



University of Pavia

Department of Economics and Management

Master Programme in Economics, Finance and International Integration

**Testing Structural Changes in Large-Dimensional Factor
Models: An Application to Russian Stock Market Data**

**Test dei cambiamenti strutturali nei modelli fattoriali ad alta
dimensionalità: un'applicazione al mercato azionario russo**

Thesis supervisor: Prof. Lorenzo Trapani

Student: Olga Chidyakina

Matr. No: 491434

Academic year 2025/2026

Contents

List of Figures	iv
List of Tables	v
List of Abbreviations	vii
List of Symbols	viii
Abstract	xi
1 Introduction	1
1.1 Motivation	1
1.2 Literature review	3
1.2.1 Large-dimensional factor models: general framework and motivation	3
1.2.2 Estimation of factors and inferential issues	5
1.2.3 Determining the number of factors	8
1.2.4 Structural breaks in factor models	14
1.2.5 Position of the present thesis in the literature	24
1.3 Research objectives and structure of the thesis	26
2 Factor models: theoretical framework	28
2.1 Factor model setup	28
2.2 Principal component estimation	29
2.3 Determining the number of factors	31

3	Structural breaks in factor models	35
3.1	Structural instability in factor models	35
3.2	Breaks in the factor loading matrix	36
3.3	Likelihood ratio approach	37
3.4	Limiting distribution of the sup-LR statistic under the null hypothesis	40
4	Simulation study	43
4.1	Data generating process	43
4.2	Estimation and testing procedure	45
4.3	Illustrative single-run results	47
4.3.1	Scenario 1. No-break benchmark	47
4.3.2	Scenario 2. Baseline break	49
4.3.3	Scenario 3. Moderate break	51
4.3.4	Scenario 4. Higher idiosyncratic noise with break	52
4.3.5	Preliminary comparison of the four illustrative cases	55
4.4	Monte Carlo evidence	55
4.4.1	Factor-number selection	56
4.4.2	Rejection frequencies and test performance	56
4.4.3	Accuracy of break-date estimation	57
4.4.4	Overall discussion of the Monte Carlo results	58
4.5	Discussion	59
5	Empirical application to the Russian stock market	61
5.1	Description of the data and preprocessing	61
5.2	Baseline empirical results	64
5.3	Robustness checks	67
5.3.1	Robustness to trimming	67
5.3.2	Robustness to the number of factors	68
5.3.3	Interpretation of the robustness results	69
5.4	Pre-break and post-break comparison	69

5.5	Empirical discussion	72
5.5.1	Statistical meaning of the detected break	72
5.5.2	Interpreting the estimated break date	73
5.5.3	Economic plausibility of the timing	74
5.5.4	Practical implications of the break	76
6	Conclusions	78
A	Data Set and Preprocessing Details	82
A.1	Empirical sample	82
A.2	Preprocessing	83
B	Additional Factor Selection and Empirical Results	85
B.1	Baseline factor selection	85
B.2	Baseline LR test result	86
B.3	Robustness to the trimming parameter	86
B.4	Robustness to the number of factors	87
B.5	Pre-break and post-break factor selection	87
B.6	Pre-break and post-break comparison	88
C	Supplementary Simulation Evidence	89
C.1	Simulation design	89
C.2	Illustrative single-run results	89
C.3	Factor-selection summary	90
C.4	Rejection frequencies	90
C.5	Break-date accuracy	91

List of Figures

4.1	Scree plot for the simulated no-break scenario	48
4.2	LR statistic in the simulated no-break scenario	49
4.3	Scree plot for an illustrative single run in the baseline break scenario	50
4.4	LR statistic in the simulated break scenario	51
4.5	Scree plot for an illustrative single run in the moderate break scenario	52
4.6	LR statistic in the simulated moderate break scenario	53
4.7	Scree plot for an illustrative single run in the noisier break scenario	54
4.8	LR statistic in the simulated noisier break scenario	54
5.1	MOEX Russia Index	62
5.2	Scree plot for the full sample	65
5.3	LR statistic over candidate break dates: full sample	66
5.4	Comparison of scree plots before and after the estimated break	70
5.5	MOEX Russia Index with estimated break date	75

List of Tables

4.1	Illustrative single-run results across the four simulation scenarios. True $r = 2$ in all scenarios	55
4.2	Simulation design	56
4.3	Factor selection and rejection frequencies in the Monte Carlo experiment	57
4.4	Accuracy of break-date estimation in break scenarios	58
5.1	Baseline empirical results for the Russian stock market data	67
5.2	Robustness of the baseline result to trimming parameter	68
5.3	Robustness of the baseline result to the number of factors	69
5.4	Comparison of the factor structure before and after the estimated break	71
A.1	Tickers included in the empirical sample	82
A.2	Summary of the empirical data set	83
B.1	Bai–Ng factor-selection results for the full sample	85
B.2	Baseline LR test result for the real-data application	86
B.3	Baseline LR decision table	86
B.4	Robustness of the LR result to the trimming parameter	87
B.5	Robustness of the LR result to the number of factors	87
B.6	Bai–Ng factor-selection results before and after the estimated break	88
B.7	Comparison of the factor structure before and after the estimated break	88
C.1	Monte Carlo simulation design	89
C.2	Illustrative single-run results across simulation scenarios	90
C.3	Monte Carlo factor-selection summary	90

C.4	Monte Carlo rejection frequencies	90
C.5	Monte Carlo break-date accuracy in break scenarios	91

List of Abbreviations

AIC	Akaike Information Criterion
APT	Arbitrage Pricing Theory
BIC	Bayesian Information Criterion
CV	Critical Value
DGP	Data Generating Process
EM	Expectation–Maximization
FAVAR	Factor-Augmented Vector Autoregression
HAC	Heteroskedasticity and Autocorrelation Consistent
IC	Information Criterion
IMOEX	Moscow Exchange Index
LM	Lagrange Multiplier
LR	Likelihood Ratio
MAE	Mean Absolute Error
ML	Maximum Likelihood
MOEX	Moscow Exchange
OLS	Ordinary Least Squares
PC	Principal Components / Panel Criterion, depending on context
PCA	Principal Component Analysis
QML	Quasi-Maximum Likelihood
SD	Standard Deviation
VAR	Vector Autoregression

List of Symbols

N	number of cross-sectional units / assets
T	number of time observations
i	cross-sectional index, $i = 1, \dots, N$
t	time index, $t = 1, \dots, T$
r	number of common factors
$\hat{r}$	estimated number of factors
k	candidate number of factors or candidate break date, depending on context
$k_{\max}$	maximum admissible number of factors
X_{it}, x_{it}	observed value or return of asset i at time t
x_t	$N \times 1$ vector of observations at time t
X_t	vector of observed variables at time t
X	$T \times N$ data matrix
F_t	vector of common factors at time t
f_t	vector of latent factors in the structural-break representation
F	matrix of common factors
λ_i	vector of factor loadings for unit i
$\lambda_{i1}, \lambda_{i2}$	pre-break and post-break loading vectors for unit i
Λ	factor loading matrix
Λ_1, Λ_2	pre-break and post-break loading matrices
e_{it}	idiosyncratic component for unit i at time t
e_t	$N \times 1$ vector of idiosyncratic errors at time t

e	matrix of idiosyncratic errors
C_{it}	common component of X_{it}
$\widehat{F}$	estimated factor matrix
$\widehat{\Lambda}$	estimated loading matrix
$\widehat{C}$	estimated common component matrix
$\widehat{C}_{it}$	estimated common component for unit i at time t
$V(k)$	residual variance when k factors are estimated
$g(N, T)$	penalty function in information or panel criteria
$IC(k)$	information criterion for choosing the number of factors
ICp_1, ICp_2, ICp_3	Bai–Ng information criteria
PCp_1, PCp_2, PCp_3	Bai–Ng panel criteria
C_{NT}	$\min\{\sqrt{N}, \sqrt{T}\}$
$\widehat{\sigma}^2$	estimate of the idiosyncratic variance
T_1	break date
$\widehat{T}_1$	estimated break date
π_1	true break fraction
π	generic break fraction
ε	trimming parameter
B, C	transformation matrices in the equivalent break representation
r_1, r_2	numbers of factors before and after the break in the equivalent representation
G	matrix of pseudo-factors
$\widehat{G}$	estimated pseudo-factor matrix
$\widehat{g}_t$	estimated pseudo-factor vector at time t
$\widehat{\Sigma}_1(k)$	estimated pre- k covariance matrix of factors
$\widehat{\Sigma}_2(k)$	estimated post- k covariance matrix of factors
$L(k)$	quasi-Gaussian log-likelihood for candidate break point k
L_0	no-break log-likelihood
$LR(k)$	likelihood ratio statistic evaluated at candidate break date k

$\sup LR$	supremum likelihood ratio statistic
H_0	null hypothesis of no structural break
H_1	alternative hypothesis of a structural break
$W(\pi)$	vector of Brownian motions
$W(\pi) - \pi W(1)$	Brownian bridge process
Ω	long-run covariance matrix
$\widehat{\Omega}$	HAC estimator of the long-run covariance matrix
η_t	normalized pseudo-factor vector
$\text{vec}(\cdot)$	vectorization operator
$\widehat{\Gamma}_j$	estimated autocovariance matrix at lag j
$K(\cdot)$	kernel function
S_T	bandwidth parameter
σ_f	standard deviation of simulated factors
σ_e	standard deviation of idiosyncratic errors
T_{sub}	number of observations in a subsample
$\lambda_1, \lambda_2, \lambda_3$	leading eigenvalues of the sample covariance matrix

Abstract

In this thesis, we study structural instability in large-dimensional factor models. We use the Russian stock market as an empirical application. The main objective of the thesis is to examine whether the factor structure of Russian stock returns remained stable over 2021–2025, or whether the data indicate a break in the factor loading matrix. The analysis is based on principal component estimation, the Bai–Ng criteria for factor selection, and a likelihood ratio test for structural change in factor models. Before proceeding to the real data, the procedure was tested on artificially generated data, where the true break date was known. After the procedure showed satisfactory results on the simulated data, we applied it to a balanced panel of weekly returns of Russian stocks. The test indicates a statistically significant break in mid-February 2024. A comparison between the subsamples before and after the estimated break date suggests that the factor structure changed. Overall, the results show that a single stable factor model is too restrictive for the Russian stock market over the period considered.

In questa tesi studiamo l'instabilità strutturale nei modelli fattoriali ad alta dimensionalità. Utilizziamo il mercato azionario russo come applicazione empirica. L'obiettivo principale della tesi è verificare se la struttura fattoriale dei rendimenti delle azioni russe sia rimasta stabile nel periodo 2021–2025, oppure se i dati indicano una rottura nella matrice dei carichi fattoriali. L'analisi si basa sulla stima tramite componenti principali, sui criteri di Bai–Ng per la selezione del numero di fattori e su un test del rapporto di verosimiglianza per i cambiamenti strutturali nei modelli fattoriali. Prima di passare ai dati reali, la procedura è stata testata su dati generati artificialmente, per i quali la vera data di rottura era nota. Dopo aver ottenuto risultati soddisfacenti sui dati simulati, abbiamo applicato la stessa procedura a un panel bilanciato di rendimenti settimanali di azioni russe. Il test indica una rottura statisticamente significativa a

metà febbraio 2024. Il confronto tra i sottocampioni prima e dopo la data di rottura stimata suggerisce che la struttura fattoriale sia cambiata. Nel complesso, i risultati mostrano che un unico modello fattoriale stabile è troppo restrittivo per il mercato azionario russo nel periodo considerato.

Chapter 1

Introduction

1.1 Motivation

Modern empirical economics and finance increasingly deal with large datasets. Financial markets generate a particularly rich form of such data: prices, returns, volumes, indices, and other market indicators are observed for many assets and over long periods of time. While this abundance of information is valuable, it also creates a methodological difficulty. A large panel of financial time series is often too complex to be interpreted directly, because individual series are affected both by market-wide forces and by asset-specific shocks. Therefore, an important task is to extract the common sources of variation from the idiosyncratic movements of individual assets.

Factor models are useful here because they give a compact way to describe the joint behavior of many financial series. Instead of treating each return series as a separate object, the model assumes that a large part of their co-movement is driven by a small number of latent common factors. In a financial context, these factors can be interpreted as unobserved sources of systematic risk or as broad market forces affecting many assets at the same time. This is the main dimension-reduction role of the model: the researcher works not with dozens of individual return series separately, but with a smaller set of estimated factors that summarizes their common dynamics.

This approach becomes particularly relevant in large panels, where both the number of assets and the number of observations are substantial. Classical factor analysis was developed

mainly for settings with a fixed or relatively small cross-sectional dimension, whereas the large-dimensional framework allows both N and T to grow. Under appropriate regularity conditions, the factor space can then be consistently estimated. This is the reason why Principal Component Analysis is widely used in this literature. In particular, Bai and Ng show that asymptotic principal components recover the true factor space consistently, which provides the theoretical justification for using PCA in large-dimensional factor models [7].

A central issue in this framework is the determination of the number of factors. This is not merely a technical detail. If the number of factors is underestimated, the model may fail to capture important sources of common variation. If it is overestimated, the model may include components that mainly reflect noise rather than genuine common structure. Bai and Ng formulate this problem as a model selection problem and propose information criteria that are specifically adapted to panels where both N and T are large [5]. This is particularly relevant for financial applications, where the number of potential assets can be large and where the researcher usually does not observe the underlying factors directly.

However, even after the number of factors has been selected and the factors have been estimated, another important question remains: whether the factor structure is stable over time. In many empirical applications, and especially in financial markets, the assumption of constant factor loadings may be restrictive. The relationship between individual assets and common market forces can change because of crises, changes in monetary policy, geopolitical shocks, sanctions, market segmentation, changes in investor behaviour, or other structural events. In such cases, a model estimated over the full sample may hide important changes in the dependence structure of the data.

For this reason, testing for structural breaks in factor models is of both theoretical and empirical interest. Bai, Duan and Han argue that structural change is common in economic variables and is more likely to occur in high-dimensional data; therefore, checking the structural stability of factor models is important for understanding the data structure and for ensuring the validity of subsequent analysis [2]. Their likelihood ratio approach is based on the idea that a break in the factor loading matrix can be studied through changes in the moments of the estimated factors. This transforms a high-dimensional structural change problem into a lower-dimensional

problem involving the estimated factor process.

The present thesis is motivated by these considerations. It applies large-dimensional factor model techniques to financial data in order to estimate the latent factor structure, determine the number of common factors, and examine whether the factor loading matrix remains stable over time. The empirical focus on financial returns is natural because such data are characterized by strong co-movement, changing volatility, and possible shifts in dependence patterns. Therefore, the analysis of structural breaks in factor models can provide useful information about changes in the underlying market structure.

1.2 Literature review

1.2.1 Large-dimensional factor models: general framework and motivation

The literature on large-dimensional factor models occupies an important place in modern econometrics and financial economics. This is due to the fact that, in many applied settings, researchers are faced with datasets in which a large number of interrelated variables are observed simultaneously. In macroeconomics, such variables may include indicators of business activity, inflation, employment, interest rates, and foreign trade; in finance, they may include asset returns, credit spreads, volatility measures, firm accounting characteristics, and other indicators of market conditions. With such a data structure, the natural question becomes whether it is possible to replace the analysis of a large number of observed series with a more compact representation in which the main part of the joint dynamics is described by a small number of latent common factors. This is precisely the task that factor models are designed to address [9, 7].

In its most general form, a factor model assumes that each observed variable can be decomposed into two components: a common component driven by a limited number of latent factors, and an individual or idiosyncratic component reflecting fluctuations specific to that variable. This formulation makes factor models especially attractive in high-dimensional environments, as they allow for significant dimensionality reduction without losing key information about the joint movement of the data. In other words, instead of analyzing dozens or hundreds of series, the

researcher can work with a small number of factors interpreted as the main sources of systematic variation [9].

As emphasized in the survey literature, the importance of large-dimensional factor models is explained not only by their ability to reduce dimensionality, but also by the breadth of their applications. They are widely used in forecasting, factor-augmented regressions, FAVAR models, covariance structure analysis, panel data analysis, and the construction of applied macroeconomic and financial indicators. In this sense, factor models should not be viewed merely as a highly specialized technical tool, but rather as a general framework for organizing information in conditions where the number of potentially relevant variables is large and the relationships between them are complex and multi-layered. This view is especially clearly presented in the review papers by Bai and Wang [9] and Bai and Ng [7], where factor models are presented as a central element of modern econometric work with large datasets.

It should be emphasized that the development of the literature on large-dimensional factor models is closely related to the transition from the classical, relatively low-dimensional statistical formulation to asymptotics, in which both the number of variables and the length of the time series grow simultaneously. It is in this formulation that the key questions for this field arise: under what conditions factors and factor loadings can be consistently estimated; what is the acceptable nature of dependence and heterogeneity in idiosyncratic errors; how to determine the number of factors; how valid statistical inference can be conducted when factors are not directly observed, but are replaced by estimates. Thus, the literature on large-dimensional factor models has developed not only in an applied direction, but also as a rigorous theoretical field, ultimately forming an independent and very rich area of research [9, 7].

Several major strands can be distinguished within this literature. The first is related to the general econometric theory of large-dimensional factor models, including the issues of identification, consistency of estimates, and asymptotic properties of estimated factors and loadings. The second strand is devoted to determining the number of factors, that is, choosing the effective dimension of the latent factor space. The third is related to the subsequent use of estimated factors in regression, forecasting, and other statistical procedures, where the problem of valid inference arises in the presence of generated regressors. Finally, more recent literature expands

the classical formulation to include weak factors, regularized methods, low-rank recovery, matrix completion, unbalanced panels, as well as problems of structural instability and temporal variability of parameters [17]. This gradual expansion of the research agenda shows that factor models have long ceased to be merely a principal components technique in applied statistics and have become a broad methodological platform for analyzing high-dimensional economic data.

Modern reviews also emphasize another important feature of this literature: factor models are increasingly viewed through the lens of low-rank structure [17]. In this interpretation, the latent factor component corresponds to the low-rank part of the observed high-dimensional system, and the task of estimating factors becomes a special case of the more general problem of recovering a low-rank structure in the presence of a noise component. This approach has proved particularly fruitful in recent years, as it has connected the classical econometrics of factor models with the modern literature on high-dimensional statistics, regularization, and econometric learning. As a result, the field of factor models has become a natural intersection point between traditional macroeconometrics, financial econometrics, and modern statistical learning methods.

Thus, large-dimensional factor models form a mature and multi-layered research area in which fundamental theoretical issues are closely intertwined with applied problems. For the purposes of this work, two areas of this literature are particularly important. On the one hand, this includes studies devoted to the estimation of factors and the determination of their number. On the other hand, it includes works analyzing the stability of the factor structure over time and the consequences of structural shifts in factor loadings and the covariance structure. The subsequent analysis is situated precisely at the intersection of these two strands of the literature.

1.2.2 Estimation of factors and inferential issues

One of the central areas of research in the literature on large-dimensional factor models is the estimation of latent factors and related model parameters. If, at the most general level, a factor model specifies a representation of observed variables in terms of a small number of latent common factors, then the next natural step is to understand how these factors can be extracted from the data, how accurate the resulting estimates are, and to what extent they can be relied upon in subsequent empirical procedures. This is why the literature on factor models has never

been limited to describing the general structure of the model itself. Instead, it has focused on developing estimation methods, analyzing their asymptotic properties, and constructing valid statistical inference in situations where the factors are not directly observed but are inferred from the sample.

The classical and most widely used approach to estimating factors in large dimensions is principal components analysis. Its popularity can be explained by several factors. First, it is relatively simple from a computational point of view and naturally relies on the spectral structure of the sample covariance matrix. Second, within the framework of approximate factor models, under fairly general conditions, principal components turn out to be a viable tool for consistently recovering both the space of common factors and the factor loadings. Third, this approach became the starting point for much of the subsequent literature, both theoretical and applied. As a result, principal components analysis has long been regarded as the standard procedure in the empirical analysis of large-dimensional factor models [17].

However, as the literature developed, it became clear that the mere consistency of the estimated factors was not enough. If a researcher uses the obtained factor estimates in subsequent regressions, forecasting equations, or hypothesis-testing procedures, a fundamentally important question arises: whether valid statistical inference remains possible in the presence of generated regressors. In other words, is it possible to treat the estimated factors as if they were observable quantities, or should their preliminary estimation be specifically taken into account in the asymptotic theory? It is precisely this question that has become the subject of a separate strand of the literature devoted to inference with estimated factors.

An important role here was played by the work of Bai and Ng (2006) [4], which examines diffusion index forecasts and factor-augmented regressions. The authors show that the use of estimated factors in regression and forecasting models requires a separate analysis, since the properties of the final estimates depend not only on the standard arguments of regression theory, but also on the accuracy of the preliminary estimation of the factors themselves. In particular, this literature emphasizes that the asymptotic behavior of estimates may depend on the combined growth of the panel dimension and the length of the time series, and therefore statistical inference in factor-augmented models cannot be reduced to the mechanical application

of familiar results for regressions with observed regressors. Thus, this paper has become one of the key contributions to understanding that the stage of factor estimation and the stage of subsequent inference are conceptually inseparable.

The next important step was the development of likelihood-based approaches to the estimation of large-dimensional factor models. Although principal components still retained the status of a basic tool, researchers increasingly turned to the question of whether more efficient procedures could be obtained if a more complete probabilistic model specification was used. A significant contribution to this area was made by the work of Bai and Li (2016) [3], which examines maximum likelihood estimation and inference for approximate factor models in high dimensions. This article is important not only as a technical extension of the PCA approach, but also as a substantive shift in the literature: it shows that the econometrics of large-dimensional factor models includes not only spectral methods, but also a full-fledged likelihood-based inferential framework. Thus, the literature moves from the question of how to extract factors to the more general question of how to estimate and test factor models within a structured probabilistic framework.

At the same time, later contributions increasingly emphasized that the classical PCA approach, although convenient, is not always optimal in terms of finite-sample properties or robustness to noise, weak factors, and other complications in the data. In this regard, the work of Bai and Ng (2017) [8] is of particular interest, where principal components are considered in conjunction with regularized estimation. The general logic of this literature is that the standard extraction of factors through principal components can be supplemented with shrinkage procedures and other forms of regularization, which, although they introduce bias, in some cases improve the overall quality of estimation, especially in the presence of weak signals or significant noise components. This strand of the literature is important for the present review for two reasons. First, it demonstrates that the modern theory of factor models is no longer limited to the classical PCA results of the early 2000s. Second, it allows us to naturally link the literature on factor models to the broader field of high-dimensional statistics and regularization methods.

