Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Disentangling Structural Breaks in High Dimensional Factor Models

CAMA Working Paper 15/2023
March 2023

Bonsoo Koo
Monash University
Centre for Applied Macroeconomic Analysis, ANU

Benjamin Wong
Monash University
Centre for Applied Macroeconomic Analysis, ANU

Ze-Yu Zhong
Monash University

Abstract

We disentangle structural breaks in dynamic factor models by establishing a projection based equivalent representation theorem which decomposes any break into a rotational change and orthogonal shift. Our decomposition leads to the natural interpretation of these changes as a change in the factor variance and loadings respectively, which allows us to formulate two separate tests to differentiate between these two cases, unlike the pre-existing literature at large. We derive the asymptotic distributions of the two tests, and demonstrate their good finite sample performance. We apply the tests to the FRED-MD dataset focusing on the Great Moderation and Global Financial Crisis as candidate breaks, and find evidence that the Great Moderation may be better characterised as a break in the factor variance as opposed to a break in the loadings, whereas the Global Financial Crisis is a break in both. Our empirical results highlight how distinguishing between the breaks can nuance the interpretation attributed to them by existing methods.

Keywords

Factor space, structural instability, breaks, principal components, dynamic factor models

JEL Classification

C12, C38, C55, E37

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

Disentangling Structural Breaks in High Dimensional Factor Models *†‡

Bonsoo Koo §, Benjamin Wong ¶, and Ze Yu Zhong ‖

Monash University

Dated: February 28, 2023

Abstract

We disentangle structural breaks in dynamic factor models by establishing a *projection based equivalent representation theorem* which decomposes any break into a rotational change and orthogonal shift. Our decomposition leads to the natural interpretation of these changes as a change in the factor variance and loadings respectively, which allows us to formulate two separate tests to differentiate between these two cases, unlike the pre-existing literature at large. We derive the asymptotic distributions of the two tests, and demonstrate their good finite sample performance. We apply the tests to the FRED-MD dataset focusing on the Great Moderation and Global Financial Crisis as candidate breaks, and find evidence that the Great Moderation may be better characterised as a break in the factor variance as opposed to a break in the loadings, whereas the Global Financial Crisis is a break in both. Our empirical results highlight how distinguishing between the breaks can nuance the interpretation attributed to them by existing methods.

Keywords: factor space, structural instability, breaks, principal components, dynamic factor models

*We acknowledge that this research was supported in part by the Monash eResearch Centre and eSolutions Research Support Services through the use of the MonARCH HPC Cluster.

†We acknowledge the financial support of the Australian Research Council under Grants LP160101038, DP210101440, and DE200100693.

‡We thank Xu Han and Fa Wang for making their code available, in addition to the comments by Xu Han, Daniele Massacci, Mirco Rubin, Matteo Barigozzi, Serena Ng, and other participants at the 2022 EC-squared conference.

§Department of Econometrics and Business Statistics, Monash University, Melbourne, Australia and Centre for Applied Macroeconomic Analysis, Australian National University Email: bonsoo.koo@monash.edu

¶Department of Econometrics and Business Statistics, Monash University, Melbourne, Australia and Centre for Applied Macroeconomic Analysis, Australian National University. Email: benjamin.wong@monash.edu

‖Corresponding author. Department of Econometrics and Business Statistics, Monash University, Melbourne, Australia. Email: ze.zhong@monash.edu

1 Introduction

High dimensional factor models are widely used in empirical macroeconomics and finance, and assume that a large panel of time series are generated according to some small number of latent factors. Thus, a large dataset can be effectively parameterized by a set of individual "loadings" and a set of common "factors," as a means of dimensional reduction, where one subsequently uses these factors for both forecasting (Stock and Watson (2002)) and structural analysis (Bernanke et al. (2005)). However, the theory underlying factor models assumes at least "mild" stability in model parameters (Stock and Watson (1998), Bai (2003)). In reality, empirical data is unlikely to maintain parameter stability, and the analysis of structural breaks in factor models presents a unique identification issue wherein breaks in the loadings and breaks in the factors cannot be easily disentangled.

Consider a factor model for $x_{it}, t = 1, \dots, T, i = 1, \dots, N$:

$$x_{it} = \lambda_i^\intercal f_t + e_{it}, \tag{1.1}$$

where λ_i is an $r \times 1$ vector of individual loadings, f_t is an $r \times 1$ vector of common factors, and e_{it} the idiosyncratic error. Suppose that there exists a structural break such that λ_i doubles in value. Because both the factors and loadings are unobserved and enter in a multiplicative relationship, this is observationally equivalent to f_t doubling in value. Consequently, it is typical for the literature to assume "strict stationarity" in the factors as an identification condition in order to pin down changes in the loadings, and necessarily interpret all "breaks" as occurring in the loadings (e.g. Chen et al. (2014), Han and Inoue (2015), Baltagi et al. (2017), and others). Such an interpretation could be misleading, because the literature has typically identified periods such as the Great Moderation (Stock and Watson (2009), Breitung and Eickmeier (2011), Baltagi et al. (2021)), the Global Financial Crisis (Ma and Su (2018), Barigozzi and Trapani (2020), Ma and Tu (2022)), and more recently the COVID-19 Pandemic (Bai et al. (2022)) as evidence of structural breaks, all of which are periods well known for the data displaying heteroscedasticity. Hence, it is unclear whether these results are capturing genuine breaks in the loadings, or simply picking up factor heteroscedasticity, two different cases with very different economic narratives, a concern initially raised by Stock and Watson (2016). In addition to their economic interpretations, differentiating these two cases is important from a mechanical viewpoint: breaks in the loadings can lead to the incorrect over-estimation of the number of factors if ignored, whereas breaks in the factors do not have this effect, a refinement of an existing concern put forth by Breitung and Eickmeier (2011).

Our contribution to the literature is a method testing whether these estimated breaks are breaks in the loadings, breaks in the factor variance, or both, thus disentangling the source of structural breaks. To this end, we propose a new *projection based* equivalent representation theorem, which decomposes any change in the factor loading matrix into a rotational break common across the entire panel, and a leftover orthogonal shift component idiosyncratic to each series. Our *projection-based* decomposition approach is motivated by the important mechanical differences of these breaks. Specifically, we observe that breaks in the factors can always be viewed as a twisting (rotation) of the same underlying factor space and are therefore absorbed into the factor estimates of the principal components (PC) estimator. Hence, this does not pose any issue for the purposes of determining the number of factors (such as using the criteria of Bai and Ng (2002)), or applying the inferential results of Bai (2003). Economically, such a break could be associated with the aforementioned Great Moderation, where the overall volatility of all series in the economy was observed to decrease. In contrast, due to their idiosyncratic nature, breaks in the factor loadings lie outside and are therefore orthogonal to the underlying factor space, and it is this orthogonality which leads to a so-called "augmentation" effect where the number of factors will be overestimated if ignored, as noted by Breitung and Eickmeier (2011). This incorrect overestimation of the number of factors can have many serious consequences, including worsening factor based forecasts in the setup of Stock and Watson (2002) as noted by Baltagi et al. (2021), or incorrect specification of factor models in a state space setup, which rely on PC based methods to estimate the number of factors.

By interpreting the rotational change as a break in factor variance, and the orthogonal shift as a break in factor loadings, we are thus able to disentangle these two effects. We emphasize that this is in contrast to other similar equivalent representation theorem based approaches in the literature, who typically assume "strict" stationarity in the factors and interpret all breaks as breaks in the loadings (Chen et al. (2014), Han and Inoue (2015), Baltagi et al. (2017)). Thus, although these have similar model setups, our *projection based* equivalent representation theorem is a further refinement of existing methods. Based on this decomposition, we then propose two separate tests: 1) a test for any evidence of rotational change, or factor variance, and 2) a test for any evidence of orthogonal shifts, or breaks in the loadings. We establish the asymptotic distributions of these two test statistics, and show that standard critical values can be used, leading to their easy implementation. Monte Carlo studies demonstrate that the tests have good size and power, and highlight the inability of existing tests to differentiate between these two types of breaks.

To the best of our knowledge, only a few contributions similarly try to disentangle structural breaks in factor models. Wang (2021) proposes an estimator for the number of breaks using eigenvalue ratios which is robust to changes in factor variance, but do not consider testing due to the difficulty in working with the distribution of eigenvalues. Our test statistics do not rely on eigenvalues, and hence we are able to derive standard asymptotic results and avoid this issue. Indeed, our test statistics converge to conventional

Chi-squared distributions, making them easy to implement for practitioners. Pelger and Xiong (2022) propose a general time varying framework, and construct a test statistic which tests whether the factor loadings in two regimes can be fully explained by a rotation via canonical correlations, essentially amounting to a test for evidence of orthogonal shifts, or legitimate breaks in loadings. Compared to their test statistic, our tests do not require the difficult estimation of a de-bias term, which is known to be non-robust to the specification of the error structure (e.g. Su and Wang (2017) and Su and Wang (2020)). Massacci (2021) proposes a test statistic for evidence of regime dependent loadings in a threshold setup that is robust against factor heteroscedasticity. All of these contributions similarly recognise that legitimate breaks in the loadings should be orthogonal to the original factor space, but their model setups, test statistics, and resulting asymptotics are all quite different. Compared to these existing papers, our setup is the only one which considers separately testing for evidence in the factor variance and loadings, and hence provides the most comprehensive framework to accurately pin down the *source* of structural breaks.

