Disentangling Structural Breaks in Factor Models for Macroeconomic Data*

Bonsoo Koo[†], Benjamin Wong[†], and Ze-Yu Zhong[†‡]

Monash University

Dated: May 2024

Abstract

Through a routine normalization of the factor variance, standard methods for estimating factor models in macroeconomics do not distinguish between breaks of the factor variance and factor loadings. We argue that it is important to distinguish between structural breaks in the factor variance and loadings within factor models commonly employed in macroeconomics as both can lead to markedly different interpretations when viewed via the lens of the underlying dynamic factor model. We then develop a projection-based decomposition that leads to two standard and easy-to-implement Wald tests to disentangle structural breaks in the factor variance and factor loadings. Applying our procedure to U.S. macroeconomic data, we find evidence of both types of breaks associated with the Great Moderation and the Great Recession. Through our projection-based decomposition, we estimate that the Great Moderation is associated with an over 60% reduction in the total factor variance, highlighting the relevance of disentangling breaks in the factor structure.

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 DP210101440, DP240100970, and DE200100693. We thank Xu Han and Fa Wang for making their code available, in addition to the comments by Matteo Barigozzi, Xu Han, Daniele Massacci, Serena Ng, Mirco Rubin, and other participants at the 2022 EC-squared, 2023 Society for Financial Econometrics (SoFiE), 2023 International Symposium on Econometric Theory and Applications (SETA) conferences, and 2023 Chinese University of Hong Kong Workshop on Econometrics.

†Department of Econometrics and Business Statistics, Monash University, Melbourne, Australia.

‡Corresponding author. Email: ze.zhong@monash.edu

1 Introduction

Dynamic factor models are increasingly used in empirical macroeconomics and finance (e.g. Aastveit et al., 2015; Alessi and Kerssenfischer, 2019; Barigozzi and Luciani, 2023) as a form of dimension reduction, and summarize the dynamics of a large set of time series through a small number of factors. There is increasing evidence of structural instability with U.S. macroeconomic time series (see, e.g. Breitung and Eickmeier, 2011; Chen et al., 2014; Stock and Watson, 2016). The analysis of structural changes in factor models presents a unique challenge because breaks in the loadings and breaks in the factor variance cannot be easily disentangled. For simplicity, consider the following representative factor model common in applied macroeconomic work for $x_{it}, t = 1, \ldots, T, i = 1, \ldots, N$:

$$x_{it} = \lambda_i^\intercal f_t + e_{it}, \tag{1}$$

where λ_i is an $r \times 1$ vector of individual loadings, f_t is an $r \times 1$ vector of factors, and e_{it} is noise. Because both the factors and loadings are unobserved and enter multiplicatively, a normalization is needed to separately identify them; often this is done on the variance of f_t.[1] While such normalizations are often innocuous, they matter for studying structural changes in factor models; if changes in the factor variance are ruled out through the normalization, these changes must manifest as changes in the factor loadings even if the loadings are stable. We note similar concerns have previously been raised (see Stock and Watson, 2016).[2]

Our contribution is a projection-based decomposition to disentangle structural breaks[3] in the factor variance and the loadings in dynamic factor models. At a high level, the projection-based decomposition reparameterizes structural change in the factor structure

[1]Other normalizations exist, (see Bai and Wang, 2016), but still serve the purpose of pinning down the scale of one quantity (e.g. the factor variance or the loadings) in order to identify the other.

[2]We argue such concerns are empirically relevant because the literature has typically identified periods such as the Great Moderation, the Great Recession and more recently the COVID-19 Pandemic as evidence of structural breaks, all periods well known for the data displaying heteroscedasticity (e.g. Breitung and Eickmeier, 2011; Baltagi et al., 2021; Bai et al., 2022).

[3]We use the term "structural break" to refer to a discrete change along the time dimension.

into a rotational and orthogonal component where each has a natural interpretation as a change in the factor covariance matrix and a change in the factor loadings respectively. This interpretation of the orthogonal component arises from recognizing that breaks in the loadings are orthogonal to the original factor space, and can therefore result in more factors being estimated over the whole sample if ignored. At the same time, rotations can be thought of as some suitable twisting or stretching of the factor space which do not result in more factors appearing if ignored, and so are associated with the factor variance.[4] Once one recognizes this insight, the reparameterization naturally leads to two easy-to-implement structural break tests: (i) a test for a break in the factor covariance matrix, and (ii) a test for a break in the factor loadings. We show that these test statistics have standard chi-squared distributions and reasonable finite sample performance.

Disentangling breaks in the factor variance and loadings is not a mere technical curiosity. From the perspective of the dynamic factor model, these breaks imply vastly different interpretations. While breaks in the loadings relate to changes in how variables relate to the factors, breaks in the factor variance imply breaks in the factor dynamics. These breaks in the factor dynamics can be in the form of a break in the dynamic process generating the factors and/or breaks in the variance of the underlying shocks to the factors. Confining breaks to just the loadings would therefore *a priori* preclude breaks in the factor dynamics as a possible interpretation. While the degree of how misleading such an interpretation is depends on context, existing applied work suggests breaks in the factor dynamics presents a more natural interpretation of at least one historical episode: the Great Moderation. Indeed, in an empirical application with U.S. macroeconomic data, our tests detect the Great Moderation as a break in the factor covariance matrix, where through our projection-based decomposition, we estimate an over 60% reduction in the total variance of the factors.[5]

[4] We also briefly touch on how the case of disappearing/emerging factors can be accommodated by a singular rotation.

[5] Henceforth, we refer to a break in the factor variance as a break in the factor covariance matrix; if we refer to a change in the "total factor variance" (i.e. the trace of the factor covariance matrix), we will make this explicit.

While this finding should be unsurprising to applied macroeconomists, it nevertheless is an effective proof-of-concept underpinning our basic argument: only by disentangling breaks in factor variances and loadings can one attribute a break in the factor variance as part of the most well-known change in volatility common across multiple macroeconomic time series. Although we still find evidence of breaks in loadings even when controlling for breaks in the factor variance, our results complement and reconcile with broader factor model work that associate large breaks in factor loadings with events such as the Great Moderation (e.g. Baltagi et al., 2021), suggesting a more nuanced interpretation of breaks in factor models.

While there are different methods to estimate dynamic factor models and forms of structural instability, our work is most closely related to work which tests for structural breaks in factor models using the principal components estimator (see Stock and Watson, 2009; Breitung and Eickmeier, 2011; Chen et al., 2014; Han and Inoue, 2015; Baltagi et al., 2017). Given breaks in either the factor variance and/or loadings will manifest as breaks in the loadings in these tests, one could first test for breaks in the factor structure using one of these aforementioned procedures to date the break date, and subsequently use our procedure to disentangle the break.

We proceed as follows. In Section 2, we first motivate how breaks in the factor variance and loadings manifest through an underlying dynamic factor model before introducing our projection-based decomposition. Section 3 presents the theory underpinning our tests followed by Monte Carlo simulations in Section 4. Section 5, presents an empirical application with 124 quarterly U.S. macroeconomic time series. Section 6 concludes.

2 Interpreting and Reparameterizing Structural Breaks in Dynamic Factor Models

To start, consider the representative dynamic factor model (see Stock and Watson, 2016)

$$X_t = \Lambda f_t + e_t \tag{2}$$

$$f_t = \sum_{j=1}^{p} \Phi_j f_{t-j} + \eta_t, \quad \eta_t \sim (0, \Sigma_\eta), \tag{3}$$

where Equation (2) stacks Equation (1) across the cross section such that X_t and e_t are $N \times 1$, $\Lambda = [\lambda_1, \ldots, \lambda_N]^\intercal$ is $N \times r$. Equation (3) describes the dynamics of the factors, where Φ_j are autoregressive coefficients for f_t, and η_t are $q \times 1$ (reduced form) innovations with covariance Σ_η.[6] Equations (2) and (3) describe the static form of the dynamic factor model, and clarify how changes in the factor variance and loadings imply vastly different interpretations of the dynamic factor model. In particular, the (unconditional) covariance matrix of f_t is both a function of the Φ_j's and Σ_η. Therefore, breaks in the factor variance require a break in Equation (3), either in the Φ_j's, Σ_η, or both. Breaks in the λ_i's, on the other hand, are isolated to breaks in the relationship between the different variables and the factors in Equation (2). As an illustrative example, consider the Great Moderation, an event marked by a reduction of volatility across many macroeconomic variables. Such an interpretation, from the perspective of the dynamic factor model, is more naturally accommodated as a reduction in the total variance (i.e. trace of covariance matrix) of f_t rather than multiple (proportional) breaks in the λ_i's.

Because the factors are identified up to rotation, one needs to impose some form of normalization. A common approach is to use the principal components estimator of the

[6] Relative to Stock and Watson (2016), we focus on the case where the number of static and dynamic factors are equal, but this does not result ina loss of generality for the discussion.

factors $\tilde{F} = [\tilde{f}_1, \ldots, \tilde{f}_T]^\intercal$ and loadings $\tilde{\Lambda} = [\tilde{\lambda}_1, \ldots, \tilde{\lambda}_N]^\intercal$, which imposes

$$\frac{1}{N}\tilde{\Lambda}^\intercal\tilde{\Lambda} = V_{NT}, \quad \frac{1}{T}\tilde{F}^\intercal\tilde{F} = I_r, \tag{4}$$

where V_{NT} is a diagonal matrix whose entries are the first r eigenvalues of the covariance matrix of X. More generally, there exist many different estimation methods and thus normalizations, but our point still holds as long as one needs to impose a normalization for estimation. Of these, it is known that the method of principal components is able to consistently estimate the space spanned by the dynamic factors under very general conditions. Therefore, the principal components estimator cab be used to estimate both λ_i and f_t, and the fitting and specification of Equation (3) can occur as a separate step (see Stock and Watson, 2016, for more details). The normalization one applies for estimation convolutes the interpretation of breaks - structural break tests in the factor loadings often first estimate the factors on the full sample, then test for breaks in $\tilde{\lambda}_i$ (e.g. Stock and Watson, 2009; Breitung and Eickmeier, 2011; Chen et al., 2014; Bai and Han, 2016). The necessary application of normalizations like Equation (4) when estimating the factors thus makes it unclear whether finding a break in $\tilde{\lambda}_i$ is a break in λ_i, or a break in the variance of factors, since the latter is typically assumed to be unchanging as an identification condition as part of constructing the test statistic, (see Stock and Watson, 2016; Chen et al., 2014).

We note that if the goal is to establish whether there are breaks in the factor *structure* (i.e. the loadings or the factor variance), current methods of normalizing the variance, and then finding breaks in the loadings are probably appropriate, since any breaks in the factor variance are subsumed into the factor loadings. However, if one wanted to appropriately interpret breaks, especially from the perspective of the dynamic factor model implied by Equations (2) and (3), it becomes important to distinguish between breaks in the loadings and breaks in the factor variance. In what follows, we present a reparameterizaton that aids in disentangling these breaks.

2.1 A Projection-based Decomposition to Disentangle Structural Changes

We now introduce structural changes[7] to the dynamic factor model in Equations (2) and (3). We only work with Equation (2) what follows since the fitting of Equation (3) can occur separately, and we can consistently estimate the space spanned by the static factors. Let k_t denote an indexing variable which partitions x_{it} into two regimes at some break point k

$$X_t = \begin{cases} \Lambda_1 f_t + e_t, & \text{for} \quad k_t \leq k, \\ \Lambda_2 f_t + e_t, & \text{for} \quad k_t > k, \end{cases} \tag{5}$$

where f_t is a $r \times 1$ vector of factors, $\Lambda_1 = (\lambda_{1,1}, \ldots, \lambda_{1,N})^\intercal$ and $\Lambda_2 = (\lambda_{2,1}, \ldots, \lambda_{2,N})^\intercal$ are $N \times r$ pre- and post-break loadings, and e_t is noise. In practice, both the number of factors r and the break k can be consistently estimated (or chosen *a priori* if desired), and thus are be treated as known throughout the remainder of the paper.

By the formulation in Equation (1), any changes in the factor variance must be common across all series, whereas changes in the loadings are by definition idiosyncratic. This motivates us to disentangle breaks in the factor loadings from breaks in the factors, by decomposing the change in Λ_2 via a projection

$$\Lambda_2 = \Lambda_1 Z + W \tag{6}$$

where Z is an $r \times r$ rotational change, and $W = (w_1, \ldots, w_N)^\intercal$ is an $N \times r$ orthogonal shift satisfying $\Lambda_1^\intercal W = O_p(1)$.[8] We briefly note that the rotational change Z can be a singular matrix if desired, which allows a new set of factors which "replace" the old factors, or a

[7] We use the term "structural changes" to refer to any kind of instability in the factor structure, including but not limited to threshold models, Markov based regime switching models, or structural break models.

[8] This implies a rate of $\Lambda_1^\intercal W/N = O_p\left(\frac{1}{N}\right)$, a definition looser than *exact* orthogonality $\Lambda_1^\intercal W = 0$ considered by Wang and Liu (2021); Pelger and Xiong (2022); Massacci (2021). It can be further loosened to *uncorrelatedness* i.e. $E[\lambda_{1i}^\intercal w_i] = 0$ with more complicated conditions on the rates between T and N.

change in the number of factors where some factors disappear (if the rank of W is unchanged or decreases respectively).[9] We elaborate on this case in the Supplementary Material, but otherwise treat the number of factors to be unchanging, as per the extant literature (e.g. Su and Wang, 2017; Massacci, 2021). Thus, the rotation Z and orthogonal shift W are naturally associated with breaks in the factor covariance matrix and loadings respectively. Heuristically, this is because a break in the factor variance can always be thought of some suitable twisting or stretching of the factors themselves, i.e. a mathematical rotation. Note that because breaks in Z are breaks in the covariance matrix of the factors, they encompass breaks in both their variances (i.e. the diagonal elements) and their correlations (off-diagonals). In contrast, a change in the loadings is idiosyncratic across series, and thus geometrically must lie outside and be orthogonal to the space spanned by the factors, (see Wang and Liu, 2021; Pelger and Xiong, 2022; Massacci, 2021, for similar interpretations).

We emphasize that the projection-based decomposition can hold for *any* generic structural change in the factor structure. In what follows, we demonstrate how the reparameterizaton can be used to disentangle a one time structural *break*, and naturally lead to test statistics which can disentangle changes in the factor variance and factor loadings.

2.2 Structural Break Setup

We apply the projection decomposition in a structural break setup. Conditional on the break fraction π satisfying $k = \lfloor \pi T \rfloor$ which splits the data into two partitions that are $T_1 = \lfloor \pi T \rfloor$ and $T_2 = T - \lfloor \pi T \rfloor$ in length, Equation (5) can stacked in matrix form:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_1 \Lambda_1^\intercal \\ F_2 \Lambda_2^\intercal \end{bmatrix} + \begin{bmatrix} e_{(1)} \\ e_{(2)} \end{bmatrix} \tag{7}$$

[9] The case of disappearing factors requires more careful specification of the model, which we detail in Appendix A.5.

where $F_1 = (f_1, \ldots, f_{T_1})^\intercal$ are $T_1 \times r$ pre-break factors, $F_2 = (f_{T_1+1}, \ldots, f_T)^\intercal$ are $T_2 \times r$ post-break factors, Λ_1, Λ_2 are $N \times r$ their respective loadings, $e_{(1)} = (e_1, \ldots, e_{T_1})$ and $e_{(2)} = (e_{T_1+1}, \ldots, e_T)$, and X_1, X_2 denote the respective partitions of X. By substituting Equation (6) into Equation (7), we can formulate an equivalent representation as follows

$$X = \begin{bmatrix} F_1\Lambda_1^\intercal \\ F_2\left[\Lambda_1 Z + W\right]^\intercal \end{bmatrix} + \begin{bmatrix} e_{(1)} \\ e_{(2)} \end{bmatrix} \tag{8}$$

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

$$X = G\Xi^\intercal + e. \tag{9}$$

Equation (9) shows that any rotational changes induced by a non-identity Z are absorbed into the factors, and any orthogonal shifts W will augment the factor space. Equation (9) re-expresses a factor model with structural breaks in its loadings into an observationally equivalent model with time invariant loadings. Equation (9) highlights that if one were to ignore the break and use the principal components estimator over the whole sample, the estimator will be consistent for an observationally equivalent model with *pseudo* factors G and time invariant loadings Ξ. Our formulation aims to complement similar formulations in the literature used to identify and estimate the break point (see, e.g. Han and Inoue, 2015; Baltagi et al., 2017).

Existing methods use an estimate of the pseudo factors G in order to either test for existence of any breaks (e.g. Han and Inoue, 2015; Chen et al., 2014), and/or estimate the break fraction (e.g. Baltagi et al., 2017, 2021; Duan et al., 2022). However, the structure of G implied by Equation (9) implies that breaks in either Z *or* W will induce a structural break in the pseudo factors G: when Z is non identity the first r columns of G will correspond to the multivariate series $[F_1^\intercal, ZF_2^\intercal]^\intercal$; when $W \neq \mathbf{0}$, the orthogonality of W will induce the last r columns of G to be $[0, F_2^\intercal]^\intercal$, corresponding to a structural break where extra factors appear. Therefore, in either case G will exhibit a break, and methods utilizing the estimated pseudo factors will necessarily have power against breaks in the factors variance, even if the loadings are time invariant.

The case of a rotational break corresponds to $Z \neq I_r$, and can be naturally interpreted a change in the factor variance. Indeed, by assuming $\Sigma_F = E(f_t f_t^\intercal)$, it follows from Equation (9) that the covariance matrix of the factors pre- and post-break are Σ_F and $Z\Sigma_F Z^\intercal$ respectively, which are in general different for non identity Z.[10]

Given that Z captures changes in the factor variance, it follows that the remaining orthogonal shift where $W \neq \mathbf{0}$ must correspond to breaks in the factor loadings.[11] Mechanically, these breaks lie outside the original factor space, which necessitates the estimation of more factors than necessary if one wanted to capture all the information while ignoring the break. It is this orthogonality of breaks in the loadings that cause the "factor augmentation" effect in the pseudo factors G raised by Breitung and Eickmeier (2011).

Disentangling these breaks naturally entails testing for changes in these parameters: a test for a break in the factor variance corresponds to $\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal$, and a test for a break in the factor loadings corresponds to $\mathcal{H}_0 : W = \mathbf{0}$. In what follows, we develop these ideas more fully, but for now, note that our tests have standard chi-square distributions. We also note that our test statistics aim to determine which type of break has occurred, for a candidate break provided either *a priori*, or estimated from the data. In the case of multiple breaks, practitioners can simply partition the data suitably and either separately or sequentially focus on different breaks, as we have in our empirical study in Section 5.

Finally, we note that while we mainly use the reparameterization to develop tests to disentangle changes in the factor variance and factor loadings, the reparameterization should have broader applications and utility in modeling structural change in factor models.

[10] Our rotational changes correspond to the type 2 breaks of Han and Inoue (2015) and Baltagi et al. (2017), and the type B breaks of Duan et al. (2022).

[11] Such breaks correspond 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.

2.3 Implementing the Test to Disentangle Structural Breaks

We outline how to implement our tests to disentangle breaks in the factor variance and loadings, relegating theoretical justifications to Section 3. We first condition on a known break fraction, which can be chosen *a priori* (e.g. Breitung and Eickmeier, 2011; Stock and Watson, 2012), or obtained from a test such as Baltagi et al. (2021). The estimation of various quantities required for constructing the test statistics then follows as:

1. Estimate the pre- and post-break factors $\tilde{F}_1$ and $\tilde{F}_2$ via principal components, subject to the normalizations $\frac{1}{T_1}\tilde{F}_1^\intercal\tilde{F}_1 = \frac{1}{T_2}\tilde{F}_2^\intercal\tilde{F}_2 = I_r$, i.e. $\tilde{F}_1$ is $\sqrt{T_1}$ times the first r eigenvectors of $X_1X_1^\intercal$, and $\tilde{F}_2$ is $\sqrt{T_2}$ times the first r eigenvectors of $X_2X_2^\intercal$.

2. Conditional on the factors, estimate the pre and post break loadings via least squares as $\tilde{\Lambda}_1 = X_1^\intercal\tilde{F}_1(\tilde{F}_1^\intercal\tilde{F}_1)^{-1} = \frac{1}{T_1}X_1^\intercal\tilde{F}_1$ and $\tilde{\Lambda}_2 = X_2^\intercal\tilde{F}_2(\tilde{F}_2^\intercal\tilde{F}_2)^{-1} = \frac{1}{T_2}X_2^\intercal\tilde{F}_2$.[12]

3. Estimate the rotational change and orthogonal shift as

$$\tilde{Z} = (\tilde{\Lambda}_1^\intercal\tilde{\Lambda}_1)^{-1}\tilde{\Lambda}_1^\intercal\tilde{\Lambda}_2, \tag{10}$$

$$\tilde{W} = \tilde{\Lambda}_2 - \tilde{\Lambda}_1\tilde{Z}. \tag{11}$$

The estimates $\tilde{Z}$ and $\tilde{W}$ absorb the effects of the normalization bases present in $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$, and hence cannot be tested directly. Instead, $\tilde{Z}$ and $\tilde{W}$ take on additional interpretations: post multiplying $\tilde{\Lambda}_1$ by $\tilde{Z}$ rotates its normalization basis to that of $\tilde{\Lambda}_2$ along with any rotational change Z; the remainder $\tilde{W}$ is the remaining idiosyncratic change.[13] We exploit this property in $\tilde{Z}$ to rotate the post-break factors onto the same basis as the pre-break factors, defined as $\widehat{F} = [\tilde{F}_1^\intercal, \tilde{Z}\tilde{F}_2^\intercal]^\intercal = [\widehat{f}_1, \dots, \widehat{f}_T]^\intercal$. The variance of the combined series $\widehat{f}_t$ reflects the effect of any rotational change Z, and thus can be used as a basis for a test.

[12]This follows because $\tilde{F}_1$ and $\tilde{F}_2$ have unit variance.

[13]The case of a disappearing factor can be accommodated by estimating $\tilde{Z}$ as a $r_1 \times r_2$ "rectangular" matrix and $\tilde{W}$ as an $N \times r_2$ matrix, where r_1 and r_2, are the pre- and post-break number of factors such that $r_2 < r_1$.

Test Statistics

We now construct the test statistics, which we label as the Z-test for breaks in the factor variance, and the W-test for breaks in the loadings following Equations (6) and (8).

We present the Z-test statistic for changes in the factor variance $\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal$ as

$$\mathscr{W}_Z(\pi, \widehat{F}) = A_Z(\pi, \widehat{F})^\intercal \widehat{S}_Z(\pi, \widehat{F})^{-1} A_Z(\pi, \widehat{F}), \tag{12}$$

where

$$A_Z(\pi, \widehat{F}) = vech\left(\sqrt{T}\left(\frac{1}{\lfloor \pi T \rfloor}\sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^\intercal - \frac{1}{T - \lfloor \pi T \rfloor}\sum_{t=\lfloor \pi T + 1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^\intercal\right)\right) \tag{13}$$

denotes the difference in the subsample means of the second moments (outer product) process of $\widehat{f}_t$ at a given break fraction π, and $vech()$ denotes the column-wise vectorization of a square matrix with the upper triangle excluded. Its long run variance estimated as a weighted average of the variance from pre- and post-break data

$$\widehat{S}_Z(\pi, \widehat{F}) = \frac{1}{\pi}\widehat{\Omega}_{Z,(1)}(\pi, \widehat{F}) + \frac{1}{1-\pi}\widehat{\Omega}_{Z,(2)}(\pi, \widehat{F}), \tag{14}$$

where $\widehat{\Omega}_{Z,(1)}$, $\widehat{\Omega}_{Z,(2)}$ are HAC estimators constructed using the respective subsamples of $vech\left(\widehat{f}_t \widehat{f}_t^\intercal - I_r\right)$. Alternatively, a bootstrap based procedure to estimate the variance can be entertained. The statistic $\mathscr{W}_Z(\pi, \widehat{F})$ is a Wald test for whether the subsample means of the second moments process of $\widehat{f}_t$ are the same at the pre-specified break point,[14] and thus has a conventional χ^2 distribution with $r(r+1)/2$ degrees of freedom.

Next, we present the W-test statistics for changes in the loadings $\mathcal{H}_0 : W = \mathbf{0}$. Defining

[14] As noted by an anonymous referee, this statistic is similar to that constructed by Han and Inoue (2015); Baltagi et al. (2021), only that we construct the statistic using estimates constructed under the alternative hypothesis.

$\tilde{w}_i$ as the ith row of $\tilde{W}$, the individual statistic for the ith series and its variance are

$$\mathscr{W}_{W,i} = T\tilde{w}_i^{\intercal}\tilde{\Omega}_{W,i}^{-1}\tilde{w}_i, \tag{15}$$

$$\tilde{\Omega}_{W,i} = \frac{1}{\pi}\tilde{\Theta}_{1,i} + \frac{1}{1-\pi}\tilde{\Theta}_{2,i} \tag{16}$$

which uses pre and post break HAC estimates of the asymptotic variance $\tilde{\Theta}_{1,i}$ and $\tilde{\Theta}_{2,i}$ respectively.[15] Next, define the joint statistic for all variables as:

$$\mathscr{W}_W = (TN)\left(\frac{\sum_{i=1}^N \tilde{w}_i}{N}\right)^{\intercal}\tilde{\Omega}_W^{-1}\left(\frac{\sum_{i=1}^N \tilde{w}_i}{N}\right), \tag{17}$$

where the matrix $\tilde{\Omega}_W = N^{-1}\sum_{i=1}^N \tilde{\Omega}_{W,i}$ is an estimate of the joint variance. Both are Wald tests based on $\tilde{w}_i$, and thus have standard χ^2 distributions with r degrees of freedom.

We emphasize that our test statistics do not maintain any assumptions on their counterparts; that is, the test for $\Sigma_F = Z\Sigma_F Z^{\intercal}$ holds irrespective of W, and conversely the test for $W = \mathbf{0}$ holds irrespective of Z. As both breaks in the factor variance and the loadings could occur simultaneously, this necessitates the practitioner to run both the Z-test and W-test to accurately pin down the source(s) of the break. A straightforward Bonferroni-Holm correction suffices to correct for the family wise error rate (see Section 4).

3 Asymptotic Theory

We discuss the asymptotic theory underpinning the test statistics discussed in Section 2.3. All limits are taken as both N and T tend to infinity simultaneously, and δ_{NT} is defined as $\min(\sqrt{T}, \sqrt{N})$. For notation, $\|\cdot\|$ denotes 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, $\lfloor\cdot\rfloor$ denotes the floor operator, M denotes generic

[15] These are constructed using the residuals $\tilde{e}_{(1),it} = x_{it} - \tilde{\lambda}_{1,i}^{\intercal}\tilde{f}_{1,t}$ and $\tilde{e}_{(2),it} = x_{it} - \tilde{\lambda}_{2,i}^{\intercal}\tilde{f}_{2,t}$ in the series $\tilde{Z}^{\intercal}\tilde{f}_{1,t}\tilde{e}_{(1),it}$ and $\tilde{f}_{2,t}\tilde{e}_{(2),it}$ respectively, (see A.4 in the Supplementary Material), or via a bootstrap.

finite constants, and $A^{-\intercal}$ denotes the inverse transpose of any invertible matrix A.

3.1 Estimation

We first establish the properties of the estimated rotational and orthogonal shift components by making the following assumptions. Let $\iota_{1t} \equiv \mathbf{1}\{t \leq \lfloor \pi T \rfloor\}$ and $\iota_{2t} \equiv \mathbf{1}\{t \geq \lfloor \pi T \rfloor + 1\}$.