Modern survey articles also emphasize that factor estimation is increasingly being viewed

through the lens of low-rank recovery. In this formulation, the factor structure is interpreted as a low-rank part of a large observed system, and the methods of factor estimation turn out to be special cases of more general procedures for recovering a low-rank signal in the presence of a noise component. It is in this context that recent developments in factor models turn out to be closely related to modern methods of statistical learning, regularization, high-dimensional optimization, and matrix recovery. Fan, Li and Liao (2021) [17] emphasize that in recent years the literature has expanded significantly due to new methods of factor estimation, factor-augmented inference and models for unbalanced panels considered from the point of view of matrix completion. Such a review is especially useful for understanding that modern methods of estimating factors can no longer be viewed as a purely classical topic of time series econometrics; on the contrary, today they are at the intersection of econometrics, high-dimensional statistics and econometric learning.

Thus, the literature on factor estimation and statistical inference in large-dimensional factor models has undergone substantial development. While early work largely focused on principal components as a basic method for extracting latent common components, subsequent research significantly expanded this perspective. It raised questions about the validity of inference when using estimated factors, about the possibility of likelihood-based estimation, about the role of regularization, as well as about the relationship between factor models and the broader low-rank methodology. For this work, this body of literature is important primarily because it provides the necessary theoretical basis for subsequent questions about the number of factors and the stability of the factor structure. In other words, before discussing how many factors should be identified and how the factor model changes with structural shifts, it is necessary to understand how factors are estimated in the first place and what limitations accompany their subsequent use in empirical analysis [17].

1.2.3 Determining the number of factors

The issue of determining the number of factors occupies a central place in the literature on large-dimensional factor models. If the very idea of a factor model is that the joint dynamics of a large number of observed variables can be described through a limited number of latent

common components, then a natural question immediately arises: how many such components should be retained in practice. This question is not purely technical. On the contrary, the choice of the number of factors determines the interpretation of the model itself, the quality of the estimation of the common component, the accuracy of forecasts, the validity of subsequent statistical inference, and, more broadly, how well the factor model reflects the underlying data structure. Therefore, the literature on determining the number of factors has emerged as one of the richest and most methodologically diverse strands of research in this field.

The difficulty here lies in the fact that the number of factors is typically unobserved and not specified in advance. Moreover, in a high-dimensional setting, one cannot simply transfer intuition from the classical low-dimensional principal components framework, where the researcher may rely on visual inspection of the scree plot or on heuristic rules concerning “significant” eigenvalues. In the context of large-dimensional factor models, the problem takes on a much more rigorous econometric character. It is necessary to distinguish between eigenvalues generated by genuinely pervasive common factors and eigenvalues that arise from idiosyncratic noise, local dependence, finite-sample distortions, or features of a particular model specification. This is why the question of the number of factors became closely related to asymptotic theory, the properties of the spectrum of large covariance matrices, and the development of formal procedures for choosing the dimension of the factor space [7].

Early contributions and classical approaches

Early work on this topic laid the foundation for the subsequent literature by showing that determining the number of factors should be treated as an independent statistical problem rather than as a secondary element of model estimation. One of the early contributions in this direction was the work of Connor and Korajczyk (1993) [15], which proposed a test for the number of factors in an approximate factor model. The significance of this article lies in the fact that it represents an early formalization of the problem itself: the number of factors should be determined not heuristically, but on the basis of a statistical procedure that takes into account the structure of dependence in the data and distinguishes the general component from the idiosyncratic one. Thus, the early literature had already outlined an important line of

research that would later be developed further: the choice of the number of factors is not merely a matter of convenience, but also one of correct model specification.

The real turning point, however, was the paper by Bai and Ng (2002) [5], which essentially set the standard framework for subsequent research. It was this work that consolidated the idea that, in large-dimensional approximate factor models, the number of factors can be determined using information criteria based on a trade-off between the quality of approximation and the penalty for model dimension. The importance of this approach can hardly be overstated. On the one hand, it gave researchers a practically feasible and theoretically justified procedure for selecting the number of factors in a setting where both the number of variables and the length of the time series grow simultaneously. On the other hand, it was after Bai and Ng that the literature began to discuss the problem of factor selection within the general asymptotic logic of large N , large T , rather than merely as a question of the applied analysis of the spectrum of the sample covariance matrix. In this sense, Bai and Ng (2002) [5] serve as a reference point for almost all subsequent literature on this topic.

The further development of the classical approach is associated with an expansion of the very concept of factor dimension. While the early literature was largely focused on the number of static factors, later work raised the question of the number of primitive shocks and the dynamic dimension of the factor structure. This is precisely the direction of Bai and Ng (2007) [6], where the discussion concerns not only the number of factors per se, but also the number of fundamental shocks underlying the observed factor dynamics. This extension is especially important because it shows that the problem of factor-model dimension can be understood at different levels. One may ask how many factors are required for an adequate representation of the data in static form, but one can also be interested in how many independent structural sources of fluctuations are behind these factors in a dynamic formulation. Thus, the literature gradually moves from the relatively narrow question of how many factors should be identified to the deeper question of the internal structure of the factor model itself.

Formal hypothesis testing approaches

Although the information criteria proposed in the classical literature proved to be extremely influential, they did not exhaust all possible approaches to determining the number of factors. In parallel, another strand of the literature developed in which the problem was formulated in terms of formal hypothesis testing. Under this approach, the researcher no longer merely minimizes a given criterion, but instead tests the null hypothesis that the number of factors is equal to a given value against an alternative with a larger number of factors. This approach brings the problem closer to the traditional statistical logic of sequential testing and allows the choice of dimension to be interpreted through the standard categories of null and alternative hypotheses.

A particularly important contribution here is Onatski (2009) [23], which develops a formal approach to testing hypotheses about the number of factors in large-dimensional factor models. The significance of this paper lies in the fact that it translates the problem of choosing the number of factors into a rigorous testing framework and focuses on the behavior of eigenvalues in high-dimensional asymptotics. Thus, the literature gains an alternative to information-criterion-based approaches: instead of looking for the minimum of a penalized objective function, one can analyze the statistical properties of the spectrum and examine whether the observed eigenvalue structure is consistent with a model with a given number of factors. This shift is especially important for the theoretical literature, as it strengthens the connection between factor selection and general high-dimensional asymptotic theory.

A similar but independent line of research is represented by Kapetanios (2010) [18], which also proposes a testing procedure for determining the number of factors in approximate factor models with large datasets. What is important here is not so much the specific technical design of the test as the broader methodological direction: the number of factors can be determined not only through information criteria, but also through procedures explicitly designed as statistical tests and aimed at robustness to various forms of dependence in the data. Collectively, this type of work has played an important role in the formation of an alternative testing strand of the literature, which exists alongside the Bai–Ng criteria and shows that the problem of factor dimension can be approached in several methodologically distinct ways.

Refinements and improved finite-sample procedures

After the basic methods for selecting the number of factors had been proposed, a significant part of the literature shifted toward their refinement and the improvement of their finite-sample properties. This development was entirely natural. Even if a criterion is asymptotically consistent, this does not mean that it will perform well in moderate-sized samples, under strong idiosyncratic noise, in the presence of dominant factors, or in settings where the spectral gap between factor and non-factor eigenvalues is not sufficiently pronounced. This is why the next stage of the literature was associated less with the creation of entirely new concepts than with the refinement of existing procedures.

An important example of such work is the paper by Alessi, Barigozzi, and Capasso (2010) [1], which proposes an improved penalization strategy compared with the original Bai–Ng criteria. Substantively, this paper is important because it shows that even very influential and theoretically sound criteria may exhibit systematic weaknesses in finite samples, in particular a tendency to overestimate the number of factors under an unfavorable noise structure. Therefore, the problem of factor selection does not end with a formal proof of consistency. Equally important is the question of how well a particular procedure behaves under realistic empirical conditions

A similar logic applies to works based on eigenvalues and their transformations. For example, Wu (2016) [30] proposes a robust determination method, aimed in particular at situations involving dominant factors. The importance of this line of research lies in the fact that it draws attention to the asymmetry of the problem of choosing the number of factors: the difficulty lies not only in separating factor eigenvalues from noise ones, but also in correctly identifying weaker factors in the presence of one or more particularly strong ones. This approach brings the problem closer to realistic empirical settings, where factors are not always of comparable strength.

The same refinement logic continues in Wu (2018) [29], which proposes the eigenvalue difference test. Such methods are important not so much because they completely replace earlier criteria, but because they demonstrate the effort within the literature to fine-tune factor-selection procedures to a range of empirical difficulties: finite-sample issues, dominant factors, instability of ratio-type statistics, sensitivity to ridge parameters, and other related features.

More generally, this strand of research shows that the problem of determining the number of factors is far from having a definitive solution in a single form; on the contrary, it requires the constant refinement of tools, especially if the goal is not only asymptotic correctness, but also good practical performance.

Cross-validation and regularization approaches

More recent literature has gone even further by proposing procedures that move beyond the classical juxtaposition of information criteria and formal tests. In particular, interest in cross-validation and regularization-based approaches has increased noticeably in recent years. The reason for this shift is quite clear. Many classical procedures were developed under relatively strict asymptotic conditions and do not always perform equally well in moderate samples, under heteroskedasticity, heavy tails, weak factors, or a particularly strong idiosyncratic component. In such circumstances, it is quite natural to turn to more flexible data-driven methods, well known from modern high-dimensional statistics and statistical learning.

It is in this context that the work of Zeng, Xia, and Zhang (2019) [31] should be understood, as it proposes a double cross-validation method for determining the number of factors in approximate factor models. The very emergence of such an approach is itself important: determining the number of factors ceases to be solely a matter of spectral asymptotics and begins to be understood as a model-selection problem in a broader statistical sense. A particularly attractive feature of procedures of this kind is that they are motivated not only theoretically, but also practically: they promise better stability in situations where classical criteria are sensitive to noise, heavy tails or unstable error structure. Thus, the literature on factor selection is increasingly moving closer to the modern culture of validation-based model choice.

A second important direction here is related to regularization. If, in the classical formulation, the number of factors is determined through penalized fit criteria or through tests based on eigenvalues, then in more recent work regularization begins to be used directly as a means of structurally selecting the effective dimension of the model. In this context, we can mention Caner and Han (2014) [12], which discusses group bridge estimators, as well as the more recent contribution of Daniele (2024) [16], which relies on ideas of group variable regularization and

adaptive hierarchical lasso. These papers are particularly important because they focus on more complex empirical situations, including weak factors and more general forms of dependence in idiosyncratic errors. In other words, the dimension of the factor space begins to be determined not only through the global behavior of the spectrum, but also through penalized selection logic, built directly into the estimation procedure.

From the same perspective, Bai and Ng (2017) [8] is also of interest, because although it relates to the broader issue of regularized estimation of factor models, it is also important for the literature on the choice of the number of factors. It shows that shrinkage and related regularization devices can lead to a more conservative estimate of the effective factor dimension, especially in the presence of weak factors or high noise load. This is especially important for understanding that, in the modern literature, the number of factors is increasingly viewed not as a fixed and directly observable characteristic of the spectrum, but as an object depending on the chosen procedure, the target task, and the balance between bias and variance. Thus, the regularization-based literature expands the classical understanding of factor selection and integrates it into the broader context of high-dimensional estimation.

Taken together, this entire strand of research shows that the problem of determining the number of factors has evolved from a relatively narrow problem of choosing a dimension in a PCA-like formulation to a multi-layered area in which information criteria, test procedures, refinements for finite-sample performance, cross-validation methods and regularization-based selection devices coexist. This development is itself an important result of the literature. It shows that the factor dimension is not a trivial technical parameter, but one of the key objects of econometric analysis in large-dimensional factor models. For the present work, this body of literature is especially important, since both the construction of an empirical factor structure and the interpretation of subsequent results related to the structural stability of the model depend on the correct determination of the number of factors.

1.2.4 Structural breaks in factor models

Another important strand of the literature on large-dimensional factor models is related to the issue of the structural stability of the factor structure over time. In the basic formulation of

a factor model, it is usually assumed that the factor loadings, and sometimes other model parameters, remain unchanged over the entire period under consideration. This assumption is convenient both from a theoretical and an applied point of view, since it allows latent factors to be interpreted as stable sources of the joint dynamics of the observed variables. However, in real macroeconomic and financial data, this assumption often proves to be too strong. Economic regimes change, the institutional environment evolves, and crises, reforms, technological shifts, and other events may alter not only the behavior of individual variables, but also the very structure of their overall dependence.

This is why the literature on structural breaks in factor models has emerged as a natural continuation of earlier studies on factor estimation and the determination of the number of factors. While one part of the literature focused primarily on the question of how many factors are required to describe the data, another gradually made it clear that the answer to this question may itself depend on the stability of the factor structure. In other words, factor selection and structural instability cannot be fully separated from one another. If the structure of factor loadings changes over time, then standard procedures that assume constant parameters may produce a distorted picture of both the number of factors and the nature of the latent common components [11].

Why structural breaks matter

Structural breaks matter in factor models because they affect the link between observed variables and latent factors. With stable loadings, the estimated factor structure can be interpreted as a common mechanism behind the joint movement of the data over the whole sample. When the loadings change, this interpretation becomes less straightforward: even with the same number of factors, their relative importance and their effect on individual variables may differ across regimes. A full-sample factor model may then average over distinct periods and hide the change in the dependence structure. For this reason, factor estimation should be followed by a stability check of the loading matrix.

This issue is especially important because ignoring structural shifts can lead not only to local errors of interpretation, but also to deeper distortions in statistical analysis. In particular,

the literature shows that breaks in factor loadings can artificially increase the number of factors selected by standard information criteria. Intuitively, this can be understood as follows: if the same sample combines several regimes with different loading structures, then the aggregated covariance matrix begins to reflect not one stable factor system, but a mixture of several different configurations. As a result, the spectral structure of the data may appear as if there were more common factors in the model than are actually required in each individual regime. This important result is emphasized, in particular, in the work of Breitung and Eickmeier [11], where it is shown that structural shifts in loadings can significantly affect procedures for selecting the number of factors and the interpretation of latent common components.

In practice, a break in the loading matrix affects more than the formal specification of the model. If the full sample contains different regimes, pooled estimation may mix them and produce factor and loading estimates that do not describe any single period well. This also matters for forecasting: parameters estimated before a regime change may perform poorly after it. Inference can be affected as well, because many asymptotic results for factor models rely on some form of parameter stability. Finally, in macroeconomic and financial applications, the break date may itself be informative, since it can correspond to crises, policy changes, deregulation, technological change, or shifts in market behavior.

Therefore, more recent literature gradually shifts the focus away from the question of whether there are factors at all toward the question of how stable the factor structure is and how possible structural changes should be accounted for. In this regard, the literature on structural breaks in factor models is developing along several directions at once. Some papers study how to test the hypothesis about the stability of factor loadings. Others compare various test statistics, in particular LR, LM and Wald-type approaches [11, 2]. Still others go beyond the single-break setup and consider an unknown number of breaks or multiple breaks [20, 21]. Finally, the most recent studies take it one step further and try to distinguish between qualitatively different types of structural changes — for example, changes in loadings themselves and changes in the variance or covariance structure of factors [19]. This makes the literature not only more technically advanced, but also more conceptually sophisticated.

Modern survey studies also play an important role here, as they show that structural changes

in high-dimensional factor models already form an independent and rich field of research. In this area, the literature discusses conditions for identifying breaks, the difference between large and small breaks, testing shifts in individual loadings and in the loading matrix as a whole, estimating break dates, and the relationship between structural breaks and model selection. Such survey work is particularly useful for the present study, as it allows the problem of structural shifts to be viewed not as a particular technical detail, but as a natural part of the general econometrics of factor models.

For the purposes of the present study, this block of literature is of fundamental importance. If the number of factors and their estimation form the first part of the analysis, then checking the stability of the factor structure constitutes the next necessary step. In other words, after the researcher has determined the factor dimension of the model and estimated its parameters, the question arises whether this structure persists throughout the entire observation period. It is at this point that the literature on structural breaks intersects with the literature on factor selection, and it is precisely at this intersection that the present study is located.

Early and classical break testing procedures

Early and classical work on structural breaks in factor models was largely focused on the question of whether it is possible to formally identify the presence of a structural shift in factor loadings and how such a shift affects the behavior of the model as a whole. At this stage, the literature had not yet sought to cover all possible forms of instability or to consider complex scenarios with multiple breaks. The main focus was on a more basic but fundamentally important question: whether there is at least one significant break in the factor model that violates the assumption of constant loadings. It was this formulation that became a natural first step in the development of the break literature, since without addressing the issue of a single structural break, it would have been difficult to move on to more complex and subtle forms of instability.

In this early literature, the idea of testing the stability of factor loadings as a key parametric object of the model occupies a central place. This is important from a substantive point of view. If one assumes that factors, as latent sources of joint dynamics, may persist while the relationship between factors and observed variables changes, then it is the loading matrix that

becomes the main carrier of the structural shift. In this case, the task is not merely to detect a change in the data in general, but to establish whether there has been a change in the factor structure in the narrow econometric sense. Thus, early tests for breaks in loadings provide a more meaningful alternative to general tests for time-series instability, since they are focused on the specific internal organization of factor models.

A typical example of this approach is the work of Bueno, de Lucas, Delgado, and Álvarez (2011) [13], which uses SupWald-type logic to test for structural shifts in factor loadings, with an application to international business cycles. The significance of this article lies not only in the testing procedure itself, but also in the fact that it demonstrates the applicability of break-testing ideas in real macroeconomic factor models. In other words, this is no longer a purely abstract problem of parameter stability, but a meaningful empirical question: whether the structure of the influence of latent factors on the observed economic series changes across different periods of time. This approach is important for the literature review, as it shows that the question of structural breaks in factor models initially developed not only as a theoretical problem, but also as a tool for empirical analysis of macroeconomic dynamics.

Another important branch of the early literature focused on large structural shifts in high-dimensional factor models. The article by Chen, Dolado, and Gonzalo (2014) [14], devoted to the detection of big structural breaks in large-dimensional factor models, is particularly notable here. Substantively, this work is important because it takes a step away from the more local problem of whether there is instability in loadings toward the question of how to detect genuinely large structural changes in a high-dimensional factor environment. This is closer to a formulation in which the factor model is treated as an object that is sensitive to regime changes at the level of the entire system, rather than merely to small deviations in individual coefficients. Accordingly, the literature is gradually beginning to shift from the idea of a local stability test to the more general task of detecting major breaks in the architecture of factor dependence itself.

This also includes the work of Breitung and Eickmeier (2009) [11], which occupies a particularly important place because it connects several key lines of literature at the same time. On the one hand, it considers structural breaks in dynamic factor models, thereby transferring the discussion of instability to a richer dynamic setting. On the other hand, it directly shows

that breaks in factor loadings can artificially increase the number of factors selected by standard information criteria. This is why the article serves as a natural bridge between the literature on structural breaks and the literature on determining the number of factors. In addition, it is important because it introduces a comparison of LR, LM, and Wald-type testing strategies at a relatively early stage, thereby paving the way for later work in which this comparison becomes a central methodological element.

More generally, the early and classical literature on structural breaks in factor models performs several functions at once. First, it formalizes the very idea that the stability of factor loadings should be the subject of separate testing, rather than a tacit assumption. Second, it shows that even a single break can have far-reaching consequences for estimating the model, choosing the number of factors, and interpreting common components. Third, it is in this literature that the basic test-based strategies first appear, and these later serve as the foundation for more modern LR-type, multiple-break, and disentangling approaches. Therefore, although many of these early works no longer represent the latest stage of methodology, they remain an essential starting point for understanding how the modern break literature in factor models was formed.

LR / LM / Wald approaches

As the literature on structural breaks in factor models has developed, the question has become increasingly important not only of whether there is a break in the factor structure, but also of how exactly such a break should be tested. At this stage, the literature begins to move from a general statement of the instability problem in factor loadings to a more specific comparison of alternative testing strategies. In particular, likelihood ratio, Lagrange multiplier, and Wald-type approaches become especially important. For the econometrics of factor models, this direction is of fundamental importance, since the choice of a specific test statistic determines not only the computational implementation of the procedure, but also its power, robustness to different types of alternatives, as well as the very way in which the structural shift is parameterized.

One of the important early contributions in this area is the article by Breitung and Eickmeier (2009) [11], which examines structural breaks in dynamic factor models and compares LR, LM,

and Wald-type testing strategies at a relatively early stage. The significance of this article lies in the fact that it takes the literature beyond the general idea that the stability of the factor structure must be tested and translates the discussion into a comparative analysis of different classes of tests. Thus, the structural break literature in factor models becomes more mature: researchers begin to be interested not only in the presence of a shift itself, but also in which statistical tool is most appropriate in this setting. In addition, it is in this work that an important result is formulated especially clearly for all subsequent literature: ignoring structural shifts in loadings can lead to an overestimation of the estimated number of factors and, consequently, to a distorted interpretation of latent common components. In this sense, the comparison of LR, LM, and Wald approaches turns out to be closely related not only to the diagnosis of breaks, but also to the broader issue of correct model specification.

A central contribution in this branch of the literature is Bai, Duan and Han (2024) [2], devoted to the likelihood ratio test for structural changes in factor models. For the present study, this article is especially important because it offers not just one of the possible test procedures, but a conceptually strong reformulation of the problem itself. The key idea of the paper is that a structural shift in factor loadings can be reparameterized as a change in the covariance structure of the factors. This formulation is fundamentally important because it allows us to move from direct testing of the instability of a large load matrix to a more convenient and econometrically transparent problem in which a break manifests itself through a change in factor variance. It is this idea that makes the LR approach particularly meaningful and distinguishes it from a number of more straightforward Wald-type constructions.

From a methodological point of view, the contribution of Bai, Duan, and Han is important for several reasons. First, it shows that testing structural shifts in factor models can be constructed not only as a test of parameter instability in the traditional sense, but also as a task that significantly uses the internal geometry of the factor model itself. Second, the LR statistic in this formulation turns out to be not just a technical alternative to other tests, but a way to use information about the model structure more effectively. Third, the authors emphasize that the corresponding LR test may have greater power than Wald-type procedures, which makes the issue of choosing the test statistic substantively important rather than merely formally relevant. In other words, in this

literature, the comparison of LR, LM, and Wald ceases to be a purely educational exercise and turns into a genuine problem of econometric test design.

