Dealing with Structural Changes in the Common Dynamic Factor Model: Deterministic Mechanism

Konstantin A. Kholodilin*

Abstract

Composite economic indicator is a very useful tool designed to trace and predict the business cycle conditions. In this paper we study possible extensions of this approach intended to cope with the potential data problems caused by various structural breaks affecting both level and volatility of the component series. The structural shifts are introduced in the composite economic indicator model via deterministic dummies capturing breaks in the observed variables' intercepts and in the residual variances of the specific factors. As an illustration the Post-World War II US monthly macroeconomic series are utilized for which different specifications of the single-factor linear and regime-switching model are evaluated.

Keywords: composite economic indicator, Markov switching, structural break, turning points, NBER dating

JEL codes: C4, C5, E3

*The author gratefully acknowledges the financial support provided by the Belgian French Community's program Action de recherches concertée 99/04-235 (IRES, Université catholique de Louvain).

1 Introduction

The composite economic indicators (both leading and coincident) can be very useful in the political and business decision-making. It can supply more timely, accurate, and complete information than some conventional indicators like real GDP (too low frequency of observation) or index of industrial production (too narrow scope).

One of the approaches to construction of this composite economic indicator is the dynamic factor model advocated in the linear context by Stock and Watson (1988, 1989, 1992) and in the Markov-switching context by Kim (1994), Kim and Nelson (1999) and Chauvet (1998). Some recent examples of applying the dynamic factor model with Markov-switching dynamics include Chauvet (2001), Fukuda and Onodera (2001).

However, quite often the practical realisation 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 changes in the statistical methodology to the natural shifts in the behaviour of economic variables.

In this paper we examine a way of dealing with the structural breaks in the observed time series using the dynamic factor analysis with linear and regime-switching dynamics. The models considered here use deterministic dummies to capture the structural breaks in different model parameters with unknown break-points different for each time series. These models' specification builds on the paper of 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 of Chauvet and Potter (2001), who make the mean and the autoregressive coefficients of the common factor change as a function of the structural break.

The remainder of the paper is organised as follows. In section 2 we set up a common dynamic factor model with deterministic structural break(s) in the observed time series. Section 3 analyses a real example — common dynamic models (both linear and with Markov switching) of the U.S. composite economic indicator with deterministic dummies capturing the structural breaks in the means and variances of the observed series and idiosyncratic components. Section 4 concludes the paper.

2 Model

We deal with a set of observed time series, whose co-movements are explained by one or several common factors which may interact in a complex temporal and/or spatial way. In contrast, the idiosyncratic dynamics of each time series in particular are captured by one specific factor per each observed time series. The observed component series are subject to a one-time structural breaks which may affect both their level and volatility. Hence the model in its general form can be written as:

$$\Delta y_t = [(\mathbf{I}_n - I_t)\delta_1 + I_t\delta_2] + [(\mathbf{I}_n - I_t)\Gamma_1 + I_t\Gamma_2]\,\Delta f_t + u_t \tag{1}$$

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 unobserved common factors in the first differences; u_t is the $n \times 1$ vector of the unobserved 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 break-point indicator function, I_t, can be written as a diagonal matrix whose diagonal elements are the individual indicator functions:

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

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 took place.

The dynamics of the latent common factors can be described in terms of a vector-autoregression model:

$$\Delta f_t = \left[(\mathbf{I}_k - I_t^f)\nu_1 + I_t^f \nu_2\right] + \left[(\mathbf{I}_k - I_t^f)\Phi_1(L) + I_t^f \Phi_2(L)\right] \Delta f_{t-1} + \varepsilon_t \tag{2}$$

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 autoregressive (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:

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

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

$$u_t = \Psi(L)u_{t-1} + \eta_t \tag{3}$$

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:

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

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

Thus, in general the means 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 simplicity that there is only one common factor. 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 changes in the accounting methodology which lead to the sudden shifts in the time series levels.