We apply our tests to the FRED-MD dataset of McCracken and Ng (2015) and focus on the two estimated break dates put forth by the literature: the Great Moderation, dated to be around 1984, and the Global Financial Crisis, dated to be around late 2008. We find that for the case of only one factor, the Great Moderation only rejects on the rotational test. Our orthogonal shift test could be interpreted as a test for evidence of breaks in the loadings while controlling for changes in the factor variance, and in this vein when compared to the tests of Breitung and Eickmeier (2011), these results suggest that for the case of one factor (supported by the estimators of Onatski (2010) and Ahn and Horenstein (2013)), the Great Moderation is more accurately described as a break in the factor variance, as opposed to a break in the loadings. Although we reject both the rotational and orthogonal shift test for the case of two or more factors, for the case of two factors, we find that most of the evidence for breaks in loadings is isolated to price series, and hence depending on the application, may not pose issues for the practitioner. In contrast, the evidence for the Global Financial Crisis tends to favour a break in both the factor variance and the loadings. These results bring nuance to how these different periods of instability can be characterised, which could lead to different implications.

In this paper, all limits are taken as both N, T tend to infinity simultaneously, and δ_{NT} is defined as $\min(\sqrt{T}, \sqrt{N})$. We use $\|\cdot\|$ to denote the Frobenius norm of a vector or matrix, $\xrightarrow{p}$ denotes convergence in probability, $\Rightarrow$ denotes weak convergence of stochastic processes, $\xrightarrow{d}$ denotes convergence in distribution, $vech()$ denotes the column-wise vectorisation of a square matrix with the upper triangle excluded, and $\lfloor \cdot \rfloor$ denotes the floor or integer part operator. We use M to denote generic constants which may take different values, and $A^{-\intercal}$ denotes the inverse transpose of any invertible matrix A.

2 Disentangling Structural Breaks in Dynamic Factor Models

2.1 A Projection Based Equivalent Representation Theorem for Structural Breaks

Let x_{it} denote the observation for the ith cross section at period t for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Let $\lfloor \pi T \rfloor$ denote the break date, where π is the break fraction which splits the data into subsample sizes of $T_1 = \lfloor \pi T \rfloor$ and $T_2 = T - \lfloor \pi T \rfloor$ respectively. Suppose that

x_{it} is generated from r common factors with the following static factor representation:

$$x_{it} = \begin{cases} \lambda_{1,i}^{\intercal} f_t + e_{it}, & \text{for} \quad t = 1, \dots, \lfloor \pi T \rfloor, \\ \lambda_{2,i}^{\intercal} f_t + e_{it}, & \text{for} \quad t = \lfloor \pi T \rfloor + 1, \dots, T, \end{cases} \tag{2.1}$$

where f_t is a $r \times 1$ vector of factors, $\lambda_{1,i}, \lambda_{2,i}$ are the corresponding $r \times 1$ loadings for series i before and after the break respectively, and e_{it} is the idiosyncratic shock. We require the number of factors r to be identical before and after the break, because our method relies on subsample estimates after splitting the sample, a common regularity condition found in many other methods utilizing subsample estimates (e.g. Ma and Su (2018) and Bai et al. (2020)). Throughout the paper, we treat both the number of factors r and the break fraction π as known, as both of these can be consistently estimated (see Remarks 1 and 2) without affecting any of our asymptotic results.

Equation (2.1) can also be written using matrix form:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_1 \Lambda_1^{\intercal} \\ F_2 \Lambda_2^{\intercal} \end{bmatrix} + e, \tag{2.2}$$

where $F_1 = (f_1, \dots, f_{\lfloor \pi T \rfloor})^{\intercal}$ is a $\lfloor \pi T \rfloor \times r$ matrix of factors before the break, $F_2 = (f_{\lfloor \pi T \rfloor + 1}, \dots, f_T)^{\intercal}$ is a $(T - \lfloor \pi T \rfloor) \times r$ matrix of factors after the break, and $\Lambda_1 = (\lambda_{1,1}, \dots, \lambda_{1,N})^{\intercal}$, $\Lambda_2 = (\lambda_{2,1}, \dots, \lambda_{2,N})^{\intercal}$ are both $N \times r$ matrices of respective loadings, and X_1, X_2 denote the respective partitions of X based on the break fraction. The matrices $F_1, F_2, \Lambda_1, \Lambda_2, e$ are all unknown. To disentangle breaks in the factor loadings from breaks in the factors, we decompose $\Lambda_2 = \Lambda_1 Z + W$, where Z represents an $r \times r$ nonsingular rotational change, and $W = (w_1, \dots, w_N)^{\intercal}$ is an $N \times r$ matrix representing the orthogonal shift that is idiosyncratic across the cross section. It follows that this decomposition can be used to yield the following equivalent representation theorem:

$$X = \begin{bmatrix} F_1 \Lambda_1^{\intercal} \\ F_2[\Lambda_1 Z + W]^{\intercal} \end{bmatrix} + \begin{bmatrix} e_{(1)} \\ e_{(2)} \end{bmatrix} \tag{2.3}$$

$$= \begin{bmatrix} F_1 & 0 \\ F_2 Z^{\intercal} & F_2 \end{bmatrix} \begin{bmatrix} \Lambda_1^{\intercal} \\ W^{\intercal} \end{bmatrix} + \begin{bmatrix} e_{(1)} \\ e_{(2)} \end{bmatrix}$$

$$X = G\Xi^{\intercal} + e \tag{2.4}$$

where $e_{(1)} = (e_1, \ldots, e_{\lfloor \pi T \rfloor})$ and $e_{(2)} = (e_{\lfloor \pi T \rfloor + 1}, \ldots, e_T)$, and each $e_t = (e_{1t}, \ldots, e_{Nt})^{\intercal}$. Equation (2.4) shows that any rotational changes induced by a non-identity Z are absorbed into the factors, and any orthogonal shifts W will result in the augmentation of the factor space. Equation (2.4) is a version of an equivalent representation theorem (ERT), and re-expresses a factor model with structural breaks in its loadings into an observationally equivalent model with time invariant loadings. ERTs were initially formulated by Han and Inoue (2015), Baltagi et al. (2017) and others, and Equation (2.4) aims to complement these. If one were to ignore the break and naively use the PC estimator over the whole subsample, it will instead be consistent for an observationally equivalent model with so called *pseudo* factors G and time invariant loadings Ξ.

Previously, it has been thought that changes in the loadings cannot be separately identified from changes in the variance of the factors, which can be equivalently be represented as a rotational change common to all loadings. This is because existing methods used the pseudo factors G in order to either test for existence of any breaks (Han and Inoue (2015), Chen et al. (2014)), and/or estimate the break fraction (Baltagi et al. (2017), Baltagi et al. (2021), Duan et al. (2022)). Methods utilizing the estimated pseudo factors from the whole sample will necessarily have power against heteroscedasticity in the factors, even if the loadings are actually time invariant.

Our projection based formulation aims to differentiate between breaks in the factor variance versus breaks in the factor loadings, and is motivated by the mechanical properties of the PC estimator. It is well known that the PC estimator only estimates the underlying factor space up to an arbitrary rotation. However, any break in the factor variance can always be thought of some suitable twisting or stretching of the factors themselves, i.e. a rotation. Because the factors still span the same underlying space, breaks in the factor variance will be absorbed by the PC estimator. In contrast, if there are breaks in the loadings, due to their idiosyncratic nature the changes in each series cannot be explained by the existing factors, and therefore lie outside the space spanned by the factors. Hence, if one were to ignore the break, extra factors need to be estimated in order to capture the same information over the whole sample. We formalize these different mechanical effects of the PC estimator of these breaks by classifying them accordingly.

We first define a "type 1" break as presence of orthogonal shifts where $W \neq \mathbf{0}$ and $Z = I_r$, and this corresponds to the type 1 break as defined by Han and Inoue (2015), Baltagi et al. (2017), and type A break by Duan et al. (2022) as per respective nomenclatures. Due to the orthogonality of W to the original factor space, these breaks cannot be absorbed into the factors, and hence can only be interpreted as a legitimate break in the loadings. With our setup, it is now more clearly understood that it is the orthogonality induced by breaks in the loadings that lead to the factor "augmentation" effect raised by Breitung and Eickmeier (2011).

We next define a "type 2" break as a rotational break where $Z \neq I_r$ and $W = \mathbf{0}$. Type 2 breaks occur when all of the loadings across the cross section are rotated in a homogeneous way, or more naturally, a change in the factor variance. Indeed, it is quite difficult to imagine or justify such changes in the loadings practically, and they are often ruled out by assumption (e.g. Chen et al. (2014) and Ma and Tu (2022)), or similarly to us, interpreted as a change in the factor variance (Wang (2021) and Pelger and Xiong (2022)). Rotational breaks correspond to type 2 breaks of Han and Inoue (2015) and Baltagi et al. (2017), and the type B breaks of Duan et al. (2022). We require Z to be non-singular, and hence this rules out the case of disappearing factors, which could be viewed as somewhat limiting compared to earlier literature which used methods based on the pseudo factors. This is because we require inferential results as estimated in each subsample before and after the break, and presently it is not clear how to do this in the presence of disappearing factors. Such a regularity assumption is quite common in the literature (e.g. see Chen et al. (2014), Su and Wang (2017), Ma and Su (2018) and Bai et al. (2020)).[1]

Finally, we define a "type 3" break as simply a combination of the two breaks where there is both a rotation and orthogonal shift, i.e. $Z \neq I_r$ and $W \neq \mathbf{0}$.

Thus, the task of disentangling breaks in the factor variance from breaks in the loadings can be expressed in the form of two hypothesis tests: 1) a test for any evidence of rotation

$$\mathcal{H}_0 : Z = I_r, \quad \mathcal{H}_1 : Z \neq I_r, \tag{2.5}$$

which does not maintain any conditions on W, and 2) a test for any evidence of orthogonal shifts:

$$\mathcal{H}_0 : W_{(N\times r)} = \mathbf{0}, \quad \mathcal{H}_1 : W_{(N\times r)} \neq \mathbf{0}, \tag{2.6}$$

which in turn does not maintain any conditions on Z. We emphasise that because these two types of breaks can occur together, *both* tests need to be run in order to tease out

[1] Similar to methods that rule out disappearing factors, our method seems to have power against the case where there is a disappearing factor, but the theoretical results are unclear, similar to Chen et al. (2014).

which *type* of break has occurred.