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 $\Sigma_F > 0$.*

Assumption 2. *For $m = 1, 2$, $E\|\lambda_{m,i}\|^4 \leq M$, $\|\Lambda_m^\intercal \Lambda_m / N\| - \Sigma_{\Lambda_m} \xrightarrow{p} 0$ for some $\Sigma_{\Lambda_m} > 0$, and $\|\Lambda_m^\intercal \Lambda_m / N - \Sigma_{\Lambda_m}\| = O_p\left(N^{-1/2}\right)$. The shift break is orthogonal such that $\Lambda_1^\intercal W = O_p(1)$.*

Assumption 3. *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 $m = 1, 2$, $\{\lambda_{m,i}\}$, $\{f_t\}$ and $\{e_{it}\}$ are mutually independent groups.*

Assumption 5. *For all T and N:*

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

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

Assumption 6. *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 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 a common identification condition (e.g. Han and Inoue, 2015; Baltagi et al., 2017, and others) which limits the factors to exhibit no heteroscedasticity, but this is not restrictive in our case as changes in Σ_F are characterized by Z. Assumption 2 is 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 relaxes the strict $\Lambda_1^{\intercal}W = 0$ condition found in Wang and Liu (2021); Pelger and Xiong (2022); Massacci (2021).[16] Assumption 3 allows for weak serial and cross sectional correlation and corresponds to Assumption C of Bai (2003). Assumption 4 is standard in the factor modeling literature, and is the subsample version of Assumption D of Bai and Ng (2006). Assumption 5 strengthens Assumption 3, and corresponds to Assumption E in Bai (2003). Assumption 6 are Assumptions F1-F2 of 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

[16] 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.

not required for estimation. Assumption 6 (c) is slightly stronger than Assumption F3 of Bai (2003), which only requires the existence of the second moments. Assumption 7 corresponds to Assumption G in Bai (2003). Assumption 8 requires that the there is infinite data pre- and post-break, and is a weaker version of Assumption 8 in Han and Inoue (2015), who assume that the terms are bounded uniformly in a range of potential π.

Recall that $\tilde{F}_1$ and $\tilde{F}_2$ satisfy $\tilde{F}_1^\intercal \tilde{F}_1/T_1 = \tilde{F}_2^\intercal \tilde{F}_2/T_2 = I_r$, and are thus estimates of $F_1 H_1$ and $F_2 H_2$, where H_1 and H_2 are the respective pre and post-break normalization bases.[17]

Theorem 3.1. *Under Assumptions 1 to 8, and as $N, T \to \infty$*

(a) $\left\| \tilde{Z} - H_1^\intercal Z H_2^{-\intercal} \right\| = O_p\left(\frac{1}{\delta_{NT}^2}\right)$,

(b) $\frac{1}{N}\left\| \tilde{W} - W H_2^{-\intercal} \right\|^2 = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

Theorem 3.1 (a) shows that $\tilde{Z}$ is consistent for Z, but is affected by the normalization matrices. Recalling that $\tilde{F}_2$ estimates $F_2 H_2$, this then implies that the combined series $\widehat{F} = [\tilde{F}_1^\intercal, \tilde{Z}\tilde{F}_2^\intercal]^\intercal$ is on the same normalization basis both before and after the break, and can thus form the basis of a test for evidence of rotational breaks. Importantly, $\widehat{F}$ is free from the effects of any possible orthogonal shifts induced by W, and thus isolates the rotational change in the factor variance. Theorem 3.1 (b) shows that $\tilde{W}$ estimates the true W up to a normalization basis, and thus can used to construct test statistics.

3.2 Z-test for Rotational Changes

We analyze the Z-test for the null of no break in the factor variance $\mathcal{H}_0 : \Sigma_F = Z \Sigma_F Z^\intercal$ that holds regardless of W, and is therefore robust to breaks in the factor loadings. We

[17] Specifically, their respective bases are $H_1 = \left(\frac{\Lambda_1^\intercal \Lambda_1}{N}\right)\left(\frac{F_1^\intercal \tilde{F}_1}{T_1}\right) V_{NT,1}^{-1}$, $H_2 = \left(\frac{Z^\intercal \Lambda_1^\intercal \Lambda_1 Z}{N} + \frac{W^\intercal W}{N}\right)\left(\frac{F_2^\intercal \tilde{F}_2}{T_2}\right) V_{NT,2}^{-1}$, where $V_{NT,1}$ and $V_{NT,2}$ are the diagonal matrices of eigenvalues of the first r eigenvalues of $(NT_1)^{-1} X_1 X_1^\intercal$ and $(NT_2)^{-1} X_2 X_2^\intercal$. There exists an alternative parameterization $H_2^\dagger = \left(\Lambda_1^\intercal \Lambda_1 + Z^{-\intercal} W^\intercal W Z^{-1}\right)\left(Z F_2^\intercal \tilde{F}_2\right)/(NT_2) V_{NT,2}^{-1}$, where the rotation Z is parameterized as part of the factors. It is straightforward to verify that either parameterization leads to same result, (see Remark 3 in the Supplementary Material).

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 $\mathscr{W}_Z(\pi, \widehat{F})$, where $H_{0,1} = \text{plim}(H_1)$,[18] 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 $\widehat{\Omega}_{Z,(1)}(\pi, FH_{0,1})$ and $\widehat{\Omega}_{Z,(2)}(\pi, FH_{0,1})$ converge to their population counterpart Ω_Z. Assumption 10 (b) is the main result of Theorem 3 of Andrews (1993), and is a necessary hyper-assumption 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. Assumption 10 (b) has been used in Han and Inoue (2015), and one can refer to Chen et al. (2014) for more primitive assumptions to see that it is satisfied for a large class of ARMA processes.

Theorem 3.2. *Under Assumptions 1 to 10, and if $\frac{\sqrt{T}}{N} \to 0$, then $\mathscr{W}_Z(\pi, \widehat{F}) \xrightarrow{d} \chi^2_{r(r+1)/2}$.*

Theorem 3.2 shows that the Z-test statistic converges to a chi-squared random variable,

[18] The definition of $H_{0,1}$ follows from Bai (2003). Lemma A3 of Bai (2003) shows that $V_{NT,1}$ converges to V_1, a diagonal matrix with the eigenvalues of $\Sigma_{\Lambda_1}^{1/2} \Sigma_F \Sigma_{\Lambda_1}^{1/2}$. Let Υ_1 denote its eigenvectors such that $\Upsilon_1^\intercal \Upsilon_1 = I_r$. Proposition 1 of Bai (2003) shows that $F_1^\intercal \tilde{F}_1 / T_1$ converges to $\Sigma_{\Lambda_1}^{-1/2} \Upsilon_1 V_1^{1/2} = H_{0,1}^{-\intercal}$. It follows that $H_1 \xrightarrow{p} \Sigma_{\Lambda_1}^{1/2} \Upsilon_1 V_1^{-1/2} = H_{0,1}$, its probability limit. One can also define $H_{0,2} = \text{plim}(H_2)$ in a similar way.

conditional on a break fraction.[19] To ensure the Z-test's power under the alternative, we make the following assumptions.

Assumption 11. $\|Z\| < \infty$*, and* $Z\Sigma_F Z^\intercal \neq \Sigma_F$*.*

Assumption 12. $plim_{T\to\infty} \inf \left(vech(C)^\intercal \left[\max(b_{\lfloor \pi T \rfloor}, b_{T-\lfloor \pi T \rfloor}) \widehat{S}(\pi, F^* H_{0,1})^{-1} \right] vech(C) \right) > 0$*, where* $F^* = [F_1^\intercal, Z F_2^\intercal]^\intercal$ *and* $C \equiv H_{0,1}^\intercal (\Sigma_F - Z\Sigma_F Z^\intercal) H_{0,1}$*.*

Assumption 11 formalizes the definition of a break in factor variance. 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., 2017, 2021, and others). Assumption 12 regulates the asymptotics of the variance matrices of the statistics under the alternative. Together, these ensure that the subsample means of $\widehat{f}_t \widehat{f}_t^\intercal$ converge to different limits, and the divergence of $\mathscr{W}_Z(\pi, \widehat{F})$, as summarized in the following theorem.

Theorem 3.3. *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. $\mathscr{W}_Z(\pi, \widehat{F}) \to \infty$ *under the alternative hypothesis that* $\Sigma_F \neq Z\Sigma_F Z^\intercal$*.*

3.3 W-test for Orthogonal Shifts

Next, we analyze the W-test for breaks the loadings $\mathcal{H}_0 : W = \mathbf{0}$ that holds regardless of Z, and is therefore robust to changes in the factor variance. First note that because W is an $N \times r$ matrix where $N \to \infty$, traditional tests are infeasible. We note that popular approaches such as Bonferroni test statistics and pooling individual test statistics

[19]It is also possible to construct an LM-like statistic with a restricted estimate of the variance using all of the data (see Appendix A.5.3 in the Supplementary Material). 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.

can suffer from significant size distortions,[20] whereas directly testing for a change in the number of factors[21] requires much stricter assumptions on the error term. This motivates us to formulate an individual test statistic $\mathscr{W}_{W,i}$ for each i, and a joint test statistic $\mathscr{W}_W$ pooled across N to overcome the infinite dimensionality problem. To analyze the test statistics, we make the following additional assumptions.

Assumption 13. *For all N, T,*

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

Assumption 14. *For all N, T, and $m = 1, 2$:*

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

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

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

[20]The Bonferroni test statistic B_w is the maximum of $\mathscr{W}_{W,i}$, and the critical value is $\mathscr{X}^{-1}(0.01-0.05/N)$ where $\mathscr{X}$ is the chi-square CDF with r degrees of freedom. The pooled statistic pools the individual test *statistics* into $P_w = \left(\sum_{i=1}^{N}\mathscr{W}_{W,i} - rN\right)/\sqrt{2Nr}$, where the corresponding critical values are ± 1.96. Both of these approaches rely on the sequential limit argument that $\mathscr{W}_{W,i} \to \chi_r^2$, and then $N \to \infty$. However, convergence of $\mathscr{W}_{W,i}$ relies on the joint asymptotics of N and T to ∞, so $\mathscr{W}_{W,i} \to \chi_r^2$, and then $N \to \infty$ cannot be separated into two separate steps. Indeed, the sequential and joint limits are not always equivalent (see Phillips and Moon, 1999), and both Bonferroni and pooled test statistics are known for potentially introducing significant size distortions (see Stock and Watson, 2005; Han, 2015).

[21]This follows by restating the null and alternative hypotheses as $\mathcal{H}_0 : r_w = 0, \mathcal{H}_1 : r_w \neq 0$, where r_w is the number of extra factors augmented by the presence of orthogonal shifts that is implied when using the *pseudo* factor representation. Existing tests such as Onatski (2009) cannot be used without imposing further restrictive assumptions on the errors of the approximate factor model.

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 the pooled version of Assumption 3. Assumption 14 (a) is Assumption 6 (a) but for 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 is required to bound the sum of the loadings by $O_p\left(\sqrt{N}\right)$, and is a slightly modified version of the Assumption 7 in Han (2015). This will hold if the loadings are centered around zero, and their sum diverges at the rate of $\sqrt{N}$ by the central limit theorem (CLT). Although this is somewhat stricter than a conventional factor model setup, it seems to hold for empirically used datasets. Assumptions 16 (a) and 16 (b) are CLTs, where the latter is the cross sectionally averaged version of the CLT in Bai (2003), which somewhat restricts the cross sectional correlation in e_{it}.

Theorem 3.4. *If* $\frac{\sqrt{T}}{N} \to 0$, *then:*

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

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

Theorem 3.4 shows that the W-test statistics[22] test statistics converge to conventional chi-squared random variables. To analyze the joint W-test under the alternative, we make 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

[22]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.

distribution to some Gaussian random variable by the CLT, 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 under the alternative hypothesis, even when if $N^{-1} \sum_{i=1}^{N} w_i \xrightarrow{p} 0$, as summarized in the following Theorem.

Theorem 3.5. *If $\frac{\sqrt{T}}{N} \to 0$, and the alternative $\mathcal{H}_1 : W \neq \mathbf{0}$ holds, then:*

(a) under Assumptions 1 to 9, 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 9 and 13 to 17, $\mathscr{W}_{W} \to \infty$ if $\frac{\sqrt{N}}{T^{1-\alpha/2}} \to 0$ as $N, T \to \infty$.

3.4 Estimation of Number of Factors and Break Fraction

Our test statistics assume that the number of factors and break fraction are known. In practice, consistent estimators can be used instead, as addressed in the following remarks.

Remark 1. *The number of factors r in either subsample can be consistently estimated conditional on consistent estimate of π (e.g. Bai and Ng, 2002; Onatski, 2010; Ahn and Horenstein, 2013; Baltagi et al., 2017). If the pre- and post-break estimates of the number of factors $\tilde{r}_1$ and $\tilde{r}_2$ differ, this can be accommodated by allowing $\tilde{Z}$ to be a $\tilde{r}_1 \times \tilde{r}_2$.*

Remark 2. *The break fraction π can be consistently estimated (e.g. Baltagi et al., 2017; Duan et al., 2022, and others). Theorem 3 of Baltagi et al. (2017) shows that 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.*

4 Monte Carlo Study

4.1 Simulation Specification

We first simulate two sets of loadings, Λ_1, Λ_2 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 $\Lambda_2 - (\Lambda_1^\intercal \Lambda_1)^{-1}\Lambda_1^\intercal \Lambda_2$. The rotation Z is set to I_3 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 we simulate from is

$$x_{it} = \begin{cases} \lambda_{1,i}^\intercal f_t + \sqrt{\theta} e_{it}, & t = 1, \ldots, \lfloor \pi T \rfloor \\ (Z\lambda_{1,i} + \omega w_i)^\intercal f_t + \sqrt{\theta} e_{it}, & t = \lfloor \pi T \rfloor + 1, \ldots, T, \end{cases} \tag{18}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The parameter $\theta = 3$ sets the signal to noise ratio to be 50%, and ω controls the "size" of the orthogonal shifts. The factors and errors are generated as:

$$f_{k,t} = \rho f_{k,t-1} + \mu_{it}, \quad \mu_{it} \sim i.i.d.\ N(0, 1 - \rho^2), \tag{19}$$

$$e_{it} = \alpha e_{i,t-1} + v_{it}, \tag{20}$$

where $\rho \in \{0, 0.7\}$ controls autocorrelation in the factors, and μ_{it}, v_{it} are mutually independent with $v_t = (v_{1,t}, \ldots, v_{N,t})^\intercal$ being i.i.d. $N(0, \Omega)$ for $t = 1, \ldots, T$. The scalar $\alpha \in \{0, 0.3\}$ allows mild serial correlation, and as in Bates et al. (2013) and Baltagi et al. (2017), $\Omega_{ij} = \beta^{|i-j|}$ with $\beta \in \{0, 0.3\}$ allowing mild cross sectional correlation. The break fraction is set to 0.5 and treated as known.[23] Disentanglement requires 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).

[23] Alternative break fractions of 0.3 and 0.7 do not qualitatively change the results, and are relegated to the Supplementary Material.

Table 1: Size of Rotation and Orthogonal Shift Tests, $N = 200, r = 3$

				Z-Test		W-Test		W Individual
T	ρ	α	β	Unadjusted	Adjusted	Unadjusted	Adjusted	
200	0.0	0.0	0.0	0.140	0.088	0.135	0.107	0.027
		0.3	0.3	0.147	0.101	0.092	0.055	0.016
500		0.0	0.0	0.100	0.053	0.003	0.001	0.007
		0.3	0.3	0.108	0.058	0.064	0.032	0.007
200	0.7	0.0	0.0	0.220	0.156	0.125	0.085	0.027
		0.3	0.3	0.215	0.154	0.087	0.062	0.029
500		0.0	0.0	0.136	0.086	0.003	0.001	0.008
		0.3	0.3	0.134	0.085	0.062	0.042	0.011

Note:
Entries report the rejection frequencies for the Z-test for break in factor variance, and W-test for break in loadings. Nominal size is 5%. The parameters α and β denote the degree of serial and cross sectional correlation in the error respectively, ρ denotes the degree of autocorrelation in the factors. The number of pseudo factors estimated over the whole sample using $IC_p(2)$ of Bai and Ng (2002) is $\tilde{r}$.

4.2 Simulation Results

We present the size analysis in Table 1. In the case of no serial correlation in the factors and $T > N$, the Z-test has a nominal size close to 5% 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 is alleviated and approaches a rejection rate of 0.09 as T increases.[24] 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 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; both have good power and are rejecting correctly only on their respective break types. This contrasts with the tests of Han and Inoue (2015) and Baltagi et al. (2021), which reject across all break types, and thus cannot discern which type of break has occurred.

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

Table 2: Power of Z and W Tests, $r = 3$, $N = 200$, $\alpha = \beta = 0.3$

				Z-Test		W-Test					
Break Type	T	ω	ρ	Unadj.	Adj.	Unadj.	Adj.	Individual	HI (2015)	BKW (2021)	$\tilde{r}$
$W \neq \mathbf{0}$	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
$Z \neq I$	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
$W \neq \mathbf{0}$ and $Z \neq I$	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

Note:
Entries denote the rejection rates across different simulated break types; a break type of $W \neq 0$ denotes a break in the loadings, $Z \neq I$ a break in the factor variance, and $W \neq 0$ and $Z \neq I$ denoting a break in both. HI denotes Han and Inoue (2015)'s test, and BKW denotes Baltagi et al (2021)'s test conducted with a pre known break fraction. The scalar ω denotes the "size" of the break in the loadings. See Table 1 for explanation of α, β, ρ and $\tilde{r}$.

5 Empirical Application

We apply our methodology to FRED-QD, a standard U.S. quarterly macroeconomic dataset (see McCracken and Ng, 2020), developed to mimic the Stock and Watson (2012) dataset. The Stock and Watson (2012) dataset is widely used in the factor modeling literature, having been used to date and test breaks in factor models (e.g. Chen et al., 2014; Baltagi et al., 2021), marking FRED-QD as well suited for our application.

We consider the sample 1959Q3-2019Q4 from the FRED-QD dataset, and adopt the suggested data cleaning and transformations, removing top level aggregates to yield a panel of 124 series.[25] We align our start date to extant work, and end before the COVID-19 pandemic (as noted by Ng, 2021; Stock and Watson, 2021, this period is unclear to deal with). We consider two breaks in 1984Q1 and 2008Q3, as estimated by the procedure of Baltagi et al. (2021) using 2-6 factors and a trimming parameter of 0.1, which we henceforth refer to as the Great Moderation and Great Recession break respectively. We note that these breaks align with events associated *a priori*, and existing evidence: the Great Moderation break is consistent with work that use a sample of 1960 to the mid

[25] Additional details of the data can be found in Table 9 in the Supplementary Material.

2000s and find a break in the early 1980s (see Stock and Watson, 2009; Breitung and Eickmeier, 2011; Chen et al., 2014; Baltagi et al., 2021), and the Great Recession break has some evidence of breaks in loadings as noted by Stock and Watson (2012) who work with a very short post-2009 sample.[26] The number of factors in each subsample differs across estimators and is often contradictory;[27] given this known instability, we consider 2 to 6 factors for our empirical analysis. In line with the extant literature, we set the number of pre and post-break factors to be the same; results allowing for a change in the number of factors are qualitatively similar and in Table 10 of the Supplementary Material.

5.1 Joint test results

Table 3 reports the p-values for the Z and joint W-tests with the Great Moderation and the Great Recession as candidate break dates. Across 2 to 6 factors, there is strong evidence of rejection of the null hypothesis of no breaks for both tests across both the Great Moderation and Great Recession, with all p values being less than 0.05, indicating that both types of structural breaks are empirically relevant for factor models of U.S. macroeconomic data.

5.2 Were breaks in the factor variance important?

A key tenet of the argument of the importance of distinguishing breaks in the factor variance and loadings is because the routine normalization applied in factor models rules out the possible interpretation that the factor variance changed; breaks in the factor variance are necessarily subsumed into breaks in the factor loadings. Given the clear rejection of the Z-test, our results suggest that rotational breaks, which may stem from changes in the factor variance, are important when modeling U.S. macroeconomic data using factors.

[26]Barigozzi et al. (2018) and Cheng et al. (2016) also find a structural break in the factor structure around the Great Recession period.

[27]The eigenvalue edge distribution estimator of Onatski (2010) and eigenvalue ratio estimator of Ahn and Horenstein (2013) often finds 2 to 4 factors in the sub-regimes, while information criteria such as Bai and Ng (2002) often suggest a large number of factors such as 5 to 9. Stock and Watson, on the other hand, always consider 6 factors when studying factor instability (e.g. Stock and Watson, 2009, 2012).

Table 3: Joint Test Results

	Z Test p values		W Test p values	
$\tilde{r}$	Unadjusted	Adjusted	Unadjusted	Adjusted
Great Moderation (1984 Q1), 1959 Q3 - 2008 Q3 Sample				
2	0.001	0.001	0.000	0.000
3	0.000	0.000	0.000	0.000
4	0.008	0.007	0.000	0.000
5	0.000	0.000	0.001	0.001
6	0.000	0.000	0.000	0.000
Great Recession (2008 Q3), 1984 Q2 - 2019 Q4 Sample				
2	0.000	0.000	0.004	0.004
3	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000

Note:
Rejection of the Z-test corresponds to a break in the factor covariance matrix, and rejection of the W-test corresponds to a break in the loadings across the entire cross section.

From our framework, rejection of the Z-test implies that the covariance matrix of the factors has changed, i.e. $\Sigma_F \neq Z\Sigma_F Z^\intercal$. However, it does not assess the economic significance of these breaks. Our projection-based decomposition also allows us to estimate how much the total factor variance has changed pre- and post-break. Denoting $tr(A)$ as the trace of a square matrix A, $tr(\Sigma_F)$ and $tr(Z\Sigma_F Z^\intercal)$ are the total variance of the factors pre and post-break. Therefore, their ratios provide an estimate of the how much the total factor variance has (de)inflated post-break. Table 4 presents our estimate of the ratio, or $tr(Z\Sigma_F Z^\intercal)/tr(\Sigma_F)$ along with the 95% confidence interval.[28] We estimate that the Great Moderation was associated with an over 60% reduction in the total variance of the factors, across the specification of 2 to 6 factors. For the Great Recession break, this ratio is close to 1, with 1 always being within the confidence interval. This suggests that despite rejection of the Z-test, breaks in the factor variance were less important for understanding the factor model relative to the Great Moderation break.[29]

[28] More precisely, we estimate $tr(H_{0,1}^\intercal Z\Sigma_F Z^\intercal H_{0,1})/tr(H_{0,1}^\intercal \Sigma_F H_{0,1})$, see Appendix C.4.

[29] Note that rejection of the Z-test and is due to a break in the factor covariance matrix, may occur even when the total factor variance (trace of the variance matrix) remains stable. This can occur if the individual factors' variances (diagonal elements) break but their sum remains the same, if the correlation between factors (off-diagonals) change, or some combination thereof. Since we can only estimate the space spanned by the factors, the Z-test cannot tell the two cases apart. All we can claim is given the trace appears similar pre and post-break with the Great Recession, breaks in factor variances are probably less important, relative to the Great Moderation.

Table 4: Estimated ratio of the factor variances

r	$tr(\tilde{Z}\tilde{Z}^{\intercal})/tr(I_r)$	95% Bootstrap Confidence Interval
Great Moderation (1984 Q1), 1959 Q3 - 2008 Q3 Sample		
2	0.255	[0.185, 0.269]
3	0.294	[0.189, 0.306]
4	0.347	[0.239, 0.353]
5	0.306	[0.23, 0.324]
6	0.289	[0.223, 0.301]
Great Recession (2008 Q3), 1984 Q2 - 2019 Q4 Sample		
2	0.893	[0.888, 1.089]
3	1.303	[0.762, 1.49]
4	1.208	[0.957, 1.375]
5	1.097	[0.913, 1.158]
6	1.030	[0.894, 1.112]

Note:
The table presents estimates of the ratio of the total factor variance pre and post-break, or $tr(\Sigma_F)/tr(Z\Sigma_F Z^{\intercal})$. Values less than 1 indicate that the estimated total variance of the factors pre-break is smaller than the total variance of the factors post-break. Confidence intervals are constructed by a block bootstrap to preserve both serial and cross sectional correlation.

Being able to associate Great Moderation break with a change in the factor variance reconciles how one can understand the underlying dynamic factor model in Equations (2) and (3). First, work by, for example, Primiceri (2005), Cogley and Sargent (2005), and Sims and Zha (2006) attach a "Good luck" interpretation of the Great Moderation. "Good luck" interprets the Great Moderation as arising from smaller shocks hitting the economy, whereas "Good policy" views the Great Moderation as arising from explicit policy choice. From the perspective of the dynamic factor model, the "Good luck" interpretation is only possible through a break in the variance of the underlying shocks, or Σ_η in Equation (3). In contrast, the "Good policy" interpretation arises from parameters not linked to the variances of shocks. That is, a break in the factor variance is a necessary condition for the dynamic factor model to attach the "Good luck" interpretation. Our estimate of an over 60% decrease in the factor variance from pre to post Great Moderation thus allows for the "Good luck" interpretation. To be clear, our tests alone cannot distinguish between the "Good luck" and "Good policy" interpretation because breaks in the factor dynamics in Equation (3) (i.e. the Φ_j's) without a corresponding break in Σ_η, also lead to a break in the factor variance, and one would probably attach a "Good policy" interpretation akin to Lubik and Schorfheide (2004) for such a case. Second, regardless of how one interprets the Great Moderation, the fact that we can reject our Z-test and estimate an over 60% decrease in the total factor variance tells us that a, probably large, break associated with the Great Moderation, when viewed via the dynamic factor model, *must* have occurred in the equations governing the factor dynamics or Equation (3). Regardless of whether breaks in Equation (2) occurred or were important, the extant literature, by only working with breaks in the loadings, rules out interpreting possible breaks in Equation (3) by construction.

5.3 Individual Test Results

Besides the joint W-test, we can also test for breaks in loadings individually. Table 5 presents the number of series where we can reject at least one of their factor loadings breaking, controlling for a possible break in the factor variance. For comparison, we also apply Breitung and Eickmeier (2011)'s test for breaks in the factor loadings; we caution however that direct comparison is not straightforward due to their use of *pseudo* factors, and hence non-robustness to changes in the factor variance. We nonetheless suggest two tentative conclusions. First, there is some evidence to suggest that by accounting for changes in the factor variance, one may find fewer breaks in the loadings. This suggests that for the Great Moderation, some of the breaks in factor loadings found when using pseudo factors may be related to breaks in the factor variance instead of the factor loadings. We base this tentative conclusion by the fact that we find fewer series with a break in their factor loadings than the Breitung and Eickmeier (2011) procedure, but note that this seems to only hold when we consider 2-4 factors. Alternatively, our higher number of rejections for 5-6 factors could be due to the loss of power of Breitung and Eickmeier (2011)'s statistic as the number of factors increases, as documented by Yamamoto and Tanaka (2015). Second, our results suggest that with a longer post Great Recession sample, we find many more breaks in the loadings than Stock and Watson (2012) and Breitung and Eickmeier (2011).

Table 5: Individual Series Loading Break Test Rejection Counts

$\tilde{r}$	Individual w_i	Breitung and Eickmeier (2011)
Great Moderation (1984 Q1), 1959 Q3 - 2008 Q3 Sample		
2	54	69
3	59	66
4	51	60
5	70	61
6	72	71
Great Recession (2008 Q3), 1984 Q2 - 2019 Q4 Sample		
2	55	25
3	48	32
4	56	42
5	67	45
6	66	52

* Numbers in cells represent the count of rejections of the null hypothesis that the loadings of an individual series broke at the given break date, (5% significance level). Total of 124 series in each subsample.

One possibility is our test is more powerful in detecting breaks in the factor loadings, though we urge caution with this interpretation given their use of the pseudo factors means that the procedures can be quite different.