We estimate the model using the maximum likelihood method. To construct the likelihood function, the model is expressed in a state-space form:

$$\Delta y_t = [\delta_1(1 - I_t) + \delta_2 I_t] + A\beta_t \tag{4}$$

$$\beta_t = \alpha + C\beta_{t-1} + v_t \tag{5}$$

where $\beta_t = (f_t|u_t)'$ is the state vector containing vector of common factor and the vector 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; 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 $g = \max\{g_1, ..., g_n\}$.

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

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

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 introduction of the regime-switching dynamics. This would allow taking into account the asymmetries existing between different phases of business cycle which, along with the co-movements of macroeconomic variables, according to Diebold and Rudebusch (1996), are the two key characteristics of business cycle. We introduce the Markov-switching dynamics through the intercept of the common factor:

$$\Delta f_t = \mu(s_t) + \Phi(L)\Delta f_{t-1} + \varepsilon_t \tag{6}$$

where $\mu(s_t)$ is the regime-dependent intercept (low in contractions and high in expansions); s_t is the state variable following first-order Markov chain process. In this paper we consider a two-regime case: one regime for expansions and another one — for recessions. To save space we do not enter further into details and simply refer our readers to the comprehensive discussion of common dynamic model with Markov switching in Kim and Nelson (1999).

In what follows the above dynamic factor model will be denoted as CF(p,q), where p is the autoregressive order of the common factor; q is the autoregressive order of the specific factors. In particular, the models with Markov-switching will be denoted as CF-MS(p,q).

3 Real example

The U.S. monthly coincident time series covering 1959:1-2002:6 are analysed — see Table 1 of Appendix. In fact, the component series we use to build our common coincident factor are those utilised by the Conference Board (USA) to construct its composite coincident index.

The first question to answer is when the structural breaks took place. To address 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:

$$\Delta y_{it} = \mu_i + \phi_i \Delta y_{it-1} + \xi_{it} \tag{7}$$

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

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

$$\sqrt{\frac{\pi}{2}} \left| \xi_{it} \right| = \alpha_{1i} D_{i1t} + \alpha_{2i} D_{i2t} + \omega_{it} \tag{8}$$

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 break-point τ_i and for each location compute the corresponding Wald statistic. The point where Wald statistic achieves its "supremum" is taken to be *the* break-point. 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 break-point in the variance of the $i-th$ time series we apply the supremum Wald statistic as Andrews (1993) proposes:

$$SupW_i = \max_{\pi} T \left[\frac{\omega'_{iR}\omega_{iR} - \omega'_{iU}\omega_{iU}}{\omega'_{iU}\omega_{iU}} \right] \tag{9}$$

where $\pi = \tau/T$; R stands for "restricted", while U means "unrestricted". In other words, for each time series we are looking for the point of time where the estimated Wald statistic attains its supremum.

The test critical values were computed using bootstrapping procedure as described in Diebold and Chen (1996). We implemented the procedure as follows:

1. Estimate model in (7) and save the residuals.

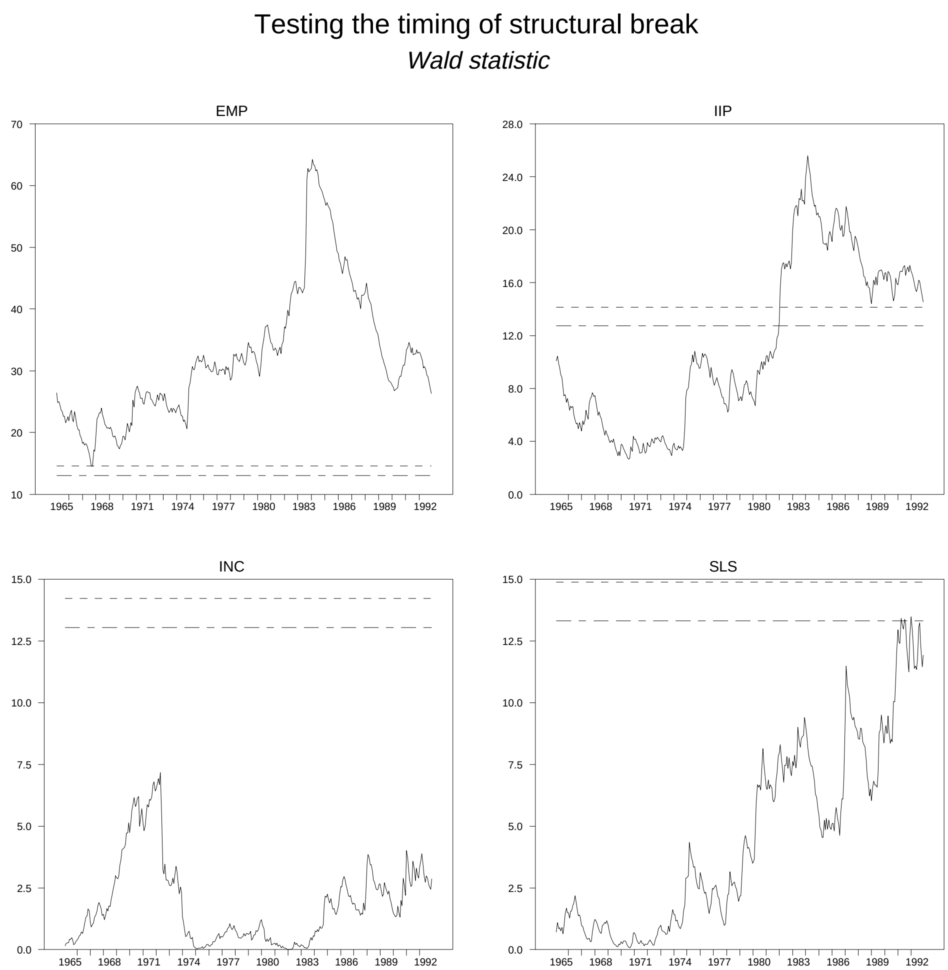
2. Estimate restricted analog of equation (8), that is, an equation where, instead of two structural break dummies, D_{i1t} and D_{i2t}, we put a constant term.

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 (8) and compute the statistic value using equation (9). Save this statistic and repeat steps 3-4 large enough number of times.

We undertook the bootstrap procedure outlined above with1000 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 sub-sample 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 1 shows the estimated Wald statistics together with corresponding bootstrap critical values at 10% and 5% significance levels (inferior and superior horizontal lines, respectively) for each of the time series. 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 3. 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, though the break-points are spread across the sample. It is worthwhile to notice, however, that two out of these four time series had the structural break in variance in the early 1984, namely EMP and IIP.



It is worth to check the intercept of equation 7 for a structural break, because the structural shifts in the residual variances may 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:

$$\Delta y_{it} = \beta_{1i} D_{i1} + \beta_{2i} D_{i2} + \phi_i \Delta y_{it-1} + \xi_{it} \tag{10}$$

Restricted regression:

$$\Delta y_{it} = \beta_i D_i + \phi_i \Delta y_{it-1} + \xi_{it} \tag{11}$$

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 break-points identified in Table 2. However, the variable subscript is still there meaning that each variable has its specific break-point.