2.2 Estimation and Consistency Results

2.2.1 Estimation

We now discuss estimation and the asymptotic properties of the estimators. Define $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$, the OLS fits from the estimates using the PC estimates $\tilde{F}_1 = (\tilde{f}_{1,t}, \dots, \tilde{f}_{1,\lfloor \pi T \rfloor})^\intercal$ and $\tilde{F}_2 = (\tilde{f}_{2,\lfloor \pi T \rfloor + 1}, \dots, \tilde{f}_{2,T})$, which are $\sqrt{T_1}$ and $\sqrt{T_2}$ times the eigenvectors corresponding to the r largest eigenvalues of $X_1 X_1^\intercal$ and $X_2 X_2^\intercal$ respectively. We define the feasible estimators for Z and W as

$$\tilde{Z} = (\tilde{\Lambda}_1^\intercal \tilde{\Lambda}_1)^{-1} \tilde{\Lambda}_1^\intercal \tilde{\Lambda}_2, \tag{2.7}$$

$$\tilde{W} = \tilde{\Lambda}_2 - \tilde{\Lambda}_1 \tilde{Z}. \tag{2.8}$$

Because $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ are estimates of Λ_1 and Λ_2 up to arbitrary rotations, $\tilde{Z}$ and $\tilde{W}$ cannot be directly interpreted, and the task of disentanglement is not straightforward. However, it turns out that $\tilde{Z}$ and $\tilde{W}$ are able to recover the true Z and W up to arbitrary rotations as well. To analyze them, we make the following assumptions. Let $\iota_{1t} \equiv \mathbf{1}\{t \leq \lfloor \pi T \rfloor\}$ and $\iota_{2t} \equiv \mathbf{1}\{t \geq \lfloor \pi T \rfloor + 1\}$.

2.2.2 Estimation Assumptions

Assumption 1. *$E\|f_t\|^4 < \infty$, $E(f_t f_t^\intercal) = \Sigma_F$ and $\frac{1}{T}\sum_{t=1}^{T} f_t f_t^\intercal \xrightarrow{p} \Sigma_F$ for some positive definite Σ_F.*

Assumption 2. *For $m = 1, 2$, there exists a positive constant M such that $E\|\lambda_{m,i}\|^4 \leq M$, $\|\Lambda_m^\intercal \Lambda_m / N\| - \Sigma_{\Lambda_m} \xrightarrow{p} 0$ for some $\Sigma_{\Lambda_m} > 0$, and $\|\Lambda_m^\intercal \Lambda_m / N - \Sigma_{\Lambda_m}\| = O_p\left(N^{-1/2}\right)$. Analogously, when $W \neq 0$, $\|W^\intercal W / N - \Sigma_W\| \xrightarrow{p} 0$ for some $\Sigma_W > 0$, and $\|W^\intercal W / N - \Sigma_W\| =$*

$O_p\left(N^{-1/2}\right)$.

Assumption 3. *There exists some positive constant $M < \infty$ such that for all N and T:*

(a) $E(e_{it}) = 0, E|e_{it}|^8 \leq M$

(b) $E(e_s^\intercal e_t/N) = E(N^{-1}\sum_{i=1}^N e_{is}e_{it}) = \gamma_N(s,t)$, $|\gamma_N(s,s)| \leq M$ *for all* s, *and* $T^{-1}\sum_{t=1}^T\sum_{s=1}^T |\gamma_N(s,t)| \leq M$.

(c) $E(e_{it}e_{jt}) = \tau_{ij,t}$, *with* $|\tau_{ij,t}| < \tau_{ij}$ *for some* τ_{ij} *and for all* t. *In addition,* $N^{-1}\sum_{i=1}^N\sum_{j=1}^N |\tau_{ij}| \leq M$.

(d) $E(e_{it}e_{js}) = \tau_{ij,ts}$, *and* $(NT)^{-1}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T\sum_{s=1}^T |\tau_{ij,ts}| \leq M$.

(e) For every (t,s), $E\left|N^{-1/2}\sum_{i=1}^N [e_{is}e_{it} - E(e_{is}e_{it})]\right|^4 \leq M$.

Assumption 4. *For $m = 1, 2$, the variables $\{\lambda_{m,i}\}$, $\{f_t\}$ and $\{e_{it}\}$ are mutually independent groups.*

Assumption 5. *There exists an $M < \infty$ such that for all T and N, and for every $t \leq T$ and $i \leq N$ such that:*

(a) $\sum_{s=1}^T |\gamma_N(s,t)| \leq M$

(b) $\sum_{k=1}^N |\tau_{ki}| \leq M$

Assumption 6. *There exists an $M < \infty$ such that for all N, T and $m = 1, 2$:*

(a) $E\left\|\frac{1}{NT}\sum_{s=1}^T\sum_{k=1}^N f_s[e_{ks}e_{kt} - E(e_{ks}e_{kt})] \cdot \iota_{ms}\right\|^2 \leq M$ *for each* t,

(b) $E\left\|\frac{1}{\sqrt{NT}}\sum_{t=1}^T\sum_{k=1}^N f_t\lambda_{m,k}^\intercal e_{kt} \cdot \iota_{mt}\right\|^2 \leq M$,

(c) For each t $E\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_{m,i}e_{it}\right\|^4 \leq M$.

Assumption 7. *The eigenvalues of $(\Sigma_{\Lambda_1}\Sigma_F)$ and $(\Sigma_{\Lambda_2}\Sigma_F)$ are distinct.*

Assumption 8. *The break fraction* π *is bounded away from 0 and 1, and*

(a) $\left\| \frac{1}{\sqrt{NT}} \sum_{t=1}^{\lfloor \pi T \rfloor} \sum_{k=1}^{N} f_t \lambda_{m,k}^{\intercal} e_{kt} \iota_{mt} \right\|^2 = O_p(1)$, $\left\| \frac{1}{\sqrt{NT}} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} \sum_{k=1}^{N} f_t \lambda_{m,k}^{\intercal} e_{kt} \iota_{mt} \right\|^2 = O_p(1)$ *for* $m = 1, 2$, *and*

(b) $\left\| \frac{\sqrt{T}}{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} (f_t f_t^{\intercal} - \Sigma_F) \right\| = O_p(1)$, *and* $\left\| \frac{\sqrt{T}}{T - \lfloor \pi T \rfloor} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} (f_t f_t^{\intercal} - \Sigma_F) \right\| = O_p(1)$

Assumptions 1 to 7 are either straight from or slight modifications of those in Bai (2003). Assumption 1 is the same as Assumption A in Bai (2003), except that we require the second moment of f_t to be time invariant. This additional "strict" stationarity assumption is common as an identification condition (e.g. Han and Inoue (2015), Baltagi et al. (2017) and others) and necessarily limited the factors to exhibit no heteroscedasticity, but this is not restrictive in our case as all changes in Σ_F are characterized by Z. Assumption 2 is the same as Assumption B in Bai (2003), except that it specifies the convergence speed of $\Lambda_m^{\intercal} \Lambda_m / N$ to be no slower than $1/\sqrt{N}$ for $m = 1, 2$. Assumption 2 allows for the loadings to be random, and although this is not required for the purposes of estimation and the Z rotation test, it is required for the W orthogonal shift tests, and we therefore combine this assumption for simplicity. Assumptions 3 and 5 correspond exactly to Assumptions C and E in Bai (2003). Assumption 3 allows for weak serial and cross sectional correlation and define the *approximate* factor model. Assumption 5 is a strengthened version of Assumption 3, but still allows for heterogeneity in time and cross-sectional dimensions. Assumption 4 is standard in the factor modeling literature, and is the subsample version of Assumption D of Bai and Ng (2006). Assumption 7 corresponds to Assumption G in Bai (2003). Assumption 6 corresponds to Assumptions F1-F2 in Bai (2003). Although we require Assumption 6 which are moment conditions in Bai (2003), asymptotic normality of $N^{-1/2} \sum_{i=1}^{N} \lambda_i e_{it}$ are not required for the purposes of estimation. Also, Assumption 6 (c) is slightly stronger that Assumption F3 of Bai (2003), which only requires the existence of the second moments.

Assumption 8 requires that the sample sizes before and after the potential break date go to infinity. It is a weaker version of Assumption 8 in Han and Inoue (2015), who assumes that the terms are bounded uniformly in a range of potential π.

Recall that $\tilde{F}_1$ and $\tilde{F}_2$ are estimates of F_1 and F_2 up to two different arbitrary rotations. Specifically, we define the rotational basis in the first subsample as $H_1 = \left(\frac{\Lambda_1^\intercal \Lambda_1}{N}\right)\left(\frac{F_1^\intercal \tilde{F}_1}{T_1}\right) V_{NT,1}^{-1}$, and in the second subsample as $H_2 = \left(\frac{Z^\intercal \Lambda_1^\intercal \Lambda_1 Z}{N} + \frac{W^\intercal W}{N}\right)\left(\frac{F_2^\intercal \tilde{F}_2}{T_2}\right) V_{NT,2}^{-1}$, where $V_{NT,1}, V_{NT,2}$ denote the diagonal matrix of eigenvalues of the first r eigenvalues of $(NT_1)^{-1} X_1 X_1^\intercal$ and $(NT_2)^{-1} X_2 X_2^\intercal$ respectively.[2]

Theorem 2.1. *Under Assumptions 1 to 8,* $\left\| \tilde{Z} - H_1^\intercal Z H_2^{-\intercal} \right\| = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

Although Theorem 2.1 shows that $\tilde{Z}$ itself is estimated up to a rotation and cannot be directly interpreted, the specific formulation of $\tilde{Z}$ allows us to present the following result.

