Journal of Econometrics 163 (2011) 71–84

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Testing for structural breaks in dynamic factor models

Jörg Breitung[a,*], Sandra Eickmeier[b]

[a] University of Bonn, Institute of Econometrics, 53113 Bonn, Germany
[b] Deutsche Bundesbank, Frankfurt, Germany

ARTICLE INFO

Article history:
Available online 15 December 2010

JEL classification:
C12
C33

Keywords:
Structural break
Factor model
LM test

ABSTRACT

In this paper we investigate the consequences of structural breaks in the factor loadings for the specification and estimation of factor models based on principal components and suggest procedures for testing for structural breaks. It is shown that structural breaks severely inflate the number of factors identified by the usual information criteria. The hypothesis of a structural break is tested by using LR, LM and Wald statistics. The LM test (which performs best in our Monte Carlo simulations) is generalized to test for structural breaks in factor models where the break date is unknown and the common factors and idiosyncratic components are serially correlated. The proposed test procedures are applied to datasets from the US and the euro area.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

In recent years dynamic factor models have become popular for analyzing and forecasting large macroeconomic datasets. These datasets include hundreds of variables and span large time periods. Thus, there is a substantial risk that the data generating process for a subset of variables or all variables has undergone structural breaks during the sampling period. Stock and Watson (2002) argue that factor models are either able to cope with breaks in the factor loadings in a fraction of the series, or can account for moderate parameter drift in all of the series. However, in empirical applications parameters may change dramatically due to important economic events, such as the collapse of the Bretton Woods system, or changes in the monetary policy regime, such as the conduct of monetary policy in the 1980s in the US or the formation of the European Monetary Union (EMU). There may also be more gradual but nevertheless fundamental changes in economic structures that may have led to significant changes in the comovements of variables, such as those related to globalization and technological progress. The common factors may become more (less) important for some of the variables and, therefore, the loading coefficients attached to the common factors are expected to become larger (smaller). If one is interested in estimating the common components or assessing the transmission of common shocks to specific variables, ignoring structural breaks may give misleading results.

Variations in dynamic factor loadings have been considered before. The study most closely related to ours is that of Stock and Watson (2008) who examine the implications of structural breaks in the factor loadings. Consequently, we will compare our testing approach with theirs. Del Negro and Otrok (2008) and Eickmeier et al. (2009) have suggested a model where the factor loadings are modelled as random walks. Finally, Banerjee and Marcellino (2008) have investigated the consequences of time variation in the factor loadings for forecasting based on Monte Carlo simulations and find it to worsen the forecasts, in particular for small samples.

In our theoretical analysis, we first consider the effects of structural breaks in Section 2. It turns out that structural breaks in the factor loadings increase the dimension of the factor space. The reason is that in the case of a single structural break, two sets of common factors are needed to represent the common components in the two subsamples before and after the break. Thus, structural breaks in the factor loadings lead not only to inconsistent estimates of the loadings but also to a larger dimension of the factor space. If we are only interested in decomposing variables into common and idiosyncratic components, it is sufficient to increase the number of factors such that the factor space is large enough to represent the different subspaces of the two regimes. However, if we are interested in a more parsimonious factor representation that allows us to recover the original factors, the estimation has to account for the structural breaks in the factor loadings.

In Section 3, we consider alternative versions of a Chow-type test for a structural break in a strict factor model, where the components are assumed to be white noise. The idea is to treat the estimated factors as if they were known. We show that under certain conditions on the relative rates of N and T the estimation error of the common factors does not affect the asymptotic distribution of the test statistic. A variant of the test procedure for an unknown break date is considered in Section 4.

* Corresponding author. Tel.: +49 228 739201; fax: +49 228 739189.
E-mail address: breitung@uni-bonn.de (J. Breitung).

0304-4076/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2010.11.008

In Section 5, the LM test procedure is generalized to allow for serially correlated factors and idiosyncratic components. By adapting the GLS estimation procedure suggested by Breitung and Tenhofen (2008) we obtain a test procedure that is robust to individual-specific dynamics of the components. The LM version of the test is shown to have reliable size properties whereas the OLS-based test statistic with robust standard errors used in Stock and Watson (2008) may exhibit severe size distortions in finite samples.

Two empirical applications of the test procedures are presented in Section 6. On the basis of a large US macroeconomic dataset provided by Stock and Watson (2005), we examine whether January 1984 (which is usually associated with the beginning of the so called Great Moderation) coincides with a structural break in the factor loadings. On the basis of the LM test, we find evidence of a break around that date. By testing for shifts in the loadings of specific variables we are able to shed some light on the possible sources of the structural break. We also apply the LM test to a large euro-area dataset used in Altissimo et al. (2007). We find evidence for breaks at the dates of the handover of monetary policy from European national central banks to the European Central Bank (ECB) (stage 3 of EMU) and – to a lesser extent – the signing of the Maastricht treaty. Breaks seem to have occurred relatively frequently in the loadings of Spanish and Italian variables around the two events. The changeover to a single monetary policy in the euro area was associated with relatively frequent structural breaks in the loadings of nominal variables, whereas evidence of structural breaks is mainly found for industrial production series around the signing of the Maastricht treaty.

2. The effect of structural breaks on the number of factors

Consider a factor model with r factors[1] $f_t = [f_{1t}, \dots, f_{rt}]'$ that is subject to a common break at time T^*:

$$y_{it} = f_t'\lambda_i^{(1)} + \varepsilon_{it} \quad \text{for } t = 1, \dots, T^* \tag{1}$$

$$y_{it} = f_t'\lambda_i^{(2)} + \varepsilon_{it} \quad \text{for } t = T^*+1, \dots, T, \tag{2}$$

where $t = 1, \dots, T$ denotes the time period and $i = 1, \dots, N$ indicates the cross-section unit.[2] The assumption of a common structural break at T^* is made for convenience only. A generalization to situations with variable-specific break dates is straightforward and is considered in the subsequent sections. The vector of idiosyncratic errors $\varepsilon_{\cdot t} = [\varepsilon_{1t}, \dots, \varepsilon_{Nt}]'$ is assumed to be i.i.d. with covariance matrix $E(\varepsilon_{\cdot t}\varepsilon_{\cdot t}') = \Sigma$, where Σ is a diagonal matrix. Furthermore f_t is assumed to be white noise with positive definite covariance matrix $E(f_tf_t') = \Phi$. Let $\Lambda^{(k)} = [\lambda_1^{(k)}, \dots, \lambda_N^{(k)}]'$, $k = 1, 2$; also $\tau = T^*/T \in (0, 1)$ denotes the relative break date. The unconditional covariance matrix of the vector $y_{\cdot t} = [y_{1t}, \dots, y_{Nt}]'$ results as

$$\begin{aligned} E\left(\frac{1}{T}\sum_{t=1}^{T} y_{\cdot t}y_{\cdot t}'\right) &= \tau\Lambda^{(1)}\Phi\Lambda^{(1)'} + (1-\tau)\Lambda^{(2)}\Phi\Lambda^{(2)'} + \Sigma \\ &\equiv \Psi + \Sigma. \end{aligned}$$

Since the matrix $\Psi = \tau\Lambda^{(1)}\Phi\Lambda^{(1)'} + (1-\tau)\Lambda^{(2)}\Phi\Lambda^{(2)'}$ is a sum of two matrices of rank r, the rank of the covariance matrix of the common component, Ψ, is $2r$ in general. This is due to the fact that a break in the factor loadings implies two linearly independent factors for the first and second subsamples. It follows that if the structural break in the factor loadings is ignored, the number of common factors is inflated by a factor of 2. More generally, if there are k structural breaks in the factor loadings of r common factors, the number of factors for the whole sample is $(k+1)r$, in general.

The practical implication of this result is that if one is only interested in a decomposition of the time series y_{it} into a common component and an idiosyncratic component, then it is sufficient to increase the number of common factors accordingly. However, if one is interested in a consistent estimator of the factors and the factor loadings, then it is important to account for the break of the factor loadings, e.g. by splitting the sample at T^* and re-estimating the factor model for the two subsamples. For illustration consider the previous example with $r = 1$, $T^* = T/2$ and $\lambda_i^{(2)} = \lambda_i^{(1)} + b$. Define an additional factor as

$$f_t^* = \begin{cases} f_t & \text{for } t = 1, \dots, T^* \\ -f_t & \text{for } t = T^*+1, \dots, T. \end{cases}$$

It is not difficult to see that the factor model with a structural break can be represented as

$$y_{it} = \lambda_{1i}^* f_t + \lambda_{2i}^* f_t^* + \varepsilon_{it} \tag{3}$$

where $\lambda_{1i}^* = \lambda_i^{(1)} + (b/2)$ and $\lambda_{2i}^* = -b/2$. Note that the factors in this representation are "orthogonal" in the sense that $E(T^{-1}\sum_{t=1}^{T} f_tf_t^*) = 0$. This example demonstrates that a factor model with a structural break admits a factor representation with a higher dimensional factor space.

To investigate the effects of a structural break on the information criteria suggested by Bai and Ng (2002) for selecting the number of common factors, a Monte Carlo experiment is performed. The data are generated by a factor model $y_{it} = \lambda_{it}f_t + \varepsilon_{it}$, where the single factor f_t and idiosyncratic components are i.i.d. with variances $E(f_t^2) = 1$, $E(\varepsilon_{it}^2) = \sigma_i^2$ and σ_i is uniformly distributed with $\sigma_i \sim U(0.5, 1.5)$. The structural break in the loadings is specified as

$$\lambda_{it} = \begin{cases} \lambda_i & \text{for } t = 1, \dots, T/2 \\ \lambda_i + b & \text{for } t = T/2+1, \dots, T \end{cases} \tag{4}$$

and λ_i is a normally distributed random variable with $\lambda_i \sim \mathcal{N}(1, 1)$. Therefore, the parameter b measures the importance of the structural break. Table 1 presents the average of the number of factors selected by the IC_{p1} criterion suggested by Bai and Ng (2002). The results show that if the break is large, the selection procedure overestimates the number of common factors. As predicted by our theoretical considerations, the information criterion indicates two factors instead of one if b gets large. Thus, ignoring a break in the factor loadings tends to identify too many factors in the sample. This may be misleading and a result of structural breaks.

It is interesting to note that the situation is comparable to the problem of estimating a dynamic factor model within a static framework. As argued by Stock and Watson (2002), lags of the original factors can be accounted for by including additional factors. If one is merely interested in a decomposition into common and idiosyncratic components (e.g. in forecasting), then it is sufficient to estimate the static representation with a larger number of factors. However, if one is interested in the original ("primitive" or "dynamic"[3]) factors, then the static factors are inappropriate as they involve linear combinations of current and lagged values of the original factors.

[1] Note that the notation does not refer to a particular normalization of the (true) common factors. In our asymptotic considerations we follow Bai (2003) and adopt a particular normalization such that $T^{-1}\sum_{t=1}^{T} f_tf_t' \xrightarrow{p} I_r$.

[2] As usual in the literature on factor models, we neglect possible deterministic terms like a constant or a linear time trend. In empirical practice the variables in the dataset are routinely de-meaned or detrended.

[3] See Bai and Ng (2007) and Amengual and Watson (2007).

Table 1
Averages of the estimated number of common factors.

	$T=50$	$T=100$	$T=200$	$T=300$
b	$N=50$			
0.0	1.000	1.000	1.000	1.000
0.3	1.003	1.001	1.000	1.000
0.5	1.100	1.197	1.325	1.398
0.7	1.436	1.729	1.894	1.945
1.0	1.804	1.965	1.999	1.999
b	$N=100$			
0.0	1.000	1.000	1.000	1.000
0.3	1.000	1.000	1.002	1.001
0.5	1.126	1.369	1.739	1.866
0.7	1.525	1.888	1.994	2.000
1.0	1.881	1.995	2.000	2.000
b	$N=200$			
0.0	1.000	1.000	1.000	1.000
0.3	1.001	1.002	1.032	1.074
0.5	1.166	1.531	1.968	1.998
0.7	1.596	1.969	2.000	2.000
1.0	1.926	2.000	2.000	2.000
b	$N=300$			
0.0	1.000	1.000	1.000	1.000
0.3	1.002	1.008	1.063	1.274
0.5	1.165	1.657	1.992	2.000
0.7	1.620	1.980	2.000	2.000
1.0	1.942	2.000	2.000	2.000

Note: This table presents the averages of the estimated number of common factors selected by the IC_{p1} criterion suggested by Bai and Ng (2002). The results are based on 1000 replications of the model with a structural break of size b.

