cycle and its phases which are widely used in the modern literature on economic fluctuations.

One of the most popular among the practitioners of the business cycle research is the rule of thumb explained in Garcia-Ferrer and Queralt (1998): "a widely used alternative to the NBER cycle-dating rules is to define a turning point as a first of at least two successive increases (declines) in the growth rate of the GNP". Nevertheless, this definition has a serious flaw — its narrowness, since it takes into account only the movements in GNP or GDP. GDP is one of the most general economic indicators, however, its level or rate of growth may change due to the partial changes, that is, fluctuations in a particular sector of the economy, while the rest of the economy does not necessarily move in the same direction.

Moreover, this definition requires only the growth rates to change, whereas according to the new classical definition of the business cycle one of the necessary conditions is the change of sign of the rate of growth. As Barro and Grilli (1994, p. 15) put it: "When real GNP falls toward a low point or trough, the economy is in a recession or an economic contraction. These are periods characterized by negative rates of growth".

The NBER's definition[1] is more consistent in this respect — it explicitly requires the downward or upward movements to be omnipresent, not restricted to a particular branch of economy: "A recession is a significant decline in activity spread across the economy, lasting more than a few months, visible in industrial production, employment, real income, and trade. A recession begins just after the economy reaches a peak of output and employment and ends as the economy reaches its trough. Between trough and peak, the economy is in an expansion. Expansion is the normal state of the economy; recessions are brief and relatively rare".

The NBER's definition, as well as the "practitioners' definition" mentioned above, are based on the comparison of the present state of the economy to its immediate past and the future. The "neoclassical definition", however, has more static nature and rests upon the notion of the potential output. As exemplified in Samuelson and Nordhaus (1998, p. 433): "A business cycle is a swing in total output, income and employment usually lasting for a period of 2 to 10 years, marked by widespread expansion or contraction in most sectors of the economy. In modern economics, business cycles are said to occur when actual GDP rises relative to potential GDP (expansion) or falls relative to potential GDP (contraction or recession)".

Stock and Watson (1989, p. 357) come with a more elaborate probabilistic definition of the business cycle phases: "The Recession Index is an estimate of the probability that the economy will be in a recession six months hence. This probability is computed using the same time series model used to calculate the proposed LEI[2], and is based on a definition (in terms of the sample path of ΔC_t[3]) of what constitutes a recession or an expansion. Unfortunately, it is difficult to quantify precisely those patterns that will be recognized as expansions or contractions. Burns and Mitchell (1946, p. 3) considered the minimum period for a full business (reference) cycle to be one year; in practice, the shortest expansions they identified were six months. The

[1] http://www.nber.org/cycles/recessions.html

[2] Leading economic indicator

[3] First differences of the common dynamic factor C_t.

Recession Index is computed by approximating a recessionary (expansionary) period in terms of negative (positive) growth of the CEI[4] that lasts at least six months. Note 6: We define a month to be in a recessionary pattern if that month is either in a sequence of six consecutive declines of C_t below some boundary b_{rt}, or is in a sequence of nine declines below the boundary with no more than one increase during the middle seven months".

Diebold and Rudebusch (1996, p. 67) stress two key characteristics of the business cycle as defined by Burns and Mitchell: "The first is the comovement among individual economic variables. The second prominent element of Burns and Mitchell's definition of business cycles is their division of business cycles into separate phases or regimes".

Thus, we can recapitulate the main features of a business cycle phase: (1) the movement should be general, embracing many sectors of the economy at the time; (2) it should last long enough to be discriminated from a casual short-lived oscillations; (3) the behavior of the economy at the upswing is different from that on the downswing[5]. Given that these three conditions hold, we observe either expansion or recession. The points on the border of two adjacent phases are the turning points.

From the practical point of view the business cycle dynamics may be translated in terms of a common factor, which reflects the overall movement of the economy, being in high or low regime for several consecutive periods. This is captured by high enough — above 0.5 or some other margin depending perhaps on the cost of false prediction imposed by the researcher — probability of the "state of the economy" approximated by the common component belonging to that regime.

1.2 The Objective of Research

The main objective of our research is to elaborate a technique of constructing a composite economic indicator or a set of such indicators which would *detecting* and *forecasting* the turning points of the business cycle. This technique should have a sound theoretical basis and adequately meet the defining characteristics implied by the business cycle definition.

[4]Composite economic indicator

[5]The expansions usually last longer than the contractions. Moreover, sometimes the volatility is higher in the recessionary phase of the cycle.

As a point of departure we choose the definition of business cycle proposed by Burns and Mitchell (1946). Therefore the technique in question should satisfy the business cycle properties exposed in the end of the previous section. We believe that the most appropriate method to capture the Burns and Mitchell's cycle would be dynamic factor analysis.

Nevertheless, dynamic factor analysis in its current state requires undoubtedly some refinements and extensions to obtain unbiased and consistent estimates of the composite economic indicators and to use the available information in the best possible way.

Our research is mostly oriented towards the practitioners who have opted for using the dynamic factor approach in the construction of the business cycle indicator both at the regional and national levels.

1.3 The Thesis Structure

The thesis is comprised of five chapters where the first and the last chapters are the introduction and conclusion delineating the objectives of the study and summarizing the results achieved during research.

Chapter two describes various approaches to the analysis of economic fluctuations proposed during the last 20 years, especially those that have to do with the construction of the composite economic indicators and identification of the turns of business cycle. On the one hand, it concentrates on models with nonlinear, namely Markov-switching, dynamics, on the other hand, it is concerned with dynamic factor models. Finally, it shows the combined techniques which unite these two principal approaches, thus, modeling common latent factor with regime-switching dynamics.

In chapter three we introduce a general multifactor dynamic model with linear and regime-switching dynamics. This model allows capturing the intertemporal (leading versus coincident) dimension of the latent common factors. Two alternative multifactor dynamic models with a leading and a coincident unobserved common factors are examined: a model where the common coincident factor is Granger-caused by the common leading factor and a model where the leading relationship is translated into a set of specific restrictions imposed on the transition probabilities matrix.

Chapter four concentrates on the supplementary devices which allow to overcome some data problems which are very frequent in the practitioner's life. Among the most prominent are the structural breaks and missing observations. It is shown that some of these troubles can be coped with by modifying the dynamic common factors models, which leads to more efficient estimates of the parameters of the models.

Each subsection of chapters three and four is accompanied by an appendix containing the tables and graphs which illustrate the argument. A list of references can be found in the very end of the thesis.

Chapter 2

Turning Points Analysis: Current State of Affairs

Since many years the measurement and forecasting of cyclical economic activity has played an important role as a tool of macroeconomic policy. The most notorious examples of such measurements are the National Bureau of Economic Research (NBER) leading, coincident, and lagging business cycle indicators (USA) calculated since 1940s, similar indicators provided by the Central Statistical Office (UK) since 1970s, and more recently — since 1980s — the indicators computed on the international bases by the Organization for Economic Cooperation and Development.

The policy makers need to know at which stage of the business cycle the economy is situated and where it is going to move next. To answer these questions we have to solve the following three problems. The first problem is to build a composite economic indicator or a set of such indicators, which would reflect the aggregate cyclical fluctuations of the economy, being a quantitative proxy of the business cycle. Secondly, there should be a way to obtain from this business conditions' indicator a business cycle chronology, that is, a sequence of turning points marking transitions from the recessions to the expansions and vice versa. This would allow us identifying the actual phase of the business cycle. Once the previous problems have been solved, one can predict the future shifts between upward and downward phases of the business cycle.

The business cycle is normally treated as the sequential upward and downward movements in the GDP. However, this approach is far from being satisfactory. One of the reasons is, as we stressed above, that GDP fails sometimes to reflect to a whole extent such a complex phenomenon as a business cycle. If we define the business cycle as a synchronized movement across all the economy, then all the changes in GDP may not be considered as attributable to the cycle, since they may be related to some sectors only. For instance, if the exports due to the lack of the purchasing power abroad are going down, GDP may decrease, however, this does not necessarily mean that the whole economy will be doing bad.

Another reason is that for the regional economies the GDP data for a more or less long period are simply unavailable. Sometimes also their frequency is too low, e.g. annual, for them to be used as an indicator of the business cycle. Therefore if one wants to investigate the movement of the economy of some region, he is compelled to look for separate pieces of appropriate information to shed light on the overall fluctuations of that economy. Historically, the second reason was the principal one for calculating the composite indices, since real GDP or GNP indices that would cover long enough periods of time were not really available until the end of 1950s.

Here we concentrate our attention on only two basic lines of business cycle research, as capturing the main features of this phenomenon: Markov-switching models, which reflect the asymmetries between different phases of business cycle, and dynamic factor models, which take into account the comovements of the observed time series, as well as a combined approach, which manages to unify these features in a single model.

The chapter is structured as follows. In the next section we will consider the Markov switching model and its modifications designed to analyze and forecast the turning points of the macroeconomic activity. In section three we discuss various linear dynamic factor models intended to model the business cycle. Section four is devoted to the models which combine the above two approaches, thus producing the dynamic factor models with regime-switching dynamics. Section five concludes the chapter.

2.1 Markov Regime Switching Models

2.1.1 Basic univariate Markov-switching model

The main references here are Hamilton (1989, 1994). The basic idea of a Markov-switching model is that the economy is thought to behave differently in different regimes it may go through. For instance, the mean and variance of the growth rates observed during recession may be different from those observed during expansion. In other words, the asymmetry in the economic dynamics is allowed for.

More specifically the Markov-switching model can be motivated as follows. Suppose we observe a time series, y_t, which is a sum of two unobserved components:

$$y_t = g_t + z_t, \tag{2.1}$$

where

$$g_t = g_{t-1} + \alpha_0 + \alpha_1 s_t \tag{2.2}$$

is the Markov trend with s_t being an unobserved state (regime) of the economy, as, for instance:

$$s_t = \begin{cases} 1, & \text{if recession in period t} \\ 0, & \text{if expansion in period} \end{cases}$$

However, there can be more than two states. For example, given the fact that in the 1990s the growth rates fell in many market economies, one has to introduce additional states (”high secular growth” and ”low secular growth”) in order to capture this structural break. Artis et al. (1999) having faced this problem, opted for constructing three-states Markov model, where one of the regimes is intended to absorb the secular growth rates deceleration, while the other two regimes are the usual ”recession” and ”expansion”.

The second component of the observable series, z_t, was postulated by Hamilton (1989) as an ARIMA(r,1,0) process, although later it was generalized to non-integrated processes case:

$$\phi(L)(1-L)z_t = \varepsilon_t \tag{2.3}$$

where ε_t is a zero-mean white noise.

When the statistical model is formulated, the objective is to find the probabilities of each regime conditional on the up-to-date information set (information about the relevant variables available at the moment). These probabilities are computed recursively as:

$$\xi_{t|t} = \frac{(\xi_{t|t} \odot \eta_t)}{\xi'_{t|t}\eta_t}$$

and updated as follows:

$$\xi_{t+1|t} = P\xi_{t|t}$$

where

$\xi'_{t|t} = \left(\begin{array}{ccc} P(s_t = 1|Y_t) & ... & P(s_t = m|Y_t) \end{array} \right)$ is the vector of the conditional state probabilities for each of the m regimes;

$\eta'_t = \left(\begin{array}{ccc} f(y_t, s_t = 1|Y_{t-1}; \theta) & ... & f(y_t, s_t = m|Y_{t-1}; \theta) \end{array} \right)$ is the vector of the densities of the $t - th$ observation for each of m states conditioned on the previous period information set; $\odot$ is the element-by-element multiplication operator; θ is the vector of parameters, and

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

is the transition matrix, or matrix of transition probabilities.

The algorithm may be started as:

$$\xi_{1|0} = \frac{1}{m}E \tag{2.4}$$

where E is an $m \times 1$ vector of ones.

If, for example, we assume Gaussian distribution of the noise, the conditional density function may be written as follows:

$$f(y_t, s_t = i, s_{t-1} = j, ..., s_{t-p} = k|Y_{t-1}; \theta) =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left\{ -\frac{1}{2\sigma_j^2} \left[y_t - \mu_{s_t} - \phi_1(y_{t-1} - \mu_{s_{t-1}}) - ... - \phi_p(y_{t-p} - \mu_{s_{t-p}}) \right]^2 \right\}$$

The filtered probabilities, $P(s_t = i, s_{t-1} = j, ..., s_{t-p} = k|Y_t)$, can be estimated by the maximum likelihood, where the log-likelihood function is constructed as:

$$L(\theta) = \sum_{t=1}^{T} \log(\xi'_{t|t}\eta_t)$$

As soon as the posterior probabilities of the states are calculated, one can determine the business cycle chronology. One of the states, having the lowest mean, is interpreted as recession, and then the recession dates may be defined as those for which the probability of being in the recession is higher compared with that of being in expansion (if we deal with the two regimes model) or that of some other state (when number of the regimes is greater than 2).

2.1.2 Forecasting with Markov-switching models

As soon as the Markov-switching model is estimated, the next important problem is to forecast both the observed variable and the predictive probabilities. The τ-step-ahead forecast of $\xi_{t+\tau}$ conditional on the information available at period t is calculated as:

$$\xi_{t+\tau|t} = P^{\tau}\xi_{t|t} \tag{2.5}$$

For example, if we are making a one-period-ahead forecast, then the forecast of y_{t+1} given the information available in t would be:

$$E(y_{t+1}|Y_t;\theta) = h_t'\xi_{t+1|t}, \tag{2.6}$$

where

$$h_t = \begin{cases} E(y_{t+1}|s_{t+1} = 1, Y_t;\theta) = \mu_1 + \phi_1(y_{t-1} - \mu_{s_t}) + ... + \phi_p(y_{t-p} - \mu_{s_{t+1-p}}) \\ \vdots \\ E(y_{t+1}|s_{t+1} = j, Y_t;\theta) = \mu_j + \phi_1(y_{t-1} - \mu_{s_t}) + ... + \phi_p(y_{t-p} - \mu_{s_{t+1-p}}) \end{cases} \tag{2.7}$$

and $\xi_{t+1|t}$ is calculated according to (2.5).

2.1.3 Multivariate Markov-switching models

Another interesting and important point is the extension of the univariate Hamilton's model to the multivariate case — the so-called Markov-switching vector autoregression (MS-VAR).

In the general case MS(m)-VAR(p) can be expressed as follows (see Krolzig (1996)):

$$y_t = \mu(s_t) + \sum_{i=1}^{p} A_i(s_t) y_{t-i} + \varepsilon_t(s_t) \tag{2.8}$$

where y_t is the $n \times 1$ vector of the observable variables; $\mu(s_t)$ is the $n \times 1$ vector of the intercept terms; $A_i(s_t)$ is the $n \times n$ matrix of the autoregressive coefficients at lag i; and $\varepsilon_t(s_t)$ is the $n \times 1$ vector of the disturbance terms. $\varepsilon_t(s_t) NID(0, \Sigma(s_t))$. The state variable $s_t = 1, ..., m$, that is, m regimes are assumed. Hence this model is called MS(m)-VAR(p): m states and p vector-autoregressive terms.

In this specification both the intercepts, and the autoregressive parameters and the disturbances are state-dependent. However, simpler specifications can be considered with the regime-invariance restrictions imposed on some of these terms or parameters.

The simplest multivariate Markov-switching model is that of Kontolemis (2001), where there are no lagged terms and the only switching components are the intercepts and covariances. Moreover, Kontolemis imposes the restriction of identical turning points across the different coincident time series, assuming that a common probability transition mechanism governs switches from one regime to another. In this case the generalization of Hamilton's univariate Markov-switching model is immediate: the only thing to be changed are the likelihood functions comprising vector η_t. The typical element of this vector is:

$$f(y_t, s_t = j | Y_{t-1}; \theta) = \frac{1}{(2\pi)^{n/2} |\Sigma_j|^{1/2}} \exp\left(- \frac{(y_t - \mu_j)' \Sigma_j^{-1} (y_t - \mu_j)}{2} \right) \tag{2.9}$$

where y_t is a $n \times 1$ vector of coincident variables; Σ_j is the variance-covariance matrix corresponding to regime j, and μ_j is the intercept term in regime j.

This approach is very promising, since it allows extracting the common information regarding the turning points of the economy which is contained in various time series. There is no need to construct any composite indicator to get the recessions chronology. Furthermore, this model can be used to forecast the turning points.

Krolzig (1996) generalizes this framework even more by elaborating an algorithm to estimate VAR Markov-switching models with cointegration — MS-VECM (Markov-switching vector error correction mechanism model) as well as MS-VARMA.

Krolzig (1996) considers a case where the time series are cointegrated. Since a valuable information is contained in the series in levels, it is more efficient to use a VECM model rather than estimating a VAR in the differences (see, e.g. Hylleberg and Mizon (1989)). The model Krolzig estimates, MS(m)-VECM(p-1), is as follows:

$$\Delta y_t = \mu(s_t) + \sum_{i=1}^{p-1} D_i \Delta y_{t-i} + \Pi y_{t-p} + \varepsilon_t \tag{2.10}$$

where $D_i = -(I_n - \sum_{j=1}^{i} A_j)$ with autoregressive matrices A_i from (2.8) but without state-dependence; $\Pi = I_n - \sum_{i=1}^{p} A_i = A(1) = BC$; B is the $n \times r$ loading matrix, and C is the $r \times n$ cointegration matrix. The rank of matrix Π, r, is called the integration rank, and $n - r$ is the number of the common stochastic trends in the system (see Stock and Watson (1988)). The only term which is regime dependent in this model is the vector of intercepts.

To estimate the above model Krolzig proposes a two-step procedure. At the first step the cointegration rank, r, is determined and the cointegration matrix, C, is estimating using Johansen maximum likelihood analysis of VAR (see Johansen (1991)). At the second stage all the remaining parameters of MS(m)-VECM(p-1) are recovered with an EM algorithm (this algorithm as applied to the Markov-switching modeling is presented in Hamilton (1990) and in particular with respect to the MS-VECM — in Krolzig (1996)).

Phillips (1991) considers a very interesting vector Markov-switching model which allows for the independent regime-switching processes as well as for various types of dependence between the unobservable states of different time series: perfect correlation (as in the Kontolemis (2001) model, which may be considered as a particular case of the Phillips' model) and one time series leading the other. The latter can be used to model the relationship between the CCI and CLI. Moreover, the Phillips' setting allows testing all these specifications against the baseline model, i.e. the model where the transition matrix is unrestricted.

2.1.4 Markov-switching models with time-varying transition probabilities

Having noticed the quite poor performance of the standard fixed transition probability (FTP) Markov-switching models in the presence of higher noised observations, Filardo (1994) had suggested to extend the standard model to the time-varying transition probabilities (TVTP) case.

He motivates this extension by three arguments. First, the probabilities of transition from one regime to the other may change depending on whether the economy is in the beginning, in the middle or in the end of a regime. Thus, for instance, probability of switch from recession to expansion may increase dramatically in the end of the contractionary period. Constant transition probabilities of the FTP models do not allow for these changes. Secondly, fixed transition probabilities imply constant expected durations of the states, while the time-varying transition probabilities lead to the varying durations, which is supported by the empirical evidence displaying wide variety of expansions and recessions durations. Thirdly, TVTP models enrich the temporal persistence picture, since to the persistence of the business cycle phases captured by the autoregressive terms and by the fixed transition probabilities they add persistence resulting from the dependence of the transition probabilities of the dynamics of the other variables.

The algorithm is basically the same as in the baseline Hamilton's model. The only difference is the way the transition probabilities are modelled. If in the baseline $p_{ij} = P(s_t = j|s_{t-1} = i)$ for any i and j, and, thus, is constant, under the TVTP framework $p_{ij} = P(s_t = j|s_{t-1} = i, Z_t)$, where $Z_t = \{z_t, z_{t-1}, ...\}$ is the past history of the variables affecting the transition variabilities.

There can be various functional forms reflecting the dependence of the transition probabilities of the variables entering Z_t. Filardo (1994) in a two-regime TVTP Markov-switching model uses logistic function to model the transition probabilities (he denotes $p = p_{11}$ and $q = p_{22}$):

$$p(Z_t) = \frac{\exp(\theta_{p0} + \sum_{j=1}^{J_1} \theta_{pj} z_{t-j})}{1 + \exp(\theta_{p0} + \sum_{j=1}^{J_1} \theta_{pj} z_{t-j})} \tag{2.11}$$

and

$$q(Z_t) = \frac{\exp(\theta_{q0} + \sum_{j=1}^{J_2} \theta_{qj} z_{t-j})}{1 + \exp(\theta_{q0} + \sum_{j=1}^{J_2} \theta_{qj} z_{t-j})} \tag{2.12}$$

When $\theta_{pj} = \theta_{qj} = 0$ for $j \neq 0$, we have a standard FTP Markov-switching model. This functional form permits checking the null hypothesis of constant transition probabilities against the alternative of the TVTP.

Filardo (1994) shows that the removing the restriction of FTP may lead to the improvement of the forecasting performance of the Markov-switching model.

Filardo and Gordon (1998) use a probit specification to model the time-varying transition probabilities. The transition probabilities at each time period are computed using a conditional cumulative distribution function. Thus, the probability of switching to the state 1, given that in the previous period the system was in the same state, may be calculated as:

$$p_t = P(U_t - \theta_0 - \theta_s - z_t'\theta_z) = 1 - \int_{-\infty}^{\bar{u}(z)} \frac{1}{\sqrt{2\pi}} \exp\left[-0.5u^2\right] du$$

where $u(z)$ is the upper limit of integration and $u = -\theta_0 - \theta_s - z_t'\theta_z$.

Similarly the probability of switching from state 2 to state 2 conditional on a vector of information variables z_t is given by:

$$q_t = P(U_t < -\theta_0 - z_t'\theta_z) = \int_{-\infty}^{\bar{u}(z)} \frac{1}{\sqrt{2\pi}} \exp\left(-0.5u^2\right) du$$

where $u = -\theta_0 - z_t'\theta_z$.

This specification stems from modeling the latent variable S_t^* as a function of the unobserved state in the previous period, s_{t-1}, and of a set of the information variables, z_t:

$$S_t^* = \theta_0 + \theta_s s_{t-1} + z_t'\theta_z$$

Note that in this model only mean of the latent variable S_t^* is state-dependent, while the coefficients with the information variables are the same for both regimes. On the other hand, the fact that an unobserved state variable enters the equations for the transition probabilities, may complicate the estimation of the model.

Another specification of the time-varying transition probabilities function was put forward by Simpson, Osborn and Sensier (2000). It is an exponential function:

$$p(Z_t) = 1 - \exp[-(\theta_{p0} + \sum_{j=1}^{J_1} \theta_{pj} z_{t-j})]$$

and

$$q(Z_t) = 1 - \exp[-(\theta_{q0} + \sum_{j=1}^{J_2} \theta_{qj} z_{t-j})]$$

As they claim and as our own experience confirms, this specification is more stable in the estimation sense compared to the logistic and probit ones.

2.2 Dynamic Common Factor Approach

2.2.1 Single factor model

Since the late 1980s the composite economic indicator (CEI) together with the leading economic indicator (LEI), estimated according to the methodology proposed in Stock and Watson (1989, 1991, 1993), are regularly published in the USA as an alternative to the official composite indicators provided by the US Department of Commerce.

In their definition of the business cycle, Burns and Mitchell stress two important and indispensable features of this economic phenomenon: the comovement of the macroeconomic variables throughout the cycle and the asymmetry between the expansions and recessions. The model of the coincident economic indicator proposed by Stock and Watson incorporates first of these features by trying to capture the common dynamics among different macroeconomic time series at the business cycle frequencies.

They assume the existence of some common dynamic factor, which underlies the comovements of the individual coincident economic variables and which, by consequence, can be defined as the state of affairs in the economy. However, this factor is unobservable and so we can make only hypothetical statements about its nature.

Stock and Watson distinguish between the common and specific (idiosyncratic) components of each of the time series they include in their model. The common component, which may depend on its own past values, and therefore is called dynamic, determines the comovements of different macroeconomic variables. The idiosyncratic components capture the specific dynamics of each of these series.

The algorithm, although having evolved over time, is basically the following one:

1. Construct CEI using the dynamic common factor model: express the model in the state space form, use Kalman filter to construct the maximum likelihood function, estimate it to get the estimates of the parameters and of the CEI.

2. Estimate VAR including CEI and a set of individual leading variables.

3. Forecast the CEI for several periods (say, 6 months) into the future, using the VAR estimated at step 2.

4. Construct the LEI as a difference between the CEI forecasted for the period $t+k$ (where k is the forecast horizon) and CEI at time t.

5. Forecast the recessions using either logit model, including the LEI and the same leading series as in the VAR at step 2, or the stochastic simulation forecasting approach, similar to that of Wecker and Kling.

The composite economic indicator is modelled as follows:

$$\Delta y_t = \delta + \gamma(L)\Delta C_t + u_t \tag{2.13}$$

where the dynamics of the common factor are:

$$\phi(L)\Delta C_t = \mu + \varepsilon_t \tag{2.14}$$

and the specific, or idiosyncratic, components are:

$$\Psi(L)u_t = \eta_t \tag{2.15}$$

where y_t is the vector of the coincident variables in levels; C_t is the common dynamic factor; δ and μ are the constant terms; $\Psi(L)$ and $\phi(L)$ are the lag polynomials; ε_t and η_t are serially and mutually uncorrelated shocks:

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \mathrm{O} \\ \mathrm{O} & \Sigma_\eta \end{pmatrix}\right)$$

To identify the model Stock and Watson set the variance of the common factor equal to one and assume that the lag polynomial matrices Ψ_j ($j = 1, ..., q$, where the q is the maximum autoregressive order of the specific components) are diagonal, thus excluding any causality relationships between the idiosyncratic components.

In principle it is possible to build an unobserved components model with the variables in levels. However, Stock and Watson use coincident variables which are integrated of order one and which are not cointegrated. Therefore the model is estimated with the logged variables in the first differences.

To estimate the model, one can express it in a state space form and then use either maximum likelihood technique — as it is done by Stock and Watson (1991) or, for example, Bayesian methods (like Gibbs sampler, see e.g. Kim and Nelson (1999c)) in order to find the unknown parameters and the estimates of the state variables.

The above model can be expressed in the following state-space form:

Measurement equation:

$$\Delta y_t = Ax_t \tag{2.16}$$

Transition equation:

$$x_t = \alpha + Cx_{t-1} + v_t \tag{2.17}$$

where $x_t = \begin{pmatrix} \Delta C_t & ... & \Delta C_{t-p+1} & u_{1t} & ... & u_{n,t-q_n+1} \end{pmatrix}'$ is the $k \times 1$ state vector, $k = p + \sum_{i=1}^{n} q_i$; $v_t = \begin{pmatrix} \varepsilon_t & ... & 0 & \varepsilon_{1t} & ... & 0 \end{pmatrix}$ is the $k \times 1$ vector of the disturbances to the state vector, and the intercept $k \times 1$ vector is defined as: $\alpha = \begin{pmatrix} \mu & 0 & ... & 0 \end{pmatrix}$.

The system matrices have the following form. The measurement matrix:

$$A = \begin{pmatrix} \Gamma_1 & i_{q_1} & O \\ \vdots & \vdots & \vdots \\ \Gamma_n & O & i_{q_n} \end{pmatrix}$$

where Γ_i is the $1 \times g$ vector of the factor loadings of the $i - th$ observed variable: $\Gamma_i = (\gamma_{i,1}, ..., \gamma_{i,g_i}, ..., 0)$ with $g = \max\{g_1, ..., g_n\}$.

The transition matrix:

$$C = \begin{pmatrix} \Phi & 0 & & 0 \\ 0 & \Psi^1 & & \\ & & \ddots & \\ 0 & 0 & & \Psi^n \end{pmatrix}$$

$$\Phi = \begin{pmatrix} \phi_C & 0 \\ I_{p_C-1} & o_{p_C-1} \end{pmatrix}$$

with $\phi_C = (\phi_{1,...,}\phi_p)$ — the $1 \times p$ vector of the AR coefficients of the common factor. The matrices $\Psi^1, ..., \Psi^n$ have the same structure as Φ^C.

The $k \times k$ variance matrix of the disturbance to the state vector, v_t, may be written as:

$$Q = \begin{pmatrix} \sigma_\varepsilon^2 & & & 0 \\ & \ddots & & \\ & & \sigma_i^2 & \\ 0 & & & 0 \end{pmatrix}$$

where σ_i^2 is the variance of the shock to the $i - th$ specific component.

Having expressed the model in the state-space form, one can apply the Kalman filter recursions to get the likelihood function and to find the filtered and smoothed estimates of the state vector. We will not reproduce here the recursive Kalman filter equations referring the reader to the chapter 13 of Hamilton (1994).

Another estimation approach is the Monte Carlo Markov Chains (MCMC), or Bayesian, method. It is especially useful when there are too many parameters to be estimated. For the details of the estimation a single dynamic factor model in the Bayesian framework see Otrok and Whiteman (1998) and Kim and Nelson (1999c).

To construct an LEI, it is necessary to predict the future values of the CEI. To do this, Stock and Watson estimate the following VAR model:

$$\Delta C_t = \mu_C + \Phi_C(L)\Delta C_{t-1} + \Phi_{CL}(L)L_{t-1} + v_{Ct} \tag{2.18}$$

$$L_t = \mu_L + \Phi_{LC}(L)\Delta C_{t-1} + \Phi_L(L)L_{t-1} + v_{Lt} \tag{2.19}$$

where L_t is the vector of the leading variables; μ_i is the constant term ($i = C, L$); $\Phi_{ij}(L)$ are the lag polynomials ($i, j = C, M$).

One can use either the two-step or one-step procedure to estimate the CEI and LEI. With the two-step procedure one uses equations (2.13)-(**??**) to put them in a state space form in order to estimate CEI ($C_{t|t} = E(C_t|I_t)$) with a Kalman filter and then, at the second step, estimate the VAR replacing ΔC_t by $\Delta C_{t|t}$. However it is also possible to do the estimation simultaneously — in one step. Then, equations (2.18) and (2.19) are substituted for the equation (2.14), and the system of equations (2.13), 2.15, (2.18), and (2.19) is used to run the Kalman filter.

Initially the forecasting was done with a logit model. The idea behind this model is the following (see for example, Amemiya (1981)).

Suppose that there exists an unobservable variable characterizing the state of the economy: either expansion or recession. However, one can code this latent variable with the observable dichotomous variable R_t :

$$R_t = \begin{cases} 1, & \text{if recession in period } t \\ 0, & \text{if expansion in period } t \end{cases}$$

Normally, in the USA one uses for R_t the dummy corresponding to the NBER business cycle chronologies. However, if we do not have any generally accepted dating of the turning points, it is not very clear what variable to use in order to code the recession/expansion latent time series.

Then the model for the recessions forecasting can be written as follows:

$$P(R_{t+k} = 1) = F(X_t'\beta)$$

where $P(R_{t+k} = 1)$ is the conditional probability of having recession in period $t+k$; $F(\cdot)$ is some cumulative distribution function; X_t is the $n \times 1$ vector of the explanatory variables (in this case it is CEI and several leading variables), and β is the $n \times 1$ vector of the unknown parameters.

In particular, when we are predicting recessions with the logit model:

$$P(R_{t+k} = 1) = \frac{\exp(X_t'\beta)}{1 + \exp(X_t'\beta)}$$

The recessions probabilities and the parameters for the above (and for any other qualitative response) model are estimated by the maximum likelihood. The log-likelihood function to maximize may be written as:

$$L(\beta) = \sum_{t=1}^{T} \{R_{t+k} \log(F(X_t'\beta)) + (1 - R_{t+k}) \log(1 - F(X_t'\beta))\}$$

Varying the forecasting horizon and the composition of the vector X_t, we can choose the model which gives the best predictions. The recession probabilities estimated from this model can be used to define the dates of the recessions and expansions, which, in turn, can be compared with the official chronology. Nevertheless, it is important to stress it once again: without any generally accepted and reliable business cycle chronology, the prediction of the recession probabilities using any qualitative response model becomes senseless.

Another method to forecast the recession probabilities, which was employed more recently by Stock and Watson (1993) is the stochastic simulation forecasting reminding that of Wecker (1979) and Kling (1987).

A very interesting extension to the dynamic common factor analysis was proposed recently by Mariano and Murasawa (2003). They consider a model where the use of the mixed-frequency data, say monthly and quarterly, in the common dynamic factor model is allowed. This is especially useful if we want to get use in constructing our composite coincident indicator of such an important coincident time series as GDP. As a rule the GDP data are released at much lower frequency than individual series characterizing specific sectors of the economy. The Mariano and Murasawa's model permits taking advantage of the valuable information contained in the lower-frequency time series. Thus, we can minimize the information losses and enhance the efficiency of our common dynamic factor model.

2.2.2 Multifactor model

This single-factor approach was extended by Kose et al. (1999) to the cross-sectional data — which automatically implies a multiplicity of the common components. Their model has a hierarchy of common factors starting from the most general world indicator, which absorbs the worldwide common fluctuations, and descending to the country-specific common factors. All these common factors are coincident.

A further step is made by Forni et al. (2001) whose generalized dynamic factor approach not only covers the cross-sectional dimension but also takes care of the intertemporal relations between the common factors allowing to estimate the leading, coincident, and lagging factors. Unlike all the above mentioned methods, this technique is nonparametric and estimates the dynamic factors in the frequency domain. Forni et al. (2001) relax the restriction of the mutual uncorrelatedness of the idiosyncratic components imposed by Stock and Watson, however, keeping intact the assumption of the common factors being orthogonal to each other. Hence the coincident and leading common factors are estimated separately: the coincident factor being extracted from the whole set of the observed time series and the leading common factor being estimated from the subset of the observed variables which turned out to be leading with respect to the coincident component found at the previous step.

The common factors are recovered by applying a set of symmetric filters to the original multivariate series. The weights of these filters are obtained by inverse Fourier transform from the eigenvectors of the spectral density matrices of the observed process computed at different frequencies. Only those eigenvectors are chosen which correspond to the largest eigenvalues of the spectral density matrices. One important disadvantage of this method, being a consequence of the way the common factors are estimated, is that the observations in the beginning and in the end of the sample are lost and hence it is necessary to fill the resulting gaps through back- and forecasting.

This technique is applied, for instance, to construct the monthly coincident indicator of the business cycle of the euro area as described in Altissimo et al. (2001).

An alternative approach known as diffusion index model has been developed recently by Stock and Watson (1998) and Watson (2000). They are using a dynamic factor analysis to estimate nonparametrically the multiple common factors based on a very large set of predictors (hundreds or even thousands of individual component series) in the time domain. Under this approach the common factors are estimated as the principal components corresponding to the largest eigenvalues of the variance-covariance matrix of the set of component series. The common factors are interpreted given their correlations with the individual component series. Thus, the common factor having the highest correlations with most of the real sector variables is thought to reflect the cyclical fluctuations caused by the real shocks. Other common factors may be related to the monetary or foreign trade sectors.

Two examples of application of this methodology are the construction of the diffusion indices for the European Union countries undertaken by Marcellino and Stock (2000) and by Angelini, Henry, and Mestre (2001).

A further extension could be a parametric bifactor model which would allow for a lead-lag relationships between the common factors: leading factor and coincident factor. Different mechanisms relating the common coincident factor to the common leading factor are feasible. One is the Granger causality between the two common factors, another is a kind of probabilistic mechanism reflecting this relationship in the transition probabilities matrix. In any case, the bifactor model would permit improving of the forecasting of the coincident factor because of the additional information coming from the leading variables. This extension will be introduced in chapter 3 of this thesis.

2.3 Regime-Switching Dynamic Factor Model

A very promising approach was introduced by Kim (1994) and Kim and Nelson (1999c). It was applied by Chauvet (1998 and 1998/9), and by Kim and Yoo (1995) to construct the US common coincident economic indicator having nonlinear (Markov-switching) dynamics.

The Markov-switching common factor (MS-CF) approach allows to estimate simultaneously both the common factor, underlying common dynamics of several macroeconomic time series, and the conditional regime probabilities corresponding to the states through which this factor evolves. In other words, this approach incorporates nonlinear dynamics into the common factor extraction by combining the unobserved component model of Stock and Watson with the Markov regime-switching methodology of Hamilton. This permits reflecting two defining features of the business cycle put forward by Burns and Mitchell (1946) and stressed by Diebold and Rudebusch (1996) in their survey of the modern turning points modeling, namely: comovement of the individual macroeconomic series within the cycle and asymmetric business cycle dynamics, when the behavior of the economy during expansions is different from that in the recessions.

The model of a single dynamic common factor with Markov switching is as follows:

$$\Delta y_t = \gamma_i(L)\Delta C_t + u_t \tag{2.20}$$

$$\Delta C_t = \mu(s_t) + \phi(L)\Delta C_{t-1} + \varepsilon_t \tag{2.21}$$

$$\Psi(L)u_t = \eta_t \tag{2.22}$$

where y_t is the $n \times 1$ vector of the observable time series; C_t is the dynamic common factor in levels; u_t is the $n \times 1$ vector of the idiosyncratic components; s_t is the regime variable taking m values, where m is the number of the regimes. Thus, for $m = 2$, $s_t = 0, 1$. The model is basically the same as that of Stock and Watson, apart from having a very important extension — that of the regime switching. In this model the intercept term, $\mu(s_t)$, and possibly the variance of the common factor disturbance, $\sigma^2_\varepsilon(s_t)$, are made state-dependent, that is, they are different for the different regimes.

As in the Stock and Watson (1991) model, when the component series, y_t, are neither integrated nor cointegrated, they enter the model in the first differences and not in the levels.