Theorem 2.2. *Under Assumptions 1 to 8, and as* $\frac{\sqrt{N}}{T} \to \infty$*, for each* t*:* $\left\| \tilde{Z} \tilde{f}_{2,t} - H_1^\intercal Z f_t \right\| = o_p(1)$.

Theorem 2.2 is a direct consequence of $\left\| \tilde{f}_{2,t} - H_2^\intercal f_t \right\| = o_p(1)$ from Lemma 1 of Bai (2003), and Theorem 2.1. Theorem 2.2 shows that post multiplying $\tilde{f}_2$ by $\tilde{Z}$ rotates it back to the same rotational basis as $\tilde{f}_1$, and maintains the rotation Z, if any. Thus, if we combine $\tilde{f}_{1,t}$ and $\tilde{Z}\tilde{f}_{2,t}$ together, we have the following.

Corollary 2.2.1. *Under Assumptions 1 to 8, and as* $\frac{\sqrt{N}}{T} \to 0$*, for each* t*:*

$$\widehat{f}_t = \begin{cases} \tilde{f}_{1,t} \xrightarrow{p} H_1^\intercal f_t & \textit{for} \quad t = 1, 2, \ldots, \lfloor \pi T \rfloor, \\ \tilde{Z}\tilde{f}_{2,t} \xrightarrow{p} H_1^\intercal Z f_t & \textit{for} \quad t = \lfloor \pi T \rfloor + 1, \ldots, T. \end{cases} \tag{2.9}$$

[2] There exists another observationally equivalent parameterization $H_2 = \left(\Lambda_1^\intercal \Lambda_1 + Z^{-\intercal} W^\intercal W Z^{-1}\right)\left(Z F_2^\intercal \tilde{F}_2\right)/(NT_2) V_{NT,2}^{-1}$, where the rotation Z is parameterized as part of the factors. It is straightforward to verify that either parameterization leads to same result stated in Theorem 2.2. For more details, see **??**

Corollary 2.2.1 shows that the combined series $\widehat{F} = (\widehat{f}_1, \ldots, \widehat{f}_T)^\intercal$ is on the same rotational basis both before and after the break, and can thus form the basis of a test for evidence of rotational breaks. Importantly, $\widehat{f}_t$ is free from the effects of any possible orthogonal shifts induced by W, and thus isolates the rotational change in the factor variance.

Theorem 2.3. *Under Assumptions 1 to 8,* $\frac{1}{N}\left\|\tilde{W} - WH_2^{-\intercal}\right\|^2 = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

Theorem 2.3 shows that $\tilde{W}$ estimates the true W up to an arbitrary rotation. Thus, if the true $W = \mathbf{0}$, then $\tilde{W}$ should also be close to zero, and can serve as a foundation for statistical tests.

2.3 Z Test for Rotational Changes

We first present the test statistic for evidence of rotational change $\mathcal{H}_0 : Z = I$ against $\mathcal{H}_1 : Z \neq I$. Recall that by combining $\tilde{F}_1$ and $\tilde{F}_2\tilde{Z}^\intercal$, we have $\widehat{F}$, an estimate of the true factors and any rotation they undergo. This motivates a Wald test statistic based on whether the subsample means of $\widehat{f}_t\widehat{f}_t^\intercal$ are equal at a predetermined[3] break date $\lfloor \pi T \rfloor$:

$$\mathscr{W}_Z(\pi, \widehat{F}) = A_Z(\pi, \widehat{F})^\intercal \widehat{S}_Z(\pi, \widehat{F})^{-1} A_Z(\pi, \widehat{F}), \tag{2.10}$$

where $A_Z(\pi, \widehat{F}) = vech\left(\sqrt{T}\left(\frac{1}{\lfloor \pi T \rfloor}\sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t\widehat{f}_t^\intercal - \frac{1}{T - \lfloor \pi T \rfloor}\sum_{t=\lfloor \pi T+1 \rfloor}^{T} \widehat{f}_t\widehat{f}_t^\intercal\right)\right)$. Its long run variance estimate is defined as $\widehat{S}_Z(\pi, \widehat{F}) = \frac{1}{\pi}\widehat{\Omega}_{Z,(1)}(\pi\widehat{F}) + \frac{1}{1-\pi}\widehat{\Omega}_{Z,(2)}(\pi, \widehat{F})$, a weighted average

[3] We treat the break fraction as known *a priori* for simplicity, but any consistent estimate of this can be used instead without affecting any of the results, as noted in Remark 2.

of the variance from pre and post break data ($m = 1, 2$ respectively)

$$\widehat{\Omega}_{Z,(m)}(\pi, \widehat{F}) = \widehat{\Gamma}_{(m),0}(\pi, \widehat{F}) + \sum_{j=1}^{T_m-1} k\left(\frac{j}{b_{T_m}}\right)\left(\widehat{\Gamma}_{(m),j}(\pi, \widehat{F}) + \widehat{\Gamma}_{(m),j}(\pi, \widehat{F}))^{\intercal}\right),$$

$$\widehat{\Gamma}_{(1),j}(\pi, \widehat{F}) = \frac{1}{T_1}\sum_{t=j+1}^{T_1} vech(\widehat{f}_t\widehat{f}_t^{\intercal} - I_r)vech(\widehat{f}_t\widehat{f}_t^{\intercal} - I_r)^{\intercal}, \tag{2.11}$$

$$\widehat{\Gamma}_{(2),j}(\pi, \widehat{F}) = \frac{1}{T_2}\sum_{t=j+T_1+1}^{T} vech(\widehat{f}_t\widehat{f}_t^{\intercal} - I_r)vech(\widehat{f}_t\widehat{f}_t^{\intercal} - I_r)^{\intercal}, \tag{2.12}$$

where $k(.)$ is a real valued kernel, and b is the bandwidth, and its subscripts denotes the size of the (sub)samples used to estimate the long run variance.

2.3.1 Z Test Asymptotics under the Null Hypothesis

We define $\mathscr{W}_Z(\pi, FH_{0,1}) = A_Z(\pi, FH_{0,1})^{\intercal}\widehat{S}_Z(\pi, FH_{0,1})^{-1}A_Z(\pi, FH_{0,1})$ as the infeasible analog of $\mathscr{W}_Z(\pi, \widehat{F})$, and make the following assumptions.

Assumption 9. *(a) The Bartlett kernel of Newey and West (1987) is used, and there exists a constant $K > 0$ such that $b_T, b_{\lfloor \pi T \rfloor}$ and $b_{T-\lfloor \pi T \rfloor}$ are less than $KT^{1/3}$; and*

(b) $\frac{T^{2/3}}{N} \to 0$ as $N, T \to \infty$.

Assumption 10. *(a) $\Omega_Z = \lim_{T\to\infty} Var\left(vech\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H_{0,1}^{\intercal} f_t f_t^{\intercal} H_{0,1} - I_r\right)\right)$ is positive definite, and $\|\Omega_Z\| < \infty$. Its estimators $\widehat{\Omega}_{Z,(m)}(\pi, FH_{0,1})$ for $m = 1, 2$ are consistent such that $\left\|\widehat{\Omega}_{Z,(m)}(\pi, FH_{0,1}) - \Omega_Z\right\| = o_p(1)$,*

(b) $\mathscr{W}_Z(\pi, FH_{0,1}) \Rightarrow Q_p(\pi)$, where $Q_p(\pi) = [B_p(\pi) - \pi B_p(1)]^{\intercal}[B_p(\pi) - \pi B_p(1)]/(\pi(1-\pi))$, and $B_p(\cdot)$ is a $p = r(r+1)/2$ vector of independent Brownian motions on $[0, 1]$.

Assumption 9 specifies conditions for the Bartlett kernel. Assumption 10 (a) is a standard HAC assumption, and states that the infeasible estimators $\widehat{\Omega}_{Z,(1)}(\pi, FH_{0,1})$, $\widehat{\Omega}_{Z,(2)}(\pi, FH_{0,1})$ and $\widehat{\Omega}_Z(\pi, FH_{0,1})$ converge to their population counterpart Ω_Z. Assumption 10 (b) is the main result of Theorem 3 of Andrews (1993), and is necessary to establish the asymptotic distributions of the test statistics. As stated by Andrews (1993), for any fixed π, $Q_p(\pi)$ is distributed as a $\chi^2_{p=r(r+1)/2}$ random variable, and therefore standard critical values can be used. Assumption 10 (b) has been used in Han and Inoue (2015), and one can refer to Chen et al. (2014) for more primitive assumptions under which similar assumptions to Assumption 10 (b) hold. Note that we do not require the convergence of $\sup_\pi \mathscr{W}_Z(\pi, FH_{0,1})$ to $\sup_\pi Q_p(\pi)$, as we are focusing on a pre-known (or estimated) break fraction π.

Theorem 2.4. *Under Assumptions 1 to 10, and if* $\frac{\sqrt{T}}{N} \to 0$, *then* $\mathscr{W}_Z(\pi, \widehat{F}) \xrightarrow{d} \chi^2_{r(r+1)/2}$.