5.4 Which variables experienced a break in their loadings?

To further understand the break associated with the Great Moderation and Great Recession, we explored which types of variables had breaks in loadings. In order to understand whether these breaks were important for understanding the variation in variables, we calculate an R^2 measure for each series subject to no restrictions, and with the restriction that there were no breaks in the loadings.[30] Thus, these R^2 statistics should have a large (small) difference if the breaks in the loadings were (un)important.

Figure 1 presents the unrestricted and $W = 0$ restricted R^2 statistics averaged across all series by category. We present the R^2 for $r = 3$ but note that these conclusions are very similar for $r = 2$ to 6. For the Great Moderation break, breaks in the loadings appear to be important for prices, earnings, exchange rates, and non-household balances. Two of these categories at least plausibly coincide with extant knowledge: the Great Inflation which

[30] This uses Equation (8) to impose the restriction of no breaks in the loadings. For more details, see the Supplementary Material.

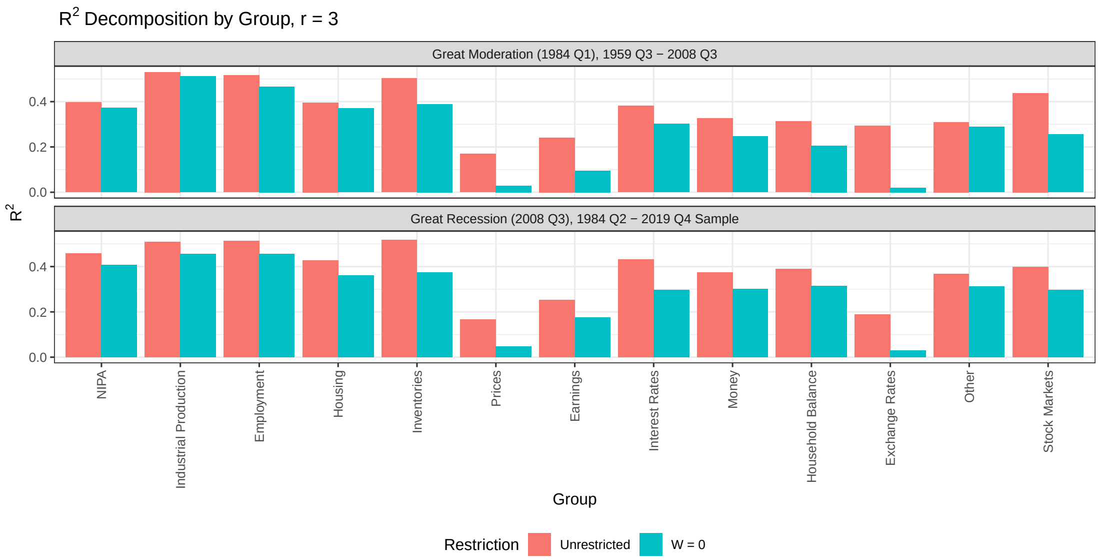

Figure 1: R^2 Statistics for unrestricted and restricted common component ($W = \mathbf{0}$) for Great Moderation Subsample, and Great Recession Subsample, for $r = 3$.

preceded the Great Moderation and affected prices; and the collapse of the Bretton Woods system in the mid-1970s which affected exchange rate variables. For the Great Recession break, while it appears that breaks in the loadings were important for many variable categories, they appear important for financial variables in categories such as exchange rate, money, and the stock market. Additionally, it appears that breaks in loadings were important for prices, also documented by Stock and Watson (2012).

6 Conclusion

The existing literature on structural breaks in factor models by and large does not distinguish between breaks between the factor variance and loadings, due to the need of a normalization during estimation. We argue it is important to distinguish them, as both can lead to different economic interpretations. To address this, we develop a projection-based decomposition of *any* structural break into a rotational and orthogonal shift component, which are naturally interpreted as a change in factor variance and loadings respectively. The estimators are simple to calculate and lead to two easy-to-implement Wald tests to disentangle structural breaks in the factor variance and loadings. Their finite sample performance is confirmed by a Monte Carlo study. Applying our procedure to U.S. macroeconomic data we find strong evidence of both types of breaks associated with the Great Moderation and Great Recession. Our projection-based decomposition allows us to estimate that the Great Moderation is associated with an over 60% reduction in the factor variance, a result precluded *a priori* if the break is not disentangled, and thus highlights the importance of doing so. Although our framework cannot distinguish between breaks in the factor dynamics and innovations and their respective "Good Policy" and "Good Luck" interpretations within the Great Moderation literature, rejection of our Z-test and the conclusion of break in the factor variance is a necessary condition for subsequent discussion regarding the factors, a conclusion which cannot be established with existing tools in the literature.

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

References

Aastveit, K. A., Bjørnland, H. C., and Thorsrud, L. A. (2015). What drives oil prices? Emerging versus developed economies. *Journal of Applied Econometrics*, 30(7):1013–1028.

Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue Ratio Test for the Number of Factors. *Econometrica*, 81(3):1203–1227.

Alessi, L. and Kerssenfischer, M. (2019). The response of asset prices to monetary policy shocks: Stronger than thought. *Journal of Applied Econometrics*, 34(5):661–672.

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

Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica*, 71(1):135–171.

Bai, J., Duan, J., and Han, X. (2022). Likelihood ratio test for structural changes in factor models. Unpublished manuscript.

Bai, J. and Han, X. (2016). Structural Changes in High Dimensional Factor Models. *Frontiers of Economics in China*, 11(1):9–39.

Bai, J. and Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 70(1):191–221.

Bai, J. and Ng, S. (2006). Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions. *Econometrica*, 74(4):1133–1150.

Bai, J. and Wang, P. (2016). Econometric Analysis of Large Factor Models. *Annual Review of Economics*, 8(1):53–80.

Baltagi, B. H., Kao, C., and Wang, F. (2017). Identification and estimation of a large factor model with structural instability. *Journal of Econometrics*, 197(1):87–100.

Baltagi, B. H., Kao, C., and Wang, F. (2021). Estimating and testing high dimensional factor models with multiple structural changes. *Journal of Econometrics*, 220(2):349–365.

Barigozzi, M., Cho, H., and Fryzlewicz, P. (2018). Simultaneous multiple change-point and factor analysis for high-dimensional time series. *Journal of Econometrics*, 206(1):187–225.

Barigozzi, M. and Luciani, M. (2023). Measuring the Output Gap using Large Datasets. *The Review of Economics and Statistics*, pages 1–15.

Bates, B. J., Plagborg-Møller, M., Stock, J. H., and Watson, M. W. (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, 177(2):289–304.

Bernanke, B. S., Boivin, J., and 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.

Breitung, J. and Eickmeier, S. (2011). Testing for structural breaks in dynamic factor models. *Journal of Econometrics*, 163(1):71–84.

Chen, L., Dolado, J. J., and Gonzalo, J. (2014). Detecting big structural breaks in large factor models. *Journal of Econometrics*, 180(1):30–48.

Cheng, X., Liao, Z., and Schorfheide, F. (2016). Shrinkage Estimation of High-Dimensional Factor Models with Structural Instabilities. *The Review of Economic Studies*, 83(4):1511–1543.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2):262–302.

Duan, J., Bai, J., and Han, X. (2022). Quasi-maximum likelihood estimation of break point in high-dimensional factor models. *Journal of Econometrics*, page S0304407622000379.

Han, X. (2015). Tests for overidentifying restrictions in Factor-Augmented VAR models. *Journal of Econometrics*, 184(2):394–419.

Han, X. and Inoue, A. (2015). Tests for Parameter Instability in Dynamic Factor Models. *Econometric Theory*, 31(5):1117–1152.

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

Lubik, T. A. and Schorfheide, F. (2004). Testing for Indeterminacy: An Application to U.S. Monetary Policy. *American Economic Review*, 94(1):190–217.

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*.

McCracken, M. W. and Ng, S. (2020). FRED-QD: A Quarterly Database for Macroeconomic Research. Technical report.

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

Ng, S. (2021). Modeling Macroeconomic Variations After COVID-19.

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

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

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

Phillips, P. C. B. and Moon, H. R. (1999). Linear Regression Limit Theory for Nonstationary Panel Data. *Econometrica*, 67(5):1057–1111.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *The Review of Economic Studies*, 72(3):821–852.

Sims, C. A. and Zha, T. (2006). Were There Regime Switches in U.S. Monetary Policy? *American Economic Review*, 96(1):54–81.

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

Stock, J. H. and Watson, M. W. (2005). Implications of Dynamic Factor Models for VAR Analysis. Working Paper 11467, National Bureau of Economic Research.

Stock, J. H. and Watson, M. W. (2009). Forecasting in Dynamic Factor Models Subject to Structural Instability. In *The Methodology and Practice of Econometrics: Festschrift in Honor of D.F. Hendry*, page 65.

Stock, J. H. and Watson, M. W. (2012). Disentangling the channels of the 2007-09 recession. *Brookings Papers on Economic Activity*, (1):81–135.

Stock, J. H. and Watson, M. W. (2016). Chapter 8 - Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2, pages 415–525. Elsevier.

Stock, J. H. and Watson, M. W. (2021). Comovement of Economic Activity During the Covid Recession. Unpublished manuscript.

Su, L. and Wang, X. (2017). On time-varying factor models: Estimation and testing. *Journal of Econometrics*, 198(1):84–101.

Wang, Y. and Liu, R. (2021). Identification and Estimation of Parameter Instability in a High Dimensional Approximate Factor Model. Unpublished manuscript.

Yamamoto, Y. and Tanaka, S. (2015). Testing for factor loading structural change under common breaks. *Journal of Econometrics*, 189(1):187–206.

A Asymptotic Proofs

A.1 Preliminary

First, recall that $V_{NT,1}, V_{NT,2}$ are the $r \times r$ diagonal matrices of the first r largest eigenvalues of the matrices $\frac{1}{T_1 N} X_1 X_1^\intercal$ and $\frac{1}{T_2 N} X_2 X_2^\intercal$ respectively. The estimated factor matrices $\tilde{F}_1, \tilde{F}_2$ are $\sqrt{T_1}, \sqrt{T_2}$ times the eigenvectors corresponding to the r largest eigenvalues of $X_1 X_1^\intercal$ and $X_2 X_2^\intercal$ respectively, we denote $T_1 = \lfloor \pi T \rfloor$ and $T_2 = T - \lfloor \pi T \rfloor$ for brevity. We therefore have for $m = 1, 2$

$$\frac{1}{NT_m} X_m X_m^\intercal \tilde{F}_m = \tilde{F}_m V_{NT,m} \tag{21}$$

Let $\delta_{NT} = \min \left\{ \sqrt{N}, \sqrt{T} \right\}$. Using $X_m = F_m \Lambda_m^\intercal + e_{(m)}$ gives:

$$\frac{1}{NT_m} \left(F_m \Lambda_m^\intercal \Lambda_m F_m^\intercal + F_m \Lambda_m^\intercal e_{(m)}^\intercal + e_{(m)} \Lambda_m f_m^\intercal + e_{(m)} e_{(m)}^\intercal \right) \tilde{F}_m V_{NT,m}^{-1} = \tilde{F}_m. \tag{22}$$

Using the fact that $H_m = (\Lambda_m^\intercal \Lambda_m / N)(F_m^\intercal \tilde{F}_m / T_m) V_{NT,m}^{-1}$ yields:

$$\tilde{F}_m - F_m H_m = \frac{1}{NT_m} \left(F_m \Lambda_m^\intercal e_{(m)}^\intercal \tilde{F}_m + e_{(m)} \Lambda_m F_m^\intercal \tilde{F}_m + e_{(m)} e_{(m)}^\intercal \right) V_{NT,m}^{-1}, \tag{23}$$

$$\begin{aligned} \tilde{f}_{m,t} - H_m^\intercal f_t = V_{NT,m}^{-1} \Bigg(& \frac{1}{T_m} \sum_{s=1}^{T} \tilde{f}_{m,s} \gamma_N(s,t) \iota_{mt} + \frac{1}{T_m} \sum_{s=1}^{T} \tilde{f}_{m,s} \zeta_{st} \iota_{mt} \\ & + \frac{1}{T_m} \sum_{s=1}^{T} \tilde{f}_{m,s} \eta_{m,st} \iota_{mt} + \frac{1}{T_m} \sum_{s=1}^{T} \tilde{f}_s \xi_{m,st} \iota_{mt} \Bigg) \end{aligned} \tag{24}$$

where $\zeta_{st} = \frac{e_s^\intercal e_t}{N} - \gamma_N(s,t), \eta_{m,st} = f_s^\intercal \Lambda_m^\intercal e_t / N$ and $\xi_{m,st} = f_t^\intercal \Lambda_m^\intercal e_s / N$.

We first present some lemmas from Bai (2003), stated for convenience.

Lemma 1. *For $m = 1, 2$:*

(a) Under Assumptions 1 to 4 and 8, $\frac{1}{T_m} \sum_{t=1}^{T} \left\| \left(\tilde{f}_{m,t} - H_m^\intercal f_t \right) \iota_{mt} \right\|^2 = O_p \left(\frac{1}{\delta_{NT}^2} \right)$.

(b) Under Assumptions 1 to 4 and 8, $\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\lambda}_{m,i} - H_m^{-\intercal} \lambda_{m,i} \right\| = O_p \left(\frac{1}{\delta_{NT}^2} \right)$.

(c) *Under Assumptions 1 to 6 and 8,* $\frac{1}{T_m}(\tilde{F}_m - F_m H_m)^\intercal e_{m,i} = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

(d) *Under Assumptions 1 to 6 and 8,* $\frac{1}{T_m}(\tilde{F}_m - F_m H_m)^\intercal F_m = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

(e) *Under Assumptions 1 to 4 and 8,* $\|H_m\| = O_p(1)$.

(f) *Under Assumptions 1 to 4, 7 and 8,* $\operatorname{plim}\left(\frac{\tilde{F}_m^\intercal F_m}{T_m}\right) = H_{0,m}^{-1}$.

Lemmas 1 (a) and 1 (c) to 1 (f) are the subsample counterparts of Lemmas A.1, B.1, B.2, B.3, A.3 and Proposition 1 of Bai (2003) respectively. Lemma 1 (b) follows from Lemma 1 (a) via symmetry.

A.2 Consistency Proofs

Proof of Theorem 3.1 (a). By the definition of $\tilde{Z}$, we have:

$$\begin{aligned}\tilde{Z} &= \left(\tilde{\Lambda}_1^\intercal \tilde{\Lambda}_1\right)^{-1} \tilde{\Lambda}_1^\intercal \tilde{\Lambda}_2 \\ &= \frac{1}{N} V_{NT,1}^{-1} \frac{1}{T_1} (\tilde{F}_1^\intercal X_1)^\intercal \frac{1}{T_2} (\tilde{F}_2^\intercal X_2),\end{aligned}$$

because for $m = 1, 2$, $\tilde{\Lambda}_m^\intercal \tilde{\Lambda}_m / N = V_{NT,m}$ by eigen-identity, and $\Lambda_m^\intercal = (\tilde{F}_m^\intercal \tilde{F})^\intercal \tilde{F}^\intercal X_m = \frac{1}{T_m} \tilde{F}_m^\intercal X_m$ via a least squares fit. Therefore

$$\begin{aligned}\tilde{Z} =& V_{NT,1}^{-1} \frac{1}{NT_1T_2} \left(\tilde{F}_1^\intercal F_1 \Lambda_1^\intercal + \tilde{F}_1^\intercal e_1\right) \left(\tilde{F}_2^\intercal F_2 Z^\intercal \Lambda_1^\intercal + \tilde{F}_2^\intercal F_2 W^\intercal + \tilde{F}_2^\intercal e_{(2)}\right)^\intercal \\ =& V_{NT,1}^{-1} \frac{1}{NT_1T_2} \left(\tilde{F}_1^\intercal F_1 \Lambda_1^\intercal e_{(2)}^\intercal \tilde{F}_2 + \tilde{F}_1^\intercal F_1 \Lambda_1^\intercal W F_2^\intercal \tilde{F}_2 + \tilde{F}_1^\intercal F_1 \Lambda_1^\intercal \Lambda_1 Z F_2^\intercal \tilde{F}_2 \right. \\ & \left. + \tilde{F}_1^\intercal e_{(1)} e_{(2)}^\intercal \tilde{F}_2 + \tilde{F}_1^\intercal e_{(1)} W F_2^\intercal \tilde{F}_2 + \tilde{F}_1^\intercal e_{(1)} \Lambda_1 Z F_2^\intercal \tilde{F}_2\right) \\ =& V_{NT,1}^{-1} (Z.I + Z.II + Z.III + Z.IV + Z.V + Z.VI).\end{aligned}$$

We shall see that $Z.III$ is the dominating term, and $Z.I, Z.II, Z.IV, Z.V, Z.VI$ are all asymptotically negligible.

Lemma 2. *Under Assumptions 1 to 6 and 8*

(a) $Z.I = O_p\left(\frac{1}{\delta_{NT}^2}\right)$,

(b) $Z.II = O_p\left(\frac{1}{N}\right)$,

(c) $Z.IV = O_p\left(\frac{1}{\delta_{NT}^2}\right)$,

(d) $Z.V = O_p\left(\frac{1}{\delta_{NT}^2}\right)$,

(e) $Z.VI = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

Proof of Lemma 2 (a):

$$\begin{aligned} Z.I &= \frac{\tilde{F}_1^\intercal F_1 \Lambda_1^\intercal e_{(2)}^\intercal \tilde{F}_2}{NT_1T_2} \\ &= \frac{\tilde{F}_1^\intercal F_1 \Lambda_1^\intercal e_{(2)}^\intercal (\tilde{F}_2 - F_2H_2)}{T_1NT_2} + \frac{\tilde{F}_1^\intercal F_1 \Lambda_1^\intercal e_{(2)}^\intercal F_2 H_2}{T_1NT_2} \\ &\leq \left\| \frac{\tilde{F}_1^\intercal F_1}{T_1} \right\| \left\| \frac{\Lambda_1^\intercal e_{(2)}^\intercal}{N\sqrt{T_2}} \right\| \left\| \frac{\tilde{F}_2 - F_2H_2}{\sqrt{T_2}} \right\| + \left\| \frac{\tilde{F}_1^\intercal F_1}{T_1} \right\| \left\| \frac{\Lambda_1^\intercal e_{(2)}^\intercal F_2}{NT_2} \right\| \|H_2\| \\ &= O_p(1) O_p\left(\frac{1}{\sqrt{N}}\right) O_p\left(\frac{1}{\delta_{NT}}\right) + O_p(1) O_p\left(\frac{1}{\delta_{NT}^2}\right) O_p(1) \\ &= O_p\left(\frac{1}{\delta_{NT}^2}\right), \end{aligned}$$

because of Lemmas 1 (d) to 1 (f) and Assumptions 1, 3 and 6 respectively.

Proof of Lemma 2 (b):

$$Z.II = \frac{\tilde{F}_1^\intercal F_1}{T_1} \frac{\Lambda_1^\intercal W}{N} \frac{F_2^\intercal \tilde{F}_2}{T_2} = O_p\left(\frac{1}{N}\right),$$

because $\frac{\Lambda_1^\intercal W}{N} = O_p\left(\frac{1}{N}\right)$. Optionally, if one is willing to assume strict orthogonality in finite sample, then $\Lambda_1^\intercal W = 0$, and it follows that $Z.II = 0$.

Proof of Lemma 2 (c)

$$\begin{aligned}
Z.IV &= \frac{\tilde{F}_1^\intercal e_{(1)} e_{(2)}^\intercal \tilde{F}_2}{NT_1T_2} \\
&= \frac{(\tilde{F}_1 - F_1H_1)^\intercal}{T_1} \frac{e_{(1)}e_{(2)}^\intercal}{N} \frac{(\tilde{F}_2 - F_2H_2)}{T_2} + \frac{(F_1H_1)^\intercal}{T_1} \frac{e_{(1)}e_{(2)}^\intercal}{N} \frac{(\tilde{F}_2 - F_2H_2)}{T_2} + \\
&\quad \frac{(\tilde{F}_1 - F_1H_1)^\intercal}{T_1} \frac{e_{(1)}e_{(2)}^\intercal}{N} \frac{(F_2H_2)}{T_2} + \frac{(F_1H_1)^\intercal}{T_1} \frac{e_{(1)}e_{(2)}^\intercal}{N} \frac{(F_2H_2)}{T_2} \\
&= Z.IV.a + Z.IV.b + Z.IV.c + Z.IV.d.
\end{aligned}$$

Analysing each of the four terms above, we have:

1.

$$\begin{aligned}
\|Z.IV.a\| &\leq \left\| \frac{(\tilde{F}_1 - F_1H_1)}{\sqrt{T_1}} \right\| \left\| \frac{e_{(1)}e_{(2)}^\intercal}{\sqrt{T_1}\sqrt{T_2}N} \right\| \left\| \frac{(\tilde{F}_2 - F_2H_2)}{\sqrt{T_2}} \right\| \\
&= O_p\left(\frac{1}{\delta_{NT}}\right) O_p\left(\frac{1}{\delta_{NT}}\right) O_p\left(\frac{1}{\delta_{NT}}\right) = O_p\left(\frac{1}{\delta_{NT}^3}\right),
\end{aligned}$$

by Lemma 1 (a) and Assumption 3,

2.

$$\begin{aligned}
\|Z.IV.b\| &\leq \left\| \frac{(F_1H_1)}{\sqrt{T_1}} \right\| \left\| \frac{e_{(1)}e_{(2)}^\intercal}{\sqrt{T_1}\sqrt{T_2}N} \right\| \left\| \frac{(\tilde{F}_2 - F_2H_2)}{\sqrt{T_2}} \right\| \\
&= O_p(1)\, O_p\left(\frac{1}{\delta_{NT}}\right) O_p\left(\frac{1}{\delta_{NT}}\right) = O_p\left(\frac{1}{\delta_{NT}^2}\right),
\end{aligned}$$

by Assumptions 1 and 3 and lemmas 1 (a) and 1 (e),

3.

$$\begin{aligned}
\|Z.IV.c\| &\leq \left\| \frac{(\tilde{F}_1 - F_1H_1)}{\sqrt{T_1}} \right\| \left\| \frac{e_{(1)}e_{(2)}^\intercal}{\sqrt{T_1}\sqrt{T_2}N} \right\| \left\| \frac{(F_2H_2)}{\sqrt{T_2}} \right\| \\
&= O_p\left(\frac{1}{\delta_{NT}}\right) O_p\left(\frac{1}{\delta_{NT}}\right) O_p(1) = O_p\left(\frac{1}{\delta_{NT}^2}\right),
\end{aligned}$$

by Lemmas 1 (a) and 1 (e) and Assumptions 1 and 3,

4.

$$
\begin{aligned}
\|Z.IV.d\| &\leq \|H_1\| \left\| \frac{F_1^\intercal e_{(2)}^\intercal}{T_1\sqrt{N}} \right\| \left\| \frac{e^\intercal F_2}{T_2\sqrt{N}} \right\| \|H_2\| \\
&= O_p\left(1\right) O_p\left(\frac{1}{\sqrt{T}}\right) O_p\left(\frac{1}{\sqrt{T}}\right) O_p\left(1\right) = O_p\left(\frac{1}{\delta_{NT}^2}\right),
\end{aligned}
$$

by Lemmas 1 (e) to 1 (f) and Assumption 6. Therefore, $Z.IV = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

Proof of Lemma 2 (d):

$$
\begin{aligned}
Z.V &= \frac{\tilde{F}_1^\intercal}{T_1} \frac{e_{(1)}W}{N} \frac{F_2^\intercal \tilde{F}_2}{T_2} \\
&\leq \left\| \frac{(\tilde{F}_1 - F_1 H_1)}{\sqrt{T_1}} \right\| \left\| \frac{e_{(1)}W}{N\sqrt{T_1}} \right\| \left\| \frac{F_2^\intercal \tilde{F}_2}{T_2} \right\| + \|H\| \left\| \frac{F_1^\intercal e_{(1)}W}{T_1 N} \right\| \left\| \frac{F_2^\intercal \tilde{F}_2}{T_2} \right\| \\
&= O_p\left(\frac{1}{\delta_{NT}}\right) O_p\left(\frac{1}{\sqrt{N}}\right) O_p\left(1\right) + O_p\left(1\right) O_p\left(\frac{1}{\delta_{NT}^2}\right) O_p\left(1\right) \\
&= O_p\left(\frac{1}{\delta_{NT}^2}\right),
\end{aligned}
$$

because of Lemmas 1 (a), 1 (e) and 1 (f) and Assumption 3. Note that $\left\| \frac{e_{(1)}W}{N\sqrt{T_1}} \right\| = O_p\left(\frac{1}{\sqrt{N}}\right)$ is implied by Assumption 2, because $\|W\| = \|\Lambda_2 - \Lambda_1 Z\| \leq \|\Lambda_2\|$.

Proof of Lemma 2 (e):

$$
\begin{aligned}
Z.VI &= \frac{\tilde{F}_1^\intercal}{T_1} \frac{e_{(1)}\Lambda_1 Z}{N} \frac{F_2^\intercal \tilde{F}_2}{T_2} \\
&\leq \left\| \frac{(\tilde{F}_1 - F_1 H_1)}{\sqrt{T_1}} \right\| \left\| \frac{e_{(1)}\Lambda_1}{N\sqrt{T_1}} \right\| \|Z\| \left\| \frac{F_2^\intercal \tilde{F}_2}{T_2} \right\| + \|H\| \left\| \frac{F_1^\intercal e_{(1)}\Lambda_1}{T_1 N} \right\| \|Z\| \left\| \frac{F_2^\intercal \tilde{F}_2}{T_2} \right\| \\
&= O_p\left(\frac{1}{\delta_{NT}}\right) O_p\left(\frac{1}{\sqrt{N}}\right) O_p\left(1\right) O_p\left(1\right) + O_p\left(1\right) O_p\left(\frac{1}{\delta_{NT}^2}\right) O_p\left(1\right) O_p\left(1\right) \\
&= O_p\left(\frac{1}{\delta_{NT}^2}\right),
\end{aligned}
$$

because of Lemmas 1 (a), 1 (e) and 1 (f) and Assumption 3, and because $\|Z\| < \infty$ is implied by Assumption 2.

Therefore, combining the terms above together, we have:

$$\begin{aligned}
\tilde{Z} &= V_{NT}^{-1}\left(\frac{\tilde{F}_1^\intercal F_1}{T_1}\right)\left(\frac{\Lambda_1^\intercal \Lambda_1}{N}\right) Z \left(\frac{F_2^\intercal \tilde{F}_2}{T_2}\right) + O_p\left(\frac{1}{\delta_{NT}^2}\right) + O_p\left(\frac{1}{\sqrt{N}}\right) \\
&= H_1^\intercal Z H_2^{-\intercal} + O_p\left(\frac{1}{\delta_{NT}^2}\right),
\end{aligned}$$

where the last line follows from the definition of H_1, and the fact that

$$\begin{aligned}
F_2 H_2 + \tilde{F}_2 - F_2 H_2 &= \tilde{F}_2 \\
\frac{1}{T_2}\tilde{F}_2^\intercal F_2 H_2 + \frac{1}{T_2}\tilde{F}_2^\intercal\left(\tilde{F}_2 - F_2 H_2\right) &= I_r \\
\frac{1}{T_2}\tilde{F}_2^\intercal F_2 H_2 &= I_r + O_p\left(\frac{1}{\delta_{NT}^2}\right) \\
\frac{1}{T_2} F_2^\intercal \tilde{F}_2 &= H_2^{-\intercal} + O_p\left(\frac{1}{\delta_{NT}^2}\right).
\end{aligned}$$

■

The consistency results for $\tilde{Z}$ allows us to extend Lemma 1 to the case of the rotated factors in the second subsample.