The shocks to the common and specific factors are assumed to be serially and mutually uncorrelated and to be normally distributed. Moreover, the variance of the common factor disturbance may be state-dependent:

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2(s_t) & O \\ O & \Sigma_\eta \end{pmatrix} \right)$$

Again, as in Stock and Watson (1991) the lag polynomial matrices for the specific components, Ψ_j $(j = 1, ..., q)$, are assumed to be diagonal.

The transition probabilities, $p_{ij} = \Pr ob(s_t = j|s_{t-1} = i)$, sum up to one when added across all the possible state for the given regime in the previous period: $\sum_{j=1}^m p_{ij} = 1$ $\forall i$ for m states.

The above model can be expressed in the following state-space form:

Measurement equation:

$$\Delta y_t = Ax_t \tag{2.23}$$

Transition equation:

$$x_t = \alpha(s_t) + Cx_{t-1} + v_t \tag{2.24}$$

where $x_t = \begin{pmatrix} \Delta C_t & ... & \Delta C_{t-p+1} & u_{1t} & ... & u_{n,t-q_n+1} \end{pmatrix}'$ is the $k \times 1$ state vector, $k = p + \sum_{i=1}^n q_i$; $v_t = \begin{pmatrix} \varepsilon_t & ... & 0 & \varepsilon_{1t} & ... & 0 \end{pmatrix}$ is the $k \times 1$ vector of the disturbances to the state vector, and the intercept $k \times 1$ vector is defined as: $\alpha(s_t) = \begin{pmatrix} \mu_1 s_t + \mu_2 & 0 & ... & 0 \end{pmatrix}$.

The system matrices have the same form as in the case of the linear common dynamic factor model — see previous section.

The Markov dynamics can be expressed as in Hamilton (1994) with slight modifications:

$$p(s_t = j, s_{t-1} = i|Y_{t-1}) = p(s_t = j|s_{t-1} = i) \sum_{h=1}^{m} p(s_{t-1} = i, s_{t-2} = h|Y_{t-1}) \tag{2.25}$$

$$p(s_t = j, s_{t-1} = i|Y_t) = \frac{p(s_t = j, s_{t-1} = i|Y_{t-1})f(y_t|s_t = j, s_{t-1} = i, Y_{t-1};\theta)}{f(y_t|Y_{t-1};\theta)} \tag{2.26}$$

where

$$f(y_t|Y_{t-1};\theta) = \sum_{j=1}^{m}\sum_{i=1}^{m} p(s_t = j, s_{t-1} = i|Y_{t-1})f(y_t|s_t = j, s_{t-1} = i, Y_{t-1};\theta) \tag{2.27}$$

and θ is a vector of unknown parameters.

The MS-CF can be estimated by the method of maximum likelihood (see, for example, Kim (1994), Kim and Nelson (1999c)). The log-likelihood function is constructed by expressing the model in the state-space form using the Kalman filter. The Kalman filter recursions would be as:

$x^{i,j}_{t|t-1} = \alpha_j + C_j x^{i}_{t-1|t-1}$
$P^{i,j}_{t|t-1} = C_j P^{i}_{t-1|t-1} C'_j + Q_j$
$u^{i,j}_t = y_t - A_j x^{i,j}_{t|t-1}$
$H^{i,j}_t = A_j P^{i,j}_{t|t-1} A'_j + R$
$x^{i,j}_{t|t} = x^{i,j}_{t|t-1} + P^{i,j}_{t|t-1} A'_j (H^{i,j}_t)^{-1} u^{i,j}_t$
$P^{i,j}_{t|t} = (I - P^{i,j}_{t|t-1} A'_j (H^{i,j}_t)^{-1} A_j) P^{i,j}_{t|t-1}$

The likelihood function for the $t-th$ observation conditional on the states in the current and previous periods is as follows:

$$f(y_t|s_t = j, s_{t-1} = i, Y_{t-1};\theta) = (2\pi)^{-n/2}|H^{i,j}_t|^{-1/2}\exp[-\frac{(u^{i,j}_t)'(H^{i,j}_t)^{-1}u^{i,j}_t}{2}] \tag{2.28}$$

However, as can be seen from the above Kalman filter recursive equations, MLE has one serious drawback: due to the dependence of the current states on the past ones, the number of the cases to consider increases m times at each iteration, making the estimation virtually impossible. Therefore to render it feasible one has to apply approximations the consequences of which are not fully understood. Kim (1994) proposes the following approximation:

$$x^{j}_{t|t} = \frac{\sum_{i=1}^{m} p(s_t = j, s_{t-1} = i|Y_{t-1}) x^{i,j}_{t|t}}{\sum_{i=1}^{m} p(s_t = j, s_{t-1} = i|Y_{t-1})} \tag{2.29}$$

$$P^j_{t|t} = \frac{\sum_{i=1}^{m} p(s_t = j, s_{t-1} = i|Y_{t-1})[P^{i,j}_{t|t} + (x^j_{t|t} - x^{i,j}_{t|t})(x^j_{t|t} - x^{i,j}_{t|t})']}{\sum_{i=1}^{m} p(s_t = j, s_{t-1} = i|Y_{t-1})} \tag{2.30}$$

One solution to the approximation problem may be the estimation based on the Gibbs sampler — an approach belonging to the family of Markov chain Monte Carlo methods, as Kim and Nelson (1998, 1999b, 1999c) propose. Gibbs sampling allows to avoid the approximations. The idea of the Gibbs sampler is to construct a joint distribution function for all the parameters of interest. Then random drawings of some parameters of this distribution are made conditional on the rest of the parameters. The estimate of the parameters is their mean value computed as an average across large enough number of drawings. For more details on use of Gibbs sampler in the estimation of the Markov-chain models see Smith and Roberts (1993).

Another extension of the MS-CF model is the modelling of more than one common factor, as Kim and Piger (2002) do, introducing both common trend and common transitory component having Markov switching dynamics, since the dynamics of the observed time series may be much more complex to be explained by a single underlying common factor.

2.4 Conclusion

In this chapter we reviewed a number of approaches to constructing the common coincident and leading indicators and to recognizing and forecasting the business cycle turning points. These tasks are very important from the government policy standpoint, since the timely and accurate estimates of the coincident economic indicators, characterizing the state of affairs in the economy, together with the prediction of the changes in the regime may allow to the decision makers to evaluate the situation correctly and to take the appropriate policies.

It seems that, out of the recession predicting techniques considered in this chapter, for the exploration and forecasting of the turning points of a regional economy the best solution is offered by the dynamic factor model with the regime switching. On the one hand, it supplies an estimate of a composite economic indicator, which may be interpreted as an index of the overall economic activity of the region or country. On the other hand, as a byproduct it provides the conditional regime probabilities without making use of the official business cycle chronologies, which are very unlikely to be available at the regional level. Moreover, the dynamic common factor model can be modified to estimate a leading composite indicator and hence makes it possible to predict the future evolution of the economy and of its cyclical movements, in particular.

From the theoretical point of view, the attractiveness of the nonlinear common dynamic factor model lies in that it combines the advantages of the dynamic factor method of Stock and Watson, which captures the comovement feature of the business cycle, with the those of the Markov regime-switching model which allows for the asymmetry between various phases of the business cycle.

Furthermore, this approach is open to introducing a rich set of fruitful modifications making the model more flexible and the estimation more efficient. Among them it is noteworthy to mention again the mixed-frequency data approach by Mariano and Murasawa (2003), the Markov-switching model with more than one common factor as exemplified by Kim and Piger (2002), and the composite indicator model with the time-varying transition probabilities as in Kim and Yoo (1995).

Chapter 3

Bifactor Models with Markov Switching

3.1 Bifactor Model with Granger-Causality Mechanism

As we have seen in sections 2.2 and 2.3, in the modern macroeconomic literature many efforts are devoted to identifying a hypothetical coincident economic indicator which represents a general economic activity and allows to trace the evolution of the business cycle. It is designed to serve as a reference time series to judge about the state of the affairs in the economy.

With respect to this common coincident indicator one can then define the leading and lagging macroeconomic variables. The former of these series are especially important since they permit to predict the changes in the state of the economy before they have occurred.

Normally, however, the leading series are not aggregated into a common leading factor. The evolution of the common coincident factor is conditioned on each of them individually, either directly through a VAR system of the common coincident factor and individual leading observed time series as in Stock and Watson (1991), Chauvet and Potter (2000) or via the time-varying transition probabilities which depend on the individual leading variables as in Kim and Yoo (1995).

In this section we introduce a bifactor model where one of the latent factors is postulated as a composite leading indicator (CLI), while the second factor is taken to be the composite coincident indicator (CCI). A one-way

Granger causality is assumed to exist coming from the former common factor to the latter one. The common leading and coincident factors are estimated from a set of the observed time series which is split into a subset of leading and a subset of coincident variables.

First, we consider a linear model with leading and coincident factor following an AR process. Next, we add a regime-switching dynamics to take into account the possible asymmetries between the recession and expansion phases of the business cycle captured by both common latent factors.

The linear specification of the two-factor model is presented in the subsection 3.1.1, while subsection 3.1.2 contains a description of the model with nonlinear dynamics. In subsection 3.1.3 we apply our models to the artificial data in order to see how well these models reflect the true data-generating process. Section 3.1.4 discusses the application of the model to the empirical data characterizing the Post-World War II evolution of the US economy. Subsection 3.1.5 summarizes the main results of this section.

3.1.1 Linear model

We consider a set of the observed time series, some of which may be defined as leading while the rest of them are treated as the coincident series. The common dynamics of the time series belonging to each of these groups are "explained" by a common factor: leading factor corresponding to the first group and coincident factor corresponding to the second group. The idiosyncratic dynamics of each time series in particular are captured by one specific factor per each observed time series. Therefore the model can be written as follows:

$$\Delta y_t = \Gamma \Delta f_t + u_t \tag{3.1}$$

where $\Delta y_t = (\Delta y_{Lt} \mid \Delta y_{Ct})'$ is the $n \times 1$ vector of the observed time series in the first differences; $\Delta f_t = (\Delta f_{Lt} \mid \Delta f_{Ct})'$ is the 2×1 vector of the latent common factors in the first differences; $u_t = (u_{Lt} \mid u_{Ct})'$ is the $n \times 1$ vector of the latent specific factors; Γ is the $n \times 2$ factor loadings matrix linking the observed series with the common factors.

The dynamics of the latent common factors can be described in terms of a VAR model:

$$\Delta f_t = \mu + \Phi(L)\Delta f_{t-1} + \varepsilon_t \tag{3.2}$$

where μ is the 2×1 vector of the constant intercepts; $\Phi(L)$ is the sequence of p ($p = \max\{p_L, p_C\}$, where p_L is the order of the AR polynomial of the leading factor, and p_C is the order of the AR polynomial of the coincident factor) 2×2 lag polynomial matrices; ε_t is the 2×1 vector of the serially and mutually uncorrelated common factor disturbances:

$$\varepsilon_t \sim NID\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_L^2 & 0 \\ 0 & \sigma_C^2 \end{array}\right)\right)$$

We assume that the leading factor Granger-causes the coincident factor but not vice versa. This assumption means that the matrices Φ_i ($i = 1, ..., p$) are diagonal or lower diagonal for all i. For simplicity we suppose that the causality from the leading to the coincident factor is transmitted only at one lag, say τ. Thus, if $i \neq \tau$,

$$\Phi_i = \left(\begin{array}{cc} \phi_{L,i} & 0 \\ 0 & \phi_{C,i} \end{array}\right)$$

and if $i = \tau$,

$$\Phi_i = \left(\begin{array}{cc} \phi_{L,i} & 0 \\ \phi_{CL,i} & \phi_{C,i} \end{array}\right)$$

The idiosyncratic factors are by definition mutually independent and are modelled as the AR processes:

$$u_t = \Psi(L)u_{t-1} + \eta_t \tag{3.3}$$

where $\Psi(L)$ is the sequence of q ($q = \max\{q_{1,...}, q_n\}$, where q_i is the order of the AR polynomial of the $i - th$ idiosyncratic factor) $n \times n$ diagonal lag polynomial matrices and η_t is the $n \times 1$ vector of the mutually and serially uncorrelated Gaussian shocks:

$$\eta_t \sim \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix} \right)$$

To estimate this model we express it in a state-space form:

$$\Delta y_t = A x_t \tag{3.4}$$

$$x_t = \alpha + C x_{t-1} + v_t \tag{3.5}$$

where $x_t = (f_t | u_t)'$ is the state vector containing stacked on top of each other vector of common factors and the vector of specific factors; v_t is the vector of the common and idiosyncratic factors' disturbances with mean zero and variance-covariance matrix Q; α is the vector of intercepts.

$$A = \begin{pmatrix} \gamma_L & O_{n_L \times (r+p_C)} & i_{q_1} & \dots & 0 \\ O_{n_L \times r} & \gamma_C & 0 & \dots & i_{q_n} \end{pmatrix}$$

where γ_L is the $n_L \times 1$ vector of the leading factor loadings; $O_{n \times m}$ is $n \times m$ matrix of zeros; i_m is the first row of the $m \times m$ identity matrix, and $r = \max\{p_L, \tau\}$.

$$C = \begin{pmatrix} \Phi^L & & & & 0 \\ \Phi^{CL} & \Phi^C & & & \\ & & \Psi^1 & & \\ & & & \ddots & \\ 0 & & & & \Psi^n \end{pmatrix}$$

where Φ^L is the $r \times r$ matrix:

$$\Phi^L = \begin{pmatrix} \phi_L & o'_{r-p_L} \\ I_{r-1} & O_{(r-1)\times(r-p_L)} \end{pmatrix}$$

where ϕ_L is the $1 \times p_L$ row vector of the AR coefficients of the leading factor, I_n is the $n \times n$ identity matrix, and o_m is the $m \times 1$ vector of zeros.

$$\Phi^C = \begin{pmatrix} \phi_C & 0 \\ I_{p_C - 1} & o_{p_C - 1} \end{pmatrix}$$

The matrices $\Psi^1, ..., \Psi^n$ have the same structure as Φ^C.

$$\Phi^{CL} = \begin{pmatrix} o'_r \\ \phi_{CL} \end{pmatrix}$$

where ϕ_{CL} is the $1 \times r$ vector of zeros with ϕ_{CL}, τ at the $\tau - th$ position.

The unknown parameters and the latent factors may be estimated using Kalman filter recursions which are presented in section 2.3. To save space we will not reproduce them here again.

3.1.2 Nonlinear model

It was observed by many authors, among them by Diebold and Rudebusch (1996), that the model of the business cycle would be incomplete if it would not take into account both the comovement of various macroeconomic variables and the asymmetries between the phases of the cycle. The linear model presented in the previous section incorporates the phenomenon of the simultaneous changes in the levels of different individual time series. However, it lacks a mechanism which would reflect the qualitatively different behavior of these series during recessions and expansions. One of the ways to introduce this mechanism in our model is to add to it the regime-switching dynamics.

The Markov-switching dynamics is introduced through the leading factor intercept:

$$\Delta f_t = \mu(s_t) + \Phi(L)\Delta f_{t-1} + \varepsilon_t \tag{3.6}$$

where $\mu(s_t) = (\mu_L(s_t), ..., 0)'$.

s_t is the unobserved regime variable. In the two-regime (expansion-recession) case it takes two values: 0 or 1. Depending on the regime, the leading factor intercept assumes different values: low in recessions and high in expansions. Thus, the common factors grow faster during the upswings and slower (or even have negative growth rate) during the downswings of the economy.

The changes in the regimes are governed by the first-order Markov chain process, which is summarized by the transition probabilities matrix:

		Expansion	Recession
		$s_{t-1}=0$	$s_{t-1}=1$
Expansion	$s_t=0$	p_{11}	$1-p_{22}$
Recession	$s_t=1$	$1-p_{11}$	p_{22}

where $p_{ij}=prob(s_t=j|s_{t-1}=i)$.

The rest of the equations of the model remains unchanged. The state-space representation of the nonlinear two-factor model may be written as:

$$\Delta y_t = A x_t \tag{3.7}$$

$$x_t = \alpha(s_t) + C x_{t-1} + v_t \tag{3.8}$$

where $\alpha(s_t)=(\mu_L(s_t),...,0)'$.

It is worthwhile to notice that, since it is the common leading factor whose dynamics equation includes the state-dependent intercept in the current period, the conditional regime probabilities predicting the occurrence of recessions or expansions of the coincident factor are simply the conditional regime probabilities computed for the leading factor shifted forward for τ periods. Hence, the conditional regime probabilities estimated using the above model provide us with the τ-periods ahead forecast of the coincident factor regimes.

All the other system matrices are as in the linear model. Thus, we have a model expressed in the state-space form and having Markov-switching dynamics. Again, we will not reproduce here all the relevant recursions which are necessary to estimate the parameters and the unobserved state vector. These can be found in Kim (1994) or Kim and Nelson (1999c).

3.1.3 Simulated examples

To see how well our models replicate the true data-generating processes where both common factors are present and the described above causal relationship is introduced, we have generated four artificial data sets and have estimated the corresponding models using as inputs the time series which may be observed. In the first two cases the dynamics are linear, while in the last two cases the common factors follow Markov-switching process. In the case one all the leading variables have the same lead time, whereas in the case two one of the leading variables leads the coincident factor at a smaller lead than the other observed leading time series. The same distinction is maintained for the cases three and four where the regime switching is added.

For the linear case one we have generated two common latent factors and five individual observable series. The first two observed time series are leading, while the three remaining are the coincident. Both the common factors (in fact, their first differences, not levels) and the idiosyncratic components are modelled as the stationary AR(1) processes. The coincident factor is positively correlated with the leading factor at the lag $\tau = 3$. The true parameters of the DGP are presented in the column two of the Table 3.1.1 of the Appendix. The length of all these series is 540 observations, which is comparable to the length of an ordinary Post World War II monthly time series for the US economy. In the case two six observed series were simulated: three leading and three coincident. The first two leading series lead the common coincident indicator by three periods, while the third leading time series has a lead of only two periods. The true parameters are reported in the column two of the Table 3.1.2 of Appendix.

To identify the model (in both cases), we set the factor loadings of the first observable variable in each subset — leading and coincident — equal to unity. Thus, we estimate only three of five factor loadings: one for the leading factor and two for the coincident factor. The model is estimated by the maximum likelihood. The estimated parameters together with the standard errors and their p-values for case one are reproduced in the Table 1, for case two — in the Table 3.1.2. The mere observation of the true and estimated parameters' values shows that the latter are sufficiently close to the former suggesting that the proposed model estimates the parameters generated process accurately enough.

The visual comparison of the common factors profiles suggests a very high degree of similarity of the simulated and estimated common factors, especially in the case of the latent leading factor. We do not display the graphs of the simulated data here in order to save space.

In the two cases with the Markov-switching dynamics the length of the series is also 540. In the case three the first two observable time series are leading, meanwhile the last three series are coincident. The coincident factor is again correlated to the leading factor with a lag of three periods. The same identifying normalization — by setting the factor loadings of the first observed time series in each group of the variables — is used. In the case four six observed series were generated with the same leading structure as in the case two. The parameters of the true DGP for the case three are presented in the second column of Table 3.1.3 and those for the case four — in the second column of Table 3.1.4 of Appendix. The estimates replicate the true parameters with a sufficiently high degree of precision. Again, as in the case of the linear model, the estimated common factors series are very similar to the simulated common factors.

The probabilities obtained from the nonlinear model are used to build the business cycle chronology. If the probability of being currently in recession exceeds some margin (for example, 0.5) we say that the economy stays in a recession. The estimated model captures the recession dates pretty well. However, the smoothed recession probabilities sometimes miss the recessions when those have a very short duration. In contrast, the filtered probabilities give sometimes false alarms by announcing the arrival of recessions which did not take place. Thus, the smoothed probabilities turn out to be a more conservative dating tool than the filtered probabilities.

3.1.4 Real example

The linear bifactor model was estimated using the US monthly data from January 1959 to December 1998. To estimate the leading common factor the data from Watson (2000) were used, while the common coincident factor was estimated based on the four real time series borrowed from Mariano and Murasawa (2003). The list of the component series of CLI and CCI and their description are contained in Table 3.1.5. In fact, the component series we use to build our common coincident factor are those utilized by The Conference Board (USA) to construct their composite coincident index[1].

The leading component series were selected by comparing them individually to the common coincident factor computed as if it did not depend on a hypothetical leading common factor. Figure 3.1.1 shows that the correlation between these series (SFYGM3, SFYGT10 in levels and the first differences of the log of HSBR and FSNCOM), on one hand, and the growth rate of the common coincident indicator, on the other hand, is relatively high at leads 4-5. It is also very important that the series are sufficiently highly correlated among each other, thus permitting to postulate existence of a latent common factor standing behind their common evolution.

One may ask whether this approach is legitimate. Indeed, we use a common coincident indicator estimated separately to select the individual leading

[1] The list of series see on www.tcb-indicators.org/GeneralInfo/serieslist.cfm

series which will be used later — along with the individual coincident series — to compute simultaneously a common leading indicator and a common coincident indicator which depends on the former one. We think that it is legitimate as soon as the independently estimated coincident indicator can be considered as a good enough approximation of the dependent coincident indicator. Anyway, we need a point of departure to figure out the leading time series which are relevant for the estimation of common leading indicator.

The estimates of the parameters of the linear and Markov-switching modifications of the two single-factor models — common leading factor model and common coincident factor model — are reported in Tables 3.1.6 and 3.1.7, respectively.

Three model combinations were estimated: LF(0,0)-CF(0,0), LF(1,1)-CF(1,1) for linear dynamics, and LF(0,0)-CF(1,1) with state-dependent intercept and common factor residual variance for the regime-switching dynamics. The LF(p,q) stands for the common leading factor model, where the common factor follows AR(p) and the specific factors follow AR(q). CF(p,q) stands for the common coincident factor. For each of these combinations different leads between the common leading and coincident factor were tried, starting from the "zero lead" (no Granger causality) and ending with a lead of six months. The results of these experiments are displayed in Table 3.1.8. The first conclusion is that introducing a Granger causality between the leading and coincident factor seems to be a meaningful exercise — there is a significant increase in the likelihood function value when a cross-regressive term is included. Secondly, even larger positive effect is achieved when the AR(1) dynamics are allowed compared to the AR(0). Finally, in the LF(0,0) case (autoregression of the zero order, that is, static leading factor model) the "optimal lead", i.e. the lead which delivers the maximum likelihood function value, is four months, while in the LF(1,1)-CF(1,1) case the "optimal lead" is three months. It should be stressed, however, that for the LF(0,0) specifications there is no big difference in the maximum likelihood function values between the cases where lead equals 3 or 4 months. The estimates of the linear two-factor model with (1,1) specification with lead equal 3 months are presented in the second and third columns of Table 3.1.9.

Before we continue one remark is due. We do not use the information criteria to compare the models where the lead is different, since the Granger causality is introduced only through one term. Hence the number of parameters is the same regardless of the lead as long as it is positive.

The common leading and coincident factors estimated with a linear model are depicted on the two left panels of Figure 3.1.2. On the upper left panel two common factors — each estimated separately in a single-factor model — are displayed. The common factors are constructed by summing up their first differences. Thus, they are represented as random walks without drift. This is done to render the cyclical movements more visible. If we were to introduce a nonzero drift as it is done normally (e.g., by Stock and Watson (1991)), it would mask the cyclical fluctuations.

The linear common factors estimated independently have the following specifications: leading factor is (1,1), that is, both common and specific factors follow AR(1), and the coincident factor is (1,1). In the case of simultaneous estimation of the two common factors, when the coincident factor depends on the leading one, the specification is also (1,1). We can observe that in terms of the turning points the two models (with and without Granger causality between the factors) are similar, differing mainly in their "vertical profile". The latter is not surprising given that the common factors were reconstructed as random walks.

One can also see clearly on Figure 3.1.2 that the CLI is systematically leading the CCI. CLI enters both recessions and expansions a few months before than CCI does.

The next exercise was to incorporate the Markovian dynamics into the multifactor model. This was done through the regime-dependent intercept of the leading common factor. Since in the bifactor model the coincident factor depends on the leading one, the Markov-switching dynamics of the latter is transferred to the former. The parameter estimates of the Markov-switching two-factors model are contained in the last two columns of Table 8. The common leading factor is specified as AR(0), with common factor's intercept and residual variance being state-dependent, while the common coincident factor follows AR(1). The lead is set equal to 3 months.

The common factors estimated assuming the regime-switching dynamics are displayed on the two right panels of Figure 3.1.2. The specification of common leading and coincident factors computed in a single-factor model are (0,0) and (1,1), respectively. In the bifactor case the leading common factor and corresponding idiosyncratic components were modeled as AR(0), while the common coincident factor and corresponding specific factors were supposed to follow AR(2) processes. Visual inspection of all four graphs depicted on Figure 3.1.2 shows that their turning points are basically the same. One important difference is that the linear models treat the recession of the early 1990s as deeper than that of the beginning of 1980s, while the nonlinear models reverse the order.

Next, we compare the CLI's and CCI's conditional recession probabilities resulting from the single-factor models to the NBER chronology. This is done on Figure 3.1.3. It is evident that there is very close correspondence between the NBER dating and the coincident factor recessions. The composite leading indicator recessions, as it is to be expected, anticipate the NBER turning points.

Figure 3.1.4 compares the recession probabilities (filtered and smoothed) of the leading and coincident common factors contrasts them to the NBER dates. The CLI's recession probabilities are computed from the bifactor model, whereas those of the CCI are estimated in the single (coincident) factor model and are basically the same as on Figure 3.1.3. One can easily see that the recession probabilities calculated for the leading factor signal the arrival of the recession phase several periods later than the coincident factor recession probabilities do. The coincident factor model suggests that there were six recessions during the January 1959 — December 1998 period, while the bifactor model discovers five recessions. The only recession which is missing is the one in the very beginning of the sample. However, given the leading nature of the recession probabilities obtained from the bifactor model, one can assume that this recession simply "does not fit the sample". In other words, it would be found, had we the had data starting a few months earlier.

Finally, we calculate the cross-correlations between the leading and coincident common factors at different lags and leads. These cross-correlations are displayed on Figure 3.1.5. The data used to plot the picture were the same as those which are displayed on Figure 3.1.1. The cross-correlations were computed for the first differences of the common factors, not their levels. The reason is that the common factors in levels are not stationary, while their growth rates are. We can see that the maximum correlation approaches 0.5 and that it is achieved at lead 4-5 months, although being pretty high in the neighborhood of this point.

3.1.5 Summary

In this section we have introduced a common dynamic factor model with two factors: leading and coincident. Each of them represents the common dynamics of a corresponding subset of the observed time series which are classified as being leading or coincident with respect to some hypothetical "state of the economy". The common leading factor Granger-causes the common coincident factor, thus allowing to use the former in the predictions of the future values of the latter. This permits to improve the forecasting of the coincident factor because of the additional information coming from the leading variables.

We consider two models: a model with the linear dynamics and a model with the regime switching. The second model allows to take care of the asymmetries which may characterize different phases of the business cycle and therefore is more complete from the point of view of the Burns and Mitchell's definition of the business cycle as interpreted by Diebold and Rudebusch (1996).

Both models are illustrated using four artificial examples (two with the identical lead time for all the observed leading series and two with the different lead times), which show a high enough "goodness-of-fit" of these models, provided that they correspond to the true data-generating process.

Quite interesting results were obtained when the model was applied to the US monthly macroeconomic data stretching from January 1959 through December 1998. A linear and a Markov-switching bifactor models were estimated. The common coincident factor is sufficiently closely related to the common leading factor, the lead time being 3-4 months. This lead is also apparent when the recession probabilities are considered: the peaks of the low state probabilities calculated for the leading factor precede those computed for the coincident factor. Moreover, there is a tight correspondence between our estimated recession dates and those provided by the NBER. The conclusion is that we can use the two-factor model to predict the evolution of the US Post-War coincident economic indicator and the business cycle turning points in the near (up to four months) future.

3.1.6 Appendix

Table 3.1.1. True and estimated parameters of the linear bifactor model (case 1: identical lead time)

Parameter	True	Estimated	St. error	p-value
γ_1	1	–	–	–
γ_2	0.9	0.91	0.03	0.0
γ_3	1	–	–	–
γ_4	2	2.06	0.03	0.0
γ_5	1.7	1.71	0.02	0.0
ϕ_L	0.8	0.79	0.03	0.0
ϕ_C	0.7	0.70	0.03	0.0
$\phi_{CL,3}$	0.5	0.48	0.05	0.0
ψ_1	-0.3	-0.36	0.05	0.0
ψ_2	-0.7	-0.67	0.04	0.0
ψ_3	-0.5	-0.47	0.05	0.0
ψ_4	-0.2	-0.22	0.07	0.0
ψ_5	-0.8	-0.79	0.03	0.0
σ_1^2	0.25	0.26	0.03	0.0
σ_2^2	0.36	0.36	0.03	0.0
σ_3^2	0.16	0.16	0.01	0.0
σ_4^2	0.49	0.48	0.05	0.0
σ_5^2	0.81	0.81	0.06	0.0
σ_L^2	0.25	0.24	0.03	0.0
σ_C^2	0.36	0.36	0.03	0.0

Table 3.1.2. True and estimated parameters of the linear bifactor model (case 2: different lead time)

Parameter	True	Estimated	St. error	p-value
γ_1	1	–	–	–
γ_2	0.9	0.86	0.03	0.0
γ_3	1.5	1.43	0.04	0.0
γ_4	1	–	–	–
γ_5	2	2.03	0.05	0.0
γ_6	1.7	1.72	0.04	0.0
ϕ_L	0.8	0.79	0.03	0.0
ϕ_C	0.7	0.67	0.02	0.0
$\phi_{CL,3}$	0.5	0.58	0.04	0.0
ψ_1	-0.3	-0.35	0.05	0.0
ψ_2	-0.7	-0.72	0.03	0.0
ψ_3	-0.5	-0.51	0.06	0.0
ψ_4	-0.2	-0.16	0.05	0.0
ψ_5	-0.8	-0.82	0.03	0.0
ψ_6	-0.3	-0.36	0.05	0.0
σ_1^2	0.25	0.25	0.02	0.0
σ_2^2	0.36	0.36	0.02	0.0
σ_3^2	0.16	0.16	0.02	0.0
σ_4^2	0.49	0.50	0.03	0.0
σ_5^2	0.81	0.79	0.08	0.0
σ_6^2	0.64	0.64	0.05	0.0
σ_L^2	0.25	0.23	0.02	0.0
σ_C^2	0.36	0.33	0.03	0.0

Table 3.1.3. True and estimated parameters of the nonlinear bifactor model (case 3: identical lead time)

Parameter	True	Estimated	St. error	p-value
p_{11}	0.95	0.97	0.01	0.0
p_{22}	0.84	0.84	0.05	0.0
μ_{L1}	0.4	0.40	0.04	0.0
μ_{L2}	-0.6	-0.70	0.06	0.0
γ_1	1	–	–	–
γ_2	0.9	0.90	0.01	0.0
γ_3	1	–	–	–
γ_4	2	1.99	0.01	0.0
γ_5	1.7	1.69	0.01	0.0
ϕ_L	0.8	0.78	0.02	0.0
ϕ_C	0.7	0.70	0.02	0.0
$\phi_{CL,3}$	0.5	0.51	0.03	0.0
ψ_1	-0.3	-0.29	0.05	0.0
ψ_2	-0.7	-0.69	0.03	0.0
ψ_3	-0.5	-0.50	0.05	0.0
ψ_4	-0.2	-0.09	0.07	0.0
ψ_5	-0.8	-0.82	0.03	0.0
σ_1^2	0.25	0.26	0.02	0.0
σ_2^2	0.36	0.35	0.03	0.0
σ_3^2	0.16	0.16	0.01	0.0
σ_4^2	0.49	0.47	0.05	0.0
σ_5^2	0.81	0.81	0.06	0.0
σ_L^2	0.16	0.18	0.03	0.0
σ_C^2	0.36	0.37	0.03	0.0

Table 3.1.4. True and estimated parameters of the nonlinear bifactor model (case 4: different lead time)

Parameter	True	Estimated	St. error	p-value
p_{11}	0.95	0.95	0.01	0.0
p_{22}	0.84	0.86	0.03	0.0
μ_{L1}	0.4	0.39	0.03	0.0
μ_{L2}	-0.6	-0.58	0.05	0.0
γ_1	1	–	–	–
γ_2	0.9	0.91	0.01	0.0
γ_3	1.5	1.51	0.02	0.0
γ_4	1	–	–	–
γ_5	2	2.00	0.02	0.0
γ_6	1.7	1.70	0.02	0.0
ϕ_L	0.8	0.79	0.02	0.0
ϕ_C	0.7	0.71	0.02	0.0
$\phi_{CL,3}$	0.5	0.49	0.03	0.0
ψ_1	-0.3	-0.27	0.05	0.0
ψ_2	-0.7	-0.73	0.03	0.0
ψ_3	-0.5	-0.46	0.06	0.0
ψ_4	-0.2	-0.22	0.05	0.0
ψ_5	-0.8	-0.83	0.03	0.0
ψ_6	-0.5	-0.49	0.05	0.0
σ_1^2	0.25	0.25	0.02	0.0
σ_2^2	0.36	0.36	0.02	0.0
σ_3^2	0.16	0.17	0.02	0.0
σ_4^2	0.49	0.49	0.03	0.0
σ_5^2	0.81	0.64	0.07	0.0
σ_6^2	0.36	0.42	0.04	0.0
σ_L^2	0.16	0.17	0.02	0.0
σ_C^2	0.36	0.36	0.03	0.0

Table 3.1.5. The component series of the US composite leading and coincident indicators
Monthly series 1959:1 – 1998:12

Series	Short-hand	Description
Composite leading indicator		
New housing	HSBR	Authorized housing, total new house units, 10^3
Spread 3-month	SFYGM3	Spread between US Treasury bills 3-month interest rate and federal funds effective annualized rate
Spread 10-year	SFYGT10	Spread between US Treasury bills 10-year interest rate and federal funds effective annualized rate
NYSE index	FSNCOM	NYSE common stock price index
Composite coincident indicator		
Employees on nonagricultural payrolls	EMP	10^3, SA
Personal income less transfer payments	INC	10^9 1992 USD, SA, annual rate
Index of industrial production	IIP	total index, 1992=100, SA
Manufacturing and trade sales	SLS	chained 10^6 1992 USD, SA

Table 3.1.6. Estimated parameters of the single leading factor model (US macroeconomic monthly data, 1959:1-1998:12)

Parameter	Linear: -1616.06		Markov-switching: -2228.76	
	Estimated	St. error	Estimated	St. error
p_{11}	–	–	0.989	0.005
$1 - p_{22}$	–	–	0.044	0.021
μ_1	–	–	0.102	0.020
μ_2	–	–	-0.380	0.079
γ_{FSNCOM}	0.611	0.228	0.78	0.244
γ_{SFYGM3}	3.56	0.581	3.74	0.725
γ_{SFYGT1}	3.09	0.534	3.45	0.669
ϕ	0.926	0.019	–	–
ψ_{HSBR}	-0.136	0.047	–	–
ψ_{FSNCOM}	0.245	0.045	–	–
ψ_{SFYGM3}	0.769	0.066	–	–
ψ_{SFYGT1}	0.981	0.012	–	–
$\sigma^2(s_t = 1)$	0.011	0.004	0.005	0.002
$\sigma^2(s_t = 2)$	–	–	0.076	0.031
σ^2_{HSBR}	0.908	0.060	0.939	0.061
σ^2_{FSNCOM}	0.910	0.059	0.962	0.062
σ^2_{SFYGM3}	0.094	0.010	0.181	0.021
$\sigma^2_{SFYGT10}$	0.009	0.005	0.301	0.025

Table 3.1.7. Estimated parameters of the single coincident factor model (US macroeconomic monthly data, 1959:1-1998:12)

Parameter	Linear: -2312.91		Markov-switching: -2296.78	
	Estimated	St. error	Estimated	St. error
p_{11}	–	–	0.976	0.010
$1-p_{22}$	–	–	0.156	0.079
μ_1	–	–	0.143	0.039
μ_2	–	–	-0.904	0.161
γ_{INC}	0.843	0.057	0.823	0.055
γ_{IIP}	0.975	0.061	0.950	0.057
γ_{SLS}	0.658	0.052	0.638	0.049
ϕ	0.629	0.047	0.407	0.066
ψ_{EMP}	0.040	0.090	-0.010	0.037
ψ_{INC}	-0.052	0.051	-0.049	0.054
ψ_{IIP}	0.006	0.069	0.037	0.056
ψ_{SLS}	-0.322	0.047	-0.311	0.047
σ^2	0.395	0.046	0.312	0.039
σ^2_{EMP}	0.343	0.035	0.32	0.035
σ^2_{INC}	0.533	0.041	0.539	0.041
σ^2_{IIP}	0.377	0.037	0.386	0.036
σ^2_{SLS}	0.619	0.046	0.631	0.046

Table 3.1.8. Optimal lead determination in the linear case Likelihood function values corresponding to different AR order combinations of common factors and different leads

Lead	Combinations		
	Linear		Markov-switching
	LF(0,0)-CF(0,0)	LF(1,1)-CF(1,1)	LF(0,0)-CF(1,1)
0	-4899.40	-3928.97	-4541.68
1	-4873.82	-3913.83	-4526.21
2	-4865.81	-3912.07	-4525.11
3	-4859.52	**-3909.28**	-4523.90
4	**-4855.32**	-3909.57	**-4523.66**
5	-4855.97	-3912.88	-4524.71
6	-4860.29	-3916.00	-4528.57