3. The static factor model

Consider a model with an individual-specific structural break at period T_i^* given by

$$y_{it} = f_t'\lambda_i^{(1)} + \varepsilon_{it} \quad \text{for } t = 1, \ldots, T_i^* \tag{5}$$

$$y_{it} = f_t'\lambda_i^{(2)} + \varepsilon_{it} \quad \text{for } t = T_i^* + 1, \ldots, T, \tag{6}$$

where f_t is an r-dimensional vector of common factors. Under the null hypothesis we assume

$$H_0 : \lambda_i^{(1)} = \lambda_i^{(2)}. \tag{7}$$

To test this null hypothesis, the usual Chow test statistics are formed by replacing the unknown vector of common factors, f_t, by its principal components (PC) estimator, $\widehat{f}_t$.[4] Applying the likelihood ratio (LR) principle for testing the ith variable gives rise to the statistic

$$\text{lr}_i = T\Big[\log(S_{0i}) - \log(S_{1i} + S_{2i})\Big],$$

where

$$S_{0i} = \sum_{t=1}^{T}(y_{it} - \widehat{f}_t'\widehat{\lambda}_i)^2,$$

$$S_{1i} = \sum_{t=1}^{T_i^*}\left(y_{it} - \widehat{f}_t'\widehat{\lambda}_i^{(1)}\right)^2,$$

$$S_{2i} = \sum_{t=T_i^*+1}^{T}\left(y_{it} - \widehat{f}_t'\widehat{\lambda}_i^{(2)}\right)^2,$$

and $\widehat{\lambda}_i$ denotes the PC estimator of the vector of factor loadings, whereas $\widehat{\lambda}_i^{(1)}$ and $\widehat{\lambda}_i^{(2)}$ denote the two estimates obtained as the OLS estimates from a regression of y_{it} on $\widehat{f}_t$ for two subsamples according to $t = 1, \ldots, T_i^*$ and $t = T_i^* + 1, \ldots, T$.

The second statistic is the Wald (W) test of the hypothesis $\psi_i = 0$ in the regression

$$y_{it} = \lambda_i'\widehat{f}_t + \psi_i'\widehat{f}_t^* + v_{it}, \quad t = 1, \ldots, T, \tag{8}$$

where

$$\widehat{f}_t^* = \begin{cases} 0 & \text{for } t = 1, \ldots, T_i^* \\ \widehat{f}_t & \text{for } t = T_i^* + 1, \ldots, T. \end{cases} \tag{9}$$

The resulting test statistic is denoted by w_i.

The Lagrange multiplier (LM) statistic, indicated by s_i, is obtained from running a regression of the form

$$\widehat{\varepsilon}_{it} = \theta_i'\widehat{f}_t + \phi_i'\widehat{f}_t^* + \widetilde{\varepsilon}_{it}\,, \tag{10}$$

where $\widehat{\varepsilon}_{it} = y_{it} - \widehat{\lambda}_i'\widehat{f}_t$ denotes the estimated idiosyncratic component. The score statistic is denoted by $s_i = T \cdot R_i^2$, where R_i^2 denotes the uncentered R^2 of the ith regression.

To study the limiting null distributions of the three test statistics we first invoke the usual assumptions of the approximate factor model.

Assumption 1. Let y_{it} be generated by the factor model $y_{it} = \lambda_i'f_t + \varepsilon_{it}$, where it is assumed that λ_i, f_t, and ε_{it} satisfy Assumptions A–G of Bai (2003).

This set of assumptions allows for some weak serial and cross-section dependence and heteroskedasticity among the idiosyncratic components ε_{it}. Furthermore, the factors and idiosyncratic components are allowed to be weakly correlated provided that

$$E\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} f_t\varepsilon_{it}\right\|^2\right) < \infty$$

for all T and N. Under Assumption 1 and $\sqrt{T}/N \to 0$ the estimation error in the regressor $\widehat{f}_t$ does not affect the asymptotic distribution of the test statistic. To establish the usual asymptotic χ^2 distribution of the Chow test, a more restrictive set of assumptions is required:

Assumption 2. (i) For all $t = 1, \ldots, T$, $E(\varepsilon_{it}^2) = \sigma_i^2$ and $E(\varepsilon_{it}\varepsilon_{is}) = 0$ for $t \neq s$. (ii) f_t is independent of ε_{is} for all i, t, s.

The null distributions of the test statistics are presented in the following theorem.

Theorem 1. *Under Assumptions 1 and 2, $T \to \infty$, $N \to \infty$, and $\sqrt{T}/N \to 0$, the statistics s_i, w_i and lr_i have a χ^2 limiting distribution with r degrees of freedom.*

Remark A. It is tempting to combine the individual statistics to obtain a pooled test of the joint null hypothesis that there is no structural break in the N loading vectors $\lambda_1, \ldots, \lambda_N$. For example, a pooled LM test may be constructed as

$$\text{LM}^* = \frac{\left(\sum_{i=1}^{N} s_i\right) - rN}{\sqrt{2rN}},$$

which is equivalent to the standardized mean-group version of the LM statistic (eg. Breitung and Pesaran, 2008). However, the application of this panel statistic would require the additional assumption that ε_{it} and ε_{jt} are independent for all $i \neq j$. Such an assumption is highly unrealistic in most empirical applications (e.g. Chamberlain and Rothschild, 1983; Stock and Watson, 2002; Bai and Ng, 2002).

[4] Some details of the estimator are considered in the Appendix.

Remark B. It is important to select the appropriate number of common factors as otherwise the test may lack power. If the number of common factors is determined from the entire sample, the identification criteria tend to select a larger number of common factors. As has been argued in Section 2, a factor model with a structural break admits a (parameter constant) factor representation with a larger number of factors. Therefore, the number of factors should be selected by applying the information criteria of Bai and Ng (2002) to the subsamples before and after the break, at least if it is assumed that the break date is (roughly) the same for all variables.[5]

Remark C. Although the three test statistics are asymptotically equivalent, i.e., they possess the same asymptotic distribution under the null hypothesis as well as a properly specified sequence of local alternatives, the power of the LM test may suffer from ignoring the break when computing the estimator of the residual variances in small samples (cf. Vogelsang, 1999).

To investigate the finite sample properties of the test statistics, a Monte Carlo experiment is performed. We simulate data according to the single-factor model $y_{it} = \lambda_i^{(k)} f_t + \varepsilon_{it}$, where the factor and idiosyncratic components are generated as in Section 2. The empirical sizes of the three different test statistics LR, LM und W are presented for various sample sizes in Table 2. It turns out that for all N and T the rejection frequencies of tests are close to the nominal size of 0.05. Among the three asymptotically equivalent tests the LM test has the best size properties. The LR test performs only slightly worse and the W test tends do be (slightly) oversized. Table 3 reports results on the empirical power of the tests. The structural break is again modelled as a shift of size b in the mean of the factor loadings (see Section 2). The results suggest that the tests have similar power. The LR and W tests seem to have slightly higher power but this is no surprise as these tests are oversized for small T. Overall, our simulation experiments (based also on models with more factors and other data generating mechanisms[6]) suggest that the performances of all three tests are similar if T and N are sufficiently large. In what follows we focus on the LM test statistic as it is computationally convenient and has superior size properties.

Table 2
Empirical sizes (average rejection frequencies).

N	LR	LM	W	LR	LM	W
	$T = 50$			$T = 100$		
20	0.055	0.049	0.057	0.055	0.052	0.057
50	0.054	0.048	0.056	0.051	0.048	0.052
100	0.056	0.049	0.058	0.052	0.049	0.054
150	0.055	0.048	0.058	0.053	0.050	0.054
200	0.056	0.048	0.058	0.052	0.049	0.053
	$T = 150$			$T = 200$		
20	0.051	0.048	0.051	0.051	0.050	0.052
50	0.052	0.049	0.052	0.050	0.048	0.050
100	0.051	0.049	0.052	0.051	0.050	0.052
150	0.052	0.050	0.053	0.051	0.049	0.052
200	0.052	0.050	0.053	0.051	0.050	0.052

Note: The entries report the average frequencies of rejection of N variable-specific tests for structural breaks within each dataset generated by a factor model without a structural break ($b = 0$). The nominal size is 0.05 and 1000 replications are used to compute the averages.

Table 3
Power against a break at $T_i^* = T/2$.

b	LR	LM	W	LR	LM	W
	$T = 50$			$T = 100$		
0.1	0.062	0.054	0.064	0.064	0.061	0.065
0.2	0.083	0.074	0.086	0.110	0.105	0.112
0.3	0.116	0.105	0.118	0.168	0.162	0.170
0.5	0.200	0.186	0.203	0.305	0.298	0.307
	$T = 150$			$T = 200$		
0.1	0.071	0.069	0.072	0.080	0.078	0.081
0.2	0.132	0.128	0.133	0.153	0.150	0.154
0.3	0.219	0.214	0.220	0.260	0.257	0.261
0.5	0.381	0.376	0.383	0.445	0.441	0.446

Note: The entries report the average frequencies of rejection of $N = 50$ variable-specific tests for structural breaks. The data are generated on the basis of (4) with a structural break of size b. See Table 2 for further information.

4. Unknown break dates

So far we have assumed that the break date T_i^* is known. In many empirical applications (such as the ones considered in Section 6) the precise date of the structural break is unknown. In this section we adapt the Andrews (1993) tests for structural breaks with an unknown break date. Let $[\cdot]$ denote the integer part operator and $W_i(a)$ be an r-dimensional vector of standard Brownian motions defined on $a \in [0, 1]$. Following Andrews (1993) and others, our test is based on the following assumption[7]:

Assumption 3. As $T \to \infty$,

$$\frac{1}{T}\sum_{t=1}^{[\tau T]} f_t f_t' \xrightarrow{p} \tau \Sigma_f$$

$$\frac{1}{T}\sum_{t=1}^{[\tau T]} f_t \varepsilon_{it} \Rightarrow \sigma_i \Sigma_f^{1/2} W_i(\tau)$$

for all i and $\tau \in [0, 1]$, where Σ_f is a positive definite matrix.

Andrews (1993) considered three asymptotically equivalent test statistics based on the supremum of the LM, LR and Wald statistics. Since the test statistics perform very similarly in our Monte Carlo experiment of Section 3 and it is particularly simple to compute, we focus on the sup-LM statistic given by

$$\mathcal{S}_{i,T}(\tau_0) = \sup_{\tau \in [\tau_0, 1-\tau_0]} (s_i^\tau), \tag{11}$$

where s_i^τ denotes the LM statistic for a structural break in the prespecified interval of relative break dates $\tau \in [\tau_0, 1 - \tau_0]$ and cross-section unit i. Andrews and Ploberger (1994) proposed optimal tests that maximize the weighted average power. However, simulation studies (e.g. Andrews et al., 1996) suggest that in most situations the power loss of the simple sup-LM statistic is small relative to that in the optimal tests and, therefore, we will only consider the sup-LM statistic (11).

In the following theorem it is stated that under assumptions similar to those in Theorem 1 the limiting distribution of the sup-LM test is the same as in Andrews (1993) and, therefore, the critical tables presented therein can be used.

Theorem 2. *Under Assumptions 1–3, $T \to \infty$, $N \to \infty$, and $\sqrt{T}/N \to 0$, the sup-LM statistic (11) is asymptotically distributed as*

$$\mathcal{S}(\tau_0) = \sup_{\tau \in [\tau_0, 1-\tau_0]} \left\{ \frac{[\tau W(1) - W(\tau)]'[\tau W(1) - W(\tau)]}{\tau(1-\tau)} \right\}$$

where $W(\cdot)$ denotes an r-dimensional vector of standard Brownian motions.

Remark D. It is not difficult to show that the same limiting distribution results if the LM statistic is replaced by the Wald or the

[5] We are grateful to Peter Boswijk who pointed out this problem during the conference.

[6] For example, if the number of common factors increases, the positive size biases of the LR and W test increase, whereas the LM test becomes slightly conservative. The results of the additional Monte Carlo simulations are included in the working paper version of this paper.

[7] See Perron (2006) for a thorough discussion of this assumption.

Table 4
Empirical sizes of the sup-LM test for unknown break dates.