The proof of Theorem 2.4 is provided in the Supplementary Material, and involves proving the convergence of $\widehat{\Omega}_{Z,(1)}(\pi, \widehat{F})$, $\widehat{\Omega}_{Z,(2)}(\pi, \widehat{F})$, $\widehat{S}_Z(\pi, \widehat{F})$ and $A_Z(\pi, \widehat{F})$ to their infeasible counterparts. Theorem 2.4 shows that the feasible Wald test statistic converges to a Chi-squared random variable, and conventional critical values can be used.[4]

2.3.2 Z Test Asymptotics under the Alternative Hypothesis

To analyse the power of the Z test under the alternative, we make the following additional assumptions on the break:

Assumption 11. *Z is a non-singular matrix, and* $Z\Sigma_F Z^\intercal \neq \Sigma_F$.

Assumption 12. $plim_{T\to\infty} \inf \left(vech(C)^\intercal \left[\max(b_{\lfloor \pi T \rfloor}, b_{T-\lfloor \pi T \rfloor}) \widehat{S}(F^* H_{0,1})^{-1} \right] vech(C) \right) > 0$, *where* $C \equiv H_{0,1}^\intercal (\Sigma_F - Z\Sigma_F Z^\intercal) H_{0,1}$.

Assumption 11 ensures that the test statistic diverges under the alternative hypothesis. It rules out the unlikely scenario where $Z = -1$, i.e. all of the loadings switch their signs after the break, and is commonly assumed (see Han and Inoue (2015), Baltagi et al.

[4] It is also possible to construct an LM-like statistic with a restricted estimate of the variance using all of the data. However, as noted by Chen et al. (2014) and Han and Inoue (2015), such LM-like statistics have much smaller power than their Wald-type counterparts. Therefore, we focus on the Wald test.

(2017), Baltagi et al. (2021), and others). We require Z to be non-singular, which implicitly rules out the case of disappearing/emerging factors. This is because unlike Han and Inoue (2015), Baltagi et al. (2021), and Duan et al. (2022) who work with the *pseudo* factors as estimated over the entire sample, we work with the subsample estimates of the factors and the appropriate definition of rotation matrices is not clear in such cases (other methods using subsample estimates similarly rule this out, see Chen et al. (2014), Massacci (2017), Ma and Su (2018), and others). Assumption 12 regulates the asymptotic property of the variance matrices of the statistics, ensures that $\mathscr{W}_Z(\pi, \widehat{F})$ diverges under the alternative.

Theorem 2.5. *Under Assumptions 1 to 9 and 12, and if Z satisfies Assumption 11, then*

1. *there exists some non-random matrix $C \neq 0$ such that*
   $\frac{1}{\pi T}\sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^{\intercal} - \frac{1}{T-\lfloor \pi T \rfloor}\sum_{t=\lfloor \pi T+1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^{\intercal} \xrightarrow{p} C,$
2. *the test statistic $\mathscr{W}_Z(\widehat{F})$ is consistent under the alternative hypothesis that $Z \neq I$.*

Theorem 2.5 shows that the subsample means of $\widehat{f}_t \widehat{f}_t^{\intercal}$ converge to different limits under the alternative, and thus result in a consistent test.

Remark 1. *The number of factors r is assumed to be known, and constant before and after the break due to Z being non-singular. Consistent estimation of the number of factors in each subsample is possible conditional on consistent estimate of π using any pre-existing estimator (e.g. Bai and Ng (2002), Onatski (2010), or Ahn and Horenstein (2013)) and π as shown in Baltagi et al. (2017). In practice, the estimate $\widehat{r}$ may not be same in either regime, and the practitioner will need to set the number of factors to be identical. Underestimation of r could omit information in rotational changes, whereas overestimation of r could result in extra noise brought about by the extra estimated factors, (Baltagi et al. (2017)). In practice, overestimation of r tends to lead to oversizing (see Section 3), so we advise a conservative estimate of r.*

Remark 2. *Similarly, the break fraction* π *needs to be estimated using a method such as Baltagi et al. (2017), Chen (2015), or Duan et al. (2022). Theorem 3 of Baltagi et al. (2017) shows that such consistent estimators of* π *are sufficient to obtain the usual* $O_p\left(\frac{1}{\delta_{NT}^2}\right)$ *consistency rate of the estimated factors and loadings, and therefore our test statistics remain valid.*

2.4 W test for Orthogonal Shifts

Next, we consider testing the null hypothesis $\mathcal{H}_0 : W_{N\times r} = \mathbf{0}$, against the alternative hypothesis $\mathcal{H}_1 : W_{N\times r} \neq \mathbf{0}$. Note that because W contains N rows and $N \to \infty$, traditional tests are infeasible. Similar to Han (2015), we re-state our null and alternative hypotheses for the type 1 break as:

$$\mathcal{H}_0 : r_w = 0, \quad \mathcal{H}_1 : r_w \neq 0, \tag{2.13}$$

where r_w is the number of extra factors augmented by the presence of orthogonal shifts. Although Equation (2.13) is essentially a problem for testing the number of factors, existing tests such as Onatski (2009) cannot be used without imposing further restrictive assumptions on the approximate factor model errors. Our strategy is to present an individual test for each i, then pool them across the cross section, thus overcoming the infinite dimensionality problem.

We define the Wald test statistic for orthogonal shifts in any individual series as:

$$\mathscr{W}_{W,i} = T\tilde{w}_i^{\intercal}\tilde{\Omega}_{W,i}^{-1}\tilde{w}_i, \tag{2.14}$$

where $\tilde{w}_i$ denotes the transpose of the ith row of $\tilde{W}$, and $\tilde{\Omega}_{W,i} = \frac{1}{1-\pi}\tilde{\Theta}_{1,i} + \frac{1}{\pi}\tilde{\Theta}_{2,i}$ is a HAC estimate of the asymptotic variance. The covariance matrices $\tilde{\Theta}_{1,i}$ and $\tilde{\Theta}_{2,i}$ are constructed using the estimated residuals $\tilde{e}_{(1),it} = x_{it} - \tilde{\lambda}_{2,i}^{\intercal}\tilde{f}_{1,t}$ and $\tilde{e}_{(2),it} = x_{it} - \tilde{\lambda}_{2,i}^{\intercal}\tilde{f}_{2,t}$ in the series $\tilde{Z}^{\intercal}\tilde{f}_{1,t} \cdot \tilde{e}_{(1),it}$ and $\tilde{f}_{2,t} \cdot \tilde{e}_{(2),it}$ respectively, and are detailed in the Supplementary Material. We define the joint Wald test statistic as:

$$\mathscr{W}_W = (TN)\left(\frac{\sum_{i=1}^{N}\tilde{w}_i}{N}\right)^{\intercal}\tilde{\Omega}_W^{-1}\left(\frac{\sum_{i=1}^{N}\tilde{w}_i}{N}\right), \tag{2.15}$$

where $\tilde{\Omega}_W = N^{-1}\sum_{i=1}^{N}\tilde{\Omega}_{W,i}$ is an estimate of the asymptotic pooled variance, detailed in the Supplementary Material.

2.4.1 W Test Asymptotics under the Null Hypothesis

To derive the properties of the test statistics, we make the following additional assumptions.

Assumption 13. *There exists a positive constant $M < \infty$ such that for all N, T, and $m = 1, 2$:*

(a) For each t, $E(N^{-1/2}\sum_{i=1}^{N} e_{it})^2 \leq M$.

Assumption 14. *There exists a positive constant $M < \infty$ such that for all N, T:*

(a) For each i and $m = 1, 2$, $E\left\|\frac{1}{\sqrt{NT_m}}\sum_{t=1}^{T}\sum_{k=1}^{N}\left(\lambda_{m,k}[e_{kt}e_{it} - E[e_{kt}e_{it}]]\right)\iota_{mt}\right\|^2 \leq M$,

(b) For each t and $m = 1, 2$, $E\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{m,i}e_{it}\right\|^2 \leq M$,

(c) For each i, $E\left\|\frac{1}{\sqrt{T_m}}\sum_{t=1}^{T} f_t e_{it}\iota_{mt}\right\|^4 \leq M$,

(d) $E\left\|\frac{1}{\sqrt{NT_m}}\sum_{t=1}^{T}\sum_{i=1}^{N} f_t e_{it}\iota_{mt}\right\|^2 \leq M$.

Assumption 15. *For $m = 1, 2$:*

(a) $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{m,i} = O_p(1)$,

(b) $E\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{m,i}e_{it}^2\right\|^2 \leq M$ for each t,

(c) $E\left\| \frac{1}{N\sqrt{T_m}} \sum_{t=1}^{T} \sum_{k\neq i} \sum_{i=1}^{N} \lambda_{m,k} e_{kt} e_{it} \cdot \iota_{mt} \right\|^2 \leq M$.

Assumption 16. *(a)* $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} f_t e_{it} \xrightarrow{d} N(0, \Phi_i)$, $(T)^{-1} \sum_{t=1}^{T} f_t f_t^{\intercal} e_{it}^2 \xrightarrow{p} \Phi_i$, *each* $\Phi_i > 0$.

(b) $\frac{1}{\sqrt{TN}} \sum_{t=1}^{T} \sum_{i=1}^{N} f_t e_{it} \xrightarrow{d} N(0, \Phi_W)$, $(TN)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{N} f_t f_t^{\intercal} e_{it}^2 \xrightarrow{p} \Phi_W$, $\Phi_W > 0$.

Assumption 13 is simply the pooled version of Assumption 3. Assumption 14 (a) is simply Assumption 6 (a) but corresponding to the loadings, Assumption 14 (b) is already implied by Assumption 6 (c), and Assumption 14 (c) is a strengthened version of Assumption 6 (a). These correspond to Assumptions 6 b), 6 d) and 6 e) in Han (2015), and are not restrictive because they involve zero mean random variables. Assumption 15 requires that the sum of factor loadings is $O_p\left(\sqrt{N}\right)$, and is a slightly modified version of the assumption initially considered by Han (2015). As explained by Han (2015), this will hold if the loadings are centered around zero, such that the sum of the loadings diverge at the rate of $\sqrt{N}$ by the central limit theorem. Although this imposes somewhat stricter restrictions compared to a conventional factor model setup, it seems to hold for empirically used datasets, as noted by Han (2015). Assumptions 16 (a) and 16 (b) are simply central limit theorems. The latter assumption somewhat strengthens the restriction on the cross sectional correlation in e_{it}, and is simply the cross sectional averaged version of the CLT assumptions introduced by Bai (2003).