Table 3.1.9. Estimated parameters of the two-factor model
(US macroeconomic monthly data, 1959:1-1998:12)

Parameter	Linear: -3909.28		Markov-switching: -4523.90	
	Estimated	St. error	Estimated	St. error
p_{11}	–	–	0.989	0.005
$1-p_{22}$	–	–	0.044	0.021
μ_1	–	–	0.103	0.021
μ_2	–	–	-0.383	0.079
γ_{FSNCOM}	0.621	0.257	0.781	0.243
γ_{SFYGM3}	3.46	0.574	3.70	0.718
γ_{SFYGT1}	2.92	0.519	3.42	0.663
γ_{INC}	0.858	0.059	0.854	0.058
γ_{IIP}	0.989	0.062	0.986	0.061
γ_{SLS}	0.675	0.054	0.672	0.052
ϕ_L	0.921	0.020	–	–
ϕ_C	0.488	0.057	0.504	0.053
$\phi_{CL,3}$	0.793	0.185	0.841	0.220
ψ_{HSBR}	-0.139	0.047	–	–
ψ_{FSNCOM}	0.243	0.045	–	–
ψ_{SFYGM3}	0.771	0.063	–	–
ψ_{SFYGT1}	0.980	0.012	–	–
ψ_{EMP}	0.092	0.104	0.081	0.087
ψ_{INC}	-0.057	0.059	-0.055	0.055
ψ_{IIP}	-0.002	0.068	0.001	0.035
ψ_{SLS}	-0.329	0.048	-0.327	0.047
$\sigma_L^2(s_t=1)$	0.012	0.004	0.005	0.002
$\sigma_L^2(s_t=2)$	–	–	0.077	0.032
σ_C^2	0.365	0.043	0.368	0.043
σ_{HSBR}^2	0.904	0.060	0.938	0.061
σ_{FSNCOM}^2	0.909	0.059	0.962	0.062
σ_{SFYGM3}^2	0.092	0.010	0.181	0.021
$\sigma_{SFYGT10}^2$	0.011	0.005	0.300	0.025
σ_{EMP}^2	0.352	0.034	0.349	0.034
σ_{INC}^2	0.526	0.041	0.528	0.041
σ_{IIP}^2	0.372	0.036	0.373	0.036
σ_{SLS}^2	0.608	0.046	0.610	0.045

Figure 3.1.1: Cross-correlation of composite coincident indicator and leading variables

US monthly data 1959:1-1998:12

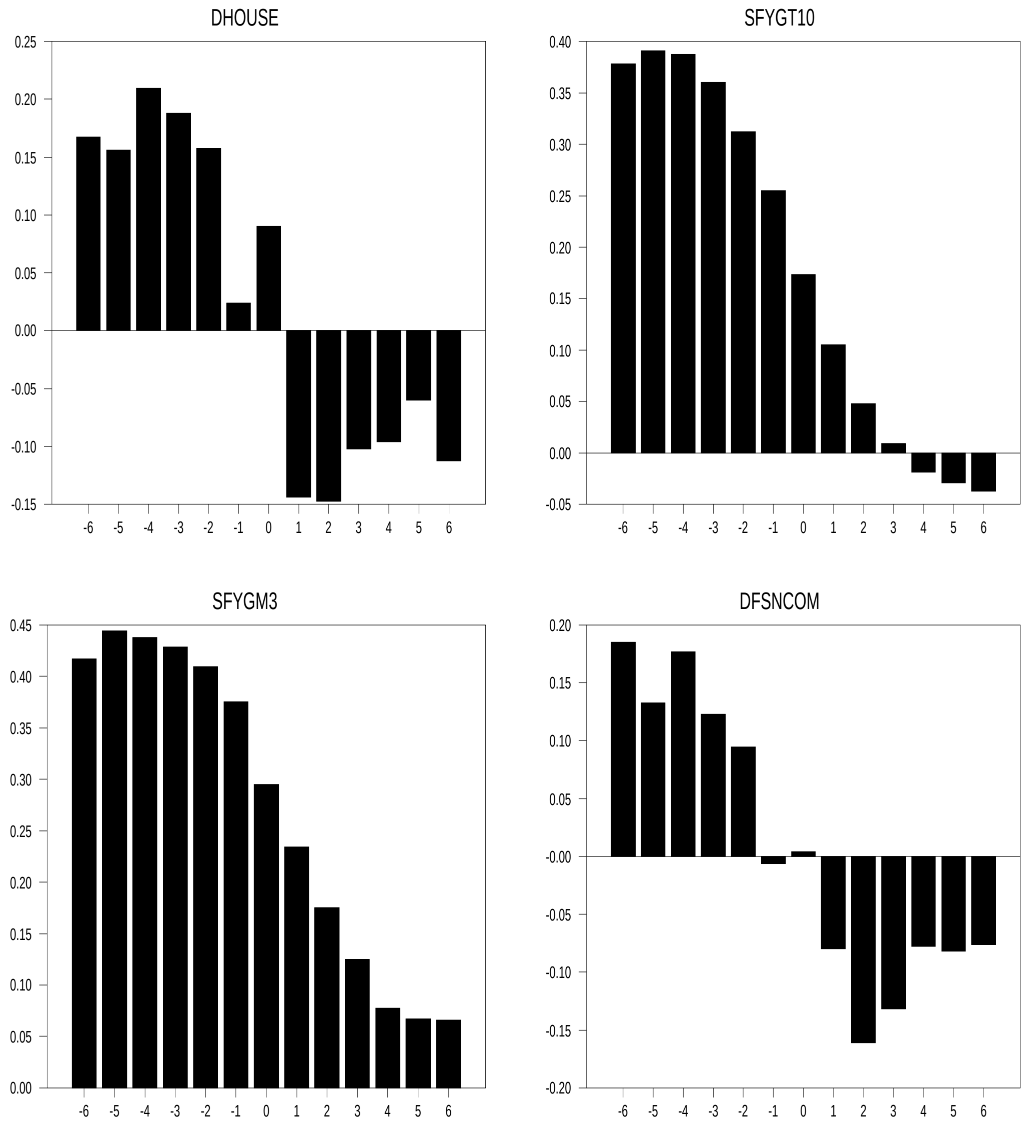


Figure 3.1.2: Composite leading and coincident indicators

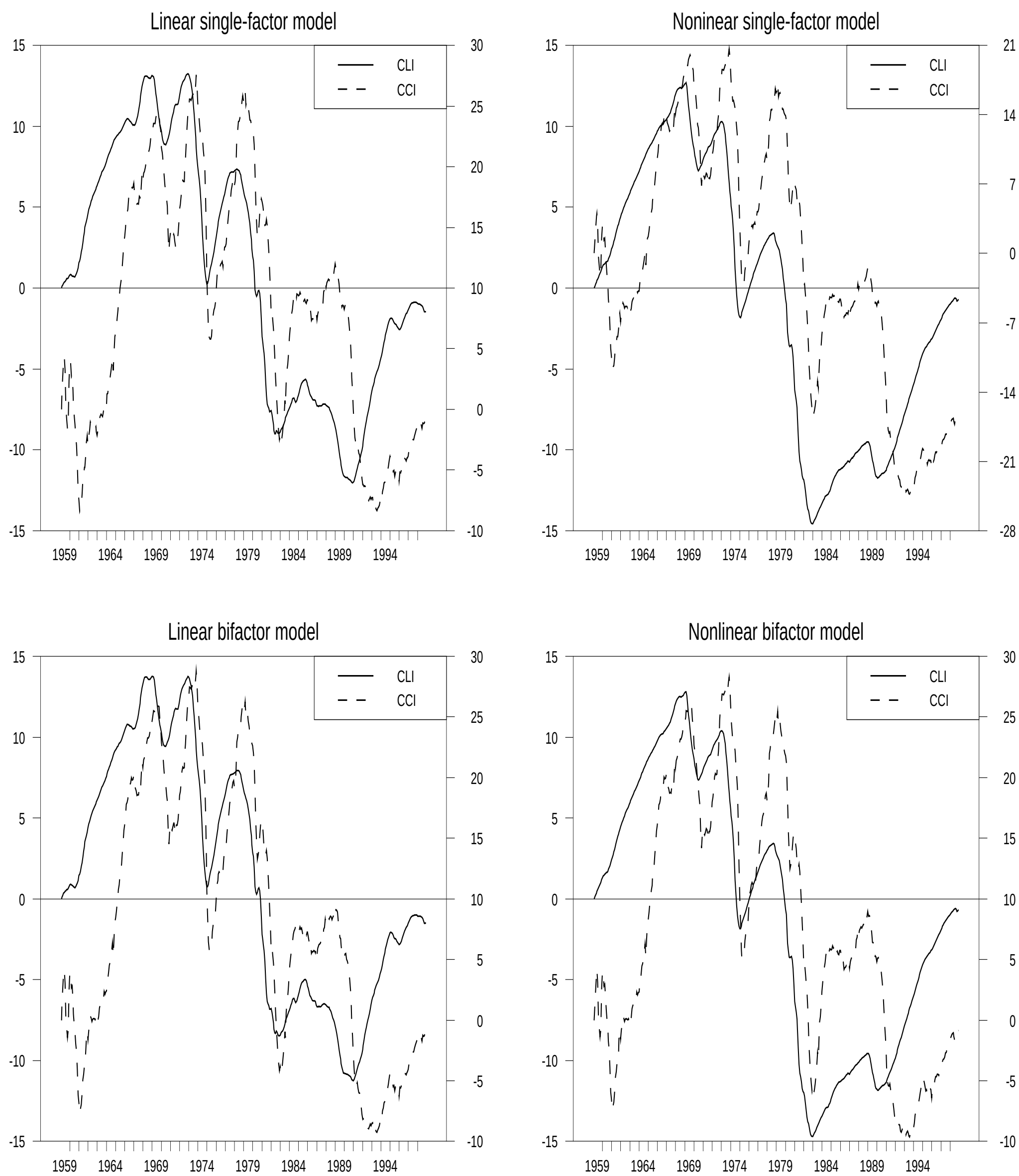


Figure 3.1.3: Single-factor model

Conditional recession probabilities vs. NBER dates

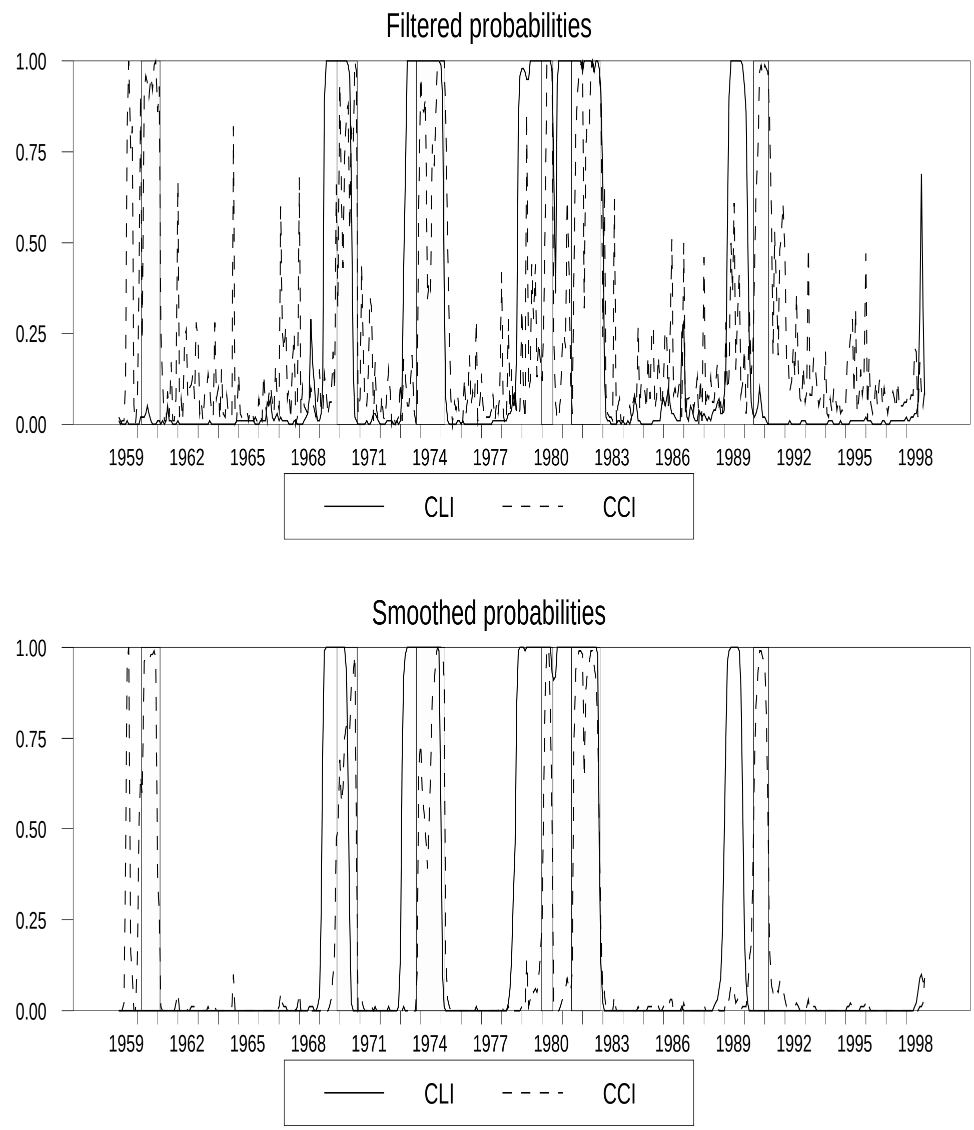


Figure 3.1.4: Bifactor model

Conditional recession probabilities vs. NBER dates

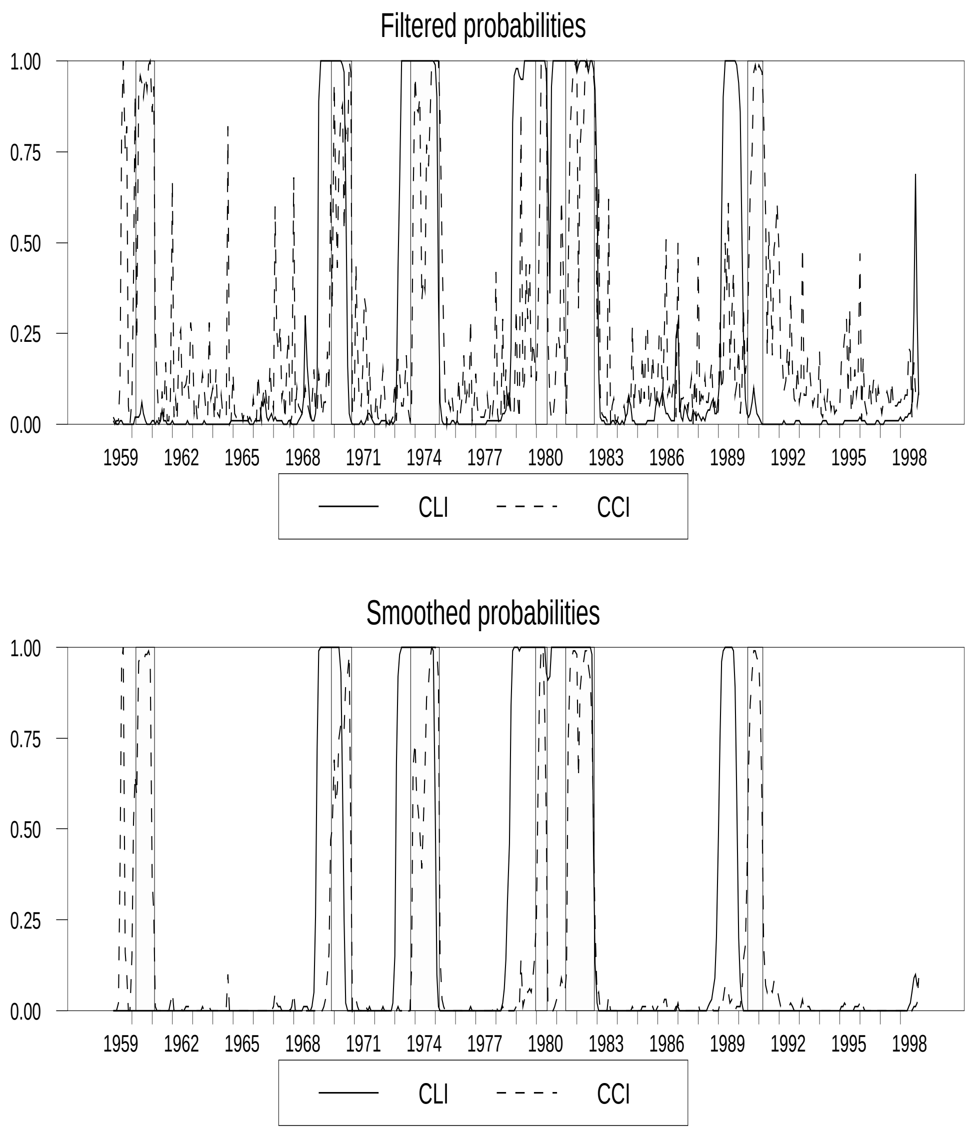
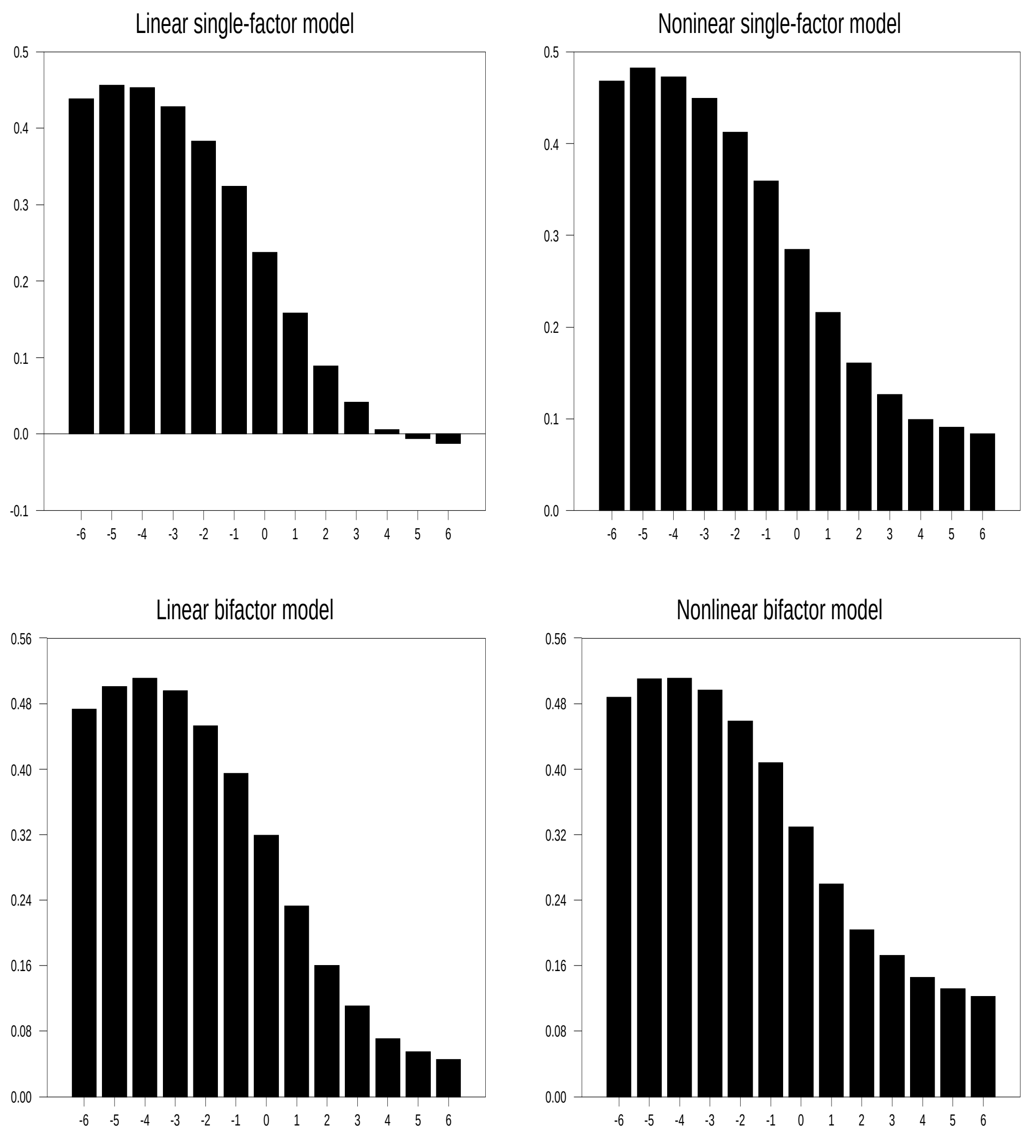


Figure 3.1.5: Cross-correlation of common leading and coincident factor

US monthly data 1959:1-1998:12



3.2 Bifactor Model with Transition Probabilities Matrix Mechanism

In this section we continue investigating the composite leading and coincident indicators. As in section 3.1 the common leading and coincident factors are estimated from two sets of observed data: leading and coincident time series. However, the angle is changed compared to the previous section: instead of trying to capture the leading relationship through a bivariate VAR, we work with the transition probabilities matrix which governs the shifts in the regimes of the two common factors. Here we base our analysis on that of Phillips (1991) who showed, using a bivariate regime switching model, how the time lags between two observed variables with Markov-switching dynamics can be expressed in terms of the transition probabilities. He applied his model to the observed variables, while we generalize his approach to the latent common factors.

The rest of the section is organized as follows. Subsection 3.2.1 contains the description of the three bifactor models with regime-switching dynamics differing in the way the cyclical evolution of both factors is defined. In section 3.2.2 these three Markov-switching models with leading and coincident common factors are estimated using the US Post-War monthly macroeconomic data. Section 3.2.3 summarizes the outcomes obtained in the section.

3.2.1 Model

If the model in section 3.1 can be endowed either with a linear or with a regime-switching dynamics, although the nonlinear dynamics fit better the business cycle definition we are using in this thesis, the model we are examining here indispensably requires Markov switching. The reason lies in the very way the leading relationship is modeled.

We consider a set of the observed time series, some of which are defined as leading while the rest of them are the coincident series. The common dynamics of the time series belonging to each of these groups are underlined by a common factor: leading corresponding to the first group and coincident corresponding to the second group. The idiosyncratic dynamics of each time series in particular are captured by one specific factor per each observed time series. Therefore the model can be written as:

$$\Delta y_t = \Gamma \Delta f_t + u_t \tag{3.9}$$

where $\Delta y_t = (\Delta y_{Lt} \mid \Delta y_{Ct})'$ is the $n \times 1$ vector of the observed time series in the first differences; $\Delta f_t = (\Delta f_{Lt} \mid \Delta f_{Ct})'$ is the 2×1 vector of the latent common factors in the first differences; $u_t = (u_{Lt} \mid u_{Ct})'$ is the $n \times 1$ vector of the latent specific factors; Γ is the $n \times 2$ factor loadings matrix linking the observed series with the common factors.

The dynamics of the latent common factors can be described in terms of a nonlinear VAR model:

$$\Delta f_t = \mu(s_t) + \Phi(L)\Delta f_{t-1} + \varepsilon_t \tag{3.10}$$

where $\mu(s_t) = \{\mu_L(s_t), \mu_C(s_t)\}$ is the 2×1 vector of the state-dependent intercepts of the common leading and coincident factors, correspondingly, which take different values depending on the regime; $\Phi(L)$ is the sequence of p ($p = \max\{p_L, p_C\}$, where p_L is the order of the AR polynomial of the leading factor, and p_C is the order of the AR polynomial of the coincident factor) 2×2 lag polynomial matrices; ε_t is the 2×1 vector of the serially and mutually uncorrelated common factor disturbances with possibly state-dependent variance:

$$\varepsilon_t \sim NID\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_L^2(s_t) & 0 \\ 0 & \sigma_C^2(s_t) \end{array}\right)\right)$$

s_t is the unobserved regime variable. In the two-regime (expansion-recession, or high-low) case it takes two values: 0 or 1. Depending on the regime, the common factor's intercept assumes different values: low in recessions and high in expansions. Thus, the common factors grow faster during the upswings and slower (or even have negative growth rate) during the downswings of the economy.

The changes in the regimes are governed by the first-order Markov chain process, which is summarized by the transition probabilities matrix, whose characteristic element is $p_{ij} = prob(s_t = j|s_{t-1} = i)$.

Since we have two common factors each of which passes through its own low and high regimes, the whole process is to be cast in a four regimes framework as it is done in Phillips (1991). Namely:

	Regime 1	Regime 2	Regime 3	Regime 4
Composite state variable	$s_t = 0$	$s_t = 1$	$s_t = 2$	$s_t = 3$
Leading state variable	$s_t^L = 0$	$s_t^L = 1$	$s_t^L = 0$	$s_t^L = 1$
Coincident state variable	$s_t^C = 0$	$s_t^C = 0$	$s_t^C = 1$	$s_t^C = 1$

where s_t^L and s_t^C are the unobserved state variables for leading factor and coincident factor, respectively.

The way the unobserved regimes of leading indicator and coincident indicator are interrelated affects the form of the 4×4 transition probabilities matrix. First, we may suppose that the state variables s_t^L and s_t^C are mutually independent. In that case the transition matrix, π, looks like:

$$\begin{pmatrix} p_{11}^L p_{11}^C & (1-p_{11}^L)p_{11}^C & p_{11}^L(1-p_{11}^C) & (1-p_{11}^L)(1-p_{11}^C) \\ (1-p_{22}^L)p_{11}^C & p_{22}^L p_{11}^C & (1-p_{22}^L)(1-p_{11}^C) & p_{22}^L(1-p_{11}^C) \\ p_{11}^L(1-p_{22}^C) & (1-p_{11}^L)(1-p_{22}^C) & p_{11}^L p_{22}^C & (1-p_{11}^L)p_{22}^C \\ (1-p_{22}^L)(1-p_{22}^C) & p_{22}^L(1-p_{22}^C) & (1-p_{22}^L)p_{22}^C & p_{22}^L p_{22}^C \end{pmatrix}$$

In fact, $\pi = \pi^C \otimes \pi^L$, where π^L and π^C are the transition probabilities matrices for state variables s_t^L and s_t^C.

A second hypothesis is that there exist not two different state variables, but only one representing a single process and that both common factors enter into each regime simultaneously, without any lags among the two factors. In other words, the recessions (expansions) of the leading factor are the recessions (expansions) of the coincident factor. In this case there is no sense to talk about a leading factor, because both factors are coincident. This case may be represented with an ordinary two-regime transition probabilities matrix:

$$\pi = \begin{pmatrix} p_{11} & 1-p_{11} \\ 1-p_{22} & p_{22} \end{pmatrix}$$

Under the third hypothesis the two unobserved processes are interrelated, with leading factor entering the recessions (expansions) several periods earlier than the coincident indicator. As Phillips (1991) remarks, the model with an integer lag exceeding one period would require a Markov process with the order higher than 1. However, the real-valued (positive) lag can be modeled with a first-order Markov process by constructing the following transition probabilities matrix:

$$\pi = \begin{pmatrix} p_{11} & 1-p_{11} & 0 & 0 \\ 0 & 1-\frac{1}{A} & 0 & \frac{1}{A} \\ \frac{1}{B} & 0 & 1-\frac{1}{B} & 0 \\ 0 & 0 & 1-p_{22} & p_{22} \end{pmatrix}$$

where A and B are the expected leads in the recession and expansion, correspondingly.

The quantity $1-\frac{1}{A}$ stands for the probability of being now in the regime (leading factor — recession, coincident factor — expansion) given that in the preceding period the economy has been in the same regime. In other words, it is a situation when the leading factor has already entered the downswing, while the coincident indicator still stays in the upswing. The expected duration of this regime ($s_t = 1$) is:

$$A = p(s_t = 1|s_t = 1) + p(s_{t+1} = 1|s_t = 1) + p(s_{t+2} = 1|s_t = 1) + ...$$

or

$$A = 1 + p(s_t = 1|s_{t-1} = 1) + p(s_t = 1|s_{t-1} = 1)^2 + ...$$

that is

$$A = \frac{1}{1-p(s_t=1|s_{t-1}=1)}$$

Similarly,

$$B = \frac{1}{1-p(s_t=2|s_{t-1}=2)}$$

where $s_t = 2$ corresponds to the leading factor being already in the new expansion and coincident factor staying still in the old contraction.

Thus, we can analyze the three above stated cases — independent cyclical evolution of leading and coincident factors (let us call it model 1), identical cyclical evolution of both factors (model 2), and similar cyclical evolution with coincident factor lagging behind the leading indicator (model 3) — and compare the resulting three hypotheses to check whether the composite leading indicator is really leading and if so, how far it is advancing the composite coincident indicator.

We assume for simplicity that the two factors are related only through the transition probabilities, no correlation and no Granger-causality coming from the leading factor to the coincident factor and vice versa as it was done in section 3.1. This assumption implies that the matrices Φ_i ($i = 1, ..., p$) are diagonal or lower diagonal for all i:

$$\Phi_i = \begin{pmatrix} \phi_{L,i} & 0 \\ 0 & \phi_{C,i} \end{pmatrix}$$

The idiosyncratic factors are by definition mutually independent and are modelled as the AR processes:

$$u_t = \Psi(L)u_{t-1} + \eta_t \tag{3.11}$$

where $\Psi(L)$ is the sequence of q ($q = \max\{q_{1,...}, q_n\}$, where q_i is the order of the AR polynomial of the $i-th$ idiosyncratic factor) $n \times n$ diagonal lag polynomial matrices and η_t is the $n \times 1$ vector of the mutually and serially uncorrelated Gaussian shocks:

$$\eta_t \sim \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix} \right)$$

To estimate this model we express it in a state-space form:

$$\Delta y_t = Ax_t \tag{3.12}$$

$$x_t = \alpha(s_t) + Cx_{t-1} + v_t \tag{3.13}$$

where $x_t = (f_t|u_t)'$ is the state vector containing the vectors of common factors and of specific factors, stacked on top of each other; v_t is the vector of the common and idiosyncratic factors' disturbances with mean zero and variance-covariance matrix Q; $\alpha(s_t) = (\mu_L(s_t), \mu_C(s_t), ..., 0)'$ is the state-dependent vector of intercepts.

$$A = \begin{pmatrix} \Gamma_L & o_{n_L} & i_{q_1} & ... & 0 \\ O_{n_L \times r} & \gamma_C & 0 & ... & i_{q_n} \end{pmatrix}$$

where Γ_L is the $n_L \times (r-1)$ matrix of the leading factor loadings:

$$\Gamma_L = \begin{pmatrix} \gamma_{L,1} & 0 & & 0 \\ \gamma_{L,2} & 0 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \gamma_{L,n_L} \end{pmatrix}$$

in which the position of each leading factor loading depends on the lead time of a corresponding observed time series.

$O_{n\times m}$ is $n \times m$ matrix of zeros; o_m is the $m \times 1$ vector of zeros; i_m is the first row of the $m \times m$ identity matrix, and $r = \max\{p_L, \tau\}$.

$$C = \begin{pmatrix} \Phi^L & O & & & 0 \\ O & \Phi^C & & & \\ & & \Psi^1 & & \\ & & & \ddots & \\ 0 & & & & \Psi^n \end{pmatrix}$$

where Φ^L is the $r \times r$ matrix:

$$\Phi^L = \begin{pmatrix} \phi_L & o'_{r-p_L} \\ I_{r-1} & O_{(r-1)\times(r-p_L)} \end{pmatrix}$$

where ϕ_L is the $1 \times p_L$ row vector of the AR coefficients of the leading factor, I_n is the $n \times n$ identity matrix, and o_m is the $m \times 1$ vector of zeros.

$$\Phi^C = \begin{pmatrix} \phi_C & 0 \\ I_{p_C-1} & o_{p_C-1} \end{pmatrix}$$

The matrices $\Psi^1, ..., \Psi^n$ have the same structure as Φ^C.

There are different ways of estimating the unknown parameters and the latent factors (maximum likelihood, EM, MCMC techniques — see Kim and Nelson (1999c) for more details). Here we applied the maximum likelihood method with log-likelihood function obtained using Kalman filter recursions.

3.2.2 Real example

The linear two-factor model was estimated using the US monthly data from January 1959 to December 1998. These are same data we have used in section 3.1 to construct the CLI and CCI. Recall that the list and the description both of the leading and of the coincident series can be found in Table 3.1.1.

The leading time series were selected following the same logic as explained in subsection 3.1.4. That is, they were compared individually to the composite coincident indicator computed as if it were not dependent of a hypothetical composite leading indicator. As we have seen already, the correlation between these series (SFYGM3, SFYGT1 in levels and the first differences of the log of HSBR and FSNCOM), on the one hand, and the growth rate of the common coincident indicator, on the other hand, is relatively high at leads 4-5. See Figure 3.1.1 in the Appendix to section 3.1.

All the three models were estimated under the identifying assumption of the first factor loading for each common factor being equal 1. The parameter estimates, together with their standard errors and corresponding p-values, of the three models can be found in Tables 3.2.1-3.2.3 of Appendix. In all the models both the common and specific factors are supposed to follow serially uncorrelated process (AR(0)).

Both the model with independent leading and coincident indicators and the model with the leading and coincident indicators having the same cyclical dynamics with coincident factor lagging behind the leading one seem to bring significant increase in the maximum likelihood compared to the model with two common factors having common cyclical dynamics. The independent factors model turns out to slightly outperform model 2. However, model 2 delivers some additional information of interest, namely the leads in low and high states. According to Table 3.2.2 the leading factor enters the recessions on average approximately five months earlier than the coincident factor and goes into the expansions roughly nine months earlier than the common coincident indicator does.

Figure 3.2.1 represents the three estimates of the common leading and coincident factors corresponding to the three models: independent composite leading and coincident indicators, two composite indicators with common dynamics, and composite leading and coincident indicators with common dynamics. The common factors were reconstructed as the partial sums of their growth rates obtained as an output of the estimation procedure. Therefore they follow random walks. Nevertheless, their profiles are quite similar regardless of the model. The coincident factors in all three cases seem to lag almost always the leading indicators.

Next, we consider the conditional recession (low intercept regime) probabilities for each of the three bifactor models we examine in this section.

Figure 3.2.2 illustrates model 1 with CLI and CCI being independent. The upper panel compares the conditional (filtered and smoothed) probabilities of CLI having low intercept (these are computed as the sum of the conditional probabilities of regimes 2 and 4) to the NBER dates. Recall that regime 2 corresponds to the leading indicator signalling the low state and coincident indicator still being in the high state, while in the regime 4 both CLI and CCI are in the low state. In four cases out of six the peaks in the probabilities precede the beginning of the NBER recessions. In the other two cases the recessions are signalled by two minor and short-lived increases in the probabilities which lead the NBER contractions.

The lower panel of Figure 3.2.2 shows the conditional recession probabilities of the composite coincident indicator compared to the NBER business cycle chronology. The former are calculated as the sum of probabilities of regime 3 (CCI is still in recession, while CLI is already in expansion) and regime 4 (both indicators are in contraction). CCI also behaves as a leading indicator entering the low-intercept state well before than NBER recessions start. This circumstance is a bit surprising, since both common factors are supposed to be independent and follow their own nonlinear dynamics.

Figure 3.2.3 shows the conditional recession probabilities of the model 2. In this case, since both composite indicators have an identical cyclical dynamics, there are only two states: low intercept and high intercept. Hence we have only one panel displaying the (filtered and smoothed) recession probabilities corresponding to these common nonlinear dynamics against the NBER dates. These probabilities are slightly leading the NBER cycle, the leading factor playing more prominent role in the determination of the conditional probabilities.

Figure 3.2.4 is illustrating model 3 where CLI and CCI are interrelated with lags depending on the cyclical phase. The layout of Figure 3.2.4 is similar to that of Figure 3.2.2. On the upper panel one can see the CLI's conditional recession probabilities (sum of regime 2 and regime 4 probabilities) versus the NBER dates. While on the lower panel the CCI's recession probabilities (sum of regime 3 and regime 4 probabiliites) are plotted versus the NBER business cycle chronology. The CLI's model-derived probabilities are unmistakenly leading the NBER's business cycle, although the lead time is not constant. The CCI recession probabilities, unlike on Figure 3.2.2, are generally coinciding with the NBER dates and are not leading the latter. While the first five contraction are well enough recognized, the sixth recession is somewhat "oversized" compared to the official NBER dating — the contractionary regime probabilities coming up with a longer recession. So, in the model 3 the correspondence between our model-derived dating and that of the NBER seems to be better than in models 1 and 2.

3.2.3 Summary

In this section we have introduced a multifactor model with two common factors (leading and coincident) having a regime-switching dynamics. Each of them represents the common dynamics of a corresponding subset of the observed time series which are classified as being leading or coincident with respect to some hypothetical "state of the economy". The common leading factor advances faster than the common coincident factor which can be reflected in the transition probabilities matrix. This permits improving the prediction of the coincident factor because of the additional information coming from the leading variables. In addition, different leads with respect to the common coincident indicator for the individual leading time series are allowed, which makes the model more flexible and realistic, since in the real life the leading time series rarely lead the coincident factor for the same periods of time.

We consider three nonlinear models: (1) a model with the independent leading and coincident factors; (2) a model with two common factors having exactly the same cyclical dynamics, and (3) a model with leading and coincident factors having the common cyclical dynamics allowing for the different lead durations in recessions and expansions.