N	$T = 50$	$T = 100$	$T = 150$	$T = 200$
50	0.024	0.030	0.034	0.038
100	0.022	0.028	0.031	0.034
150	0.020	0.025	0.030	0.033
200	0.021	0.027	0.033	0.034

Note: The entries report the average frequencies of rejection of variable-specific tests for structural breaks assuming unknown break dates. The data are generated by the factor model with $r = 1$ and no structural break (the null hypothesis). The critical value presented in Andrews (1993) for $\tau_0 = 0.1$ is applied. The nominal size is 0.05.

LR statistics. However, as noted by Andrews (1993), the sequence of LM statistics is particularly easy to compute.

Remark E. So far we have assumed that under the alternative there is a single structural break in the factor loadings. To allow for multiple breaks, the efficient search procedure proposed by Bai and Perron (2003) may be employed for finding the candidate break dates. Alternatively, a sequential estimation and testing procedure for the break dates may be entertained (cf. Bai and Perron, 1998).

To investigate whether the asymptotic distribution presented in Theorem 2 yields a reliable approximation for small samples, we repeated the Monte Carlo experiment of Section 3. However, it is assumed that the break point is unknown and, therefore, the sup-LM statistic is employed. The test procedures search for a structural break within the interval $\tau \in [0.1, 0.9]$ (i.e. $\tau_0 = 0.1$) and the (asymptotic) critical values provided by Andrews (1993) are applied. Table 4 presents the actual size of the sup-LM test for various sample sizes. The nominal size is 0.05. It turns out that for small samples the tests tend to be conservative, yet the actual sizes tend slowly to 0.05 as T increases.

5. Dynamic factor models

In the previous section we have considered the framework of a static factor model, where the common and idiosyncratic components are white noise. In many practical situations, however, the variables are generated by dynamic processes. In this section we therefore generalize the factor model and assume that the idiosyncratic components in the model $y_{it} = \lambda_i' f_t + u_{it}$ are generated by individual-specific AR(p_i) processes:

$$u_{it} = \varrho_{i,1} u_{i,t-1} + \cdots + \varrho_{i,p_i} u_{i,t-p_i} + \varepsilon_{it} \tag{12}$$

$$\varrho_i(L) u_{it} = \varepsilon_{it}, \tag{13}$$

where $\varrho_i(L) = 1 - \varrho_{i,1} L - \cdots - \varrho_{i,p_i} L^{p_i}$. To analyze the asymptotic properties of the tests in a dynamic factor model we make the following assumption.

Assumption 4. (i) The idiosyncratic components are generated by (13), where all roots of the autoregressive polynomial $\varrho_i(z)$ are outside the unit circle. (ii) For all t, $E(\varepsilon_{it}^2) = \sigma_i^2$ and $E(\varepsilon_{it}\varepsilon_{is}) = 0$ for $t \neq s$. (iii) f_t is independent of ε_{it} for all i and t.

The dynamic process of the vector of common factors is left unspecified. We only assume that the second moments are finite, i.e., the probability limit $T^{-1}\sum_{t=1}^T f_t f_t' \xrightarrow{p} \Sigma_f$ is a finite positive definite matrix (see Assumption A in Bai (2003)).

To test for structural breaks, Stock and Watson (2008) suggest applying conventional Chow tests for each variable y_{it}, where the unobserved factors are replaced by estimates obtained from applying principal components. A possible serial correlation of the errors is accounted for by using heteroskedasticity and autocorrelation consistent (HAC) estimators for the standard errors of the coefficients (cf. Newey and West, 1987). This approach has, however, two important drawbacks. First, since the OLS estimator is inefficient in the presence of autocorrelated errors, the resulting test suffers from a loss of power relative to a test based on a GLS estimator. Second, it is well known that the HAC estimator may perform poorly for small samples.

To sidestep these difficulties, we follow Breitung and Tenhofen (2008) and compute the test statistic based on a GLS estimation of the model. The GLS transformed model results as

$$\varrho_i(L) y_{it} = \lambda_i'[\varrho_i(L)\widehat{f}_t] + \psi_i'[\varrho_i(L)\widehat{f}_t^*] + \varepsilon_{it}^*, \tag{14}$$

where $\widehat{f}_t$ denotes the PC estimator of the common factors, $\widehat{f}_t^* = \widehat{f}_t$ for $t = T_i^* + 1, \ldots, T$ and $\widehat{f}_t^* = 0$ otherwise. The lag polynomials $\varrho_i(L)$, $i = 1, \ldots, N$, can be estimated by running least squares regressions

$$\widehat{u}_{it} = \varrho_{i,1}\widehat{u}_{i,t-1} + \cdots + \varrho_{i,p_i}\widehat{u}_{i,t-p_i} + \widetilde{\epsilon}_{it}\ , \tag{15}$$

where $\widehat{u}_{it}$ is the PC estimator of the idiosyncratic component. The lag length p_i can be determined by employing the usual information criteria. To test the hypothesis of no structural break at T_i^*, the LM statistic for $\psi_i = 0$ is computed. The resulting test statistic is denoted by $\widetilde{s}_i$. We focus on the LM statistic as this statistic possesses the best size properties among all tests considered in Section 3. The following theorem states that the asymptotic null distribution of the resulting LM test statistic is the same as in Theorem 1.

Theorem 3. *Let $\widetilde{s}_i$ denote the LM statistic for $\psi_i = 0$ in the regression*

$$\widehat{\varrho}_i(L) y_{it} = \lambda_i'[\widehat{\varrho}_i(L)\widehat{f}_t] + \psi_i'[\widehat{\varrho}_i(L)\widehat{f}_t^*] + \widetilde{\varepsilon}_{it}^*, \qquad t = p_i + 1, \ldots, T. \tag{16}$$

Under Assumptions 1 and 4, $T \to \infty$, $N \to \infty$, and $\sqrt{T}/N \to 0$, $\widetilde{s}_i$ is asymptotically χ^2 distributed with r degrees of freedom.

Remark F. Assumption 3 rules out temporal heteroskedasticity of the idiosyncratic components. It is well known that the Chow test is not robust against a break in the variances. To obtain a robust statistic in the case of serial heteroskedasticity, the approach of White (1980) can be adopted. Alternatively, a GLS variant of the test statistic that is robust against a break in the variance at T_i^* can be constructed as

for $t = p_i + 1, \ldots, T_i^*$:

$$\frac{1}{\widehat{\sigma}_i^{(1)}}\varrho_i(L) y_{it} = \lambda_i'\left[\frac{1}{\widehat{\sigma}_i^{(1)}}\varrho_i(L)\widehat{f}_t\right] + \psi_i'\left[\frac{1}{\widehat{\sigma}_i^{(1)}}\varrho_i(L)\widehat{f}_t^*\right] + \widetilde{\varepsilon}_{it}^*$$

for $t = T_i^* + 1, \ldots, T$:

$$\frac{1}{\widehat{\sigma}_i^{(2)}}\varrho_i(L) y_{it} = \lambda_i'\left[\frac{1}{\widehat{\sigma}_i^{(2)}}\varrho_i(L)\widehat{f}_t\right] + \psi_i'\left[\frac{1}{\widehat{\sigma}_i^{(2)}}\varrho_i(L)\widehat{f}_t^*\right] + \widetilde{\varepsilon}_{it}^*.$$

Remark G. It is possible to construct the sup-LM statistic of Section 4 based on the GLS transformed series. Using arguments similar to those in the proof of Theorem 2 it can be shown that the resulting test statistic possesses the same limiting distribution as $S(\tau_0)$ defined in Theorem 2.

To investigate the small sample properties of the test, we generate the factor as $f_t = 0.5 f_{t-1} + \nu_t$.[8] The idiosyncratic errors are generated by using the model

$$u_{it} = \varrho u_{i,t-1} + \varepsilon_{it}$$

[8] Since the data generating process for f_t is irrelevant for the asymptotic properties of the test, we do not present the results for other values of the autoregressive coefficient.

Table 5
Empirical sizes in the dynamic model.

ϱ	LM(stat)	LM(dyn)	HAC(4)	HAC(12)	$\text{HAC}_0(4)$	$\text{HAC}_0(12)$
	$N = 100, T = 100$					
0.2	0.089	0.049	0.105	0.147	0.052	0.030
0.5	0.187	0.052	0.142	0.175	0.072	0.037
0.9	0.405	0.049	0.258	0.245	0.156	0.055
−0.2	0.023	0.050	0.078	0.129	0.037	0.026
−0.5	0.004	0.049	0.058	0.114	0.028	0.023
ϱ	$N = 100, T = 500$					
0.2	0.095	0.050	0.070	0.081	0.055	0.050
0.5	0.193	0.049	0.083	0.088	0.065	0.052
0.9	0.425	0.049	0.141	0.110	0.115	0.065
−0.2	0.022	0.051	0.057	0.076	0.045	0.046
−0.5	0.003	0.050	0.047	0.068	0.038	0.043
ϱ	$N = 100, T = 1000$					
0.2	0.094	0.049	0.062	0.069	0.053	0.050
0.5	0.197	0.049	0.070	0.072	0.060	0.052
0.9	0.426	0.050	0.109	0.088	0.094	0.062
−0.2	0.021	0.049	0.053	0.064	0.046	0.047
−0.5	0.003	0.049	0.046	0.061	0.040	0.045

Note: Entries report the average frequencies of rejection of N variable-specific tests for a structural break at $T_i^* = T/2$ computed from 1000 replications of the dynamic model without a structural break. The nominal size is 0.05. The column LM(stat) presents the rejection rates for an LM test that ignores the serial correlation in the idiosyncratic component. LM(dyn) indicates the test based on a GLS regression considered in Theorem 3. HAC(k) denotes an OLS-based test using robust (HAC) standard errors with the truncation lag computed from (17). $\text{HAC}_0(k)$ is the LM variant of the test statistic based on the residuals of the restricted regression.

for all $i = 1, \ldots, N$. For the variances we set $E(v_t^2) = 1$ and $E(\varepsilon_{it}^2) = \sigma_i^2$, where $\sigma_i \sim U(0.5, 1.5)$. The factor loadings are obtained from independent draws of a $\mathcal{N}(1, 1)$ distribution. Table 5 reports the average rejection rates for the individual tests $\widetilde{s}_i$. The tests assume that the break occurs at period $T_i^* = T/2$.

To assess the size bias that results from ignoring the serial correlation of the idiosyncratic component we first present the ordinary LM statistic that assumes white noise errors. As can be seen from the first column of Table 5, the rejection rates of the test are far from the nominal size of 0.05 even if the autoregressive coefficient is fairly small. In contrast, the actual size of the LM statistic computed from the GLS regression is close to the nominal size for all values of ϱ. The columns labelled as HAC(k) report the actual sizes of the OLS-based t-statistics employing robust standard errors, where the truncation lag is specified by applying the rule

$$\ell_T(k) = k(T/100)^{2/5} \quad \text{with } k \in \{4, 12\}. \tag{17}$$

Since we found that the sizes are more reliable if the test is computed using the LM principle, we also compute the HAC standard errors from the residuals of the restricted regression (i.e. where we have imposed the null hypothesis). The resulting test statistics are indicated by $\text{HAC}_0(k)$.

From the results presented in Table 5 it turns out that the test statistics based on HAC standard errors perform poorly for small samples. The test based on the restricted residuals ($\text{HAC}_0(k)$) performs much better but still exhibits some size distortions. To demonstrate that the size bias of HAC tests is indeed a small sample phenomenon, we repeat the simulations for $T = 500$ and $T = 1000$. The results show that if T increases, the empirical sizes of the original HAC(k) slowly tend to the nominal size.

6. Empirical applications

Our test procedure is applied to two settings. In Section 6.1, we investigate whether the mid-1980s in the US can be associated with structural breaks in the loadings. In Section 6.2, we consider possible breaks in the euro-area economies due to the two major events in the 1990s, the signing of the Maastricht treaty and the handover of monetary policy from national central banks to the ECB. We will address important issues that typically arise in applications.