Another important aspect of this paper is that it helps to conceptually organize the preceding literature. If early research mainly raised the question of how to detect a break in loadings, then the LR approach of Bai, Duan and Han shows that the structure of the problem itself allows for deeper parameterization. This is especially important in a high-dimensional setting, where direct instability testing in a large coefficient matrix can be cumbersome from both a computational and theoretical point of view. Transferring the problem to the space of the factor covariance structure makes the analysis more compact and allows for better alignment of the testing procedure with the internal logic of the approximate factor model. In this sense, the paper appears not merely as a new contribution in an already existing line of tests, but as a methodological step forward.

For the literature review, it is also important to emphasize that the LR/LM/Wald comparison in factor models cannot be understood in isolation from the more general task of model selection and structural interpretation. If different tests react differently to shifts in loadings, then this affects not only the formal decision about the presence of a break, but also how the researcher will interpret the evolution of the factor structure itself, how many factors the researcher considers appropriate to retain and which changes in the data will be interpreted as a regime change. Therefore, the branch of literature comparing LR, LM, and Wald-type procedures performs a broader function: it links break detection, factor selection, and interpretation of latent common components into a single econometric problem.

For the present study, this strand of the literature is of central importance. If early studies provide a general background and show why structural breaks matter, then it is the LR-based approach that allows us to move on to a more focused analysis of the stability of the factor structure. In particular, the work of Bai, Duan and Han sets a methodological framework in which a structural shift is considered not just as an external violation of the constancy of parameters, but as a change in the internal factor structure of the data. This is why it serves as the main methodological foundation for the subsequent analysis in the present study.

Unknown or multiple breaks

As the literature developed, it became clear that a framework with a single structural shift, although a natural starting point, is in many cases too narrow for real data. In macroeconomic and financial applications, the factor structure can change not once, but sequentially, responding to a series of crises, institutional reforms, policy regime changes, technological shifts, or other major events. Therefore, later literature began to move away from the single-break framework toward a more general problem in which the number of breaks is unknown in advance or may potentially exceed one. This transition was fundamentally important because it brought the econometrics of factor models closer to a real empirical environment in which instability rarely manifests itself in the form of a single break.

One of the important steps in this direction was the work of Ma and Su (2018) [21], devoted to the estimation of large-dimensional factor models with an unknown number of breaks. The significance of this article lies in the fact that it extends the break literature from a simple test of the hypothesis of a single break to a more general formulation in which several interrelated tasks must be solved simultaneously: to determine whether there are any structural changes at all, how many of them have occurred, and where exactly they are located in the sample. Thus, the literature begins to move from testing a break as a separate event to the full-fledged problem of sample segmentation, where the factor model allows for several regimes with different parameters. This is especially important econometrically, because with an unknown number of breaks, the problem no longer boils down to comparing two subsamples or testing a single candidate break; it becomes one of joint model selection and break dating.

The next important step is related to the literature on multiple structural breaks in high-dimensional covariance structures. Particularly notable here is the article by Li, Li, and Fryzlewicz (2023) [20], which examines the detection of multiple structural shifts in large covariance matrices in the context of approximate factor models. The significance of this work lies not only in the transition from single break to multiple breaks, but also in the fact that it focuses on the distinction between changes in the common and idiosyncratic parts of the model. This is especially important because in real data, instability can affect not only factor loadings or the factor component itself, but also the dependence structure in idiosyncratic errors. Consequently,

the task of detecting breaks becomes more complicated: it is necessary not only to fix that the covariance structure has changed, but also to understand which part of the model is the source of this change. In this sense, the literature on multiple breaks turns out to be closely related to the more general problem of decomposition of structural change.

An even more conceptually subtle step is taken by the work of Koo, Wong, and Zhong (2025) [19], which proposes distinguishing between structural breaks in factor loadings and structural breaks in factor variance. This idea is of great importance for the modern literature, as it shows that not every break in a factor model should be interpreted in the same way. A change in loadings and a change in factor variance can lead to different economic and statistical consequences, even if at the level of observed data both types of instability manifest themselves as a change in the overall covariance structure. Thus, the problem of structural shifts becomes more profound: the researcher must not only discover the fact of the break itself, but also correctly classify its nature. This is especially important in applied work, since different types of instability lead to different interpretations of what is happening in the data. In one case, one may speak of a change in how latent factors affect the observed variables; in the other, about a change in the strength or volatility of the latent factors themselves while maintaining the same loading system.

Thus, the more recent literature on unknown and multiple breaks significantly expands the earlier formulation of the problem. Whereas the classic articles mostly asked whether there was at least one break in loadings, modern work raises much more complex questions: how many breaks have occurred exactly, how they should be located, whether they relate to the common or idiosyncratic part of the model, and whether it is possible to distinguish between breaks in loadings and breaks in factor variance. This extension is extremely important for understanding modern break literature, as it shows that structural instability in factor models is no longer considered as a single deviation from the basic model, but as an independent, multi-layered object of econometric analysis.

For this work, this body of literature is important for two reasons. First, it shows the natural direction of development of the entire break literature: from simple testing of a single structural break to more flexible formulations focused on an unknown number of breaks and more complex forms of instability. Second, it emphasizes that the very nature of the break is no less important

than the fact of its presence. Even if the main focus of this study is on the more classical formulation of LR-based testing for structural change, the modern literature on unknown and multiple breaks creates the necessary broader context and shows the direction in which the methodology for analyzing structural instability in factor models is evolving.

1.2.5 Position of the present thesis in the literature

The literature review shows that modern research on large-dimensional factor models constitutes a broad and already fairly mature field of econometric analysis. Within this field, at least two major and closely related strands can be identified. The first is devoted to estimating latent factors, determining their number, and constructing valid statistical inference in models where the factors are not directly observed, but are estimated from the data. The second area is related to the analysis of structural instability in factor models, including testing shifts in factor loadings, comparing various break-detection procedures, and studying the consequences of such shifts for model specification and interpretation. Collectively, these strands define the intellectual context in which the present study is situated.

On the one hand, this work is based on the literature on factor estimation and factor selection. It is this literature that provides the initial theoretical basis for understanding how a common component can be identified in large panels, how to determine the dimension of the factor space, and how to interpret the resulting factor structure. Without this block of research, it would have been impossible to even set a subsequent task on structural stability, since any analysis of breaks in factor models assumes that the factor model itself has already been defined, its dimension selected, and latent common components estimated. In this sense, the literature on factor estimation and the determination of the number of factors plays the role of a fundamental first level of analysis in the present study [9, 17].

On the other hand, the present work directly relates to the literature on structural breaks in factor models. This literature shows that the assumption of constant factor loadings is not a neutral technical detail: the number of factors identified by standard procedures, the interpretation of latent common components, and the quality of subsequent empirical analysis depend on it. Particularly important for this study is the result according to which ignoring

structural shifts can lead to an overestimation of the number of factors and a distorted view of the internal structure of the data. This is where a natural point of intersection arises between the literature on determining the number of factors and the literature on structural instability [11].

More specifically, the present study is situated in the part of the break literature where the LR-based approach to testing for structural change in factor models plays a central role. In contrast to the more general background of literature devoted either to the choice of the number of factors or to early tests for the stability of loadings, this approach allows structural break to be understood as a change in the internal factor structure of the model in a narrower and econometrically meaningful sense. The work of Bai, Duan and Han is particularly important in this context, as it provides a methodological framework in which a break in loadings can be reparameterized as a change in the covariance structure of the factors, and the likelihood ratio test becomes a natural analysis tool. Thus, this study is related not just to the general topic of structural instability, but more specifically to a specialized strand of the literature focused on LR-based testing strategies and their advantages over alternative approaches [2].

Thus, this study is situated at the intersection of two strands of the literature: research on the definition of factor dimension and research on the structural instability of factor models. This intersection is especially important because structural shifts can affect the observed factor dimension itself and, consequently, the conclusions that the researcher draws about the nature of the latent common components. In this sense, the present study is not located on the periphery of the existing literature, but, on the contrary, is embedded in one of its most important points of contact, where the issues of factor selection, model specification and structural break testing are inextricably linked.

From the point of view of its contribution to the existing literature, the present study does not aim to develop a new general theory of factor models. Its aim is different: based on methods and results already developed in the literature, to show how the relevant tools can be used consistently to analyze the structural stability of a factor model and how the results of such an analysis can be interpreted in an empirical context. Therefore, the place of the present study in the literature can be described as applied and methodologically oriented: it is located within the established econometric framework, but at the same time it is aimed at the substantive use of

this framework to study the specific problem of structural change in factor models.

This is why the literature review does not serve merely as external background, but as a direct theoretical foundation for the subsequent analysis. The literature on factor estimation explains how the initial factor structure is constructed; the literature on determining the number of factors shows why choosing a dimension is a non-trivial task; the literature on structural breaks demonstrates that this structure can be unstable over time; and the LR-based strand of the break literature provides a specific methodological language in which such instability can be formally tested [2]. Collectively, all these strands determine the place of the present study in the modern literature on large-dimensional factor models.

1.3 Research objectives and structure of the thesis

This thesis uses large-dimensional factor models to study financial time series, with particular attention to the stability of the estimated factor structure. The empirical question is whether Russian stock returns over 2021–2025 can be described by one stable factor model, or whether the loading matrix changes during the sample period. The analysis is organized around three steps: estimating latent common factors, choosing the number of factors, and testing for a structural break in the factor loading matrix.

The first objective is to present the large-dimensional approximate factor model and to explain how the common component of a large panel can be estimated by Principal Component Analysis. Since the factors are not directly observed, PCA provides a practical way to recover the factor space from the observed data. This step is essential for the subsequent analysis, because both factor selection and structural break testing are based on the estimated factor representation.

The second objective is to determine the number of common factors in the data. The thesis considers the information criteria proposed by Bai and Ng, which are designed for large-dimensional panels and take into account both the cross-sectional dimension N and the time dimension T . This part of the analysis is important because the estimated number of factors determines the dimension of the factor space used in the remaining empirical procedures.

The third objective is to examine whether the factor loading matrix is stable over time. In the context of financial returns, a structural break in the loadings can be interpreted as a change

in the way individual assets are exposed to common market forces. The thesis therefore applies a likelihood ratio procedure for detecting a break in the factor structure. The purpose is not only to obtain a formal statistical indication of instability, but also to interpret the estimated break date in relation to the empirical behavior of the financial market.

The empirical part applies this sequence of steps to Russian stock market returns. I first prepare a balanced panel of weekly returns, then select the number of factors, estimate the factor structure, and compute the LR statistic over admissible candidate break dates. I also report graphical diagnostics, including scree plots and the profile of the LR statistic, because they help to interpret both the factor selection results and the location of the estimated break.

The rest of the thesis is organized as follows. Chapter 1 presents the motivation, literature review, and research objectives. Chapter 2 introduces the static approximate factor model, the PCA estimator, and the Bai–Ng criteria. Chapter 3 discusses structural breaks in factor models and presents the likelihood ratio procedure. Chapter 4 uses simulated data to examine the behavior of the procedure when the true factor structure and break date are known. Chapter 5 applies the method to Russian stock returns and discusses the baseline result, robustness checks, and the pre-break and post-break comparison. Chapter 6 concludes.

The applied contribution of the thesis is to put these elements together in one empirical workflow. Rather than studying factor estimation, factor selection, and break testing separately, the thesis uses them as consecutive steps of the same analysis. This makes it possible to move from the theoretical factor-model literature to a concrete empirical question: whether the dependence structure of Russian stock returns remained stable over the period considered.

Chapter 2

Factor models: theoretical framework

2.1 Factor model setup

Various studies have shown that the common dynamics of large datasets, such as asset returns, is often driven by a relatively small number of latent factors. Instead of analyzing each of the time series separately, a factor model allows one to represent the data as the sum of a common component and an idiosyncratic component.

In a stationary factor model, the observed variable $x_{i,t}$ can be written as

$$x_{i,t} = \lambda_i' F_t + e_{i,t}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad (2.1)$$

where F_t is a vector of latent common factors, λ_i is a vector of factor loadings, and $e_{i,t}$ is an idiosyncratic error.

Equivalently, for a given time period t , the model can be rewritten in vector form as

$$X_t = \Lambda F_t + e_t, \quad t = 1, \dots, T. \quad (2.2)$$

For the whole sample, the matrix representation is

$$X = F\Lambda' + e. \quad (2.3)$$

The idea of the model is to decompose the observed variable as the sum of two components.

The first part, $\lambda_i' F_t$, is called the common component and reflects the impact of factors that affect many time series simultaneously. The second part, $e_{i,t}$, is called the idiosyncratic component, and it captures fluctuations specific to each individual series.

In financial applications, this decomposition has a clear interpretation. Stock returns are usually affected by a few broad sources of risk, but individual assets do not react to these sources in the same way. The factor loadings capture precisely these differences in exposure: two assets may depend on the same latent factor, but with different intensity.

The main difficulty is that neither the factors nor the loadings are observed. The data contain only the matrix X , so both parts of the factor representation have to be recovered from the sample. In this thesis, this is done by Principal Component Analysis, which is the standard estimation method in large-dimensional factor models.

The analysis is conducted in a large-dimensional framework, where both the cross-sectional dimension N and the time dimension T are allowed to increase. This framework is appropriate for panels of financial returns, because it treats the common structure as something that can be estimated from a large number of assets observed over time. In this sense, the factor model is used both as a dimension-reduction device and as a way to describe the common component of the data.

2.2 Principal component estimation

As has been stated before, factors and factor loadings are not directly observable. Hence, for the practical implementation of the model, we must use some method that allows us to estimate the latent factor structure using the data matrix.

The most widely used method in the literature on factor models is principal component analysis. This method makes it possible to approximate the panel of observations with a low-rank representation, that is, to replace a large number of the original time series with a small number of factors that capture most of the common variation in the data.

Principal components analysis searches for those linear combinations of the observable variables that explain the largest share of the variation in the data. In the context of factor models, these linear combinations are interpreted as estimates of latent factors. In other words,

PCA identifies the directions along which the strongest comovement of the data is observed.

Let X be the $T \times N$ matrix of observations. Then the principal component estimator is built as the best low-rank approximation of this matrix. Formally, the problem can be formulated as the minimization of the sum of squared residuals between the observed matrix X and its factor representation $F\Lambda'$:

$$\min_{F, \Lambda} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (x_{i,t} - \lambda'_i F_t)^2 \quad (2.4)$$

or, equivalently, in matrix form:

$$\min_{F, \Lambda} \frac{1}{NT} \|X - F\Lambda'\|^2. \quad (2.5)$$

Since the factors are latent, the problem suffers from the identification problem: for any nonsingular matrix an equivalent representation with different factors and loadings can be obtained. Hence, it is necessary to impose the normalization conditions in estimation. One commonly used normalization is $\frac{F'F}{T} = I_r$, while $\Lambda'\Lambda$ is taken to be diagonal.

In practice, factor and loading estimates are extracted from the eigenvectors and eigenvalues of the covariance matrix of the data. In particular, factor estimates can be obtained from the eigenvectors of the matrix XX' , and the loading estimators from the corresponding projection of the original data onto the estimated factors. Thus, PCA transforms the problem of estimating latent factors into the problem of spectral decomposition of the data matrix.

If the factor matrix $\hat{F}$ is estimated under the normalization $\hat{F}'\hat{F}/T = I_r$, the estimated matrix $\hat{F}$ is obtained as $\sqrt{T}$ times the eigenvectors associated with the r largest eigenvalues of the matrix XX' and the corresponding loading estimator is given by

$$\hat{\Lambda} = \frac{X'\hat{F}}{T}. \quad (2.6)$$

The estimated common component can then be written as

$$\hat{C} = \hat{F}\hat{\Lambda}', \quad (2.7)$$

or, equivalently, in scalar form

$$\hat{C}_{it} = \hat{\lambda}'_i \hat{F}_t. \quad (2.8)$$

The importance of PCA in the literature on large-dimensional factor models is related to the fact that PCA provides consistent estimation of the factor space when both N and T are large. This result makes it possible to use the estimated factors for determining the number of factors, analyzing the data structure and testing for structural changes.

2.3 Determining the number of factors

After the factor model has been stated and the method of estimating the latent factors has been chosen, the question arises as to how many factors should be included in the model. This is a fundamental issue, since an incorrect choice of the dimensionality of the factor structure affects the quality of the subsequent analysis.

If the number of factors is underestimated, a significant part of the data variance remains unexplained. On the other hand, components reflecting idiosyncratic noise rather than the common structure may be included in the model if the number of factors is excessive. Thus, the problem of determining the number of factors is crucial in the empirical implementation of factor models.

Intuitively, the number of factors can be assessed from the behavior of the eigenvalues of the covariance matrix: the first several eigenvalues usually are noticeably larger than the others, since they reflect the common component, while the remaining ones mostly reflect the idiosyncratic part. However, this visual approach does not always provide an unambiguous answer, especially in finite samples. For this reason, several formal information criteria for determining the number of factors have been proposed in the literature [5].

The most influential approach was introduced by Bai and Ng [5]. The authors treat the determination of the number of factors as a model selection problem and propose criteria that balance the quality of the approximation and a penalty for excessive model dimensionality. Their result is of great importance, since standard criteria like AIC and BIC are not suitable for this problem, because the penalty must depend on both the cross-sectional and the time dimensions

of the panel in case of large N and T .

The idea of the Bai–Ng criteria is as follows. The factor model is estimated for each possible number of factors k . Then the value of the residual variation is calculated, which is the measure of what share of the variation remains unexplained after the extraction of k factors. This residual variation is then combined with a penalty term that increases with k . The optimal number of factors is chosen as the value that minimizes the corresponding criterion.

Let $V(k)$ denote the minimized normalized sum of squared residuals for a factor model with k factors. In the notations of Bai and Ng it is written as

$$V(k) = \min_{\Lambda^k, F^k} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left(x_{it} - \lambda_i^{k'} F_t^k \right)^2. \quad (2.9)$$

In other words, for each given k , $V(k)$ is the smallest residual variance that can be achieved by approximating the data with a k -factor representation. The minimization is taken over all admissible factor matrices F^k and loading matrices Λ^k . Thus, $V(k)$ measures the best possible fit of a factor model with k factors.

The value of $V(k)$ decreases as k grows, because a model with more factors can fit the data better. However, reducing the residual variance itself cannot serve as a selection criterion, since a model with too many factors begins to describe not only the common structure, but also noise. For this reason a penalty term is added to $V(k)$. In general, the information criteria can be written as

$$IC(k) = \log V(k) + k \cdot g(N, T), \quad (2.10)$$

where $g(N, T)$ is a penalty function that depends on the dimensionality of the panel.

Accordingly, the estimate of the number of factors is defined as

$$\hat{r} = \arg \min_{0 \leq k \leq k_{\max}} IC(k), \quad (2.11)$$

where $k_{\max}$ is a predefined maximum admissible number of factors.

Bai and Ng propose several concrete criteria with different forms of the penalty function.

The most commonly used information criteria are the following:

$$IC_{p1}(k) = \log V(k) + k \left(\frac{N+T}{NT} \right) \log \left(\frac{NT}{N+T} \right), \quad (2.12)$$

$$IC_{p2}(k) = \log V(k) + k \left(\frac{N+T}{NT} \right) \log C_{NT}^2, \quad (2.13)$$

$$IC_{p3}(k) = \log V(k) + k \left(\frac{\log C_{NT}^2}{C_{NT}^2} \right), \quad (2.14)$$

where

$$C_{NT} = \min\{\sqrt{N}, \sqrt{T}\}. \quad (2.15)$$

These criteria are constructed in such a way as to provide consistent estimation of the number of factors with large N and T .

Along with the information criteria Bai and Ng also propose panel criteria in the form of

$$PC(k) = V(k) + k \cdot g(N, T), \quad (2.16)$$

where the penalty is added not to the logarithm of the residual variance but directly to the value of $V(k)$ itself. In particular, the paper proposes the following options:

$$PC_{p1}(k) = V(k) + k \cdot \hat{\sigma}^2 \left(\frac{N+T}{NT} \right) \log \left(\frac{NT}{N+T} \right), \quad (2.17)$$

$$PC_{p2}(k) = V(k) + k \cdot \hat{\sigma}^2 \left(\frac{N+T}{NT} \right) \log C_{NT}^2, \quad (2.18)$$

$$PC_{p3}(k) = V(k) + k \cdot \hat{\sigma}^2 \left(\frac{\log C_{NT}^2}{C_{NT}^2} \right), \quad (2.19)$$

where $\hat{\sigma}^2 = V(k_{max})$ serves as an estimate of the idiosyncratic component variance.

In practice, both the IC_j family and PC_j family are used to choose the number of factors. The final choice can be checked for robustness with respect to the particular criterion used.

The theoretical importance of the Bai and Ng approach lies in the fact that the proposed criteria are consistent. In other words, provided that the relevant conditions are satisfied, the probability that the estimated number of factors coincides with the true one converges to one as

N and T tend to infinity. In particular, consistency is obtained under suitable conditions on the penalty term, notably $g(N, T) \rightarrow 0$ and $C_{NT}^2 g(N, T) \rightarrow \infty$. This property makes the Bai–Ng criteria a standard tool for the large-dimensional factor analysis.

The formal proof of consistency is derived under the large- N , large T asymptotic framework under assumptions on the factors, factor loadings, and idiosyncratic errors. In particular, the analysis requires regularity conditions ensuring that the common and idiosyncratic components are sufficiently well behaved as both dimensions of the panel increase [5]. Since the present thesis focuses on the implementation and application of the procedure, only the main criteria and the corresponding consistency result are reported here.

For the purposes of this thesis, determining the number of factors is of crucial significance because the subsequent analysis is based on the estimated factors. In particular, testing for structural breaks in a factor model requires a preliminary choice of the dimensionality of the factor space. Therefore, correct estimation of the number of factors is a necessary step prior to the analysis of the stability of the factor structure.

Chapter 3

Structural breaks in factor models

3.1 Structural instability in factor models

In the standard formulation of a factor model, it is usually assumed that the factor loading matrix remains constant throughout the sample. This assumption is convenient from both a theoretical and a computational point of view, since it allows one to describe the observable data within a single stable factor structure. However, in empirical applications, especially in financial data, the hypothesis of constant loadings may turn out to be too strong.

Financial time series are not usually generated by a single stable environment. Crises, sanctions, changes in monetary conditions and other systemic shocks may change both individual asset behaviour and the way assets move together. In the language of factor models, this means that the exposure of observed variables to latent factors may vary over time. Such a change is represented by a non-constant factor loading matrix.