These models were applied to the US monthly macroeconomic data stretching from January 1959 to December 1998. The common coincident factor, as model 2 reveals, is lagging behind the common leading factor 5 months when entering into the recessions and around 9 months when going into the expansions. Moreover, there exists a tight correspondence between our estimated recession chronologies and those provided by the NBER. The conclusion is that it is feasible to use the bifactor model to forecast the evolution of the US Post-World War II coincident economic indicator and its turning points in the near (5 months for downswings and 9 months for upswings) future.

3.2.4 Appendix

Table 1. Estimated parameters of model 1
Log-likelihood: -4633.5

Parameter	Estimated	St. error	p-value
$p_{L.11}$	0.988	0.01	0.0
$p_{L.22}$	0.921	0.04	0.0
$p_{C.11}$	0.980	0.01	0.0
$p_{L.22}$	0.892	0.04	0.0
μ_{L1}	0.084	0.02	0.0
μ_{L2}	-0.527	0.11	0.0
μ_{C1}	0.255	0.04	0.0
μ_{L2}	-1.330	0.132	0.0
γ_{FSNCOM}	0.745	0.241	0.0
γ_{SFYGM3}	3.520	0.693	0.0
γ_{SFYGT1}	3.46	0.672	0.0
γ_{INC}	0.786	0.05	0.0
γ_{IIP}	0.898	0.06	0.0
γ_{SLS}	0.655	0.06	0.0
σ^2_{HSBR}	0.936	0.06	0.0
σ^2_{FSNCOM}	0.963	0.06	0.0
σ^2_{SFYGM3}	0.228	0.03	0.0
σ^2_{SFYGT1}	0.256	0.03	0.0
σ^2_{EMP}	0.279	0.04	0.0
σ^2_{INC}	0.554	0.04	0.0
σ^2_{IIP}	0.419	0.04	0.0
σ^2_{SLS}	0.689	0.05	0.0
σ^2_{L}	0.018	0.007	0.01
σ^2_{C}	0.380	0.042	0.0

Table 2. Estimated parameters of model 2
Log-likelihood: -4682.6

Parameter	Estimated	St. error	p-value
p_{11}	0.988	0.01	0.0
$1-p_{22}$	0.066	0.03	0.01
μ_{L1}	0.084	0.02	0.0
μ_{L2}	-0.434	0.09	0.0
μ_{C1}	0.134	0.04	0.0
μ_{L2}	-0.689	0.10	0.0
γ_{FSNCOM}	0.746	0.25	0.0
γ_{SFYGM3}	3.980	0.82	0.0
γ_{SFYGT1}	3.61	0.75	0.0
γ_{INC}	0.829	0.06	0.0
γ_{IIP}	0.998	0.06	0.0
γ_{SLS}	0.738	0.06	0.0
σ^2_{HSBR}	0.945	0.06	0.0
σ^2_{FSNCOM}	0.969	0.06	0.0
σ^2_{SFYGM3}	0.164	0.03	0.0
σ^2_{SFYGT1}	0.311	0.03	0.0
σ^2_{EMP}	0.358	0.04	0.0
σ^2_{INC}	0.558	0.04	0.0
σ^2_{IIP}	0.360	0.04	0.0
σ^2_{SLS}	0.649	0.05	0.0
σ^2_{L}	0.016	0.01	0.01
σ^2_{C}	0.548	0.06	0.0

Table 3. Estimated parameters of model 3
Log-likelihood: -4641.1

Parameter	Estimated	St. error	p-value
p_{11}	0.983	0.01	0.0
p_{22}	0.873	0.05	0.0
A	4.950	2.18	0.01
B	8.840	3.5	0.01
μ_{L1}	0.090	0.02	0.0
μ_{L2}	-0.467	0.09	0.0
μ_{C1}	0.278	0.04	0.0
μ_{L2}	-1.040	0.10	0.0
γ_{FSNCOM}	0.760	0.25	0.0
γ_{SFYGM3}	3.590	0.71	0.0
γ_{SFYGT1}	3.530	0.69	0.0
γ_{INC}	0.791	0.05	0.0
γ_{IIP}	0.889	0.06	0.0
γ_{SLS}	0.656	0.06	0.0
σ^2_{HSBR}	0.938	0.06	0.0
σ^2_{FSNCOM}	0.963	0.06	0.0
σ^2_{SFYGM3}	0.229	0.03	0.0
σ^2_{SFYGT1}	0.254	0.03	0.0
σ^2_{EMP}	0.277	0.04	0.0
σ^2_{INC}	0.548	0.04	0.0
σ^2_{IIP}	0.429	0.04	0.0
σ^2_{SLS}	0.688	0.05	0.0
σ^2_{L}	0.018	0.01	0.01
σ^2_{C}	0.433	0.04	0.0

Figure 3.2.1: Composite leading and coincident indicators

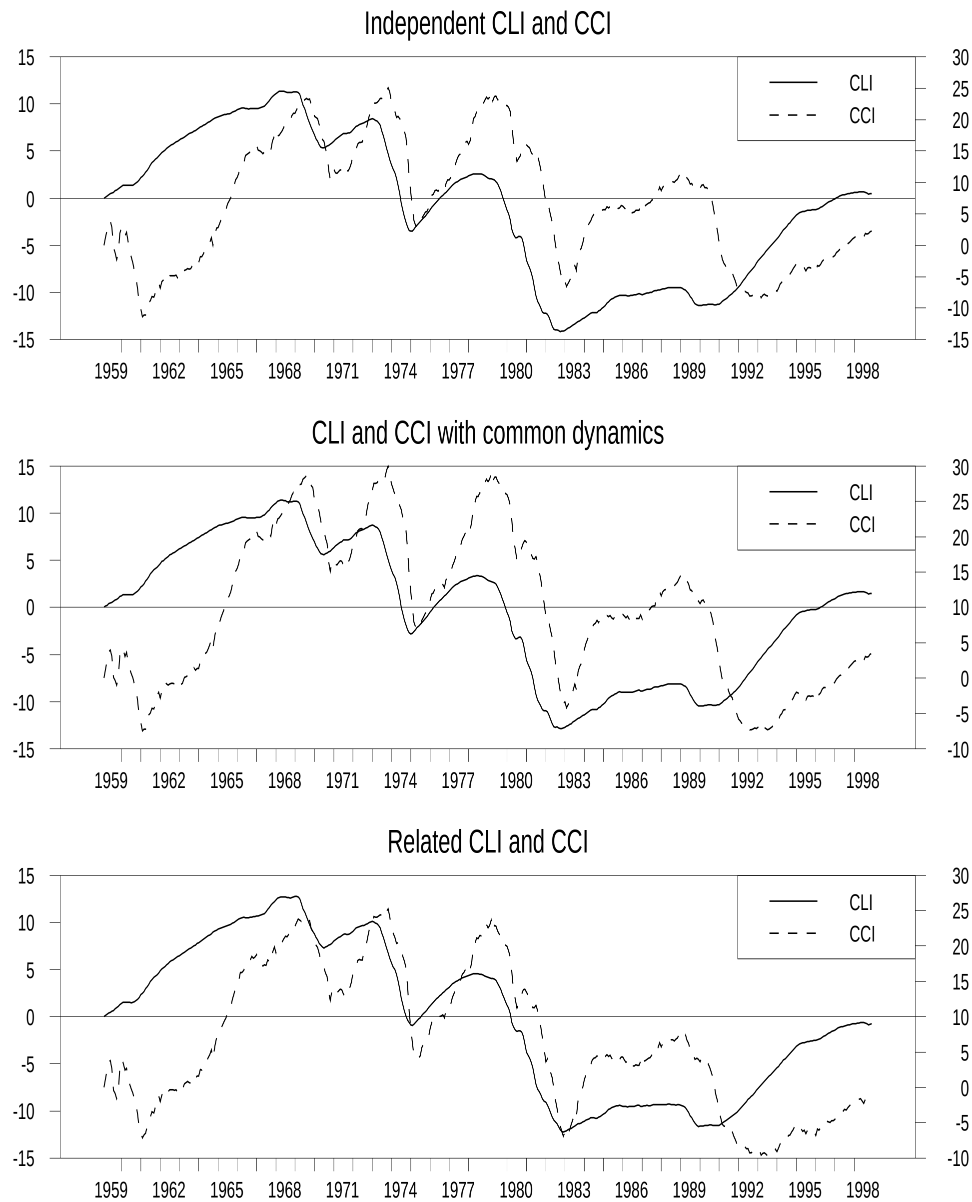


Figure 3.2.2: Model 1. Conditional recession probabilities vs. NBER chronology

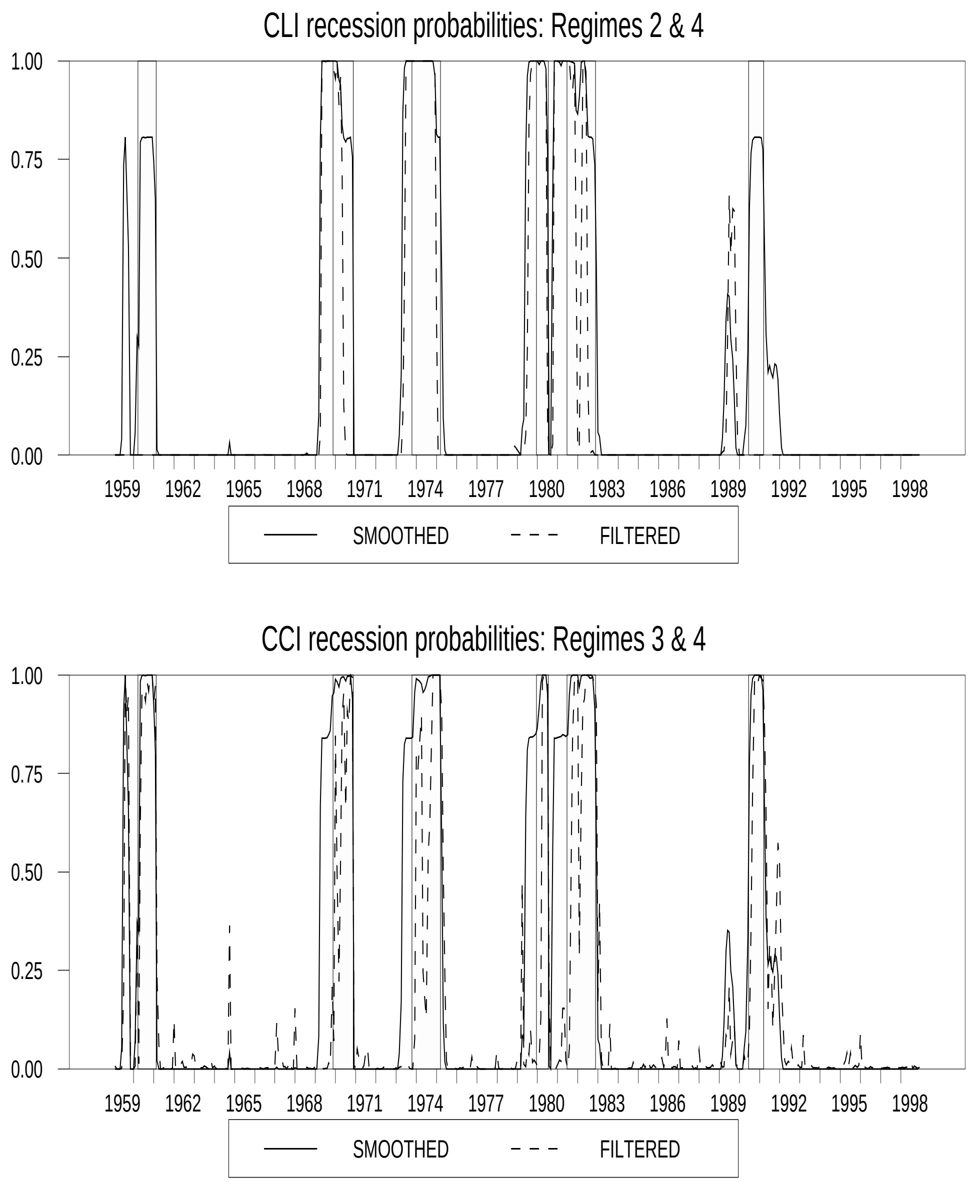


Figure 3.2.3: Model 2. Conditional recession probabilities vs. NBER chronology

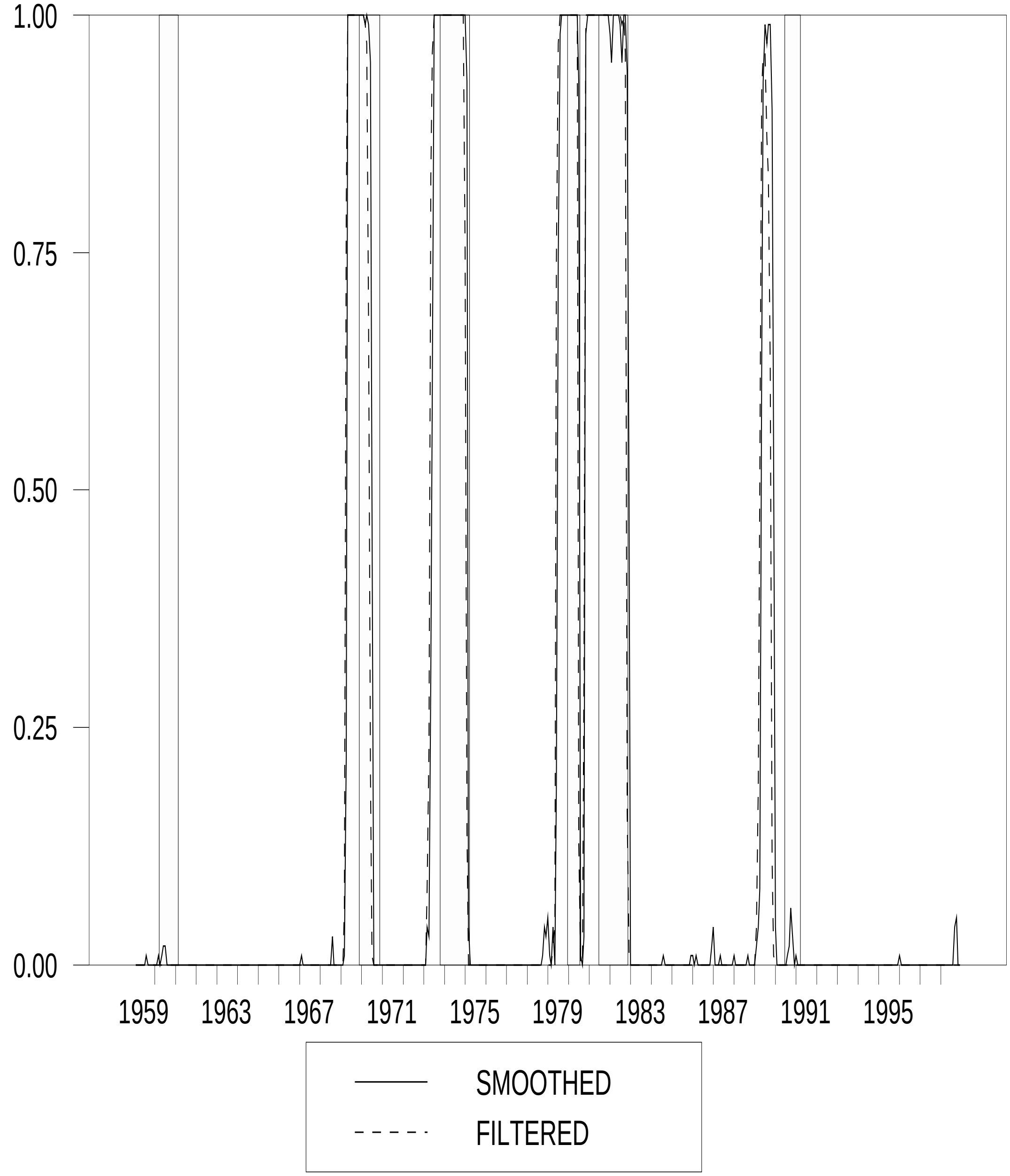
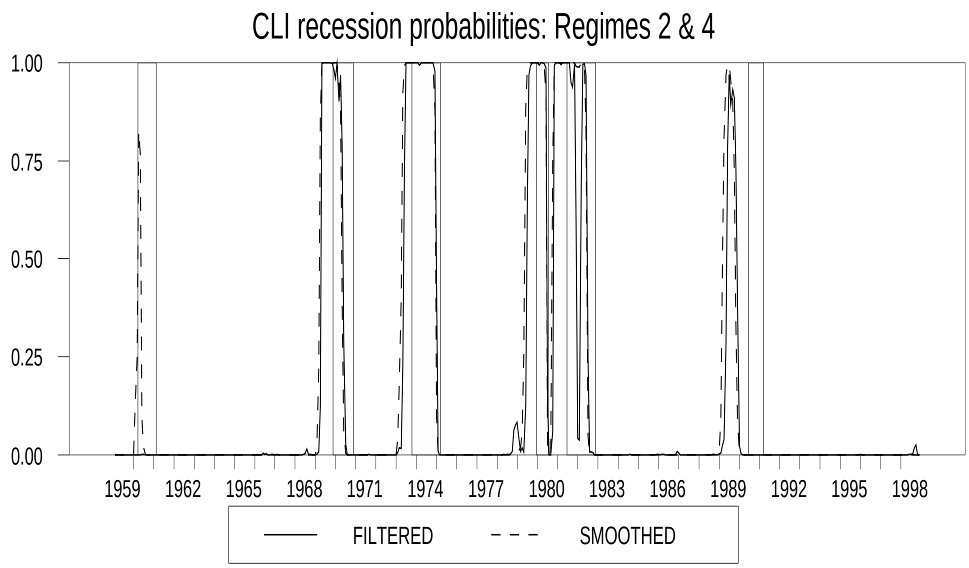
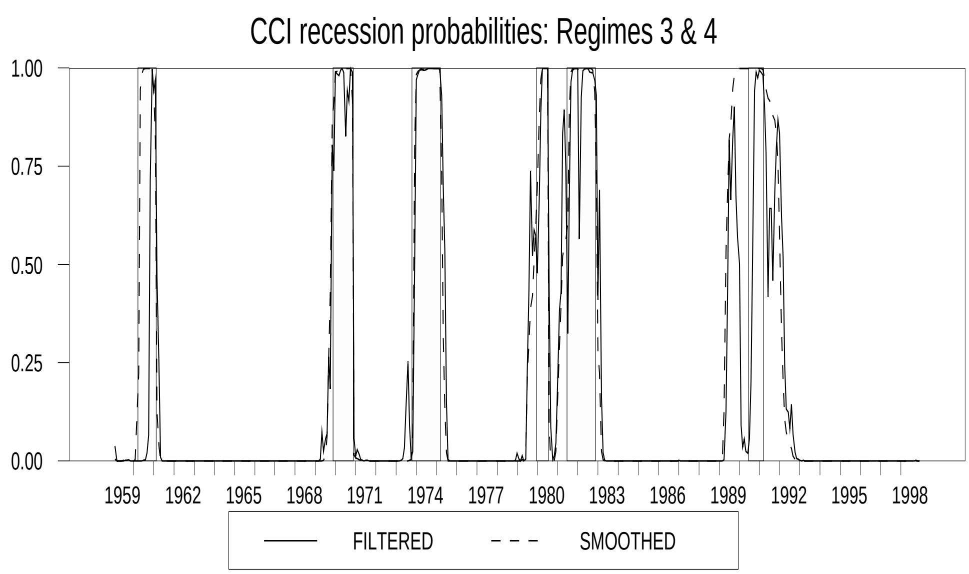


Figure 3.2.4: Model 3. Conditional recession probabilities vs. NBER chronology





3.3 Conclusion

This chapter has been devoted to an important extension of the standard single factor dynamic model — a bifactor dynamic model. The two common factors are distinguished across the temporal dimension: a composite leading indicator and a composite coincident indicator. This distinction may be extremely useful for the forecasting, since, once the relationship (if any) between the CLI and the CCI is figured out, we can use this information to predict the future values of the common coincident indicator as well as its cyclical phases more efficiently.

We have examined two general models: (1) a model where the link between the common leading factor and the common coincident one is determined by the Granger causality, and (2) a model where this lead-lag relationship between the two common factors has a probabilistic nature.

Unfortunately, the statistical comparison of these two models is impossible due to the presence of the nuisance parameters — the parameters which are not present under the null hypothesis, as, for instance, transition probabilities, state-dependent intercepts and variances — unless we are using some simulation method of model estimation, like Monte Carlo Markov Chain. For this reason, for example, we cannot contrast a model having two regimes against that with four regimes. Therefore we have to rely upon the qualitative judgements, like, for instance, the applicability of each model given particular situation or some indirect measures of performance of the models in question.

The model with Granger causality seems to deliver a significant improvement in the performance, as measured by the value of the log-likelihood function, compared to the model where the common leading and coincident indicators are supposed to be independent. In the case of the model with probabilistic mechanism the submodel with two common factors being independent slightly outperforms the model where the coincident indicator cycle lags behind the cycle of the leading indicator.

Furthermore, the model with Granger causality is somewhat less flexible, since it assumes that the lag between the leading and coincident indicators is fixed. While in the model with probabilistic mechanism this lag may vary depending on the cyclical phase. One can treat the availability of the Granger-causality model both with linear and regime-switching dynamics as its advantage: it might be useful, especially in the cases when the hypothesis of the regime switching is rejected. Moreover, it is more parsimonious in terms of the number of parameters to estimate than the probabilistic mechanism model.

One of indirect measures of the model performance is the so-called quadratic probability score (QPS) introduced by Diebold and Rudebusch (1989) and employed by Layton and Katsuura (2001). This method compares the recession probabilities estimated using some model to a generally accepted business cycle dating. In the case of the US economy one normally takes advantage of the NBER's dates as such "official dating".

Formally the QPS is defined as follows (see Layton and Katsuura (2001, p.408)):

$$QPS = \frac{1}{T}\sum_{t=1}^{T}(P_t - D_t)^2 \tag{3.14}$$

where T is the number of observations; P_t is the model-derived probability for the t-th observation; D_t is the binary variable taking value of 1 during the NBER recessions and of 0 during the NBER expansions.

The table below represents the QPS estimated for the various models analyzed in this chapter together with the Diebold-Mariano statistic[2] used to compare the accuracy of prediction of the NBER turning points by these models to the accuracy of prediction of the NBER dates by a benchmark model. The benchmark model is defined as single coincident factor model with specification CF-MS(0,0). The idea is to find out whether the QPS of a compared model is significantly different from the QPS of the benchmark model, since we are using QPS as a measure of predictive accuracy.

[2]See Diebold and Mariano (1995). The authors claim that their statistic allows for the forecast errors which are potentially non-Gaussian, non-zero mean, serially and/or contemporaneously correlated.

Lead	Filtered probabilities	Smoothed probabilities
	Single leading factor model	
8	0.123 (2.45)	0.127 (3.02)
	Single coincident factor model	
0	0.061 (0.50)	0.023 (0.79)
	Bifactor model with Granger-causality mechanism	
8	0.121 (2.48)	0.125 (3.01)
9	0.126 (2.57)	0.124 (3.12)
	Bifactor models with transition-probabilities mechanism	
	Model 1: Independent common factors	
	Leading factor's probabilities	
1	0.113 (2.64)	0.072 (2.23)
5	0.077 (1.67)	0.090 (5.79)
	Coincident factor's probabilities	
0	0.047 (1.11)	0.055 (2.32)
	Model 2: Perfectly correlated probabilities	
5	0.097 (2.80)	0.093 (3.67)
	Model 3: Common factors related with lead	
	Leading factor's probabilities	
5	0.083 (1.75)	0.083 (3.21)
6	0.085 (1.90)	0.079 (2.87)
	Coincident factor's probabilities	
0	0.062 (0.42)	0.054 (1.0)

Numbers in parentheses are the absolute values of the Diebold-Mariano statistic. The statistic was computed using a rectangular spectral window of length 101. Here the forecast accuracy of each model listed in the table is compared to that of the benchmark model. The test statistic is standardized and hence is asymptotically distributed as N(0,1).

Six models are presented in the above table: two single-factor models (leading factor model LF(0,0) with state-dependent intercept and variance and coincident factor model CF(1,1)) and four bifactor models (one with Granger-causality mechanism and three with transition probabilities matrix mechanism). The first column "Leads" stands for the time displacement which renders the best correspondence between the particular dating and the NBER chronology. Thus, the leading factor recession probabilities should be shifted backwards to get closer to the NBER cycle, given the fact that they lead the coincident indicator whose turning points are reflected by the NBER dates. The second and the third columns display the QPS characterizing conformity between the filtered and smoothed recession probabilities, on the one hand, and the NBER business cycle chronology, on the other hand. In the bifactor models with transition probabilities matrix mechanism in two cases (model 1 and model 2) we have two datings: one for the leading factor and another for the coincident factors. In the former case the recession probabilities are obtained as a sum of conditional probabilities of regime 2 and 4 ("low CLI's intercept — high CCI's intercept" and "low CLI's intercept — low CCI's intercept"), while in the latter case it is a sum of the probabilities of regimes 3 and 4 (see section 3.2).

Leading factor recession probabilities appear to fit the NBER dates worse than the recession probabilities of the coincident factors. This is quite natural since the common leading factor serves as a proxy of the common coincident factor within a few months. However, despite the worse *ex post* performance, the leading factor's recession probabilities permit forecasting the arrival of future recessions and expansions.

In the case of single leading factor model and bifactor model with Granger mechanism the lead is 8-9 months, while in the case of the bifactor model with transition matrix mechanism the lead is 5-6 months.

The QPS of the coincident recession probabilities stays almost unchanged regardless of the model, while the leading recession probabilities estimated using the bifactor model with probabilistic mechanism fit the NBER dates much better than those of the single leading factor model and of the bifactor model with Granger-causality mechanism.

Thus, to our opinion, the choice between each of the two models can be made depending on the particular data-generating process which we encounter in our research. Neither of the models is absolutely superior to its counterpart, each having its advantages which may be useful in one situation and useless in another.

Finally, one could possibly combine the two approaches allowing for both the Granger causality and transition probabilities matrix mechanism. In such a way we would be able to consider a case where there exist two independent business cycles — one for each common indicator — and where common coincident indicator's dynamics depend not only on its own state variable but also on that of the common leading factor.

Chapter 4

Dynamic Factor Analysis and Data Problems

In the previous chapter we have been discussing some important extensions of the common dynamic factor model towards developing bifactor models having both linear and Markov-switching dynamics. We have shown how one can introduce a multiplicity of factors in the time dimension, namely: leading and coincident common factors. This renders the forecasts more efficient, since they are based not only on the past values of the coincident economic indicator but also on the past and present values of the leading series. Moreover, two different ways of modeling the relationships between these common factors were examined: the Granger-causality mechanism and the mechanism of the transition probabilities matrix.

However, quite often the practical realization of the dynamic factor approach is impeded by the lack of the relevant data measured at high (say, monthly) frequencies. Another source of the problems are the various structural breaks which introduce discontinuities in the time series, thus, shortening already not very long contemporaneous macroeconomic time series. This is especially the case of most European countries and their regions, not to mention the developing economies whose statistical databases may be even worse. The causes of these breaks are very diversified ranging from the changes in the statistical methodology to the secular volatility shifts.

In this chapter we propose some devices to cope with these problems. In section 4.1 a model with mixed-frequency data is considered, while section 4.2 and section 4.3 deal with the structural breaks problem. Section 4.2 employs the deterministic dummies to model the structural shifts both in the means and in the variances of the observed variables, whereas section 4.3 introduces additional state variable to take into account the structural break in volatility.

4.1 Markov-Switching Dynamic Factor Model with Mixed-Frequency Data

A lot of valuable information is lost because many important time series are only available at the quarterly or annual frequencies. For instance, the CEI estimated with the monthly data does not take into account the information contained in the GDP series which is available only at quarterly or lower frequencies. This problem is especially severe at the regional level, since the regional statistical databases are much poorer than the national ones.

Fortunately, the problem of discrepancy in the frequency of observations seems to be solved. The solution was proposed recently by Mariano and Murasawa (2003). They consider a model where different frequencies, say monthly and quarterly, for different variables entering the model are allowed. This is especially useful if we want our coincident indicator to be a proxy for some aggregate observable variable, e.g. GDP. As a rule the GDP data are released at much lower frequency than individual series characterizing specific sectors of the economy. The Murasawa and Mariano's model enables us to take advantage of the valuable information contained in the lower-frequency time series.

Our idea is to apply this approach to the Markov-switching common dynamic factor model so that to be able to estimate CEI which incorporates both the comovement of the macroeconomic variables and the asymmetry of the different business cycle phases without losing the important information which is otherwise wasted because of the discrepancies in the observation spacing.

The rest of the section is structured as follows. In the next subsection we discuss the technical details of construction and estimation of the Markov-switching common factor models. In subsection 4.1.2 we consider application of this methodology to the real data. Subsection 4.1.3 concludes the section.

4.1.1 Model

The model of the common factor with nonlinear (Markov-switching) dynamics as the one estimated by Kim and Nelson can be expressed as follows:

$$\Delta y_t = \delta + \Gamma \Delta c_t + u_t \tag{4.1}$$

where Δy_t is the $n \times 1$ vector of the first differences of the observed time series in logs; Δc_t is first difference of the unobserved common factor having a regime-switching dynamics; u_t is the $n \times 1$ vector of the specific, or idiosyncratic, components characterizing the individual dynamics of each of the observed series, and Γ is the matrix of factor loadings.

Assume now that we deal with the mixed-frequency data. In other words, some of the series are observed at a lower frequency (say, yearly or quarterly) and some other series are observed at a higher frequency (say, monthly or weekly). This amounts to saying that we dispose of $n = n_1 + n_2$ observable component series[1]. The first n_1 component series, y_1, are observed at lower frequency (each $f > 1$ periods), while the remaining n_2 series, y_2, are measured at a higher frequency which we may normalize to 1. Thus, if we have quarterly and monthly data, $f = 3$ and we observe $y_1 = \{y_{13}, y_{16}..., y_{1.T-3}, y_{1.T}\}$ and $y_2 = \{y_{21}, y_{22}..., y_{2.T-1}, y_{2.T}\}$. Denote by y^*_{1t} the values of the first n_1 component series that we might have observed if these series were measured at the same frequency as y_2, that is, $y^*_1 = \{y^*_{11}, y^*_{12}, ..., y^*_{1.T-1}, y^*_{1.T}\}$. The observed lower-frequency series can be expressed in terms of these unobserved values as follows:

$$y_{1t} = \frac{1}{f} \sum_{i=0}^{f-1} L^i y^*_{1t} \tag{4.2}$$

Hence after taking the first differences of the observable lower frequency series, the growth rates of these series would be as:

$$(1 - L^f) y_{1t} = \frac{1}{f} (\sum_{i=0}^{f-1} L^i)^2 (1 - L) y^*_{1t} \tag{4.3}$$

where $(\sum_{i=0}^{f-1} L^i)^2 = \sum_{i=0}^{2f-1} (f + 1 - |i - f|) L^i$ or simpler

[1] Here we consider only the case of the flow variables.

$$(\sum_{i=0}^{f-1} L^i)^2 = 1 + L + 2L^2 + 3L^3 + ... + 3L^{2f-4} + 2L^{2f-3} + L^{2f-2} + L^{2f-1} \quad (4.4)$$

To estimate the model at the higher frequency, the unobserved values of the lower-frequency time series are treated as missing. As Mariano and Murasawa (2003) have shown, they can be replaced by any random variable as long as it is independent of the parameters of the model. In particular, these missing observations may be substituted by zeros. Thus, the growth rates of the first n_1 variables expressed at the higher frequency can be constructed as:

$$(1 - L)\tilde{y}_{1t}^{*} = \begin{cases} (1 - L^f)y_{1t}, & \text{if } t = f, 2f, ..., T \\ 0 & \text{otherwise} \end{cases}$$

In principle, we can do this kind of substitution not only for the observations between the observed values of the lower-frequency time series, but also in case of the series which are shorter than the others. In the general case we may define the constructed higher frequency series as:

$$\tilde{y}_t^{*} = \begin{cases} y_t, & \text{if } t \in \Xi \\ 0, & \text{otherwise} \end{cases}$$

where Ξ is the set of dates for which the shortest time series is observable. For instance, when the t_1 initial observations are missing, the set Ξ will be defined as:

$$\Xi = \{t|\ tt_1\}$$

In the case when the same variable is also the one which is measured at the lower frequency, the definition of Ξ will be as:

$$\Xi = \{t|\ tt_1 \text{ and } t = f, 2f, ..., T\}$$

Therefore the vector of the growth rates of all the n observed series, Δy_t, measured at the higher frequency may be decomposed as:

$$\begin{pmatrix} (1 - L^f)\tilde{y}_{1t}^{*} \\ (1 - L)y_{2t} \end{pmatrix} = \delta + \Gamma \begin{pmatrix} \frac{1}{f}(\sum_{i=0}^{f-1} L^i)^2 I_t \\ 1 \end{pmatrix} (1-L)c_t + \begin{pmatrix} \frac{1}{f}(\sum_{i=0}^{f-1} L^i)^2 I_t \\ 1 \end{pmatrix} u_t \quad (4.5)$$

where I_t is the indicator function: $I_t = 1$ when $t = f, 2f, ..., T$ and $I_t = 0$ otherwise.

The common dynamic factor is modeled as:

$$\phi(L)\Delta c_t = \mu(s_t) + \varepsilon_t \tag{4.6}$$

where $\phi(L)$ is the AR(p) lag polynomial; $\mu(s_t)$ is the common factor intercept depending on the state variable s_t following a first-order Markov-chain process, and $\varepsilon_t \sim NID(0, \sigma^2(s_t))-$ thus the variance of the common factor shock may also be state-dependent. In a more general specification the coefficients of the autoregressive polynomial $\phi(L)$ may depend on the state too.

The vector of the idiosyncratic components can be represented as follows:

$$\psi(L)u_t \ = \eta_t \tag{4.7}$$

where $\eta_t \sim NID(0, \Sigma)$ and both the lag polynomial $\psi(L)$ and variance-covariance matrix Σ have a diagonal structure. Each idiosyncratic component is modelled as $AR(q_i)$ where $i = 1, ..., n$. In principle, the autoregressive order may be different across the specific components and may be equal to zero.

In order to be estimated using Kalman filter, this model can be expressed in the state-space form.

The measurement equation:

$$\Delta y_t = Ax_t + w_t \tag{4.8}$$

Transition equation:

$$x_t = \alpha(s_t) + Cx_{t-1} + v_t \tag{4.9}$$

where $\Delta y_t = \begin{pmatrix} (1-L^f)y_{1t} & (1-L)y_{2t} \end{pmatrix}'$
is the $n \times 1$ vector of observed variables in differences;

$x_t = \begin{pmatrix} \Delta c_t^* & u_t \end{pmatrix}'$ is the $m \times 1$ state vector containing the common dynamic factor vector $\Delta c_t^* = \begin{pmatrix} \Delta c_{t-1} & \Delta c_{t-2} & ... & \Delta c_{t-r} \end{pmatrix}'$, with $r = \max\{p, 2f-1\}$,
and the specific components vector

$u_t = \begin{pmatrix} u_{1t} & ... & u_{1t-l} & ... & u_{nt} & ... & u_{nt-q_n} \end{pmatrix}'$, with $l = \max\{q_1, 2f-1\}$;

$\alpha(s_t) = \begin{pmatrix} \mu(s_t) & 0 & ... & 0 \end{pmatrix}'$ is the vector of intercepts, and finally $v_t = \begin{pmatrix} \varepsilon_t & 0 & ... & \eta_{1t} & ... & \eta_{nt} & ... & 0 \end{pmatrix}'$ is the vector of disturbances. The dimension of the state vector, m, is determined as:

$$m = r + n_1 * l + \sum_{i=n_1+1}^{n} q_i$$

The system matrices have the following structure:
The measurement $n \times m$ matrix:

$$A = \begin{pmatrix} \gamma_1 \Lambda I_t & & & & 0 \\ \gamma_2 & o_{r-1} & \Lambda I_t & & \\ & & & \ddots & \\ \gamma_n & & & & i_{q_n} \end{pmatrix}$$

where Λ is the $1 \times (2f-1)$ vector of coefficients of the $(\sum_{i=0}^{f-1} L^i)^2$; o_k is the $k \times 1$ vector of zeros, and i_k is the first row of the $k \times k$ identity matrix.

The $m \times m$ transition matrix:

$$C = \begin{pmatrix} \Phi & o_r & & & & & 0 \\ I_{r-1} & & & & & & \\ & & \Psi_1 & o_l & & & \\ & & I_{l-1} & & & & \\ & & & & \ddots & & \\ & & & & & \Psi_n & o_{q_n} \\ 0 & & & & & I_{q_n-1} & \end{pmatrix}$$

where Φ and Ψ_i $(i = 1, ..., n)$ are the row vectors of the autoregressive coefficients; I_k is the $k \times k$ identity matrix.

The $n \times n$ variance-covariance matrix of the disturbances to the measurement equation:

$$R = \begin{pmatrix} I(f) & O \\ O & O \end{pmatrix}$$

where $I(f)$ is the diagonal $n_1 \times n_1$ matrix with the indicator functions, I_t, on the main diagonal.

The $m \times m$ variance-covariance matrix of the disturbances to the transition equation:

$$Q = \begin{pmatrix} \sigma^2(s_t) & & & & 0 \\ & \ddots & & & \\ & & \sigma_1^2 & & \\ & & & \ddots & \\ 0 & & & & \sigma_n^2 \end{pmatrix}$$

We introduce three identifying assumptions in this specification of model. First, the variance-covariance matrix Q is diagonal. Secondly, we may set either $\gamma_1 = 1$ or $\sigma^2(s_t = 1) = 1$. We chose the first option.