6.1. The US economy in the mid-1980s

In this subsection we apply our test procedure to the dataset constructed by Stock and Watson (2005) and provided on Mark Watson's web page to investigate whether the mid-1980s in the US can be associated with structural breaks in the factor loadings. The dataset contains 132 monthly US series including measures of real economic activity, prices, interest rates, money and credit aggregates, stock prices, and exchange rates. It spans 1960–2003.[9]

We start by considering a single break in 1984:01. That date has been associated with the beginning of the so called Great Moderation, i.e. the decline in the volatility of output growth and inflation (Kim and Nelson, 1999; McConnell and Perez-Quiros, 2000; Stock and Watson, 2008). One motivation for our empirical application is that we will be able to compare our results to those of Stock and Watson (2008) who also test for structural breaks in the factor loadings in 1984:01 and use a very similar dataset.[10] Another motivation of our application is that the sources of the Great Moderation are still controversial. Previous papers have applied structural break tests to univariate linear and univariate Markov-switching models or, more recently, structural VAR models with time-varying parameters to tackle this question. They have come up with various explanations, and it is still unclear to what extent either "good luck" or structural changes including "good policy" have contributed to the volatility decline (cf. Galí and Gambetti, 2008 as well as Stock and Watson, 2003 and references therein). "Good luck" is based on the observation that smaller shocks hit the economy after the break date considered (cf. Benati and Mumtaz, 2007). "Good policy" on the other hand emphasizes the fact that monetary policy has put more weight on inflation relative to output stabilization since the 1980s (Clarida et al., 2000). Other structural changes that may have played a role include improved inventory management mainly in the durable goods sector (McConnell and Perez-Quiros, 2000; Davis and Kahn, 2008) as well as financial innovation and better risk sharing spurred on by financial deregulation (IMF, 2008). Given this ongoing controversy in the discussion we find it useful to analyze the mid-1980s in the US with a new methodology. Our data-rich framework enables us not only to test for breaks in the factor loadings associated with many variables and thus to identify "dramatic" changes in the economy, but also to test where breaks have occurred, i.e. loadings of which variables or groups of variables have changed. This may help to shed some light on the sources of possible structural changes.

Factor analysis requires some pre-treatment of the data. We proceed exactly as in Stock and Watson (2005). Non-stationary raw data (which were already available to us in seasonally adjusted form) are differenced until they are stationary. In our baseline, we remove outliers.[11] To assess whether removing outliers from the data affects our results we also consider below the case where

9 The original dataset is provided for the period 1959–2003. Some observations are, however, missing in 1959. We therefore decided to use a balanced dataset starting in 1960.

10 The main difference is that their dataset is quarterly and also covers more recent years (up to 2006).

11 Outliers are defined as observations of each (stationary) variable with absolute median deviations larger than six times the interquartile range. They are replaced by the median value of the preceding five observations.

Table 6
Tests for structural breaks (US data).

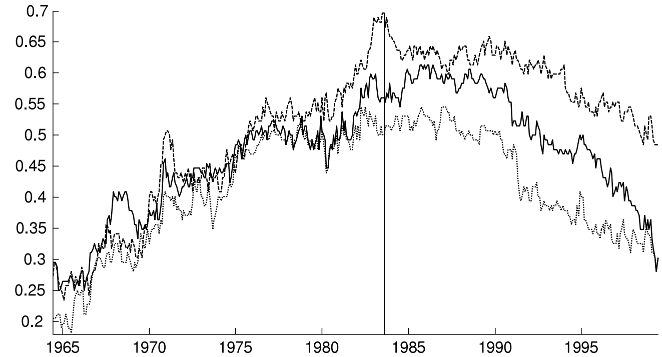
	$r = 6$	$r = 7$	$r = 8$	$r = 9$
	With outlier adjustment			
rej % LM (1984:01)	0.48	0.52	0.54	0.55
rej % HAC (1984:01)	0.50	0.58	0.64	0.62
rej % sup-LM	0.66	0.67	0.75	0.70
	Without outlier adjustment			
rej % LM (1984:01)	0.61	0.64	0.65	0.67
rej % HAC (1984:01)	0.61	0.66	0.67	0.65
rej % sup-LM	0.73	0.73	0.80	0.76

Note: "rej % LM" is the relative rejection rate for the N individual LM statistics and "rej % HAC" is the respective rejection rate for the OLS-based test procedure with HAC standard errors, where the truncation lag results from (17) with $k = 4$. "rej % sup-LM" indicates the rejection rates for the test with unknown break date.

data were not outlier adjusted.[12] Finally we normalize the series to have means of zero and unit variances. The reader is referred to Stock and Watson (2005) for details on the composition and the treatment of the dataset. Following Stock and Watson (2005), our benchmark estimation is based on $r = 9$ factors. The Bai and Ng (2002) IC_{p1} criterion only indicates $r = 7$, but, as already pointed out in Stock and Watson (2005), we find the criterion to be flat for $r = 6$–10. We therefore also consider $r = 6$–8 factors below.[13]

As argued in the previous sections, we focus on the LM test in our application. We test the null hypothesis of no break in the factor loadings in 1984:01. We generally allow for a break in the variance of the idiosyncratic components as suggested in Remark F. Table 6 shows the rejection rates, i.e. the shares of the 132 variables for which a structural break is found, indicated by the LM test and, in comparison, by the OLS-based test statistic with HAC (robust) standard errors. For the former test, we allow for six autoregressive lags of the idiosyncratic components, and for the latter test, the number of autoregressive lags for the Newey–West correction is set to 7 according to the formula (17) with $k = 4$. This allows us to concentrate on structural changes in the common component as a source of the Great Moderation as opposed to "good luck" which would at least partly be reflected in the error variances.

A clear structural break is identified for the majority of the variables at 1984:01. On the basis of $r = 9$ and outlier-adjusted data, the LM test yields a rejection rate of 0.55. The rejection rate suggested by the HAC test procedure considered in Section 4 is even larger (0.62), consistent with our simulation results which have illustrated that the HAC test procedure tends to reject too often the null hypothesis of no structural breaks. That share also exceeds the share estimated by Stock and Watson (2008), who find that 35% of the variables exhibited structural breaks in the loadings. The reason is that Stock and Watson (2008) rely on fewer (three or four) factors in that paper. When we re-do the tests on the basis of fewer factors, we obtain rejection rates comparable to those presented by the authors. Interestingly, the shares increase to 0.67 (for the LM test) and 0.65 (for the HAC test procedure) when outliers are not removed prior to the estimation, suggesting that our outlier adjustment already takes care of breaks in a subset of variables.



Note: The vertical line presents the supposed starting date of the Great Moderation. Solid line: LM statistic that assumes a break in variances (based on outlier adjusted data). Dotted line: LM statistic based on constant variances (based on outlier adjusted data). Dashed line: LM statistic that assumes a break in variances (based on data without outlier adjustment).

Fig. 1. Relative rejection frequencies (US data).

As shown in Section 2, the number of common factors may be overestimated in the case of a structural break. We therefore split the sample into two subsamples: 1960:01–1983:12 and 1984:01–2003:12 and re-estimated r for each subsample (and for our baseline with outlier-adjusted data). The Bai and Ng (2002) IC_{p1} test suggests $r = 4$ for the first subsample and $r = 6$ factors for the second subsample supporting our theoretical considerations and our finding of a structural break based on $r = 9$. Unlike in the simulations, the estimated numbers of factors in the two subsamples are not equal, nor are they equal to half the number of factors estimated on the basis of the total sample. One reason may be that the loadings of some of the variables or those associated with some of the factors does not exhibit a structural break. Other explanations may be that the size of the break is moderate (see our Monte Carlo simulations of Section 2) or that variables' loadings shift at different points in time. If we were interested in estimating the factors, we would need to split the sample and estimate the factors on the basis of smaller r. However, our objective is to test for a structural break. In order to consider all factors, we keep on working with nine factors.

We next investigate whether the break has occurred exactly in 1984:01 and whether it is the only structural break during the sample period. We apply the LM test for each possible break point in the interval $\tau = [0.1, 0.9]$, i.e. $\tau_0 = 0.1$. The solid line in Fig. 1 shows the relative rejection frequencies for the individual LM tests for each point in time. Between the beginning of the 1980s and the beginning of the 1990s the test rejects the null hypothesis of no structural break for more than half of the variables, and particularly high rejection rates (around 60%) are found around 1985. Fig. 1 also shows that it may matter whether one allows for a break in the variance of the idiosyncratic components. The test that assumes a constant variance is represented by the dotted lines. This version of the test tends to yield smaller rejection rates compared to the robust version. We have also applied the sup-LM test for unknown break dates as suggested in Section 4 to our dataset and reject the null hypothesis of no structural break at any point in time for $r = 9$ for 70% of the variables (Table 6). The average of the estimated break dates obtained from the maxima of the LM statistics is 1981:01. Finally, Fig. 1 also reveals a discrepancy between rejection rates obtained from outlier-adjusted data and from the unadjusted data (already apparent in Table 6). The latter rates are represented by the dashed line and exceed the rejection rates from our baseline over most of the sample period. Interestingly, the dashed line reaches its peak exactly in 1984:01.

12 We are grateful to Jean-Pierre Urbain for suggesting this exercise to us.

13 As noted in Remark B, the number of factors should be determined by using the subsamples before and after the break. Indeed we found that the information criteria tend to suggest a smaller number of factors for the subsamples than for the whole sample. However, since the test for structural breaks is applied to a range of possible break dates, this would mean that the numbers of factors have to be re-estimated for all time periods under consideration. Furthermore, the information criteria tend to choose different numbers of factors for the two subsamples. We therefore decided to employ the same number of factors as was used in the earlier literature. Note that if the number of factors is overspecified, the tests tend to have low power. Since in our applications the great majority of the tests reject the null hypothesis, we conclude that a possible loss of power is not a problem in our case.

Table 7
Tests for specific variables (US data).

Variable	*p*-value	Commonality
Industrial production (IP)	0.06	1.00
IP durable cons. goods	0.04	1.00
IP non-dur. cons. goods	0.72	0.99
IP durable mat. goods	0.00	1.00
IP non-dur. mat. goods	0.04	0.98
Inventory	0.00	0.50
Consumption	0.15	1.00
CPI	0.00	0.99
FFR	0.01	0.74
Cons. expectations	0.00	0.67
10 y gvt bond yields	0.01	0.71
S&P 500	0.03	0.95
Effective exch. rate	0.00	0.76
Commodity prices	0.12	0.49

Note: The *p*-values are the marginal significance levels of the LM test for a structural break at 1984:01. The commonality is equivalent to the R^2 of the regression of the variable on the common factors. Variables were transformed as in Stock and Watson (2005). Outlier-adjusted data are used.

Giordani (2007) has pointed out that, although some series may be I(1) in the total period, they may be stationary in subperiods and differencing them would result in an overdifferencing. To avoid overdifferencing, we consider an alternative dataset where inflation, interest rates, money growth, capacity utilization and the unemployment rate enter in levels rather than in growth rates as before (and as in Stock and Watson, 2005, 2008). The results do not change much, and these are available upon request.

To investigate where structural breaks have occurred, it may be instructive to look at test results for individual variables. We focus on several key macroeconomic variables which are of general interest, but also on variables which are particularly interesting against the background of the Great Moderation and its possible sources such as a monetary policy instrument, inventories, the production of durable and non-durable goods as well as consumption and several financial variables. Breaks or the lack of breaks in the loadings of these variables may support some of the conjectures on the sources of the Great Moderation discussed above. We provide results for the heteroskedasticity-robust version of the test. Table 7 suggests that there is evidence of a break in the loadings of some but not all key macroeconomic variables in 1984:01. There seems to be a break in the loadings of the CPI and consumer expectations (transformed accordingly), but not in the loadings of commodity prices. For total industrial production the test rejects the null hypothesis only at the 10% significance level. Among the variables which may provide some information on the sources of the changes, breaks are found in the loadings of inventories, the production of material goods and durable consumer goods, but not of the production of non-durable consumer goods. The LM test also rejects the null hypothesis of no structural break in the loadings of the Federal funds rate and of most financial variables (long-term interest rates, stock prices, and effective exchange rates). Table 7, however, also reveals no evidence for changes in the loadings of consumption, although a popular hypothesis is that financial integration leads to consumption smoothing and therefore reduces the responsiveness of consumption to shocks. Notice also that the commonality is high for all variables shown in Table 7: the factors explain at least half of the variation in each variable and almost all of the variance in (the stationary versions of) industrial production variables, consumption, and CPI.