When the loading matrix changes, a factor model estimated on the whole sample may become difficult to interpret. The pre-break and post-break observations can correspond to different regimes, even if the number of latent factors is kept fixed. This matters for the estimation of factors and loadings, for the selected number of factors, and for the interpretation of the common component. For this reason, testing the stability of the loading matrix is an important step before drawing conclusions from the estimated factor structure.

For this reason, the problem of detecting structural breaks occupies an important place in the literature on factor models. In this thesis we concentrate on a possible break in the factor

loading matrix. Such a setting makes it possible to examine whether the factor structure of the Russian stock market remains stable over the entire sample, or whether the effect of common latent factors on the returns of individual assets changes at some moment in time.

In what follows, a factor model with a possible single structural break in the loading matrix is considered, and the hypothesis of factor-structure stability is tested using the likelihood ratio approach.

3.2 Breaks in the factor loading matrix

For the formal analysis of structural instability, we consider a factor model in which at some unknown point in time a single change in the factor loading matrix occurs. At the same time, it is assumed that before and after the break point the data are described by the same factor structure with the same number of factors, but with different loadings.

A factor model with a break can be written as follows:

$$x_t = \begin{cases} \Lambda_1 f_t + e_t, & t = 1, \dots, T_1, \\ \Lambda_2 f_t + e_t, & t = T_1 + 1, \dots, T, \end{cases} \quad (3.1)$$

where x_t is an $N \times 1$ vector of observed variables, f_t is a $r \times 1$ vector of latent factors, Λ_1 and Λ_2 are the pre-break and post-break loading matrices, e_t is an idiosyncratic component, and T_1 is an unknown break date.

Equivalently, in scalar form, the model can be written as:

$$x_{i,t} = \begin{cases} \lambda'_{i1} f_t + e_{it}, & t = 1, \dots, T_1, \\ \lambda'_{i2} f_t + e_{it}, & t = T_1 + 1, \dots, T. \end{cases} \quad (3.2)$$

For convenience, the break fraction π_1 is also used, which represents the relative position of the break within the sample:

$$\pi_1 = \frac{T_1}{T}, \quad \pi_1 \in (0, 1), \quad (3.3)$$

Equation (3.1) means that before time T_1 the observed variables are related to the latent

factors through the matrix Λ_1 , whereas after the break this relationship is governed by the matrix Λ_2 . If $\Lambda_1 \neq \Lambda_2$, the way in which latent factors affect the observed data changes, and this is interpreted as a structural break.

The aim of the test is to examine the hypothesis of stability of the factor loading matrix throughout the sample. The corresponding hypotheses can be written as:

$$H_0 : \Lambda_1 = \Lambda_2 \quad (3.4)$$

against

$$H_1 : \Lambda_1 \neq \Lambda_2. \quad (3.5)$$

The null hypothesis corresponds to a stable factor structure, while the alternative hypothesis implies a single structural break in the factor loading matrix.

The problem of testing (3.4) against (3.5) is high-dimensional, because the number of parameters in the loading matrix increases with the dimensionality of the panel. However, a key result in the literature is that the factor model with a structural break in the loading matrix can be rewritten in an equivalent form as a model with constant loadings but a change in the moments of an expanded vector of the factors. This idea makes it possible to transform the original high-dimensional problem into a low-dimensional one and to construct a likelihood ratio approach for testing structural instability [2].

3.3 Likelihood ratio approach

The key idea of the likelihood ratio approach is that a factor model with a break in the factor loading matrix can be represented in an equivalent form as a model with constant loadings but with a change in the moments of an expanded vector of factors. Owing to this equivalence, the original problem of testing for a structural break in a high-dimensional loading matrix is reduced to the analysis of low-dimensional characteristics of the pseudo-factors. This idea underlies the LR approach proposed by Bai, Duan, and Han [2].

For the model (3.1) with one break in the loading matrix, the two regimes can be represented

as

$$\Lambda_1 = \Lambda B, \quad \Lambda_2 = \Lambda C.$$

Here Λ is an $N \times r$ loading matrix, common to both regimes and with full column rank. The matrices B and C are regime-specific transformation matrices such that

$$\text{rank}(B) = r_1 \leq r, \quad \text{rank}(C) = r_2 \leq r.$$

The quantities r_1 and r_2 therefore correspond to the numbers of factors in the two regimes, whereas r denotes the number of pseudo-factors in the equivalent representation.

For a given break point T_1 , if $X_{T_1}^{(1)} = [x_1, \dots, x_{T_1}]'$ and $X_{T_1}^{(2)} = [x_{T_1+1}, \dots, x_T]'$ denote the pre-break and post-break blocks of observations and $F_{T_1}^{(1)} = [f_1, \dots, f_{T_1}]'$ and $F_{T_1}^{(2)} = [f_{T_1+1}, \dots, f_T]'$ define the pre-break and post-break blocks of factors, the model (3.1) can be written in block form as:

$$\begin{bmatrix} X_{T_1}^{(1)} \\ X_{T_1}^{(2)} \end{bmatrix} = \begin{bmatrix} F_{T_1}^{(1)} B' \\ F_{T_1}^{(2)} C' \end{bmatrix} \Lambda' + e = G \Lambda' + e, \quad (3.6)$$

where G is the matrix of the so-called pseudo-factors. The matrix G is obtained by combining the transformed factor blocks before and after the break:

$$G = \begin{bmatrix} F_{T_1}^{(1)} B' \\ F_{T_1}^{(2)} C' \end{bmatrix}, \quad (3.7)$$

In such a representation, the loadings become constant, and the information about the structural change is transferred to the properties of the factors G , in particular to their second moments.

Thus, the hypothesis of the absence of a structural break in the loading matrix is equivalent to the hypothesis of no change in the covariance structure of the pseudo-factors. If there is no break, the pre-break and post-break covariance matrices coincide. If there is a break, these matrices differ. As a result, the problem of testing whether $\Lambda_1 = \Lambda_2$ is transformed into the problem of comparing the second moments before and after the candidate break date.

Since the true pseudo-factors are unobserved, the principal components estimator of G is used in practice. Let $\hat{G} = [\hat{g}_1, \dots, \hat{g}_T]'$ denote the principal components estimator of G , normalized so that $T^{-1} \hat{G}' \hat{G} = I_r$.

For each candidate break point k , covariance matrix estimates are constructed on the two subsamples:

$$\hat{\Sigma}_1(k) = \frac{1}{k} \sum_{t=1}^k \hat{g}_t \hat{g}_t', \quad (3.8)$$

$$\hat{\Sigma}_2(k) = \frac{1}{T-k} \sum_{t=k+1}^T \hat{g}_t \hat{g}_t'. \quad (3.9)$$

These matrices represent the estimated second moments of the factors before and after the candidate break point.

Based on these matrices, a quasi-Gaussian log-likelihood is constructed for a fixed k :

$$L(k) = k \log |\hat{\Sigma}_1(k)| + (T-k) \log |\hat{\Sigma}_2(k)|. \quad (3.10)$$

Under the null hypothesis of no break, the full-sample covariance matrix of the normalized estimated factors is equal to the identity matrix. Therefore, the corresponding no-break log-likelihood is

$$L_0 = T \log \left| \frac{1}{T} \sum_{t=1}^T \hat{g}_t \hat{g}_t' \right| = T \log |I_r| \equiv 0. \quad (3.11)$$

The likelihood ratio statistic for the candidate break point k is then defined as $LR(k) = L_0 - L(k)$.

Thus, the likelihood ratio statistic for the candidate k takes the following form:

$$LR(k) = -k \log |\hat{\Sigma}_1(k)| - (T-k) \log |\hat{\Sigma}_2(k)|. \quad (3.12)$$

If the covariance structure of the estimated factors differs substantially across the two subsamples, the value of $LR(k)$ becomes large, which indicates a possible structural break near this date.

Since the date of the break is not known in advance, the testing procedure considers all the

admissible values of k within a certain interval, which excludes excessively small subsamples at the boundaries. The corresponding sup-LR statistic is defined as follows:

$$\sup-LR = \sup_{[\varepsilon T] \leq k \leq [(1-\varepsilon)T]} LR(k), \quad (3.13)$$

where $\varepsilon \in (0, 0.5)$ is a trimming parameter.

The estimated break date is defined as the value of k at which the statistic $LR(k)$ reaches its maximum. It should be emphasized that the maximization in (3.13) always produces a candidate break date. Therefore, the existence of a maximizer alone cannot be interpreted as evidence of a structural break. A break is detected only if the observed value of the sup-LR statistic exceeds the corresponding critical value under the null hypothesis.

From the theoretical point of view the important property of the LR approach is that it exploits a special property of a factor model with a break, that is, under the alternative hypothesis (3.5) the pre-break and post-break covariance matrices of the pseudo-factors may become singular. This particular property leads to a faster growth of the LR statistic and, as a consequence, greater power than Wald- and LM-type approaches based on the same estimated factors. The authors of the article [2] demonstrate that under the null hypothesis (3.4) the LR test is asymptotically equivalent to the sup-Wald test, but under the alternative it enjoys more favorable power properties.

For the purposes of the present thesis, the LR test is of particular interest because it allows not only the detection of a break, but also the estimation of its date. This makes it a natural tool for empirical analysis of the stability of the factor structure of the Russian stock market.

3.4 Limiting distribution of the sup-LR statistic under the null hypothesis

After defining the statistic $LR(k)$ for a fixed candidate break point and the corresponding sup-LR procedure, it is necessary to characterize its limiting distribution under the null hypothesis of no structural break. This distribution is fundamental for constructing the critical values and, therefore, practical implementation of the test.

The asymptotic theory of the LR test is developed in a large- N , large- T asymptotic framework

for the approximate factor models. The formal result requires regularity conditions on the factor moments, loading matrix, dependence structure of the idiosyncratic errors, and a functional central limit theorem for the processes built on the second moments of the factors.

In addition, in the research of Bai, Duan and Han the following condition is of crucial importance:

$$\frac{\sqrt{T}}{N} \rightarrow 0, \quad (3.14)$$

since it ensures the validity of the limiting argument for the LR statistic under the null hypothesis [2].

If the corresponding conditions are satisfied, the sup-LR statistic has a nontrivial limiting distribution. That is, under H_0 , it holds that

$$\sup_{[\varepsilon T] \leq k \leq [(1-\varepsilon)T]} LR(k) \Rightarrow \frac{1}{2} \sup_{\pi \in [\varepsilon, 1-\varepsilon]} \frac{[W(\pi) - \pi W(1)]' \Omega [W(\pi) - \pi W(1)]}{\pi(1-\pi)}, \quad (3.15)$$

where $W(\pi)$ is a vector of independent Brownian motions, and Ω is a long-run covariance matrix associated with the process built on the second moments of the normalized factors [2]. Therefore, critical values cannot be read directly from standard chi-square tables and must be obtained from the corresponding limiting distribution or its numerical approximation.

The expression $W(\pi) - \pi W(1)$ represents a multidimensional Brownian bridge. Thus, the limiting distribution (3.15) has the standard form of a sup-type structural break statistic: the statistic depends on the maximum quadratic functional of a Brownian bridge on the interval $[\varepsilon, 1 - \varepsilon]$. However, in this particular problem the matrix Ω , which reflects the long-run dependence in the process generated by the second moments of the normalized factors, plays an additional role.

The following normalized process is introduced to define the matrix Ω more precisely:

$$\eta_t = \Sigma_G^{-1/2} g_t, \quad (3.16)$$

where Σ_G is the covariance matrix of the pseudo-factors. Thus, the matrix Ω is defined as

follows:

$$\Omega = \text{plim}_{T \rightarrow \infty} \text{Var} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \text{vec}(\eta_t \eta_t' - I_r) \right]. \quad (3.17)$$

Thus, Ω characterizes the long-run variance structure of the process associated with the deviations of the second moments of the factors from their value under the null hypothesis [2].

Since Ω is unknown, its consistent estimate is necessary for the practical implementation of the test. For this purpose, Bai, Duan, and Han use a HAC (heteroskedasticity and autocorrelation consistent) estimator of the form

$$\hat{\Omega} = \hat{\Gamma}_0 + \sum_{j=1}^{T-1} K\left(\frac{j}{S_T}\right) (\hat{\Gamma}_j + \hat{\Gamma}_j'), \quad (3.18)$$

where $K(\cdot)$ is a kernel function used to weight sample autocovariances at different lags, while S_T is a bandwidth parameter that determines how many lags are effectively taken into account in the HAC estimation, and

$$\hat{\Gamma}_j = \frac{1}{T} \sum_{t=j+1}^T \text{vec}(\hat{g}_t \hat{g}_t' - I_r) \text{vec}(\hat{g}_{t-j} \hat{g}_{t-j}' - I_r)'. \quad (3.19)$$

This estimator allows one to substitute the unknown matrix Ω in the limiting distribution of the statistic with its sample analogue. In practical implementation, the choice of the kernel and the bandwidth affects the estimation of Ω and, therefore, the finite-sample approximation of the null distribution.

In the special case where the factors are i.i.d. Gaussian, under the null hypothesis, the limiting distribution simplifies and takes a more familiar form similar to that of standard sup-type structural break tests. In this case, the LR test is asymptotically equivalent to the sup-Wald test under the null. However, under the alternative hypothesis the LR approach enjoys more favorable power properties, which is due to the possible singularity of the pre-break and post-break covariance matrices of the pseudo-factors. This particular property leads to a faster growth of the LR statistic in the presence of a break [2].

Thus, the limiting theory of the LR test under the null provides the basis for constructing the critical values, while its behavior under the alternative explains why this approach is particularly attractive for the empirical analysis of the structural instability of factor models.

Chapter 4

Simulation study

4.1 Data generating process

In this section, a simulation study is conducted in order to assess the performance of the procedure in a controlled environment. The use of artificially generated data allows us to specify the true factor structure, to determine in advance whether a structural break is present, and to compare the estimated results with the known parameters of the model. In particular, the simulations make it possible to evaluate whether the procedure determines the number of factors correctly, detects a break, and estimates its date accurately.

As the basic data-generating framework, we consider the static large-dimensional factor model, namely:

$$x_t = \Lambda f_t + e_t, \quad t = 1, \dots, T \quad (4.1)$$

where x_t is an $N \times 1$ vector of observed variables, f_t is an $r \times 1$ vector of latent factors, Λ is an $N \times r$ matrix of factor loadings, and e_t is an $N \times 1$ vector of idiosyncratic errors.

In the simulations, the true number of factors is fixed at $r = 2$, allowing us to directly compare it with the number of factors chosen by the Bai–Ng information criteria. In all the considered scenarios, the sample dimensions are set to $T = 300$ and $N = 200$. Thus, we study a panel with a moderate time dimension and a relatively large cross-sectional dimension which is consistent with the general framework of large-dimensional factor models. Latent factors are

generated as random variables with zero mean and variance $\sigma_f^2 = 1$, while idiosyncratic errors have zero mean and variance σ_e^2 , whose value varies across scenarios. The baseline value is taken as $\sigma_e = 0.5$ while in one of the scenarios a noisier environment with $\sigma_e = 1$ is studied.

The main focus of the simulation study is the behavior of the procedure in the presence and absence of a structural break in the factor loading matrix. For this purpose we consider four scenarios. The first scenario is a control scenario and corresponds to the absence of a break: the loading matrix is constant throughout the entire sample. In the second scenario, a structural break is introduced at time $T_1 = 200$. In the third scenario, the break remains, but it becomes weaker. Finally, in the fourth scenario the break remains, but the variance of the idiosyncratic error increases, which makes it possible to assess the robustness of the procedure in noisier data. The four scenarios are designed to cover both the null case and several alternative configurations differing in the magnitude of the break and the level of idiosyncratic noise.

In the break scenarios, the loading matrix changes after T_1 . The pre-break observations are generated using Λ_1 , while the post-break observations are generated using Λ_2 . The parameter `loading_shift` controls the magnitude of this change. Hence, the simulations do not change the number of latent factors; instead, they change the way these factors enter the observed variables after the break date. This corresponds to the structural-break specification introduced in Chapter 3.

The set of the considered scenarios is constructed so as to evaluate several aspects of the performance of the procedure. The scenario without a break makes it possible to assess its behavior under the null hypothesis and verify whether a false signal of structural instability arises. The basic scenario with a break demonstrates how the method works in a favorable environment, where the change in the factor structure is sufficiently pronounced. The third scenario with a moderate break allows us to evaluate the sensitivity of the procedure to a lower magnitude of the structural change. Finally, the scenario with a higher level of idiosyncratic noise is intended to analyze how a lower signal-to-noise ratio influences the choice of the number of factors and the ability of the LR approach to detect a break.

For each of the four scenarios we consider both illustrative single runs and full Monte Carlo experiments. Within the Monte Carlo design, 100 independent replications are performed for

each scenario. For each replication the results of factor-number selection using Bai–Ng criteria, the final value of $\hat{r}$ used in the PCA/LR procedure, the value of the sup-LR statistic, the estimated break date, the corresponding HAC critical values, and the formal decision of rejecting or not rejecting the null hypothesis on the standard levels of significance are saved. This design makes it possible to move from a single illustrative run to a systematic evaluation of the properties of the method.

For each scenario, the following summary characteristics are analyzed: the accuracy of factor-number selection, the rejection frequency under the null hypothesis of no break, and the accuracy of break-date estimation in the scenarios where a break is present. This makes it possible to simultaneously study the size and power properties of the procedure, as well as its sensitivity to different parameters of the simulation design.

4.2 Estimation and testing procedure

After defining the data generating mechanism, it is necessary to describe the procedure applied to each artificially generated sample. In all the considered scenarios, the estimation and testing are performed according to the same scheme, which makes it possible to compare the results of different configurations of the simulation design.

In the first stage, the number of factors is determined for each generated sample. For this purpose, the Bai–Ng criteria are used, that is, the information criteria (2.12)–(2.14) and the panel criteria (2.17)–(2.19). For each criterion the corresponding value of $\hat{r}$ is calculated as in (2.11), and after that, one final number of factors is chosen from the six computed values. In the implemented computer program the IC_{p2} (2.13) criterion is used as the main one, and the corresponding value of $\hat{r}$ is used in the sequential stages of PCA estimation and LR testing. Such an approach makes it possible both to preserve the link with the full Bai–Ng system of criteria and to ensure the uniformity of the final procedure in the Monte Carlo experiment.

In the second stage, the factor structure is then estimated by principal component analysis using the selected value of $\hat{r}$. For each sample estimates of factor and loading matrices are obtained, and the share of explained variance is recorded. This makes it possible not only to implement the subsequent testing procedure, but also to control how well the estimated factor

model explains the generated data in each scenario.

In the third stage the LR statistic of the structural break detection is constructed for each sample. For this purpose all admissible values of k within a trimmed interval are considered. The interval is defined by the trimming parameter. In the basic simulation study configuration the trimming parameter is set to 0.15, so that candidates close to the boundaries of the sample are excluded. For each admissible value of k the pre-break and post-break covariance matrices of the estimated factors are calculated, then the statistic $LR(k)$ (3.12) is constructed, and, finally, its maximum over all k is determined. Thus, the sup-LR statistic (3.13) is computed for each sample, and the corresponding break date is defined as the value of k , where the statistic reaches its maximum.

In the fourth stage, for each replication the critical values of the test are calculated. Since the limiting distribution of the sup-LR statistic does not reduce to classical tables like χ^2 , critical values are obtained separately on the basis of the asymptotic theory and HAC estimation of the long-run covariance matrix. In the implemented simulation framework a numerical approximation is used, based on simulating the limiting distribution on a Brownian motion grid and on a HAC estimate of the matrix Ω . In the baseline configuration of the Monte Carlo experiment the following values are used: the trimming parameter is set to 0.15, the grid length for the Brownian motion approximation is 2000, and the number of simulations used to estimate the critical values is 1000 at the 10%, 5% and 1% significance levels.

After obtaining the observed value of the sup-LR statistic and the corresponding critical values, a formal decision regarding the presence or absence of a structural break is made. The null hypothesis (3.4) of no structural break is rejected if the observed value of sup-LR statistic exceeds the critical value at the chosen significance level. Thus, for each replication not only the value of the statistic is saved, but also the formal decision to reject or not to reject the null hypothesis at the levels 10%, 5% and 1%.

Such an organization of the calculations allows us to obtain, for each generated sample, a complete set of results, including: the values of the criteria for determining the number of factors, the chosen number of factors, the corresponding PCA output, the value of the sup-LR statistic, the estimated break date, the critical values, and the final decision of the test. Aggregating these

results over 100 replications makes it possible to pass from individual examples to a systematic evaluation of the properties of the procedure in finite samples.

Thus, the simulation study is constructed as the repeated application of the same procedure to different artificially specified data configurations. This ensures comparability across scenarios and makes it possible to analyze the behavior of the test under the null hypothesis, in the presence of a clear break, when the structural break is weaker and when the level of idiosyncratic noise is higher.

4.3 Illustrative single-run results

For an initial illustration of the behavior of the procedure, we first consider illustrative runs in four basic configurations: without a break, with a baseline break, with a moderate break, and with higher idiosyncratic noise in the presence of a break. Such single runs do not allow conclusions to be drawn about the systematic properties of the procedure, but they do show how factor-selection criteria, sup-LR statistic, and the estimated break date behave in a specific sample. The summary results for the four scenarios are reported in Table 4.1.

All four illustrative scenarios are generated with $T = 300$, $N = 200$, and the true factor dimension $r = 2$.

4.3.1 Scenario 1. No-break benchmark

In the first illustrative example, we generate data without a structural break, that is, the factor loading matrix remains constant throughout the sample. In this scenario, the idiosyncratic noise level is set to $\sigma_e = 0.5$, and the parameter `loading_shift` is equal to 0, which corresponds to the null hypothesis of no break.

The results of this run demonstrate that all the Bai–Ng criteria correctly choose the number of factors equal to 2. Thus, the final value of $\hat{r}$ used for PCA estimation and LR testing is equal to 2 as well. This is consistent with the true structure of the model and provides a first indication that, in the absence of a break, the procedure correctly recovers the factor dimension.

The estimated factor model in this case explains a substantial share of the variation in the

data: the explained variation ratio is approximately 0.886. The spectrum of the first eigenvalues also corresponds to the expected pattern: the two leading eigenvalues significantly exceed the others, while the subsequent values remain small. This confirms the presence of two pronounced common components and the absence of artificial inflation of the factor structure.