Theorem 2.6. *If* $\frac{\sqrt{T}}{N} \to 0$, *then:*

1. *Under Assumptions 1 to 9, and additionally Assumptions 13, 14 and 16,* $\mathscr{W}_{W,i} \xrightarrow{d} \chi_r^2$ *for each* i, *and*

2. *Under Assumptions 1 to 9, and additionally Assumptions 13 to 16,* $\mathscr{W}_W \xrightarrow{d} \chi_r^2$.

Theorem 2.6 shows that the Wald[5] test statistics converge to conventional Chi-squared

[5]It is also possible to construct an LM-like test statistic by imposing the null hypothesis of no break, but this results in a statistic with lower power, so we focus on the Wald test again.

random variables. The detailed proof of Theorem 2.6 is provided in the Supplementary Material. The basic idea of the proof is to recognize that $\tilde{w}_i$ is a suitable weighted average of $\tilde{Z}^\intercal\tilde{\lambda}_{1,i}$, and $\tilde{\lambda}_{2,i}$, both of which have asymptotic normal expansions following Theorem 2 of Bai (2003). The asymptotic normality of $\tilde{w}_i$ and $N^{-1}\sum_{i=1}^{N}\tilde{w}_i$ follows, implying the form of the Wald tests under the null hypothesis.

2.4.2 W test Asymptotics under the Alternative Hypothesis

To analyze the behavior of the pooled W test under the alternative hypothesis, we introduce some further assumptions.

Assumption 17. *There exist constants* $0 < \alpha \leq 0.5$ *and* $C > 0$ *such that as* $N, T \to \infty$, $Pr\left(\left\|\frac{T^{\alpha/2}}{\sqrt{N}}\sum_{i=1}^{N} w_i\right\| > C\right) \to 1$.

Assumption 17 requires $\left\|\frac{T^{\alpha/2}}{\sqrt{N}}\sum_{i=1}^{N} w_i\right\|$ to be bounded away from zero asymptotically. Note that if $N^{-1}\sum_{i=1}^{N} w_i \xrightarrow{p} 0$ under the alternative, then $N^{-1/2}\sum_{i=1}^{N} w_i$ converges in distribution to some Gaussian random variable by the Law of Large Numbers, and hence $\left\|N^{-1/2+\epsilon}\sum_{i=1}^{N} w_i\right\|$ is diverging as $N \to \infty$ for any positive ϵ. In order for $\left\|\frac{T^{\alpha/2}}{\sqrt{N}}\sum_{i=1}^{N} w_i\right\|$ to be bounded away from zero, any $\alpha \in (0, 0.5]$ such that $T^{\alpha/2} \geq N$ is required, which is not difficult. Assumption 17 therefore ensures that the joint test statistic diverges to infinity under the alternative hypothesis, even when if $N^{-1}\sum_{i=1}^{N} w_i \xrightarrow{p} 0$.

Theorem 2.7. *Suppose that* $\frac{\sqrt{T}}{N} \to 0$, *and the alternative hypothesis* $\mathcal{H}_1 : r_w \neq 0$ *holds. Then:*

(a) under Assumptions 1 to 8, 13, 14 and 16, and if $w_i \neq 0$, *then* $\mathscr{W}_{W,i} \to \infty$ *as* $N, T \to \infty$,

(b) under Assumptions 1 to 8 and 13 to 17, $\mathscr{W}_W \to \infty$ *if* $\frac{\sqrt{N}}{T^{1-\alpha/2}} \to 0$ *as* $N, T \to \infty$.

Theorem 2.7 shows that both the individual test $\mathscr{W}_{W,i}$ and joint test $\mathscr{W}_W$ diverge to infinity asymptotically under the alternative, and are thus consistent tests.

3 Monte Carlo Simulations

3.1 Simulation Specification

We first simulate two sets of arbitrary loadings, Λ_1, Λ_2 both of which are distributed as a multivariate $N(\mathbf{0}_3, I_3)$, focusing on the case of $r = 3$ factors. Then, we set W to be the residuals of the projection $\Lambda_2 - (\Lambda_1^\intercal \Lambda_1)^{-1}\Lambda_1^\intercal \Lambda_2$ to ensure that it is orthogonal to Λ_1. The rotation break Z is set to the identity matrix in the case of no break, or a lower triangular matrix with $[2.5, 1.5, 0.5]$ on the main diagonal and its lower triangular entries drawn from $N(0, 1)$, as in Duan et al. (2022). The overarching model from we simulate can then be formulated as below.

$$x_{it} = \begin{cases} \lambda_{1,i}^\intercal f_t + \sqrt{\theta} e_{it}, & t = 1, \dots, \lfloor \pi T \rfloor \\ (Z\lambda_{1,i} + \omega w_i)^\intercal f_t + \sqrt{\theta} e_{it}, & t = \lfloor \pi T \rfloor + 1, \dots, T, \end{cases} \tag{3.1}$$

for $i = 1, \dots, N$ and $t = 1, \dots, T$. The parameter θ is set to 3 in order to calibrate the signal to noise ratio to be 50%, and the scalar ω controls the "size" of the orthogonal shifts.

The factors and errors are generated as follows:

$$f_{k,t} = \rho f_{k,t-1} + \mu_{it}, \mu_{it} \sim i.i.d. N(0, 1 - \rho^2), \tag{3.2}$$

$$e_{it} = \alpha e_{i,t-1} + v_{it}, \tag{3.3}$$

where $\rho \in \{0, 0.7\}$ captures the serial correlation in the factors, and μ_{it}, v_{it} are mutually independent with $v_t = (v_{1,t}, \dots, v_{N,t})^\intercal$ being i.i.d. $N(0, \Omega)$ for $t = 1, \dots, T$. For $t = 1$,

Table 1: Size of Rotation and Orthogonal Shift Tests, $N = 200$

				Z Test		W Test		W Individual
T	ρ	α	β	Unadj.	Adj.	Unadj.	Adj.	
200	0.0	0.0	0.0	0.108	0.072	0.005	0.003	0.013
		0.3	0.3	0.115	0.076	0.088	0.053	0.014
500		0.0	0.0	0.068	0.031	0.003	0.001	0.004
		0.3	0.3	0.070	0.037	0.061	0.027	0.004
200	0.7	0.0	0.0	0.243	0.160	0.003	0.002	0.017
		0.3	0.3	0.239	0.174	0.096	0.066	0.027
500		0.0	0.0	0.151	0.093	0.000	0.000	0.005
		0.3	0.3	0.144	0.095	0.074	0.047	0.009

$e_{.t} = (e_{1,1}, \ldots, e_{N,1})^\intercal$ is $N(0, \frac{1}{1-\alpha^2}\Omega)$ to initialize the errors at their stationary distributions. The scalar α captures the serial correlation in the errors, and as in Bates et al. (2013) and Baltagi et al. (2017), $\Omega_{ij} = \beta^{|i-j|}$ captures the cross sectional correlation in the errors. We consider $\alpha \in \{0, 0.3\}$ and $\beta \in \{0, 0.3\}$ to consider up to mild serial and cross sectional correlation. The true break fraction is set to 0.5 and treated as known.

Disentanglement necessitates the practitioner running both the Z and W tests, which could lead to a higher family wise error rate, and to this end we report the unadjusted p values, in addition to the adjusted p values using Holm (1979).

3.2 Simulation Results

We present the size analysis in Table 1. In the case of no serial correlation in the factors, and large T relative to N, the Z test has a nominal size close to the desired 5%, and this tends to hold regardless of the serial or cross sectional correlation in the errors. The Z test seems to be oversized when there is serial correlation in the factors, but this issue is alleviated and approaches a rejection rate of 0.15 as T increases.[6] The W test does not seem to be affected by serial correlation in the factors, and also seems to be overly conservative when there is no serial correlation in the error, but otherwise seems to have good size. Implementation of the Bonferroni-Holm procedure to adjust the p values also

[6]Increasing T further does seem to make the size approach 5% (see Table 1 in Supplementary Material).

Table 2: Power of Z and W Tests, $r = 3$, $\alpha = \beta = 0.3$

					Z Test		W Test					
Type	T	N	ω	ρ	Unadj.	Adj.	Unadj.	Adj.	Individual	HI	BKW	$\tilde{r}$
Type 1	200	200	1	0.0	0.136	0.129	0.860	0.821	0.849	1.000	1.000	5.928
				0.7	0.244	0.233	0.916	0.896	0.908	1.000	1.000	6.000
	500			0.0	0.079	0.076	0.950	0.939	0.947	1.000	1.000	6.000
				0.7	0.146	0.144	0.968	0.965	0.968	1.000	1.000	6.000
Type 2	200		0	0.0	1.000	1.000	0.100	0.100	0.026	1.000	1.000	3.000
				0.7	1.000	1.000	0.106	0.106	0.035	1.000	1.000	3.000
	500			0.0	1.000	1.000	0.094	0.094	0.009	1.000	1.000	3.000
				0.7	1.000	1.000	0.096	0.096	0.012	1.000	1.000	3.000
Type 3	200		1	0.0	1.000	1.000	0.804	0.803	0.765	1.000	1.000	4.206
				0.7	1.000	1.000	0.867	0.867	0.846	1.000	1.000	5.047
	500			0.0	1.000	1.000	0.919	0.919	0.901	1.000	1.000	4.772
				0.7	1.000	1.000	0.946	0.946	0.938	1.000	1.000	5.511

seems to correct the oversizing issue, so we advocate for its use. Table 2 presents the power of the Z and W tests across all types of breaks. It can be seen that both the Z and W test have good power and are rejecting correctly only on their respective break types. This is in contrast to HI and BKW tests, which consistently reject across all break types, and are thus unable to discern which type of break has occurred.