Lemma 3. *(a) Under Assumptions 1 to 4 and 8,* $\frac{1}{T_2}\left\|\tilde{F}_2\tilde{Z}^\intercal - F_2 Z^\intercal H_1)\right\|^2 = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

(b) Under Assumptions 1 to 6 and 8, $\frac{1}{T}(\tilde{F}_2\tilde{Z}^\intercal - F_2 Z^\intercal H_1)^\intercal e_{(1),i} = O_p\left(\frac{1}{\delta_{NT}^2\sqrt{T}}\right)$,

(c) Under Assumptions 1 to 6 and 8, $\frac{1}{T}(\tilde{F}_2\tilde{Z}^\intercal - F_2 Z^\intercal H_1)^\intercal F_2 = O_p\left(\frac{1}{\delta_{NT}^2}\right)$.

Proof of Lemma 3. For part a), we have

$$\begin{aligned}
\frac{1}{\sqrt{T_2}}\left\|\tilde{F}_2\tilde{Z}^\intercal - F_2 Z^\intercal H_1\right\| &= \frac{1}{\sqrt{T_2}}\left\|\tilde{F}_2\left(\tilde{Z} - H_1^\intercal Z H_2^{-\intercal}\right)^\intercal + \left(\tilde{F}_2 - F_2 H_2\right)\left(H_1^\intercal Z H_2^{-\intercal}\right)^\intercal - F_2 Z^\intercal H_1\right\| \\
&= \left(O_p\left(\frac{1}{\delta_{NT}^2}\right) + O_p\left(\frac{1}{\delta_{NT}}\right)\right) = O_p\left(\frac{1}{\delta_{NT}}\right),
\end{aligned}$$

where the last line follows by Theorem 3.1 (a) and Lemma 1 (a). Squaring both sides proves the result.

For part b), we have

$$
\begin{aligned}
&\frac{1}{T}\left(\tilde{F}_2\tilde{Z}^\intercal - F_2 Z^\intercal H_1\right)^\intercal e_{(2),i}\\
=&\frac{1}{T}\left[\tilde{F}_2\left(\tilde{Z}^\intercal - H_2^{-1}Z^\intercal H_1\right) + \left(\tilde{F}_2 - F_2H_2\right)H_2^{-1}Z^\intercal H_1\right]^\intercal e_{(2),i}\\
=&\frac{1}{T}\left(\tilde{Z}^\intercal - H_2^{-1}Z^\intercal H_1\right)^\intercal \tilde{F}_2^\intercal e_{(2),i} + \frac{1}{T}H_1^\intercal Z H_2^{-\intercal}(\tilde{F}_2 - F_2H_2)^\intercal e_{(2),i}\\
=&\left(\tilde{Z}^\intercal - H_2^{-1}Z^\intercal H_1\right)^\intercal\left[\frac{(\tilde{F}_2 - F_2H_2)^\intercal e_{(2),i}}{T} + \frac{F_2^\intercal e_{(2),i}}{T}\right] + \frac{1}{T}H_1^\intercal Z H_2^{-\intercal}(\tilde{F}_2 - F_2H_2)^\intercal e_{(2),i}\\
=&\left[O_p\left(\frac{1}{\delta_{NT}^2}\right)\right]\left[O_p\left(\frac{1}{\delta_{NT}^2}\right) + O_p\left(\frac{1}{\sqrt{T}}\right)\right]\\
=&O_p\left(\frac{1}{\delta_{NT}^2}\right).
\end{aligned}
$$

For part c), we have

$$
\begin{aligned}
&\frac{1}{T}\left(\tilde{F}_2\tilde{Z}^\intercal F_2 Z^\intercal H_1\right)^\intercal F_2\\
=&\frac{1}{T}\left[\tilde{F}_2\left(\tilde{Z}^\intercal - H_2^{-1}Z^\intercal H_1\right) + (\tilde{F}_2 - F_2)H_2 Z^{-1}Z^\intercal H_1\right]^\intercal F_2\\
=&\frac{1}{T}(\tilde{Z}^\intercal - H_2^{-1}Z^\intercal H_1)^\intercal \tilde{F}_2^\intercal F_2 + \frac{1}{T}H_1^\intercal Z H_2^{-\intercal}\left(\tilde{F}_2 - F_2H_2\right)^\intercal F_2\\
=&O_p\left(\frac{1}{\delta_{NT}^2}\right).
\end{aligned}
$$

■

Before we prove Theorem 3.1 (b), we need the following lemmas.

Lemma 4. *Under Assumptions 1 to 5 and 8*

$$
\frac{1}{\sqrt{N}}\left\|\tilde{\Lambda}_1\tilde{Z} - \Lambda_1 Z H_2^{-\intercal}\right\| = O_p\left(\frac{1}{\delta_{NT}}\right)
$$

Proof of Lemma 4.

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\left\|\tilde{\Lambda}_1\tilde{Z}-\Lambda_1 Z H_2^{-\intercal}\right\| \\
=&\frac{1}{\sqrt{N}}\left\|(\Lambda_1-\Lambda_1 H_1^{-\intercal})\tilde{Z}+\Lambda_1 H_1^{-\intercal}\tilde{Z}-\Lambda_1 Z H_2^{-\intercal}\right\| \\
=&\frac{1}{\sqrt{N}}\left\|(\tilde{\Lambda}_1-\Lambda_1 H_1^{-\intercal})\tilde{Z}+\Lambda_1 H_1^{-\intercal}\left(\tilde{Z}-H_1^{\intercal} Z H_2^{-\intercal}\right)+\ \Lambda_1 H_1^{-\intercal} H_1^{\intercal} Z H_2^{-\intercal}-\Lambda_1 Z H_2^{-\intercal}\right\| \\
\leq&\frac{1}{\sqrt{N}}\left\|(\tilde{\Lambda}_1-\Lambda_1 H_1^{-\intercal})\right\|\left\|\tilde{Z}\right\|+\frac{1}{\sqrt{N}}\left\|\Lambda_1 H_1^{-\intercal}\right\|\left\|\tilde{Z}-H_1^{\intercal} Z H_2^{-\intercal}\right\| \\
=&O_p\left(\frac{1}{\delta_{NT}}\right)O_p\left(1\right)+O_p\left(1\right)\left[O_p\left(\frac{1}{\delta_{NT}^2}\right)\right] \\
=&O_p\left(\frac{1}{\delta_{NT}}\right),
\end{aligned}
$$

where the second last line follows because the first term is $O_p\left(\frac{1}{\delta_{NT}}\right)$ by Lemma 1 (a) via symmetry, and the second term is $O_p\left(\frac{1}{\delta_{NT}^2}\right)$ by Theorem 3.1 (a). ■

Proof of Theorem 3.1 (b). Next, we prove the consistency of $\tilde{W}$. First, recall that

$$\tilde{W}=\tilde{\Lambda}_2-\tilde{\Lambda}_1\tilde{Z}, \tag{25}$$

$$W=\Lambda_2-\Lambda_1 Z, \tag{26}$$

which we can rearrange to form

$$\tilde{W}-WH_2^{-\intercal}=\tilde{\Lambda}_2-\Lambda_2 H_2^{-\intercal}-(\tilde{\Lambda}_1\tilde{Z}-\Lambda_1 Z H_2^{-\intercal}) \tag{27}$$

Taking the norm of both sides, and dividing by $\sqrt{N}$, we have

$$\begin{aligned}
\frac{1}{\sqrt{N}}\left\|\tilde{W} - WH_2^{-\intercal}\right\| &= \frac{1}{\sqrt{N}}\left\|\tilde{\Lambda}_2 - \Lambda_2 H_2^{-\intercal} - (\tilde{\Lambda}_1\tilde{Z} - \Lambda_1 Z H_2^{-\intercal})\right\| \\
&\leq \frac{1}{\sqrt{N}}\left(\left\|\tilde{\Lambda}_2 - \Lambda_2 H_2^{-\intercal}\right\| + \left\|(\tilde{\Lambda}_1\tilde{Z} - \Lambda_1 Z H_2^{-\intercal})\right\|\right) \\
&\leq \frac{1}{\sqrt{N}}\left\|\tilde{\Lambda}_2 - \Lambda_2 H_2^{-\intercal}\right\| + \frac{1}{\sqrt{N}}\left\|\tilde{\Lambda}_1\tilde{Z} - \Lambda_1 Z H_2^{-\intercal}\right\| \\
&= O_p\left(\frac{1}{\delta_{NT}}\right),
\end{aligned}$$

where the last line follows by Lemma 4. Squaring both sides results in the usual $O_p\left(\frac{1}{\delta_{NT}^2}\right)$ rate, mirroring the result of Bai (2003). ■

Remark 3. *We detail how there exists an alternative observationally equivalent parameterization of the rotation matrix H_2, and how this ultimately does not matter.*

Recall that we originally define $H_2 = \left(\frac{Z^\intercal\Lambda_1^\intercal\Lambda_1 Z}{N} + \frac{W^\intercal W}{N}\right)\left(\frac{F_2^\intercal\tilde{F}_2}{T_2}\right)V_{NT,2}^{-1}$, and this parameterises all of the change in terms of the loadings, and is what the literature at large does (see Baltagi et al., 2017; Han and Inoue, 2015). It is also possible parameterize the rotational change explicitly as part of the factors by defining:

$$\begin{aligned}
H_2^\dagger &= \left(\frac{\Lambda_1^\intercal\Lambda_1}{N} + \frac{Z^{-\intercal}W^\intercal W Z^{-1}}{N}\right)\left(\frac{ZF_2^\intercal\tilde{F}_2}{T_2}\right)V_{NT,2}^{-1} \\
&= \left(\frac{\Lambda_1^\intercal\Lambda_1}{N} + \frac{Z^{-\intercal}W^\intercal W Z^{-1}}{N}\right)\left(\frac{G_2^\intercal\tilde{F}_2}{T_2}\right)V_{NT,2}^{-1} \qquad (28)
\end{aligned}$$

where $G_2 = F_2 Z^\intercal$, the pseudo factor representation.

With $H_2^\dagger$, the consistency result $\tilde{Z}$ in Theorem 3.1 (a) changes to:

$$\left\|\tilde{Z} - H_1^\intercal H_2^{\dagger-\intercal}\right\| = O_p\left(\frac{1}{\delta_{NT}^2}\right) + O_p\left(\frac{1}{\sqrt{N}}\right) \qquad (29)$$

where the Z is absorbed into $H_2^\dagger$. Note that this does not affect following results, because we can simply replace $\tilde{F}_2 - F_2 H_2$ with $\tilde{F}_2 - F_2 Z^\intercal H_2^\dagger$ in all the proofs. Doing so, we get the *same results for $\tilde{F}_2\tilde{Z}$. Thus, it does not matter which parameterization of H_2 we use.*

Next, present some lemmas necessary for the proofs of the Z and W-tests.

Lemma 5. *For $m = 1, 2$:*

(a) under Assumptions 1 to 4 and 8,

$$\frac{1}{T_m}\sum_{t=1}^{T}\left\|(\tilde{f}_{m,t} - H_m^\intercal f_{m,t})\iota_{mt}\right\|^4 = O_p\left(\frac{1}{\delta_{NT}^4}\right),$$
$$\frac{1}{T_2}\sum_{t=\lfloor \pi T+1\rfloor}^{T}\left\|(\tilde{Z}\tilde{f}_{2,t} - H_1^\intercal Z f_t)\right\|^4 = O_p\left(\frac{1}{\delta_{NT}^4}\right),$$

(b) under Assumptions 1 to 4 and 8, $\frac{1}{T_m}\sum_{t=1}^{T}\left\|\tilde{f}_{m,t}\iota_{mt}\right\|^4 = O_p(1)$ and $\frac{1}{T_2}\sum_{t=\lfloor \pi T+1\rfloor}^{T}\left\|\tilde{Z}\tilde{f}_{2,t}\right\|^4 = O_p(1)$,

(c) under Assumptions 1 to 8, $\|H_m - H_{m,0}\| = O_p\left(\frac{1}{\delta_{NT}}\right)$.

(d) under Assumptions 1 to 4 and 8 $\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\lambda}_{m,i} - H_m^{-1}\lambda_{m,i}\right\|^4 = O_p\left(\frac{1}{\delta_{NT}^4}\right)$

Proof of Lemma 5. Lemmas 5 (a) to 5 (c) are just the subsample counterparts of Lemmas 5.i), 5.ii), 5.iii) and 6 of Han and Inoue (2015), but with the addition of the use of the rotated set of factors.

The second part of Lemma 5 (a) follows by

$$\begin{aligned}
&\frac{1}{T_2}\sum_{t=\lfloor \pi T+1\rfloor}^{T}\left\|(\tilde{Z}\tilde{f}_{2,t} - H_1^\intercal Z f_t)\right\|^4 \\
=&\frac{1}{T_2}\sum_{t=\lfloor \pi T+1\rfloor}^{T}\left\|\tilde{Z}\left(\tilde{f}_{2,t} - H_2^\intercal f_t\right) + \left(\tilde{Z} - H_1^\intercal Z H_2^{-\intercal} H_2^\intercal f_t\right)\right\|^4 \\
\leq&\left\|\tilde{Z}\right\|^4\frac{1}{T_2}\sum_{t=\lfloor \pi T+1\rfloor}^{T}\left\|\tilde{f}_{2,t} - H_2^\intercal f_t\right\|^4 + \left\|\tilde{Z} - H_1^\intercal Z H_2^{-\intercal}\right\|\frac{1}{T_2}\sum_{t=\lfloor \pi T+1\rfloor}^{T}\|H_2^\intercal f_t\|^4 \\
=&O_p\left(\frac{1}{\delta_{NT}^4}\right) + \left[O_p\left(\frac{1}{\delta_{NT}^2}\right) + O_p\left(\frac{1}{\sqrt{N}}\right)\right]^4 \\
=&O_p\left(\frac{1}{\delta_{NT}^4}\right),
\end{aligned}$$

using Assumption 1, Lemma 1 (a), and Theorem 3.1 (a).

The second part of Lemma 5 (b) follows by the fact that $\tilde{Z} = O_p(1)$.

Lemma 5 (d) follows by symmetry from Lemma 5 (a).

■

A.3 Z test Proofs

Define the long run variance estimate of the $A_Z(\pi, \widehat{F})$ 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 of the variance from pre and post break data ($m = 1, 2$ respectively)

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

$$\widehat{\Gamma}_{(1),j}(\pi, \widehat{F}) = \frac{1}{T_1}\sum_{t=j+1}^{T_1} vech(\widehat{f}_t\widehat{f}_t^\intercal - I_r)vech(\widehat{f}_t\widehat{f}_t^\intercal - I_r)^\intercal, \tag{30}$$

$$\widehat{\Gamma}_{(2),j}(\pi, \widehat{F}) = \frac{1}{T_2}\sum_{t=j+T_1+1}^{T} vech(\widehat{f}_t\widehat{f}_t^\intercal - I_r)vech(\widehat{f}_t\widehat{f}_t^\intercal - I_r)^\intercal, \tag{31}$$

where $\mathbf{k}(.)$ 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.

The proof of Theorem 3.2 requires proving the following lemmas:

Lemma 6. *(a) Under Assumptions 1 to 8, if $\frac{\sqrt{T}}{N} \to \infty$, then*

$$\left\|A_Z(\pi, (\widehat{F})) - A_Z(\pi, FH_{0,1})\right\| \xrightarrow{p} 0.$$

(b) Under Assumptions 1 to 8, and if the conditions in Assumption 9 hold, then

$$\left\|\widehat{S}(\pi, \widehat{F}) - \widehat{S}(\pi, FH_{0,1})\right\| \xrightarrow{p} 0.$$

Proof of Lemma 6 (a). Taking the norm of $A_Z(\pi, \widehat{F}) - A_Z(\pi, FH_{0,1})$, we have

$$\left\| A_Z(\pi, \widehat{F}) - A_Z(\pi, FH_{0,1}) \right\| = \left\| vech\sqrt{T} \left(\frac{1}{\pi T} \sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^\intercal - \frac{1}{(1-\pi)T} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^\intercal \right) \right.$$
$$\left. - \, vech\sqrt{T} \left(\frac{1}{\pi T} \sum_{t=1}^{\lfloor \pi T \rfloor} H_{0,1}^\intercal f_t f_t^\intercal H_{0,1} - \frac{1}{(1-\pi)T} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} H_{0,1}^\intercal f_t f_t^\intercal H_{0,1} \right) \right\|.$$

Because $\widehat{f}_t$ is consistent for $H_1^\intercal f_t$, and H_1 is consistent for $H_{0,1}$, it suffices to prove that

$$A_Z.I = \sqrt{T} \left\| vech \left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^\intercal}{\pi T} - \frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^\intercal}{(1-\pi)T} \right) - vech \left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor} H_1^\intercal f_t f_t^\intercal H_1}{\pi T} - \frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T} H_1^\intercal f_t f_t^\intercal H_1}{(1-\pi)T} \right) \right\|,$$

and

$$A_Z.II = \sqrt{T} \left\| vech \left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor} H_1^\intercal f_t f_t^\intercal H_1}{\pi T} - \frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T} H_1^\intercal f_t f_t^\intercal H_1}{(1-\pi)T} \right) \right.$$
$$\left. - \, vech \left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor} H_{0,1}^\intercal f_t f_t^\intercal H_{0,1}}{\pi T} - \frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T} H_{0,1}^\intercal f_t f_t^\intercal H_{0,1}}{(1-\pi)T} \right) \right\|,$$

are both $o_p(1)$. The first term $A_Z.I$ is bounded by

$$A_Z.I \le \sqrt{T} \left\| vech \left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^\intercal - H_1^\intercal f_t f_t^\intercal H_1}{\pi T} \right) \right\| + \sqrt{T} \left\| vech \left(\frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^\intercal - H_1^\intercal f_t f_t^\intercal H_1}{(1-\pi)T} \right) \right\|.$$

We focus on proving that the first term is $o_p(1)$, as the second term can be proved very similarly. The first term is bounded by

$$
\begin{aligned}
&\sqrt{T}\left\| vech\left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}\widehat{f}_t(\widehat{f}_t^{\intercal}-H_1^{\intercal}f_t)^{\intercal}+\widehat{f}_t(H_1^{\intercal}f_t)-H_1^{\intercal}f_tf_t^{\intercal}H_1}{\pi T}\right)\right\| \\
=&\sqrt{T}\left\| vech\left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}\widehat{f}_t(\widehat{f}_t^{\intercal}-H_1^{\intercal}f_t)+(\widehat{f}_t^{\intercal}-f_tH_1)f_t^{\intercal}H_1}{\pi T}\right)\right\| \\
=&\sqrt{T}\left\| vech\left(\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}(\widehat{f}_t-H_1^{\intercal}f_t)(\widehat{f}_t^{\intercal}-f_t^{\intercal}H_1)+H_1^{\intercal}f_t(\widehat{f}_t^{\intercal}-f_t^{\intercal}H_1)+(\widehat{f}_t-H_1^{\intercal}f_t)f_t^{\intercal}H_1}{\pi T}\right)\right\| \\
\leq&\sqrt{T}\left(\left\|\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}(\widehat{f}_t-H_1^{\intercal}f_t)(\widehat{f}_t^{\intercal}-f_t^{\intercal}H_1)}{\pi T}\right\|+2\left\|\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}H_1^{\intercal}f_t(\widehat{f}_t^{\intercal}-f_t^{\intercal}H_1)}{\pi T}\right\|\right) \\
=&O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right)+O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right) \\
=&o_p(1)
\end{aligned}
$$

as $\frac{\sqrt{T}}{N}\to 0$, where each of the terms on the second last line are $O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right)$ by Lemma 1 (a). Next, $A_Z.II$ can be bounded by

$$
\begin{aligned}
&\sqrt{T}\left\|\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}H_1^{\intercal}f_tf_t^{\intercal}H_1-H_{0,1}^{\intercal}f_tf_t^{\intercal}H_{0,1}}{\pi T}-\frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T}H_1^{\intercal}f_tf_t^{\intercal}H_1-H_{0,1}^{\intercal}f_tf_t^{\intercal}H_{0,1}}{(1-\pi)T}\right\| \\
=&\sqrt{T}\left\|\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}(H_1-H_{0,1})^{\intercal}f_tf_t^{\intercal}H_1+H_{0,1}^{\intercal}f_tf_t^{\intercal}(H_1-H_{0,1})}{\pi T}\right. \\
&\left.-\frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T}(H_1-H_{0,1})^{\intercal}f_tf_t^{\intercal}H_1+H_{0,1}^{\intercal}f_tf_t^{\intercal}(H_1-H_{0,1})}{(1-\pi)T}\right\| \\
\leq&\sqrt{T}\|H_1-H_{0,1}\|\left\|\frac{\sum_{t=1}^{\lfloor \pi T \rfloor}f_tf_t^{\intercal}}{\pi T}-\frac{\sum_{t=\lfloor \pi T+1 \rfloor}^{T}f_tf_t^{\intercal}}{(1-\pi)T}\right\|(\|H_1\|+\|H_{0,1}\|) \\
=&\sqrt{T}o_p(1)O_p\left(\frac{1}{\sqrt{T}}\right)O_p\left(1\right),
\end{aligned}
$$

because we use Assumption 8 (b) and $\|H_1-H_{0,1}\|=o_p(1)$ which is implied by Assumptions 1 to 4 and 7 (see Bai (2003)). ■

To prove Lemma 6 (b), we need to prove the following lemmas first.

Lemma 7. *Under Assumptions 1 to 5, for $m = 1, 2$:*

$$\left\|\widehat{\Gamma}_{(m),j}(\pi, \widehat{F}) - \widehat{\Gamma}_{(m),j}(\pi, FH_1)\right\| = O_p\left(\frac{1}{\delta_{NT}}\right).$$

Proof of Lemma 7. We shall focus on the case of $m = 1$, as the case for $m = 2$ is analogous and thus omitted.

$$\begin{aligned}
&\left\|\widehat{\Gamma}_{(1),j}(\pi, \widehat{F}) - \widehat{\Gamma}_{(1),j}(\pi, FH_1)\right\| \\
\leq&\left\|\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor} vech(\widehat{f}_t\widehat{f}_t^\intercal - I_r)vech(\widehat{f}_{t-j}\widehat{f}_{t-j}^\intercal - I_r)^\intercal - vech(H_1^\intercal f_t f_t^\intercal H_1 - I_r)vech(H_1^\intercal f_{t-j} f_{t-j}^\intercal H_1 - I_r)^\intercal\right\| \\
\leq&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|vech(\widehat{f}_t\widehat{f}_t^\intercal - I_r)vech(\widehat{f}_{t-j}\widehat{f}_{t-j}^\intercal - H_1^\intercal f_{t-j} f_{t-j}^\intercal H_1)\right\| + \\
&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|vech(\widehat{f}_t\widehat{f}_t^\intercal - H_1^\intercal f_t f_t^\intercal H_1)vech(H_1^\intercal f_{t-j} f_{t-j}^\intercal H_1 - I_r)^\intercal\right\| \\
\leq&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_t\widehat{f}_t^\intercal\right\|\left\|\widehat{f}_{t-j}\widehat{f}_{t-j}^\intercal - H_1^\intercal f_{t-j} f_{t-j}^\intercal H_1\right\| + \\
&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor} r\left\|\widehat{f}_{t-j}\widehat{f}_{t-j}^\intercal - H_1^\intercal f_{t-j} f_{t-j}^\intercal H_1\right\| + \frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_t\widehat{f}_t^\intercal - H_1^\intercal f_t f_t^\intercal H_1\right\|\left\|H_1^\intercal f_{t-j} f_{t-j}^\intercal H_1\right\| \\
&+ \frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor} r\left\|\widehat{f}_j\widehat{f}_j^\intercal - H_1^\intercal f_t f_t^\intercal H_1\right\| \\
=&\Gamma.I + \Gamma.II + \Gamma.III + \Gamma.IV.
\end{aligned}$$

We proceed by bounding Term $\Gamma.I$:

$$
\begin{aligned}
\Gamma.I &= \frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_t\widehat{f}_t^{\intercal}\right\|\left\|\widehat{f}_{t-j}\widehat{f}_{t-j}^{\intercal}-H_1^{\intercal}f_{t-j}f_{t-j}^{\intercal}H_1\right\| \\
&\leq \frac{1}{\pi T}\left(\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_t\right\|^4\right)^{1/2}\left(\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_{t-j}(\widehat{f}_{t-j}^{\intercal}\widehat{f}_{t-j}^{\intercal}H_1)+(\widehat{f}_{t-j}-H_1^{\intercal}f_{t-j})f_{t-j}^{\intercal}H_1\right\|^2\right)^{1/2} \\
&\leq \left(\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_t\right\|^4\right)^{1/2}\left(\frac{2}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_{t-j}(\widehat{f}_{t-j}^{\intercal}-\widehat{f}_{t-j}^{\intercal}H_1)\right\|^2+\frac{2}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|(\widehat{f}_{t-j}-\widehat{f}_{t-j}H_1)f_{t-j}^{\intercal}H_1\right\|^2\right)^2 \\
&\leq \left(\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_t\right\|^4\right)^{1/2} \\
&\quad\times\left[\left(\frac{2}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_{t-j}\right\|^4\frac{2}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_{t-j}-f_{t-j}H_1\right\|^4\right)^{1/2}\right. \\
&\qquad\left.+\left(\frac{2}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_{t-j}-H_1^{\intercal}f_{t-j}\right\|^4\frac{2}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|f_{t-j}H_1\right\|^4\right)^{1/2}\right]^{1/2} \\
&= O_p\left(\frac{1}{\delta_{NT}}\right),
\end{aligned}
$$

under Assumptions 1 to 5, where $\frac{1}{T}\sum_{t=1}^{T}\left\|\widehat{f}_t\right\|^4 = O_p(1)$ by Lemma 5 (a). Using similar arguments, terms $\Gamma.II$, $\Gamma.III$, and $\Gamma.IV$ can be shown to be $O_p\left(T^{-1/4}\right)+O_p\left(\frac{1}{\sqrt{N}}\right)$. ■

Lemma 8. *Under Assumptions 1 to 5, for $m=1,2$,*

$$
\left\|\widehat{\Gamma}_{m,j}(\pi, F_mH_1)-\widehat{\Gamma}_{m,j}(\pi, F_mH_{0,1})\right\| = O_p\left(\frac{1}{\delta_{NT}}\right).
$$

Proof of Lemma 8. We shall only prove the lemma for $m=1$, because the proof for $m=2$

is similar and thus omitted.

$$
\begin{aligned}
&\left\|\widehat{\Gamma}_{1,j}(\pi, FH_1) - \widehat{\Gamma}_{1,j}(\pi, FH_{0,1})\right\| \\
=&\left\|\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\Big[vech(H_1^\intercal f_t f_t^\intercal H_1 - I_r)vech(H_1^\intercal f_{t-j}f_{t-j}^\intercal H_1 - I_r)\right. \\
&\quad \left. -vech(H_{0,1}^\intercal f_t f_t^\intercal H_{0,1} - I_r)vech(H_{0,1}^\intercal f_{t-j}f_{t-j}^\intercal H_{0,1} - I_r)^\intercal\Big]\right\| \\
\leq&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\|H_1^\intercal f_t f_t^\intercal H_1\|\left\|H_1^\intercal f_{t-j}f_{t-j}^\intercal H_1 - H_{0,1}^\intercal f_{t-j}f_{t-j}^\intercal H_{0,1}\right\| + \\
&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|H_1^\intercal f_{t-j}f_{t-j}^\intercal H_1 - H_{0,1}^\intercal f_{t-j}f_{t-j}^\intercal H_{0,1}\right\| + \\
&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|\widehat{f}_t\right\|^2\left\|H_1^\intercal f_t f_t^\intercal H_1 - H_{0,1}^\intercal f_t f_t^\intercal H_{0,1}\right\| + \\
&\frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\left\|H_1^\intercal f_t f_t^\intercal H_1 - H_{0,1}^\intercal f_t f_t^\intercal H_{0,1}\right\|\left\|H_{0,1}^\intercal f_{t-j}f_{t-j}^\intercal H_{0,1}\right\| \\
=&\Gamma.V + \Gamma.VI + \Gamma.VII + \Gamma.VIII.
\end{aligned}
$$