To summarize, we find some support for substantial changes in the US economy around the date that is generally associated with the Great Moderation in the US, 1984:01. Our analysis further suggests that various structural changes can probably explain this result, including changes in the conduct of monetary policy, ongoing financial integration and better inventory management (possibly in the durable goods sector).

6.2. Have the Maastricht treaty and the handover of monetary policy to the ECB led to structural breaks in the euro area?

Our second application is concerned with possible changes in comovements that may have occurred in the euro area in the 1990s due to two major events. The first event is the Maastricht treaty, which was signed in 1992:02. With the treaty, a timetable for EMU was prepared and conditions for countries to become EMU members were fixed. These include low inflation rates, converged interest rates, stable exchange rates, and solid fiscal budgets. The second event was stage 3 of EMU, i.e. the changeover to a single monetary policy and the fixing of exchange rates, in 1999:01. This setting is particularly interesting, since these events may have altered the comovement between variables, and this will just be reflected in breaks in the loadings.

It is still not entirely clear how these two events have affected the comovements of business cycles and other variables in euro-area countries. Some arguments point to greater comovements, some to smaller comovements. Also it is unclear whether changes have occurred at exactly the dates of or before or after these two events. On the one hand, the Maastricht treaty and accession prospects have forced countries to improve their fiscal situation and to carry out structural reforms in order to qualify for EMU membership. Greater structural and political similarity could lead to long-run convergence and a greater synchronization of business cycles, possibly already before the handover of monetary policy from national central banks to the ECB. On the other hand, these requirements have limited the scope for national fiscal policy to stabilize the economy. Similarly, the handover of monetary policy from the national central banks to the ECB implied a loss for individual EMU member countries of an important stabilization tool, which they could previously apply in response to asymmetric shocks. Both effects may have lowered business cycle synchronization before and after the events, respectively. There is, however, an argument stressing the "endogeneity of optimum currency-area criteria" (including the synchronization of business cycles) (Frankel and Rose, 1998): as a consequence of the events, transaction costs have declined, and this should spur the processes of greater trade and financial integration and hence greater business cycle comovements (cf. Imbs, 2004; Kose et al., 2003; Baxter and Kouparitsas, 2005).

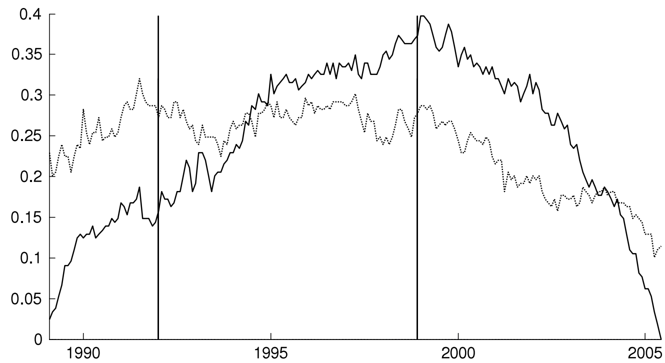
Given the ambiguity of these arguments, it remains to be tackled empirically whether and to what extent the two events have led to structural breaks and what has been the exact timing of structural breaks if there were any. Our empirical application is most closely related to Canova et al. (2009), who also investigate to what extent these two events have affected business cycles and their comovements in the euro area. On the basis of a panel VAR index model, the authors find some changes over time, but no evidence of clear structural breaks that coincide with the events.[14]

We apply the LM test procedure presented in Sections 4 and 5 to a monthly dataset used in Altissimo et al. (2007).[15] The dataset spans 1987:01–2007:06 and includes 209 macroeconomic variables from EMU member countries, the euro area as a whole, and a few external variables.[16] Series which were not already in seasonally adjusted form were seasonally adjusted by using the

14 Canova et al. (2009) assess changes around slightly different points in time, namely 1993:04, i.e. when the Maastricht treaty became effective, 1998:03, i.e. the date of the ECB creation, and 2002:01, i.e. the date of the euro changeover.

15 We are grateful to Giovanni Veronese for providing us with an updated version of that dataset.

16 The New Eurocoin indicator suggested in Altissimo et al. (2007) is constructed on the basis of 145 variables. The underlying dataset is larger. We use a subset of this larger dataset to obtain a balanced panel.



Note: The first vertical line indicates the signing date of the Maastricht treaty and the second vertical line marks the starting date of the EMU. Dotted line: LM statistic based on constant variances. Solid line: LM statistic that assumes a break in variances.

Fig. 2. Relative rejection frequencies (EMU data).

Census X12 procedure. Non-stationary series were transformed to stationary series as in Altissimo et al. (2007). Variables such as inflation and interest rates enter in levels. Therefore, there is no need to consider an additional transformation of the data as in the previous application. Outliers were removed as before. Unlike in the previous application, results barely depend on whether data were outlier adjusted or not. Therefore we only present results based on the adjusted dataset. Finally, as before, the series were de-meaned and divided by their standard deviations. For details on the data and the transformations, see Altissimo et al. (2007).

On the basis of the entire dataset and the IC_{p1} criterion of Bai and Ng (2002), r is estimated to be 9. We also split the dataset into three subsamples, pre-Maastricht, post-Maastricht and pre-EMU, and post-EMU. The IC_{p1} criterion selects $r = 3$ for the first, $r = 4$ for the second, and $r = 5$ for the third subsample, which is perhaps a first indication of a structural break. The autoregressive order of the idiosyncratic components is, again, set to 6, and the lag length for the Newey–West correction to 5.

The rejection rates of the test for structural breaks are 0.18 and 0.63 for the Maastricht treaty and 0.40 and 0.60 for the changeover to a single monetary policy in the euro area when the tests are based on the LM and HAC test procedures, respectively.

Have linkages become tighter or looser? We compare the commonality of the pre-Maastricht, post-Maastricht and pre-EMU and the post-EMU periods and find no major change between the first and the second period when nine factors explain 53.7% and 53.8% of the total variance, respectively. By contrast, the commonality increases to 55.7% in the third period which also supports our finding of a break being more likely in 1999:01 than in 1992:02.

We can, again, assess whether the breaks have occurred only at the dates of the two specific events or before or after these dates. As shown in Fig. 2, the heteroskedasticity-robust version of the test indicates that the rejection rate is indeed highest (at 0.40) in 1999:01. The unknown break point LM test suggests a rejection rate of 45% of the variables at any point in time. On average, over all variables, the break point is most likely to have occurred in 1995:02; the dispersion is, however, large. One possible interpretation is that reforms and other public measures in the run-up of EMU may have altered comovements. Also, EMU has been anticipated and private agents may have adjusted their behaviour prior to the event. A third explanation is that the mid-1990s are also associated with a general worldwide acceleration of globalization, which may have tightened linkages between countries. Finally, as in the previous application, we find evidence for considerable heteroskedasticity in the idiosyncratic components as indicated by a marked difference between the solid and the dotted line in Fig. 2.

Table 8
Tests for specific variables.

Country	Maastricht	EMU	# variables
DEU	0.14	0.31	42.
BEL	0.13	0.19	16.
ESP	0.25	0.67	24.
FRA	0.03	0.36	33.
ITA	0.26	0.48	27.
NLD	0.24	0.38	21.
Variables			
Ind. prod.	0.24	0.31	62.
Inflation	0.21	0.44	43.
Mon. and fin. var.	0.15	0.53	59.
Labor markets	0.17	0.39	23.
Surveys	0.05	0.23	22.

Note: This table presents the rejection frequencies for various groups of variables. The last column presents the number of variables in the group.

Next, we investigate whether the events have affected certain countries more than others. We have also formed groups of variables with similar economic content[17] and examine whether certain groups of variables have experienced structural breaks in the loadings while the loadings of other variables' groups have remained stable. Table 8 shows the rejection rates for individual countries. We only consider countries for which more than 10 variables were included in the dataset. Rejection rates are relatively high for both events for Spain and Italy, which are the countries with the lowest initial (1992) incomes[18] and the highest inflation and long-term interest rates[19] of the countries considered and, hence, the greatest needs to converge. Italy's public debt was, in addition, quite elevated, compared to the public debt of other countries.[20] Table 8 also reports rejection rates for groups of variables. As for the overall tests, rejections rates for all groups are higher for stage 3 of EMU than for Maastricht. At the date of the changeover to a single monetary policy in the euro area, rejection rates are relatively high for inflation as well as monetary and financial variables. After all, the changeover to a single monetary policy conducted by the ECB is a monetary event, and this result may therefore not be surprising. Maastricht has mainly caused breaks in the loadings of industrial production series. Unlike Canova et al. (2009), we have detected clear structural breaks in lots of variables at the two dates. The fact that their dataset does not include nominal variables may explain this difference between our finding and theirs.

7. Conclusions

Analyzing datasets with a large number of variables and time periods involves a severe risk that some of the model parameters

17 "Industrial production" includes, besides industrial production, also retail sales, orders, export, imports, inventories, and car registrations. The "Inflation" group summarizes PPI as well as export and import price inflation. "Monetary and financial variables" contain interest rates, monetary aggregates, exchange rates, and stock prices. "Labor market" summarizes employment variables and wages as well as unit labor costs. Finally, survey expectations form the group "Surveys".

18 GDP per capita amounted to 25,536 and 21,103 US$ for Italy and Spain in 1992 and to 27,725, 26,608, 27,116, 28,168 US$ for Germany, France, Belgium, and the Netherlands, respectively, according to The Conference Board and Groningen Growth and Development Centre, Total Economy Database, January 2008.

19 In 1992, year-on-year CPI inflation was at 5.3% and 5.9% in Italy and Spain and at 5.1%, 2.4%, 2.4%, 3.2% in Germany, France, Belgium, and the Netherlands, respectively. In 1992, the long-term interest rates were at 13.3% and 11.7% for Italy and Spain and at 7.9%, 8.6%, 8.7% and 8.1% for Germany, France, Belgium, and the Netherlands, respectively. Source: Economic Outlook, OECD.

20 In 1992 the gross public debt as a percentage of GDP according to the Maastricht criterion was at 105.3% for Italy and at 45.9%, 42.1%, 38.8%, 128.5%, 77.4% for Spain, Germany, France, Belgium and the Netherlands, respectively. Source: Economic Outlook, OECD.

are subject to structural breaks. We show that structural breaks in the factor loadings may inflate the number of factors identified by the usual information criteria. Furthermore, we propose Chow-type tests for structural breaks in factor models. It is shown that under the assumptions of an approximate factor model and if the number of variables is sufficiently large, the estimation error of the common factors does not affect the asymptotic distribution of the Chow statistics. In other words, the PC estimator of the common factors is "super-consistent" with respect to the estimation of the factor loadings and, therefore, the usual Chow test can be applied to the factor model in a regression, where the unknown factors are replaced by principal components. We also show that the Andrews (1993) tests for a structural break with an unknown break date can be used in a factor model. Furthermore, these tests can be generalized to dynamic factor models by adopting a GLS version of the test. This approach assumes a finite order autoregressive process for the idiosyncratic components, whereas no specific dynamic process needs to be specified for the common factors. Our Monte Carlo simulations suggest that the LM version outperforms the other variants of the test.

The LM test procedures are applied to two different settings. Our first empirical application uses a large US macroeconomic dataset provided by Stock and Watson (2005). We have tested whether the so called Great Moderation in the US (assuming the first quarter of 1984 as the starting date) coincides with structural breaks in the factor loadings. A lot of the attention of researchers and policy makers has recently been directed to the Great Moderation, and there is still some controversy about the sources ("good luck" versus structural changes including "good policy"). We find evidence of "dramatic changes" in the economy, reflected in significant breaks in the factor loadings, in the early 1980s. By testing for breaks in the loadings of individual variables such as the Federal funds rate, inventories, industrial production in the durable and non-durable sectors, personal consumption expenditure and financial variables, we can assess the underlying sources of the structural change. We find support for the hypothesis that not a single but various factors have played an important role. These factors are, according to our analysis, changes in the conduct of monetary policy and in inventory management as well as financial integration.

In the second application we take a large euro-area dataset used in Altissimo et al. (2007) to test whether structural breaks have occurred in the euro area around two major events: the signing of the Maastricht treaty in the second quarter of 1992 and the handover of monetary policy to the ECB (stage 3 of EMU) in the first quarter of 1999. This setting is particularly interesting, since these events may have altered comovements between variables as noted, and this will just be reflected in structural breaks in the factor loadings. We find evidence of structural breaks around both dates, with higher rejection rates for stage 3 of EMU than for the signing of the Maastricht treaty. It is not fully clear whether breaks have occurred exactly in 1999 or a few years before, possibly due to prior adjustments. Breaks finally seem to have occurred relatively frequently in the loadings of Spanish and Italian variables around the two events. The changeover to a single monetary policy in the euro area was associated with relatively frequent structural breaks in the loadings of nominal variables, whereas the signing of the Maastricht treaty seems to coincide with breaks in the factor loadings of industrial production series.