4 Empirical Application

4.1 Data and Methodology

For our empirical application, we apply our tests to the FRED-MD dataset (see McCracken and Ng (2015) for data cleaning and preparation). We focus on two candidate break dates: 1984 February, corresponding to the Great Moderation (Baltagi et al. (2021), Ma and Su (2018) and Breitung and Eickmeier (2011)); and 2008 November, corresponding to the Global Financial Crisis (Baltagi et al. (2021), Ma and Su (2018) and Duan et al. (2022)).

Our tests aim to *differentiate* the type of break once they have been estimated, and were formulated under the assumption that there is only one break. As argued by Bai (1997), Bai and Perron (1998) and others, tests formulated for the case of one break can

Table 3: Subsample $\tilde{r}$ Estimates

Sample	Onatski (2010)	Ahn and Horenstein (2013) Eigenvalue Ratio
Great Moderation (1984 February) Sample		
Whole	6	1
Pre-break	4	1
Post-break	3	3
Global Financial Crisis (2008 November) Sample		
Whole	4	1
Pre-break	2	1
Post-break	2	1

be expected to have power against multiple breaks. To this end, we consider a sample of 1975 January to 2000 January for the Great Moderation (GM) break, and a sample of 2003 January to 2013 January for the Global Financial Crisis (GFC) break, in order to ensure that there is only one break in each sample. These samples were chosen because there is mixed evidence that there could be a break in the mid to late 1990s (Hansen (2001) and Ma and Tu (2022)), or in 2000 associated with the early 2000s recession (Ma and Su (2018) and Ma and Tu (2022)). This is not restrictive, because the case of multiple breaks can be dealt with partitioning the data, and running the tests on each break separately.

The number of factors in each subsample is estimated by the eigenvalue edge distribution estimator of Onatski (2010), and the eigenvalue ratio estimator of Ahn and Horenstein (2013).[7] As seen in Table 3, these estimators do not typically agree with one another on empirical data, and we therefore report the results of the tests for $\tilde{r} = 1, \dots 4$, which is the maximum subsample r estimated.

4.2 Joint Test Results

Table 4 reports the results of the Z and pooled W tests when the Great Moderation and the Global Financial Crisis are candidate break dates. For the Great Moderation, a higher

[7] We also consider using the information criteria IC_{p1} and IC_{p2} of Bai and Ng (2002), but do not report them here as they are well known to overestimate the number of factors. Ahn and Horenstein (2013)'s eigenvalue growth ratio estimator also tends to produce similar results to their eigenvalue ratio and is hence omitted. For more comprehensive results, see Table 7 in Section 3.2.

Table 4: Joint Test Results

	Z Test p values		W Test p values			
$\tilde{r}$	Unadj.	Adj.	Unadj.	Adj.	Han and Inoue (2015)	Baltagi et al. (2021)
Great Moderation (1984 February) Sample						
1	0.000	0.001	0.596	0.596	0.001	0.000
2	0.000	0.000	0.000	0.000	0.001	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.001
Global Financial Crisis (2008 November) Sample						
1	0.175	0.175	0.005	0.011	0.688	0.116
2	0.009	0.019	0.486	0.486	0.354	0.116
3	0.010	0.010	0.002	0.005	0.009	0.004
4	0.021	0.021	0.000	0.000	0.000	0.007

r leads to significant rejection on both tests. However, for the case of only one factor, we fail to reject the null of orthogonal shifts. This is in stark contrast to the tests of Han and Inoue (2015) and Baltagi et al. (2021), which strongly reject for all values of r. In contrast, for the Global Financial Crisis, we see mixed evidence: in the case of one factor, we only reject the W test; in the case of two factors, we only reject the Z test; and we reject both tests when the number of factors is three or more. At present, it is unclear why this mixed evidence occurs, but given that the tests of Han and Inoue (2015) and Baltagi et al. (2021) fail to reject for the case of one to two factors, we therefore conclude that the Global Financial Crisis is best characterized as a type 3 break for the case of three factors, and would lead to a factor augmentation effect if ignored. This is most clearly seen with Onatski (2010)'s estimator, which estimates exactly double the number of factors over the whole sample, compared to each subsample.

4.3 Individual Test Results

In order to aid economic interpretation in the precise nature of the breaks, we also report the results for the individual w_i test results. For comparative purposes, we also report the individual loading break tests of Breitung and Eickmeier (2011) (BE) using the same candidate break dates. The rejection frequencies are visualized in Figure 1. For the Great

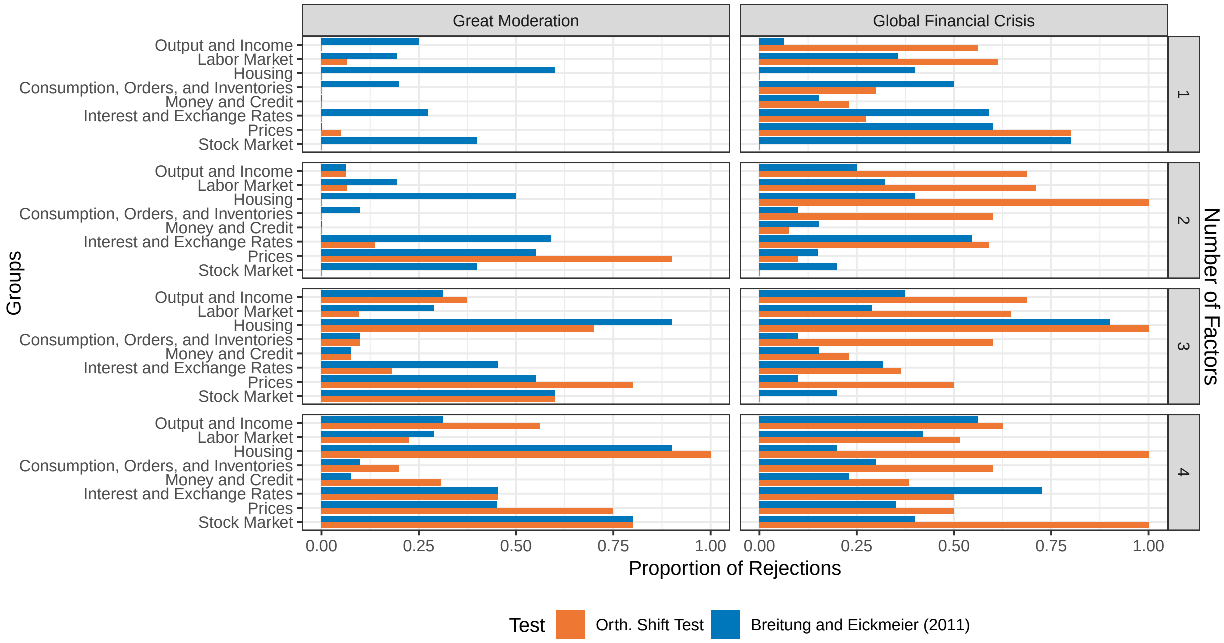

Figure 1: Individual Loading Rejection Proportions for Great Moderation Sample (1975 to 2000, break date of 1984 February) and Global Financial Crisis Samples (2003 to 2013, break date of 2008 November)

Moderation, the w_i test rejects far less often than BE's tests for the case of 1-2 factors. An interpretation of this result is that our w_i test controls for possible changes in the factor variance, as opposed to BE, whose test statistics are constructed used the *pseudo* factors and therefore do not control for the possibility of the factor variance breaking.

Further inspection of which *specific* series are breaking in Figure 1 reveals that for the case of two factors, the statistically significant joint break is actually mostly isolated to variables in the Prices group. For the case of three factors, the W test rejects a much smaller fraction of Interest and Exchange Rate variables compared to BE. The precise implications of this are beyond the scope of this paper, but this suggests that if the practitioner is not concerned with price series, the augmentation effect could be safely ignored.

This is in contrast to the results when the GFC is used as a candidate break. Instead, we are in general able to reject a higher proportion of series than BE. Although BE did not consider the GFC as a common break date, it is interesting to see that the use of pseudo factors seems to be confounding and reducing the power in detecting legitimate breaks in the loadings. A specific look into *which* specific variables are breaking reveals some differences across groups, for the case of one factor. Our orthogonal shift test fails to reject any housing or stock market variables, compared to BE, which report a rejection of a significant fraction in both of these groups.

4.4 Variance Decomposition

Our projection based equivalent representation theorem provides a natural framework to decompose structural breaks and quantify the proportion of variance change due to a change in factor variance and the change in factor loadings. We relegate the technical details of this to the Supplementary Material, and report the various restricted and unrestricted R^2 values in Table 5.[8] The results for Great Moderation match up with the results of the formal joint tests - for the case of $r = 1$ factor, restricting $Z = I$ results in a decrease of in sample R^2 from 17.8% to 12.6%, compared to the the restriction of $W = 0$, which only decreases in sample R^2 to 16.9%. The results for the Global Financial Crisis at first glance seem contradictory to the results of the joint test - there appears to be negligible decreases in the in sample R^2 from imposing $Z = I$. However, this is because the nature of rotational change during the Global Financial Crisis is that the ordering of the factors has changed.[9] Due to the limitation of our variance decomposition methodology in controlling for this, we interpret this this as meaning that the statistical evidence from the joint test was simply picking up on this "re-ordering" of the factors. The restriction of $W = 0$ results in large decreases R^2 and in consistent with the results of the joint tests.