Term $\Gamma.V$ is bounded by:

$$
\begin{aligned}
\Gamma.V &= \frac{1}{\pi T}\sum_{t=j+1}^{\lfloor \pi T\rfloor}\|H_1^\intercal f_t f_t^\intercal H_1\|\left\|H_1^\intercal f_{t-j}f_{t-j}^\intercal H_1 - H_{0,1}^\intercal f_{t-j}f_{t-j}^\intercal H_{0,1}\right\| \\
&\leq \left(\frac{1}{\pi T}\sum_{t=1}^{T}\|f_t H_1\|^4\right)^{1/2}\left(\frac{1}{\pi T}(\|H_1\|^2 + \|H_{0,1}\|^2)\sum_{t=1}^{T}\|f_t\|^4\right)^{1/2}\|H_1 - H_{0,1}\| \\
&= O_p(1)\, O_p\left(\frac{1}{\delta_{NT}}\right)
\end{aligned}
$$

by Assumption 1 and Lemma 5 (c). The proofs of terms $\Gamma.VI, \Gamma.VII, \Gamma.VIII$ are similar and thus omitted. ■

Proof of Lemma 6 (b). It suffices to show that

$$
\left\|\widehat{\Omega}_{Z,m}(\pi, \widehat{F}) - \widehat{\Omega}_{Z,m}(\pi, FH_{0,1})\right\| \xrightarrow{p} 0, \quad \text{for} \quad m = 1, 2.
$$

For brevity, we will only prove the case for $m = 1$, as the case for $m = 2$ can be proved similarly. First, see that

$$\begin{aligned}\left\|\widehat{\Omega}_{Z,1}(\pi,\widehat{F})-\widehat{\Omega}_{Z,1}(\pi,FH_{0,1})\right\| \leq& \left\|\widehat{\Omega}_{Z,1}(\pi,\widehat{F})-\widehat{\Omega}_{Z,1}(\pi,FH_1)\right\|+\left\|\widehat{\Omega}_{Z,1}(\pi,\widehat{FH_1})-\widehat{\Omega}_{Z,1}(\pi,FH_{0,1})\right\| \\ =&\Omega_Z.I+\Omega_Z.II\end{aligned}$$

For term $\Omega_Z.I$, we have:

$$\begin{aligned}&\left\|\widehat{\Omega}_{Z,1}(\pi,\widehat{F})-\widehat{\Omega}_{Z,1}(\pi,FH_1)\right\| \\ \leq&\left\|\widehat{\Gamma}_{(1),0}(\pi,\widehat{F})+\sum_{t=1}^{\lfloor \pi T\rfloor}\mathbf{k}\left(\frac{j}{b_{\lfloor \pi T\rfloor}}\right)(\widehat{\Gamma}_{(1),j}(\pi,\widehat{F})+\widehat{\Gamma}_{(1),j}(\pi,\widehat{F}))^\intercal)\right. \\ &\left.-\widehat{\Gamma}_{(1),0}(\pi,FH_1))+\sum_{t=1}^{\lfloor \pi T\rfloor}\mathbf{k}\left(\frac{j}{b_{\lfloor \pi T\rfloor}}\right)(\widehat{\Gamma}_{(1),j}(\pi,FH_1)+\widehat{\Gamma}_{(1),j}(\pi,FH_1))^\intercal)\right\|.\end{aligned}$$

Recall that $\left|\mathbf{k}\left(\frac{j}{b_{\lfloor \pi T\rfloor}}\right)\right| \leq 1$ and $\mathbf{k}\left(\frac{j}{b_{\lfloor \pi T\rfloor}}\right)=0$ if $j > b_{\lfloor \pi T\rfloor}$ for the Bartlett kernel. Thus,

$$\Omega_Z.I \leq \left\|\widehat{\Gamma}_{1,0}(\pi,\widehat{F})-\widehat{\Gamma}_{1,0}(\pi,FH_1)\right\|+2\sum_{j=1}^{b_{\lfloor \pi T\rfloor}}\left\|\widehat{\Gamma}_{1,j}(\pi,\widehat{F})-\widehat{\Gamma}_{1,j}(\pi,FH_1)\right\|.$$

In the case of the Bartlett kernel, $\Omega_Z.I$ is $O_p\left(\frac{T^{1/3}}{\delta_{NT}}\right)$ by Lemma 8 and the condition in Assumption 9 (a) which states that $b_{\lfloor \pi T\rfloor} \leq KT^{1/3}$, so $\Omega_Z.I$ is $o_p(1)$ if $\frac{T^{2/3}}{N} \to 0$ as $N, T \to \infty$.

The term $\Omega_Z.II$ can be also be shown to be $o_p(1)$ with similar arguments. ■

To prove the consistency of the Wald test statistic in Theorem 3.2, we present the following lemmas:

Lemma 9. *Under Assumptions 1 to 9,*

(a) $\left\|\widehat{S}(\pi,FH_{0,1})^{-1}\right\|=O_p(1)$ *and* $\left\|\widehat{S}(\pi,\widehat{F})^{-1}\right\|=O_p(1)$,

(b) $\left\|\widehat{S}(\pi,FH_{0,1})^{-1}-\widehat{S}(\pi,\widehat{F})^{-1}\right\|=o_p(1)$.

Proof of Lemma 9. For Lemma 9 (a), because $0 < \pi < 1$, this implies that $\left\|\widehat{S}(\pi, FH_{0,1}) - \left(\frac{1}{\pi} + \frac{1}{1-\pi}\Omega\right)\right\| = o_p(1)$. Let ρ_{min}, ρ_{max} denote the minimum and maximum eigenvalues of a symmetric matrix respectively. Since Ω is positive definite, $\left|\rho_{min}(\widehat{S}(\pi, FH_{0,1})) - \rho_{min}\left(\left(\frac{1}{\pi} + \frac{1}{1-\pi}\right)\Omega\right)\right| \leq \left\|\widehat{S}(\pi, FH_{0,1})\left(\frac{1}{\pi} + \frac{1}{1-\pi}\right)\Omega\right\| = o_p(1)$. This means that the eigenvalues of $\widehat{S}(\pi, FH_{0,1})$ are bounded away from zero, so $\left\|\widehat{S}(\pi, FH_{0,1})\right\| = O_p(1)$. For the second part of this lemma, we have $\left\|\widehat{S}(\pi, \widehat{F}) - \left(\frac{1}{\pi} + \frac{1}{1-\pi}\right)\Omega\right\| \leq \left\|\widehat{S}(\pi, \widehat{F}) - \widehat{S}(\pi, FH_{0,1})\right\| + \left\|\widehat{S}(\pi, FH_{0,1}) - \left(\frac{1}{\pi} + \frac{1}{1-\pi}\right)\Omega\right\| = o_p(1)$ by Lemma 6 (b) and Assumption 9 (a). Therefore,

$$\left|\rho_{min}(\widehat{S}(\pi, \widehat{F})) - \rho_{min}\left(\left(\frac{1}{\pi} + \frac{1}{1-\pi}\right)\Omega\right)\right| \leq \left\|\widehat{S}(\pi, \widehat{F}) - \left(\frac{1}{\pi} + \frac{1}{1-\pi}\right)\Omega\right\| = o_p(1),$$

which means that the eigenvalues of $\widehat{S}(\pi, \widehat{F})$ are also bounded away from zero, which subsequently implies that $\widehat{S}(\pi, \widehat{F})^1 = O_p(1)$. For Lemma 9 (b):

$$\begin{aligned}\left\|\widehat{S}(\pi, \widehat{F})^{-1} - \widehat{S}(\pi, FH_{0,1})^{-1}\right\| =& \left\|\widehat{S}(\pi, FH_{0,1})^{-1}\left(\widehat{S}(\pi, FH_{0,1}) - \widehat{S}(\pi, \widehat{F})\right)\widehat{S}(\pi, \widehat{F})^{-1}\right\| \\ \leq& \left\|\widehat{S}(\pi, FH_{0,1})^{-1}\right\|\left\|\widehat{S}(\pi, FH_{0,1}) - \widehat{S}(\pi, \widehat{F})\right\|\left\|\widehat{S}(\pi, \widehat{F})^{-1}\right\| \\ =& O_p(1)\, o_p(1) O_p(1) = o_p(1),\end{aligned}$$

by Lemma 6 (b) and Assumption 9 (a). ■

Proof of Theorem 3.2.

$$\begin{aligned}\left|\mathscr{W}_Z(\pi, \widehat{F}) - \mathscr{W}_Z(\pi, FH_{0,1})\right| \leq& \left|A_Z(\pi, \widehat{F})^\intercal\left[\widehat{S}(\pi, \widehat{F})^{-1} - \widehat{S}(\pi, FH_{0,1})^{-1}\right]A_Z(\pi, \widehat{F})\right| \\ &+ \left|\left[A_Z(\pi, \widehat{F}) - A_Z(\pi, FH_{0,1})\right]^\intercal \widehat{S}(\pi, FH_{0,1})^{-1}A_Z(\pi, \widehat{F})\right| \\ &+ \left|A_Z(\pi, FH_{0,1})^\intercal\widehat{S}(\pi, FH_{0,1})^{-1}\left[A_Z(\pi, \widehat{F}) - A_Z(\pi, FH_{0,1})\right]\right| \\ =& o_p(1),\end{aligned}$$

using the results of Lemma 9 and Lemma 6 (a). ■

Proof of Theorem 3.3. Under the alternative hypothesis, $Z \neq I$, so we have:

$$
\begin{aligned}
&\frac{1}{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^{\intercal} - \frac{1}{T - \lfloor \pi T \rfloor} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^{\intercal} \\
=& \left(\frac{1}{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} H_1^{\intercal} f_t f_t^{\intercal} H_1 - \frac{1}{T - \lfloor \pi T \rfloor} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} H_1^{\intercal} Z f_t f_t^{\intercal} Z^{\intercal} H_1 \right) \\
&+ \frac{1}{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\widehat{f}_t \widehat{f}_t^{\intercal} - H_1^{\intercal} f_t f_t^{\intercal} H_1 \right) - \frac{1}{T - \lfloor \pi T \rfloor} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} \left(\widehat{f}_t \widehat{f}_t^{\intercal} - H_1^{\intercal} Z f_t f_t^{\intercal} Z^{\intercal} H_1 \right).
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\frac{1}{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\widehat{f}_t \widehat{f}_t^{\intercal} - H_1^{\intercal} f_t f_t^{\intercal} H_1 \right) \\
=& \frac{1}{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} \left[(\widehat{f}_t - H_1^{\intercal} f_t) f_t^{\intercal} H_1 + (\widehat{f}_t - H_1^{\intercal} f_t)(\widehat{f}_t^{\intercal} - f_t^{\intercal} H_1) + H_1^{\intercal} f_t (\widehat{f}_t^{\intercal} - f_t^{\intercal} H_1) \right] \\
=& O_p\left(\frac{1}{\delta_{NT}^2} \right),
\end{aligned}
$$

by the arguments in the proof of Lemma 6 (a). Similarly, $\frac{1}{T-\lfloor \pi T \rfloor} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} \left(\widehat{f}_t \widehat{f}_t^{\intercal} - H_1^{\intercal} Z f_t f_t^{\intercal} Z^{\intercal} H_1 \right) = O_p\left(\frac{1}{\delta_{NT}^2} \right)$.

Under the alternative hypothesis where there is a rotational break, we have:

$$
H_1^{\intercal} \left(\frac{1}{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} f_t f_t^{\intercal} - \frac{1}{T - \lfloor \pi T \rfloor} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} Z f_t f_t^{\intercal} Z^{\intercal} \right) H_1 \xrightarrow{p} H_{0,1}^{\intercal} (\Sigma_F - Z \Sigma_F Z^{\intercal}) H_{0,1} \equiv C,
$$

by Assumption 8, and the definitions of H_1 and $H_{0,1}$. Matrix C contains non-zero entries because $\Sigma_F - Z\Sigma_F Z^{\intercal}$ is not zero by Assumption 11, and the fact that $H_{0,1}$ is non-singular. Note that Assumptions 1 to 8 still hold under the alternative hypothesis, and hence Lemma 6 (b) still hold for the equivalent models under the alternative. Finally, putting the above together we have:

$$\begin{aligned}
\mathscr{W}_Z(\widehat{F}) &= A_Z(\pi,\widehat{F})^\intercal \widehat{S}(\pi,\widehat{F})^{-1} A_Z(\pi,\widehat{F}) \\
&= \frac{T}{\max(b_{\lfloor \pi T\rfloor}, b_{T-\lfloor \pi T\rfloor})}\left[\frac{1}{\sqrt{T}} A_Z(\pi,\widehat{F})^\intercal\right]\left[\max(b_{\lfloor \pi T\rfloor}, b_{T-\lfloor \pi T\rfloor})\tilde{S}(\pi,\widehat{F})^{-1}\right]\left[\frac{1}{\sqrt{T}} A_Z(\pi,\widehat{F})\right] \\
&= \frac{T}{\max(b_{\lfloor \pi T\rfloor}, b_{T-\lfloor \pi T\rfloor})}\left[vech(C)^\intercal + o_p(1)\right]\left[\max(b_{\lfloor \pi T\rfloor}, b_{T-\lfloor \pi T\rfloor})\tilde{S}(\pi,\widehat{F})^{-1}\right]\left[vech(C) + o_p(1)\right] \\
&\to \infty
\end{aligned}$$

by Assumptions 11 and 12. ■

A.4 W test Proofs

We first recall the following identity for the factor loadings for $m = 1, 2$:

$$\tilde{\lambda}_{m,i} - H_m^{-1}\lambda_{m,i} = \frac{1}{T_m} H_m^\intercal F_m^\intercal e_{(m),i} + \frac{1}{T_m}\tilde{F}^\intercal (F_m - \tilde{F}_m H_m^{-1})\lambda_{m,i} + \frac{1}{T_m}(\tilde{F}_m - F_m H_m)^\intercal e_{(m),i}, \tag{32}$$

where Equation (32) is the subsample version of the asymptotic expansion of the factor loadings considered by Bai (2003) (see the proof of their Theorem 2). The last two terms are both $O_p\left(\frac{1}{\delta_{NT}^2}\right)$ by Lemma 1 (d) and Lemma 1 (c), and therefore we have

Lemma 10. *Under Assumptions 1 to 6, for $m = 1, 2$,*

$$\tilde{\lambda}_{m,i} - H_m^{-1}\lambda_{m,i} = H_m^\intercal \frac{1}{T_m}\sum_{t=1}^{T} f_{m,t} e_{it} \iota_{mt} + O_p\left(\frac{1}{\delta_{NT}^2}\right) \tag{33}$$

for each i.

Lemma 10 is simply the subsample counterpart of Equation B.2 in Bai (2003).

Proof of Theorem 3.4 (a). Recall that $\tilde{W} = \tilde{\Lambda}_2 - \tilde{\Lambda}_1 \tilde{Z}$, which implies that

$$\begin{aligned} \tilde{\lambda}_{2,i} &= \tilde{Z}^{\intercal} \tilde{\lambda}_{1,i} + \tilde{w}_i, \\ \tilde{w}_i &= \tilde{\lambda}_{2,i} - \tilde{Z}^{\intercal} \tilde{\lambda}_{1,i}. \end{aligned} \tag{34}$$

Substituting in the decompositions in Lemma 10, we have

$$\begin{aligned} \tilde{\lambda}_{2,i} - H_2^{-1}\lambda_{2,i} &= H_2^{\intercal} \frac{1}{(1-\pi)T} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} f_t e_{it} + O_p(\delta_{NT}^{-2}) \\ (\tilde{Z}^{\intercal}\tilde{\lambda}_{1,i} + \tilde{w}_i) - H_2^{-1}(\lambda_{1,i} + w_i) &= H_2^{\intercal} \frac{1}{(1-\pi)T} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} f_t e_{it} + O_p(\delta_{NT}^{-2}) \\ (\tilde{w}_i - H_2^{-1} w_i) &= H_2^{\intercal} \frac{1}{(1-\pi)T} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} f_t e_{it} - (\tilde{Z}^{\intercal}\tilde{\lambda}_{1,i} - H_2^{-1} Z^{\intercal} \lambda_{1,i}) + O_p(\delta_{NT}^{-2}). \end{aligned}$$

We now focus on the asymptotic expansion of $\tilde{Z}^{\intercal}\tilde{\lambda}_{1,i} - H_2^{-1} Z^{\intercal}\lambda_{1,i}$:

$$\begin{aligned} &\tilde{Z}^{\intercal}\tilde{\lambda}_{1,i} - H_2^{-1} Z^{\intercal}\lambda_{1,i} \\ =&\tilde{Z}^{\intercal}(\tilde{\lambda}_{1,i} - H_1^{-1}\lambda_{1,i}) + \left(\tilde{Z} - H_1^{\intercal} Z H_2^{-\intercal}\right)^{\intercal} H_1^{-1}\lambda_{1,i} + \left(H_1^{\intercal} Z H_2^{-\intercal}\right)^{\intercal} H_1^{-1}\lambda_{1,i} - H_2^{-1} Z^{\intercal}\lambda_{1,i} \\ =&\tilde{Z}^{\intercal}(\tilde{\lambda}_{1,i} - H_1^{-1}\lambda_{1,i}) + O_p\left(\frac{1}{\delta_{NT}^2}\right) O_p(1) \\ =&\tilde{Z}^{\intercal}(\tilde{\lambda}_{1,i} - H_1^{-1}\lambda_{1,i}) + O_p\left(\frac{1}{\delta_{NT}^2}\right), \end{aligned}$$

by Theorem 3.1 (a).

Applying Lemma 10 to expand $(\tilde{\lambda}_{1,i} - H_1^{-1}\lambda_{1,i})$, we have

$$(\tilde{w}_i - H_2^{-1} w_i) = H_2^{\intercal} \frac{1}{(1-\pi)T} \sum_{t=\lfloor \pi T+1 \rfloor}^{T} f_t e_{it} - \tilde{Z}^{\intercal} \frac{1}{\pi T} \sum_{t=1}^{\lfloor \pi T \rfloor} H_1^{\intercal} f_t e_{it} + O_p\left(\frac{1}{\delta_{NT}^2}\right).$$

Multiplying both sides by $\sqrt{T}$ then yields

$$\sqrt{T}(\tilde{w}_i - H_2^{-1}w_i) = H_2^{\intercal}\frac{1}{(1-\pi)\sqrt{T}}\sum_{t=\lfloor \pi T+1\rfloor}^{T} f_t e_{it} - \tilde{Z}^{\intercal}\frac{1}{\pi\sqrt{T}}\sum_{t=1}^{\lfloor \pi T\rfloor} H_1^{\intercal} f_t e_{it} + O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right),$$

where the remainder term is $o_p(1)$ as $\frac{\sqrt{T}}{N} \to 0$. Recognising the CLT random variable terms in Assumption 16, we have

$$\sqrt{T}(\tilde{w}_i - H_2^{-1}w_i) \xrightarrow{d} N(0, \Omega_{W,i})$$

where

$$\begin{aligned}\Omega_{W,i} &= \left(\frac{1}{1-\pi}\right) H_{0,2}^{\intercal}\Phi_i H_{0,2} + \left(\frac{1}{\pi}\right) H_{0,2}^{-1} Z' \Sigma_F^{-1}\Phi_i \Sigma_F^{-1} Z H_{0,2}^{-\intercal} \\ &= \left(\frac{1}{1-\pi}\right)\Theta_{1,i} + \left(\frac{1}{\pi}\right)\Theta_{2,i}.\end{aligned}$$

The form of $\Theta_{2,i}$ comes from the fact that $\tilde{Z}$ estimates $H_1 Z H_2^{-\intercal}$ by Theorem 3.1 (a). By the convergence of H_1 to its limit $H_{0,1}$, we have

$$\begin{aligned}\tilde{Z}^{\intercal} H_1^{\intercal} &\xrightarrow{p} H_{0,2}^{-1} Z^{\intercal} H_{0,1} H_{0,1}^{\intercal} \\ &= H_{0,2}^{-1} Z^{\intercal}\Sigma_F^{-1}.\end{aligned}$$

where the last line follows from the identity $H_{0,1}H_{0,1}^{\intercal} = \Sigma_F^{-1}$. To see this, recall that $H_{0,1}^{-1} = V^{1/2}\Upsilon_1^{\intercal}\Sigma_{\Lambda_1}^{1/2}$. This means that $H_{0,1}^{-1}\Sigma_F^{-1}H_{0,1}^{\intercal} = V_1^{1/2}\Upsilon_1^{\intercal}\left(\Sigma_{\Lambda_1}^{1/2}\Sigma_F^{-1}\Sigma_{\Lambda_1}^{1/2)}\right)\Upsilon_1 V_1^{1/2} = V_1^{1/2}V_1^{-1}V_1^{1/2} = I$ by eigen-identity, which can then be re-arranged as required.

Their estimators $\tilde{\Theta}_{1,i}, \tilde{\Theta}_{2,i}$ are discussed in Bai (2003), and are given by HAC estimators 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:

$$\tilde{\Theta}_{1,i} = D_{0,1,i} + \sum_{v=1}^{\lfloor \pi T \rfloor - 1} \mathbf{k}\left(\frac{v}{b_{\lfloor \pi T \rfloor}}\right)(D_{1,vi} + D_{1,vi}^\intercal) \tag{35}$$

$$\tilde{\Theta}_{2,i} = D_{0,2,i} + \sum_{v=1}^{T - \lfloor \pi T \rfloor - 1} \mathbf{k}\left(\frac{v}{b_{T - \lfloor \pi T \rfloor}}\right)(D_{2,vi} + D_{2,vi}^\intercal), \tag{36}$$

where $D_{1,vi} = (T_1)^{-1} \sum_{t=v+1}^{\lfloor \pi T \rfloor} \tilde{f}_{1,t} \tilde{e}_{it} \tilde{e}_{i,t-v} \tilde{f}_{1,t-v}^\intercal$, $D_{2,vi} = (T_2)^{-1} \sum_{t=T_1+v+1}^{T} \tilde{f}_{2,t} \tilde{e}_{it} \tilde{e}_{i,t-v} \tilde{f}_{2,t-v}^\intercal$, and $\mathbf{k}(.)$ is a real valued kernel, such as the Bartlett kernel, satisfying Assumption 9. The consistency of $\tilde{\Theta}_{1,i}$ and $\tilde{\Theta}_{2,i}$ for $\Theta_{1,i} = H_{1,0}^\intercal \Phi_{i,1} H_{1,0}$ and $\Theta_{2,i} = H_{2,0}^\intercal \Phi_{i,2} H_{2,0}$ respectively can be proved using the argument of Newey and West (1987), as stated by Bai (2003). Recalling that $\tilde{Z}$ estimates $H_1^\intercal Z H_2^{-\intercal}$ from Theorem 3.1 (a), it follows that we can estimate $\Omega_{W,i}$ using

$$\widehat{\Omega}_{W,i} = \frac{1}{1-\pi} \tilde{\Theta}_{2,i} + \frac{1}{\pi} \tilde{Z}^\intercal \tilde{\Theta}_{1,i} \tilde{Z}.$$

The asymptotic Chi-squared distribution then follows. ■

Before we prove Theorem 3.4 (b), we prove some lemmas that need to be used.

Lemma 11. *Under Assumptions 1 to 5, 8 and 15, for $m = 1, 2$ we have:*

(a) $\frac{\tilde{F}_m^\intercal (F_m - F_m H_m)}{T_m} \frac{\sum_{i=1}^N \lambda_{m,i}}{\sqrt{N}} = O_p\left(\frac{1}{\delta_{NT}^2}\right)$,

(b) $\frac{(\tilde{F}_m - F_m H_m)^\intercal}{\sqrt{T_m}} \frac{\sum_{i=1}^N e_{m,i}}{\sqrt{T_m N}} = O_p\left(\frac{1}{\delta_{NT}^2}\right)$,

(c) Under Assumptions 1 to 8, 13 and 14, and if $\frac{\sqrt{T}}{N} \to 0$, *then for* $m = 1, 2$, $\left\|\tilde{\Theta}_m - \Theta_m\right\| = o_p(1)$, *where* $\Theta_m = \text{plim}(N)^{-1} \sum_{i=1}^N \Theta_{m,i}$.

Proof of Lemma 11 (a). The first term can be bounded by

$$\frac{\tilde{F}_m^\intercal (F_m - F_m H_m)}{T_m} \frac{\sum_{i=1}^N \lambda_{m,i}}{\sqrt{N}}$$
$$= O_p\left(\frac{1}{\delta_{NT}^2}\right) O_p(1)$$

by Lemma 1 (d), and Assumption 15. ■

Proof of Lemma 11 (b). It suffices to show that $\frac{1}{T_m\sqrt{N}}\sum_{t=1}^{T}(\tilde{f}_{m,t}-H_m^\intercal f_{m,t})\sum_{i=1}^{N}e_{it} = O_p\left(\frac{1}{\delta_{NT}^2}\right)$. We focus on the case of $m=1$, as the proof for $m=2$ is similar and thus omitted. From Equation (24), we have

$$\begin{aligned}
&\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}(\tilde{f}_{1,t}-H_1^\intercal f_t)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\\
&=V_{NT,1}^{-1}\left(\frac{1}{T_1^2}\sum_{s=1}^{\lfloor \pi T\rfloor}\sum_{t=1}^{\lfloor \pi T\rfloor}\tilde{f}_{1,s}\gamma_N(s,t)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}+\frac{1}{T_1^2}\sum_{s=1}^{\lfloor \pi T\rfloor}\sum_{t=1}^{\lfloor \pi T\rfloor}\tilde{f}_{1,s}\left(\frac{e_s^\intercal e_t}{N}-\gamma_N(s,t)\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}+\right.\\
&\left.\frac{1}{NT_1^2}\sum_{s=1}^{\lfloor \pi T\rfloor}\sum_{t=1}^{\lfloor \pi T\rfloor}\tilde{f}_{1,t}f_t^\intercal\Lambda_1 e_t\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}+\frac{1}{NT_1^2}\sum_{s=1}^{\lfloor \pi T\rfloor}\sum_{t=1}^{\lfloor \pi T\rfloor}\tilde{f}_{1,t}e_{(1)}^\intercal\Lambda_1 f_t\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\right)\\
&=V_{NT}^{-1}(a+b+c+d),
\end{aligned}$$

where we shall prove that each of a,b,c,d are asymptotically negligible.

$$\begin{aligned}
a&=\frac{1}{T_1^2}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}H_1^\intercal f_{1,s}\gamma_N(s,t)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}+\frac{1}{T_1^2}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}(\tilde{f}_{1,s}-H_1^\intercal f_{1,s})\gamma_N(s,t)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\\
&=a.I+a.II
\end{aligned}$$

We shall prove that each of $a.I$ and $a.II$ are asymptotically negligible. First, the term $a.I$ can be bounded by

$$\begin{aligned}
a.I&\le\frac{1}{T_1^2}E\left(\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}f_{1,s}\gamma_N(s,t)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\right)\\
&=\frac{1}{T_1^2}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}|\gamma_N(s,t)|\,E\left(\|f_{1,s}\|^2E\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\right)^2\right)\\
&\le\frac{1}{T_1^2}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}|\gamma_N(s,t)|\,M=O_p\left(\frac{1}{T}\right),
\end{aligned}$$

by Assumptions 1, 3 (a) and 3 (c), and $E\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\right)\le M$ by Assumption 13. Next, the term $a.II$ can be bounded by

$$
\begin{aligned}
a.II \leq & \frac{1}{\sqrt{T_1}} \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,s} - H_1 f_s \right\|^2 \right)^{\frac{1}{2}} \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \sum_{s=1}^{\lfloor \pi T \rfloor} |\gamma_N(s,t)|^2 \frac{1}{T_1} E \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{it} \right)^2 \right)^{\frac{1}{2}} \\
= & \frac{1}{\sqrt{T_1}} O_p \left(\frac{1}{\delta_{NT}} \right) O_p(1),
\end{aligned}
$$

where the $O_p(1)$ term follows from Assumption 3 (a), and because $T^{-1} \sum_{t=1}^{\lfloor \pi T \rfloor} \sum_{s=1}^{\lfloor \pi T \rfloor} |\gamma_N(s,t)|^2 \leq M$ by Lemma 1(i) of Bai and Ng (2002). Therefore, it follows that $a = O_p\left(\frac{1}{\delta_{NT}\sqrt{T}}\right)$.