As always in this thesis we estimate the model using the method of maximum likelihood.

4.1.2 Real example

Having tested the performance of our model on the artificial data, we applied it to the actual Post World War II macroeconomic time series, the only difference being the time span: they use the series covering the period of 1959-1998, while we utilize the series stretching from 1959 through 2002. For more details on the time series used in this study see Table 4.1.2. The series were logged, then their first differences were taken and normalized.

To select the lag order, we applied Akaike information criterion (AIC) and Schwartz Bayesian information criterion (SBIC) computed as follows:

Akaike information criterion:

$$AIC = 2\log[L(\theta)] - 2[n_1 p + n_2 q] \tag{4.10}$$

where $L(\theta)$ is the likelihood function value at maximum; n_1 number of the low-frequency series (in this case we have only one such time series — quarterly GDP); n_2 is the number of the high-frequency series; p and q are the orders of the AR polynomials of the low- and high-frequency series, respectively.

Schwartz Bayesian information criterion:

$$SBIC = 2\log[L(\theta)] - [n_1 p + n_2 q]\log(T) \tag{4.11}$$

where T is the number of observations.

The values of the log-likelihoods for the various autoregressive order combinations (p, q) as well as the two information criteria are presented in Table 4.1.3 of Appendix. The AIC chooses (3,3) while SBIC selects (1,2) combination as the optimal one. We are going to use the latter combination as a more parsimonious. This is the same combination which was suggested by the SBIC in the linear case (see Mariano and Murasawa (2003)).

We represent the estimates of the parameters of the linear and Markov-switching common factor model with single- and mixed-frequency data in Tables 4.1.4 and 4.1.5, respectively. The estimated parameters for the linear and nonlinear models are very similar, with the exception of the autoregressive parameter of the common dynamic factor which is slightly smaller when the Markov switching is introduced.

Based on the parameter estimates of the nonlinear common factor model with different observation frequencies, we calculated the estimate of the common factor by taking the partial sums:

$$c_t = c_{t-1} + \Delta c_t + \delta \tag{4.12}$$

Figure 4.1.1 shows the profiles of the composite coincident indicators constructed using the single- and mixed-frequency component series.

The behavior of the linear and Markov-switching composite indicators is very similar, regardless of whether single- or mixed-frequency data are used. On the other hand, the volatility of the mixed-frequency indicator is substantially smaller than that of the single-frequency indicator. One can observe this also by comparing the common factor's residual variance in Tables 4.1.4 and 4.1.5: in the single-frequency case it varies between 0.274 and 0.324, whereas in the mixed-frequency case it is between 0.074 and 0.09. Moreover, nonlinear composite indicators appear to be less volatile than linear ones.

Figure 4.1.2 displays the conditional recession probabilities obtained from the estimation of two Markov-switching model — with single- and mixed-frequency data — plotted against the National Bureau of Economic Research (NBER) recession dates, where the latter are represented by the shading.

The correspondence between the filtered and smoothed recession probabilities, on the one hand, and the NBER recession dates, on the other hand, is striking. The only exception is the recession detected by our model in the very beginning of the sample and absent in the NBER chronology. The probabilities obtained for the single- and mixed-frequency models are almost indistinguishable.

The formal analysis of the in-sample forecasting performance of single-frequency and mixed-frequency CEI models was conducted using the quadratic probability score (QPS).

To test whether the differences in the QPS of different models are statistically significant we use the Diebold-Mariano statistic (with the rectangular spectral window of length 15) proposed by Diebold and Mariano (1995). The DM-statistic is computed by comparing the loss differentials (with respect to the binary coded NBER dating) of the filtered and smoothed regime probabilities of the single-frequency CF-MS, on the one hand, to the loss differentials of the regime probabilities of mixed-frequency CF-MS, on the other hand.

The results of the comparison of performance of both nonlinear models are presented in Table 4.1.6. The second column displays the QPS statistic. The columns 3 through 5 report the Diebold-Mariano (DM) statistics for different pairwise comparisons of QPS.

The first conclusion that can be drawn from Table 4.1.6 is that the smoothed probabilities are much better predictors of the NBER dating than the filtered probabilities. This difference is statistically significant at 1% level. By contrast, we cannot say that the mixed-frequency CF-MS forecasts are significantly better than the forecasts supplied by the single-frequency model.

4.1.3 Summary

In this section we have introduced a Markov-switching common dynamic factor model with mixed-frequency observations. Until now only the data of the same frequency and with the same length were used to estimate the latent common factor models with Markov-switching dynamics. Building on the extension of the linear common factor model to the case of the data with different observation frequencies proposed by Mariano and Murasawa (2003), we offer a solution to the problem of missing observations in the nonlinear case.

This would allow preventing the losses of valuable information concerning the evolution of the common dynamic factor which may be contained in the lower-frequency time series and, in general, in the time series with any type of missing values.

The comparison of the ability to predict the NBER business cycle chronology, though, conducted in this paper, does not provide any evidence of the statistically important superiority of the mixed-frequency model with regime switching over the single-frequency CF-MS.

The results presented here will be used later on in the thesis to investigate the structural break problem — see section 4.3.

4.1.4 Appendix

Table 4.1.1. Simulated example: true and estimated parameters

Parameter	True	Estimated	St. error	p-value
p_{11}	0.95	0.93	0.02	0.0
p_{22}	0.84	0.87	0.03	0.0
μ_1	0.4	0.43	0.03	0.0
μ_2	-0.6	-0.68	0.05	0.0
γ_2	0.5	0.44	0.05	0.0
γ_3	0.8	0.81	0.02	0.0
γ_4	2.0	2.01	0.06	0.0
γ_5	1.7	1.73	0.05	0.0
ϕ	0.6	0.56	0.03	0.0
ψ_1	-0.5	-0.61	0.16	0.0
ψ_2	0.6	0.59	0.04	0.0
ψ_3	-0.1	-0.06	0.05	0.12
ψ_4	-0.2	-0.17	0.06	0.0
ψ_5	-0.8	-0.84	0.02	0.0
σ_1^2	0.25	0.99	0.16	0.0
σ_2^2	0.36	0.38	0.02	0.0
σ_3^2	0.16	0.16	0.01	0.0
σ_4^2	0.49	0.53	0.05	0.0
σ_5^2	0.81	0.81	0.06	0.0
σ_c^2	0.16	0.15	0.02	0.0

Table 4.1.2. The component series of the US composite coincident indicator

Series	Short-hand	Description
Monthly series 1959:1 – 2002:6		
Employees on nonagricultural payrolls	EMP	10^3, SA
Personal income less transfer payments	INC	10^9 1996 \$, SA, annual rate
Index of industrial production	IIP	total index, 1996=100, SA
Manufacturing and trade sales	SLS	chained 10^6 1996 \$, SA
Quarterly series 1959:1-2002:2		
Real GDP	GDP	chained 10^9 1996 \$, SA

Table 4.1.3. Lag selection analysis

(p,q)	LogLik	AIC	SBIC
(0,0)	-1643.52	-3287.04	-3287.04
(0,1)	-1605.68	-3221.36	-3242.82
(0,2)	-1565.85	-3151.7	-3194.62
(0,3)	-1555.95	-3141.9	-3204.48
(1,0)	-1626.37	-3254.74	-3259.03
(1,1)	-1589.89	-3191.78	-3217.53
(1,2)	-1550.31	-3122.62	-3169.83
(1,3)	-1539.81	-3111.62	-3178.37
(2,0)	-1625.38	-3254.76	-3263.34
(2,1)	-1589.3	-3192.60	-3222.64
(2,2)	-1549.68	-3123.36	-3174.86
(2,3)	-1539.81	-3113.62	-3184.54
(3,0)	-1625.29	-3256.58	-3269.10
(3,1)	-1589.26	-3194.52	-3227.89
(3,2)	-1545.73	-3117.46	-3171.69
(3,3)	-1534.41	-3104.82	-3179.91

Table 4.1.4. Estimated parameters of linear and Markov-switching models with single-frequency data 1959:1-2002:6

Parameter	Linear LL=-2474.77		Nonlinear LL=-2464.20	
	Coefficient	St. error	Coefficient	St. error
p_{11}	–	–	0.975	0.009
$1-p_{22}$	–	–	0.125	0.052
μ_1	–	–	0.136	0.037
μ_2	–	–	-0.670	0.112
γ_{INC}	0.931	0.070	0.926	0.069
γ_{IIP}	1.19	0.078	1.16	0.077
γ_{SLS}	0.755	0.057	0.741	0.057
ϕ	0.579	0.045	0.358	0.063
$\psi_{EMP.1}$	0.114	0.044	0.108	0.044
$\psi_{EMP.2}$	0.466	0.048	0.462	0.050
$\psi_{INC.1}$	-0.008	0.047	-0.020	0.069
$\psi_{INC.2}$	0.051	0.050	0.041	0.047
$\psi_{IIP.1}$	-0.105	0.078	-0.069	0.070
$\psi_{IIP.2}$	-0.093	0.072	-0.073	0.066
$\psi_{SLS.1}$	-0.415	0.049	-0.411	0.048
$\psi_{SLS.2}$	-0.192	0.047	-0.189	0.047
σ^2_C	0.324	0.038	0.274	0.033
σ^2_{EMP}	0.310	0.029	0.300	0.030
σ^2_{INC}	0.570	0.041	0.561	0.041
σ^2_{IIP}	0.306	0.035	0.321	0.034
σ^2_{SLS}	0.583	0.041	0.585	0.041

LL = the value of loglikelihood function

Table 4.1.5. Estimated parameters of linear and Markov-switching models with mixed-frequency data 1959:1-2002:6

Parameter	Linear LL=-2945.31		Nonlinear LL=-2934.72	
	Coefficient	St. error	Coefficient	St. error
p_{11}	–	–	0.975	0.009
$1-p_{22}$	–	–	0.124	0.052
μ_1	–	–	0.069	0.018
μ_2	–	–	-0.341	0.061
γ_{EMP}	1.89	0.145	1.92	0.151
γ_{INC}	1.81	0.137	1.82	0.142
γ_{IIP}	2.24	0.130	2.24	0.133
γ_{SLS}	1.46	0.095	1.46	0.096
ϕ	0.573	0.045	0.354	0.061
$\psi_{GDP.1}$	-0.050	0.090	-0.052	0.074
$\psi_{GDP.2}$	-0.834	0.064	-0.835	0.063
$\psi_{EMP.1}$	0.119	0.043	0.116	0.043
$\psi_{EMP.2}$	0.467	0.048	0.467	0.048
$\psi_{INC.1}$	-0.028	0.056	-0.034	0.099
$\psi_{INC.2}$	0.036	0.048	0.031	0.060
$\psi_{IIP.1}$	-0.056	0.070	-0.042	0.092
$\psi_{IIP.2}$	-0.061	0.060	-0.058	0.071
$\psi_{SLS.1}$	-0.432	0.048	-0.427	0.048
$\psi_{SLS.2}$	-0.204	0.047	-0.201	0.047
σ^2_C	0.090	0.013	0.074	0.011
σ^2_{GDP}	0.240	0.050	0.243	0.045
σ^2_{EMP}	0.313	0.029	0.306	0.030
σ^2_{INC}	0.556	0.040	0.552	0.040
σ^2_{IIP}	0.324	0.031	0.327	0.031
σ^2_{SLS}	0.569	0.040	0.570	0.040

LL = the value of loglikelihood function

Table 4.1.6. Comparison of the turning points detection performance

	QPS	DM-statistic		
		Single-FreqSm	Mixed-FreqF	Mixed-FreqSm
Single-FreqF	0.0544	1.68**	0.75	1.51*
Single-FreqSm	0.0364		1.71**	1.36*
Mixed-FreqF	0.0533			1.54*
Mixed-FreqSm	0.0375			

Column 2 contains the QPS of different models; columns 3 through 5 represent the absolute values of Diebold-Mariano (DM) statistics measuring the statistical significance of the forecasting accuracy: * stands for 10% significant difference, ** stands for 5% significant difference.
Single-FreqF = single-frequency MS-CF(1,2) filtered recession probabilities
Single-FreqSm = single-frequency MS-CF(1,2) smoothed recession probabilities
Mixed-FreqF = mixed-frequency MS-CF(1,2) filtered recession probabilities
Mixed-FreqSm = mixed-frequency MS-CF(1,2) smoothed recession probabilities

Figure 4.1.1: US composite indicators based on single- & mixed-frequency data

US monthly and quarterly data 1959:1-2002:6

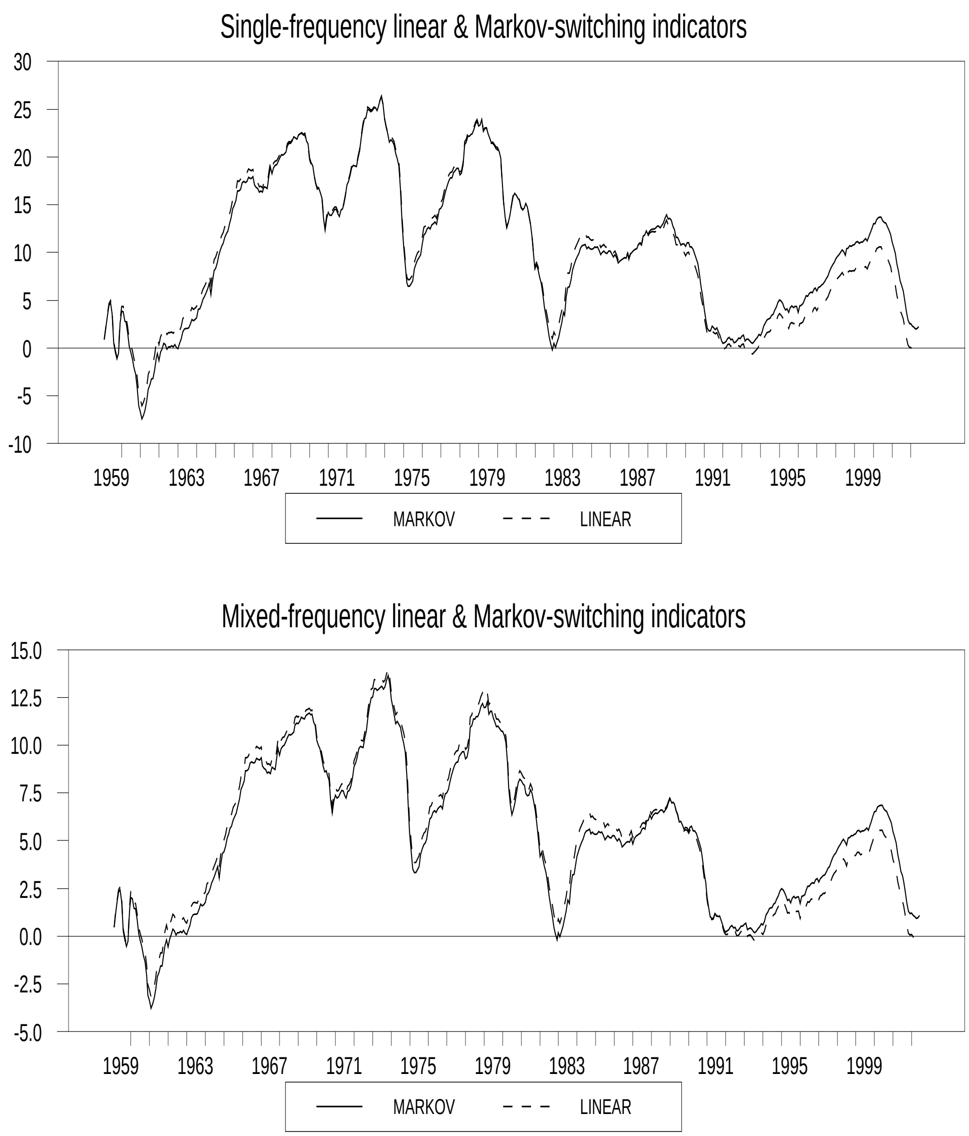
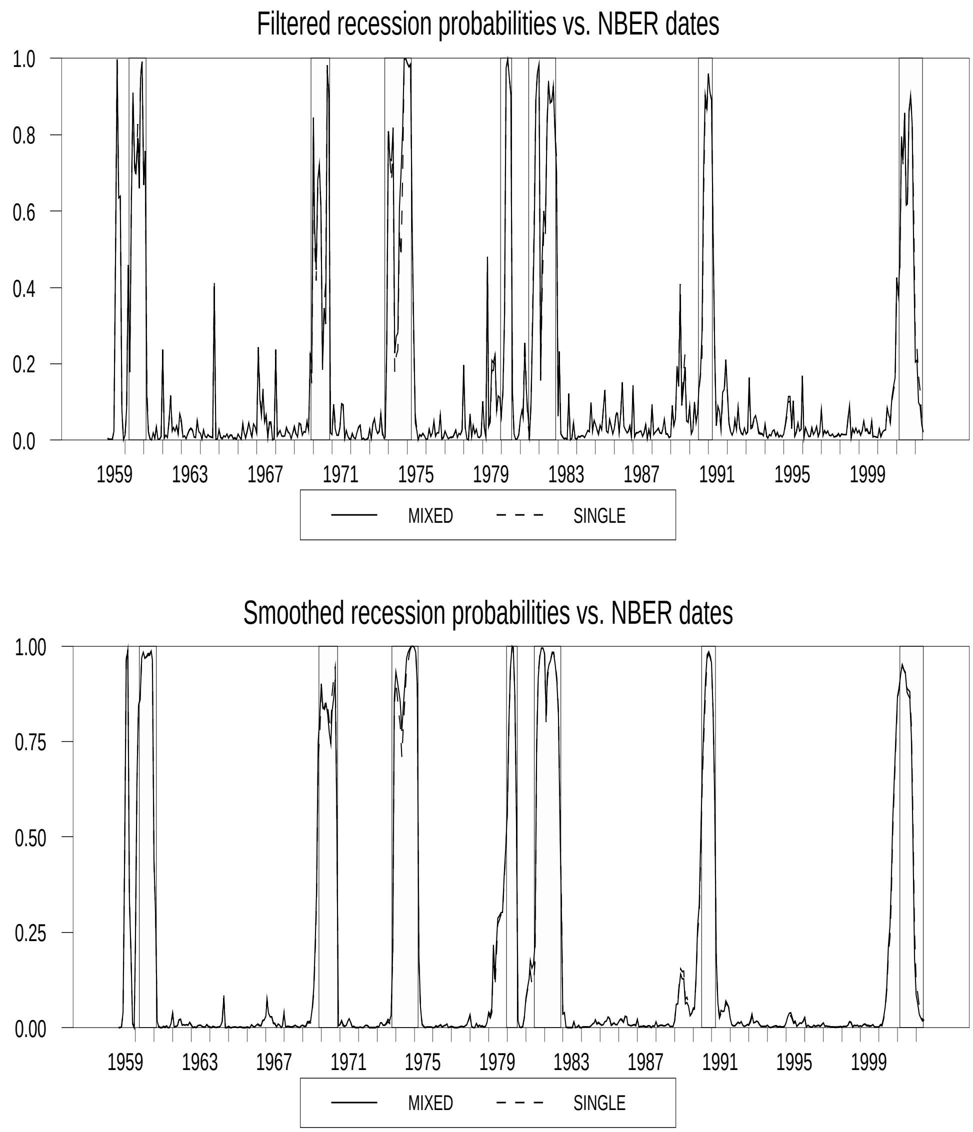


Figure 4.1.2: Conditional recession probabilities

US monthly and quarterly data 1959:1-2002:6



4.2 Structural Breaks in Dynamic Factor Model: Deterministic Mechanism

In this section we will consider one of the ways of dealing with the structural breaks in the observed time series which may be caused both by the "natural" shifts in the behavior of the economic variables and by the changes in the statistical methodologies. The models examined here use the deterministic dummies to capture the structural breaks in different model parameters with unknown breakpoints different for each time series. The specification proposed here builds on the paper by Krane and Wascher (1999), who introduce the seasonal dummies in the means of the observed time series and of the common factor as well as in the factor loadings in order to take account of the deterministic seasonality in the common dynamic factor model, and on that by Chauvet and Potter (2001), who make the intercept and the autoregressive coefficients of the common factor change as a function of the structural break.

In subsection 4.2.1 we present the linear common dynamic factor model with the structural break(s) in the observed time series. In subsection 4.2.2 an artificial example is considered. Subsection 4.2.3 comes up with a real example — common dynamic model of the US composite economic indicator with deterministic dummies capturing the structural breaks in the means and variances of the observed series and idiosyncratic components. Subsection 4.2.4 concludes the section.

4.2.1 Model

We consider a set of the observed time series, whose common dynamics are underlined by one or several common factors which may interact in a complex temporal and/or spatial way.

The idiosyncratic dynamics of each time series in particular are captured by one specific factor per each observed time series. Therefore the model in the general form can be written as follows:

$$\Delta y_t = [(\mathbf{I}_n - I_t)\delta_1 + I_t\delta_2] + [(\mathbf{I}_n - I_t)\Gamma_1 + I_t\Gamma_2]\,\Delta f_t + u_t \tag{4.13}$$

where Δy_t is the $n \times 1$ vector of the logged observed time series in the first differences (growth rates); Δf_t is the $k \times 1$ vector of the latent common factors in the first differences; u_t is the $n \times 1$ vector of the latent specific factors; δ_1 and δ_2 are the $n \times 1$ vectors of the means of the observed time series; Γ_1 and Γ_2 are the $n \times k$ factor loadings matrices linking the observed series with the common factors, $\mathbf{I}_n$ is the $n \times n$ identity matrix, and I_t is the structural break indicator function. Since, in principle, there is no reason to suppose that all the observed time series were subject to the structural break and that, if any, the structural breaks took place in the same moment, the breakpoint indicator function, I_t, can be written as a diagonal matrix whose diagonal elements are the individual indicator functions:

$$I_t = I_n \otimes \begin{pmatrix} I_{1t} \\ I_{2t} \\ \cdots \\ I_{nt} \end{pmatrix}$$

where

$$I_{it} = \begin{cases} 0, & \text{if } t < \tau_i \\ 1, & \text{otherwise} \end{cases}$$

where τ_i is the period when the structural break in the $i-th$ observed time series has taken place.

The dynamics of the latent common factors can be described in terms of a VAR model:

$$\Delta f_t = \left[(\mathbf{I}_k - I_t^f)\nu_1 + I_t^f \nu_2\right] + \left[(\mathbf{I}_k - I_t^f)\Phi_1(L) + I_t^f \Phi_2(L)\right] \Delta f_{t-1} + \varepsilon_t \quad (4.14)$$

where ν_1 and ν_2 are the $k \times 1$ vectors of the constant intercepts; I_t^f is the $k \times k$ diagonal matrix having the structure similar to that of I_t; $\Phi_1(L)$ and $\Phi_2(L)$ are the sequences of p ($p = \max\{p_{f_1}, ..., p_{f_k}\}$, where p_{f_j} is the order of the AR polynomial of the $j-th$ common factor) $k \times k$ lag polynomial matrices; ε_t is the $k \times 1$ vector of the serially and mutually uncorrelated common factor disturbances:

$$\varepsilon_t \sim NID\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{f_1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{f_k}^2 \end{pmatrix}\right)$$

The specific factors are assumed to be mutually independent but serially correlated:

$$u_t = \Psi(L)u_{t-1} + \eta_t \tag{4.15}$$

where $\Psi(L)$ is the sequence of q ($q = \max\{q_{1,...}, q_n\}$, where q_i is the order of the autoregressive polynomial of the $i-th$ idiosyncratic factor) $n \times n$ diagonal lag polynomial matrices and η_t is the $n \times 1$ vector of the mutually and serially uncorrelated Gaussian shocks:

$$\eta_t \sim \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{1t} & & 0 \\ & \ddots & \\ 0 & & \sigma^2_{nt} \end{pmatrix} \right)$$

where $\sigma^2_{it} = \lambda_i(1 - I_{it}) + \sigma^2_i$, $i = 1, ..., n$.

Thus, in general the intercepts of the observed variables, their factor loadings, and the residual variances of the specific components may be subject to the deterministic structural breaks.

Assume for the sake of simplicity that we have only one common factor (the extensions to the multi-factor case are straightforward). Furthermore, suppose that only observed variables are subject to the structural change which affects their means but not their factor loadings. These assumptions would seem realistic especially in the case of the changes in the accounting methodology which lead to the sudden shifts in the time series levels.

The model is estimated by the maximum likelihood method. To do this it is expressed in the state-space form:

$$\Delta y_t = [\delta_1(1 - I_t) + \delta_2 I_t] + Ax_t \tag{4.16}$$

$$x_t = \alpha + Cx_{t-1} + v_t \tag{4.17}$$

where $x_t = (f_t|u_t)'$ is the state vector containing stacked on top of each other vector of common factor and the vector of specific factors; v_t is the vector of the common and idiosyncratic factors' disturbances with mean zero and variance-covariance matrix Q; α is the vector of intercepts.

$$A = \begin{pmatrix} \Gamma_1 & i_{q_1} & O \\ \vdots & \vdots & \vdots \\ \Gamma_n & O & i_{q_n} \end{pmatrix}$$

where Γ_i is the $1 \times g$ vector of the factor loadings of the $i-th$ observed variable: $\Gamma_i = (\gamma_{i,1}, ..., \gamma_{i,g_i}, ..., 0)$ with $g = \max\{g_1, ..., g_n\}$.

$$C = \begin{pmatrix} \Phi & 0 & & 0 \\ 0 & \Psi^1 & & \\ & & \ddots & \\ 0 & 0 & & \Psi^n \end{pmatrix}$$

$$\Phi = \begin{pmatrix} \phi_C & 0 \\ I_{p_C-1} & o_{p_C-1} \end{pmatrix}$$

with $\phi_C = (\phi_{1,...,}\phi_p)$ being the $1 \times p$ vector of the AR coefficients of the common factor. The matrices $\Psi^1, ..., \Psi^n$ have the same structure as Φ^C.

One immediate extension of this model is the introduction of the regime-switching dynamics. This would allow taking care of the asymmetries which may characterize different phases of the business cycle and therefore is more complete from the standpoint of the Burns and Mitchell's definition of the business cycle as interpreted by Diebold and Rudebusch (1996). We are not discussing here the way it is to be done, the issue of the regime-switching dynamics being studied at length elsewhere in the thesis.

4.2.2 Simulated example

Following the pattern chosen above, we have generated an artificial data set, where some of the observed time series are subject to the one-time structural break, and have estimated the corresponding model using as inputs the observable time series. The dynamics in this model are linear, but one should keep in mind that it is straightforward to extend the model to a case of the Markov-switching dynamics.

For the linear case we have generated one common latent factor and four individual observable series. Both the common factors (in fact, their first differences, not levels) and the idiosyncratic components are modelled as the stationary AR(1) processes. The true parameters of the DGP are presented in the column two of Table 1 of the Appendix. The length of all these time series is 500 observations, which is comparable to the length of an ordinary Post World War II monthly time series for the US economy. The structural break takes place in the period 350, thus $\tau = 350$. Only the time series two is characterized by the structural break which makes its mean go up, the rest of the observed time series being exempt of the structural changes.

To identify the model, we set the factor loadings of the first observable variable equal to unity. Hence we estimate only three of four factor loadings. However, this implies that the common factor residual variance must be left unrestricted.

The model is estimated by the maximum likelihood method. The estimated parameters together with the standard errors and the p-values are reproduced in Table 4.2.1. The comparison of the true and estimated parameters' values shows that the latter are sufficiently close to the former suggesting that the proposed model estimates the parameters of the data-generating process with a high enough precision. The largest deviations from the true values are found in the case of the autoregressive parameters of the specific components. Nevertheless, even these deviations are not significant.

4.2.3 Real example

We use U.S. monthly coincident time series covering 1959:1-2002:6 — see Table 4.2.2. The data are logged, their first differences are taken and are demeaned and standardized.

The first question to answer when we are using the deterministic mechanism of structural breaks modeling is when these structural breaks have taken place. To answer this question we employed the following procedure as in McConnell and Perez Quiros (2000). Firstly, for the growth rates of each of the time series in question an AR(1) model with a constant was estimated:

$$\Delta y_{it} = \mu_i + \phi_i \Delta y_{it-1} + \xi_{it} \tag{4.18}$$

where $i = 1, ..., n$.

Secondly, the residuals of these models were used to estimate the following model:

$$\sqrt{\frac{\pi}{2}}\,|\xi_{it}| = \alpha_{1i} D_{i1t} + \alpha_{2i} D_{i2t} + \omega_{it} \tag{4.19}$$

where D_{i1t} and D_{i2t} are the dummies capturing structural break in the variance of the $i-th$ time series.

$$D_{i1t} = \begin{cases} 0, & \text{if } t < \tau_i \\ 1, & \text{otherwise} \end{cases}$$
and
$$D_{i2t} = \begin{cases} 0, & \text{if } t > \tau_i \\ 1, & \text{otherwise} \end{cases}$$

The idea is to change the location of the breakpoint τ_i and for each location compute the corresponding Wald statistics. The point where Wald statistics achieves its "supremum" is taken to be *the* breakpoint. However, not all the points of the sample were considered — only those between $0.15T$ and $0.85T$, where T is the sample size, as suggested by Andrews (1993).

To test for the structural break with unknown breakpoint in the variance of the $i-th$ time series we apply the supremum Wald statistic as Andrews (1993) proposes:

$$SupW_i = \max_{\pi} T \left[\frac{\omega_{iR}' \omega_{iR} - \omega_{iU}' \omega_{iU}}{\omega_{iU}' \omega_{iU}} \right] \tag{4.20}$$

where $\pi = \tau / T$. In words, for each time series we are looking for the point of time where the estimated Wald statistic attains its supremum. This Wald statistic's value is taken to be the test statistic.

The test critical values were computed using bootstrapping procedure as described in Diebold and Chen (1996). Our implementation of the procedure was as follows:

1. Estimate model in (4.18) and save the residuals.

2. Estimate restricted analog of equation (4.19), that is, an equation where, instead of two structural break dummies, D_{i1t} and D_{i2t}, we put the constant term.

3. Construct the pseudo-observations of the standard errors, $\sqrt{\frac{\pi}{2}} \left| \hat{\xi}_{it} \right|$, using the estimate of the coefficient with the constant term of the model from step 2 as well as its residuals drawn with replacement.

4. Estimate equation (4.19) and compute the statistic value using equation (4.20). Save this statistic and repeat steps 3-4 large enough number of times.

We undertook a bootstrap with 1000 iterations. The resulting critical values corresponding to 10%, 5%, and 1% significance levels are reported in Table 2.

Then we estimated the Wald statistics for each time series within the subsample obtained from the original sample when 15% of the observations were left out in the beginning and 15% in the end of the sample. Figure 4.2.1 shows the estimated Wald statistics for each of the time series in question. The results of the test aiming at determination of the timing of the structural breaks in the variances of these series are reported in Table 4.2.4. Three time series out of four — for INC the null hypothesis of no structural break could not be rejected — seem to have experienced a structural break in the variance, although the breakpoints are spread across the sample. It is worthwhile to notice, however, that two out of these four time series have had the structural break in variance in the beginning of 1984, namely EMP and IIP. The importance of this circumstance will become more clear later on when we will be discussing the evidence of the decline in the volatility of some US macroeconomic time series in section 4.3.

It makes sense to check the intercept of equation (4.18) for a structural break, because the structural shifts we just identified in the residual variances may, in fact, be due to the changes in the intercepts. Therefore the following unrestricted and restricted regressions were estimated to test the null of no structural breaks in the intercepts:

Unrestricted regression:

$$\Delta y_{it} = \beta_{1i} D_{i1} + \beta_{2i} D_{i2} + \phi_i \Delta y_{it-1} + \xi_{it} \tag{4.21}$$

Restricted regression:

$$\Delta y_{it} = \beta_i D_i + \phi_i \Delta y_{it-1} + \xi_{it} \tag{4.22}$$

where $i = 1, ..., n$. The null hypothesis is formulated as: $D_i = D_{i1} = D_{i2}$ for all i, the dummies being defined as above. Notice that the time subscripts are suppressed, since the structural break locations were fixed at the breakpoints identified in Table 4.2.4. However, the variable subscript is still there meaning that each variable has its specific breakpoint.

The parameter estimates of regressions (4.21 and (4.22 are reported in Tables 4.2.5 and 4.2.6, respectively. The results of the F test conducted using these two regressions are displayed in Table 4.2.7. Only for INC the null hypothesis of no structural break in intercept may be rejected at the usual 0.05 significance level. This implies that most probably the structural breaks detected for the other three variables are due entirely to the breaks in the residual variances.

The results of Table 4.2.4 were used to construct the structural break dummies in the common dynamic factor model. We have estimated eight models: four linear and four with regime-switching dynamics. Under each dynamics assumption the following modifications of the model were considered: (1) no structural break, (2) structural break only in the observed variables' intercepts, (3) structural break only in the residual variance of the specific components, and (4) structural break both in intercepts and in variances. Only specification (0,0) was used. The estimated parameters of these models together with their standard errors are reported in Tables 4.2.8-4.2.11. The respective log-likelihood function values are presented in the header after the specification of the model "linear" or "Markov-switching".

It can be seen (Tables 4.2.8 through 4.2.11) that the structural-break-in-intercept dummies in most cases (except sometimes the INC variable) are not significantly different from zero. This is not the case, however, of the residual variances which are almost always significant at 0.05 significance level. This implies that it was rather the volatility of the growth rates of the US macroeconomic variables that went down during the last 30 years, whereas the average growth rates level did not experience any noteworthy changes.

The linear estimates of the composite economic indicators with and without breaks in intercepts and/or variances are plotted on Figure 4.2.2 of Appendix. The estimates with deterministic dummies are strikingly close to CF(0,0) with no breaks. This is especially true in the case of CF(0,0) with breaks in the intercepts.

The profiles of the nonlinear composite economic indicators with and without deterministic dummies are shown on Figure 4.2.3. It appears that the CF-MS with no break and CF-MS with break in the residual variances have the closest profiles. In contrast, the model with breaks both in intercepts and residual variances of the specific factors displays quite distinct behavior — it has a clearly expressed downward trend which is not the case of other CF-MS.

Different specifications of the model are compared in Table 4.2.12. The likelihood ratio (LR) test is used to conduct this comparison. Each cell of the table contains the double difference between the log-likelihoods of the unrestricted and restricted models. Numbers in the parentheses stand for the degrees of freedom. The asterisks show the test statistics values which exceed the critical $\chi^2_{0.95}$ values. The upper triangular matrix contains the LR-statistics for the linear models, while the lower triangular matrix displays those corresponding to the Markov-switching models. The models 2 and 3 are not nested and therefore cannot be compared using the LR test. Therefore the corresponding cells are left empty.

The model with the structural break only in the intercepts of the observed variables does not lead to an important improvement of the log-likelihood: in the linear case the difference between the model with no structural breaks and the breaks only in the intercepts is not significant, while in the regime-switching case it is rather significant at 5% level. The introduction of the structural breaks in the residual variances of the idiosyncratic components, however, substantially improves the performance of the model. On the other hand, in the linear case there is no significant gain of introducing the structural shifts both in intercepts and variances as compared to the model with breaks in variances only. In the Markov-switching case the estimated test statistic again is very close to the critical value $\chi^2_{0.95}(8) = 21.03$. This implies that the bulk of the improvements in the model stem from including the deterministic dummies taking account of the structural breaks in the residual variances of the specific factors. This can be regarded as an evidence of the negative structural shift in the volatility which has affected the four US macroeconomic time series we are using in our analysis.

Figures 4.2.4 through 4.2.6 illustrate the low-intercept regime probabilities for each of the models with deterministic dummies. These probabilities are superimposed on the NBER business cycle chronology. The simple "eyeball analysis" of the pictures permits concluding that the it is the CF-MS(0,0) with break in the intercepts that replicates the NBER dates the best. The other two models, while capturing good enough most of the recessions, exaggerate the last one which took place in early 1990s. They make it last twice as long as the "official" contraction had lasted. The formal analysis of the similarities between the model-derived datings and the NBER's chronology, however, is postponed to the concluding section of this chapter.

Finally we compare performance of the non-linear models examined in this paper in terms of the in-sample prediction of NBER turning points. To evaluate the performance we use quadratic probability score (QPS) of Diebold and Rudebusch (1989) which is reported in Table 4.2.13. The forecast accuracy of each structural break model is compared to that of the benchmark (no structural break) model using statistic of Diebold and Mariano (1995) with rectangular window of length 201. The null hypothesis states no difference between the predictive accuracy of the two models. The test statistic is standardized and asymptotically distributed as N(0,1).

Among the deterministic break models only model with break in mean gives better estimates of the NBER turning points than the benchmark model. Nevertheless, if we take into account the Diebold-Mariano statistic we have to conclude that there is no significant difference in the in-sample turning points forecast accuracy between the benchmark model and the best model with structural break (deterministic break in the means of the observed variables).

The smoothed conditional probabilities are normally closer to the NBER dates.

4.2.4 Summary

In this section we have introduced a common dynamic factor model with the one-time deterministic shift in the mean of the observed variables. This is often a case when the accounting methodology used in constructing of the statistical indicators is changed. Usually this introduces a discontinuity into the observed time series not allowing to compare the dynamics before and after the structural break.

We consider the models with both the linear and regime-switching dynamics having a single common dynamic factor. An illustrative artificial example for a linear dynamics case with a structural break in the intercept of one of the observed series was estimated. It shows a high enough fitting ability of these models when they correspond to the true data-generating process.