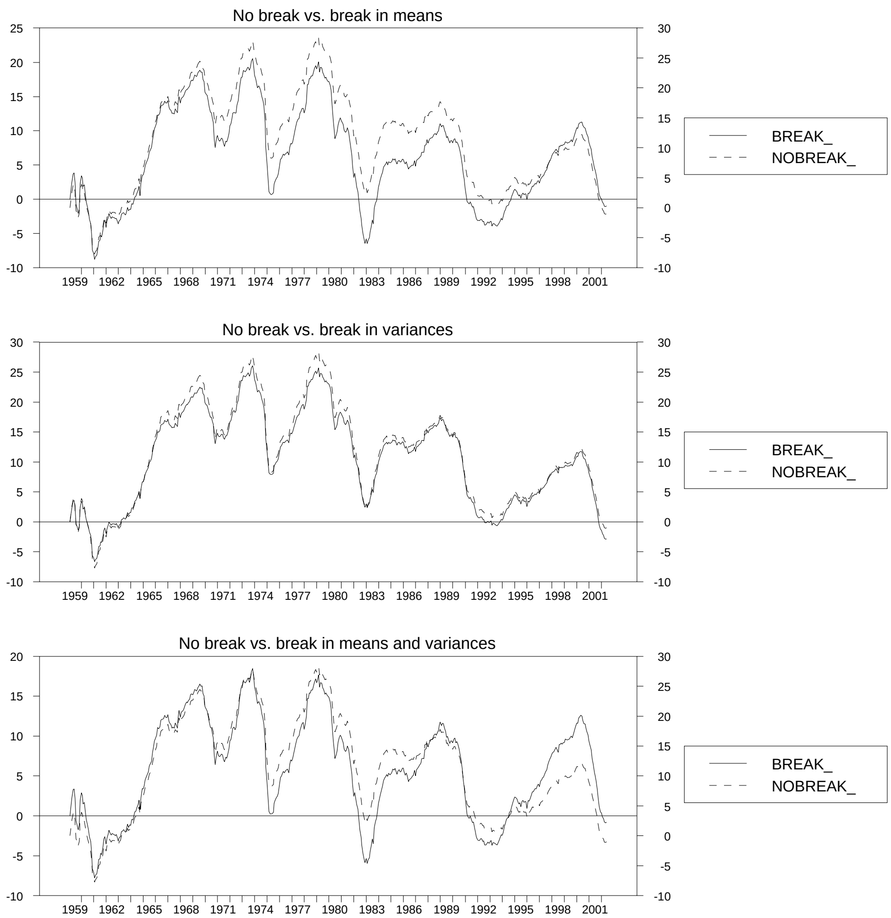
The parameter estimates of regressions (10) and (11) are reported in Tables 4 and 5, respectively. The results of the F test conducted using these two regressions are displayed in Table 6. Only for INC the null hypothesis of no structural break in intercept may be rejected at the usual 0.05 significance 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 3 were used to construct the structural break dummies in the common dynamic factor model. We 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' means, (3) structural break only in the residual variance of the specific components, and (4) structural break both in means and in variances. Only specification CF(0,0) — both for linear and regime-switching models — was used. The estimated parameters of these models together with their standard errors are reported in Tables 7-10. 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 that the structural-break-in-means 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 volatility of the rates of growth of the U.S. macroeconomic variables that decreased during the last 30 years, whereas the average growth rates level did not experience any noteworthy changes.