Next, term b can be bounded by

$$
\begin{aligned}
b &= \frac{1}{T_1^2} \sum_{s=1}^{\lfloor \pi T \rfloor} \sum_{t=1}^{\lfloor \pi T \rfloor} \tilde{f}_{1,s} \left(\frac{e_s^{\intercal} e_t}{N} - \gamma_N(s,t) \right) \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{it} \\
&= b.I + b.II
\end{aligned}
$$

For $b.I$, we shall define $z_{1,t} = \frac{\sum_{s=1}^{\lfloor \pi T \rfloor} \sum_{i=1}^{N} f_{1,s}[e_{is}e_{it} - E(e_{is}e_{it})]}{\sqrt{TN}}$. By Assumption 6 (a), $E\|z_{1,t}\|^2 < M$, and thus $\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \|z_{1,t}\|^2 = O_p(1)$ by Assumption 13. This implies:

$$
\frac{1}{\sqrt{NT_1}} \|H_1\| \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \|z_{1,t}\|^2 \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{it} \right)^2 \right)^{\frac{1}{2}} = \frac{1}{\sqrt{T_1 N}} O_p(1),
$$

by Assumptions 3 (a) and 6 (a), and $\|H_1\| = O_p(1)$ due to Lemma 1 (e). For $b.II$, we can bound it by

$$
\begin{aligned}
& \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t} - H_1^{\intercal} f_{1,s} \right\|^2 \right)^{\frac{1}{2}} \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\frac{1}{T_1} \sum_{i=1}^{N} \left(\frac{e_s^{\intercal} e_t}{N} - \gamma_N(s,t) \right) \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{it} \right)^2 \right)^{\frac{1}{2}} \\
\leq & O_p\left(\frac{1}{\delta_{NT}} \right) \left(\frac{1}{T_1^2} \sum_{t=1}^{\lfloor \pi T \rfloor} \sum_{i=1}^{N} \left(\frac{e_s^{\intercal} e_t}{N} - \gamma_N(s,t) \right)^2 \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{it} \right)^2 \right)^{\frac{1}{2}} \\
= & O_p\left(\frac{1}{\delta_{NT}} \right) \frac{1}{\sqrt{N}} \left(\frac{1}{T_1^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{is} e_{it} - E(e_{is} e_{it}) \right)^2 \right) \frac{1}{T_1} \left(\sum_{i=1}^{N} \left(\frac{1}{\sqrt{N}} \left(\sum_{i=1}^{N} e_{it} \right)^2 \right) \right)^{\frac{1}{2}} \\
= & O_p\left(\frac{1}{\delta_{NT}} \right) \frac{1}{\sqrt{N}} O_p(1),
\end{aligned}
$$

because of Lemma 1 (a), the $O_p(1)$ term comes from Assumptions 3 (a) and 13. Next, term c can be bounded by

$$
\begin{aligned}
c &= \frac{1}{NT_1^2}\sum_{s=1}^{\lfloor \pi T\rfloor}\sum_{t=1}^{\lfloor \pi T\rfloor}\tilde{f}_{1,s}f_{1,s}^{\intercal}\Lambda_1 e_t \frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it} \\
&= O_p(1)\left(\frac{1}{T_1 N}\sum_{i=1}^{N}\left(\frac{1}{\sqrt{N}}\sum_{k=1}^{N}\lambda_{1,k}e_{kt}e_{it}\right)+\frac{1}{T_1 N\sqrt{N}}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}\sum_{k\neq i}\lambda_{1,k}e_{kt}e_{it}\right) \\
&= O_p(1)\left(O_p\left(\frac{1}{N}\right)+O_p\left(\frac{1}{\sqrt{TN}}\right)\right)=O_p\left(\frac{1}{\delta_{NT}\sqrt{N}}\right) \qquad (37)
\end{aligned}
$$

by Assumptions 6 (b) and 15 (b). Term d can be bounded by

$$
\begin{aligned}
d &= \frac{1}{NT_1^2}\sum_{s=1}^{\lfloor \pi T\rfloor}\sum_{t=1}^{\lfloor \pi T\rfloor}\tilde{f}_{1,s}e_s^{\intercal}\Lambda_1 f_t \frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it} \\
&= \frac{1}{NT_1^2}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}\tilde{f}_{1,s}e_s^{\intercal}\Lambda_1 f_t\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}+\frac{1}{NT_1^2}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}(\tilde{f}_{1,s}-H_1^{\intercal}f_{1,s})e_s^{\intercal}\Lambda_1 f_t\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it} \\
&= d.I + d.II.
\end{aligned}
$$

The first term $d.I$ can be bounded by

$$
\begin{aligned}
d.I &\leq \frac{1}{\sqrt{T_1 N}}\left\|\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}H_1^{\intercal}f_{1,s}e_{it}\lambda_{1,i}^{\intercal}\right\|\left\|\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}f_t\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\right\| \\
&\leq \frac{1}{\sqrt{T_1 N}}O_p(1)\left(\frac{1}{T_1}\sum_{t=1}^{T}\|f_t\|^2\frac{1}{T_1}\sum_{i=1}^{N}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it}\right)^2\right)^2 \\
&= O_p\left(\frac{1}{\sqrt{T_1 N}}\right),
\end{aligned}
$$

by Assumptions 1, 6 (a) and 13.

The second term $d.II$ can be bounded by:

$$d.II \leq \frac{1}{\sqrt{N}} \left\| \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} (\widehat{f}_{1,s} - H_1^{\intercal} f_{1,s}) \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \lambda_{1,i}^{\intercal} e_{is} \right) \right\| \left\| \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} f_t \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{it} \right\|$$

$$\leq \frac{1}{\sqrt{N}} \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \widehat{f}_s - H_1^{\intercal} f_s \right\|^2 \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \lambda_{1,i}^{\intercal} e_{is} \right\|^2 \right)^{1/2} \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \| f_t \|^2 \frac{1}{T_1} \sum_{i=1}^{N} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{it} \right)^2 \right)^{1/2}$$

$$= \frac{1}{\sqrt{N}} \left(O_p \left(\frac{1}{\delta_{NT}^2} \right) O_p (1) \right)^{1/2} O_p (1) = \frac{1}{\sqrt{N}} O_p \left(\frac{1}{\delta_{NT}} \right),$$

by Assumptions 1 to 6 (c) and Lemma 1 (a). Therefore, a, b, c, d in the remainder term are all asymptotically negligible. ■

The proof of $\tilde{\Theta}_{m,i} \to \Theta_{m,i}$ for $m = 1, 2$ has been briefly illustrated in Bai (2003), and can be proved by applying a HAC estimator using $\tilde{f}_t \cdot \tilde{e}_{it}$. We therefore focus on the consistency of the pooled version $\tilde{\Theta}_m \to \Theta_m$. Before we present the main proofs for the W covariance matrices, we present some lemmas to be used.

Lemma 12. *Under Assumptions 1 to 6, 8, 13 and 14, and if $\sqrt{T}/N \to 0$, then for $m = 1, 2$:*

(a) $\frac{1}{T_m} \sum_{t=1}^{T} \left| \left(\tilde{e}_{(m),it} - e_{it} \right) \iota_{mt} \right|^4 = o_p(1)$ *for all i,*

(b) If additionally, Assumption 15 holds, then $\frac{1}{T_m N} \sum_{i=1}^{N} \sum_{t=1}^{T} \left| \left(\tilde{e}_{(m),it} - e_{it} \right) \iota_{mt} \right|^4 = o_p(1)$

(c) $\frac{1}{T_m} \sum_{t=1}^{T} (\tilde{e}_{(m),it} \iota_{mt})^4 = O_p(1)$

(d) If additionally, Assumption 15 holds, $\frac{1}{T_m N} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{e}_{(m),it}^4 \iota_{mt} = O_p(1)$

Proof of Lemma 12 (a). For brevity, we focus on $m = 1$, as the case for $m = 2$ is very similar. We have by definition:

$$
\begin{aligned}
\left|\tilde{e}_{(1),it} - e_{it}\right| &= \left|\tilde{\lambda}_{1,i}^{\intercal}\tilde{f}_{1,t} - \lambda_{1,i}^{\intercal}f_t\right| \\
&= \left|\tilde{\lambda}_{1,i}^{\intercal}\tilde{f}_{1,t} - \lambda_{1,i}^{\intercal}H_1^{-\intercal}H_1^{\intercal}f_t\right| \\
&= \left|\tilde{\lambda}_{1,i}^{\intercal}(\tilde{f}_{1,t} - H_1^{\intercal}f_t) + \tilde{\lambda}_{1,i}^{\intercal}H_1^{\intercal}f_t - (\lambda_{1,i}^{\intercal}H_1^{-\intercal}H_1^{\intercal}f_t)\right| \\
&= \left|\tilde{\lambda}_{1,i}^{\intercal}(\tilde{f}_{1,t} - H_1^{\intercal}f_t) + (\tilde{\lambda}_{1,i}^{\intercal} - \lambda_{1,i}^{\intercal}H_1^{-\intercal})H_1^{\intercal}f_t\right| \\
&= |E.I_t + E.II_t|.
\end{aligned}
$$

Noting that $\left|\tilde{e}_{(1),it} - e_{it}\right|^4 \leq 64|E.I_t^4 + E.II_t^4|$, it therefore suffices to consider the bounds of $\frac{1}{T_m}\sum_{t=1}^{T} E.I_t^4 \iota_{mt}$ and $\frac{1}{T_m}\sum_{t=1}^{T} E.II_t^4 \iota_{mt}$.

First, $\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} E.I_t^4$ can be bounded by:

$$
\begin{aligned}
\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} E.I_t^4 &\leq \left\|\tilde{\lambda}_{1,i}\right\|^4 \frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor}\left\|\tilde{f}_{1,t} - H_1^{\intercal}f_t\right\|^4 \\
&= O_p(1)\, O_p\left(\frac{1}{\delta_{NT}^4}\right),
\end{aligned}
$$

where $\left\|\tilde{\lambda}_{1,i}\right\|^4 = O_p(1)$ because each $\tilde{\lambda}_{1,i}$ is bounded by normalisation, and $\frac{1}{T}\left\|\tilde{f}_{1,t} - H_1^{\intercal}f_t\right\|^4 = O_p\left(\frac{1}{\delta_{NT}^4}\right)$ by Lemma 5 (a).

Next, $\frac{1}{T_1}\sum_{t=1}^{T} E.II_t^4$ can be bounded by:

$$
\begin{aligned}
\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} E.II_t^4 &\leq \left\|\tilde{\lambda}_{1,i}^{\intercal} - \lambda_{1,i}^{\intercal}H_1^{-\intercal}\right\|^4 \|H_1\|^4 \frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} f_t^4 \\
&= O_p\left(\frac{1}{\delta_{NT}^4}\right) O_p(1)\, O_p(1),
\end{aligned}
$$

where $\left\|\tilde{\lambda}_i^{\intercal} - \lambda_i^{\intercal}H^{-\intercal}\right\|^4 = O_p\left(\frac{1}{\delta_{NT}^4}\right)$ by Lemma 5 (d), and $\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} f_t^4 = O_p(1)$ by Assumption 1. ■

Proof of Lemma 12 (b). The proof is similar to that of Lemma 12 (a) - it suffices to show that $\frac{1}{TN}\sum_{t=1}^{T}\sum_{i=1}^{N} E.I_t^4$ and $\frac{1}{TN}\sum_{t=1}^{T}\sum_{i=1}^{N} E.II_t^4$ are both negligible. For brevity, we will focus on the case of $m=1$, as the case for $m=2$ is similar. First, $\frac{1}{T_1 N}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N} E.II_t^4$ can be bounded by

$$\begin{aligned}
\frac{1}{T_1 N}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N} E.I_t^4 =&\frac{1}{T_1 N}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}\tilde{\lambda}_{1,i}^{\intercal}(\tilde{f}_{1,t}-H_1^{\intercal}f_t)\\
=&\frac{1}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}\left[\left(\tilde{\lambda}_{1,i}-\lambda_{1,i}H_1^{-\intercal}\right)\left(\tilde{f}_{1,t}-H_1^{\intercal}f_t\right)+\lambda_{1,i}H_1^{\intercal}\left(\tilde{f}_{1,t}-H_1^{\intercal}f_t\right)\right]^4\\
\leq&\frac{64}{T_1 N}\sum_{t=1}^{\lfloor \pi T\rfloor}\left(\tilde{f}_{1,t}-H_1^{\intercal}f_t\right)^4\frac{1}{N}\left(\tilde{\lambda}_{1,i}-\lambda_i H_1^{-\intercal}\right)^4+\frac{64}{T_1 N}\sum_{i=1}^{N}(\lambda_{1,i}H_1^{\intercal})^4\sum_{t=1}^{\lfloor \pi T\rfloor}\left(\tilde{f}_{1,t}-H_1^{\intercal}f_t\right)^4\\
\leq&\frac{64}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}\left\|\tilde{f}_{1,t}-H_1^{\intercal}f_t\right\|^4\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\lambda}_{1,i}-\lambda_{1,i}H_1^{-\intercal}\right\|^4\\
&\quad+64\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}\left\|\tilde{f}_{1,t}-H_1^{\intercal}f_t\right\|^4\frac{1}{N}\sum_{i=1}^{N}\|\lambda_{1,i}\|^4\|H_1\|^4\\
=&o_p(1)o_p(1)+o_p(1)O_p(1)\\
=&o_p(1),
\end{aligned}$$

where the first $o_p(1)$ term comes from Lemmas 5 (a) and 5 (d), and the second term is $o_p(1)$ from Lemmas 5 (a) and 1 (e) and Assumption 2. The second term $\frac{1}{TN}\sum_{t=1}^{T}\sum_{i=1}^{N} E.II_t^4$ can be bounded by

$$\begin{aligned}
\frac{1}{T_1 N}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N} E.II_t^4 &= \frac{1}{T_1 N}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}(\tilde{\lambda}_{1,i}^{\intercal}-\lambda_{1,i}^{\intercal}H^{-\intercal})H_1^{\intercal}f_t\\
&\leq \frac{1}{T_1 N}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}\left\|\tilde{\lambda}_{1,i}-\lambda_{1,i}H_1^{-\intercal}\right\|^4\|H_1\|^4\|f_t\|^4\\
&= \frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\lambda}_{1,i}-\lambda_{1,i}H_1^{-\intercal}\right\|^4\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}f_t^4\|H_1\|^4\\
&= o_p(1)O_p(1)\,O_p(1)\\
&= o_p(1),
\end{aligned}$$

because of Lemmas 5 (d) and 1 (e) and Assumption 1. ■

Proof of Lemmas 12 (c) and 12 (d). Lemmas 12 (c) and 12 (d) are implications and Lemmas 12 (a) and 12 (b) and can be proven in a similar way. Focusing on $m = 1$ as the case for $m = 2$ is similar, we have

$$\begin{aligned}\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} \tilde{e}_{it}^4 &\leq \frac{8}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} (\tilde{e}_{it} - e_{it})^4 + \frac{8}{T_1}\sum_{t=1}^{\lfloor \pi T \rfloor} e_{it}^4 \\ &= O_p(1), \\ \frac{1}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T \rfloor} \tilde{e}_{it}^4 &\leq \frac{8}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T \rfloor} (\tilde{e}_{it} - e_{it})^4 + \frac{8}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T \rfloor} e_{it}^4 \\ &= O_p(1).\end{aligned}$$

■

Lemma 13. *Under Assumptions 1 to 8 and 13 to 15, for $m = 1, 2$,*

$$\left\| \frac{1}{N}\sum_{i=1}^{N} D_{j,m,i} - D_{j,m}^* \right\| = o_p(1),$$

where $D_{j,m}^* = \text{plim}\, \frac{1}{N}\sum_{i=1}^{N} D_{j,m,i}^* = \text{plim}\, \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\lfloor \pi T \rfloor}\sum_{t=\lfloor \pi T+1 \rfloor}^{T} H_m^\intercal f_t e_{it} e_{i,t-v}^\intercal f_{t-v}^\intercal H_m$, its infeasible counterpart.

Proof of Lemma 13. We focus on the case of $m = 1$, as the proof for $m = 2$ is similar and thus omitted. We have

$$\begin{aligned}&\left\| \frac{1}{N}\sum_{i=1}^{N} D_{j,m,i} - D_{j,m}^* \right\| \\ =&\left\| \frac{1}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T \rfloor} \tilde{f}_{1,t}\tilde{f}_{1,t-j}^\intercal \tilde{e}_{(1),it}\tilde{e}_{(1),i,t-j} - H_1^\intercal \left(\frac{1}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T \rfloor} f_t f_{t-j}^\intercal e_{it} e_{i,t-j} \right) H_1 \right\| \\ \leq&\frac{1}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T \rfloor} \left(\tilde{f}_{1,t}\tilde{f}_{1,t-j}^\intercal - H_1^\intercal f_t f_{t-j} H_1 \right) \left(\tilde{e}_{it}\tilde{e}_{i,t-j} \right) \\ &\quad + \frac{1}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T \rfloor} \left(H_1^\intercal f_t f_{t-j} H_1 \right) \left(\tilde{e}_{(1),it}\tilde{e}_{(1),i,t-j} - e_{it} e_{i,t-j} \right) \\ =&D.I + D.II.\end{aligned}$$

The first term $D.I$ is bounded by

$$
\begin{aligned}
& D.I \\
=& \frac{1}{T_1 N} \sum_{i=1}^{N} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\tilde{f}_{1,t} \tilde{f}_{1,t-j}^{\intercal} - H_1^{\intercal} f_t f_{t-j} H_1 \right) \left(\tilde{e}_{(1),it} \tilde{e}_{(1),i,t-j} \right) \\
\leq & \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t} \tilde{f}_{1,t-j}^{\intercal} - H_1^{\intercal} f_t f_{t-j} H_1 \right\|^2 \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\frac{1}{N} \sum_{i=1}^{N} \tilde{e}_{(1),it} \tilde{e}_{(1),i,t-j} \right)^2 \right)^{\frac{1}{2}} \\
\leq & \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t} (\tilde{f}_{1,t-j}^{\intercal} - f_{t-j} H_1) + (\tilde{f}_{1,t} - H_1^{\intercal} f_t)(\tilde{f}_{1,t-j}) \right\|^2 \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\frac{1}{N} \sum_{i=1}^{N} \tilde{e}_{(1),it} \tilde{e}_{(1),i,t-j} \right)^2 \right)^{\frac{1}{2}} \\
\leq & \left(\frac{2}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t} (\tilde{f}_{1,t-j}^{\intercal} - f_{t-j} H_1) \right\|^2 + \left\| (\tilde{f}_{1,t} - H_1^{\intercal} f_t)(\tilde{f}_{1,t-j}) \right\|^2 \right)^{\frac{1}{2}} \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left(\frac{1}{N} \sum_{i=1}^{N} \tilde{e}_{(1),it} \tilde{e}_{(1),i,t-j} \right) \right)^{\frac{1}{2}} \\
\leq & \left(2 \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t} \right\|^4 \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t-j} - f_{t-j}^{\intercal} H_1 \right\|^4 \right)^{\frac{1}{2}} + 2 \left(\frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t-j} \right\|^4 + \frac{1}{T_1} \sum_{t=1}^{\lfloor \pi T \rfloor} \left\| \tilde{f}_{1,t} - H_1^{\intercal} f_t \right\|^4 \right)^{\frac{1}{2}} \right)^{\frac{1}{2}} \\
& \times \left(\frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} \tilde{e}_{(1),it}^2 \tilde{e}_{i,t-j}^2 \right)^{1/2} \\
\leq & o_p(1) \left(\frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} \tilde{e}_{it}^2 \tilde{e}_{i,t-j}^2 \right)^{1/2} \\
\leq & o_p(1) \left[\left(\frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} \tilde{e}_{(1),it}^4 \right)^{\frac{1}{2}} \left(\frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} \tilde{e}_{(1),i,t-j}^4 \right)^{\frac{1}{2}} \right]^{1/2} \\
=& o_p(1) O_p(1),
\end{aligned}
$$

where the first term is $o_p(1)$ follows from applying Lemmas 5 (a) and 5 (c), and the second term follows from Lemma 12 (d). The second term $D.II$ is bounded by

$$
\begin{aligned}
D.II =& \frac{1}{T_1 N}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor \pi T\rfloor}\left(H_1^{\intercal} f_t f_{t-j} H_1\right)\left(\tilde{e}_{(1),it}\tilde{e}_{(1),i,t-j} - e_{it}e_{i,t-j}\right) \\
\leq & \left(\frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}\|H_1^{\intercal} f_t f_{t-j} H_1\|^2 \frac{1}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\tilde{e}_{(1),it}\tilde{e}_{(1),i,t-j} - e_{it}e_{i,t-j}\right)\right)^2\right)^{\frac{1}{2}} \\
\leq & O_p(1)\left(\frac{2}{T_1}\sum_{t=1}^{\lfloor \pi T\rfloor}\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{e}^2_{(1),it}(\tilde{e}_{(1),i,t-j} - e_{i,t-j}) + e^2_{i,t-j}(\tilde{e}_{(1),it} - e_{it})^2\right)\right)^{\frac{1}{2}} \\
\leq & O_p(1)\left(\frac{2}{T_1 N}\left(\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}\tilde{e}^4_{(1),it}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}(\tilde{e}_{(1),i,t-j} - e_{t,i-j})^4\right)^{\frac{1}{2}}\right. \\
& \qquad \left. + \frac{2}{T_1}\left(\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}e^4_{i,t-j}\sum_{t=1}^{\lfloor \pi T\rfloor}\sum_{i=1}^{N}(\tilde{e}_{(1),it} - e_{it})^4\right)^{\frac{1}{2}}\right) \\
=& o_p(1),
\end{aligned}
$$

where the $O_p(1)$ term comes from Assumption 1 and lemma 1 (e) and the $o_p(1)$ term comes from Lemma 12 (b). ■

Proof of Lemma 11 (c). It suffices to prove that $\left\|\frac{1}{N}\sum_{i=1}^{N}\tilde{\Theta}_{m,i} - \Theta_m\right\| = o_p(1)$ for $m = 1, 2$. For brevity, we will only prove the case for $m = 1$, as the case for $m = 2$ is similar. See that

$$
\begin{aligned}
& \left\|\frac{1}{N}\sum_{i=1}^{N}\tilde{\Theta}_{m,i} - \Theta_m\right\| \\
& \leq \left\|\frac{1}{N}\sum_{i=1}^{N}D_{0,1,i} + \frac{1}{N}\sum_{i=1}^{N}\sum_{v=1}^{\lfloor \pi T-1\rfloor}\mathbf{k}\left(\frac{v}{b_{\lfloor \pi T\rfloor}}\right)D_{1,vi} - D^*_{0,1} - \sum_{v=1}^{\lfloor \pi T-1\rfloor}\mathbf{k}\left(\frac{v}{b_{\lfloor \pi T\rfloor}}\right)D^*_{vi}\right\| \\
& \leq \left\|\frac{1}{N}\sum_{i=1}^{N}D_{0,1,i} - D^*_{0,1}\right\| + 2\sum_{v=1}^{b_{\lfloor \pi T\rfloor}}\left\|\frac{1}{N}\sum_{i=1}^{N}D_{1,vi} - D^*_{vi}\right\| \\
= & \; o_p(1),
\end{aligned}
$$

by Lemma 13. ■

Proof of Theorem 3.4 (a). We are now considering the asymptotic expansion of $\tilde{\lambda}_{m,i}$ in

Lemma 10, but averaged across the cross section and inflated by $\sqrt{N}$:

$$\frac{\sum_{i=1}^{N}\tilde{\lambda}_{m,i}-H_1^{-1}\lambda_{m,i}}{\sqrt{N}}=\frac{1}{T_m\sqrt{N}}\sum_{i=1}^{N}\tilde{F}_m^{\intercal}(F_m-F_mH_m)\lambda_{m,i}+\frac{1}{T_m\sqrt{N}}\sum_{i=1}^{N}\tilde{F}_m-F_mH_m{}^{\intercal}e_{(m),i}, \tag{38}$$

where the last two terms are asymptotically negligible because of Lemmas 11 (a) and 11 (b).