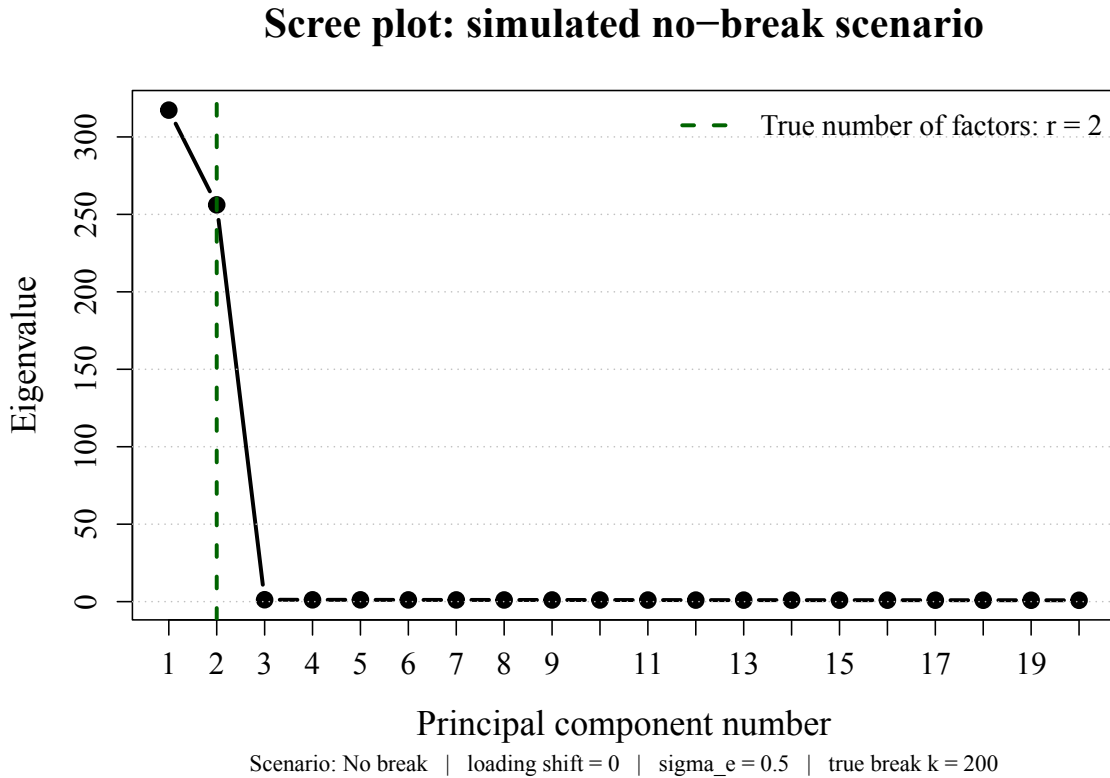


Figure 4.1: Scree plot for the simulated no-break scenario

In this run, the value of the sup-LR statistic is relatively small, amounting to approximately 5.76. The corresponding HAC-critical values at levels 10%, 5% and 1% exceed this value, and therefore the null hypothesis of no structural break is not rejected. Formally, the algorithm still returns the value of k at which $LR(k)$ attains its maximum, and in this run this date corresponds to the 143rd observation. However, this result should not be interpreted as a genuine break date, since the test does not provide grounds for rejecting the null hypothesis. Thus, this example clearly shows an important practical point: the existence of a maximum of the LR statistics is a mechanical property of the procedure and does not by itself imply the presence of a structural change.

The corresponding scree plot and LR path are reported in Figures 4.1 and 4.2, respectively.

LR statistic in the simulated no-break scenario

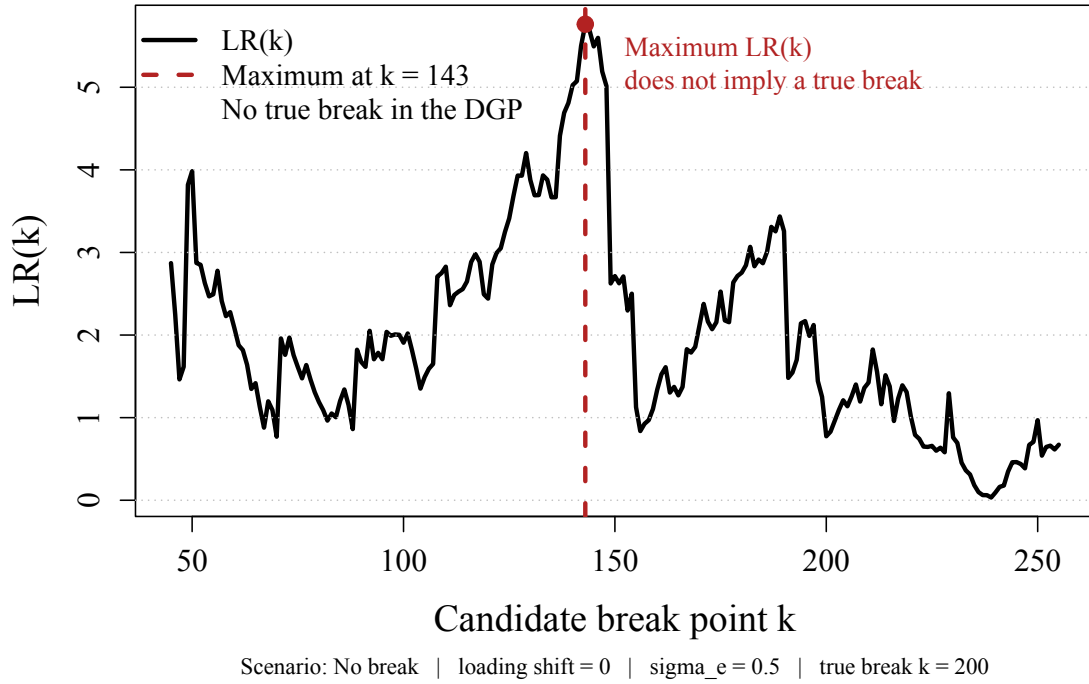


Figure 4.2: LR statistic in the simulated no-break scenario

All in all, in the no-break scenario the test behaves in an expected way: the number of factors is determined correctly, LR statistic remains below the critical values, and the null hypothesis is not rejected. This result is important as a basic verification of the correctness of the method in the absence of structural instability.

4.3.2 Scenario 2. Baseline break

In the second illustrative example, a structural break in the factor loading matrix is introduced into the model. The idiosyncratic noise is again set to $\sigma_e = 0.5$. The break is introduced at time $T_1 = 200$, and its magnitude is determined by the parameter `loading_shift` = 1. Therefore, in this configuration, the alternative hypothesis of a structural change is true.

For the baseline break scenario, the output is quite different from the no-break case. Although the data are generated with $r = 2$ latent factors, all Bai–Ng criteria select three factors. This upward shift is consistent with the idea that a break in the loading matrix makes the pooled sample look more complex. When the two regimes are combined, the change in loadings is partly

captured as an additional common component rather than only as a change in the relationship between factors and observed variables.

Scree plot: simulated baseline break scenario

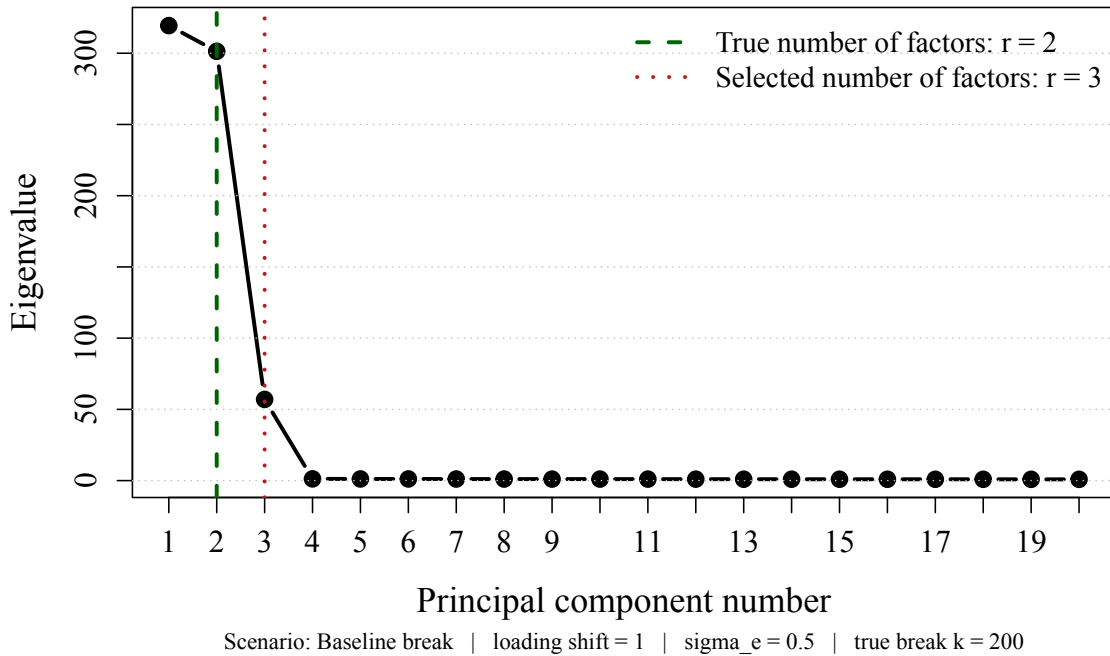


Figure 4.3: Scree plot for an illustrative single run in the baseline break scenario

The LR procedure nevertheless identifies the break very clearly. In this run, the estimated break date is exactly $T_1 = 200$, which coincides with the true break date used in the simulation. The observed sup-LR statistic is approximately 1381.77, far above the critical values at the conventional significance levels. Therefore, the null hypothesis of no structural break is rejected.

It should be noted, that in this scenario the share of the explained variation remains high and amounts to about 0.895. This means that the estimated factor model still describes the data well, but in a more complex pooled representation, where the presence of a structural break affects both the choice of the number of factors and the behavior of LR statistics.

The behavior of the eigenvalues and of the LR statistic in this illustrative run is shown in Figures 4.3 and 4.4.

This example demonstrates that in a favorable configuration with a sufficiently strong break, the test in question not only rejects the null hypothesis, but also almost exactly recovers the break

LR statistic in the simulated baseline break scenario

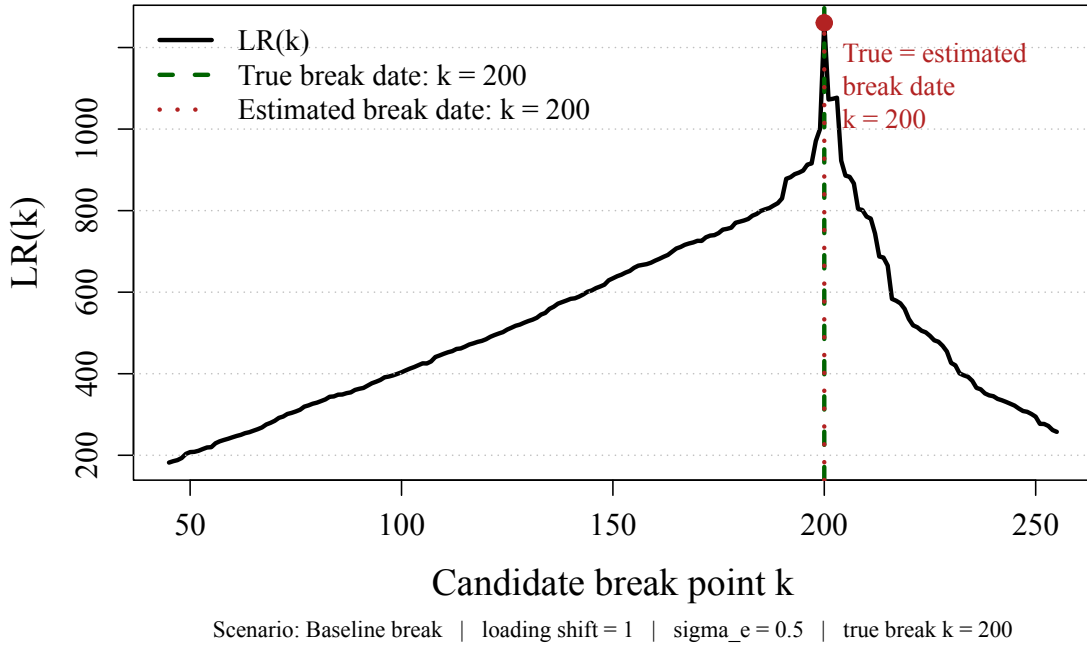


Figure 4.4: LR statistic in the simulated break scenario

date. At the same time, it shows that the presence of a break can distort the standard procedure for selecting the number of factors, leading to their overestimation. This makes it especially important to consider not only individual illustrative runs, but also systematic Monte Carlo results. Hence, even in a single realization, the break scenario already suggests that structural instability may distort factor-number selection while remaining clearly detectable by the LR procedure.

4.3.3 Scenario 3. Moderate break

The third scenario considers a weaker structural change. In contrast to the baseline break case, the parameter `loading_shift` is reduced from 1 to 0.5, while the level of idiosyncratic noise remains unchanged at $\sigma_e = 0.5$. Thus, the break in the loading matrix is still present, but the difference between the pre-break and post-break regimes is less pronounced.

The scree plot in Figure 4.5 shows that the Bai–Ng selection procedure still chooses $\hat{r} = 3$, although the true number of factors in the data-generating process is $r = 2$. This is the same

pattern as in the baseline break scenario. The additional selected factor can again be interpreted as a consequence of pooling observations from two regimes with different loading matrices.

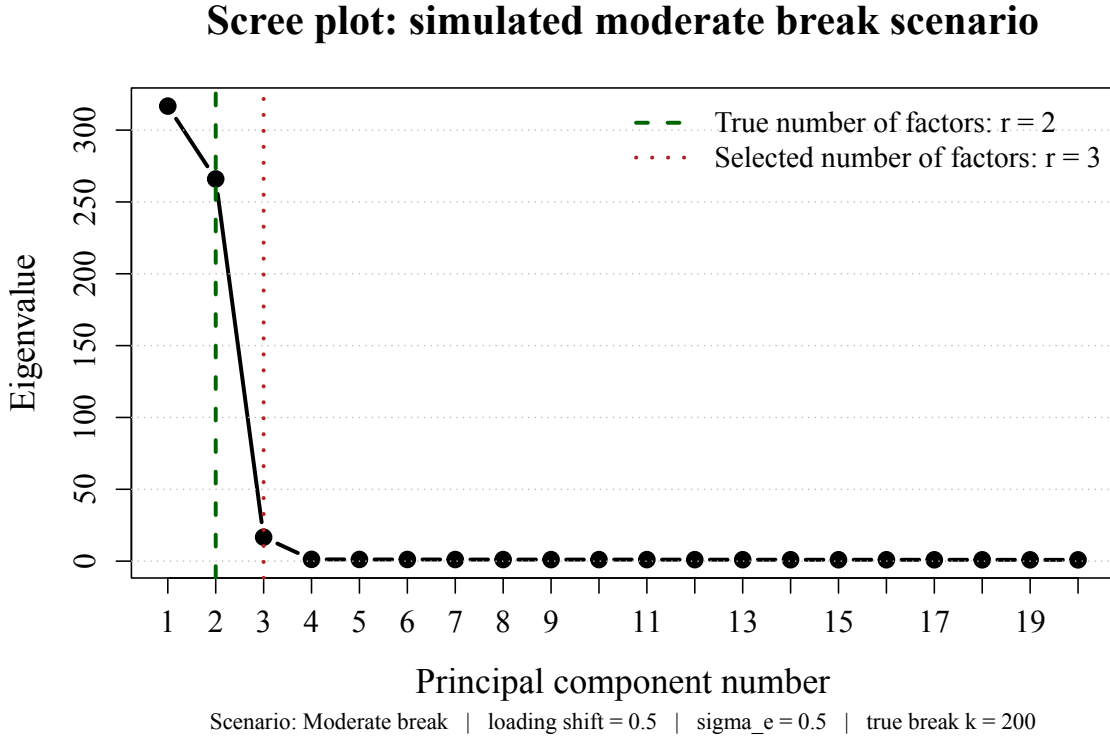


Figure 4.5: Scree plot for an illustrative single run in the moderate break scenario

The LR path in Figure 4.6 confirms that the weaker break is still clearly detected. The maximum of the statistic is attained at $k = 200$, which coincides with the true break date used in the simulation. Although the loading shift is smaller than in the baseline case, the peak of the LR statistic remains visually sharp and well separated from the surrounding values. This indicates that, in this illustrative run, the procedure remains sensitive even to a moderate change in the loading matrix.

4.3.4 Scenario 4. Higher idiosyncratic noise with break

The fourth scenario keeps the stronger loading shift, `loading_shift = 1`, but increases the standard deviation of the idiosyncratic component to $\sigma_e = 1$. Compared with the baseline break scenario, the common component is therefore observed in a noisier environment. This setting is useful because it checks whether the LR procedure still identifies the break when the

LR statistic in the simulated moderate break scenario

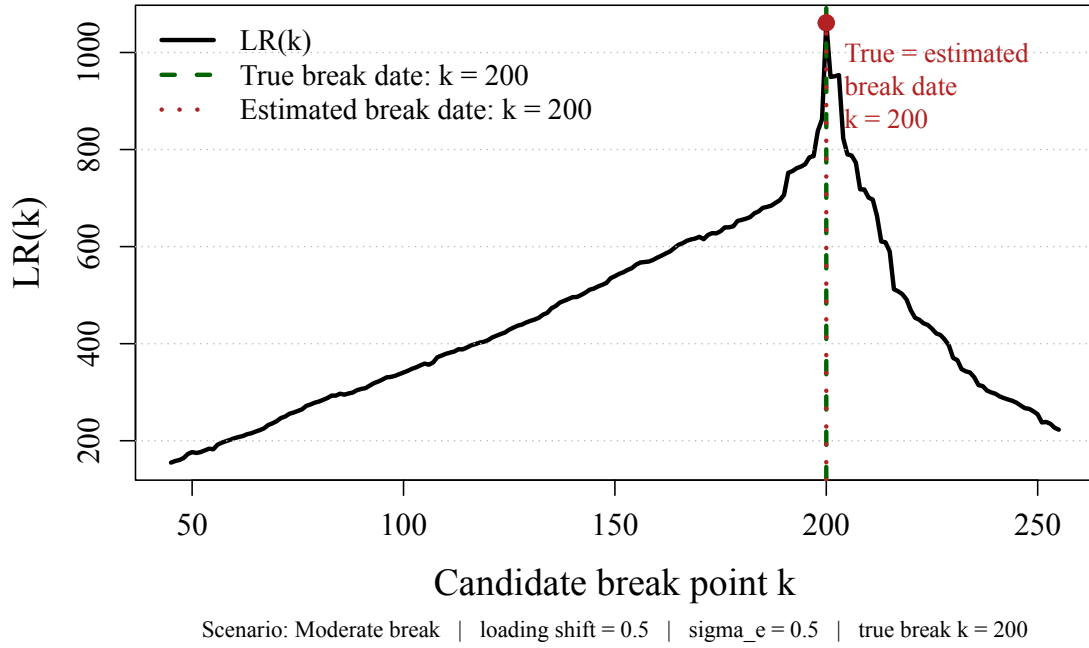


Figure 4.6: LR statistic in the simulated moderate break scenario

signal-to-noise ratio is lower.

The scree plot in Figure 4.7 again leads to the selection of $\hat{r} = 3$, while the true number of latent factors is $r = 2$. At the same time, the spectrum is less clean than in the lower-noise cases, which is consistent with the larger idiosyncratic variance. The increase in noise makes the separation between common and idiosyncratic variation less favorable, but it does not eliminate the upward shift in the selected factor dimension caused by the break.

The LR statistic in Figure 4.8 still reaches its maximum at the true break date $k = 200$. This means that, even with higher idiosyncratic noise, the break in the loading matrix remains visible in the estimated factor structure. The result is important for the interpretation of the procedure, since real financial data are usually noisy and rarely correspond to the idealized low-noise setting. In this illustrative case, the LR approach continues to localize the break correctly despite the less favorable simulation environment.