[8] For results with higher r, see Table 10 in Section 3.2.

[9] See Section 4.1 in Supplementary Material.

Table 5: R^2 Comparisons for $Z = I, W = 0$, $Z = I$ and $W = 0$ restrictions. R^2 values for Whole Sample PCA represent the fit from the *pseudo* factors and loadings and are thus not directly comparable.

		Restricted R^2			
$\tilde{r}$	Unrestricted R^2	$Z = I$	$W = 0$	$Z = I, W = 0$	Whole Sample PCA
Great Moderation (1984 February) Sample					
1	0.178	0.126	0.169	0.117	0.172
2	0.274	0.221	0.225	0.173	0.241
3	0.344	0.289	0.277	0.222	0.302
4	0.398	0.342	0.313	0.257	0.359
Global Financial Crisis (2008 November) Sample					
1	0.228	0.228	0.141	0.140	0.182
2	0.342	0.341	0.223	0.222	0.291
3	0.424	0.422	0.309	0.307	0.370
4	0.489	0.487	0.372	0.371	0.434

5 Conclusion

We propose a *projection based equivalent representation theorem* to decompose any structural break in dynamic factors into a rotational change and orthogonal shift. By interpreting these two changes as a break in factor variance and a break in factor loadings respectively, we are able to subsequently propose two separate tests: 1) a test for evidence of rotational change, and 2) a test for evidence of orthogonal shifts. Monte Carlo studies demonstrate their good finite sample performance, as well as the inability of existing methods to differentiate between these different break types. We apply the tests to the FRED-MD dataset using the Great Moderation and Global Financial Crisis as candidate break dates, and find evidence that the Great Moderation may be better characterised as a break in the factor variance, as opposed to a break in the loadings, whereas the Global Financial Crisis is a break in both. Our results highlight the limitations of existing methods in differentiating between these break types and nuance the discussion surrounding structural breaks in dynamic factor models.

Our framework provides a potential foundation to explore the precise practical and theoretical implications of structural breaks in dynamic factor models. For example, a natural question to consider is how the different break types can affect the estimation and subsequent use of factors, such as the factor augmented forecasts of Stock and Watson (2002) and Bai and Ng (2006), and factor augmented vector auto-regressions of Bernanke et al. (2005). Indeed, although there have been many suggestions for how to use factors in forecasting when a structural break is present (see Stock and Watson (2009) and Baltagi et al. (2021)), there is still no formal treatment of this in the literature.

SUPPLEMENTARY MATERIAL

Appendices Appendices containing proofs, and additional simulation and empirical results (PDF)

R Code R Code including FREDMD vintage available on request (.zip file)

References

Ahn, S. C., & Horenstein, A. R. (2013). Eigenvalue Ratio Test for the Number of Factors. *Econometrica*, *81*(3), 1203–1227. https://doi.org/10.3982/ECTA8968

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change With Unknown Change Point. *Econometrica*, *61*(4), 821–856. https://doi.org/10.2307/2951764

Bai, J. (1997). Estimation of a Change Point in Multiple Regression Models. *Review of Economics and Statistics*, *79*(4), 551–563. https://doi.org/10.1162/003465397557132

Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica*, *71*(1), 135–171. https://doi.org/10.1111/1468-0262.00392

Bai, J., Duan, J., & Han, X. (2022, June 16). Likelihood ratio test for structural changes in factor models. https://doi.org/10.48550/arXiv.2206.08052

Bai, J., Han, X., & Shi, Y. (2020). Estimation and inference of change points in high-dimensional factor models. *Journal of Econometrics*, *219*(1), 66–100. https://doi.org/10.1016/j.jeconom.2019.08.013

Bai, J., & Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, *70*(1), 191–221. https://doi.org/10.1111/1468-0262.00273

Bai, J., & Ng, S. (2006). Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions. *Econometrica*, *74*(4), 1133–1150. https://doi.org/10.1111/j.1468-0262.2006.00696.x

Bai, J., & Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, *66*(1), 47. https://doi.org/10.2307/2998540

Baltagi, B. H., Kao, C., & Wang, F. (2017). Identification and estimation of a large factor model with structural instability. *Journal of Econometrics*, *197*(1), 87–100. https://doi.org/10.1016/j.jeconom.2016.10.007

Baltagi, B. H., Kao, C., & Wang, F. (2021). Estimating and testing high dimensional factor models with multiple structural changes. *Journal of Econometrics*, *220*(2), 349–365. https://doi.org/10.1016/j.jeconom.2020.04.005

Barigozzi, M., & Trapani, L. (2020). Sequential testing for structural stability in approximate factor models. *Stochastic Processes and their Applications*, *130*(8), 5149–5187. https://doi.org/10.1016/j.spa.2020.03.003

Bates, B. J., Plagborg-Møller, M., Stock, J. H., & Watson, M. W. (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, *177*(2), 289–304. https://doi.org/10.1016/j.jeconom.2013.04.014

Bernanke, B. S., Boivin, J., & Eliasz, P. (2005). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach*. *The Quarterly Journal of Economics*, *120*(1), 387–422. https://doi.org/10.1162/0033553053327452

Breitung, J., & Eickmeier, S. (2011). Testing for structural breaks in dynamic factor models. *Journal of Econometrics*, *163*(1), 71–84. https://doi.org/10.1016/j.jeconom.2010.11.008

Chen, L. (2015). Estimating the common break date in large factor models. *Economics Letters*, *131*, 70–74. https://doi.org/10.1016/j.econlet.2015.03.037

Chen, L., Dolado, J. J., & Gonzalo, J. (2014). Detecting big structural breaks in large factor models. *Journal of Econometrics*, *180*(1), 30–48. https://doi.org/10.1016/j.jeconom.2014.01.006

Duan, J., Bai, J., & Han, X. (2022). Quasi-maximum likelihood estimation of break point in high-dimensional factor models. *Journal of Econometrics*, S0304407622000379. https://doi.org/10.1016/j.jeconom.2021.12.011

Han, X. (2015). Tests for overidentifying restrictions in Factor-Augmented VAR models. *Journal of Econometrics*, *184*(2), 394–419. https://doi.org/10.1016/j.jeconom.2014.04.024

Han, X., & Inoue, A. (2015). TESTS FOR PARAMETER INSTABILITY IN DYNAMIC FACTOR MODELS. *Econometric Theory*, *31*(5), 1117–1152. https://doi.org/10.1017/S0266466614000486

Hansen, B. E. (2001). The New Econometrics of Structural Change: Dating Breaks in U.S. Labour Productivity. *Journal of Economic Perspectives*, *15*(4), 117–128. https://doi.org/10.1257/jep.15.4.117

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Ma, C., & Tu, Y. (2022). Group fused Lasso for large factor models with multiple structural breaks. *Journal of Econometrics*, S0304407622000331. https://doi.org/10.1016/j.jeconom.2022.02.003

Ma, S., & Su, L. (2018). Estimation of large dimensional factor models with an unknown number of breaks. *Journal of Econometrics*, *207*(1), 1–29. https://doi.org/10.1016/j.jeconom.2018.06.019

Massacci, D. (2017). Least squares estimation of large dimensional threshold factor models. *Journal of Econometrics*, *197*(1), 101–129. https://doi.org/10.1016/j.jeconom.2016.11.001

Massacci, D. (2021). Testing for Regime Changes in Portfolios with a Large Number of Assets: A Robust Approach to Factor Heteroskedasticity. *Journal of Financial Econometrics*, nbaa046. https://doi.org/10.1093/jjfinec/nbaa046

McCracken, M. W., & Ng, S. (2015). *FRED-MD: A Monthly Database for Macroeconomic Research*. https://doi.org/10.20955/wp.2015.012

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703–708. https://doi.org/10.2307/1913610

Onatski, A. (2009). Testing Hypotheses About the Number of Factors in Large Factor Models. *Econometrica*, *77*(5), 1447–1479. https://doi.org/10.3982/ECTA6964

Onatski, A. (2010). Determining the Number of Factors from Empirical Distribution of Eigenvalues. *The Review of Economics and Statistics*, *92*(4), 1004–1016. https://doi.org/10.1162/REST_a_00043

Pelger, M., & Xiong, R. (2022). State-Varying Factor Models of Large Dimensions. *Journal of Business & Economic Statistics*, *40*(3), 1315–1333. https://doi.org/10.1080/07350015.2021.1927744

Stock, J. H., & Watson, M. W. (1998, August). *Diffusion Indexes* (Working Paper No. 6702). National Bureau of Economic Research. https://doi.org/10.3386/w6702

Stock, J. H., & Watson, M. W. (2002). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179. https://doi.org/10.1198/016214502388618960

Stock, J. H., & Watson, M. W. (2009). Forecasting in Dynamic Factor Models Subject to Structural Instability, 65.

Stock, J. H., & Watson, M. W. (2016, January 1). Chapter 8 - Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of Macroeconomics* (pp. 415–525). Elsevier. https://doi.org/10.1016/bs.hesmac.2016.04.002

Su, L., & Wang, X. (2017). On time-varying factor models: Estimation and testing. *Journal of Econometrics*, *198*(1), 84–101. https://doi.org/10.1016/j.jeconom.2016.12.004

Su, L., & Wang, X. (2020). TESTING FOR STRUCTURAL CHANGES IN FACTOR MODELS VIA A NONPARAMETRIC REGRESSION. *Econometric Theory*, *36*(6), 1127–1158. https://doi.org/10.1017/S0266466619000446

Wang, Y. (2021, October 8). *Identification and Estimation of Parameter Instability in a High Dimensional Approximate Factor Model* [Unpublished Manuscript].