Similarly, considering the asymptotic expansion of $\tilde{w}_i$ in Equation (35), taking its cross sectional mean, and then inflating by $\sqrt{TN}$ on both sides, we have:

$$\begin{aligned}\sqrt{TN}\frac{\sum_{i=1}^{N}(\tilde{w}_i-H_2^{-1}w_i)}{N}=&H_2^{\intercal}\frac{1}{(1-\pi)\sqrt{TN}}\sum_{i=1}^{N}\sum_{t=\lfloor\pi T+1\rfloor}^{T}f_te_{it}\\&-\tilde{Z}^{\intercal}\frac{1}{\pi\sqrt{TN}}\sum_{i=1}^{N}\sum_{t=1}^{\lfloor\pi T\rfloor}H_1^{\intercal}f_te_{it}+O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right)\\&\xrightarrow{d}N(0,\Omega_W),\end{aligned}$$

where $\Omega_W=\frac{1}{(1-\pi)TN}H_{0,2}^{\intercal}\sum_{i=1}^{N}\Phi_{i,2}H_{0,2}+\frac{1}{\pi TN}H_{0,2}^{-1}Z'\Sigma_F^{-1}\sum_{i=1}^{N}\Phi_{i,1}\Sigma_F^{-1}ZH_{0,2}^{-\intercal}$, and the remainder terms are asymptotically negligible by Lemmas 11 (a) and 11 (b). The asymptotic distribution then follows by Assumption 16, the convergence of H_1,H_2 and $\tilde{Z}$ to their probability limits, and the consistency of $\tilde{\Omega}_W$ for Ω_W. ■

Proof of Theorem 3.5. We will show that $\mathscr{W}_{W,i}\to\infty$ as $N,T\to\infty$ when $w_i\neq 0$. Recalling the asymptotic expansion of $\tilde{w}_i$, we have:

$$\begin{aligned}\tilde{w}_i=&(H_2^{-1}w_i)+H_2^{-1}\frac{1}{(1-\pi)T}\sum_{t=\lfloor\pi T+1\rfloor}^{T}f_te_{it}-\tilde{Z}^{\intercal}\frac{1}{\pi T}\sum_{t=1}^{\lfloor\pi T\rfloor}H_1^{\intercal}f_te_{it}+O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right)\\=&H_2^{-1}w_i+O_p\left(\frac{1}{\sqrt{T}}\right)+O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right),\\=&H_2^{-1}w_i+o_p(1).\end{aligned}$$

Because $\tilde{\Omega}_{W,i}\xrightarrow{p}\Omega_{W,i}$ and $\Omega_{W,i}$ is positive definite, it follows that $\tilde{\Omega}_{W,i}^{-1}=O_p(1)$ and we have the desired result:

$$\text{plim}_{N,T\to\infty}\inf\left(\frac{1}{\sqrt{T}}W_{W_i}\right)=\text{plim}_{N,T\to\infty}\inf\left(\tilde{w}_i^\intercal\Omega_{W,i}^{-1}\tilde{w}_i\right)>0$$

which implies the desired divergence under the alternative hypothesis. We will show that $\mathscr{W}_W\to\infty$ as $N,T\to\infty$ under the alternative. We have:

$$\begin{aligned}
&(T^{\alpha/2})\left(\frac{\sum_{i=1}^N\tilde{w}_i}{\sqrt{N}}\right)\\
=&T^{\alpha/2}\frac{\sum_{i=1}^N(H_2^{-1}w_i)}{\sqrt{N}}+H_2^\intercal\frac{T^{\alpha/2}}{(1-\pi)T\sqrt{N}}\sum_{i=1}^N\sum_{s=\lfloor\pi T+1\rfloor}^T f_se_{is}-\tilde{Z}^\intercal\frac{T^{\alpha/2}}{\pi T\sqrt{N}}\sum_{i=1}^N\sum_{s=1}^{\lfloor\pi T\rfloor}H_1^\intercal f_se_{is}+O_p\left(\frac{T^{\alpha/2}}{\delta_{NT}^2}\right)\\
=&H_2^\intercal T^{\alpha/2}\frac{\sum_{i=1}^N(H_2^{-1}w_i)}{\sqrt{N}}+H_2^\intercal\frac{1}{(1-\pi)T^{1-\alpha/2}}O_p\left(1\right)-\tilde{Z}^\intercal H_1^\intercal\frac{1}{\pi T^{1-\alpha/2}}O_p\left(1\right)+O_p\left(\frac{T^{\alpha/2}}{\delta_{NT}^2}\right)\\
=&H_2^\intercal T^{\alpha/2}\frac{\sum_{i=1}^N(H_2^{-1}w_i)}{\sqrt{N}}+O_p\left(\frac{1}{T^{1-\alpha/2}}\right)+O_p\left(\frac{T^{\alpha/2}}{\delta_{NT}^2}\right),
\end{aligned}$$

where the last two terms are $o_p(1)$ because $0<\alpha\le 0.5$ and $\sqrt{T}/N\to 0$ as $N,T\to\infty$. Since $\tilde{\Omega}_W\xrightarrow{p}\Omega_W$ and Ω_W is positive definite, it follows that $\tilde{\Omega}_W^{-1}=O_p\left(1\right)$ and we have the desired result by Assumption 17:

$$\begin{aligned}
&\text{plim}_{N,T\to\infty}\inf\left(\frac{T^\alpha}{T}\mathscr{W}_W\right)\\
=&\text{plim}_{N,T\to\infty}\inf\left[(T^\alpha)\left(\frac{\sum_{i=1}^N\tilde{w}_i}{\sqrt{N}}\right)^\intercal(\tilde{\Omega}_W)^{-1}\left(\frac{\sum_{i=1}^N\tilde{w}_i}{\sqrt{N}}\right)\right]>0,
\end{aligned}$$

which implies the desired divergence under the alternative hypothesis. ■

A.5 Singular Z

The case of a singular Z can be further classified into two cases, depending on the column rank of W. In this section, we show that with some suitable adjustments, our test statistics can accommodate these cases.

A.5.1 Replacement of factors: $rank(Z) < rank(W)$

If the column rank of is still r, this represents the case where some of the original factors are "replaced" by an entirely new set of factors. In this case, $\Lambda_2 = \Lambda_1 Z + W$ is still of full rank, and the existing theory can still go through.

A.5.2 Changing number of factors: $rank(Z) = rank(W)$

If Z is singular and has identical (column) rank to W, then this represents the case of a disappearing factor. Note that the case of an emerging factor can always be parameterized in by reversing the pre- and post-break samples, and thus our method can be extended to accommodate a changing number of factors.

Existing work tends to parameterize a disappearing factor by allowing for a singular Z, (e.g. Han and Inoue, 2015; Baltagi et al., 2017; Bai et al., 2022). However, these approaches work by using the *pseudo* factors - the case of split sample estimation estimation is more difficult. The main issue is to ensure that H_2 has valid limiting behavior - once this this done, the proofs for the split sample factors and rotated factors can follow on without major adjustments.

Without loss of generality, suppose that the $r - r_2$th factors disappear. To avoid Λ_2 not being of full column rank, we instead parameterize Λ_2 as an $N \times (r - r_2)$ matrix:

$$\begin{aligned} \Lambda_2 &= (\Lambda_1 + W) \begin{bmatrix} I_{r-r_2} \\ 0 \end{bmatrix} \\ &= \Lambda_1 Z_0 + W_0. \end{aligned} \tag{39}$$

This allows us to write

$$
\begin{aligned}
X_2 &= F_2\Lambda_2^{\intercal} + e_{(2)} \\
&= F_2 \begin{bmatrix} I_{r-r_2} \\ 0 \end{bmatrix} \left((\Lambda_1 Z + W) \begin{bmatrix} I_{r-r_2} \\ 0 \end{bmatrix} \right)^{\intercal} + e_{(2)} \\
&= F_{2,r-r_2} \left(\Lambda_1 Z_0 + W_0\right)^{\intercal} + e_{(2)},
\end{aligned} \tag{40}
$$

which expresses the post break data as a factor structure with $r - r_2$ factors. We can therefore apply the usual framework of Bai (2003) and use

$$
H_{2,r-r_2} = \frac{\Lambda_2^{\intercal}\Lambda_2}{N} \frac{F_{2,r-r_2}^{\intercal}\tilde{F}_2}{T_2} V_{NT,2,r-r_2}^{-1} \tag{41}
$$

where we can use the first $r - r_2$ post-break factors denoted by $\tilde{F}_{2,r-r_2}$. All of the above quantities exhibit full rank, and hence $H_{2,r-r_2}$ is an $(r - r_2) \times (r - r_2)$ square matrix.

Lemma 14. *Under Assumptions 1 to 8, as $N, T \to \infty$*

(a) $\frac{1}{T}\left\|\tilde{F}_{2,r-r_2} - F_{2,r-r_2}H_{2,r-r_2}\right\|^2 = O_p\left(\frac{1}{\delta_{NT}^2}\right)$,

(b) $\frac{1}{T}\left(\tilde{F}_{2,r-r_2} - F_{2,r-r_2}H_{2,r-r_2}\right)^{\intercal} F_{2,r-r_2} = O_p\left(\frac{1}{\delta_{NT}^2}\right)$

(c) $\frac{1}{T}\left(\tilde{F}_{2,r-r_2} - F_{2,r-r_2}H_{2,r-r_2}\right)^{\intercal} e_{i,(2)} = O_p\left(\frac{1}{\delta_{NT}^2}\right)$

Proof of Lemma 14. These correspond to Theorem of Bai and Ng (2002) and Lemmas B.1 and B.2 of Bai (2003). ■

Lemma 14 can also be used to prove analogous results for the rotated factors, where $\tilde{Z}$ is now an $r \times (r - r_2)$ matrix.

Lemma 15. *Under Assumptions 1 to 8, as $N, T \to \infty$,*

$$
\tilde{Z} = H_1^{\intercal} Z_0 H_{2,r-r_2}^{-\intercal} + O_p\left(\frac{1}{\delta_{NT}^2}\right).
$$

Proof of Lemma 15. Lemma 15 is analogous to Theorem 3.1 (a) and can be proved in a similar way. ■

A.5.3 LM-like Z-test

In either cases of a singular Z, the variance of the Z-test can also be estimated with an LM-like estimator which uses the whole data and hence leads to more numerical stability. This is because the variance estimate using the post break data, i.e. the outer product of $\tilde{F}_2\tilde{Z}^\intercal$ is singular and therefore a standard HAC estimator applied to this process will fail. Define

$$\widehat{\Omega}_Z(\widehat{F}) = \widehat{\Gamma}(FH_{1,0}) + \sum_{j=1}^{T} \mathbf{k}\left(\frac{j}{S_T}\right)\left(\widehat{\Gamma}_j\left(FH_{1,0}\right) + \widehat{\Gamma}_j\left(FH_{1,0}\right)^\intercal\right),$$
$$\widehat{\Gamma}_j(FH_{1,0}) = \frac{1}{T}\sum_{t=j+1}^{T} vech\left(H_{1,0}^\intercal f_t f_t^\intercal H_{1,0} - I_r\right) vech\left(H_{1,0}^\intercal f_{j-1} f_{j-1}^\intercal H_{1,0} - I_r\right)^\intercal,$$

which can be proven to be consistent for their respective infeasible counterparts $\widehat{\Omega}_Z(FH_{1,0})$ and $\widehat{\Gamma}(FH_{1,0})$ in a similar way to Lemmas 6 (b), 7 and 8. These lead to the infeasible estimator using all of the data

$$\tilde{S}_Z(\pi, \widehat{F}) = \left(\frac{1}{\pi} + \frac{1}{1-\pi}\right)\widehat{\Omega}_Z(\widehat{F}),$$

which can also be proven to be consistent for its infeasible counterpart in a similar way to Lemma 6 (b). Together with $A_Z(\pi, \widehat{F})$, these can be used to define the LM-like test statistic

$$LM_Z(\pi, \widehat{F}) = A_Z(\pi, \widehat{F})^\intercal \tilde{S}_Z(\pi, \widehat{F})^{-1} A_Z(\pi, \widehat{F}). \tag{42}$$

Its consistency to its infeasible counterpart and power under the alternative hypotheis can be proven in a similar way to Theorems 3.2 and 3.3.

B Additional Tables

B.1 Additional Simulation Results

Table 6: Size of Rotation and Orthogonal Shift Tests, $r = 3$

						Z Test		W Test		W Individual
T	N	τ	ρ	α	β	Unadjusted	Adjusted	Unadjusted	Adjusted	
200	200	0.3	0.0	0.0	0.0	0.368	0.257	0.001	0.000	0.009
200	200				0.3	0.374	0.280	0.038	0.029	0.010
200	200			0.3	0.0	0.491	0.371	0.004	0.001	0.012
200	200				0.3	0.504	0.397	0.054	0.042	0.014
200	200		0.7	0.0	0.0	0.204	0.147	0.001	0.000	0.007
200	200				0.3	0.206	0.145	0.035	0.020	0.008
200	200			0.3	0.0	0.221	0.156	0.006	0.001	0.021
200	200				0.3	0.225	0.167	0.069	0.047	0.021
200	200	0.5	0.0	0.0	0.0	0.286	0.204	0.004	0.003	0.013
200	200				0.3	0.292	0.215	0.082	0.060	0.014
200	200			0.3	0.0	0.345	0.256	0.004	0.003	0.015
200	200				0.3	0.352	0.281	0.101	0.067	0.017
200	200		0.7	0.0	0.0	0.219	0.144	0.002	0.002	0.012
200	200				0.3	0.221	0.152	0.059	0.039	0.013
200	200			0.3	0.0	0.241	0.167	0.009	0.006	0.026
200	200				0.3	0.238	0.176	0.098	0.072	0.027
200	200	0.7	0.0	0.0	0.0	0.353	0.275	0.000	0.000	0.005
200	200				0.3	0.365	0.282	0.024	0.018	0.006
200	200			0.3	0.0	0.410	0.313	0.000	0.000	0.007
200	200				0.3	0.417	0.335	0.036	0.028	0.007
200	200		0.7	0.0	0.0	0.377	0.289	0.000	0.000	0.005
200	200				0.3	0.385	0.297	0.031	0.020	0.005
200	200			0.3	0.0	0.393	0.318	0.005	0.003	0.014
200	200				0.3	0.398	0.330	0.072	0.051	0.015

500	200			0.0	0.0	0.182	0.110	0.001	0.001	0.002
500	200		0.0		0.3	0.184	0.112	0.031	0.015	0.003
500	200	0.3		0.3	0.0	0.209	0.136	0.001	0.000	0.003
500	200				0.3	0.224	0.135	0.031	0.020	0.003
500	200			0.0	0.0	0.160	0.101	0.000	0.000	0.002
500	200		0.7		0.3	0.158	0.105	0.027	0.018	0.002
500	200			0.3	0.0	0.151	0.093	0.002	0.001	0.005
500	200				0.3	0.148	0.097	0.039	0.029	0.006
500	200			0.0	0.0	0.125	0.079	0.007	0.004	0.007
500	200		0.0		0.3	0.129	0.085	0.082	0.051	0.007
500	200	0.5		0.3	0.0	0.143	0.076	0.007	0.001	0.007
500	200				0.3	0.148	0.098	0.095	0.059	0.008
500	200			0.0	0.0	0.156	0.080	0.005	0.001	0.006
500	200		0.7		0.3	0.155	0.086	0.075	0.052	0.006
500	200			0.3	0.0	0.155	0.088	0.011	0.005	0.010
500	200				0.3	0.159	0.094	0.096	0.068	0.011
500	200			0.0	0.0	0.198	0.146	0.000	0.000	0.002
500	200		0.0		0.3	0.200	0.150	0.026	0.013	0.002
500	200	0.7		0.3	0.0	0.204	0.145	0.000	0.000	0.002
500	200				0.3	0.221	0.156	0.034	0.015	0.002
500	200			0.0	0.0	0.220	0.148	0.000	0.000	0.002
500	200		0.7		0.3	0.220	0.149	0.020	0.012	0.002
500	200			0.3	0.0	0.231	0.151	0.002	0.001	0.004
500	200				0.3	0.232	0.157	0.034	0.020	0.004
1000	200			0.0	0.0	0.096	0.061	0.001	0.001	0.001
1000	200		0.0		0.3	0.096	0.063	0.027	0.016	0.001
1000	200	0.3		0.3	0.0	0.108	0.070	0.000	0.000	0.001
1000	200				0.3	0.118	0.070	0.031	0.013	0.001
1000	200			0.0	0.0	0.108	0.060	0.000	0.000	0.001
1000	200		0.7		0.3	0.104	0.059	0.016	0.005	0.001
1000	200			0.3	0.0	0.091	0.049	0.000	0.000	0.002
1000	200				0.3	0.093	0.054	0.024	0.011	0.002

1000	200	0.5	0.0	0.0	0.0	0.080	0.045	0.005	0.002	0.005
1000	200				0.3	0.081	0.049	0.075	0.047	0.005
1000	200			0.3	0.0	0.086	0.050	0.005	0.002	0.005
1000	200				0.3	0.084	0.055	0.086	0.044	0.005
1000	200		0.7	0.0	0.0	0.091	0.054	0.004	0.002	0.004
1000	200				0.3	0.093	0.055	0.061	0.035	0.004
1000	200			0.3	0.0	0.092	0.053	0.003	0.002	0.006
1000	200				0.3	0.095	0.059	0.080	0.048	0.007
1000	200	0.7	0.0	0.0	0.0	0.122	0.083	0.000	0.000	0.001
1000	200				0.3	0.124	0.087	0.022	0.011	0.001
1000	200			0.3	0.0	0.138	0.091	0.001	0.000	0.001
1000	200				0.3	0.145	0.094	0.031	0.015	0.001
1000	200		0.7	0.0	0.0	0.162	0.121	0.000	0.000	0.001
1000	200				0.3	0.162	0.117	0.015	0.010	0.001
1000	200			0.3	0.0	0.166	0.122	0.000	0.000	0.001
1000	200				0.3	0.165	0.118	0.027	0.009	0.002

Note:

Entries denote the rejection rates for a nomail size of 5%. HI denotes Han and Inoue (2015)'s test, and BKW denotes Baltagi et al (2021)'s test conducted with a pre known break fraction. The scalar ω denotes the "size" of the break in the loadings. See Table 1 for explanation of α, β and ρ.

Table 7: Power of Z Rotation Test, $r = 3$

						Z Test		W Test			
T	N	τ	ρ	α	β	Unadjusted	Adjusted	Unadjusted	Adjusted	HI Test	BKW Test
200	100	0.3	0.0	0.0	0.0	1.000	1.000	0.000	0.000	1.000	0.914
200	100				0.3	1.000	1.000	0.038	0.038	1.000	0.913
200	100			0.3	0.0	1.000	0.999	0.001	0.001	1.000	0.915
200	100				0.3	1.000	1.000	0.036	0.036	1.000	0.915
200	100		0.7	0.0	0.0	0.982	0.958	0.000	0.000	0.994	0.739
200	100				0.3	0.982	0.956	0.023	0.023	0.994	0.737
200	100			0.3	0.0	0.978	0.950	0.002	0.002	0.994	0.736
200	100				0.3	0.980	0.954	0.053	0.053	0.995	0.734

200	100	0.5	0.0	0.0	0.0	1.000	1.000	0.003	0.003	1.000	1.000
200	100				0.3	1.000	1.000	0.144	0.144	1.000	1.000
200	100			0.3	0.0	1.000	1.000	0.011	0.011	1.000	1.000
200	100				0.3	1.000	1.000	0.162	0.162	1.000	1.000
200	100		0.7	0.0	0.0	1.000	1.000	0.003	0.003	1.000	0.999
200	100				0.3	1.000	1.000	0.088	0.088	1.000	1.000
200	100			0.3	0.0	1.000	1.000	0.007	0.007	1.000	0.999
200	100				0.3	1.000	1.000	0.146	0.146	1.000	0.999
200	100	0.7	0.0	0.0	0.0	0.998	0.994	0.003	0.003	1.000	0.935
200	100				0.3	0.998	0.993	0.054	0.054	0.999	0.934
200	100			0.3	0.0	0.999	0.994	0.004	0.004	0.998	0.936
200	100				0.3	0.999	0.993	0.092	0.092	0.999	0.937
200	100		0.7	0.0	0.0	0.970	0.927	0.002	0.002	0.979	0.794
200	100				0.3	0.971	0.928	0.036	0.036	0.981	0.795
200	100			0.3	0.0	0.968	0.927	0.004	0.004	0.975	0.791
200	100				0.3	0.972	0.932	0.074	0.073	0.977	0.792
200	200	0.3	0.0	0.0	0.0	1.000	1.000	0.002	0.002	1.000	0.926
200	200				0.3	1.000	1.000	0.026	0.026	1.000	0.922
200	200			0.3	0.0	1.000	1.000	0.000	0.000	1.000	0.926
200	200				0.3	1.000	1.000	0.036	0.036	1.000	0.921
200	200		0.7	0.0	0.0	0.982	0.953	0.001	0.001	0.995	0.760
200	200				0.3	0.981	0.953	0.026	0.026	0.995	0.760
200	200			0.3	0.0	0.979	0.947	0.004	0.004	0.995	0.753
200	200				0.3	0.974	0.952	0.049	0.049	0.996	0.764
200	200	0.5	0.0	0.0	0.0	1.000	1.000	0.005	0.005	1.000	1.000
200	200				0.3	1.000	1.000	0.086	0.086	1.000	1.000
200	200			0.3	0.0	1.000	1.000	0.006	0.006	1.000	1.000
200	200				0.3	1.000	1.000	0.101	0.101	1.000	1.000
200	200		0.7	0.0	0.0	1.000	1.000	0.002	0.002	1.000	1.000
200	200				0.3	1.000	1.000	0.069	0.069	1.000	1.000
200	200			0.3	0.0	1.000	1.000	0.008	0.008	1.000	1.000
200	200				0.3	1.000	1.000	0.103	0.103	1.000	1.000

200	200	0.7	0.0	0.0	0.0	1.000	0.999	0.000	0.000	1.000	0.928
200	200				0.3	1.000	0.999	0.041	0.041	1.000	0.931
200	200			0.3	0.0	1.000	0.999	0.001	0.001	1.000	0.932
200	200				0.3	1.000	0.999	0.048	0.048	1.000	0.930
200	200		0.7	0.0	0.0	0.961	0.933	0.000	0.000	0.970	0.799
200	200				0.3	0.961	0.931	0.035	0.034	0.971	0.801
200	200			0.3	0.0	0.958	0.926	0.003	0.003	0.970	0.798
200	200				0.3	0.962	0.928	0.055	0.055	0.970	0.800
500	100	0.3	0.0	0.0	0.0	1.000	1.000	0.001	0.001	1.000	1.000
500	100				0.3	1.000	1.000	0.030	0.030	1.000	1.000
500	100			0.3	0.0	1.000	1.000	0.001	0.001	1.000	1.000
500	100				0.3	1.000	1.000	0.034	0.034	1.000	1.000
500	100		0.7	0.0	0.0	1.000	1.000	0.000	0.000	1.000	0.998
500	100				0.3	1.000	1.000	0.016	0.016	1.000	0.998
500	100			0.3	0.0	1.000	1.000	0.001	0.001	1.000	0.998
500	100				0.3	1.000	1.000	0.027	0.027	1.000	0.998
500	100	0.5	0.0	0.0	0.0	1.000	1.000	0.008	0.008	1.000	1.000
500	100				0.3	1.000	1.000	0.153	0.153	1.000	1.000
500	100			0.3	0.0	1.000	1.000	0.009	0.009	1.000	1.000
500	100				0.3	1.000	1.000	0.175	0.175	1.000	1.000
500	100		0.7	0.0	0.0	1.000	1.000	0.004	0.004	1.000	1.000
500	100				0.3	1.000	1.000	0.083	0.083	1.000	1.000
500	100			0.3	0.0	1.000	1.000	0.012	0.012	1.000	1.000
500	100				0.3	1.000	1.000	0.114	0.114	1.000	1.000
500	100	0.7	0.0	0.0	0.0	1.000	1.000	0.000	0.000	1.000	1.000
500	100				0.3	1.000	1.000	0.075	0.075	1.000	1.000
500	100			0.3	0.0	1.000	1.000	0.001	0.001	1.000	1.000
500	100				0.3	1.000	1.000	0.093	0.093	1.000	1.000
500	100		0.7	0.0	0.0	0.999	0.998	0.002	0.002	1.000	0.995
500	100				0.3	0.999	0.999	0.042	0.042	1.000	0.995
500	100			0.3	0.0	0.999	0.998	0.001	0.001	1.000	0.993
500	100				0.3	0.999	0.999	0.064	0.064	1.000	0.993

500	200	0.3	0.0	0.0	0.0	1.000	1.000	0.001	0.001	1.000	1.000
500	200				0.3	1.000	1.000	0.019	0.019	1.000	1.000
500	200			0.3	0.0	1.000	1.000	0.000	0.000	1.000	1.000
500	200				0.3	1.000	1.000	0.018	0.018	1.000	1.000
500	200		0.7	0.0	0.0	1.000	1.000	0.000	0.000	1.000	1.000
500	200				0.3	1.000	1.000	0.013	0.013	1.000	1.000
500	200			0.3	0.0	1.000	1.000	0.001	0.001	1.000	1.000
500	200				0.3	1.000	1.000	0.025	0.025	1.000	1.000
500	200	0.5	0.0	0.0	0.0	1.000	1.000	0.005	0.005	1.000	1.000
500	200				0.3	1.000	1.000	0.075	0.075	1.000	1.000
500	200			0.3	0.0	1.000	1.000	0.003	0.003	1.000	1.000
500	200				0.3	1.000	1.000	0.083	0.083	1.000	1.000
500	200		0.7	0.0	0.0	1.000	1.000	0.001	0.001	1.000	1.000
500	200				0.3	1.000	1.000	0.067	0.067	1.000	1.000
500	200			0.3	0.0	1.000	1.000	0.003	0.003	1.000	1.000
500	200				0.3	1.000	1.000	0.090	0.090	1.000	1.000
500	200	0.7	0.0	0.0	0.0	1.000	1.000	0.000	0.000	1.000	1.000
500	200				0.3	1.000	1.000	0.028	0.028	1.000	1.000
500	200			0.3	0.0	1.000	1.000	0.000	0.000	1.000	1.000
500	200				0.3	1.000	1.000	0.032	0.032	1.000	1.000
500	200		0.7	0.0	0.0	1.000	0.999	0.000	0.000	1.000	0.996
500	200				0.3	1.000	0.999	0.030	0.030	1.000	0.996
500	200			0.3	0.0	1.000	0.999	0.000	0.000	1.000	0.995
500	200				0.3	1.000	0.999	0.044	0.044	1.000	0.995

Note:

Entries denote the rejection rates across different simulated break types; a break type of $W \neq 0$ denotes a break in the factor loadings, $Z \neq I$ a break in the factor variance, and $W \neq 0$ and $Z \neq I$ denoting a break in both. HI denotes Han and Inoue (2015)'s test, and BKW denotes Baltagi et al (2021)'s test conducted with a pre known break fraction. The scalar ω denotes the "size" of the break in the loadings. See Table 1 for explanation of α, β and ρ.