Scree plot: simulated noisier break scenario

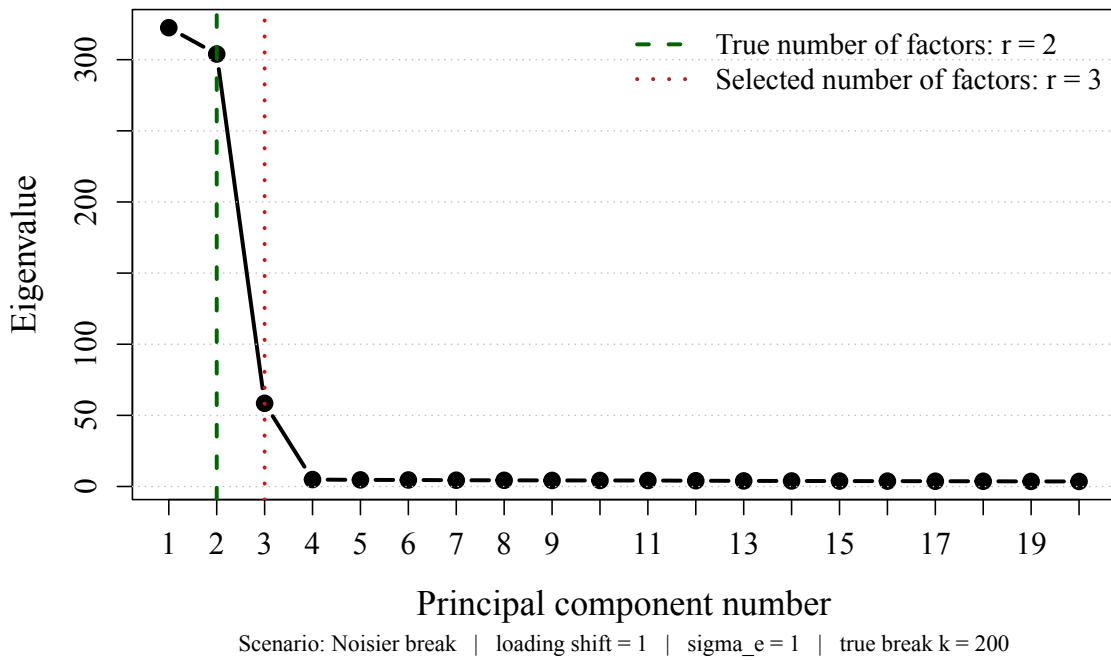


Figure 4.7: Scree plot for an illustrative single run in the noisier break scenario

LR statistic in the simulated noisier break scenario

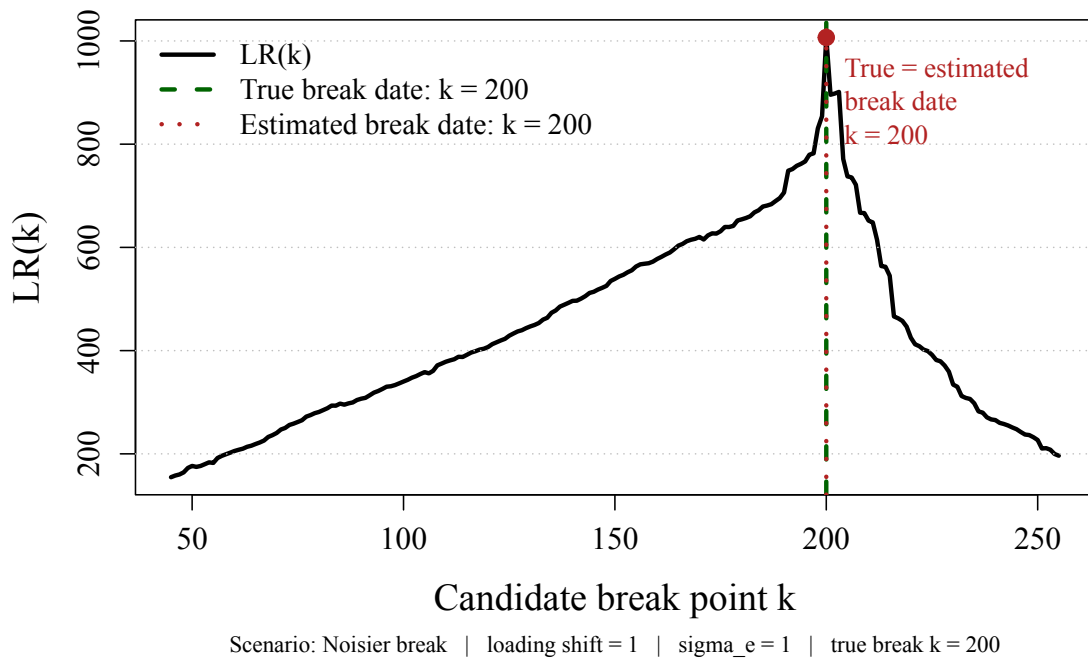


Figure 4.8: LR statistic in the simulated noisier break scenario

4.3.5 Preliminary comparison of the four illustrative cases

A comparison of the four illustrative cases allows us to draw several preliminary conclusions. Firstly, in the absence of a break, the procedure correctly recovers the true number of factors and does not produce a false signal of structural instability. Secondly, in all scenarios with a break, the test confidently rejects the null hypothesis and accurately localizes the break date. Finally, the presence of a break is systematically accompanied by an overestimation of the number of factors in the pooled sample: instead of the true value of $r = 2$, the criteria choose $\hat{r} = 3$. This effect, already visible at the level of single runs, indicates that the structural shift can be interpreted by factor selection procedures as an additional latent component.

At the same time, individual illustrative runs do not allow a systematic evaluation of the finite-sample properties of the method. For this reason, the Monte Carlo results are considered next, in which the same patterns are examined on the basis of 100 independent replications in each scenario.

Table 4.1: Illustrative single-run results across the four simulation scenarios. True $r = 2$ in all scenarios

Scenario	Final $\hat{r}$	True break	Estimated break	sup-LR	CV (5%)	Reject at 5%
No break	2	–	143	5.764	12.743	No
Baseline break	3	200	200	1381.770	79.092	Yes
Moderate break	3	200	200	1070.463	77.614	Yes
Noisier break	3	200	200	980.607	84.634	Yes

4.4 Monte Carlo evidence

Illustrative single runs allow us to evaluate the behavior of the procedure on specific realizations. However, a systematic Monte Carlo analysis is necessary for a meaningful assessment of its finite-sample properties. This subsection examines the results of 100 independent replications for each of the four simulation design scenarios. The main focus is on three aspects: the choice of the number of factors, the rejection frequency of the null hypothesis of no break, and the accuracy of estimating the date of the break in those scenarios where the break is actually present. A compact summary of the Monte Carlo simulation design is provided in Table 4.2 and in Appendix C, Table C.1.

Table 4.2: Simulation design

Scenario	T	N	r	σ_f	σ_e	loading shift	Break date
No break	300	200	2	1	0.5	0.0	—
Baseline break	300	200	2	1	0.5	1.0	200
Moderate break	300	200	2	1	0.5	0.5	200
Noisier break	300	200	2	1	1.0	1.0	200

4.4.1 Factor-number selection

The results of the Monte Carlo experiment show that the behavior of the procedure for determining the number of factors depends strongly on the presence of a structural break. In the no-break scenario, the final criterion used in the subsequent PCA/LR procedure consistently recovers the true number of factors. The average and median values of $\hat{r}$ coincide with the true value of $r = 2$, and the frequency of correct selection is equal to 1. Thus, under the null hypothesis, the procedure for selecting the number of factors demonstrates correct behavior.

A different picture is observed in break scenarios. In all three configurations, where the structural change is actually present, the selected number of factors is systematically higher than the true value. In these experiments the average and median values of $\hat{r}$ are 3, and the frequency of selection of $\hat{r} > 2$ reaches unity. This means that in the presence of a break, the factor dimension in the pooled sample is overestimated. This result can be interpreted as follows: a structural change in the loading matrix leads to the fact that the combined sample appears more complex in terms of the overall covariance structure, and the information criteria interpret this increased complexity as the presence of an additional factor component.

This pattern helps explain why factor selection and break testing should not be treated as completely separate steps. Even when the true number of latent factors is fixed, a break in the loading matrix can make the pooled sample appear to have an additional common component. In this sense, structural instability may affect the selected factor dimension before it is formally detected by the LR test.

4.4.2 Rejection frequencies and test performance

The rejection frequencies provide direct information about the size and power properties of the considered LR approach. In the no-break scenario, the rejection frequency is 0.04 at the

5% level and 0.05 at the 10% level. Thus, the procedure is well calibrated at the 5% level and somewhat conservative at the 10% level. In other words, the procedure does not exhibit a noticeable tendency toward false positives in the absence of structural instability.

In the three scenarios with a structural break, the rejection frequencies are equal to one. The null hypothesis is rejected in all replications not only at the 5% level, but also at the 1% level. This result is obtained for the baseline break, the moderate break, and the noisier break scenario. Hence, for the parameter values used in the Monte Carlo experiment, the LR procedure has very high power against the considered alternatives.

It is especially noteworthy that even in a scenario with a moderate break, the procedure retains full power. This means that with the selected panel dimensions and generation parameters, even a relatively less pronounced change in the loading matrix remains a strong enough signal for the LR approach. Similarly, increasing the variance of the idiosyncratic noise to be equal to 1 does not lead to any loss in the ability of the test to detect a break, although it makes the task statistically more difficult.

The average values of the sup-LR statistic also reflect the difference between the scenarios well. In the no break case, they remain small, whereas in all break scenarios they take very large values that significantly exceed the corresponding critical levels. The Monte Carlo results therefore confirm that the procedure successfully distinguishes between cases with and without structural change.

Table 4.3: Factor selection and rejection frequencies in the Monte Carlo experiment

Scenario	Mean $\hat{r}$	Median $\hat{r}$	$\Pr(\hat{r} = 2)$	Rej. 10%	Rej. 5%	Rej. 1%	Mean sup-LR
No break	2	2	1.00	0.05	0.04	0.00	7.353
Baseline break	3	3	0.00	1.00	1.00	1.00	1336.687
Moderate break	3	3	0.00	1.00	1.00	1.00	1073.042
Noisier break	3	3	0.00	1.00	1.00	1.00	1022.535

4.4.3 Accuracy of break-date estimation

For scenarios in which a break is present, not only the rejection of the null hypothesis is important, but also the accuracy of locating the date of the break. The results of the Monte Carlo experiment show that the date of the structural change is estimated with very high accuracy. In

all three break scenarios, the mean estimated break date almost coincides with the true value $T_1 = 200$, and the median estimate in each case is also equal to 200.

The standard deviation of the estimated date remains extremely small, and the average absolute error is only a small fraction of one period. Additionally, the proportion of replications in which the estimate falls within ± 5 periods of the true break is equal to one; the same holds for the wider ± 10 -period neighborhood. This means that in the scenarios under consideration, the break-date estimator performs almost flawlessly.

From a practical point of view, this result is particularly important. In many empirical problems, it is not enough to simply detect the presence of structural instability; it is also necessary to establish the moment at which the regime change occurs. The Monte Carlo study shows that in the chosen parameter configuration, the LR approach not only has high power, but also provides extremely accurate localization of the break point.

The accuracy of break-date estimation in the break scenarios is summarized in Table 4.4.

Table 4.4: Accuracy of break-date estimation in break scenarios

Scenario	True break	Mean estimate	Median est.	SD	MAE	$P(\hat{T}_1 - T_1 \leq 5)$
Baseline break	200	200.07	200	0.326	0.11	1.00
Moderate break	200	200.00	200	0.376	0.12	1.00
Noisier break	200	200.03	200	0.460	0.13	1.00

4.4.4 Overall discussion of the Monte Carlo results

The summary of the results obtained allows us to draw several general conclusions about the properties of the procedure in question. First, in the absence of a break, the test behaves satisfactorily under the null hypothesis: the frequency of false rejections remains low, with good calibration at the 5% level, and a somewhat conservative behavior at the 10% level. Secondly, in the presence of a break, the method shows exceptionally high power even in less favorable scenarios with a weaker structural shift or a higher level of idiosyncratic noise. Finally, the estimate of the break date turns out to be extremely accurate in all the cases considered.

Additional Monte Carlo results, including the single-run summary, factor-selection frequencies, rejection frequencies, and break-date accuracy measures, are reported in Appendix C, Tables C.2 – C.5.

Along with this, Monte Carlo analysis reveals another important feature: a structural break systematically affects the procedure for selecting the number of factors, inducing an upward shift in the selected factor dimension. Thus, the presence of a break in loadings can manifest itself not only through an increase in the LR statistic, but also through a distortion of the standard pooled estimate of the factor dimension. This observation will also be useful for subsequent analysis of real data, where the instability of the factor structure may also be accompanied by ambiguity in the choice of the number of factors.

Overall, the Monte Carlo results confirm that the procedure under consideration has good finite-sample properties and is a suitable tool for analyzing structural instability in large-dimensional factor models.

4.5 Discussion

The results of the simulation study allow us to draw several general conclusions about the properties of the procedure under consideration. First, in a scenario without a structural break, the test demonstrates correct behavior under the null hypothesis: the frequency of false rejections remains low, and the choice of the number of factors coincides with the true dimension of the model. This means that in the absence of a structural break, the procedure does not create an artificial signal of structural instability.

Secondly, in all considered scenarios with a structural break, the LR approach shows very high power. The null hypothesis is systematically rejected, and the estimated date of the break turns out to be almost identical to the true one. This indicates that in the selected simulation design configurations, the method under consideration is able not only to detect a structural change, but also to accurately localize the moment of transition between regimes.

The simulations also show how the break affects factor selection. In the break scenarios, the pooled sample leads to $\hat{r} = 3$, although the true number of factors is $r = 2$. This suggests that the change in the loading matrix is partly captured as an additional latent component. The same mechanism may be relevant for the empirical application: if the Russian stock return panel contains a structural break, the Bai–Ng criteria may give a less clear signal about the appropriate factor dimension.

In general, the simulation study confirms that the procedure in question has good finite-sample properties under conditions considered in this study. Therefore, it seems natural to proceed to its application to real data of the Russian stock market.

Chapter 5

Empirical application to the Russian stock market

After analyzing the artificially generated data, the natural next step is to apply the procedure to a real financial sample. The purpose of the empirical part is to examine whether a structural break can be detected in the factor structure of the Russian stock market and how robust this result is with reasonable changes in the specification.

The choice of the Russian stock market over the period 2021–2025 is motivated by the possibility of substantial structural changes in the economic and financial environment. Against the background of geopolitical tensions, the military conflict, and the successive rounds of Western sanctions, it is natural to expect changes in the dependence structure of Russian asset returns.

5.1 Description of the data and preprocessing

The panel of weekly returns on Russian stocks, constructed from securities traded on the Moscow Exchange, is used as a broad proxy for the Russian stock market. At the same time, the empirical sample does not represent a literal reconstruction of the historical composition of the Moscow Exchange Index (IMOEX) at each point in time. Instead, it should be understood as a working panel of liquid Russian stocks, constructed from Moscow Exchange data and then filtered according to data availability and sufficient time-series coverage.

MOEX Russia Index (IMOEX), 2021—2025

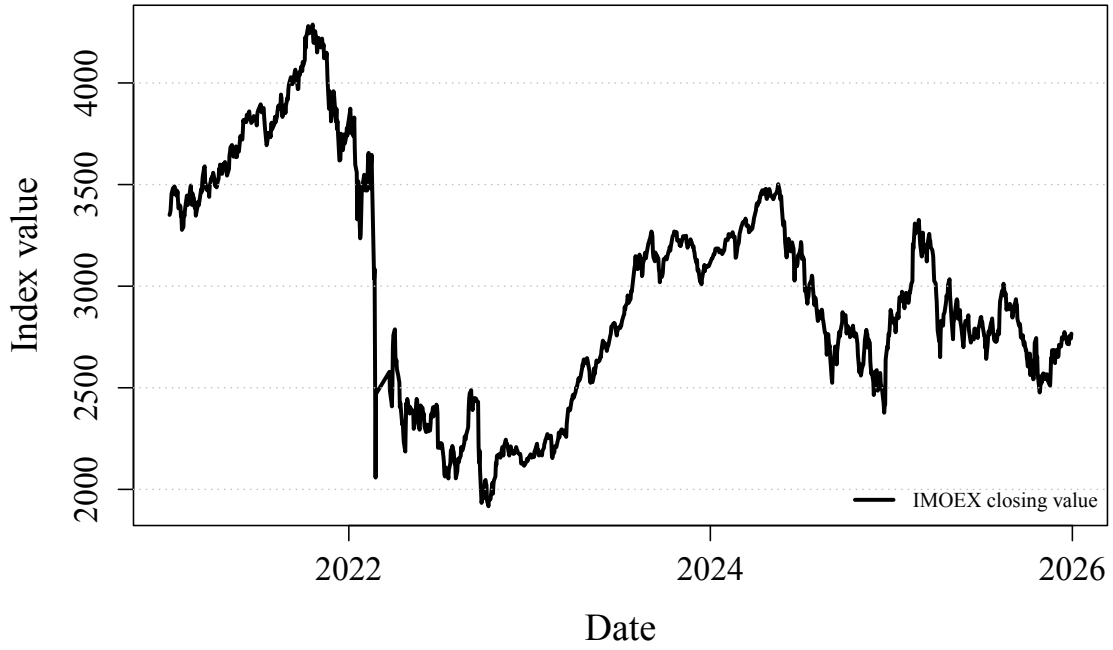


Figure 5.1: MOEX Russia Index

The IMOEX is used primarily as a natural reference point for describing the most important segment of the Russian stock market. It is a capitalization-weighted index that includes the most liquid shares of the largest issuers traded on the Moscow Exchange. Historically, companies in the oil and gas, financial, and metallurgical sectors have played a significant role in the index, so it serves as a convenient starting point for characterizing the structure of the Russian stock market as a whole. The initial data were obtained from the official website of the Moscow Exchange [22].

The general evolution of the MOEX Russia Index over the sample period is shown in Figure 5.1. The complete list of tickers included in the empirical sample is reported in Appendix A, Table A.1.

The empirical analysis is based on a panel of weekly stock returns, which are stored in the file `returns_weekly.csv`. The final sample contains $T = 257$ observations for $N = 71$ assets and covers the period from January 2021 to December 2025. Thus, this panel combines a moderate time length with a fairly significant cross-sectional dimension. This data structure

is well aligned with the general formulation of large-dimensional factor models, in which the overall dynamics of a large number of time series is approximated by a relatively small number of latent factors.

Working with real financial data requires a separate stage of preprocessing. The original panel contains missing values, which is typical of financial return panels due to differences in listing dates, incomplete trading histories of individual securities, and other observational features. Before applying the factor model procedures, the empirical panel of weekly returns was made balanced by filling missing observations.

For this purpose, we used an iterative PCA-based imputation procedure following the EM-type approach discussed by Bai and Ng (2002) [5], who refer to Rubin and Thayer (1982) [27] and Stock and Watson (1998) [28]. Although the main theoretical analysis in Bai and Ng (2002) focuses on balanced panels, they note that missing observations can be handled by an iterative EM algorithm. The idea is to keep observed entries fixed and to replace missing entries by the values predicted by the factor model using the parameters estimated in the previous iteration.

More precisely, if X_{it} is observed, it is kept unchanged. If X_{it} is missing, it is replaced by

$$X_{it}^*(j-1) = \widehat{\lambda}_i^{(j-1)'} \widehat{F}_t^{(j-1)},$$

where $\widehat{F}_t^{(j-1)}$ and $\widehat{\lambda}_i^{(j-1)}$ are the factor and loading estimates from the previous iteration. The completed matrix is then used to re-estimate the factors and loadings by principal components, and the procedure is repeated until convergence.

In the empirical implementation, the initial missing values were replaced by column means only to initialize the algorithm. After that, missing entries were updated using the low-rank factor reconstruction, while the originally observed values were kept unchanged. Before the imputation procedure, the data contained 91 missing values. The implemented procedure converged in 11 iterations, after which no missing values remained in the final data matrix.

After completion of preprocessing, the dimension of the final observation matrix was 257×71 . This matrix is used at all subsequent stages of empirical analysis: selecting the number of factors using the Bai–Ng criteria, estimating the factor structure by principal component analysis, and constructing LR statistics for structural break testing.

Additional details on the raw data set, the balanced panel, and the preprocessing steps are summarized in Appendix A, Table A.2.

It should be emphasized that in an empirical application, the main object of interest is not the interpretation of individual factors as specific economic variables but the stability of the factor structure itself over time. In other words, the task is not to assign a specific macroeconomic meaning to each estimated factor, but to establish whether the relationship between stock returns and the common latent component remains stable throughout the sample under consideration.

From a practical point of view, this means that this chapter is mainly focused on three issues. First, it is necessary to determine the number of factors used in the subsequent analysis. Second, it is necessary to compute the LR statistics and the corresponding critical values in order to test the hypothesis of the absence of a structural break. Third, it is important to investigate the stability of the obtained result to changes in the trimming parameter and to small changes in the chosen factor dimension, as well as to compare the properties of the sample before and after the estimated break date.

Thus, the empirical chapter is organized as a sequential application of the previously developed procedure to the real panel of Russian stock returns. First, we consider the basic result for the initial specification, then we perform the robustness checks, and after that, we compare the factor structure of the pre-break and post-break subsamples.

5.2 Baseline empirical results

After describing and preprocessing the data, we proceed to the main empirical result. In the first stage, the number of factors is determined for the complete sample. This number is then used in the PCA estimation and in the construction of the LR statistic for structural break testing.

The results of applying the Bai–Ng criteria to the full panel of weekly returns show noticeable dispersion in the estimated factor dimension. In particular, the criteria PC_{p1} , PC_{p2} , and PC_{p3} select 16, 15, and 19 factors respectively, while the criteria IC_{p1} , IC_{p2} , and IC_{p3} suggest the values 10, 8 and 14. Thus, the data do not provide a completely unambiguous signal regarding the exact number of factors. Since the Bai–Ng criteria do not produce a fully identical estimate of the factor dimension in the empirical sample, a single baseline criterion is required in order

to implement the PCA/LR procedure in a consistent way. Since the PC_p criteria tend to select relatively high-dimensional specifications in this empirical sample, while IC_{p2} gives the most parsimonious specification among the information criteria, the baseline analysis uses $\hat{r} = 8$. The sensitivity of the result to nearby values of r is assessed in the robustness section.

The scree plot for the full sample is presented in Figure 5.2. The complete Bai–Ng factor-selection results for the full sample are reported in Appendix B, Table B.1.

Scree plot for the full sample

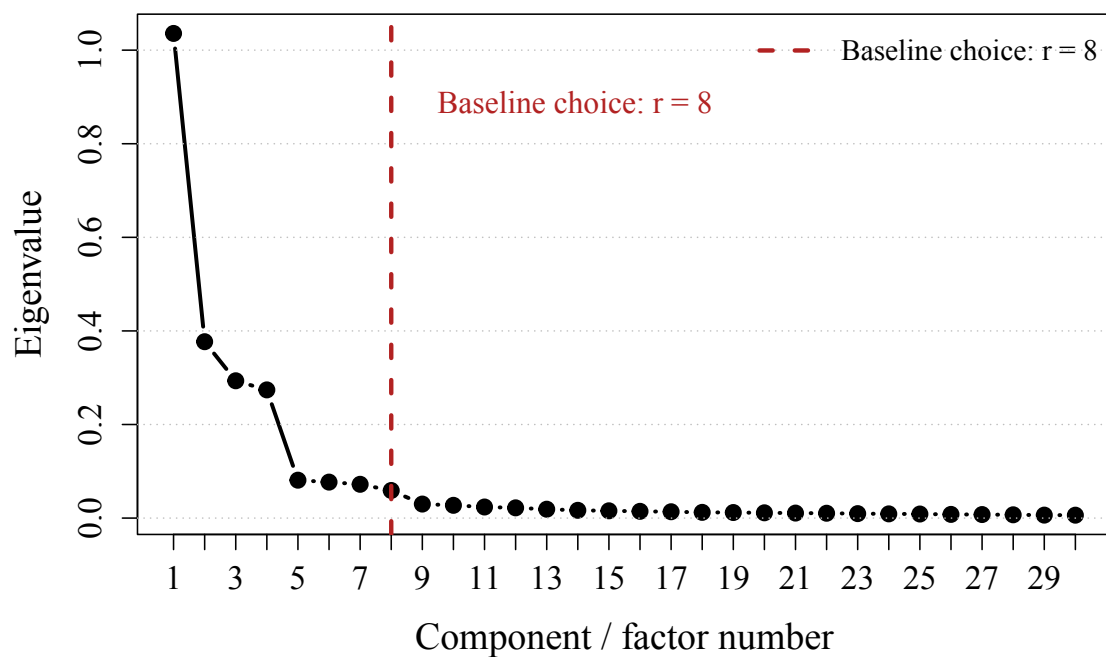


Figure 5.2: Scree plot for the full sample

Using $\hat{r} = 8$, the estimated factor matrix is 257×8 and the loading matrix is 71×8 . The corresponding share of explained variation is about 0.848, which suggests that the selected factors capture a large part of the common dynamics of returns. The leading eigenvalues give a similar indication: the first eigenvalue is much larger than the subsequent ones, although the remaining eigenvalues show that the variation is not concentrated in a single component only.