Acknowledgements

The views expressed in this paper do not necessarily reflect the views of the Deutsche Bundesbank. This paper was presented at the Workshop on Panel Methods and Open Economies, Frankfurt/Main, May 21, 2008 and at the International Conference on Factor Structures for Panel and Multivariate Time Series Data, Maastricht, September 19–20, 2008. The authors would like to thank Jörn Tenhofen, Jean-Pierre Urbain, Joakim Westerlund, and two anonymous referees for many helpful comments and suggestions.

Appendix

Preliminaries

Following Bai (2003) the true values of the factors and factor loadings are indicated by a zero, so the model is written as

$$y_{it} = \lambda_i^{0\prime} f_t^0 + \varepsilon_{it}.$$

Let $Y = (y_{it})$ be the $T \times N$ matrix of observations. The columns of the PC estimator $\widehat{F} = [\widehat{f}_1, \ldots, \widehat{f}_T]'$ are obtained as $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of the matrix YY' obeying the normalization $T^{-1}\widehat{F}'\widehat{F} = I_r$. The estimated matrix of factor loadings results as $\widehat{\Lambda} = [\widehat{\lambda}_1, \ldots, \widehat{\lambda}_T]' = Y'\widehat{F}/T$.

Since the factors are identified only up to an arbitrary rotation we will apply the normalization

$$\begin{aligned} y_{it} &= \lambda_i^{0\prime} H'^{-1} H' f_t^0 + \varepsilon_{it} \\ &= \lambda_i' f_t + \varepsilon_{it}, \end{aligned}$$

where $f_t = H' f_t^0$, $\lambda_i = H^{-1}\lambda_i^0$. Furthermore, let $F^0 = [f_1^0, \ldots, f_T^0]'$, $F = [f_1, \ldots, f_T]' = F^0 H$, $\Lambda^0 = [\lambda_1^0, \ldots, \lambda_T^0]'$, $\Lambda = [\lambda_1, \ldots, \lambda_T]' = \Lambda^0 H'^{-1}$. Following Bai (2003) we employ the rotation matrix

$$H = T\Lambda^{0\prime}\Lambda^0 F^{0\prime}\widehat{F}(\widehat{F}'YY'\widehat{F})^{-1}.$$

Using this normalization, Bai and Ng (2002) showed that the PC estimator $\widehat{F}$ is a consistent estimator for F.

Proof of Theorem 1

First, consider the LM statistic. Let $\varepsilon_i = [\varepsilon_{i1}, \ldots, \varepsilon_{iT}]'$. The residuals are obtained as $M_{\widehat{F}}\varepsilon_i$, where $\widehat{F} = [\widehat{f}_1, \ldots, \widehat{f}_T]'$ and $M_{\widehat{F}} = I_T - \widehat{F}(\widehat{F}'\widehat{F})^{-1}\widehat{F}' = I_T - T^{-1}\widehat{F}\widehat{F}'$. The individual LM statistic results as

$$s_i = \frac{\varepsilon_i' M_{\widehat{F}} \widehat{F}_2 (\widehat{F}_2' M_{\widehat{F}} \widehat{F}_2)^{-1} \widehat{F}_2' M_{\widehat{F}} \varepsilon_i}{\varepsilon_i' M_{\widehat{F}} \varepsilon_i / T}, \qquad (18)$$

where $\widehat{F}_2 = [0, \ldots, 0, \widehat{f}_{T_i^*+1}, \ldots, \widehat{f}_T]'$.

Using Lemma B.3 of Bai (2003) and Lemma A.1(ii) of Breitung and Tenhofen (2008) it follows that

$$T^{-1}\widehat{F}'\widehat{F} = I_r = T^{-1}F'F + O_p(\delta_{NT}^{-2}),$$

where $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$. The following Lemma shows that a similar result holds for $T^{-1}\widehat{F}_2'\widehat{F}_2$:

Lemma A.1. *Let* $F_2 = [0, \ldots, 0, f_{T_i^*+1}, \ldots, f_T]'$. *Under assumptions* A–F *of* Bai (2003) *we have*

(i) $\dfrac{1}{T}\widehat{F}_2'\widehat{F}_2 - \dfrac{1}{T}F_2'F_2 = \dfrac{1}{T}\widehat{F}_2'\widehat{F} - \dfrac{1}{T}F_2'F = O_p(\delta_{NT}^{-2})$

(ii) $\dfrac{1}{T}(\widehat{F}_2 - F_2)'\varepsilon_i = O_p(\delta_{NT}^{-2})$.

Proof. (i) Since the upper block of F_2 is a matrix of zeros we have $F_2'F = F_2'F_2$ and $\widehat{F}_2'\widehat{F} = \widehat{F}_2'\widehat{F}_2$. Consider

$$\begin{aligned} &\frac{1}{T}\left(\widehat{F}_2'\widehat{F}_2 - F_2'F_2\right) \\ &= \frac{1}{T}(\widehat{F} - F)'F_2 + \frac{1}{T}F'(\widehat{F}_2 - F_2) + \frac{1}{T}(\widehat{F} - F)'(\widehat{F}_2 - F_2) \\ &= I + II + III. \end{aligned}$$

Following Bai (2003) we start from the representation

$$\widehat{f}_t - f_t = \frac{1}{NT} V_{NT}^{-1} \left(\widehat{F}' F \Lambda' \varepsilon_{\cdot t} + \widehat{F}' \varepsilon \Lambda f_t + \widehat{F}' \varepsilon \varepsilon_{\cdot t} \right),$$

where $\varepsilon_{\cdot t} = [\varepsilon_{1t}, \ldots, \varepsilon_{Nt}]'$, $\varepsilon = [\varepsilon_{\cdot 1}, \ldots, \varepsilon_{\cdot T}]$, and V_{NT} is an $r \times r$ diagonal matrix of the r largest eigenvalues of $(NT)^{-1}YY'$. We first analyze

$$\frac{1}{T}(\widehat{F} - F)' F_2 = \frac{1}{NT^2} V_{NT}^{-1} \left(\widehat{F}' F \Lambda' \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} f_t' + \widehat{F}' \varepsilon \Lambda \sum_{t=T_i^*+1}^{T} f_t f_t' + \widehat{F}' \varepsilon \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} f_t' \right) = a + b + c.$$

From Assumption F(2) of Bai (2003) it follows that

$$\Lambda' \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} f_t' = O_p(\sqrt{NT}),$$

and $T^{-1}\widehat{F}'F - T^{-1}F'F = T^{-1}(\widehat{F} - F)'F = O_p(\delta_{NT}^{-2})$ (cf. Bai, Lemma A.2). Thus, we obtain

$$a = V_{NT}^{-1}(T^{-1}\widehat{F}'F) \left(\frac{1}{\sqrt{NT}} \Lambda' \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} f_t' \right) \frac{1}{\sqrt{NT}} = O_p\left(\frac{1}{\sqrt{NT}} \right).$$

Next consider

$$\Lambda' \varepsilon' \widehat{F} = \Lambda' \sum_{t=1}^{T} \varepsilon_{\cdot t} f_t' + \Lambda' \sum_{t=1}^{T} \varepsilon_{\cdot t} (\widehat{f}_t - f_t)'.$$

As shown by Bai (2003, p. 160),

$$\frac{1}{NT} \Lambda' \sum_{t=1}^{T} \varepsilon_{\cdot t} f_t' = O_p\left(\frac{1}{\sqrt{NT}} \right)$$

$$\frac{1}{NT} \Lambda' \sum_{t=1}^{T} \varepsilon_{\cdot t} (\widehat{f}_t - f_t)' = O_p\left(\frac{1}{\delta_{NT}\sqrt{N}} \right).$$

Using $T^{-1}F_2'F_2 = O_p(1)$ we obtain

$$b = V_{NT}^{-1} \left(\frac{1}{NT} \widehat{F}' \varepsilon \Lambda \right) \left(\frac{1}{T} \sum_{t=T_i^*+1}^{T} f_t f_t' \right) = \left[O_p\left(\frac{1}{\sqrt{NT}} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}\sqrt{N}} \right) \right] O_p(1).$$

As in Bai (2003, p. 164f), we obtain for the remaining term

$$\frac{1}{NT^2} \widehat{F}' \varepsilon \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} f_t' = \frac{1}{NT^2} \sum_{s=1}^{T} \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot s}' \varepsilon_{\cdot t} \widehat{f}_s f_t'$$
$$= \frac{1}{T^2} \sum_{s=1}^{T} \sum_{t=T_i^*+1}^{T} \widehat{f}_s f_t' \zeta_{NT}(s,t) + \frac{1}{T^2} \sum_{s=1}^{T} \sum_{t=T_i^*+1}^{T} \widehat{f}_s f_t' \gamma_N(s,t)$$
$$= O_p\left(\frac{1}{\delta_{NT}\sqrt{T}} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}} \right)$$

where

$$\zeta_N(s,t) = \varepsilon_{\cdot s}' \varepsilon_{\cdot t}/N - \gamma_N(s,t)$$
$$\gamma_N(s,t) = E(\varepsilon_{\cdot s}' \varepsilon_{\cdot t}/N).$$

Thus,

$$c = V_{NT}^{-1} \left[O_p\left(\frac{1}{\delta_{NT}\sqrt{T}} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}} \right) \right].$$

Collecting these results we obtain

$$I = a + b + c = O_p\left(\frac{1}{\sqrt{NT}} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{T}} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}} \right) = O_p\left(\frac{1}{\delta_{NT}^2} \right).$$

Using the same arguments it follows that $II = O_p(\delta_{NT}^{-2})$. Finally, following closely the proof of Theorem 1 in Bai and Ng (2002) we obtain $III = O_p(\delta_{NT}^{-2})$.

(ii) The proof is similar to (i). We therefore present the main steps only. Consider

$$\frac{1}{T} \sum_{t=T_i^*+1}^{T} (\widehat{f}_t - f_t) \varepsilon_{it} = \frac{1}{NT^2} V_{NT}^{-1} \left(\widehat{F}' F \Lambda' \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} \varepsilon_{it} + \widehat{F}' \varepsilon \Lambda \sum_{t=T_i^*+1}^{T} f_t \varepsilon_{it} + \widehat{F}' \varepsilon \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} \varepsilon_{it} \right) = a_i + b_i + c_i.$$

The term a_i results as

$$a_i = V_{NT}^{-1} \left(\frac{1}{T} \widehat{F}' F \right) \left(\frac{1}{NT} \Lambda' \sum_{t=T_i^*+1}^{T} \varepsilon_{\cdot t} \varepsilon_{it} \right) = O_p(1) \left[O_p\left(\frac{1}{\sqrt{NT}} \right) + O_p\left(\frac{1}{N} \right) \right]$$

(cf. Bai, 2003, B.1). For the second term we obtain

$$b_i = V_{NT}^{-1} \left(\frac{1}{NT} \widehat{F}' \varepsilon \Lambda \right) \left(\frac{1}{T} \sum_{t=T_i^*+1}^{T} f_t \varepsilon_{it} \right) = O_p\left(\frac{1}{\delta_{NT}\sqrt{N}} \right) O_p\left(\frac{1}{\sqrt{T}} \right).$$

Finally we have

$$c_i = V_{NT}^{-1} \left(\frac{1}{T^2} \sum_{s=1}^{T} \sum_{t=T_i^*+1}^{T} f_s \varepsilon_{it} (N^{-1} \varepsilon_{\cdot s}' \varepsilon_{\cdot t}) \right) = O_p\left(\frac{1}{\delta_{NT}\sqrt{T}} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}} \right)$$

(cf. Bai, 2003, p. 163). Collecting these results we have

$$a_i + b_i + c_i = O_p\left(\frac{1}{\sqrt{NT}} \right) + O_p\left(\frac{1}{N} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{T}} \right) + O_p\left(\frac{1}{\delta_{NT}\sqrt{N}} \right) = O_p\left(\frac{1}{\delta_{NT}^2} \right). \quad \square$$