Next, the model allowing structural breaks both in the intercepts of the observed time series and in the residual variances of the idiosyncratic factors was estimated for the real US post-World War II monthly data. A model without structural breaks is compared to (1) model with structural breaks in the observed variables intercepts; (2) model with structural breaks in the residual variances of the specific factors, and (3) model with structural breaks both in the intercepts and variances. It turns out that the hypothesis of structural break in the means finds no support in the real data, whereas the hypothesis of the structural breaks in the variances, although occurred at different points of time for different observed variables, is likely to be confirmed by the empirical evidence.

4.2.5 Appendix

Table 4.2.1. True and estimated parameters of the linear common factor model with structural break

Coefficient	True	Estimated	St. error	p-value
γ_1	1	–	–	–
γ_2	0.9	0.89	0.02	0.0
γ_3	1.2	1.19	0.02	0.0
γ_4	0.7	0.68	0.02	0.0
δ_{21}	0	-0.01	0.02	0.33
δ_{22}	5	5.03	0.03	0.0
ϕ	0.8	0.82	0.03	0.0
ψ_1	-0.3	-0.35	0.06	0.0
ψ_2	-0.7	-0.71	0.03	0.0
ψ_3	-0.5	-0.51	0.07	0.0
ψ_4	-0.2	-0.13	0.05	0.0
σ_C^2	0.81	0.76	0.06	0.0
σ_1^2	0.25	0.25	0.02	0.0
σ_2^2	0.36	0.36	0.03	0.0
σ_3^2	0.16	0.16	0.02	0.0
σ_4^2	0.49	0.48	0.03	0.0

Table 4.2.2. The component series of the US composite economic indicator
Monthly data, January 1959 – June 2002

Series	Short-hand	Description
Employees on nonagricultural payrolls	EMP	10^3, SA
Personal income less transfer payments	INC	10^9 1996 \$, SA, annual rate
Index of industrial production	IIP	total index, 1996=100, SA
Manufacturing and trade sales	SLS	chained 10^6 1996 \$, SA

Notation: 10^9 1996 \$ = billions of 1996 dollars; SA = seasonally adjusted.
Note: Chained (1996) dollar series are calculated as the product of the chain-type quantity index and the 1996 current-dollar value of the corresponding series, divided by 100.
Source: NBER (www.nber.org/cycles/hall.xlw): the industrial production series has an erroneous entry in December 1985 which was replaced by the figure taken from the index of industrial production with 1992 base, given that the neighboring values (before and after 1985:12) are exactly the same for both indices.

Table 4.2.3. Bootstrap critical values of the Wald statistics
1000 iterations

Variable	Significance level		
	10%	5%	1%
EMP	13.071	14.620	17.609
INC	13.053	14.220	16.826
IIP	12.726	14.119	17.987
SLS	13.324	14.880	17.270

Table 4.2.4. Structural breaks timing

Variable	Wald statistic	Date
EMP	64.25**	1984:3
INC	7.171	1972:6
IIP	25.59**	1984:4
SLS	13.48*	1992:2

The statistic has superscripts * and ** if it exceeds 10% and 5% critical value, respectively.

Table 4.2.5. Unrestricted regression
Breakpoints: as suggested by Table 2

Variable	D_1	D_2	Δy_{t-1}
EMP	0.0031	-0.0057	0.4363
INC	0.0540	-0.0281	0.2623
IIP	0.0089	-0.0249	0.3765
SLS	-0.0273	0.1170	-0.0894

Table 4.2.6. Restricted regression
Breakpoint: none

Variable	D	Δy_{t-1}
EMP	-0.0002	0.4369
INC	-0.0006	0.2752
IIP	-0.0036	0.3771
SLS	-0.0025	-0.0871

Null hypothesis: $D = D_1 = D_2$

Table 4.2.7. F-test for structural break in intercept
Based on results of the regressions in Tables 4 and 5

Variable	F-statistic	p-value
EMP	0.188	0.664
INC	4.233	0.040
IIP	0.188	0.665
SLS	1.282	0.258

Table 4.2.8. Estimated parameters of the linear and regime-switching models with no structural break 1959:1-2002:6

	Linear LL=-2627.9		Markov-switching LL=-2547.8	
Parameter	Estimated	St.error	Estimated	St.error
p_{11}	–	–	0.979	0.008
$1 - p_{22}$	–	–	0.085	0.031
μ_1	–	–	0.296	0.042
μ_2	–	–	-1.16	0.104
γ_{INC}	0.825	0.057	0.778	0.052
γ_{IIP}	1.00	0.065	0.887	0.056
γ_{SLS}	0.710	0.062	0.619	0.056
σ^2	0.639	0.066	0.385	0.040
σ^2_{EMP}	0.359	0.039	0.270	0.037
σ^2_{INC}	0.563	0.042	0.557	0.040
σ^2_{IIP}	0.353	0.040	0.426	0.039
σ^2_{SLS}	0.676	0.047	0.720	0.049

LL = the value of loglikelihood function

Table 4.2.9. Estimated parameters of the linear and regime-switching models with structural break in means 1959:1-2002:6

	Linear LL=-2624.4		Markov-switching LL=-2538.5	
Parameter	Estimated	St.error	Estimated	St.error
p_{11}	–	–	0.979	0.008
$1 - p_{22}$	–	–	0.104	0.036
μ_1	–	–	0.364	0.073
μ_2	–	–	-1.21	0.114
γ_{INC}	0.816	0.057	0.773	0.052
γ_{IIP}	1.01	0.066	0.914	0.056
γ_{SLS}	0.713	0.062	0.632	0.056
σ^2	0.634	0.066	0.352	0.038
σ^2_{EMP}	0.358	0.039	0.288	0.035
σ^2_{INC}	0.561	0.042	0.558	0.040
σ^2_{IIP}	0.353	0.040	0.408	0.037
σ^2_{SLS}	0.676	0.047	0.716	0.048
$\delta_{1.EMP}$	0.054	0.054	0.026	0.060
$\delta_{1.INC}$	0.127	0.070	0.103	0.075
$\delta_{1.IIP}$	0.025	0.058	0.0	0.036
$\delta_{1.SLS}$	-0.002	0.032	-0.044	0.067
$\delta_{2.EMP}$	-0.074	0.061	-0.252	0.091
$\delta_{2.INC}$	-0.056	0.047	-0.147	0.068
$\delta_{2.IIP}$	-0.035	0.059	-0.199	0.090
$\delta_{2.SLS}$	0.006	0.032	-0.101	0.093

LL = the value of loglikelihood function

Table 4.2.10. Estimated parameters of the linear and regime-switching models with structural break in variances 1959:1-2002:6

	Linear LL=-2586.7		Markov-switching LL=-2492.0	
Parameter	Estimated	St.error	Estimated	St.error
p_{11}	–	–	0.978	0.008
$1 - p_{22}$	–	–	0.073	0.026
μ_1	–	–	0.290	0.034
μ_2	–	–	-1.13	0.078
γ_{INC}	0.900	0.066	0.816	0.056
γ_{IIP}	0.978	0.072	0.795	0.056
γ_{SLS}	0.769	0.072	0.615	0.059
σ^2	0.553	0.065	0.301	0.034
σ^2_{EMP}	0.175	0.034	0.029	0.024
σ^2_{INC}	0.551	0.043	0.568	0.040
σ^2_{IIP}	0.183	0.033	0.293	0.035
σ^2_{SLS}	0.601	0.083	0.664	0.086
λ_{EMP}	0.363	0.069	0.426	0.058
λ_{INC}	0.0	0.0	0.0	0.0
λ_{IIP}	0.365	0.070	0.364	0.080
λ_{SLS}	0.089	0.100	0.113	0.104

LL = the value of loglikelihood function

Table 4.2.11. Estimated parameters of the linear and regime-switching models with structural break in intercepts and variances 1959:1-2002:6

	Linear LL=-2582.9		Markov-switching LL=-2484.8	
Parameter	Estimated	St.error	Estimated	St.error
p_{11}	–	–	0.978	0.008
$1-p_{22}$	–	–	0.075	0.028
μ_1	–	–	0.267	0.080
μ_2	–	–	-1.17	0.124
γ_{INC}	0.893	0.067	0.799	0.057
γ_{IIP}	0.983	0.073	0.793	0.058
γ_{SLS}	0.776	0.073	0.613	0.060
σ^2	0.545	0.064	0.292	0.034
σ^2_{EMP}	0.173	0.034	0.026	0.026
σ^2_{INC}	0.549	0.042	0.569	0.041
σ^2_{IIP}	0.183	0.033	0.295	0.036
σ^2_{SLS}	0.599	0.082	0.660	0.086
$\delta_{1.EMP}$	0.055	0.056	0.132	0.090
$\delta_{1.INC}$	0.130	0.070	0.180	0.091
$\delta_{1.IIP}$	0.027	0.053	0.090	0.080
$\delta_{1.SLS}$	-0.002	0.036	0.029	0.069
$\delta_{2.EMP}$	-0.075	0.050	-0.064	0.073
$\delta_{2.INC}$	-0.058	0.050	-0.023	0.041
$\delta_{2.IIP}$	-0.037	0.047	-0.031	0.067
$\delta_{2.SLS}$	0.005	0.024	0.034	0.091
λ_{EMP}	0.366	0.069	0.427	0.058
λ_{INC}	0.0	0.0	0.0	0.0
λ_{IIP}	0.365	0.070	0.362	0.080
λ_{SLS}	0.090	0.099	0.120	0.105

LL = the value of loglikelihood function

Table 4.2.12. Comparison of different modifications of the model
Specifications CF(0,0) and CF-MS(0,0). Likelihood ratio test

Linear models			
No SB	7.0 (8)	82.4* (4)	90.1* (12)
18.7* (8)	SB in M		83.1* (4)
111.7* (4)		SB in V	7.7 (8)
126.0* (12)	107.3* (4)	14.3 (8)	SB in M & V
Markov-switching models			

"No SB" = no structural break; "SB in M" = structural break in the mean of observed variables; "SB in V" = structural break in the residual variance of specific components; "SB in M & V" = structural break both in the mean and in the residual variance.
The asterisks show the LR-statistics exceeding the critical $\chi^2_{0.95}$ values.

Table 4.2.13. In-sample forecasting performance of the Markov-switching common factor models with deterministic structural break
Model-derived recession probabilities compared to the NBER business cycle chronology, 1959:1-2002:6

Model	QPS	DM	p-value
Filtered probabilities			
No break (benchmark model)	0.050	—	—
Break in mean	0.047	0.378	0.353
Break in variance	0.060	-1.27	0.102
Break in mean and variance	0.056	-0.862	0.194
Smoothed probabilities			
No break (benchmark model)	0.037	—	—
Break in mean	0.029	0.959	0.169
Break in variance	0.051	-1.89	0.029
Break in mean and variance	0.044	-0.779	0.218

QPS = quadratic probability score
DM = Diebold-Mariano statistic testing the hypothesis of equality of the forecast accuracy of two alternative models — see Diebold and Mariano (1995)
p-value = significance value of DM-statistic

Figure 4.2.1: Testing the timing of structural break

Wald statistic

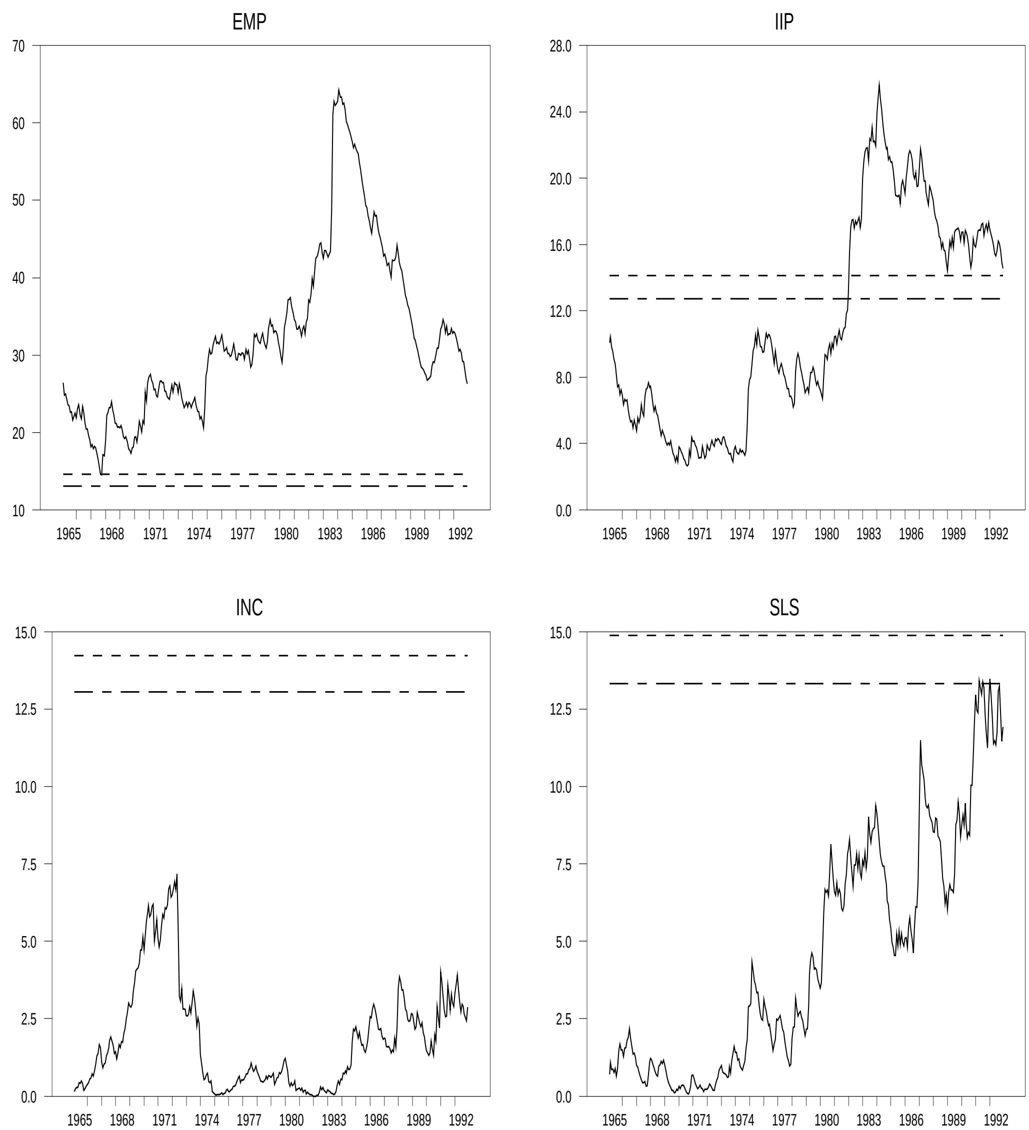


Figure 4.2.2: Linear CF(0,0) with and without dummies

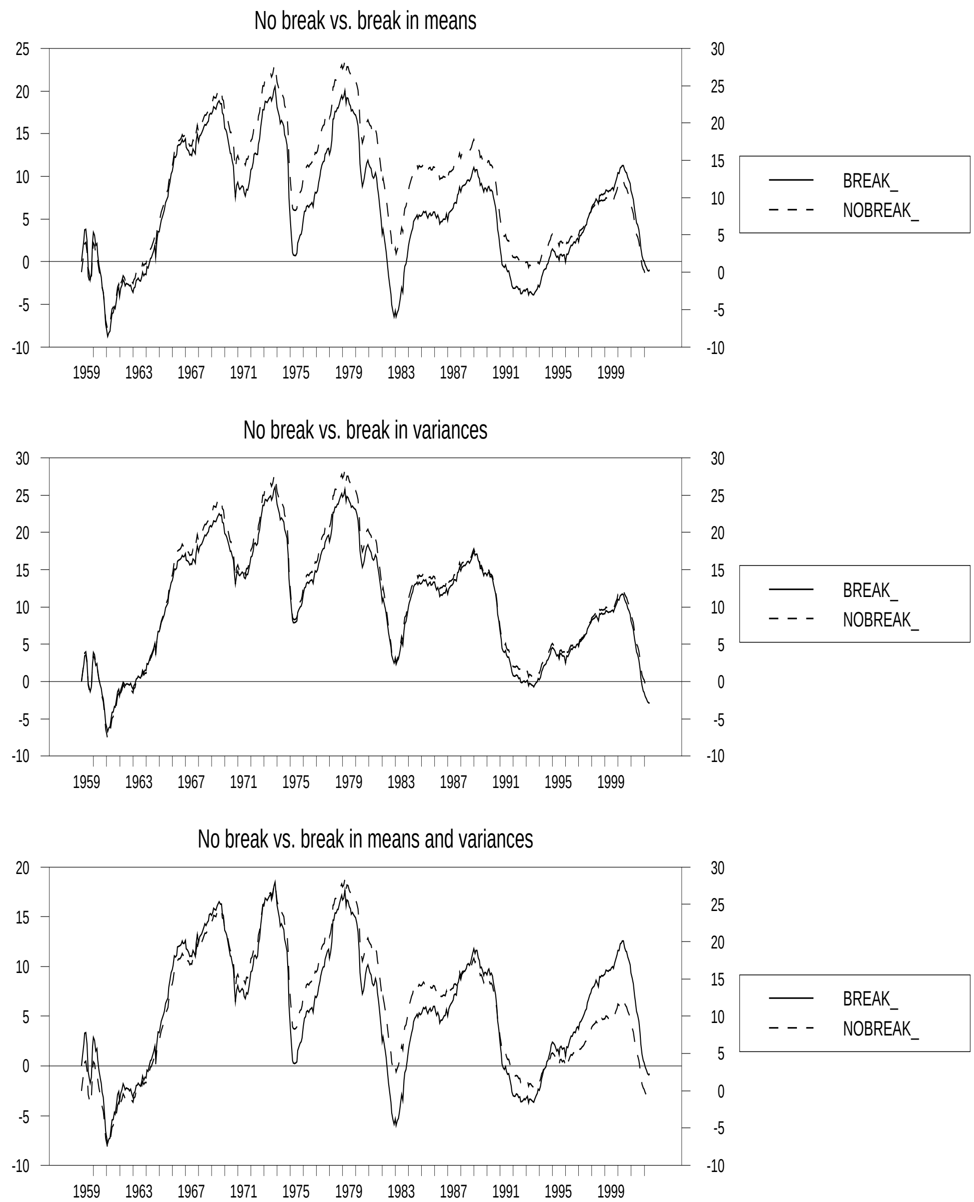


Figure 4.2.3: CF-MS(0,0) with and without dummies

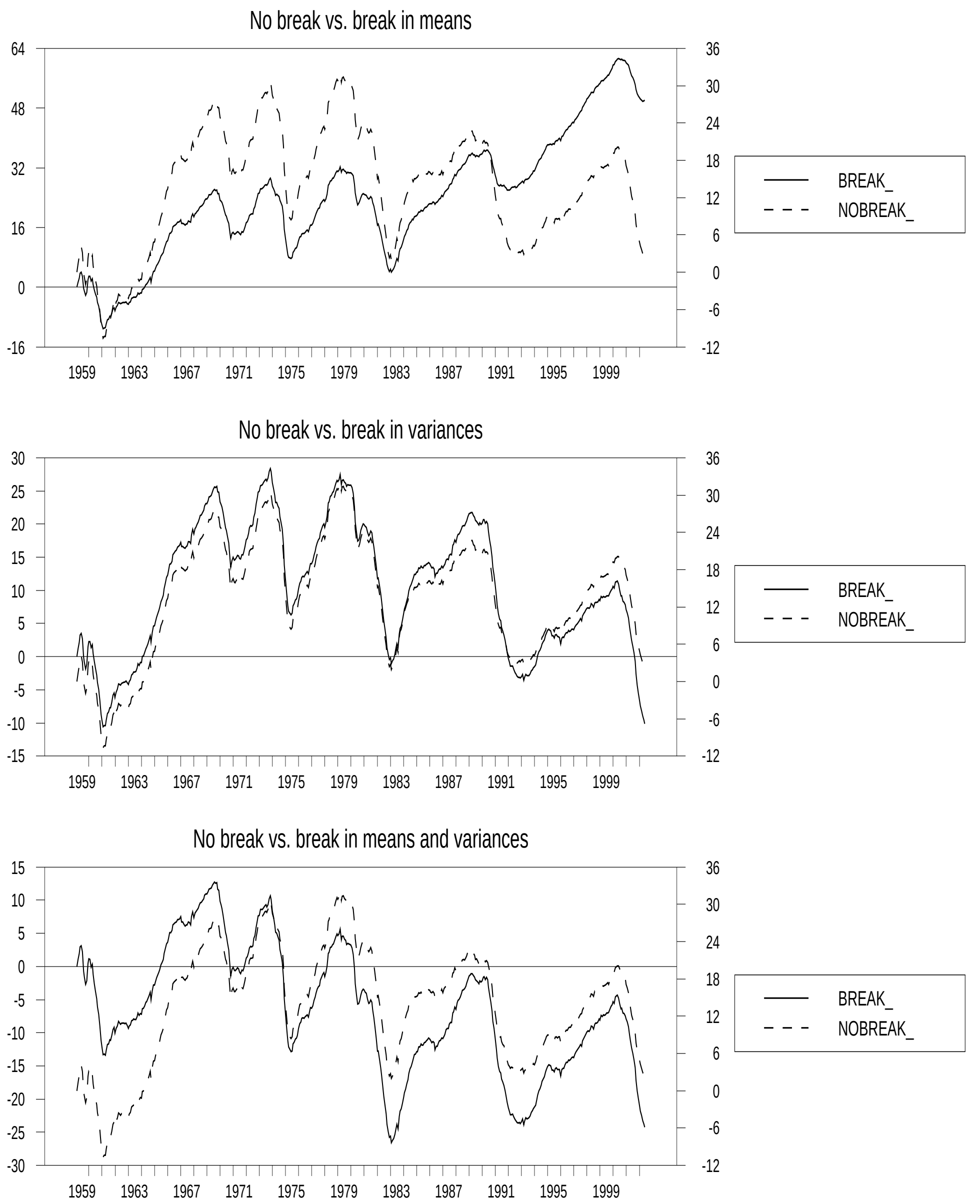


Figure 4.2.4: Recession probabilities vs. NBER dates

Model with break in means

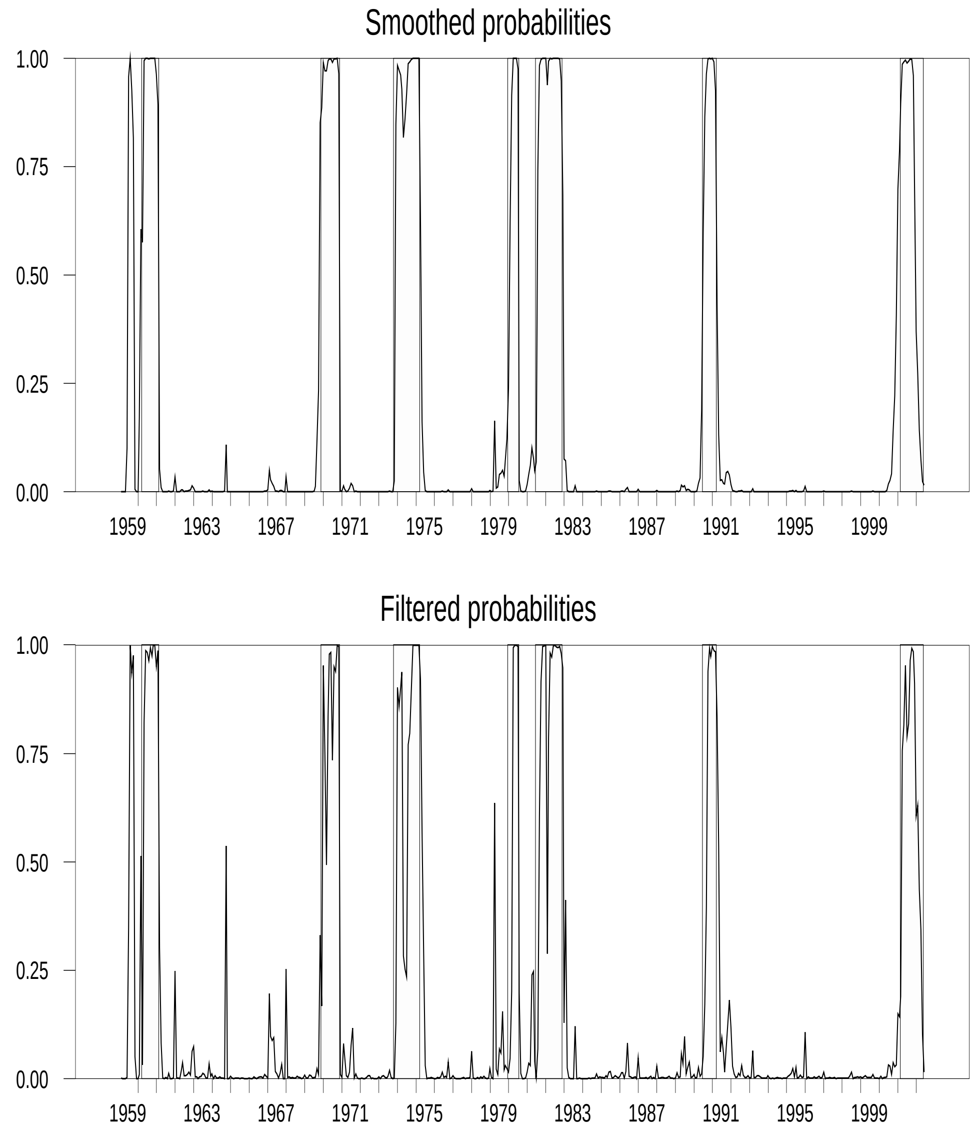


Figure 4.2.5: Recession probabilities vs. NBER dates

Model with break in variances

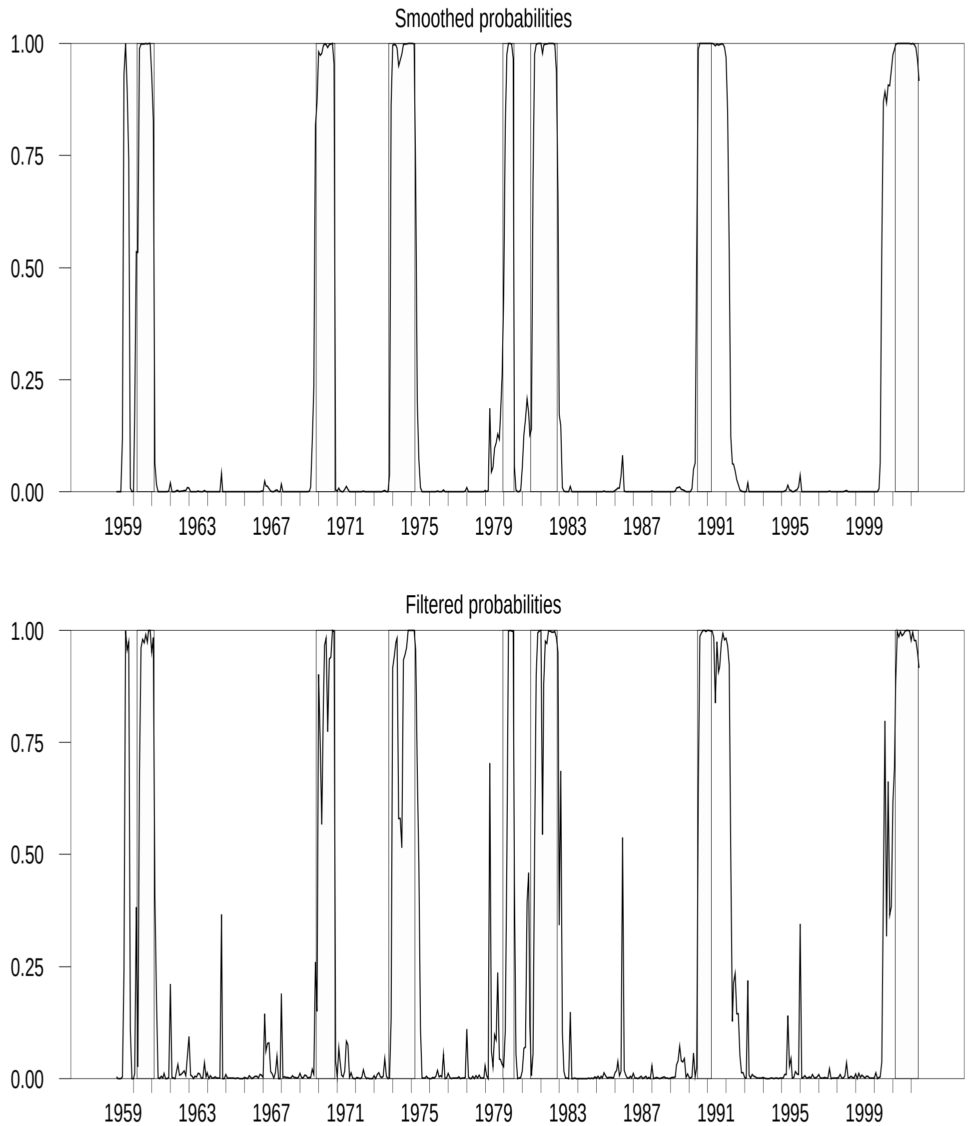
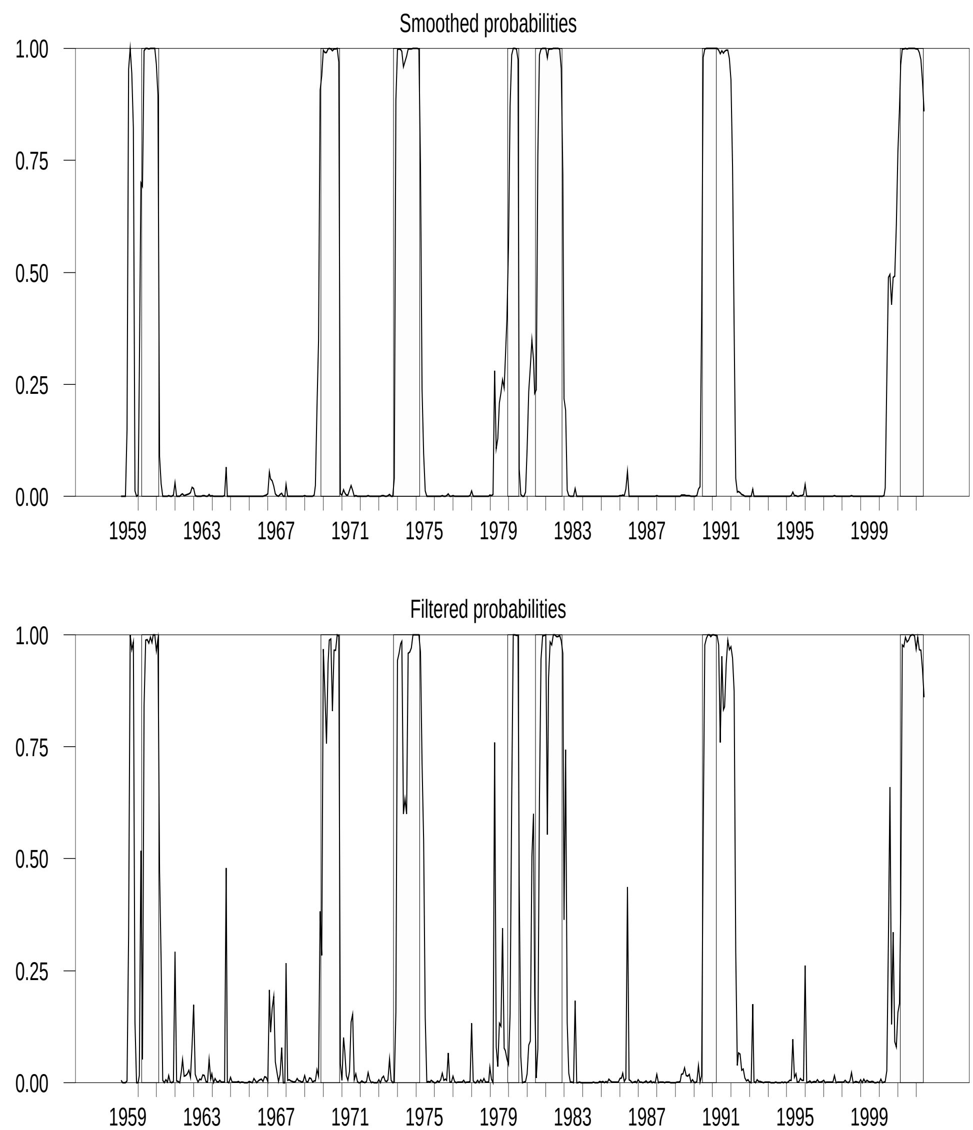


Figure 4.2.6: Recession probabilities vs. NBER dates

Model with break in means and variances



4.3 Structural Breaks in Dynamic Factor Model: Probabilistic Mechanism

This section studies a probabilistic model of structural break. Unlike the previous section, here we concentrate only on the structural break in volatility. However, given the empirical evidence presented both above, in subsection 4.2.3, and below, in section 4.3.2, for the US data there is little evidence to support the hypothesis of structural change in the mean, while there is strong support for the hypothesis of the structural shifts in the variance.

Another difference is that in this section we consider a single structural break which hit the common dynamic factor, not the idiosyncratic factors as in the preceding section. In this sense the model we are going to present is somewhat less flexible. Nevertheless, it has an important advantage since it does not impose the predefined breakpoints as the deterministic model of section 4.3, rather estimating the moment when the structural break had presumably taken place.

4.3.1 Model

We consider a set of the coincident time series which are supposed to evolve at the same pace as a current state-of-affairs indicator (e.g. real GDP). The common dynamics of the coincident time series are explained by a latent common factor. The idiosyncratic dynamics of each time series are captured by one specific factor per each observed time series. Formally:

$$\Delta y_t = \Gamma \Delta f_t + u_t \tag{4.23}$$

where Δy_t is the $n \times 1$ vector of the observed time series in the first differences; Δf_t is the scalar representing the latent common factor in the first differences; u_t is the $n \times 1$ vector of the latent specific factors; Γ is the $n \times 1$ factor loadings vector linking the observed series with the common factors.

The dynamics of the latent common factor can be described in terms of a nonlinear autoregressive (AR) model:

$$\Delta f_t = \mu(s_t) + \phi(L)\Delta f_{t-1} + \varepsilon_t \tag{4.24}$$

where $\mu(s_t)$ is the state-dependent intercept of the common coincident factor which takes different values depending on the regime; $\phi(L)$ is the common factor AR lag polynomial of order p; ε_t is the serially uncorrelated common factor disturbance term with possibly state-dependent variance:

$$\varepsilon_t \sim NID\left(0, \sigma^2(s_t)\right)$$

s_t is the unobserved regime variable. In the two-regime (expansion-recession, or high-low) case it takes two values: 0 or 1. Depending on the regime, the common factor's intercept assumes different values: low in recessions and high in expansions. Thus, the common factor grows faster during the upswings and slower (or even have negative growth rate) during the downswings of the economy. Here we introduce yet another dimension to the problem by allowing the common factor residual variance, $\sigma^2(s_t)$, to have its own low and high regimes — regime of low volatility and regime of high volatility. We assume that the "mean" state variable, s_t^μ, is independent of the residual variance state variable, s_t^σ. This is a kind of model which was used by McConnell and Perez Quiros (2000) to examine the Post-World War II evolution of the US quarterly real GDP.

The changes in the regimes are governed by the first-order Markov chain process, which is summarized by the transition probabilities matrix, whose characteristic element is $p_{ij} = prob(s_t = j|s_{t-1} = i)$.

Since we have two parameters — intercept and variance — each of which passes through its own low and high regimes, the whole process should be cast in a four regimes framework as it is done in McConnell and Perez Quiros (2000). Namely:

	Regime 1	Regime 2	Regime 3	Regime 4
Composite state variable	$s_t = 0$	$s_t = 1$	$s_t = 2$	$s_t = 3$
Intercept state variable	$s_t^\mu = 0$	$s_t^\mu = 1$	$s_t^\mu = 0$	$s_t^\mu = 1$
Variance state variable	$s_t^\sigma = 0$	$s_t^\sigma = 0$	$s_t^\sigma = 1$	$s_t^\sigma = 1$

where s_t^μ and s_t^σ are the unobserved state variables for common factor intercept and common factor residual variance, respectively. Each state variable has its own 2×2 transition probabilities matrix:

		s_t^j	
		High	Low
s_{t-1}^j	High	p_{11}^j	$1 - p_{11}^j$
	Low	$1 - p_{22}^j$	p_{22}^j

where $j = \{\mu, \sigma\}$.

Given that the state variables s_t^{μ} and s_t^{σ} staying behind the evolution of the common factor intercept and residual variance are independent, the 4×4 transition probabilities matrix, π, governing the behavior of the "composite" state variable, s_t, would look as:

$$\begin{pmatrix} p_{11}^{\mu} p_{11}^{\sigma} & (1-p_{11}^{\mu})p_{11}^{\sigma} & p_{11}^{\mu}(1-p_{11}^{\sigma}) & (1-p_{11}^{\mu})(1-p_{11}^{\sigma}) \\ (1-p_{22}^{\mu})p_{11}^{\sigma} & p_{22}^{\mu} p_{11}^{\sigma} & (1-p_{22}^{\mu})(1-p_{11}^{\sigma}) & p_{22}^{\mu}(1-p_{11}^{\sigma}) \\ p_{11}^{\mu}(1-p_{22}^{\sigma}) & (1-p_{11}^{\mu})(1-p_{22}^{\sigma}) & p_{11}^{\mu} p_{22}^{\sigma} & (1-p_{11}^{\mu})p_{22}^{\sigma} \\ (1-p_{22}^{\mu})(1-p_{22}^{\sigma}) & p_{22}^{\mu}(1-p_{22}^{\sigma}) & (1-p_{22}^{\mu})p_{22}^{\sigma} & p_{22}^{\mu} p_{22}^{\sigma} \end{pmatrix}$$

In fact, $\pi = \pi^{\mu} \otimes \pi^{\sigma}$, where π^{μ} and π^{σ} are the transition probabilities matrices for the state variables s_t^{μ} and s_t^{σ}.