Table 8: Power of Z and W Tests, $r = 3$, $\omega = 1$

						Z Test		W Test					
T	N	τ	ρ	α	β	Unadj.	Adj.	Unadj.	Adj.	Individual	HI	BKW	$\tilde{r}$

200	100			0.0	0.0	1.000	1.000	0.770	0.770	0.796	1.000	1.000	4.069
200	100		0.0		0.3	1.000	1.000	0.814	0.814	0.796	1.000	1.000	4.088
200	100			0.3	0.0	1.000	1.000	0.806	0.806	0.776	1.000	1.000	4.026
200	100				0.3	1.000	1.000	0.793	0.793	0.779	1.000	1.000	4.030
200	100			0.0	0.0	1.000	1.000	0.868	0.868	0.901	1.000	1.000	4.857
200	100		0.7		0.3	1.000	1.000	0.916	0.916	0.905	1.000	1.000	4.872
200	100			0.3	0.0	1.000	1.000	0.830	0.830	0.856	1.000	1.000	4.827
200	100				0.3	1.000	1.000	0.839	0.839	0.859	1.000	1.000	4.846
200	200			0.0	0.0	1.000	1.000	0.817	0.817	0.792	1.000	1.000	4.310
200	200		0.0		0.3	1.000	1.000	0.801	0.801	0.789	1.000	1.000	4.318
200	200	0.3		0.3	0.0	1.000	1.000	0.787	0.787	0.769	1.000	1.000	4.188
200	200				0.3	1.000	1.000	0.815	0.815	0.769	1.000	1.000	4.206
200	200			0.0	0.0	1.000	1.000	0.879	0.879	0.898	1.000	1.000	5.046
200	200		0.7		0.3	1.000	1.000	0.915	0.915	0.899	1.000	1.000	5.053
200	200			0.3	0.0	1.000	1.000	0.870	0.870	0.850	1.000	1.000	5.039
200	200				0.3	1.000	1.000	0.860	0.860	0.851	1.000	1.000	5.047
500	100			0.0	0.0	1.000	1.000	0.896	0.896	0.920	1.000	1.000	4.330
500	100		0.0		0.3	1.000	1.000	0.931	0.931	0.921	1.000	1.000	4.365
500	100			0.3	0.0	1.000	1.000	0.889	0.889	0.910	1.000	1.000	4.168
500	100				0.3	1.000	1.000	0.892	0.892	0.912	1.000	1.000	4.191
500	100			0.0	0.0	1.000	1.000	0.965	0.965	0.968	1.000	1.000	5.049
500	100		0.7		0.3	1.000	1.000	0.985	0.985	0.968	1.000	1.000	5.065
500	100			0.3	0.0	1.000	1.000	0.960	0.960	0.947	1.000	1.000	4.999
500	100				0.3	1.000	1.000	0.929	0.929	0.946	1.000	1.000	5.012
500	200			0.0	0.0	1.000	1.000	0.916	0.916	0.917	1.000	1.000	4.858
500	200		0.0		0.3	1.000	1.000	0.935	0.935	0.916	1.000	1.000	4.863
500	200			0.3	0.0	1.000	1.000	0.911	0.911	0.904	1.000	1.000	4.769
500	200				0.3	1.000	1.000	0.918	0.918	0.907	1.000	1.000	4.772
500	200			0.0	0.0	1.000	1.000	0.959	0.959	0.967	1.000	1.000	5.568
500	200		0.7		0.3	1.000	1.000	0.942	0.942	0.967	1.000	1.000	5.583
500	200			0.3	0.0	1.000	1.000	0.935	0.935	0.944	1.000	1.000	5.491
500	200				0.3	1.000	1.000	0.954	0.954	0.944	1.000	1.000	5.511

200	100		0.0	0.0	0.0	1.000	1.000	0.784	0.784	0.792	1.000	1.000	4.822
200	100				0.3	1.000	1.000	0.803	0.803	0.792	1.000	1.000	4.360
200	100			0.3	0.0	1.000	1.000	0.764	0.764	0.771	1.000	1.000	4.990
200	100				0.3	1.000	1.000	0.785	0.785	0.773	1.000	1.000	4.140
200	100		0.7	0.0	0.0	1.000	1.000	0.882	0.882	0.898	1.000	1.000	4.538
200	100				0.3	1.000	1.000	0.893	0.893	0.898	1.000	1.000	4.205
200	100			0.3	0.0	1.000	1.000	0.834	0.834	0.852	1.000	1.000	4.120
200	100				0.3	1.000	1.000	0.840	0.840	0.853	1.000	1.000	4.085
200	200		0.0	0.0	0.0	1.000	1.000	0.801	0.800	0.785	1.000	1.000	4.935
200	200				0.3	1.000	1.000	0.820	0.820	0.785	1.000	1.000	4.370
200	200	0.5		0.3	0.0	1.000	1.000	0.773	0.772	0.764	1.000	1.000	4.047
200	200				0.3	1.000	1.000	0.804	0.803	0.765	1.000	1.000	4.987
200	200		0.7	0.0	0.0	1.000	1.000	0.899	0.899	0.894	1.000	1.000	4.006
200	200				0.3	1.000	1.000	0.909	0.909	0.893	1.000	1.000	4.046
200	200			0.3	0.0	1.000	1.000	0.860	0.860	0.846	1.000	1.000	4.024
200	200				0.3	1.000	1.000	0.867	0.867	0.846	1.000	1.000	4.468
500	100		0.0	0.0	0.0	1.000	1.000	0.903	0.903	0.916	1.000	1.000	4.225
500	100				0.3	1.000	1.000	0.917	0.917	0.916	1.000	1.000	4.170
500	100			0.3	0.0	1.000	1.000	0.898	0.898	0.905	1.000	1.000	4.318
500	100				0.3	1.000	1.000	0.909	0.909	0.906	1.000	1.000	4.028
500	100		0.7	0.0	0.0	1.000	1.000	0.960	0.960	0.964	1.000	1.000	4.541
500	100				0.3	1.000	1.000	0.960	0.960	0.964	1.000	1.000	4.799
500	100			0.3	0.0	1.000	1.000	0.936	0.936	0.940	1.000	1.000	4.438
500	100				0.3	1.000	1.000	0.938	0.938	0.940	1.000	1.000	4.759
500	200		0.0	0.0	0.0	1.000	1.000	0.924	0.924	0.911	1.000	1.000	4.448
500	200				0.3	1.000	1.000	0.927	0.927	0.912	1.000	1.000	4.393
500	200			0.3	0.0	1.000	1.000	0.913	0.913	0.901	1.000	1.000	4.524
500	200				0.3	1.000	1.000	0.919	0.919	0.901	1.000	1.000	4.331
500	200		0.7	0.0	0.0	1.000	1.000	0.964	0.964	0.962	1.000	1.000	4.263
500	200				0.3	1.000	1.000	0.965	0.965	0.962	1.000	1.000	4.155
500	200			0.3	0.0	1.000	1.000	0.946	0.945	0.938	1.000	1.000	4.346
500	200				0.3	1.000	1.000	0.946	0.946	0.938	1.000	1.000	4.176

200	100			0.0	0.0	1.000	1.000	0.764	0.764	0.798	1.000	1.000	4.069
200	100		0.0		0.3	1.000	1.000	0.825	0.825	0.797	1.000	1.000	4.088
200	100			0.3	0.0	1.000	1.000	0.772	0.772	0.776	1.000	1.000	4.026
200	100				0.3	1.000	1.000	0.759	0.759	0.777	1.000	1.000	4.030
200	100			0.0	0.0	1.000	1.000	0.886	0.886	0.903	1.000	1.000	4.857
200	100		0.7		0.3	1.000	1.000	0.922	0.922	0.903	1.000	1.000	4.872
200	100			0.3	0.0	1.000	1.000	0.828	0.828	0.856	1.000	1.000	4.827
200	100				0.3	1.000	1.000	0.821	0.821	0.858	1.000	1.000	4.846
200	200			0.0	0.0	1.000	1.000	0.811	0.811	0.788	1.000	1.000	4.310
200	200		0.0		0.3	1.000	1.000	0.823	0.823	0.789	1.000	1.000	4.318
200	200	0.7		0.3	0.0	1.000	1.000	0.779	0.779	0.767	1.000	1.000	4.188
200	200				0.3	1.000	1.000	0.812	0.812	0.772	1.000	1.000	4.206
200	200			0.0	0.0	1.000	1.000	0.869	0.869	0.898	1.000	1.000	5.046
200	200		0.7		0.3	1.000	1.000	0.917	0.917	0.898	1.000	1.000	5.053
200	200			0.3	0.0	1.000	1.000	0.823	0.823	0.851	1.000	1.000	5.039
200	200				0.3	1.000	1.000	0.865	0.865	0.850	1.000	1.000	5.047
500	100			0.0	0.0	1.000	1.000	0.918	0.918	0.922	1.000	1.000	4.330
500	100		0.0		0.3	1.000	1.000	0.924	0.924	0.922	1.000	1.000	4.365
500	100			0.3	0.0	1.000	1.000	0.897	0.897	0.911	1.000	1.000	4.168
500	100				0.3	1.000	1.000	0.909	0.909	0.911	1.000	1.000	4.191
500	100			0.0	0.0	1.000	1.000	0.971	0.971	0.968	1.000	1.000	5.049
500	100		0.7		0.3	1.000	1.000	0.938	0.938	0.969	1.000	1.000	5.065
500	100			0.3	0.0	1.000	1.000	0.940	0.940	0.947	1.000	1.000	4.999
500	100				0.3	1.000	1.000	0.945	0.945	0.946	1.000	1.000	5.012
500	200			0.0	0.0	1.000	1.000	0.935	0.935	0.918	1.000	1.000	4.858
500	200		0.0		0.3	1.000	1.000	0.897	0.897	0.917	1.000	1.000	4.863
500	200			0.3	0.0	1.000	1.000	0.927	0.927	0.905	1.000	1.000	4.769
500	200				0.3	1.000	1.000	0.906	0.906	0.907	1.000	1.000	4.772
500	200			0.0	0.0	1.000	1.000	0.954	0.954	0.968	1.000	1.000	5.568
500	200		0.7		0.3	1.000	1.000	0.966	0.966	0.969	1.000	1.000	5.583
500	200			0.3	0.0	1.000	1.000	0.928	0.928	0.942	1.000	1.000	5.491
500	200				0.3	1.000	1.000	0.955	0.955	0.943	1.000	1.000	5.511

Note:

Entries denote the rejection rates across different simulated breaks where both $W \neq 0$ and $Z \neq I$. HI denotes Han and Inoue (2015)'s test, and BKW denotes Baltagi et al (2021)'s test conducted with a pre known break fraction. The scalar ω denotes the "size" of the break in the loadings. See Table 1 for explanation of α, β and ρ.

C Empirical

C.1 Data Description

Table 9: Data Description

Mnemonic	Description	Group	Trans.
PCDGx	Real personal consumption expenditures: Durable goods (Billions of Chained 2012 Dollars), deflated using PCE	NIPA	5
PCESVx	Real Personal Consumption Expenditures: Services (Billions of 2012 Dollars), deflated using PCE	NIPA	5
PCNDx	Real Personal Consumption Expenditures: Nondurable Goods (Billions of 2012 Dollars), deflated using PCE	NIPA	5
Y033RC1Q027SBEAx	Real Gross Private Domestic Investment: Fixed Investment: Nonresidential: Equipment (Billions of Chained 2012 Dollars), deflated using PCE	NIPA	5
PNFIx	Real private fixed investment: Nonresidential (Billions of Chained 2012 Dollars), deflated using PCE	NIPA	5
PRFIx	Real private fixed investment: Residential (Billions of Chained 2012 Dollars), deflated using PCE	NIPA	5
A014RE1Q156NBEA	Shares of gross domestic product: Gross private domestic investment: Change in private inventories (Percent)	NIPA	1
A823RL1Q225SBEA	Real Government Consumption Expenditures and Gross Investment: Federal (Percent Change from Preceding Period)	NIPA	1
FGRECPTx	Real Federal Government Current Receipts (Billions of Chained 2012 Dollars), deflated using PCE	NIPA	5
SLCEx	Real government state and local consumption expenditures (Billions of Chained 2012 Dollars), deflated using PCE	NIPA	5
EXPGSC1	Real Exports of Goods & Services, 3 Decimal (Billions of Chained 2012 Dollars)	NIPA	5
IMPGSC1	Real Imports of Goods & Services, 3 Decimal (Billions of Chained 2012 Dollars)	NIPA	5
IPDMAT	Industrial Production: Durable Materials (Index 2012=100)	Industrial Production	5
IPNMAT	Industrial Production: Nondurable Materials (Index 2012=100)	Industrial Production	5
IPDCONGD	Industrial Production: Durable Consumer Goods (Index 2012=100)	Industrial Production	5
IPB51110SQ	Industrial Production: Durable Goods: Automotive products (Index 2012=100)	Industrial Production	5

Table 9: Data Description *(continued)*

Mnemonic	Description	Group	Trans.
IPNCONGD	Industrial Production: Nondurable Consumer Goods (Index 2012=100)	Industrial Production	5
IPBUSEQ	Industrial Production: Business Equipment (Index 2012=100)	Industrial Production	5
IPB51220SQ	Industrial Production: Consumer energy products (Index 2012=100)	Industrial Production	5
TCU	Capacity Utilization: Total Industry (Percent of Capacity)	Industrial Production	1
CUMFNS	Capacity Utilization: Manufacturing (SIC) (Percent of Capacity)	Industrial Production	1
DMANEMP	All Employees: Durable goods (Thousands of Persons)	Employment	5
USCONS	All Employees: Construction (Thousands of Persons)	Employment	5
USEHS	All Employees: Education & Health Services (Thousands of Persons)	Employment	5
USFIRE	All Employees: Financial Activities (Thousands of Persons)	Employment	5
USINFO	All Employees: Information Services (Thousands of Persons)	Employment	5
USPBS	All Employees: Professional & Business Services (Thousands of Persons)	Employment	5
USLAH	All Employees: Leisure & Hospitality (Thousands of Persons)	Employment	5
USSERV	All Employees: Other Services (Thousands of Persons)	Employment	5
USMINE	All Employees: Mining and logging (Thousands of Persons)	Employment	5
USTPU	All Employees: Trade, Transportation & Utilities (Thousands of Persons)	Employment	5
USTRADE	All Employees: Retail Trade (Thousands of Persons)	Employment	5
USWTRADE	All Employees: Wholesale Trade (Thousands of Persons)	Employment	5
CES9091000001	All Employees: Government: Federal (Thousands of Persons)	Employment	5
CES9092000001	All Employees: Government: State Government (Thousands of Persons)	Employment	5
CES9093000001	All Employees: Government: Local Government (Thousands of Persons)	Employment	5
LNS14000012	Unemployment Rate - 16 to 19 years (Percent)	Employment	2
LNS14000025	Unemployment Rate - 20 years and over, Men (Percent)	Employment	2
LNS14000026	Unemployment Rate - 20 years and over, Women (Percent)	Employment	2
UEMPLT5	Number of Civilians Unemployed - Less Than 5 Weeks (Thousands of Persons)	Employment	5
UEMP5TO14	Number of Civilians Unemployed for 5 to 14 Weeks (Thousands of Persons)	Employment	5
UEMP15T26	Number of Civilians Unemployed for 15 to 26 Weeks (Thousands of Persons)	Employment	5

Table 9: Data Description *(continued)*

Mnemonic	Description	Group	Trans.
UEMP27OV	Number of Civilians Unemployed for 27 Weeks and Over (Thousands of Persons)	Employment	5
LNS13023621	Unemployment Level - Job Losers (Thousands of Persons)	Employment	5
LNS13023557	Unemployment Level - Reentrants to Labor Force (Thousands of Persons)	Employment	5
LNS13023705	Unemployment Level - Job Leavers (Thousands of Persons)	Employment	5
LNS13023569	Unemployment Level - New Entrants (Thousands of Persons)	Employment	5
LNS12032194	Employment Level - Part-Time for Economic Reasons, All Industries (Thousands of Persons)	Employment	5
AWHMAN	Average Weekly Hours of Production and Nonsupervisory Employees: Manufacturing (Hours)	Employment	1
AWHNONAG	Average Weekly Hours Of Production And Nonsupervisory Employees: Total private (Hours)	Employment	2
AWOTMAN	Average Weekly Overtime Hours of Production and Nonsupervisory Employees: Manufacturing (Hours)	Employment	2
PERMIT	New Private Housing Units Authorized by Building Permits (Thousands of Units)	Housing	5
HOUSTMW	Housing Starts in Midwest Census Region (Thousands of Units)	Housing	5
HOUSTNE	Housing Starts in Northeast Census Region (Thousands of Units)	Housing	5
HOUSTS	Housing Starts in South Census Region (Thousands of Units)	Housing	5
HOUSTW	Housing Starts in West Census Region (Thousands of Units)	Housing	5
RSAFSx	Real Retail and Food Services Sales (Millions of Chained 2012 Dollars), deflated by Core PCE	Inventories	5
AMDMNOx	Real Manufacturers' New Orders: Durable Goods (Millions of 2012 Dollars), deflated by Core PCE	Inventories	5
ACOGNOx	Real Value of Manufacturers' New Orders for Consumer Goods Industries (Million of 2012 Dollars), deflated by Core PCE	Inventories	5
AMDMUOx	Real Value of Manufacturers' Unfilled Orders for Durable Goods Industries (Million of 2012 Dollars), deflated by Core PCE	Inventories	5
ANDENOx	Real Value of Manufacturers' New Orders for Capital Goods: Nondefense Capital Goods Industries (Million of 2012 Dollars), deflated by Core PCE	Inventories	5
INVCQRMTSPL	Real Manufacturing and Trade Inventories (Millions of 2012 Dollars)	Inventories	5
GPDICTPI	Gross Private Domestic Investment: Chain-type Price Index (Index 2009=100)	Prices	6

Table 9: Data Description *(continued)*

Mnemonic	Description	Group	Trans.
IPDBS	Business Sector: Implicit Price Deflator (Index 2009=100)	Prices	6
DMOTRG3Q086SBEA	Personal consumption expenditures: Durable goods: Motor vehicles and parts (chain-type price index)	Prices	6
DFDHRG3Q086SBEA	Personal consumption expenditures: Durable goods: Furnishings and durable household equipment (chain-type price index)	Prices	6
DREQRG3Q086SBEA	Personal consumption expenditures: Durable goods: Recreational goods and vehicles (chain-type price index)	Prices	6
DODGRG3Q086SBEA	Personal consumption expenditures: Durable goods: Other durable goods (chain-type price index)	Prices	6
DFXARG3Q086SBEA	Personal consumption expenditures: Nondurable goods: Food and beverages purchased for off-premises consumption (chain-type price index)	Prices	6
DCLORG3Q086SBEA	Personal consumption expenditures: Nondurable goods: Clothing and footwear (chain-type price index)	Prices	6
DGOERG3Q086SBEA	Personal consumption expenditures: Nondurable goods: Gasoline and other energy goods (chain-type price index)	Prices	6
DONGRG3Q086SBEA	Personal consumption expenditures: Nondurable goods: Other nondurable goods (chain-type price index)	Prices	6
DHUTRG3Q086SBEA	Personal consumption expenditures: Services: Housing and utilities (chain-type price index)	Prices	6
DHLCRG3Q086SBEA	Personal consumption expenditures: Services: Health care (chain-type price index)	Prices	6
DTRSRG3Q086SBEA	Personal consumption expenditures: Transportation services (chain-type price index)	Prices	6
DRCARG3Q086SBEA	Personal consumption expenditures: Recreation services (chain-type price index)	Prices	6
DFSARG3Q086SBEA	Personal consumption expenditures: Services: Food services and accommodations (chain-type price index)	Prices	6
DIFSRG3Q086SBEA	Personal consumption expenditures: Financial services and insurance (chain-type price index)	Prices	6
DOTSRG3Q086SBEA	Personal consumption expenditures: Other services (chain-type price index)	Prices	6
WPSFD49502	Producer Price Index by Commodity for Finished Consumer Goods (Index 1982=100)	Prices	6

Table 9: Data Description *(continued)*

Mnemonic	Description	Group	Trans.
WPSFD4111	Producer Price Index by Commodity for Finished Consumer Foods (Index 1982=100)	Prices	6
PPIIDC	Producer Price Index by Commodity Industrial Commodities (Index 1982=100)	Prices	6
WPSID61	Producer Price Index by Commodity Intermediate Materials: Supplies & Components (Index 1982=100)	Prices	6
WPU0531	Producer Price Index by Commodity for Fuels and Related Products and Power: Natural Gas (Index 1982=100)	Prices	5
WPU0561	Producer Price Index by Commodity for Fuels and Related Products and Power: Crude Petroleum (Domestic Production) (Index 1982=100)	Prices	5
COMPRMS	Manufacturing Sector: Real Compensation Per Hour (Index 2009=100)	Earnings	5
COMPRNFB	Nonfarm Business Sector: Real Compensation Per Hour (Index 2009=100)	Earnings	5
RCPHBS	Business Sector: Real Compensation Per Hour (Index 2009=100)	Earnings	5
OPHMFG	Manufacturing Sector: Real Output Per Hour of All Persons (Index 2009=100)	Earnings	5
OPHNFB	Nonfarm Business Sector: Real Output Per Hour of All Persons (Index 2009=100)	Earnings	5
ULCMFG	Manufacturing Sector: Unit Labor Cost (Index 2009=100)	Earnings	5
ULCNFB	Nonfarm Business Sector: Unit Labor Cost (Index 2009=100)	Earnings	5
UNLPNBS	Nonfarm Business Sector: Unit Nonlabor Payments (Index 2009=100)	Earnings	5
FEDFUNDS	Effective Federal Funds Rate (Percent)	Interest Rates	2
TB3MS	3-Month Treasury Bill: Secondary Market Rate (Percent)	Interest Rates	2
BAA10YM	Moody's Seasoned Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity (Percent)	Interest Rates	1
MORTG10YRx	30-Year Conventional Mortgage Rate Relative to 10-Year Treasury Constant Maturity (Percent)	Interest Rates	1
TB6M3Mx	6-Month Treasury Bill Minus 3-Month Treasury Bill, secondary market (Percent)	Interest Rates	1
GS1TB3Mx	1-Year Treasury Constant Maturity Minus 3-Month Treasury Bill, secondary market (Percent)	Interest Rates	1

Table 9: Data Description *(continued)*

Mnemonic	Description	Group	Trans.
GS10TB3Mx	10-Year Treasury Constant Maturity Minus 3-Month Treasury Bill, secondary market (Percent)	Interest Rates	1
CPF3MTB3Mx	3-Month Commercial Paper Minus 3-Month Treasury Bill, secondary market (Percent)	Interest Rates	1
BUSLOANSx	Real Commercial and Industrial Loans, All Commercial Banks (Billions of 2009 U.S. Dollars), deflated by Core PCE	Money	5
CONSUMERx	Real Consumer Loans at All Commercial Banks (Billions of 2009 U.S. Dollars), deflated by Core PCE	Money	5
NONREVSLx	Total Real Nonrevolving Credit Owned and Securitized, Outstanding (Billions of Dollars), deflated by Core PCE	Money	5
REALLNx	Real Real Estate Loans, All Commercial Banks (Billions of 2009 U.S. Dollars), deflated by Core PCE	Money	5
REVOLSLx	Total Real Revolving Credit Owned and Securitized, Outstanding (Billions of 2012 Dollars), deflated by Core PCE	Money	5
DRIWCIL	FRB Senior Loans Officer Opions. Net Percentage of Domestic Respondents Reporting Increased Willingness to Make Consumer Installment Loans	Money	1
TLBSHNOx	Real Total Liabilities of Households and Nonprofit Organizations (Billions of 2012 Dollars), deflated by Core PCE	Household Balance	5
TNWBSHNOx	Real Net Worth of Households and Nonprofit Organizations (Billions of 2012 Dollars), deflated by Core PCE	Household Balance	5
TARESAx	Real Assets of Households and Nonprofit Organizations excluding Real Estate Assets (Billions of 2012 Dollars), deflated by Core PCE	Household Balance	5
HNOREMQ027Sx	Real Real Estate Assets of Households and Nonprofit Organizations (Billions of 2012 Dollars), deflated by Core PCE	Household Balance	5
TFAABSHNOx	Real Total Financial Assets of Households and Nonprofit Organizations (Billions of 2012 Dollars), deflated by Core PCE	Household Balance	5
VXOCLSx	CBOE S&P 100 Volatility Index: VXO	Stock Markets	1
USSTHPI	All-Transactions House Price Index for the United States (Index 1980 Q1=100)	Housing	5
SPCS10RSA	S&P/Case-Shiller 10-City Composite Home Price Index (Index January 2000 = 100)	Housing	5
SPCS20RSA	S&P/Case-Shiller 20-City Composite Home Price Index (Index January 2000 = 100)	Housing	5

Table 9: Data Description *(continued)*

Mnemonic	Description	Group	Trans.
TWEXAFEGSMTHx	Trade Weighted U.S. Dollar Index: Major Currencies (Index March 1973=100)	Exchange Rates	5
EXUSEU	U.S. / Euro Foreign Exchange Rate (U.S. Dollars to One Euro)	Exchange Rates	5
EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	Exchange Rates	5
EXJPUSx	Japan / U.S. Foreign Exchange Rate	Exchange Rates	5
EXUSUKx	U.S. / U.K. Foreign Exchange Rate	Exchange Rates	5
EXCAUSx	Canada / U.S. Foreign Exchange Rate	Exchange Rates	5
UMCSENTx	University of Michigan: Consumer Sentiment (Index 1st Quarter 1966=100)	Other	1
USEPUINDXM	Economic Policy Uncertainty Index for United States	Other	2

Note:

Transformation codes correspond to: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1.0)$

C.2 R^2 Variance Decomposition Exercise

We present the details for how to calculate the R^2 figures in Section 5.4. Equation (8) can be used to decompose the common component in the second regime $\widehat{X}_2$ as follows:

$$\widehat{X}_2 = \tilde{F}_2 \left(\tilde{\Lambda}_1 \tilde{Z}\right)^\intercal + \tilde{F}_2 \tilde{W}^\intercal. \tag{43}$$

which allows us to study the effect of the change in the factor loadings where we set $\tilde{W} = \mathbf{0}$ to yield

$$\widehat{X}_{2,W=\mathbf{0}} = \tilde{F}_2 \left(\tilde{\Lambda}_1 \tilde{Z}\right)^\intercal. \tag{44}$$

Combined with $\widehat{X}_1 = \tilde{F}_1 \tilde{\Lambda}_1^\intercal$ in order to produce estimates of the common components with $W = \mathbf{0}$ restriction yields

$$\widehat{X}_{W=\mathbf{0}} = \begin{bmatrix} \widehat{X}_1 \\ \widehat{X}_{2,W=\mathbf{0}} \end{bmatrix}$$

For a given estimate of $\widehat{X}$, restricted or otherwise, the corresponding R^2 is calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^N (\widehat{X}_i - X_i)^2}{\sum_{i=1}^N X_i^2}. \tag{45}$$

The intuition is as follows. If breaks in the factor loading were important for understanding variation in particular variables, this will induce a large discrepancy between the restricted and unrestricted R^2. Note that in some cases, this metric can be negative if $\sum_{i=1}^N (\widehat{X}_i - X_i)^2 > \sum_{i=1}^N X_i^2$, i.e. the fit is so poor that the residual variation is larger than the variation in the data itself.

C.3 Robustness Checks

Table 10: Great Recession (2008 Q3), 1984 Q2 - 2019 Q4 Sample, allowing changing r

		Z Test p values		W Test p values		
$\tilde{r}_1$	$\tilde{r}_2$	Unadjusted	Adjusted	Unadjusted	Adjusted	w_i test Reject Count
2	4	0.000	0.000	0.080	0.080	56

Note:

Rejection of the Z-test corresponds to a break in the factor covariance matrix, and rejection of the W-test corresponds to a break in the loadings across the entire cross section. $\tilde{r}_1$ and $\tilde{r}_2$ are estimated using the Bai and Ng (2002) criteria and Onatski (2010)'s estimator

Table 10 shows the results of the proposed test statistics allowing for a change in the number of factors, for the Great Recession subsample, where $\tilde{r}_1$ and $\tilde{r}_2$ are estimated to be 2 and 4 respectively according to the criteria of Bai and Ng (2002) and estimator of Onatski (2010). Results for the Great Moderation subsample are omitted, as $\tilde{r}_1$ and $\tilde{r}_2$ are both estimated to be 3 by Bai and Ng (2002) and Onatski (2010).

Figure 2: R^2 Statistics for unrestricted and restricted common component ($W = 0$) for Great Moderation Subsample, and Global Financial Crisis Subsample, for $r = 2$ to 6 factors.

C.4 Estimation of $tr(H_{0,1}^{\intercal}\Sigma_F H_{0,1})$ and $tr(H_{0,1}^{\intercal}Z\Sigma_F Z^{\intercal}H_{0,1})$

We detail the of how to estimate the quantities $tr(H_{0,1}^{\intercal}\Sigma_F H_{0,1})$ and $tr(H_{0,1}^{\intercal}Z\Sigma_F Z^{\intercal}H_{0,1})$.

Similar to the case of the factors themselves, cannot estimate Σ_F directly and can only estimate $H_{0,1}^{\intercal}\Sigma_F H_{0,1}$ which is an r dimensional identity matrix due to the normalization effect of $H_{0,1}$. This follows from the fact that we estimate $\tilde{f}_{1t}$ which is an estimate of $H_1^{\intercal}f_{1t}$, and hence $\frac{1}{T_1}\tilde{F}_1^{\intercal}\tilde{F}_1 = \frac{1}{\lfloor \pi T \rfloor}\sum_{t=1}^{\lfloor \pi T \rfloor}\tilde{f}_{1t}\tilde{f}_{1t}^{\intercal} = I_r$. Hence, $tr(\frac{1}{T_1}\tilde{F}_1^{\intercal}\tilde{F}_1) = tr(I_r)$ is an estimate of $tr(H_{0,1}^{\intercal}\Sigma_F H_{0,1})$.

Similarly, we cannot estimate $Z\Sigma_F Z^{\intercal}$ directly, and only up to a normalization basis. By Lemma 3 (a), $\tilde{F}_2\tilde{Z}^{\intercal}$ is an estimate of $F_2 Z^{\intercal}H_1$, and hence $\frac{1}{T_2}\tilde{Z}\tilde{F}_2^{\intercal}\tilde{F}_2\tilde{Z} = \tilde{Z}\tilde{Z}^{\intercal}$ is an estimate of $\frac{1}{T-\lfloor \pi T \rfloor}\sum_{t=\lfloor \pi T+1 \rfloor}^{T} H_{0,1}^{\intercal}Zf_{2t}f_{2t}^{\intercal}Z^{\intercal}H_{0,1}$ which converges to $Z\Sigma_F Z^{\intercal}$. Therefore, $tr(\tilde{Z}\tilde{Z}^{\intercal})$ is an estimate of $tr(H_{0,1}^{\intercal}Z\Sigma_F Z^{\intercal}H_{0,1})$.