Using these results, we obtain

$$T^{-1}\widehat{F}_2' M_{\widehat{F}} \widehat{F}_2 = T^{-1} F_2' M_F F_2 + O_p(\delta_{NT}^{-2}),$$

where $M_F = I_T - F(F'F)^{-1}F'$. Using Lemma A.1(i) and (ii) and Lemma B.1 of Bai (2003) we obtain in a similar manner

$$\begin{aligned} T^{-1/2}\varepsilon_i' M_{\widehat{F}} \widehat{F}_2 &= T^{-1/2}\varepsilon_i' F_2 + T^{-1/2}\varepsilon_i'(\widehat{F}_2 - F_2) \\ &\quad - \left(T^{-1/2}\varepsilon_i'\widehat{F}\right)\left(\frac{1}{T}\widehat{F}'\widehat{F}_2\right) \\ &= T^{-1/2}\varepsilon_i' F_2 - \left(T^{-1/2}\varepsilon_i'\widehat{F}\right)\left(\frac{1}{T}F'F_2\right) \\ &\quad + O_p(\sqrt{T}/\delta_{NT}^2) \\ &= T^{-1/2}\varepsilon_i' M_F F_2 + O_p(\sqrt{T}/\delta_{NT}^2). \end{aligned}$$

Finally Eq. 10 of Bai and Ng (2002) implies

$$T^{-1}\varepsilon_i' M_{\widehat{F}}\varepsilon_i = T^{-1}\varepsilon_i' M_F \varepsilon_i + O_p(\delta_{NT}^{-2}).$$

From these results it follows that

$$\begin{aligned} s_i &= \frac{\varepsilon_i' M_F F_2 (F_2' M_F F_2)^{-1} F_2' M_F \varepsilon_i}{\varepsilon_i' M_F \varepsilon_i / T} + O_p(\sqrt{T}/\delta_{NT}^2) \\ &= s_i^0 + O_p(\sqrt{T}/\delta_{NT}^2). \end{aligned}$$

Note that s_i^0 is the LM statistic obtained from the (infeasible) regression that uses F instead of $\widehat{F}$. Under Assumption 2 s_i^0 has a χ^2 limiting distribution as $T \to \infty$.

To derive the limiting distribution of the Wald statistic w_i we first note that the only difference from the LM statistic is that the variance estimator in the denominator of (18) is computed by using the sum of squared residuals from a regression of $M_{\widehat{F}}\varepsilon_i$ on $M_{\widehat{F}}F_2$. Denote the resulting residual vector as $\widehat{\varepsilon}_i^*$. From standard regression theory it is well known that

$$\varepsilon_i' M_{\widehat{F}}\varepsilon_i = \widehat{\varepsilon}_i^{*\prime}\widehat{\varepsilon}_i^* + \varepsilon_i' M_{\widehat{F}}\widehat{F}_2(\widehat{F}_2' M_{\widehat{F}}\widehat{F}_2)^{-1}\widehat{F}_2' M_{\widehat{F}}\varepsilon_{it}.$$

Using the same results as were obtained for the LM statistic, we have

$$T^{-1}(\varepsilon_i' M_{\widehat{F}}\varepsilon_i - \widehat{\varepsilon}_i^{*\prime}\widehat{\varepsilon}_i^*) = T^{-1}\varepsilon_i' M_F F_2 (F_2' M_F F_2)^{-1} F_2' M_F \varepsilon_i + O_p(\delta_{NT}^{-2}).$$

The first term on the r.h.s. is $+O_p(T^{-1})$ and therefore the difference between the variance estimators based on the restricted and unrestricted models is positive and $O_p(T^{-1})$. Therefore, $w_i \geq s_i$ and $w_i = s_i^0 + O_p(T^{-1}) + O_p(\sqrt{T}/\delta_{NT}^2)$.

Using a first-order Taylor expansion, we obtain for the LR statistic

$$\begin{aligned} T[\log(S_{0i}) - \log(S_{1i} + S_{2i})] &= \frac{S_{0i} - S_{1i} - S_{2i}}{(S_{1i} + S_{2i})/T} + O_p(T^{-1}) \\ &= \frac{\varepsilon_i' M_{\widehat{F}}\widehat{F}_2(\widehat{F}_2' M_{\widehat{F}}\widehat{F}_2)^{-1}\widehat{F}_2' M_{\widehat{F}}\varepsilon_i}{\varepsilon_i' M_{\widehat{F}}\varepsilon_i/T + O_p(T^{-1})} + O_p(T^{-1}) \\ &= s_i + O_p(T^{-1}). \end{aligned}$$

Therefore,

$$lr_i = s_i^0 + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right).$$

Proof of Theorem 2

Let $F_2^\tau = [0, \ldots, 0, f_{[\tau T]+1}, \ldots, f_T]'$ and write the LM statistic for a structural break at $T_i^* = [\tau T]$ as

$$s_i^\tau = \frac{\varepsilon_i' M_{\widehat{F}}\widehat{F}_2^\tau (\widehat{F}_2^{\tau\prime} M_{\widehat{F}}\widehat{F}_2^\tau)^{-1}\widehat{F}_2^{\tau\prime} M_{\widehat{F}}\varepsilon_i}{\varepsilon_i' M_{\widehat{F}}\varepsilon_i/T}. \tag{19}$$

Using Lemma A.1 and Assumption 3 we obtain

$$\begin{aligned} T^{-1/2}\widehat{F}_2^{\tau\prime}\varepsilon_i &= T^{-1/2}\sum_{t=[\tau T]+1}^{T}\varepsilon_{it} f_t + O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right) \\ &\xrightarrow{d} \sigma_i[W_i(1) - W_i(\tau)] \end{aligned}$$

$$T^{-1}\widehat{F}_2^{\tau\prime}\widehat{F}_2^\tau = T^{-1}\sum_{t=[\tau T]+1}^{T} f_t f_t' + O_p\left(\frac{1}{\delta_{NT}^2}\right) \xrightarrow{p} (1-\tau) I_r$$

$$T^{-1}\widehat{F}'\widehat{F} = T^{-1}\sum_{t=1}^{T} f_t f_t' + O_p\left(\frac{1}{\delta_{NT}^2}\right) \xrightarrow{p} I_r$$

where we have used the fact that $\Omega_F = I_r$ in Assumption 3 due to the normalization of the factors. It follows that

$$\begin{aligned} T^{-1/2}\widehat{F}_2^{\tau\prime} M_{\widehat{F}}\varepsilon_i &= T^{-1/2}\varepsilon_i'\widehat{F}_2^\tau - T^{-1/2}\varepsilon_i'\widehat{F}(\widehat{F}'\widehat{F})^{-1}T^{-1/2}\widehat{F}'\widehat{F}_2^\tau \\ &\xrightarrow{d} \sigma_i[\tau W_i(1) - W_i(\tau)] \end{aligned}$$

$$T^{-1}\widehat{F}_2^{\tau\prime} M_{\widehat{F}}\widehat{F}_2^\tau \xrightarrow{p} \tau(1-\tau) I_r.$$

Using these results, the limiting distribution of s_i^τ is obtained as

$$\begin{aligned} s_i^\tau &= \mathcal{Z}_{T,i}^{\tau\prime}\mathcal{Z}_{T,i}^\tau + O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right) \\ &\xrightarrow{d} \frac{[\tau W_i(1) - W_i(\tau)]'[\tau W_i(1) - W_i(\tau)]}{\tau(1-\tau)}, \end{aligned}$$

where

$$\mathcal{Z}_{T,i}^\tau = \frac{1}{\sqrt{\tau(1-\tau)}}\left(\frac{\tau}{\sigma_i\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it} f_t - \frac{1}{\sigma_i\sqrt{T}}\sum_{t=1}^{[\tau T]}\varepsilon_{it} f_t\right).$$

Note that the latter sequence can be represented as a vector of (weighted) partial sums of the form $\mathcal{Z}_{T,i}^\tau = T^{-1/2}\sum_{t=1}^{T} w(f_t,\tau)\varepsilon_{it}$, where the weights are functions of F_t and τ.

By letting $\mathcal{Z}_{T,i}(\tau) = \mathcal{Z}_{T,i}^{\overline{\tau}}$ with $\overline{\tau} = [\tau T]/T$ we embed the observed sequence in the space $D[\tau_0, 1-\tau_0]$. The finite dimensional distributions for all choices of $\tau_1, \ldots, \tau_n$ with $\tau_\kappa \in [\tau_0, 1-\tau_0]$ converge to the corresponding finite dimensional distributions of $[\tau W_i(1) - W_i(\tau)]/\sqrt{\tau(1-\tau)}$. In order to obtain weak convergence in $D[\tau_0, 1-\tau_0]$, the tightness of the sequence $\mathcal{Z}_{T,i}(\tau)$ has to be shown. To this end we closely follow Billingsley (1968, Section 6). Since $\{F_t\}$ is independent of $\{\varepsilon_{it}\}$ we may condition on $\{F_t\}$ and obtain a sum of independent random variables with uniformly bounded second moments for which the inequality (16.5) in Billingsley (1968) remains valid modulo a constant.

Having established weak convergence of $\mathcal{Z}_{T,i}(\tau)$ to $[\tau W_i(1) - W_i(\tau)]/\sqrt{\tau(1-\tau)}$ we invoke the continuous mapping theorem together with standard results on weak convergence yielding

$$s_i^\tau \Rightarrow \frac{[\tau W_i(1) - W_i(\tau)]'[\tau W_i(1) - W_i(\tau)]}{\tau(1-\tau)}.$$

Another application of the continuous mapping theorem yields

$$\sup_{\tau\in[\tau_0,1-\tau_0]} s_i^\tau \Rightarrow \sup_{\tau\in[\tau_0,1-\tau_0]}\left\{\frac{[\tau W_i(1) - W_i(\tau)]'[\tau W_i(1) - W_i(\tau)]}{\tau(1-\tau)}\right\}.$$

Since the asymptotic distribution is the same for all i we drop the index i in Theorem 2.

Proof of Theorem 3

To derive the limiting distribution of the feasible GLS version of the LM test, we make use of the following two lemmas:

Lemma A.2. *It holds for any fixed m and $k \leq m$ that*

$$\text{(i) } T^{-1}\sum_{t=m+1}^{T}(\widehat{f}_t - f_t)f'_{t-k} = O_p(\delta_{NT}^{-2}),$$

$$T^{-1}\sum_{t=m+1}^{T}(\widehat{f}_t - f_t)\widehat{f}'_{t-k} = O_p(\delta_{NT}^{-2})$$

$$\text{(ii) } T^{-1}\sum_{t=m+1}^{T}\widehat{f}_t\widehat{f}'_{t-k} = T^{-1}\sum_{t=m+1}^{T}f_tf'_{t-k} + O_p(\delta_{NT}^{-2})$$

$$\text{(iii) } T^{-1}\sum_{t=m+1}^{T}(\widehat{f}_t - f_t)u_{i,t-k} = O_p(\delta_{NT}^{-2}).$$

Proof. For $m = p_i$ these results are shown in Breitung and Tenhofen (2008, Lemma A.1). For $m = T_i^*$ the proof can be modified straightforwardly according to Lemma A.1. □

Lemma A.3. *Let $\varrho^{(i)} = [\varrho_{i,1}, \ldots, \varrho_{i,p_i}]'$ and $\widehat{\varrho}^{(i)} = [\widehat{\varrho}_{i,1}, \ldots, \widehat{\varrho}_{i,p_i}]'$ denote the least squares estimates from (15). Under Assumption 1 we have as $(N, T) \to \infty$*

$$\widehat{\varrho}^{(i)} = \varrho^{(i)} + O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}).$$

Proof. The proof is given in Breitung and Tenhofen (2008, Lemma 1). □

To simplify the notation, we focus on the AR(1) model $u_{it} = \varrho_i u_{i,t-1} + \varepsilon_{it}$. The extension to AR($p$) models is straightforward but implies a considerable additional notational burden.