Thus, in our four-regime model we have four state-dependent means, μ_{ij}, where $i = \{s_t^{\mu} = 0, s_t^{\mu} = 1\}$ and $j = \{s_t^{\sigma} = 0, s_t^{\sigma} = 1\}$, and two state-dependent residual variances, σ_j^2, $j = \{high, low\}$.

A restricted version of the above model was also considered. It is a variant proposed by Kim and Nelson (1999a) for the univariate US real GDP data. They regard the low volatility regime as an absorbing state. In other words, whenever the system attains this state, it remains there forever. This assumption translates into the following constraint imposed on the transition probabilities matrix π^{σ}:

$$\pi^{\sigma} = \begin{pmatrix} p_{11}^{\sigma} & 1-p_{11}^{\sigma} \\ 0 & 1 \end{pmatrix}$$

The quantity $\frac{1}{1-p_{11}^{\sigma}}$ measures the expected duration of the high volatility regime and hence indirectly indicates the approximate location of the break point. The two models — unrestricted and restricted — can be compared using the standard likelihood ratio (LR) test.

The unobserved idiosyncratic factors are by definition mutually independent and are modelled as the AR processes:

$$u_t = \Psi(L)u_{t-1} + \eta_t \tag{4.25}$$

where $\Psi(L)$ is the sequence of q ($q = \max\{q_{1,\ldots}, q_n\}$, where q_i is the order of the AR polynomial of the $i-th$ idiosyncratic factor) $n \times n$ diagonal lag polynomial matrices and η_t is the $n \times 1$ vector of the mutually and serially uncorrelated normally distributed shocks:

$$\eta_t \sim NID\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}\right)$$

The state-space form of this model would be as follows:

$$\Delta y_t = A x_t \tag{4.26}$$

$$x_t = \alpha(s_t) + C x_{t-1} + v_t \tag{4.27}$$

where $x_t = (f_t | u_t)'$ is the state vector containing stacked on top of each other vector of common factor and the vector of specific factors; v_t is the vector of the common and idiosyncratic factors' disturbances with mean zero and variance-covariance matrix Q; $\alpha(s_t) = (\mu(s_t), 0, ..., 0)'$ is the state-dependent vector of intercepts.

$$A = \begin{pmatrix} \Gamma & i_{q_1} & \dots & 0 \\ o_{n\times 1} & 0 & \dots & i_{q_n} \end{pmatrix}$$

where Γ is the $n \times 1$ vector of the common factor loadings:

$$\Gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}$$

which, to simplify the matters, we represent as a column vector, although when some of the observed time series are lagging with respect to the common factor we would need to include additional columns with the factor loadings.

$O_{n\times m}$ is $n \times m$ matrix of zeros; o_m is the $m \times 1$ vector of zeros; i_m is the first row of the $m \times m$ identity matrix.

$$C = \begin{pmatrix} \Phi & 0 & & 0 \\ 0 & \Psi^1 & & \\ & & \ddots & \\ 0 & & & \Psi^n \end{pmatrix}$$

where Φ is the $p \times p$ companion matrix of the AR polynomial of the common dynamic factor:

$$\Phi = \begin{pmatrix} \phi \\ I_{p-1} \end{pmatrix}$$

where ϕ is the $1 \times p$ row vector of the AR coefficients of the leading factor, I_n is the $n \times n$ identity matrix, and o_m is the $m \times 1$ vector of zeros.

The matrices $\Psi^1, ..., \Psi^n$ have the same structure as Φ.

There are different ways of estimating the unknown parameters and the latent factors (maximum likelihood, EM, MCMC techniques — see Kim and Nelson (1999b) for more details). Here we applied the maximum likelihood method with log-likelihood function obtained using Kalman filter recursions. To save space we will not present them here, referring the reader, for instance, to Hamilton (1994) who gives very clear and systematic explanation of the Kalman filter methodology.

4.3.2 Real example

The previous research (e.g. McConnell and Perez Quiros (2000)) suggests that a structural break in the US economy had presumably taken place in early 1984. For the sake of illustration we analyze the behavior of the major US macroeconomic series before and after that date (see Table 4.2.1 in Appendix to section 4.2 which lists the component series of our composite economic indicator). Recall that only two of the four series had their structural breaks close to the early 1984, as we have seen in section 4.2. Figures 4.3.1-4.3.2 show the evolution of the growth rates of the US real GDP (quarterly data) and the four monthly component series of the US composite economic indicator listed in Table 4.2.1. Besides the time series themselves, their means and the two standard deviations band for two subperiods — 1959:1-1983:12 for monthly data (1959:2-1983:4 for GDP) and 1984:1-2002:6 for monthly data (1984:1-2002:2 for GDP) — are plotted on the graphs. This allows evaluating the scale of shifts in the mean and variance of each time series. The means apparently have not changed much, while the variances, especially those of GDP, nonagricultural employment and industrial production have undergone an important decline.

The quantitative characterization of these changes may be found in Tables 4.3.1a-4.3.1b of Appendix which contain means, standard deviations (St.dev.), coefficients of variation (CV), minima (Min) and maxima (Max) of the time series in question before and after the beginning of 1984. As comparison of Tables 1a and 1b shows, the means have decreased, although not significantly. The variances have gone down, especially those of employment and industrial production which in the second subsample have experienced almost double reduction. The coefficients of variation fell down too, save for the case of industrial production where the variance diminished faster than the mean. It seems also that the growth rates have changed somewhat asymmetrically: if in two cases out of five (INC and SLS) the lowest growth rates where attained in the second subperiod, only in one case out of five (SLS) the highest growth rate had been achieved after January 1984.

The formal test[2] of the differences between the means and variances in the two subsamples is contained in Table 4.3.2. The columns two and four represent the test statistics values — Z distributed as a normal and F following $F(n_1, n_2)$ distribution — for means and variances, respectively. n_1 and n_2 stand for the sizes of each of two subsamples. The p-values (see columns three and five of Table 4.3.2) computed using these test statistics allow testing the null hypothesis of no difference between the moments of the two subperiods. One can see that only for the mean of the variable INC the null hypothesis may be rejected at significance level of 10%, while the rest of the means seem not to change. What for the variances, here the structural break appears to happen for all the time series under inspection. The largest decrease in the growth rate variance took place in nonagricultural employment, GDP, and industrial production. Thus, it makes sense to talk about a structural break in 1983-1984 which turned out to affect mainly the volatility of growth but not the average level of the growth rates.

Given the fact that the GDP has experienced large volatility structural break as our own calculations show and as was discovered, e.g., by McConnell and Perez Quiros (2000), we decided to estimate both a single-frequency model based only on the monthly time series and a mixed-frequency model which uses, in addition to the monthly series, the quarterly real GDP data.

The data used are as in section 4.1. Namely, for the estimation of the common coincident factor model with single-frequency data the four US monthly time series stretching from January 1959 through June 2002 were used: EMP, INC, IIP, and SLS (see Table 4.2.1). The time series in levels were logged, then their first differences were taken and multiplied by 100. Finally, all the component series were demeaned and normalized.

Six models were estimated: (1) single-frequency model with no structural break (2-regime model); (2) mixed-frequency model with no structural

[2] For details on inferences based on two samples see Devore (1987).

break (2-regime model); (3) unrestricted single-frequency model with structural break (4-regime model); (4) unrestricted mixed-frequency model with structural break (4-regime model); (5) restricted single-frequency model with structural break (4-regime model), and finally (6) restricted mixed-frequency model with structural break (4-regime model).

All the models were estimated under the identifying assumption of the first factor loading being equal 1. The results of estimation of models 1 and 2 are reported in Table 4.3.4. The parameter estimates, together with their standard errors, of the unrestricted and restricted single-frequency models with structural break are in Table 4.3.4, while those of the unrestricted and constrained mixed-frequency models with structural break - in Table 4.3.5. In all the models both the common factors and idiosyncratic factors are specified as AR(0).

The LR test of the restriction imposed on the transition probabilities matrix of the variance state variable shows that the null hypothesis of $p^{\sigma}_{22} = 1$ can be rejected at 5% significance level: the estimated test value for the single-frequency model is equal 24.8 and for the mixed-frequency model it is 22.4 (see the first rows of Tables 4.3.4 and 4.3.5) against $\chi^2_{0.95}(1) = 3.84$.

The low residual standard error of the unobserved common factor is about 10 times smaller than the high residual standard error of the common factor in case of both restricted and unrestricted single-frequency models. For the 4-regime mixed-frequency models this ratio is approximately 3.5. This difference is somewhat surprising, given that the pre-1984 to post-1984 ratio of standard deviations of the individual component time series does not exceed 2.

The state-dependent intercepts corresponding to the high variance regime (μ_{11} and μ_{21}) are much greater in the absolute value than their counterparts in the low-volatility regime. This implies that the shift in the volatility was accompanied by a "stabilization" of the growth rates. Both the recessions and expansions became milder.

The expected duration of the high volatility state computed using the estimate of the transition probability, p^{σ}_{11}, in both cases (single- and mixed-frequency models) is equal 333 months. If we assume that the high variance state commences at the very beginning of the sample (February 1959), the 333rd period will correspond to October 1986. The date is somewhat late compared to the beginning of 1984 proposed as the date of the start of volatility decline.

Figure 4.3.3 displays the estimates of the common coincident factor obtained using the single-frequency models and mixed-frequency models plotted against the NBER's US business cycle dates. The common factor in both cases was reconstructed as the partial sums of its growth rates obtained as an output of the estimation procedure. Unlike Stock and Watson (1991), we are not adding the drift in order to be able to see the composite economic indicator's fluctuations more clearly. Seven major recessions can be observed on the picture which fit quite well the NBER recession dating (shaded areas), with the composite economic indicator leading sometimes the dating. The common factor resulting from the mixed-frequency model has a far smaller variance than that of the single-frequency model. The profiles of the common factor from the models with and without structural break do not display any perceivable differences.

Figure 4.3.4 illustrates the conditional recession (low-intercept) probabilities of the benchmark model, that is, the model without structural break. Two cases are considered: with single-frequency and mixed-frequency data. The recession probabilities appear to correspond quite well to the NBER business cycle chronology represented by the shaded areas as contractions. There are three noteworthy differences: (1) the conditional probabilities detect a recession in late 1950s which is not treated as such by the NBER; (2) the "model-derived" recession of early 1990s is twice as long as the "official" one, especially in the single-frequency case; (3) the last contraction which, according to NBER started in March 2001, is signalled well in advance by the conditional probabilities.

On Figures 4.3.5-4.3.6 the conditional (smoothed) probabilities of the low intercept regime (sum of the conditional probabilities corresponding to the regimes 2 and 4) and low intercept — high variance regime (conditional probabilities corresponding to the regime 2) for both the single- and mixed-frequency model are depicted. Both conditional probabilities are contrasted against the NBER's business cycle chronology. However, only the conditional probabilities of the low intercept — high variance regime display fairly high degree of conformity to this dating — see Figure 4.3.6. The low regime probabilities detect quite a bit of false signals: two false recessions in 1960s, one in 1980s, and one in 1990s. The picture is the same for the two models, although the mixed-frequency model attenuates slightly the false alarms compared to the single-frequency model. Nevertheless, the situation is improved when only the conditional probabilities of the low intercept — high variance regime are considered: they display fairly high degree of conformity to NBER dating (see Figure 4.3.6).

Figure 4.3.7 displays the conditional (smoothed) probabilities of the low variance regime (sum of the conditional probabilities corresponding to the regimes 1 and 2) for both unrestricted models, too. It seems that, regardless of model, this regime becomes much more probable since February-March of 1984. From that period on the conditional probabilities of the coincident economic indicator having low residual variance are almost always — except for two short interruptions — exceeding 0.7. This evidence is in accordance with the finding of McConnell and Perez Quiros (2000) who using the quarterly GDP data point out to the first quarter of 1984 as the beginning of "low volatility era".

The low-intercept and low-variance regime probabilities corresponding to the restricted models — with single and mixed frequency — are displayed on Figures 4.3.8 and 4.3.9, respectively. The recession (low intercept) probabilities obtained from the constrained model do not differ from those resulting from the unconstrained estimation. The restricted model low volatility regime probabilities are much smoother thanks to the restriction imposed on the transition probability, p^{σ}_{22}. The smoothed probabilities signal the arrival of the low volatility regime earlier than the filtered conditional probabilities do. One can clearly see the frontier between the high and low variance regimes which passes through the middle of 1984. This does not contradict the results of the previous studies.

We conclude the analysis of the models examined above by comparing their performance in terms of the in-sample prediction of the NBER turning points. We use the quadratic probability score (QPS) suggested by Diebold and Rudebusch (1989) to evaluate the forecasting accuracy. QPS compares the recession probabilities derived from some model to a generally accepted business cycle dating. In the USA case it is normally the NBER's recession dates.

Table 4.3.6 reports QPS both for the models with structural break and those without it. Columns 2 and 3 display QPS corresponding to the filtered and smoothed regime probabilities. The smoothed probabilities seem to outperform the filtered ones. This is due to the fact that smoothing eliminates the smaller spikes which are very typical for the filtered probabilities and which reflect nothing but the noise. The low-intercept regime probabilities computed for the unrestricted models with break have very poor performance. However, the low intercept — high volatility regime probabilities derived from these models appear to fit better. This confirms what we saw on Figures 4.3.5 and 4.3.6.

The fact that QPS of different models are different is not sufficient to conclude that one model is superior with respect to another. The difference may be statistically insignificant. In order to check this we computed the statistic proposed by Diebold and Mariano (1995) to test the hypothesis of equality of the forecast accuracy of two alternative models. The Diebold-Mariano (DM) statistic was calculated using the rectangular spectral window of length 21. The forecast accuracy of one model-derived chronology is compared to that of other model-derived dating. The null hypothesis states no difference between the predictive accuracy of the two models. The test statistic is standardized and hence it is asymptotically distributed as N(0,1).

Since the DM-statistic is pairwise and we have 8 model-derived chronologies to compare, there can be many different combinations. To save space we do not report all of them, only two tables comparing the models without structural break serving as a benchmark to those with break (Tables 4.3.7a and 4.3.7b). In Table 4.3.7a filtered conditional probabilities of all the models are compared to the filtered and smoothed probabilities of two benchmark models (with single- and mixed-frequency data). These two models are endowed with Markov-switching with two regimes: low and high intercept. Table 4.3.7b compares the smoothed conditional probabilities of all the models to the filtered and smoothed conditional probabilities of the benchmark models. The following conclusions can be drawn from the prediction accuracy comparisons which are partly reported in Tables 4.3.7a and 4.3.7b.

First, the smoothed conditional probabilities are normally closer to the NBER dates. In other words, their forecasting accuracy is statistically superior to that of the filtered regime probabilities.

Secondly, there is no evidence of better performance of the models with mixed-frequency data compared to the single-frequency models, no matter how many regimes are assumed in those models.

Finally, we did not find any strong statistical evidence of the structural break (4-regime) models performing the in-sample forecasting better than the models with no structural break.

The reason why neither adding the mixed-frequency nor including another independent state variable improves the prediction of the NBER business cycle chronology may be that there is no much room for such an improvement, since the two-regime model replicates the chronology sufficiently well.

4.3.3 Summary

An analysis of the five US macroeconomic time series shows that in the beginning of 1984 a structural break in the volatility of the growth rates of the series in question had taken place. There is no strong evidence of any break in the average growth rate, however.

In this section we considered the problem of declining volatility of the US economy. Six single-factor models with Markov switching were examined: two models without structural break with single (monthly) and mixed (monthly and quarterly) observation frequencies, two unrestricted structural break models with single and mixed frequencies, and two restricted structural break models with single and mixed frequencies.

In the restricted models a constraint is imposed on the transition probabilities matrix of the variance state variable forcing the low variance regime to be an absorbing state. The restriction results in smoother conditional regime probabilities but, according to likelihood ratio test, the constrained models are beaten by the unrestricted models — the difference between the log-likelihood function values of the restricted and unrestricted models being statistically insignificant.

The models were estimated using the US monthly and quarterly macroeconomic data covering the period 1959-2002. The shift in the residual variance of the composite economic indicator appears to have happened in the early 1984. This is compliant to the previous findings, e.g. McConnell and Perez Quiros (2000), Chauvet and Potter (2001).

A strong link between our models' recession probabilities and the NBER chronology is evident. As the formal forecasting accuracy tests show, the four-regime models do not contribute new information in forecasting the NBER dates but do allow detecting the secular structural break in the volatility of the US economy.

4.3.4 Appendix

Table 4.3.1a. The component series statistics in first subperiod

Series	Mean	St.dev.	CV	Min	Max
Monthly data 1959:2-1983:12					
EMP	0.189	0.283	1.50	-0.86	1.23
INC	0.293	0.430	1.47	-1.05	1.61
IIP	0.296	1.048	3.54	-4.25	6.00
SLS	0.286	1.128	3.94	-3.11	3.12
Quarterly data 1959:2-1983:4					
GDP	0.853	1.082	1.27	-2.06	3.78

Table 4.3.1b. The component series statistics in second subperiod

Series	Mean	St.dev.	CV	Min	Max
Monthly data 1984:1-2002:6					
EMP	0.158	0.148	0.94	-0.31	0.56
INC	0.246	0.370	1.50	-1.10	1.25
IIP	0.235	0.534	2.27	-1.34	2.06
SLS	0.254	0.945	3.72	-3.21	3.54
Quarterly data 1984:1-2002:2					
GDP	0.776	0.556	0.72	-0.83	2.16

Table 4.3.2. Testing significance of differences between the means and variances of two subsamples

Series	Mean Z	Mean p-value	Variance F	Variance p-value
EMP	1.620	0.053	3.66	0.0
INC	1.340	0.091	1.35	0.009
IIP	0.866	0.193	3.85	0.0
SLS	0.352	0.363	1.42	0.003
GDP	0.609	0.271	3.79	0.0

Table 4.3.3. Estimated parameters of the 2-regime models with single- and mixed-frequency data

Parameter	Single-frequency: -2547.1		Mixed-frequency: -3041.1	
	Estimated	St. error	Estimated	St. error
p_{11}^{μ}	0.98	0.007	0.979	0.008
$1 - p_{22}^{\mu}$	0.085	0.030	0.091	0.032
μ_1	0.300	0.041	0.121	0.019
μ_2	-1.150	0.099	-0.484	0.058
γ_{GDP}	–	–	1	–
γ_{EMP}	1	–	2.28	0.188
γ_{INC}	0.774	0.052	1.85	0.163
γ_{IIP}	0.880	0.055	2.16	0.170
γ_{SLS}	0.615	0.056	1.52	0.15
σ^2	0.389	0.040	0.073	0.012
σ_{GDP}^2	–	–	0.192	0.024
σ_{EMP}^2	0.265	0.036	0.315	0.037
σ_{INC}^2	0.558	0.040	0.549	0.040
σ_{IIP}^2	0.430	0.039	0.385	0.038
σ_{SLS}^2	0.721	0.049	0.695	0.047

Table 4.3.4. Estimated parameters of the unrestricted and restricted models with single-frequency data

Parameter	Unrestricted: -2470.7		Restricted: -2481.9	
	Estimated	St. error	Estimated	St. error
p_{11}^{μ}	0.959	0.013	0.977	0.008
p_{22}^{μ}	0.919	0.027	0.919	0.030
p_{11}^{σ}	0.937	0.022	0.997	0.003
p_{22}^{σ}	0.951	0.017	1	–
μ_{1high}	0.677	0.089	0.469	0.067
μ_{2high}	-1.460	0.131	-1.430	0.190
μ_{1low}	0.223	0.044	0.110	0.036
μ_{2low}	-0.405	0.075	-0.878	0.079
γ_{INC}	0.697	0.046	0.688	0.050
γ_{IIP}	0.827	0.048	0.803	0.054
γ_{SLS}	0.540	0.050	0.541	0.052
σ_{high}^2	0.661	0.086	0.633	0.074
σ_{low}^2	0.014	0.023	0.053	0.021
σ_{EMP}^2	0.194	0.030	0.178	0.039
σ_{INC}^2	0.607	0.041	0.610	0.043
σ_{IIP}^2	0.448	0.037	0.469	0.042
σ_{SLS}^2	0.764	0.050	0.758	0.050

Table 4.3.5. Estimated parameters of the unrestricted and restricted models with mixed-frequency data

Parameter	Unrestricted: -2966.3		Restricted: -2978.8	
	Estimated	St. error	Estimated	St. error
p_{11}^{μ}	0.957	0.014	0.975	0.009
p_{22}^{μ}	0.917	0.029	0.916	0.030
p_{11}^{σ}	0.932	0.025	0.997	0.003
p_{22}^{σ}	0.947	0.018	1	–
μ_{1high}	0.274	0.042	0.189	0.027
μ_{2high}	-0.536	0.061	-0.534	0.064
μ_{1low}	0.081	0.018	0.043	0.015
μ_{2low}	-0.148	0.037	-0.316	0.036
γ_{EMP}	2.50	0.199	2.47	0.201
γ_{INC}	1.85	0.165	1.82	0.165
γ_{IIP}	2.23	0.179	2.18	0.177
γ_{SLS}	1.46	0.155	1.48	0.153
σ_{high}^2	0.101	0.020	0.097	0.017
σ_{low}^2	0.001	0.004	0.008	0.003
σ_{GDP}^2	0.196	0.024	0.201	0.025
σ_{EMP}^2	0.247	0.034	0.247	0.034
σ_{INC}^2	0.588	0.040	0.587	0.040
σ_{IIP}^2	0.404	0.037	0.410	0.038
σ_{SLS}^2	0.742	0.049	0.729	0.048

Table 4.3.6. In-sample performance of the probabilistic structural break models: QPS of filtered and smoothed low-intercept regime probabilities

Model	Filtered	Smoothed
No structural break (two-regime) models		
Single-frequency (regime 2)	0.069	0.037
Mixed-frequency (regime 2)	0.063	0.031
Structural break (four-regime) models		
Single-frequency unrestricted (regime 2 + regime 4)	0.121	0.128
Mixed-frequency unrestricted (regime 2 + regime 4)	0.122	0.128
Single-frequency unrestricted (regime 2)	0.059	0.043
Mixed-frequency unrestricted (regime 2)	0.056	0.037
Single-frequency restricted (regime 2 + regime 4)	0.071	0.053
Mixed-frequency restricted (regime 2 + regime 4)	0.064	0.052

Note: recall that regime 2 in the two-regime (no break) model stands for the low intercept regime, while in the four-regime structural break model it stands for the low intercept - high variance regime. The sum of regime 2 and regime 4 probabilities corresponds in the latter model to the low intercept regime.

Table 4.3.7a. Comparing forecasting accuracy of filtered regime probabilities: pairwise DM-statistic

Model	Benchmark			
	SFNoB		MFNoB	
	filtered	smoothed	filtered	smoothed
SFNoB (regime 2)	–	-3.08**	-1.35*	-2.46**
MFNoB (regime 2)	1.35*	-3.19**	–	-2.70**
SFUnR (regime 2 + regime 4)	-2.90**	-3.90**	-2.97**	-3.53**
MFUnR (regime 2 + regime 4)	-3.01**	-4.09**	-3.17**	-3.72**
SFUnR (regime 2)	0.42	-1.10	0.16	-1.61*
MFUnR (regime 2)	0.61	-1.00	0.38	-1.60*
SFRes (regime 2 + regime 4)	-0.31	-3.15**	-0.92	-2.59**
MFRes (regime 2 + regime 4)	0.67	-3.74**	-0.27	-2.82**

Notation: SF = single frequency; MF = mixed frequency; NoB = no structural break; UnR = unrestricted structural break model; Res = restricted structural break model; * and ** — means that difference of accuracy is significant at 10% and 5%, respectively.

Table 4.3.7b. Comparing forecasting accuracy of smoothed regime probabilities: pairwise DM-statistic

Model	Benchmark			
	SFNoB		MFNoB	
	filtered	smoothed	filtered	smoothed
SFNoB (regime 2)	3.08**	–	3.19**	-0.89
MFNoB (regime 2)	2.46**	0.89	2.70**	–
SFUnR (regime 2 + regime 4)	-2.50**	-3.35**	-2.60**	-3.13**
MFUnR (regime 2 + regime 4)	-2.46**	-3.34**	-2.60**	-3.14**
SFUnR (regime 2)	1.06	-0.27	0.93	-0.66
MFUnR (regime 2)	1.38*	-0.01	1.31*	-0.41
SFRes (regime 2 + regime 4)	1.26	-1.74**	0.75	-1.47*
MFRes (regime 2 + regime 4)	1.29*	-1.88**	0.82	-1.58*

Notation: SF = single frequency; MF = mixed frequency; NoB = no structural break; UnR = unrestricted structural break model; Res = restricted structural break model; * and ** means that difference of accuracy is significant at 10% and 5%, respectively.

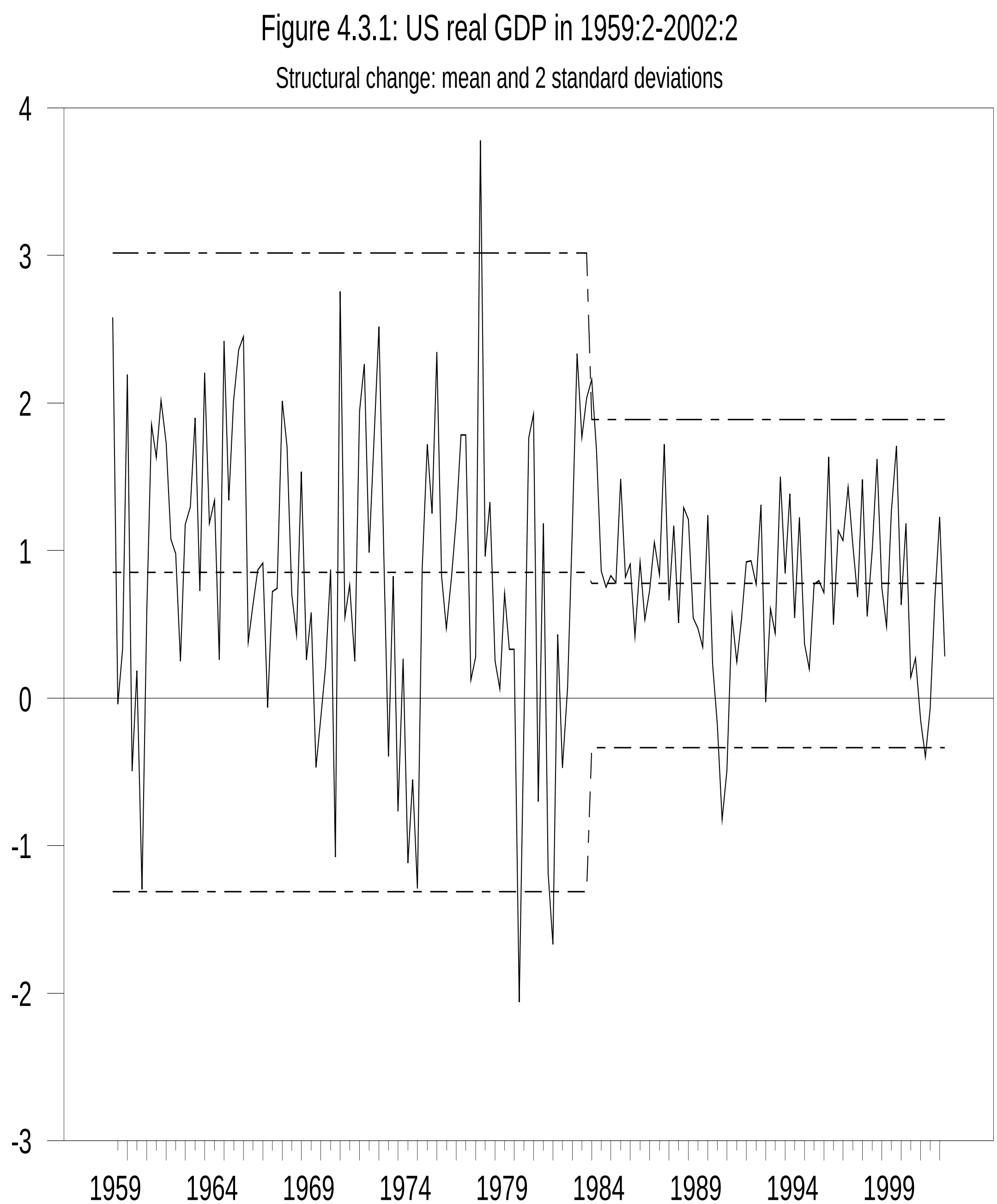


Figure 4.3.1: US real GDP in 1959:2-2002:2
Structural change: mean and 2 standard deviations
4
3
2
1
0
-1
-2
-3
1959
1964
1969
1974
1979
1984
1989
1994
1999

Figure 4.3.2: Component series of the US coincident indicator, 1959:2-2002:6

Structural change: mean and 2 standard deviations

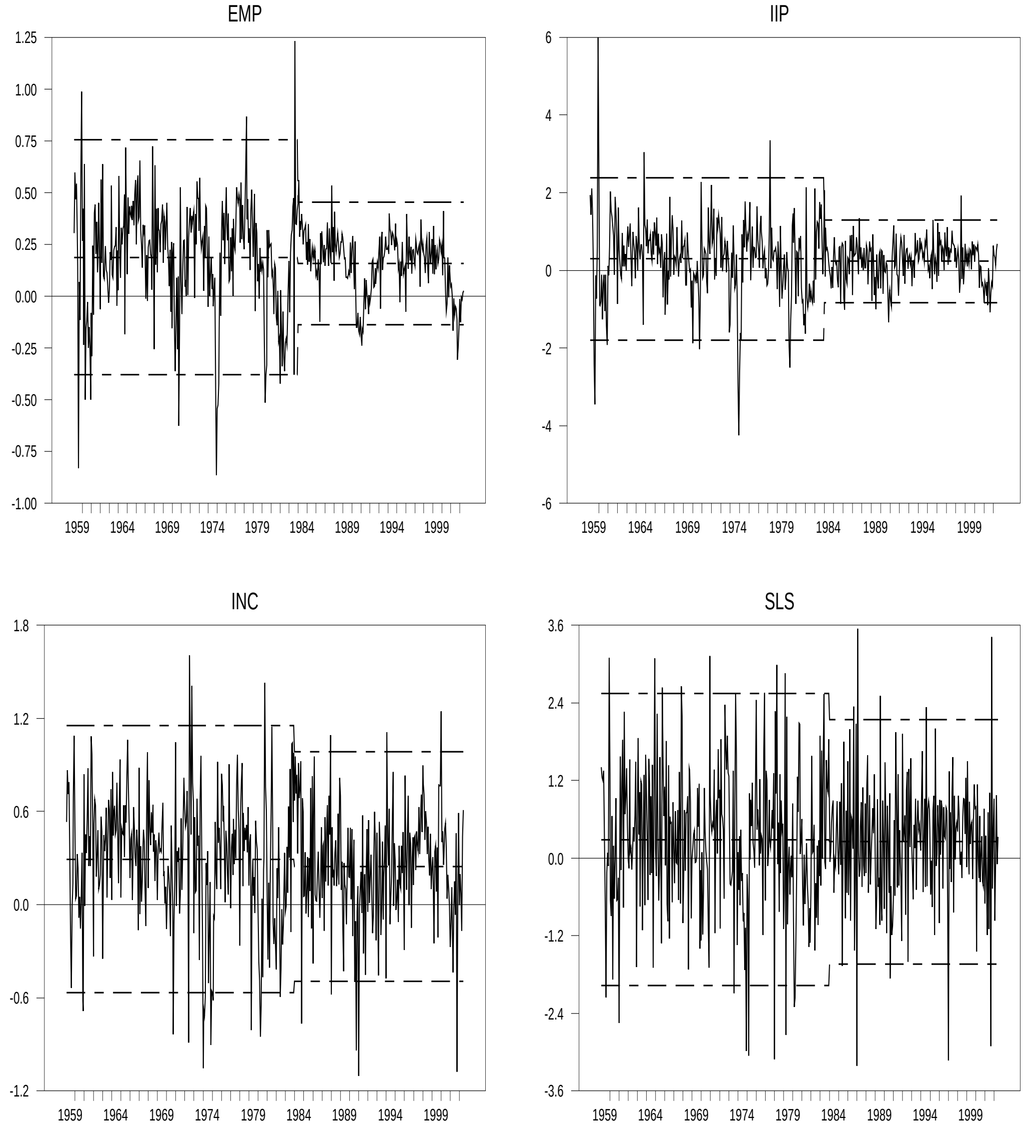


Figure 4.3.3: US coincident economic indicator vs. NBER dates

Models with & without structural break 1959:1-2002:6

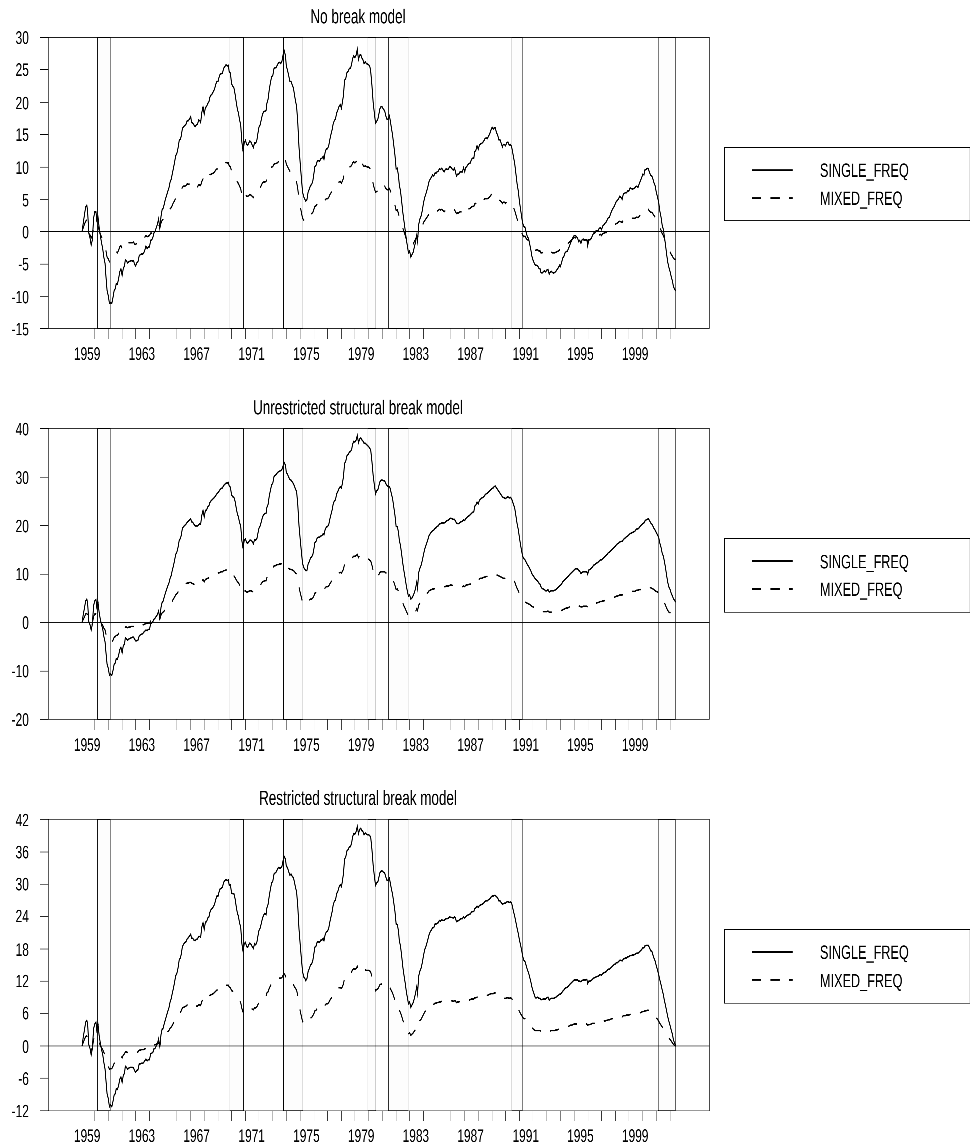


Figure 4.3.4: Low intercept regime probabilities vs. NBER dates

Benchmark (no break) model

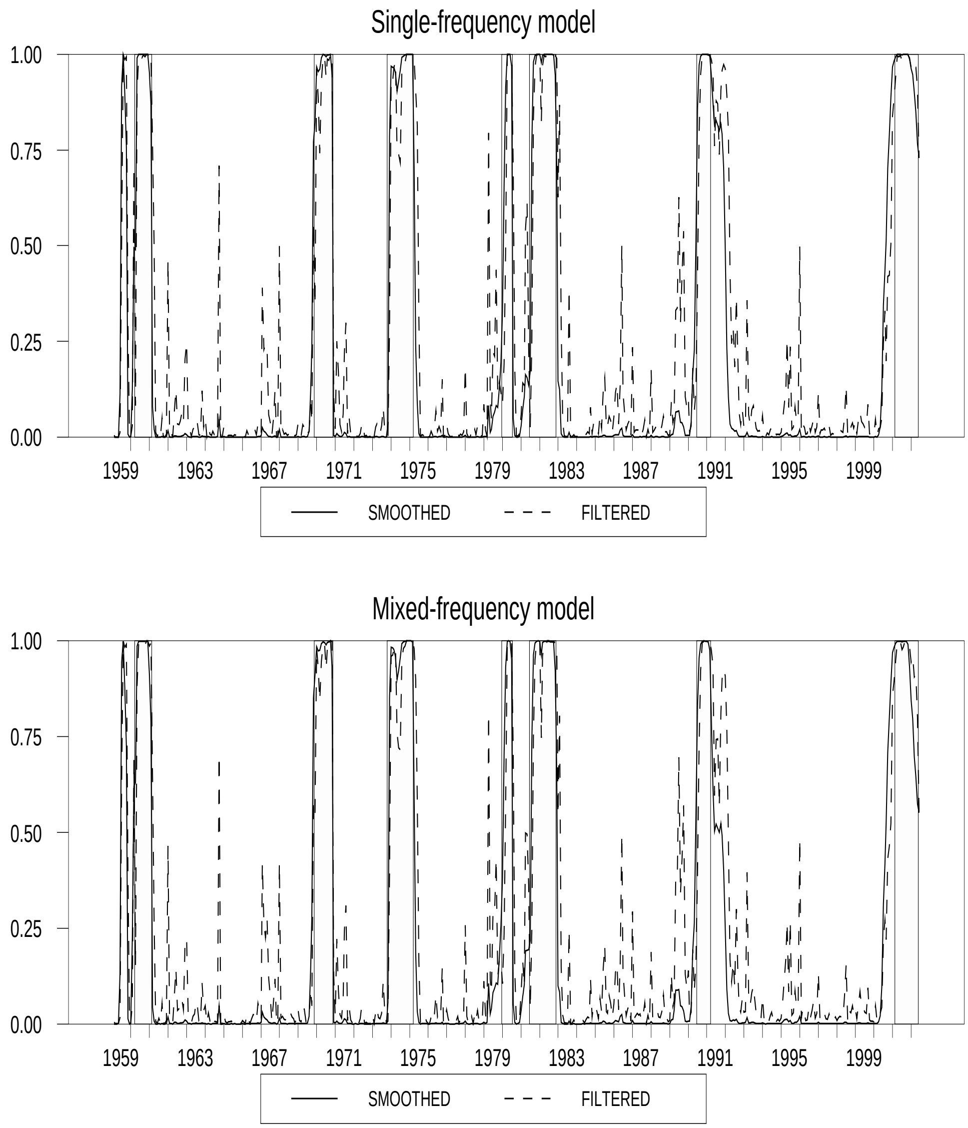
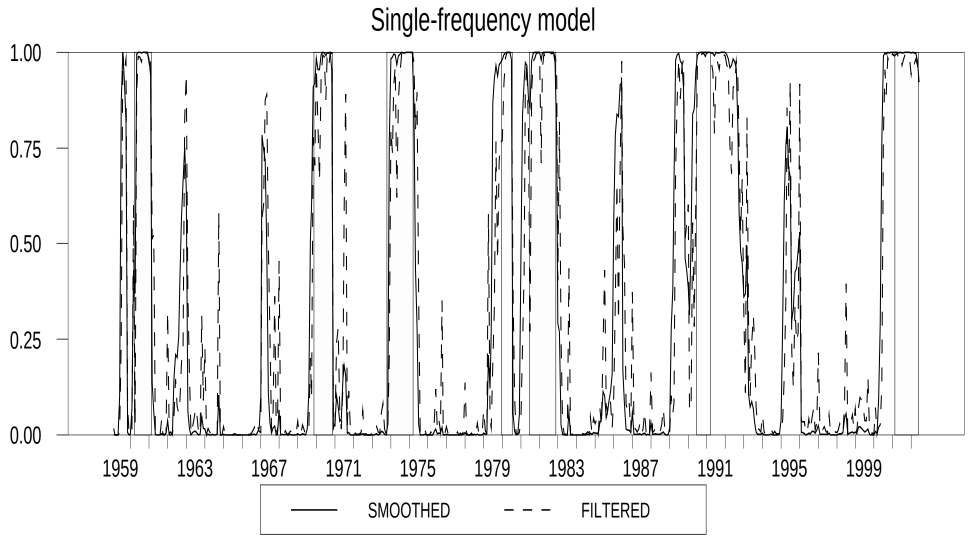