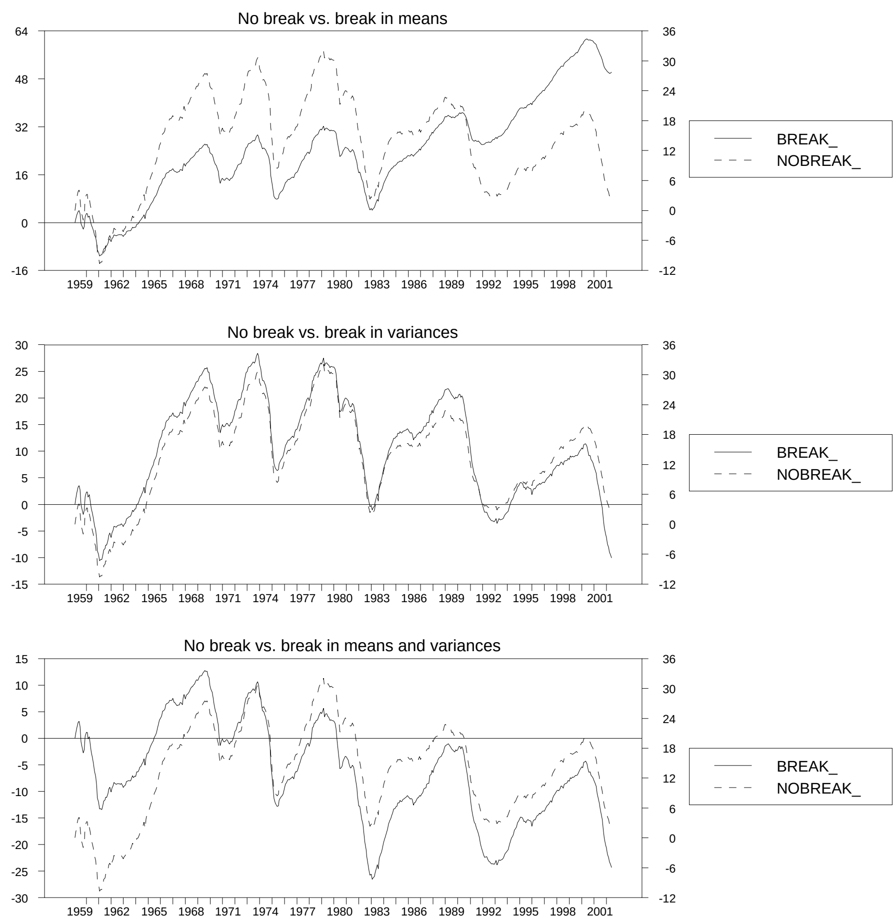
The linear estimates of the composite economic indicators with and without breaks in intercepts and/or variances are plotted on Figure 2. The estimates with deterministic dummies are strikingly close to CF(0,0) with no breaks. This is especially true in the case of CF(0,0) with breaks in the means.

Figure 2: Linear CF(0,0) with and without dummies



The profiles of non-linear composite economic indicators with and without deterministic dummies are shown on Figure 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 means and residual variances of the specific factors displays quite distinct behaviour — it has a clearly expressed downward trend which is not the case of other CF-MS, whereas the CF-MS with breaks in means produces a CEI with an upward growing profile.

Figure 3: CF-MS(0,0) with and without dummies



Different specifications of the model are compared in Table 11. 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 more 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 means of the observed variables does not lead to an important improvement of the loglikelihood: in the linear case the difference between the model with no structural breaks and with breaks only in the means is not significant, while in the regime-switching case it is significant at 5% level. The introduction of the structural breaks in the residual variances of idiosyncratic components 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 means 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 accounting for the structural breaks in the residual variances of specific factors. This can be regarded as an evidence of the negative structural shift in the volatility which has affected the four U.S. macroeconomic time series in question.

Figures 4 through 6 illustrate the low-intercept regime probabilities for each of the models with deterministic dummies. These probabilities are superimposed on the National Bureau of Economic Research (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 means 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.