The LM statistic can be written as

$$\widetilde{s}_i = \psi'_{i,21}\Psi_{i,22}^{-1}\psi_{i,21}/\psi_{i,11},$$

where

$$\psi_{i,21} = \widetilde{G}'_{i,2}M_{\widetilde{G}_i}\widetilde{\varepsilon}_i$$
$$\Psi_{i,22} = \widetilde{G}'_{i,2}M_{\widetilde{G}_i}\widetilde{G}_{i,2}$$
$$\psi_{i,11} = \widetilde{\varepsilon}'_iM_{\widetilde{G}_i}\widetilde{\varepsilon}_i/T,$$

and

$$\widetilde{G}_i = [\widehat{f}_2 - \widehat{\varrho}_i\widehat{f}_1, \ldots, \widehat{f}_t - \widehat{\varrho}_i\widehat{f}_{T-1}]'$$
$$\widetilde{G}_{i,2} = [0, \ldots, 0, \widehat{f}_{T_i^*+1} - \widehat{\varrho}_i\widehat{f}_{T_i^*}, \ldots, \widehat{f}_t - \widehat{\varrho}_i\widehat{f}_{T-1}]'$$
$$\widetilde{\varepsilon}_i = [u_2 - \widehat{\varrho}_iu_1, \ldots, u_T - \widehat{\varrho}_iu_{T-1}]$$
$$M_{\widetilde{G}_i} = I_{T-1} - \widetilde{G}_i(\widetilde{G}'_i\widetilde{G}_i)^{-1}\widetilde{G}'_i.$$

Using Lemma A.2(ii) and Lemma A.3, we obtain

$$\begin{aligned}\frac{1}{T}\widetilde{G}'_{i,2}\widetilde{G}_i &= \frac{1}{T}\sum_{t=T_i^*+1}^{T}(\widehat{f}_t - \widehat{\varrho}_i\widehat{f}_{t-1})(\widehat{f}_t - \widehat{\varrho}_i\widehat{f}_{t-1})' \\ &= \frac{1}{T}\sum_{t=T_i^*+1}^{T}(\widehat{f}_t - \varrho_i\widehat{f}_{t-1})(\widehat{f}_t - \varrho_i\widehat{f}_{t-1})' \\ &\quad + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\delta_{NT}^2}\right) \\ &= \frac{1}{T}\sum_{t=T_i^*+1}^{T}(f_t - \varrho_if_{t-1})(f_t - \varrho_if_{t-1})' \\ &\quad + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\delta_{NT}^2}\right),\end{aligned}$$

$$\begin{aligned}\frac{1}{T}\widetilde{G}'_i\widetilde{G}_i &= \frac{1}{T}\sum_{t=2}^{T}(f_t - \varrho_if_{t-1})(f_t - \varrho_if_{t-1})' \\ &\quad + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\delta_{NT}^2}\right),\end{aligned}$$

and

$$\begin{aligned}\frac{1}{\sqrt{T}}\widetilde{G}'_i\widetilde{\varepsilon}_i &= \frac{1}{\sqrt{T}}\sum_{t=2}^{T}(\widehat{f}_t - \widehat{\varrho}_i\widehat{f}_{t-1})(u_{it} - \widehat{\varrho}_iu_{i,t-1}) \\ &= \frac{1}{\sqrt{T}}\sum_{t=2}^{T}[\widehat{f}_t - \varrho_i\widehat{f}_{t-1} + (\varrho_i - \widehat{\varrho}_i)\widehat{f}_{t-1}] \\ &\quad \times (\varepsilon_{it} + (\varrho_i - \widehat{\varrho}_i)u_{t-1}) \\ &= \frac{1}{\sqrt{T}}\sum_{t=2}^{T}(\widehat{f}_t - \varrho_i\widehat{f}_{t-1})\varepsilon_{it} + A + B_1 + B_2 + C,\end{aligned}$$

where $\widehat{u}_{\cdot t} = [\widehat{u}_{1t}, \ldots, \widehat{u}_{Nt}]'$,

$$\begin{aligned}A &= \frac{1}{\sqrt{T}}\sum_{t=2}^{T}(\varrho_i - \widehat{\varrho}_i)\widehat{f}_{t-1}\varepsilon_{it} \\ &= \sqrt{T}(\varrho_i - \widehat{\varrho}_i)\left(\frac{1}{T}\sum_{t=2}^{T}f_{t-1}\varepsilon_{it} + (\widehat{f}_{t-1} - f_{t-1})\varepsilon_{it}\right) \\ &= \sqrt{T}(\varrho_i - \widehat{\varrho}_i)\left(\frac{1}{T}\sum_{t=2}^{T}f_{t-1}\varepsilon_{it} + O_p(\delta_{NT}^{-2})\right) \\ &= O_p(1)\left[O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2})\right].\end{aligned}$$

Next, using Lemma A.2(iii) we obtain

$$\begin{aligned}B_1 &= \frac{1}{\sqrt{T}}(\varrho_i - \widehat{\varrho}_i)\sum_{t=2}^{T}\widehat{f}_tu_{t-1} \\ &= \sqrt{T}(\varrho_i - \widehat{\varrho}_i)\left[\frac{1}{T}\sum_{t=2}^{T}f_tu_{i,t-1} + O_p(\delta_{NT}^{-2})\right] \\ &= O_p(1)\left[O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2})\right]\end{aligned}$$

and, similarly,

$$\begin{aligned}B_2 &= -\frac{1}{\sqrt{T}}\varrho_i(\varrho_i - \widehat{\varrho}_i)\sum_{t=2}^{T}\widehat{f}_{t-1}u_{t-1} \\ &= O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}).\end{aligned}$$

For the last term we have

$$\begin{aligned}C &= \frac{1}{\sqrt{T}}(\varrho_i - \widehat{\varrho}_i)^2\sum_{t=2}^{T}\widehat{f}_{t-1}u_{t-1} \\ &= \left[O_p(T^{-1}) + O_p(T^{-1/2}\delta_{NT}^{-2}) + O_p(\delta_{NT}^{-4})\right] \\ &\quad \times \left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\widehat{f}_{t-1}u_{t-1}\right) \\ &= O_p(\delta_{NT}^{-2}).\end{aligned}$$

Using Lemma A.2(iii) it follows that

$$\frac{1}{\sqrt{T}}\sum_{t=2}^{T}(f_t - \varrho_if_{t-1})\varepsilon_{it} + O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right).$$

Collecting these results gives

$$T^{-1/2}\psi_{1,21} = \frac{1}{\sqrt{T}}G'_{i,2}M_{G_i}\varepsilon_i + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right),$$

where

$$G_i = [F_2 - \varrho_iF_1, \ldots, f_t - \varrho_if_{T-1}]'$$
$$G_{i,2} = [0, \ldots, 0, f_{T_i^*+1} - \varrho_if_{T_i^*}, \ldots, f_t - \varrho_if_{T-1}]'$$
$$M_{G_i} = I_{T-1} - G_i(G'_iG_i)^{-1}G'_i.$$

Furthermore

$$\frac{1}{T}\Psi_{i,22} = \frac{1}{T}G'_{i,2}M_{G_i}G_{i,2} + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\delta^2_{NT}}\right)$$

and

$$\frac{1}{T}\psi_{i,11} = \frac{1}{T}\varepsilon'_i M_{G_i}\varepsilon_i + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\delta^2_{NT}}\right).$$

It follows that

$$\widetilde{s}_i = \widetilde{s}^0_i + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{\sqrt{T}}{\delta^2_{NT}}\right),$$

where

$$\widetilde{s}^0_i = \frac{\varepsilon'_i M_{G_i}G_{i,2}(G'_{i,2}M_{G_i}G_{i,2})^{-1}G'_{i,2}M_{G_i}\varepsilon_i}{\varepsilon'_i M_{G_i}\varepsilon_i/T}.$$

Under Assumption 2 we therefore have $\widetilde{s}_i \xrightarrow{d} \chi^2_{(r)}$ as $N, T \to \infty$ and $\sqrt{T}/N \to 0$.

References

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856.
Andrews, D.W.K., Lee, I., Ploberger, W., 1996. Optimal change point tests for normal linear regression. Journal of Econometrics 70, 9–38.
Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.
Altissimo, F., Forni, M., Lippi, M., Veronese, G., Cristadoro, R., 2007. New Eurocoin: tracking economic growth in real time. Bank of Italy Working Paper 631.
Amengual, D., Watson, M.W., 2007. Consistent estimation of the number of dynamic factors in a large N and T panel. Journal of Business & Economic Statistics 25, 91–96.
Bai, J., 2003. Inferential theory for factor models of large dimensions. Econometrica 71, 135–172.
Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70, 191–221.
Bai, J., Ng, S., 2007. Determining the number of primitive shocks in factor models. Journal of Business and Economic Statistics 25, 52–60.
Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47–78.
Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. Journal of Applied Econometrics 18, 1–22.
Banerjee, A., Marcellino, M., 2008. Forecasting macroeconomic variables using diffusion indexes in short samples with structural change. CEPR Working Paper 6706.
Baxter, M., Kouparitsas, M.A., 2005. Determinants of business cycle comovement: a robust analysis. Journal of Monetary Economics 52 (1), 113–157.
Benati, L., Mumtaz, H., 2007. US evolving macroeconomic dynamics—a structural investigation. ECB Working Paper 746.
Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.
Breitung, J., Pesaran, M., 2008. Unit roots and cointegration in panels. In: Matyas, L., Sevestre, P. (Eds.), The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice. Kluwer Academic Publishers, pp. 279–322.
Breitung, J., Tenhofen, J., 2008. GLS estimation of dynamic factor models. Working Paper, University of Bonn.
Canova, F., Ciccarelli, M., Ortega, E., 2009. Do institutional changes affect business cycles? Evidence from Europe, Mimeo, Universitat Pompeu Fabra, Barcelona.
Chamberlain, G., Rothschild, M., 1983. Arbitrage, factor structure and mean–variance analysis in large asset markets. Econometrica 51, 1305–1324.
Clarida, R., Galí, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. The Quarterly Journal of Economics 115, 147–180.
Davis, S.J., Kahn, J.A., 2008. Interpreting the great moderation: changes in the volatility of economic activity at the macro and micro levels. Journal of Economic Perspectives 22, 155–180.
Del Negro, M., Otrok, C., 2008. Dynamic factor models with time-varying parameters: measuring changes in international business cycles. Revised Version of Federal Reserve Bank of New York Staff Report 326 (2005).
Eickmeier, S., Lemke, W., Marcellino, M., 2009. Classical time-varying FAVAR models—estimation, forecasting and structural analysis. Mimeo, Deutsche Bundesbank.
Frankel, J.A., Rose, A.K., 1998. The endogeneity of the optimum currency area criteria. Economic Journal 108, 1009–1025.
Galí, J., Gambetti, L., 2008. On the sources of the great moderation. CEPR Discussion Paper 6632.
Giordani, P., 2007. Discussion of "forecasting in dynamic factor models subject to structural instability" by James Stock and Mark Watson, December 2007. http://www.ecb.int/events/pdf/conferences/ftworkshop2007/Giordani.pdf.
Imbs, J., 2004. Trade, finance, specialization and synchronization. Review of Economics and Statistics 86, 723–734.
IMF, 2008. World Economic Outlook: Housing and the Business Cycle (Chapter 3).
Kim, C., Nelson, C.R., 1999. Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. Review of Economics and Statistics 81, 608–616.
Kose, A.M., Prasad, E.S., Terrones, M.E., 2003. How does globalization affect the synchronization of business cycles?. American Economic Review 93, 57–62.
McConnell, M., Perez-Quiros, G., 2000. Output fluctuations in the United States: what has changed since the early 1980s?. American Economic Review 90, 1464–1476.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Perron, P., 2006. Dealing with structural breaks. In: Mills, T.C., Patterson, K. (Eds.), Econometric Theory. In: Palgrave Handbook of Econometrics, vol. 1. Palgrave Macmillan, New York, pp. 278–352.
Stock, J.H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.
Stock, J.H., Watson, M.W., 2003. Has the business cycle changed? Evidence and explanations. Forthcoming FRB Kansas City Symposium, Jackson Hole, Wyoming, August 28–30.
Stock, J.H., Watson, M.W., 2005. Implications of dynamic factor models for VAR analysis. NBER Working Paper No. 11467.
Stock, J.H., Watson, M.W., 2008. Forecasting in dynamic factor models subject to structural instability. In: Castle, J., Shephard, N. (Eds.), The Methodology and Practice of Econometrics, A Festschrift in Honour of Professor David F. Hendry. Oxford University Press, Oxford.
Vogelsang, T.J., 1999. Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. Journal of Econometrics 88, 283–299.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.