Figure 4.3.5: Low intercept regime probabilities vs. NBER dates

Unrestricted structural break model



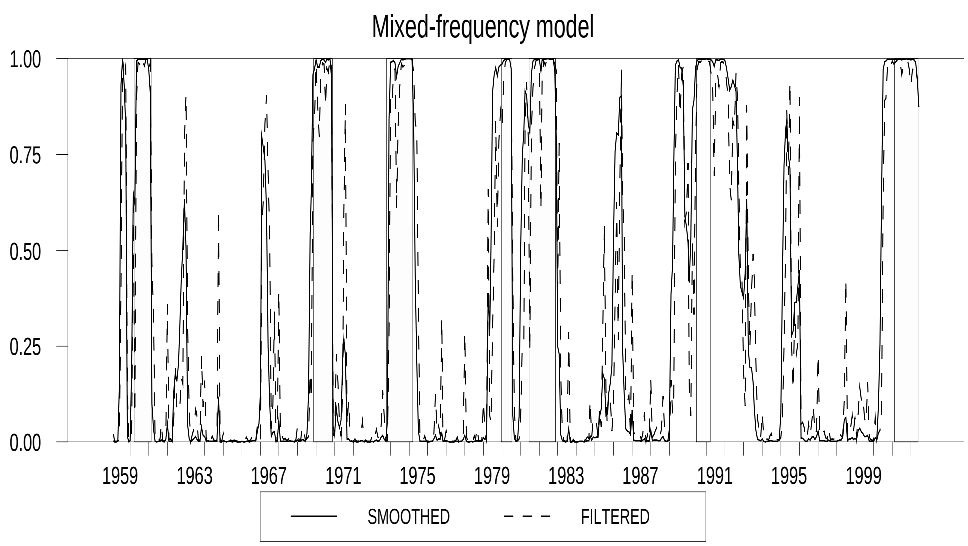


Figure 4.3.6: Low intercept - high variance regime probabilities vs. NBER dates

Unrestricted structural break model

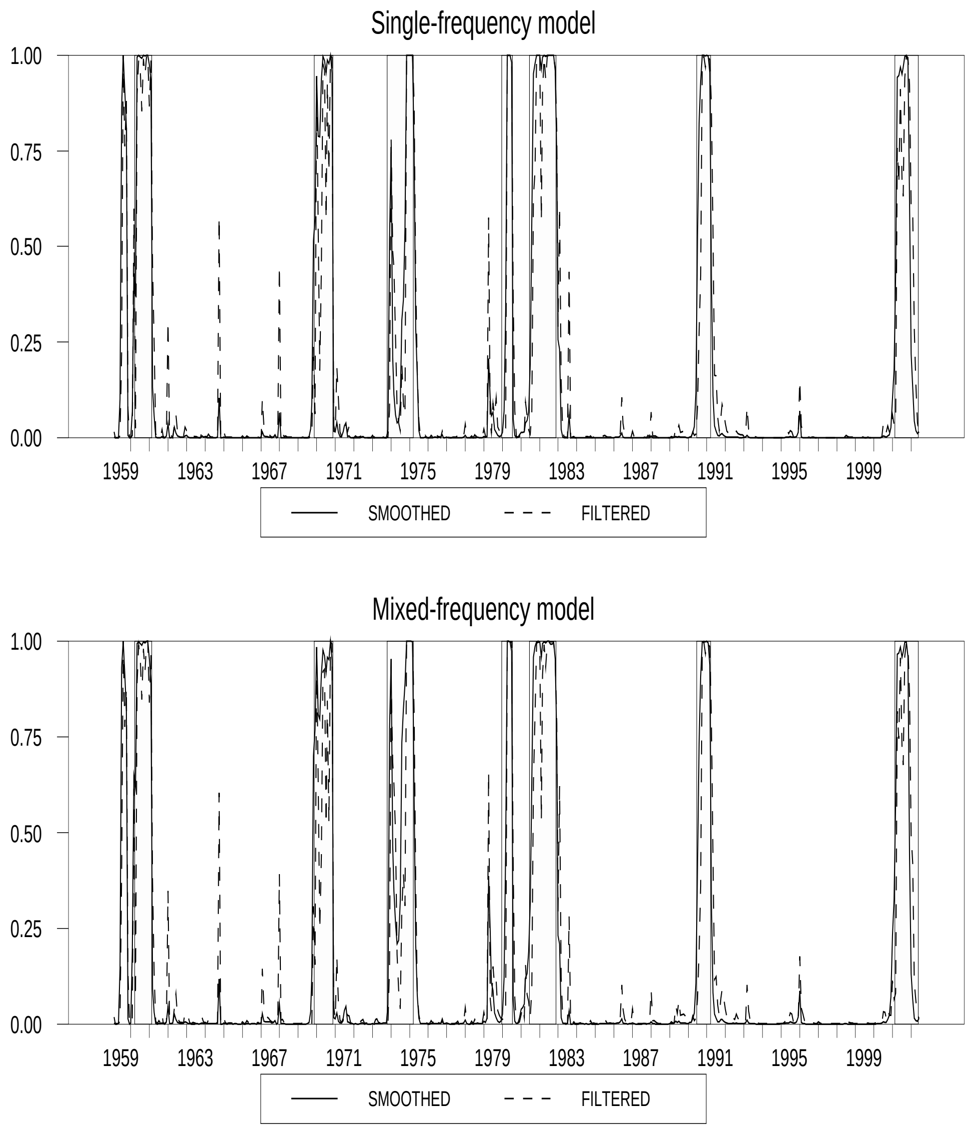


Figure 4.3.7: Low variance regime probabilities

Unrestricted structural break model

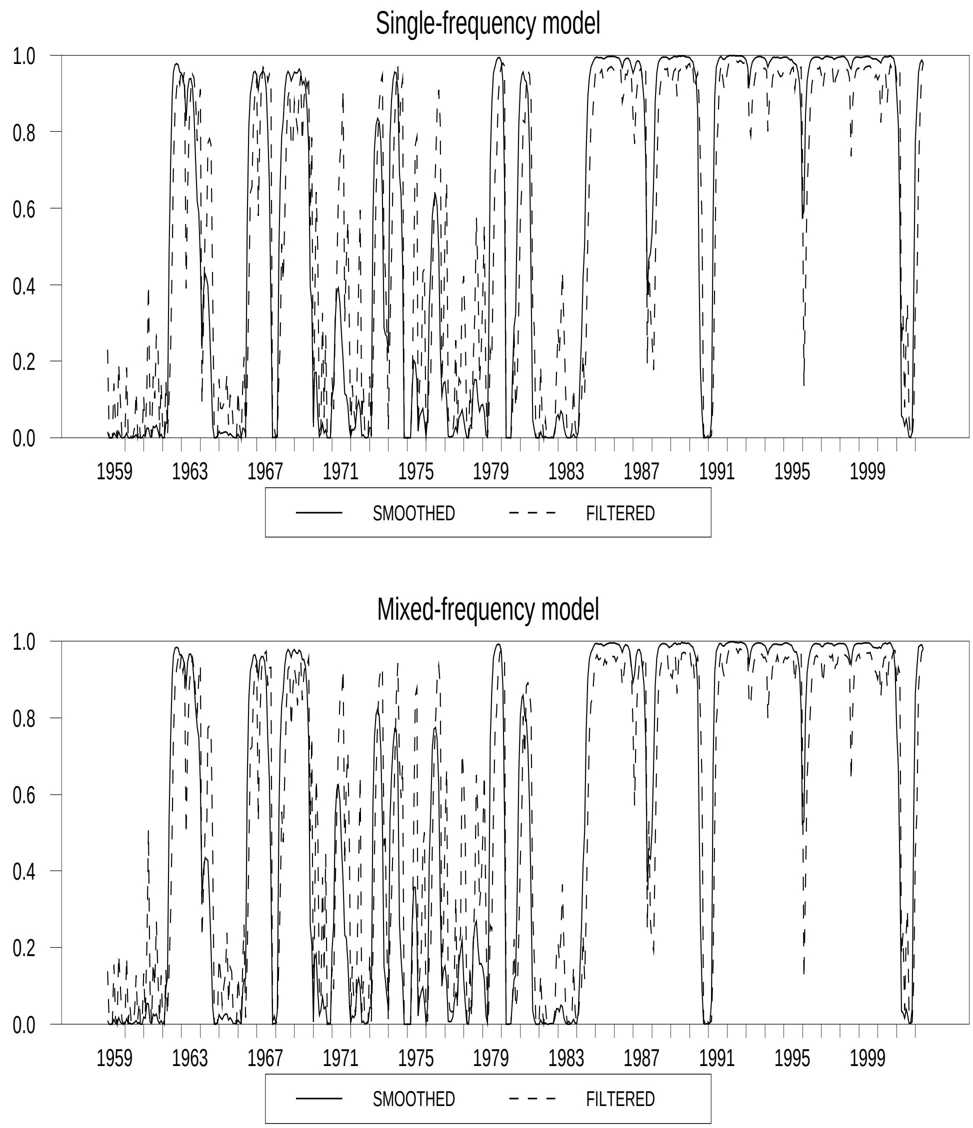


Figure 4.3.8: Low intercept regime probabilities vs. NBER dates

Restricted structural break models

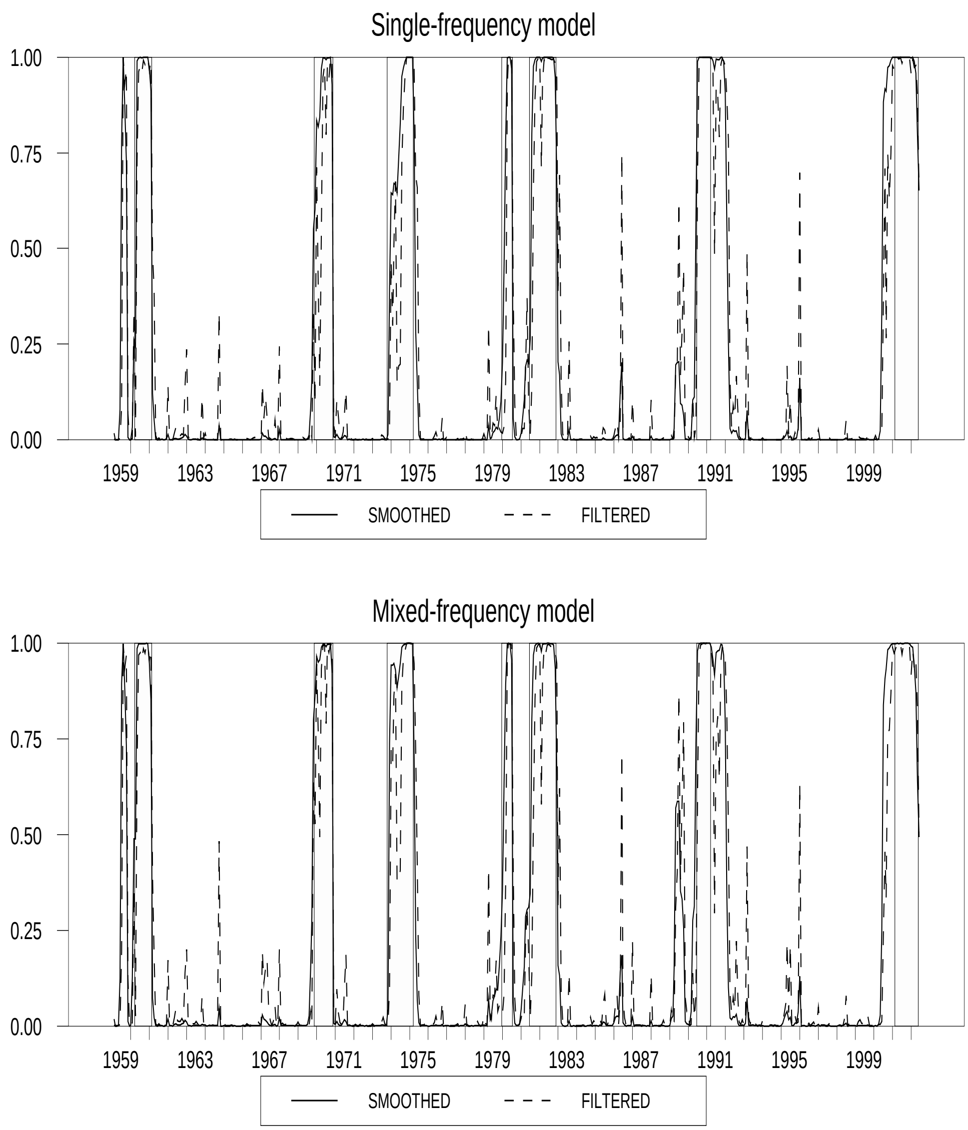
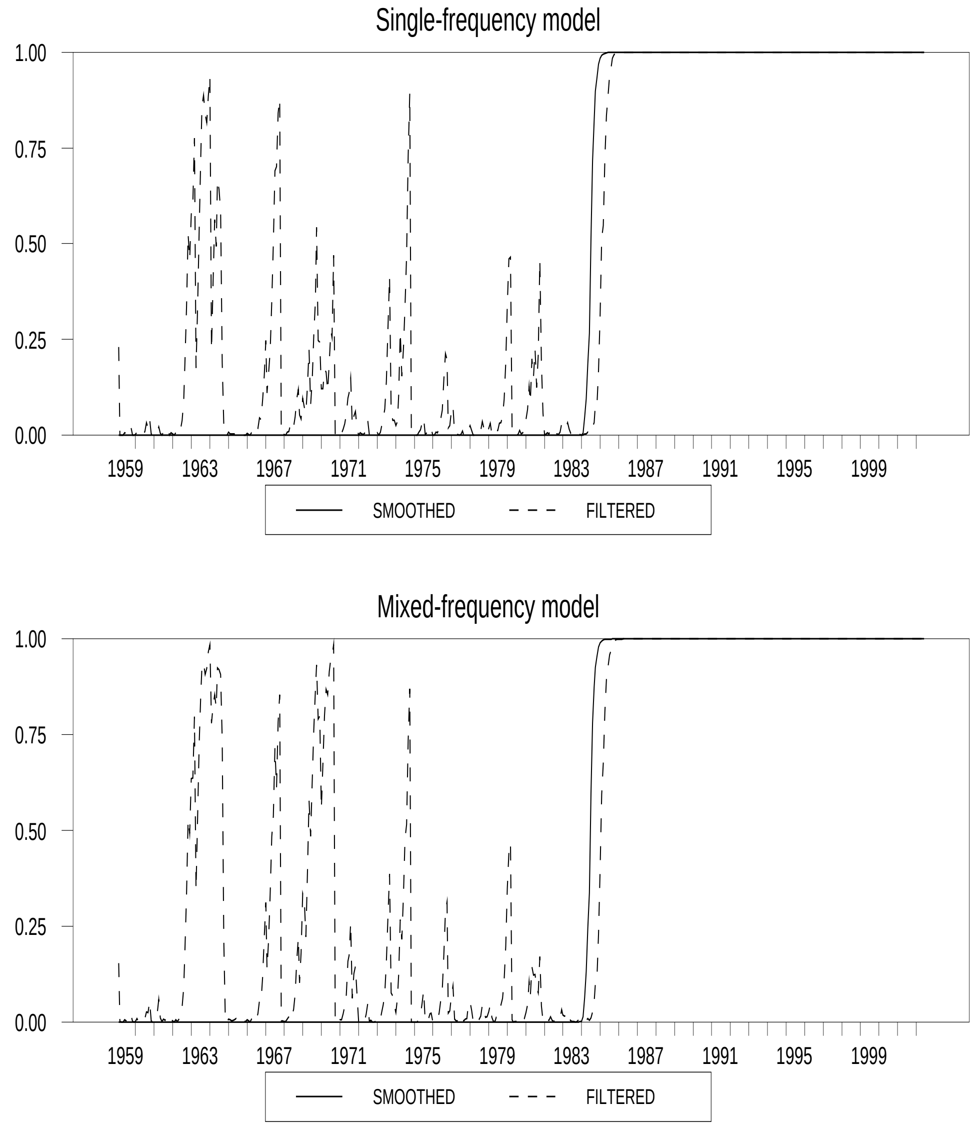


Figure 4.3.9: Low variance regime probabilities

Restricted structural break models



4.4 Conclusion

We conclude the analysis of the models examined in this chapter by comparing their performance in terms of the in-sample prediction of the NBER turning points. As in chapter 3 we use the quadratic probability score to evaluate the performance. The below table compares the models with deterministic and probabilistic structural break mechanisms as defined in sections 4.2 and 4.3, correspondingly. Remind that in the former case the structural break is captured by the deterministic dummies added to the idiosyncratic factors' intercepts and/or residual variances. In the latter case the structural break is taken into account by introducing second state variable responsible for the secular change in the common factor's volatility.

All the models in this table have both the common factor and the idiosyncratic factors specified as AR(0). The benchmark model is that with no break. It is endowed with the Markov-switching dynamics represented by a state variable with two regimes: low intercept (contraction) and high intercept (expansion).

In-sample performance: QPS 1959:1-2002:6

Model	Filtered	Smoothed
Deterministic structural break model		
No break (benchmark model)	0.050	0.037
Break in mean	0.047	0.029
Break in variance	0.060	0.051*
Break in mean and variance	0.056	0.044
Probabilistic structural break model		
Single-frequency unrestricted	0.121*	0.128*
Mixed-frequency unrestricted	0.121*	0.127*
Single-frequency restricted	0.071*	0.053*
Mixed-frequency restricted	0.064*	0.052*

The asterisks denote the QPS which are significantly — at 5significance level — different from the benchmark model's QPS. The significance level is determined using the Diebold-Mariano statistic. The null hypothesis of the Diebold-Mariano test states no difference between the predictive accuracy of the two models.

First conclusion that can be drawn from this table is that among the deterministic break models only model with break in intercept gives better estimates of the NBER turning points than the benchmark model. Among the probabilistic break models only the last one outperforms the benchmark model and only in terms of the filtered conditional probabilities. However, if we take into account the Diebold-Mariano statistics we have to conclude that there is no significant difference in the in-sample turning points forecast accuracy between the benchmark model and the best model with structural break (deterministic break in the intercepts of the specific factors).

Secondly, the smoothed conditional probabilities are normally closer to the NBER dates.

Thirdly, the models with deterministic structural break are generally outperforming those with stochastic break, especially in terms of the smoothed probabilities. The two models with restricted transition matrix perform better than the two models with deterministic dummies for the structural break in the variances of the specific factors.

Fourthly, the models with mixed-frequency series appear to predict better the NBER business cycle chronology than the single-frequency models. This may be due to the additional information contained in the quarterly real GDP data which have been used in these models. The gain is especially noticeable for the models with restricted transition probabilities matrix.

Chapter 5

Conclusion

Dynamic factor analysis is a very rich methodology which can be extended in many ways to get a closer approximation to complex economic reality. In particular, it is a good approach to the investigation of the business cycle, since, especially when augmented with Markov switching, it captures the basic features of this macroeconomic phenomenon: the comovements of different macroeconomic time series and their asymmetrical evolution over contraction and expansion phases of the business cycle.

In this thesis we have introduced several models aimed at increasing the forecasting ability of the common dynamic factor model and improving its efficiency. The models were analyzed using both the simulated time series with known data-generating process and the real macroeconomic variables characterizing the Post-World War II evolution of the US economy.

Firstly, two bifactor models, including common leading factor and common coincident factor, were examined. The models differ in the way the intertemporal relation between these two common factors is defined. The first model uses the Granger causality concept to link the coincident factor to the leading one and thus to transmit the cyclical fluctuations of the latter to the former. The second model assumes no Granger causality between the two factors, working through the transition probabilities matrix which permits considering various hypotheses concerning the intertemporal relation in question: mutually independent common factors, common factors having exactly the same cyclical dynamics, and the common factors whose recession probabilities are correlated with some lag.

Secondly, we have elaborated a number of tools to allow a better use of the information available to the researcher. Here we have addressed two issues: missing data and structural breaks. In principle, both can be unified under the heading of lacking information, since in both cases the usable dataset is reduced. One of these tools is the model combining the use of mixed-frequency data and the Markov-switching dynamics, which allows dealing with data having different (for example, monthly and quarterly) observation frequencies. Another tool is implemented in two alternative models incorporating the structural breaks. The first model introduces deterministic structural break dummies in the intercepts and residual variances of each of the observed variables. The breakpoints may be specific for each observed time series. Their timing is determined from the structural break analysis preceding the model estimation. The second model deals with a single one-time structural break hitting the residual variance of the common unobserved factor. This requires additional state variable in the regime-switching model tracing uniquely the variance shifts. In this case the breakpoint timing is estimated endogenously based on the conditional regime probabilities and expected duration of high volatility regime.

In fact, the aforementioned models can be combined depending on the conditions of a particular problem, the main objective being the efficient utilization of all available information to obtain the reliable results which can be used in the macroeconomic analysis, forecasting, and decision-making.

There is still a lot of room for making useful extensions and refinements of the dynamic factor models. Firstly, the nonlinear dynamics other than regime switching can be employed. One example might be the threshold autoregression (TAR) or smooth transition autoregression (STAR) applied to the unobserved common factor. Secondly, the hierarchical common dynamic factor models as that of Kose et al. (1999), including the common factors of many levels, can be endowed with nonlinear dynamics which may be allowed to be different for different common factors. In this case, however, the MLE would be hardly applicable. Instead of MLE one would use some of the MCMC estimation techniques.

Bibliography

[1] Altissimo, F., Bassanetti, A., Cristadoro, R., Forni, M., Hallin, M., Lippi, M., Reichlin, L. and Veronese, G. (2001) "EuroCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle" *CEPR (Centre for Economic Policy Research) Discussion Paper* no. 3108.

[2] Amemiya T. (1981) "Qualitative Response Models: A Survey" *Journal of Economic Literature* **19**,1483-1536.

[3] Andrews D.W.K. (1993) "Test for Parameter Instability and Structural Change with Unknown Change Point" *Econometrica* **61**, 821-856.

[4] Angelini E., Henry J., Mestre R. (2001) "Diffusion Index-Based Inflation Forecasts for the Euro Area" *European Central Bank Working Paper* 61.

[5] Artis M.J., Bladen-Hovell R.C., Osborn D.R., Smith G. and Zhang W. (1995) "Predicting Turning Points in the UK Inflation Cycle" *The Economic Journal* **105**, 1145-64.

[6] Artis M.J., Kontolemis Z.G., Osborn D. R. (1997) "Business Cycles for G7 and European Countries" *Journal of Business* **70**, 249-279.

[7] Artis M.Z., Krolzig H.-M., Toro J. (1999) "The European Business Cycle" European University Institute working paper ECO 99.

[8] Barro R.J., Grilli V. (1994) *European Macroeconomics*. London.

[9] Bodart V., Candelon B. (2000) "Appréhender la conjoncture à l'aide de la méthode de Stock-Watson: une application à l'économie belge" *Economie et Prevision* **146.**

[10] Bollen K.A. (1989) *Structural Equations with Latent Variables.* John Wiley and Sons.

[11] Burns A.F., Mitchell W.C. (1946) *Measuring Business Cycles*. New York: National Bureau of Economic Research.

[12] Carter C.K., Kohn R. (1996) "Markov Chain Monte Carlo in Conditionally Gaussian State Space Models" *Biometrika* **83**, 589-601.

[13] Chauvet M. (1998) "An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching" *International Economic Review* **39**, 969-96.

[14] Chauvet M. (1998/9) "Stock Market Fluctuations and the Business Cycle" *Journal of Economic and Social Measurement* **25**, 235-258.

[15] Chauvet M., Potter S. (2000) "Coincident and Leading Indicators of the Stock Market", *Journal of Empirical Finance* **7**, 87-111.

[16] Chauvet M., Potter S. (2001) "Recent Changes in the US Business Cycle", *The Manchester School* **69**, 481-508.

[17] Chin D., Geweke J., Miller P. (2000) "Predicting Turning Points" Federal Reserve Bank of Minneapolis Research Department Staff Report *267*.

[18] Devore J.L. (1987) *Probability and Statistics for Engineering and Science*. Monterey: Brooks / Cole Publishing Company.

[19] Diebold F.X., Chen C. (1996) "Testing Structural Stability with Endogenous Breakpoint. A Size Comparison of Analytic and Bootstrap Procedures" *Journal of Econometrics* **70**, 221-241.

[20] Diebold F.X., Lee J.-H., Weinbach G.C. (1994) "Regime Switching with Time-Varying Transition Probabilities", in Hargreaves C.P. (ed.) *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press.

[21] Diebold F.X., Mariano R.S. (1995) "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics* **13**, 253-263.

[22] Diebold F.X., Rudebusch G.D. (1989) "Scoring the Leading Indicators" *Journal of Business* **62**, 369-402.

[23] Diebold F.X., Rudebusch G.D. (1991) "Forecasting Output with the Composite Leading Index: A Real-Time Analysis" *Journal of American Statistical Association* **86**, 603-10.

[24] Diebold F.X., Rudebusch G.D. (1996) "Measuring Business Cycles: A Modern Perspective" *The Review of Economics and Statistics* **78**, 67-77.

[25] Estrella A., Mishkin F.S. (1998) "Predicting the US Recessions: Financial Variables as Leading Indicators" *The Review of Economics and Statistics* **80**, 45-61.

[26] Filardo A.J. (1994) "Business Cycle Phases and Their Transitional Dynamics" *Journal of Business and Economic Statistics* **12**, 299-308.

[27] Filardo A.J. (1999) "How Reliable Are Recession Prediction Models" *Federal Reserve Bank of Kansas City Economic Review* **84**, 35-55.

[28] Filardo A.J., Gordon S.F. (1998) "Business Cycle Durations" *Journal of Econometrics* **85**, 99-123.

[29] Forni M., Hallin M., Lippi M., Reichlin L. (2001) "Coincident and Leading Indicators for the EURO Area" *The Economic Journal* **111**, C62-C85.

[30] Garcia-Ferrer A., Queralt R.A. (1998) "Using Long-, Medium, and Short-Term Trends to Forecast Turning Points in the Business Cycle: Some International Evidence" *Studies in Nonlinear Dynamics and Econometrics*, Vol. 3, Number 2.

[31] Grégoir S., Lenglart F. (2000) "Measuring the Probability of a Business Cycle Turning Point by Using a Multivariate Hidden Markov Model" *Journal of Forecasting* **19**, 81-102.

[32] Gregory A.W., Head A.C. (1999) "Common and Country-Specific Fluctuations in Productivity, Investment, and the Current Account" *Journal of Monetary Economics* **44**, 423-451.

[33] Hall S.G., Psaradakis Z., Solà M. (1997) "Cointegration and Changes in Regime: The Japanese Consumption Function" *Journal of Applied Econometrics* **12**, 151-68.

[34] Hamilton J.D. (1989) "A New Approach to the Economic Analysis of Non-Stationary Time Series and the Business Cycle" *Econometrica* **57**, 357-84.

[35] Hamilton J.D. (1990) "Analysis of Time Series Subject to Changes in Regime" *Journal of Econometrics* **45**, 39-70.

[36] Hamilton J.D. (1994) *Time Series Analysis*. New Jersey: Princeton University Press.

[37] Harvey D., Leybourne S., Newbold P. (1997) "Testing the Equality of Prediction Mean Squared Errors" *International Journal of Forecasting* **13**. 281-291.

[38] Hylleberg S., Mizon G.E. (1989) "Cointegration and Error Correction Mechanisms" *The Economic Journal (Supplement: Conference Papers)* **99**, 113-25.

[39] Johansen S. (1991) "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models" *Econometrica* **59**, 1551-80.

[40] Kim C.-J. (1994) "Dynamic Linear Models with Markov-Switching" *Journal of Econometrics* **60**, 1-22.

[41] Kim C.-J., Nelson C.R. (1998) "Business Cycle Turning Points, a New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching" *The Review of Economics and Statistics* **80**, 188-201.

[42] Kim C.-J., Nelson C.R. (1999a) "Has the US Economy Become More Stable? A Bayesian Approach Based on Markov-Switching Model of the Business Cycle" *The Review of Economics and Statistics* **81**, 608-16.

[43] Kim C.-J., Nelson C.R. (1999b) "A Bayesian Approach to Testing for Markov Switching in Univariate and Dynamic Factor Model" *International Economic Review*. Forthcoming.

[44] Kim C.-J., Nelson C.R. (1999c) *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge and London: MIT Press.

[45] Kim C.-J., Piger J. (2002) "Common Stochastic Trends, Common Cycles, and Asymmetry in Economic Fluctuations", *Journal of Monetary Economics*, **49**, 1189-1211.

[46] Kim M.-J., Yoo J.-S. (1995) "New Index of Coincident Indicators: A Multivariate Markov Switching Factor Model Approach" *Journal of Monetary Economics* **36**, 607-30.

[47] Kling J.L. (1987) "Predicting the Turning Points of Business and Economic Time Series" *Journal of Business* **60**, 201-38.

[48] Kontolemis Z.G. (2001) "Analysis of the US Business Cycle with a Vector-Markov-Switching Model" *Journal of Forecasting* **20**, 47-61.

[49] Kose M.A., Otrok C., Whiteman C.H. (1999) "International Business Cycles: World, Region, and Country-Specific Factors" University of Virginia.

[50] Krane S., Wascher W. (1999) "The Cyclical Sensitivity of Seasonality in US Unemployment" *Journal of Monetary Economics* **44**, 523-553.

[51] Krolzig H.-M. (1996) "Statistical Analysis of Cointegrated VAR Processes with Markovian Regime Shifts" Humboldt Universität zu Berlin SFB 373 Discussion Paper *25/1996*.

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

[53] Layton A.P., Katsuura M. (2001) "Comparison of Regime Switching, Probit and Logit Models in Dating and Forecasting US Business Cycles" *International Journal of Forecasting* **17**, 403-417.

[54] Li D.T., Dorfman J.H. (1996) "Predicting Turning Points through the Integration of Multiple Models" *Journal of Business & Economic Statistics* **14,** 421-28.

[55] Luginbuhl R., de Vos A. (1999) "Bayesian Analysis of An Unobserved-Component Time Series Model of GDP with Markov-Switching and Time Varying Growths" *Journal of Business and Economic Statistics* **17**, 456-65.

[56] Marcellino M., Stock J., Watson M. (2000) "Macroeconomic Forecasting in the Euro-Area: Country Specific versus Area-Wide Information" *Mimeo.*

[57] Mariano R.S., Murasawa Y. (2003) "A New Coincident Index of Business Cycle Based on Monthly and Quarterly Series" *Journal of Applied Econometrics*, forthcoming.

[58] McConnell M.M., Perez Quiros G. (2000) "Output Fluctuations in the United States: What Has Changed Since the Early 1980s?" *The American Economic Review* **90**, 1464-1476.

[59] Otrok C., Whiteman C.H. (1998) "Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa" *International Economic Review* **39**, 997-1014.

[60] Phillips K.L. (1991) "A Two-Country Model of Stochastic Output with Changes in Regime" *Journal of International Economics* **31**, 121-42.

[61] Pindyck R.S., Rubinfeld D.L. (1991) *Econometric Models and Economic Forecasts.* McGraw-Hill.

[62] Samuelson P.A., Nordhaus W.D. (1998) *Economics.*

[63] Shephard N., Pitt M.K. (1997) "Likelihood analysis of Non-Gaussian Measurement Time Series" *Biometrika* **84**, 653-67.

[64] Simpson P.W., Osborn D.R., Sensier M. (2000) "Modelling Business Cycle Movements in the UK Economy" University of Manchester School of Economic Studies discussion paper Series 9908.

[65] Smith A.F.M., Roberts G.O. (1993) "Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods" *Journal of Royal Statistical Society Series B* **55**, 3-23.

[66] Stock J.H., Watson M.W. (1988) "Testing for Common Trends" *Journal of the American Statistical Association* **83**, 1097-1107.

[67] Stock J.H., Watson M.W. (1989) "New Indexes of Coincident and Leading Economic Indicators" in Blanchard O.-J., Fischer S. (eds.) *NBER Macroeconomics Annual.* Cambridge, Mass. and London: MIT Press.

[68] Stock J.H., Watson M.W. (1991) "A Probability Model of the Coincident Economic Indicators" in Moore G., Lahiri K. (eds.) *The Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge: Cambridge University Press, 63-90.

[69] Stock J.H., Watson M.W. (1993) "A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience" in Stock J.H, Watson M.W. (eds.) *Business Cycles, Indicators and Forecasting*. University of Chicago Press for the NBER.

[70] Stock J.H., Watson M.W. (1998) "Diffusion Indexes" NBER Working Paper 6702.

[71] Swets J.A., Pickett R.M. (1982) *Evaluation of Diagnostic Systems Methods from Signal Detection Theory*. New York Academic Press.

[72] Watson M.W. (2000) "Macroeconomic Forecasting Using Many Predictors", *Journal of Business and Economic Statistics*, forthcoming.

[73] Wecker W.E. (1979) "Predicting the Turning Points of a Time Series" *Journal of Business* **52**, 35-50.

[74] Zellner A. (1971) *An Introduction to Bayesian Inference in Econometrics*. John Wiley and Sons.

[75] Zellner A., Tobias J., Ryu H. (1999) "BMOM Analysis of Time Series Models with an Application to Forecasting Turning Points in Output Growth Rates" *Estadistica* (forthcoming).