Figure 4: Recession probabilities vs. NBER dates

Model with break in means

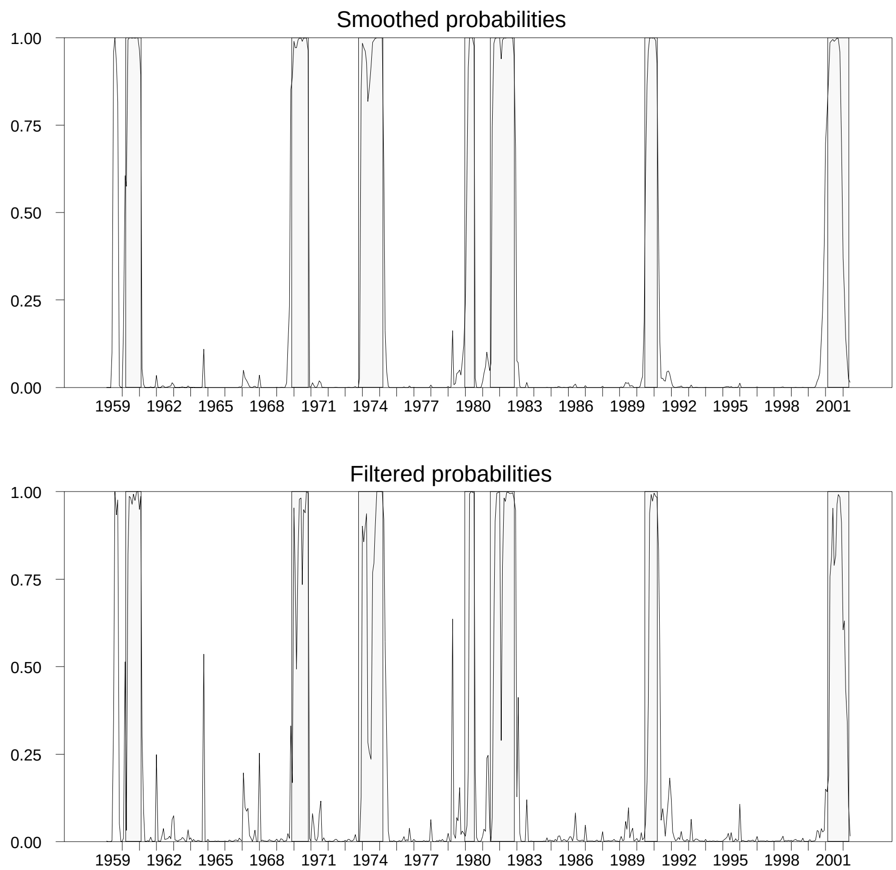


Figure 5: Recession probabilities vs. NBER dates

Model with break in variances

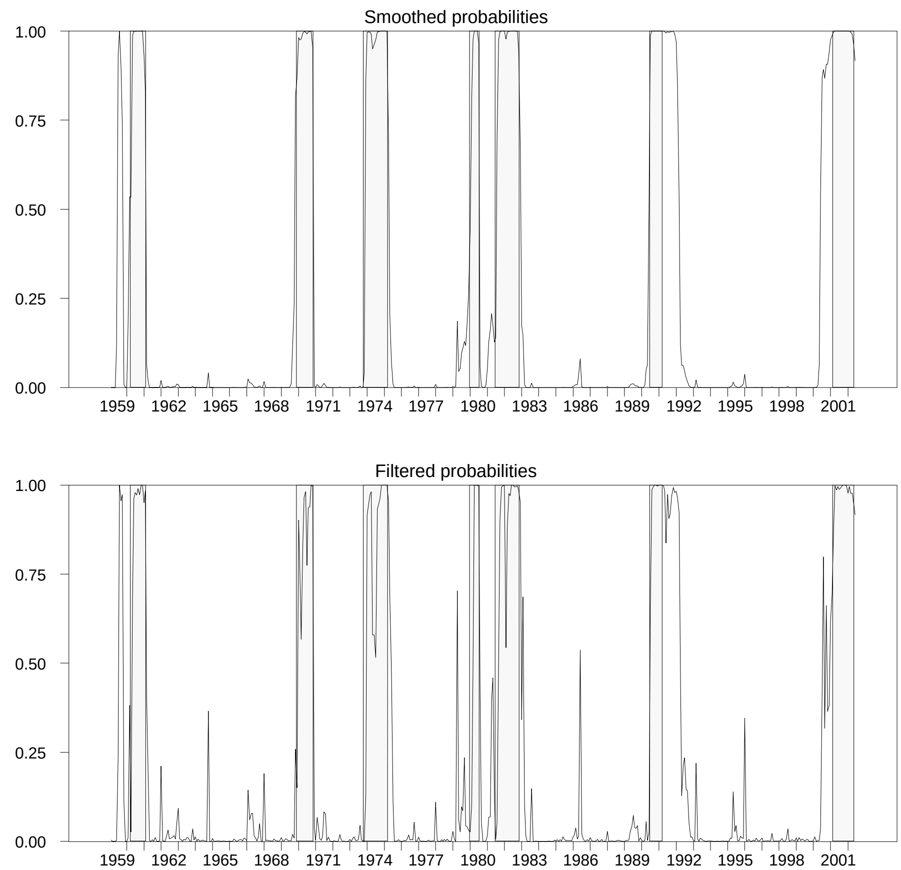
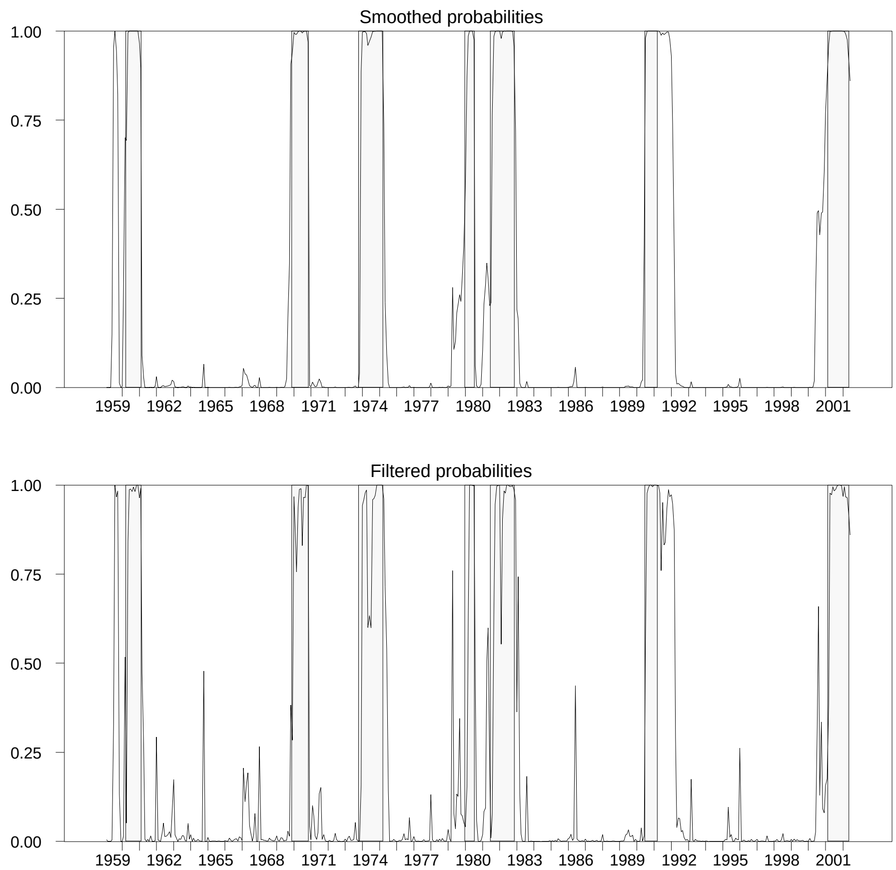


Figure 6: Recession probabilities vs. NBER dates

Model with break in means and variances



We conclude the analysis of the non-linear models examined in this paper by comparing their performance in terms of the in-sample prediction of the NBER turning points. We use the quadratic probability score (QPS) proposed by Diebold and Rudebusch (1989) to evaluate the performance, which is reported in Table 12. Numbers in the third column of the table denoted by DM are the values of the statistic introduced by Diebold and Mariano (1994) to test the hypothesis of equality of the forecast accuracy of two alternative models. The Diebold-Mariano (DM) statistic was computed using the rectangular spectral window of length 201. The forecast accuracy of each structural break model is compared to that of the benchmark (no structural break) model . 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).

First conclusion that can be drawn from this table is that among the deterministic break models only model with break in mean gives better estimates of the NBER turning points than the benchmark model. 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 means of the observed variables).