After selecting the number of factors, the LR testing procedure for structural break detection is applied. In the baseline specification, the observed value of the sup-LR statistic is approximately 2474.56. The estimated break date corresponds to the 159th observation of the sample, which,

LR statistic over candidate break dates: full sample

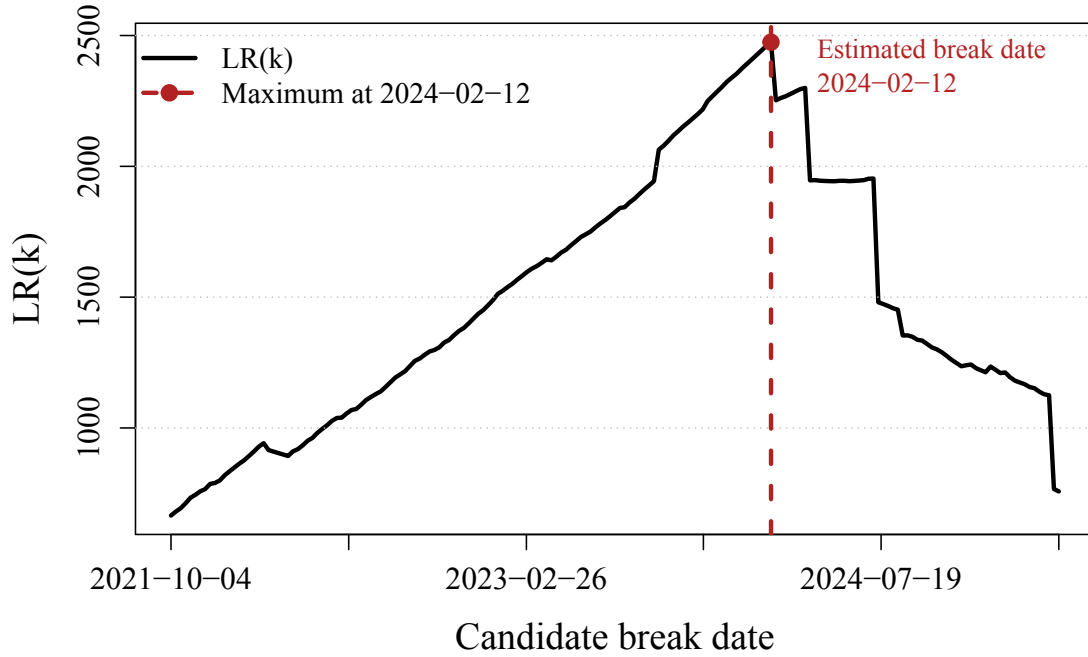


Figure 5.3: LR statistic over candidate break dates: full sample

in calendar time, corresponds to the week of February 12, 2024. Taking into account the discreteness of the weekly data, the estimated break is located in the interval between February 12, 2024 and February 19, 2024.

To assess the statistical significance of the sup-LR result, the statistic is compared with the corresponding critical values. In the baseline run, the critical value at the 10% level is approximately 1988.99, at the 5% level it is 2253.47, and at the 1% level it equals 2724.29. The observed value of sup-LR exceeds the critical values at the 10% and 5% levels, but remains below the critical value at the 1% level. Therefore, the null hypothesis of no structural break is rejected at the 10% and 5% significance levels, but is not rejected at the 1% level. This suggests that the evidence is statistically meaningful, but should not be overstated. The main baseline result is summarized in Table 5.1. More detailed baseline LR test outputs, including the decision table based on simulated critical values, are provided in Appendix B, Tables B.2 - B.3.

This result indicates the presence of statistically significant evidence in favor of a structural change in the factor structure of the Russian stock market in the sample that we consider. At

Table 5.1: Baseline empirical results for the Russian stock market data

Chosen $\hat{r}$	sup-LR	Break date	Break frac.	CV 10%	CV 5%	CV 1%	Rej. 10%	Rej. 5%	Rej. 1%
8	2474.556	2024-02-12	0.6187	1988.990	2253.466	2724.285	Yes	Yes	No

the same time, it is important to emphasize that the result cannot be interpreted as completely invariant to the specification: since the Bai–Ng criteria give markedly different estimates of the factor dimension, the baseline choice $\hat{r} = 8$ itself should be treated as a working specification rather than as the only possible value. It is for this reason that a series of robustness checks is carried out, which makes it possible to assess how stable the found break date and the test outcome are to changes in the trimming parameter and to small changes in the selected number of factors.

In general, the baseline empirical result already at this stage provides a meaningful conclusion: for the full panel of weekly returns of Russian stocks, evidence of a break in the factor structure is found, localized in mid-February 2024. However, the final interpretation of this result requires an assessment of its stability, which we perform in the next subsection.

5.3 Robustness checks

After obtaining the baseline empirical result, it is necessary to check how robust the conclusion about the presence of a structural break is to reasonable changes in the specification. This assessment is particularly important in the considered problem, since the choice of the number of factors according to the Bai–Ng criteria is not completely unambiguous, and the LR procedure itself depends on the trimming parameter. For this reason, two types of robustness checks are considered: the robustness of the result to changes in the trimming parameter and to changes in the number of factors used in the PCA/LR procedure.

5.3.1 Robustness to trimming

The first robustness check concerns the trimming parameter. This parameter defines the range of admissible candidate break dates by excluding observations too close to the beginning or the end of the sample. The baseline value is 0.15. To check whether the result depends on this

choice, I repeat the LR procedure with two alternative values, 0.10 and 0.20.

The estimated break is unchanged across these three specifications. For trim values 0.10, 0.15, and 0.20, the maximum of the LR statistic is reached at the same date, corresponding to the week of February 12, 2024. The test conclusion is also the same in all cases: the null hypothesis of no structural break is rejected at the 5% significance level.

This result provides important evidence in favor of the stability of the baseline conclusion. It demonstrates that the found break is not an artifact of a particular choice of the trimming parameter and persists under reasonable changes in the boundaries of the search interval. In other words, the localization of the break turns out to be robust to natural variations in the specification of the testing procedure.

The results of this robustness check are reported in Table 5.2. The full set of robustness results for alternative trimming parameters is reported in Appendix B, Table B.4.

Table 5.2: Robustness of the baseline result to trimming parameter

Trim	Chosen $\hat{r}$	sup-LR	Break date	Break frac.	CV 5%	Rej. 5%
0.10	8	2474.556	2024-02-12	0.6187	2388.253	Yes
0.15	8	2474.556	2024-02-12	0.6187	2203.090	Yes
0.20	8	2474.556	2024-02-12	0.6187	2245.034	Yes

5.3.2 Robustness to the number of factors

The second type of robustness analysis concerns the choice of factor dimension. Since the Bai–Ng criteria provide markedly different estimates of the number of factors in the full sample, it is necessary to check whether the main conclusion about the presence of a break persists under small changes in $\hat{r}$ used in the estimation. For this purpose, the baseline choice $\hat{r} = 8$ is compared with close values $\hat{r} = 7$ and $\hat{r} = 9$.

The results demonstrate the high robustness of the baseline finding to changes in the factor dimension. In all three specifications — with $\hat{r} = 7$, $\hat{r} = 8$, and $\hat{r} = 9$ — the observed sup-LR statistic remains above the corresponding critical value at the 5% level, so the null hypothesis of the absence of a break is rejected in all cases. Moreover, the estimated break date in all three specifications is completely the same and corresponds to February 12, 2024.

Therefore, the main empirical result does not depend on the exact choice of the factor

dimension in a sufficiently small neighborhood of the baseline value. This point is especially important because it shows that the conclusion about the presence of a structural change is not driven by the arbitrary choice of a single value of $\hat{r}$, but persists under local changes in the dimension of the factor space.

The corresponding results are summarized in Table 5.3. The detailed robustness results for alternative values of r are reported in Appendix B, Table B.5.

Used r	sup-LR	Break date	Break frac.	CV 5%	Rej. 5%
7	2302.996	2024-02-12	0.6187	2168.947	Yes
8	2474.556	2024-02-12	0.6187	2270.731	Yes
9	2522.158	2024-02-12	0.6187	2205.130	Yes

5.3.3 Interpretation of the robustness results

Taken together, the two robustness checks allow us to draw a general conclusion about the high stability of the baseline empirical result. Changing the trimming parameter does not affect the localization of the break and does not change the final test outcome. Similarly, small changes in the chosen number of factors likewise do not eliminate the statistical significance of the result or alter the estimated break date.

Thus, the break found in the factor structure of the Russian stock market cannot be considered as a fragile result that depends on a single specific specification. On the contrary, within natural variations of the key settings of the procedure — the trimming parameter and factor dimension — the conclusion about the presence of a structural change in mid-February 2024 remains unchanged. This creates a fairly reliable basis for the next step of the analysis, namely, the comparison of the factor structure in the pre-break and post-break subsamples.

5.4 Pre-break and post-break comparison

After checking the stability of the baseline empirical result, the natural next step is to compare the factor structure in more detail before and after the estimated break date. If the found structural change really reflects a regime change in the data, then the differences should manifest

Comparison of scree plots before and after the estimated break

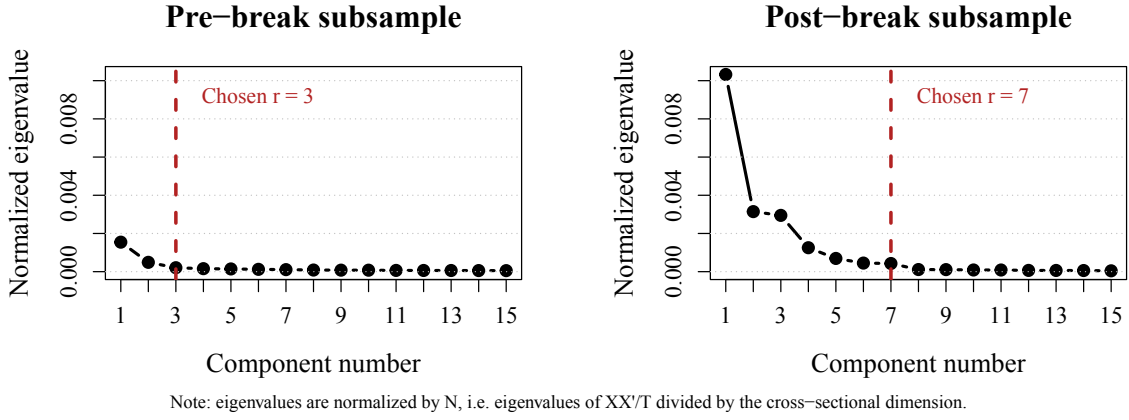


Figure 5.4: Comparison of scree plots before and after the estimated break

themselves not only in the value of the LR statistic, but also in the characteristics of the factor structure in the pre- and post-break subsamples.

The baseline estimated break date is used as the split point, which lies in the interval between February 12, 2024 and February 19, 2024. Accordingly, the sample is divided into two parts: a pre-break subsample covering the period from January 11, 2021 to February 12, 2024, and a post-break subsample covering the period from February 19, 2024 to December 29, 2025. The size of the pre-break subsample is 159 observations, whereas the post-break part contains 98 observations.

The comparison starts with factor-number selection in the two subsamples. Using IC_{p2} , the pre-break part gives $\hat{r} = 3$, while the post-break part gives $\hat{r} = 7$. This difference suggests that the dependence structure of Russian stock returns becomes more complex after the estimated break in mid-February 2024.

The complete Bai–Ng factor-selection results for the pre-break and post-break subsamples are reported in Appendix B, Table B.6.

The share of explained variation changes in the same direction. In the pre-break subsample, the estimated factor model explains about 0.538 of the total variation. In the post-break subsample, the corresponding value rises to about 0.940. Thus, after the break, the common component becomes much more pronounced.

The eigenvalues provide another way to see the same shift. Before the break, the first three leading eigenvalues are approximately 0.246, 0.078, and 0.033. After the break, they increase to about 1.015, 0.310, and 0.290. The post-break spectrum is therefore more concentrated in the first components, which is consistent with a stronger factor structure in the second part of the sample.

For additional verification, it is useful to compare two subsamples not only with separately selected r values, but also with the same fixed number of factors. In the implemented program, such a comparison is performed for $r = 8$, that is, with the baseline factor dimension selected for the full sample. Even in this case, the difference between the two periods remains: the proportion of the explained variation before the break is about 0.686, whereas after the break it increases to about 0.946. This means that the detected shift is not merely a consequence of the separately selected number of factors, but reflects a deeper change in the data structure itself.

The results of the comparison are summarized in Table 5.4. Additional numerical details of the pre-break and post-break comparison are provided in Appendix B, Table B.7.

Table 5.4: Comparison of the factor structure before and after the estimated break

Subsample	Start date	End date	T_{sub}	Chosen $\hat{r}$	Explained var.	λ_1	λ_2	λ_3
Pre-break	2021-01-11	2024-02-12	159	3	0.538	0.246	0.078	0.033
Post-break	2024-02-19	2025-12-29	98	7	0.940	1.015	0.310	0.290

The results show that the break detected by the LR test is associated not only with the formal rejection of the null hypothesis, but also with a substantial restructuring of the market's factor structure. After the estimated break date, the number of factors selected by the baseline criterion increases, the share of explained variation rises sharply, and the spectrum of the covariance matrix becomes more concentrated in the leading components. This suggests that after mid-February 2024, the joint dynamics of Russian stocks became more strongly driven by a common latent structure.

It is especially important that this conclusion remains valid when comparing two subsamples with the same fixed number of factors. Consequently, the detected difference cannot be reduced merely to the mechanical effect of selecting different values of r before and after the break, but instead points to a deeper change in the dependence structure of the data itself. In other words,

the result should not be interpreted as a local statistical irregularity, but rather as evidence of a transition to a different regime of joint return dynamics.

5.5 Empirical discussion

The empirical analysis carried out in this chapter allows us to draw several main conclusions regarding the stability of the factor structure of the Russian stock market in the sample under consideration.

5.5.1 Statistical meaning of the detected break

The structural break detected in this study should be interpreted primarily as evidence of instability in the internal relationship between the returns of individual stocks and the common latent component. In other words, the LR test result indicates not merely a possible change in the level of returns at some point in time, but a deeper shift in the factor structure of the sample itself. This suggests that the relationships between individual securities and common latent market factors have ceased to be stable over the entire sample period (see Figure 5.3).

However, this result requires careful interpretation. The LR test used in this study does not allow us to fully disentangle the change in factor loadings and the change in the volatility of the latent factors themselves. This is due to the nature of the factor model: factors and loadings are not observed separately but only through their product. Therefore, a change in the second moments of the estimated pseudo-factors can be consistent with both a change in factor loadings and with a change in the strength or volatility of the common market component. Bai, Duan, and Han exploit this observational equivalence in their approach to reduce the high-dimensional problem of structural loading variation to the low-dimensional problem of second-moment instability of the estimated factors. In this regard, the break detected in this research is interpreted primarily as evidence of a change in the factor structure and the nature of the joint movement of Russian stocks, rather than as a definitive empirical separation between changes in loadings and changes in factor volatility.

This understanding of the result is especially important in the context of factor models. If the

loading structure remained constant, then the full sample could be considered homogeneous, and a single factor representation would be a natural approximation for the entire period. However, the results indicate that this assumption is too strong for the data considered here. In this sense, the detected break should be understood as evidence that the full sample of 2021–2025 probably combines at least two different regimes of joint return dynamics (see Table 5.4).

Additional support for this interpretation is provided by the comparison of the pre-break and post-break subsamples (see Table 5.4 and Figure 5.4). After the estimated break date, the number of factors selected by the baseline criterion increases, the share of explained variation increases, and the leading eigenvalues become significantly more pronounced. Consequently, after mid-February 2024, the common latent structure of the market became more pronounced than in the pre-break period. This means that the detected break is reflected not only in the test statistic, but also in the properties of the data themselves.

From a practical point of view, this result is also important because it calls into question the use of a single factor model estimated over the entire period as a whole, without taking into account possible structural heterogeneity. If there really are two regimes in the sample, then pooled estimation can average across different dependence structures and thus partially hide the features of each regime. Therefore, the detected break should be regarded as an argument in favor of a more cautious interpretation of the full-sample factor structure and as an indication that the pre-break and post-break parts of the sample may require separate analysis.

5.5.2 Interpreting the estimated break date

The break date identified in this study should not be interpreted too literally as a single calendar moment at which a specific event took place that completely changed the market. Since the empirical analysis is based on weekly returns, it is more appropriate to understand the result as an indication of an interval of regime change between February 12 and February 19, 2024. In this sense, the LR procedure does not identify a “single turning point day”, but rather a period in which the structure of the joint dynamics of Russian stock returns ceases to be homogeneous (see Figure 5.3).

This interpretation is especially important for understanding the economic significance of

the result. In structural break analysis, the estimated date does not necessarily coincide with the exact moment when the new regime emerged. Rather, it usually localizes the part of the sample around which the difference between the two parts of the sample becomes statistically most pronounced. Therefore, the break identified in this study should be understood not as a pointwise reference to a single news item or a single trading day, but as an empirical indication of a period of noticeable restructuring of market dependence structures.

An additional argument in favor of this interpretation is provided by the very shape of the LR statistic across possible candidate dates. The maximum is attained around mid-February 2024, and the estimated break date itself remains unchanged under reasonable changes in the trimming parameter and the number of factors (see Figure 5.3 and Tables 5.2 – 5.3). This means that the result does not appear to be a random outlier related to one specific technical setting of the procedure, but rather reflects a consistently identified boundary between two regimes in the data.

Thus, the estimated date should be used primarily as a guideline for economic interpretation and for the subsequent comparison of the pre-break and post-break periods. It provides a natural point for dividing the sample into two parts, within which it becomes possible to analyze separately the factor dimension, the share of explained variation, and the strength of the common component. It is in this sense that the estimated break date plays a central role in the empirical part of the thesis: not as a literal indication of a single point in time, but as a statistically justified identification of the period in which the regime change occurred.

5.5.3 Economic plausibility of the timing

The structural break by itself does not prove that it was caused by a single specific event. Nevertheless, its location in mid-February 2024 appears economically plausible, given the broader institutional and news background of that period. Since the empirical analysis is based on weekly returns, it is more appropriate to relate the estimated break date not to one isolated news item, but to a cluster of events that could have changed the nature of the joint movement of Russian stocks. In this sense, mid-February 2024 can be regarded as a period of increasing uncertainty and repricing of market relationships, which is consistent with the estimated break

interval shown in Figure 5.3. Figure 5.5 places the estimated break date on the time series of the MOEX Russia Index, which helps to relate the statistical result to the market background.

First, an important element of this background was the decision of the Bank of Russia on 16 February 2024 to keep the key rate at 16% [10]. By itself, maintaining the rate does not necessarily mechanically generate a structural break. However, in the presence of continued inflationary pressures and tight monetary conditions, it could affect a wide range of Russian assets through the cost of funding, expectations about domestic demand, and the repricing of future cash flows. For the stock market, this meant the continuation of a high-interest-rate environment and, consequently, a possible change in the sensitivity of different securities to common macro-financial factors.



Figure 5.5: MOEX Russia Index with estimated break date

Second, the estimated break interval coincides with an intensification of the sanctions-related background. On 14 February 2024, Reuters reported that European Union countries were close to agreeing on the 13th sanctions package against Russia [25], and on 21 February the same source reported its formal approval [24]. Even if the measures themselves did not take effect at a

single moment, such a news environment could increase uncertainty, change market participants' expectations, and affect the perception of external economic risks. This is important for the Russian market not only in a political sense, but also in terms of the dependence structure of returns, since sanctions-related expectations can affect exporters, the financial sector, and domestically oriented companies in different ways.

Third, the same period was accompanied by additional pressure on external payment channels and trade settlement mechanisms. On 19 February 2024, Reuters reported that the threat of U.S. secondary sanctions had already disrupted some trade flows between Turkey and Russia and slowed some cross-border payments [26]. This is especially important in the context of the market factor structure, since such restrictions can affect not only individual companies, but also broader channels of shock transmission through exports, imports, settlements, and financial intermediation. As a result, they could contribute to a broader change in the nature of the joint movement of Russian stocks.

Thus, mid-February 2024 can be interpreted as a period in which several sources of market instability intensified simultaneously: a restrictive monetary-policy signal, sanctions pressure, and difficulties in external payment channels. The present work, of course, does not claim to establish a strict causal relationship between these events and the detected break. However, their temporal coincidence with the results of the LR procedure makes the interpretation of the detected break as a market regime shift economically plausible.

5.5.4 Practical implications of the break

The structural break detected in this study is important not only as a formal statistical result, but also from a practical point of view. If the factor structure of the market changes over time, then this means that the dependence patterns estimated on the full sample do not necessarily remain equally informative for all parts of the sample. In other words, a model estimated over the entire 2021–2025 period as if it were a homogeneous sample can partially average across different regimes and thereby smooth out the specific features of each regime (see Table 5.4).

This is especially important for empirical modeling. In many applications, the factor model is used to describe the overall structure of the market, to measure systematic risk, to reduce the

dimensionality of the data, and to analyze dependence relationships between assets. However, if factor loadings change after a structural break, then a single model for the full period may provide a less accurate picture of the nature of the common component than a separate analysis of the pre-break and post-break subsamples. In this case, the very assumption of the stability of the factor structure becomes questionable.

The practical significance of this result is also evident in the context of risk management. If the pattern of joint stock movements changes after the break, then the covariance structure, correlations, and sensitivity of individual securities to common factors may no longer match what was observed in the earlier part of the sample. Consequently, risk estimates based on a long historical sample without taking into account a possible regime change may be less reliable. This applies not only to formal risk models, but also to more general conclusions about diversification, portfolio resilience, and the behavior of different groups of assets during periods of instability.

In addition, the detected break is important for interpreting the number of factors. As shown in the simulation study, the presence of a structural shift can lead to an overestimation of the factor dimension in a pooled sample, since the regime change itself is partially perceived by the selection criteria as an additional latent component (see Table 4.3). Therefore, the distinction between pre-break and post-break periods is important not only in itself, but also as an argument in favor of a cautious interpretation of factor selection results in the full sample. In other words, if the sample is structurally heterogeneous, then the question of the number of factors can no longer be considered completely separately from the question of the stability of factor loadings.