Secondly, the smoothed conditional probabilities are as a rule closer to the NBER dates than the filtered conditional probabilities. The former take into account both past and future information which allows detecting better the signal, i.e., turning points of the business cycle.

4 Concluding remarks

In this paper we examined a common dynamic factor model with the one-time deterministic shifts in the means of the observed variables and in the residual variance of the idiosyncratic components. The structural shifts in the mean often occur when the accounting methodology used to construct the statistical indicators is changed. Usually this causes a discontinuity in the observed time series not allowing to compare the dynamics before and after the structural break. Alternatively, the structural breaks, especially those hitting the variance, may reflect some "real" economic phenomenon like stabilisation of the economic system when the cyclical fluctuations become less wild.

If we want to investigate the business cycle which is approximated by the fluctuations of the composite economic indicator, these structural breaks may cause a lot of troubles distorting the cyclical signal. Therefore we need some mechanism to isolate the cyclical component. The models of composite economic indicator considered here offer a solution to this problem.

We consider the models with both linear and regime-switching dynamics having a single common dynamic factor. The model allowing structural breaks both in the means of the observed time series and in the residual variances of the idiosyncratic factors was estimated for the real U.S. post-World War II monthly data. A model without structural breaks is compared to (1) model with structural breaks in the observed variables means; (2) model with structural breaks in the residual variances of the specific factors, and (3) model with structural breaks both in the means 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.

References

[1] Andrews, D.W.K. (1993) Test for Parameter Instability and Structural Change with Unknown Change Point, Econometrica, 61, 821-56.

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

[3] Chauvet, M. (2001) A Monthly Indicator of Brazilian GDP, The Brazilian Review of Econometrics (Revista de Econometria), 21.

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

[5] Diebold, F.X., Chen, C. (1996) Testing Structural Stability with Endogenous Breakpoint. A Size Comparison of Analytic and Bootstrap Procedures, Journal of Econometrics, 70, 221-41.

[6] Diebold, F.X., Mariano, R.S. (1994) Comparing Predictive Accuracy, NBER Technical Working Paper N 169.

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

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

[9] Fukuda, S.-I., Onodera, T. (2001) A New Composite Index of Coincident Economic Indicators in Japan: How Can We Improve Forecast Performance?, International Journal of Forecasting, 17, 483-98.

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

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

[12] Krane, S., Wascher, W. (1999) The Cyclical Sensitivity of Seasonality in U.S. Employment, Journal of Monetary Economics, 44, 523-53.

[13] McConnell, M.M., Perez Quiros, G. Output Fluctuations in the United States: What Has Changed Since the Early 1980s?, The American Economic Review, 90, 1464-76.

[14] Stock, J.H., Watson, M.W. (1988) A Probability Model of the Coincident Economic Indicators, NBER Working Paper 2772.

[15] Stock, J.H., Watson, M.W. (1989) New Indexes of Coincident and Leading Economic Indicators, in NBER Macroeconomics Annual

(Eds.) O.-J.Blanchard, S.Fischer, MIT Press, Cambridge, Massachussets, and London.
[16] Stock, J.H., Watson, M.W. (1992) A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience, NBER Working Paper 4014.

5 Appendix

Table 1. 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 2. 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 3. 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. Unrestricted regression
Break-points: 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 5. Restricted regression
Break-point: 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 6. 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 7. Estimated parameters of the linear and regime-switching models with no structural break 1959:1-2002:6

Parameter	Linear LL=-2627.9		Markov-switching LL=-2547.8	
	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 8. 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 9. Estimated parameters of the linear and regime-switching models with structural break in variances 1959:1-2002:6

Parameter	Linear LL=-2586.7		Markov-switching LL=-2492.0	
	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 10. Estimated parameters of the linear and regime-switching models with structural break in intercepts and variances 1959:1-2002:6

Parameter	Linear LL=-2582.9		Markov-switching LL=-2484.8	
	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 11. 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 12. 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 (1994)
p-value = significance value of DM-statistic