Finally, from a more general economic point of view, the result shows that the Russian stock market over the period under review cannot be described as a system with an unchanged internal structure of dependence relationships. This does not mean that after the estimated break date the market became completely different in all respects, but it does mean that the assumption of its homogeneity over the entire observation period appears to be too strong. That is why the detected break should be regarded as a practically significant result: it indicates the need for a more careful interpretation of full-sample estimates and supports the idea that a separate analysis of different regimes may be more meaningful than a single pooled specification.

Chapter 6

Conclusions

This work was devoted to the analysis of structural changes in large-dimensional factor models using the likelihood ratio approach, with an application to the Russian stock market. The research focused on three interrelated tasks: describing and estimating the factor model, determining the number of factors, and testing the stability of the factor structure over time.

In the theoretical part of the work, a large-dimensional static factor model was considered, in which the observed variables are represented as the sum of a common component explained by a small number of latent factors and an idiosyncratic component. It has been shown that the principal components method is a natural tool for estimating the factor structure in conditions where both the cross-sectional and time dimensions of the panel are large. Special attention was paid to the task of choosing the number of factors. To address this task, the Bai–Ng information and panel criteria were considered, which provide consistent estimates of the factor dimension in large-dimensional panels.

The next important step was to consider the problem of structural instability in the matrix of factor loadings. It has been shown that in the presence of a structural break, the standard assumption of the constancy of factor loadings over the entire observation period may be too strong, especially when analyzing financial data sensitive to crises, sanctions, changes in the institutional environment, and other systemic shocks. In this regard, the likelihood ratio approach to testing a structural break in factor models was considered, based on an equivalent representation of the model in terms of pseudo-factors. This representation makes it possible to reduce the original high-dimensional problem to the analysis of the second moments of the

estimated factors and thereby make the testing procedure practically feasible.

The theoretical part also discussed the limiting distribution of the sup-LR statistic under the null hypothesis and the role of the long-run covariance matrix in the construction of critical values. This made it possible to present the testing procedure as a sequence of concrete steps: specification of the factor model, estimation of the factors, selection of the factor dimension, and evaluation of the LR statistic for possible break dates.

The simulation part examined how the procedure behaves when the data-generating process is known. We considered a no-break design and several designs with a break, varying the size of the shift in the loadings and the variance of the idiosyncratic component. In the no-break case, the procedure did not produce evidence of structural instability. In the break designs, the LR statistic reached its maximum close to the true break date and led to the rejection of the stability hypothesis. The simulations also showed a feature that is relevant for the empirical application: when a break is present, factor-selection criteria applied to the pooled sample may choose more factors than in each stable regime. This happens because part of the regime change is absorbed as an additional latent component.

In the empirical part of the work, the procedure was applied to a panel of weekly returns on Russian stocks for the period 2021–2025. The baseline result showed statistically significant evidence in favor of a structural break in the factor structure, localized in mid-February 2024. At the same time, it is more appropriate to interpret the estimated date not as a single calendar moment but as an interval of regime change between 12 and 19 February 2024. Robustness checks have shown that this conclusion persists both under changes in the trimming parameter and under small changes in the selected number of factors. A comparison of the pre-break and post-break subsamples additionally confirmed that the detected break is accompanied by a substantial restructuring of the factor structure itself: after the estimated break date, the number of factors selected by the baseline criterion increases, the share of explained variation increases, and the spectrum becomes more concentrated in the leading eigenvalues. In other words, after the break, the overall latent structure of the market became more pronounced than in the pre-break period.

At the same time, the empirical interpretation of this result should be formulated carefully.

The LR procedure reveals instability in the second-moment structure of the estimated pseudo-factors. This is consistent with a break in the factor loadings structure within the framework of the Bai, Duan, and Han approach. However, the test itself does not distinguish between changes in factor loadings and changes in factor volatility. Therefore, the main empirical conclusion of the thesis is formulated as evidence of a change in the factor structure and in the nature of the joint movement of Russian stocks. A natural extension of the present analysis would be to apply recent projection-based decomposition methods, such as Koo, Wong, and Zhong [19], in order to distinguish more explicitly between breaks in factor volatility and breaks in factor loadings.

It should also be noted that the LR procedure used in the thesis is based on a single-break specification. Therefore, the estimated break date should be interpreted as the most pronounced boundary between regimes within a one-break framework, rather than as evidence that no other changes in the factor structure occurred during the sample period. Identifying an unknown number of breaks would require a different methodology and represents a natural direction for further research.

A meaningful interpretation of this result is that the detected break reflects not merely a local deviation of one test statistic, but a change in the internal structure of the joint movement of Russian stocks. This means that the full sample probably combines at least two different regimes of joint return dynamics, and therefore the assumption of complete homogeneity of the sample is too strong. In this sense, the result obtained is important not only for formal testing but also for the interpretation of the common component of the market, the choice of the number of factors, and the construction of empirical factor representations on a pooled sample.

From an economic point of view, the detected break also seems plausible. Its timing coincides with a period of noticeably increased market uncertainty in mid-February 2024, when a restrictive monetary-policy signal, sanctions-related pressure, and difficulties in external payment channels were all present at the same time. The present work, of course, does not claim to establish a strict causal relationship between specific events and the detected break. However, the temporal coincidence of these circumstances with the results of the LR procedure makes the interpretation of the detected break as a market regime shift economically plausible.

The practical implication is that estimates based on the full 2021–2025 sample should be

interpreted with caution. If factor loadings are not stable over time, a pooled factor model may mix different regimes and produce dependence patterns that do not describe either regime well. This matters for the interpretation of correlations, diversification, and portfolio risk in the Russian market. For this reason, the detected break is relevant not only as a formal econometric result, but also as a warning against treating the whole sample as homogeneous.

Thus, the work shows that the likelihood ratio approach is a useful tool for analyzing structural instability in high-dimensional factor models. It makes it possible not only to test the hypothesis of a break, but also to localize the period of structural change, which is especially important in applied financial research. The results obtained can serve as a basis for the further study of unstable factor structures in longer samples, for the analysis of alternative markets, as well as for possible extensions of the model in the case of multiple structural breaks, a time-varying factor structure, or more complex dynamics of latent factors.

Appendix A

Data Set and Preprocessing Details

This appendix reports additional details on the empirical data set used in Chapter 5.

A.1 Empirical sample

The empirical analysis is based on weekly returns of $N = 71$ Russian stocks over $T = 257$ weeks, covering the period from 11 January 2021 to 29 December 2025. The final data matrix used in the empirical analysis has dimension 257×71 .

Table A.1: Tickers included in the empirical sample

Ticker	Ticker	Ticker	Ticker
VTBR	SBER	GAZP	GMKN
PLZL	MAGN	LKOH	NVTK
MOEX	MGNT	AFKS	IRAO
MTSS	AFLT	ROSN	PIKK
SBERP	RUAL	SNGSP	CHMF
NLMK	TRNFP	TATN	ALRS
SNGS	RTKM	PHOR	ENPG
HYDR	TATNP	CBOM	BSPB

Continued on the next page

Table A.1: Tickers included in the empirical sample – continued

Ticker	Ticker	Ticker	Ticker
SMLT	MSNG	MTLR	SIBN
TGKA	UPRO	NMTP	MRKV
SELG	BELU	LSNGP	FLOT
RNFT	TGKN	TRMK	MRKP
SFIN	LNZLP	BANEP	MRKC
MTLRP	KMAZ	PMSBP	IRKT
LNZL	ROLO	RTKMP	PMSB
FESH	MRKU	MSRS	APTK
AQUA	FEES	MDMG	OZON
ETLN	BSPBP	SGZH	

A.2 Preprocessing

The original weekly return panel contained 91 missing observations. Since the empirical procedure requires a balanced panel, missing values were filled using an iterative PCA-based imputation procedure. The algorithm converged after 11 iterations. After imputation, the final data set contained no missing values.

Table A.2: Summary of the empirical data set

Quantity	Value
Sample period	11 January 2021 – 29 December 2025
Number of weekly observations, T	257
Number of stock return series, N	71

Continued on the next page

Table A.2: Summary of the empirical data set – continued

Quantity	Value
Dimension of the raw data matrix	257×71
Missing values before imputation	91
Missing values after imputation	0
Imputation converged	Yes
Number of imputation iterations	11
Dimension of the final data matrix	257×71

Appendix B

Additional Factor Selection and Empirical Results

This appendix reports additional factor-selection results and empirical robustness checks for the real-data application.

B.1 Baseline factor selection

Table B.1: Bai–Ng factor-selection results for the full sample

Criterion	Selected number of factors
PCp1	16
PCp2	15
PCp3	19
ICp1	10
ICp2	8
ICp3	14

In the baseline empirical specification, the final number of factors used in the PCA and LR procedure is selected according to the ICp2 criterion, giving $\hat{r} = 8$.

B.2 Baseline LR test result

Table B.2: Baseline LR test result for the real-data application

Quantity	Value
Selected number of factors, $\hat{r}$	8
Explained variation ratio	0.848046
Observed sup LR statistic	2474.556
Estimated break index	159
Estimated break fraction	0.618677
Estimated break date	12 February 2024
Estimated break interval	12 February 2024 – 19 February 2024
HAC bandwidth	4

Table B.3: Baseline LR decision table

Level	Critical value	Reject H_0
10%	1988.990	Yes
5%	2253.466	Yes
1%	2724.285	No

Thus, in the baseline specification, the null hypothesis of no structural break is rejected at the 5% level.

B.3 Robustness to the trimming parameter

Table B.4: Robustness of the LR result to the trimming parameter

Trim	$\hat{r}$	$\sup LR$	Break date	Reject at 5%
0.10	8	2474.556	12 February 2024	Yes
0.15	8	2474.556	12 February 2024	Yes
0.20	8	2474.556	12 February 2024	Yes

The estimated break date remains unchanged across the considered trimming parameters.

B.4 Robustness to the number of factors

Table B.5: Robustness of the LR result to the number of factors

Used r	$\sup LR$	Break date	Break fraction	Reject at 5%
7	2302.996	12 February 2024	0.618677	Yes
8	2474.556	12 February 2024	0.618677	Yes
9	2522.158	12 February 2024	0.618677	Yes

The estimated break date is the same for all three values of r , and the null hypothesis is rejected at the 5% level in each case.

B.5 Pre-break and post-break factor selection

Table B.6: Bai–Ng factor-selection results before and after the estimated break

Criterion	Pre-break sample	Post-break sample
PCp1	15	18
PCp2	15	15
PCp3	20	20
ICp1	6	11
ICp2	3	7
ICp3	20	20

Using ICp2, the selected number of factors increases from 3 before the estimated break to 7 after the estimated break.

B.6 Pre-break and post-break comparison

Table B.7: Comparison of the factor structure before and after the estimated break

Subsample	Period	T_{sub}	$\hat{r}$	Explained variation	λ_1
Pre-break	11 Jan 2021 – 12 Feb 2024	159	3	0.538478	0.246042
Post-break	19 Feb 2024 – 29 Dec 2025	98	7	0.940412	1.015337

Appendix C

Supplementary Simulation Evidence

This appendix reports supplementary results from the Monte Carlo simulation study.

C.1 Simulation design

Table C.1: Monte Carlo simulation design

Scenario	T	N	r	σ_f	σ_e	Shift	Break date
No break	300	200	2	1	0.5	0.0	–
Baseline break	300	200	2	1	0.5	1.0	200
Moderate break	300	200	2	1	0.5	0.5	200
Noisier break	300	200	2	1	1.0	1.0	200

The Monte Carlo experiment uses 100 replications for each scenario. The final factor-selection criterion used in the PCA and LR procedure is ICp2. The maximum admissible number of factors is $k_{\max} = 20$, and the trimming parameter is $\varepsilon = 0.15$.

C.2 Illustrative single-run results

Table C.2: Illustrative single-run results across simulation scenarios

Scenario	$\hat{r}$	Expl. var.	sup LR	Estimated break	CV 5%	Reject 5%
No break	2	0.885833	5.764	143	12.743	No
Baseline break	3	0.895384	1381.770	200	79.092	Yes
Moderate break	3	0.892763	1070.463	200	77.614	Yes
Noisier break	3	0.695664	980.607	200	84.634	Yes

C.3 Factor-selection summary

Table C.3: Monte Carlo factor-selection summary

Scenario	Mean $\hat{r}$	Median $\hat{r}$	$\Pr(\hat{r} = r)$	$\Pr(\hat{r} > r)$
No break	2	2	1.00	0.00
Baseline break	3	3	0.00	1.00
Moderate break	3	3	0.00	1.00
Noisier break	3	3	0.00	1.00

C.4 Rejection frequencies

Table C.4: Monte Carlo rejection frequencies

Scenario	10%	5%	1%	Mean sup LR
No break	0.05	0.04	0.00	7.353
Baseline break	1.00	1.00	1.00	1336.687

Continued on the next page

Table C.4: Monte Carlo rejection frequencies – continued

Scenario	10%	5%	1%	Mean sup LR
Moderate break	1.00	1.00	1.00	1073.042
Noisier break	1.00	1.00	1.00	1022.535

C.5 Break-date accuracy

Table C.5: Monte Carlo break-date accuracy in break scenarios

Scenario	True	Mean	Median	MAE	Within 5
Baseline break	200	200.07	200	0.11	1.00
Moderate break	200	200.00	200	0.12	1.00
Noisier break	200	200.03	200	0.13	1.00

Note: “Within 5” denotes the fraction of replications for which $|\widehat{T}_1 - T_1| \leq 5$.

Bibliography

- [1] Lucia Alessi, Matteo Barigozzi, and Marco Capasso. “Improved Penalization for Determining the Number of Factors in Approximate Factor Models”. In: *Statistics & Probability Letters* 80.23–24 (2010), pp. 1806–1813. DOI: [10.1016/j.spl.2010.07.004](https://doi.org/10.1016/j.spl.2010.07.004).
- [2] Jushan Bai, Jiangtao Duan, and Xu Han. “The Likelihood Ratio Test for Structural Changes in Factor Models”. In: *Journal of Econometrics* 244.2 (2024), p. 105745. DOI: [10.1016/j.jeconom.2023.105631](https://doi.org/10.1016/j.jeconom.2023.105631).
- [3] Jushan Bai and Kunpeng Li. “Maximum Likelihood Estimation and Inference for Approximate Factor Models of High Dimension”. In: *Review of Economics and Statistics* 98.2 (2016), pp. 298–309. DOI: [10.1162/REST_a_00536](https://doi.org/10.1162/REST_a_00536).
- [4] Jushan Bai and Serena Ng. “Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions”. In: *Econometrica* 74.4 (2006), pp. 1133–1150. DOI: [10.1111/j.1468-0262.2006.00696.x](https://doi.org/10.1111/j.1468-0262.2006.00696.x).
- [5] Jushan Bai and Serena Ng. “Determining the Number of Factors in Approximate Factor Models”. In: *Econometrica* 70.1 (2002), pp. 191–221. DOI: [10.1111/1468-0262.00273](https://doi.org/10.1111/1468-0262.00273).
- [6] Jushan Bai and Serena Ng. “Determining the Number of Primitive Shocks in Factor Models”. In: *Journal of Business & Economic Statistics* 25.1 (2007), pp. 52–60. DOI: [10.1198/0735001060000000413](https://doi.org/10.1198/0735001060000000413).
- [7] Jushan Bai and Serena Ng. “Large Dimensional Factor Analysis”. In: *Foundations and Trends in Econometrics* 3.2 (2008), pp. 89–163. DOI: [10.1561/08000000002](https://doi.org/10.1561/08000000002).

- [8] Jushan Bai and Serena Ng. “Principal Components and Regularized Estimation of Factor Models”. arXiv preprint arXiv:1708.08137. 2017. URL: <https://arxiv.org/abs/1708.08137>.
- [9] Jushan Bai and Peng Wang. “Econometric Analysis of Large Factor Models”. In: *Annual Review of Economics* 8 (2016), pp. 53–80. DOI: [10.1146/annurev-economics-080315-015356](https://doi.org/10.1146/annurev-economics-080315-015356).
- [10] Bank of Russia. *Statement by Bank of Russia Governor Elvira Nabiullina in follow-up to Board of Directors meeting on 16 February 2024*. Bank of Russia. Feb. 16, 2024. URL: <https://www.cbr.ru/eng/press/event/?id=18424> (visited on 05/16/2026).
- [11] Jörg Breitung and Sandra Eickmeier. *Testing for Structural Breaks in Dynamic Factor Models*. Discussion Paper Series 1: Economic Studies 2009,05. Deutsche Bundesbank, 2009. URL: <https://hdl.handle.net/10419/27659>.
- [12] Mehmet Caner and Xu Han. “Selecting the Correct Number of Factors in Approximate Factor Models: The Large Panel Case With Group Bridge Estimators”. In: *Journal of Business & Economic Statistics* 32.3 (2014), pp. 359–374. DOI: [10.1080/07350015.2014.880349](https://doi.org/10.1080/07350015.2014.880349).
- [13] José Luis Cendejas Bueno et al. “Testing for Structural Breaks in Factor Loadings: An Application to International Business Cycle”. In: *Economic Modelling* 28.1–2 (2011), pp. 259–263. DOI: [10.1016/j.econmod.2010.09.004](https://doi.org/10.1016/j.econmod.2010.09.004).
- [14] Liang Chen, Juan J. Dolado, and Jesús Gonzalo. “Detecting Big Structural Breaks in Large Factor Models”. In: *Journal of Econometrics* 180.1 (2014), pp. 30–48. DOI: [10.1016/j.jeconom.2014.01.006](https://doi.org/10.1016/j.jeconom.2014.01.006).
- [15] Gregory Connor and Robert A. Korajczyk. “A Test for the Number of Factors in an Approximate Factor Model”. In: *The Journal of Finance* 48.4 (1993), pp. 1263–1291. DOI: [10.1111/j.1540-6261.1993.tb04754.x](https://doi.org/10.1111/j.1540-6261.1993.tb04754.x).
- [16] Massacci Daniele. “Selecting the Number of Factors in Approximate Factor Models Using Group Variable Regularization”. In: *Econometric Reviews* 43.10 (2024), pp. 796–823. DOI: [10.1080/07474938.2024.2365795](https://doi.org/10.1080/07474938.2024.2365795).

- [17] Jianqing Fan, Kunpeng Li, and Yuan Liao. “Recent Developments in Factor Models and Applications in Econometric Learning”. In: *Annual Review of Financial Economics* 13 (2021), pp. 401–430. DOI: [10.1146/annurev-financial-091420-011735](https://doi.org/10.1146/annurev-financial-091420-011735).
- [18] George Kapetanios. “A Testing Procedure for Determining the Number of Factors in Approximate Factor Models With Large Datasets”. In: *Journal of Business & Economic Statistics* 28.3 (2010), pp. 397–409. DOI: [10.1198/jbes.2009.07239](https://doi.org/10.1198/jbes.2009.07239).
- [19] Bonsoo Koo, Benjamin Wong, and Ze-Yu Zhong. “Disentangling Structural Breaks in Factor Models for Macroeconomic Data”. arXiv preprint arXiv:2303.00178. 2025. URL: <https://arxiv.org/abs/2303.00178>.
- [20] Yu-Ning Li, Degui Li, and Piotr Fryzlewicz. “Detection of Multiple Structural Breaks in Large Covariance Matrices”. In: *Journal of Business & Economic Statistics* 41.3 (2023), pp. 846–861. DOI: [10.1080/07350015.2022.2076686](https://doi.org/10.1080/07350015.2022.2076686).
- [21] Shujie Ma and Liangjun Su. “Estimation of Large Dimensional Factor Models With an Unknown Number of Breaks”. In: *Journal of Econometrics* 207.1 (2018), pp. 1–29. DOI: [10.1016/j.jeconom.2018.06.019](https://doi.org/10.1016/j.jeconom.2018.06.019).
- [22] Moscow Exchange. *Market Data*. 2025. URL: <https://www.moex.com/>.
- [23] Alexei Onatski. “Testing Hypotheses About the Number of Factors in Large Factor Models”. In: *Econometrica* 77.5 (2009), pp. 1447–1479. DOI: [10.3982/ECTA6964](https://doi.org/10.3982/ECTA6964).
- [24] Reuters. “EU approves 13th sanctions package against Russia”. In: *Reuters* (Feb. 21, 2024). URL: <https://www.reuters.com/world/europe/eu-approves-13th-sanctions-package-against-russia-eu-sources-2024-02-21/> (visited on 05/16/2026).
- [25] Reuters. “EU countries close to agreement on new Russia sanctions, diplomats say”. In: *Reuters* (Feb. 14, 2024). URL: <https://www.reuters.com/world/europe/eu-countries-close-agreement-new-russia-sanctions-diplomats-say-2024-02-14/> (visited on 05/16/2026).

- [26] Reuters. “Exclusive: Turkish-Russian trade hit by fresh U.S. sanctions threat”. In: *Reuters* (Feb. 19, 2024). URL: <https://www.reuters.com/markets/turkish-russian-trade-hit-by-fresh-us-sanctions-threat-2024-02-19/> (visited on 05/16/2026).
- [27] Donald B. Rubin and Dorothy T. Thayer. “EM Algorithms for ML Factor Analysis”. In: *Psychometrika* 47.1 (1982), pp. 69–76. DOI: [10.1007/BF02293851](https://doi.org/10.1007/BF02293851).
- [28] James H. Stock and Mark W. Watson. *Diffusion Indexes*. NBER Working Paper 6702. National Bureau of Economic Research, 1998.
- [29] Jianhong Wu. “Eigenvalue Difference Test for the Number of Common Factors in the Approximate Factor Models”. In: *Economics Letters* 169 (2018), pp. 63–67. DOI: [10.1016/j.econlet.2018.05.009](https://doi.org/10.1016/j.econlet.2018.05.009).
- [30] Jianhong Wu. “Robust Determination for the Number of Common Factors in the Approximate Factor Models”. In: *Economics Letters* 144 (2016), pp. 102–106. DOI: [10.1016/j.econlet.2016.04.026](https://doi.org/10.1016/j.econlet.2016.04.026).
- [31] Xianli Zeng, Yingcun Xia, and Linjun Zhang. “Double Cross Validation for the Number of Factors in Approximate Factor Models”. arXiv preprint arXiv:1907.01670. 2019. URL: <https://arxiv.org/abs/1907.01670>.