DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

Testing for time-varying loadings in dynamic factor models

Jakob Guldbæk Mikkelsen

CREATES Research Paper 2017-22

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Testing for time-varying loadings in dynamic factor models*

Jakob Guldbæk Mikkelsen†

August 17, 2016

Abstract

In this paper we develop a test for time-varying factor loadings in factor models. The test is simple to compute and is constructed from estimated factors and residuals using the principal components estimator. The hypothesis is tested by regressing the squared residuals on the squared factors. The squared correlation coefficient times the sample size has a limiting χ^2 distribution. The test can be made robust to serial correlation in the idiosyncratic errors. We find evidence for factor loadings variance in over half of the variables in a dataset for the US economy, while there is evidence of time-varying loadings on the risk factors underlying portfolio returns for around 80% of the portfolios.

Keywords: Factor models, principal components, LM test.
JEL classification: C12, C33.

*The author acknowledges support from The Danish Council for Independent Research (DFF 4003-00022) and CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

†CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. E-mail: jmikkelsen@econ.au.dk.

1 Introduction

Large factor models have become an important tool for analysing and forecasting large economic datasets spanning several hundred of variables. Factors extracted from large panels of disaggregated macroeconomic variables explain a large part of the comovement in the series. Estimated factors are successful in summarizing the predictive content in large datasets when used in predictive regression, see e.g. Stock and Watson (2002). Examples of structural analysis using factor models are Bernanke et al. (2005) and Giannone et al. (2006). In financial applications, an underlying factor structure is often assumed, and factors can be extracted from portfolios of asset returns to form risk factors that separate the returns into to systematic and non-systematic risk. When the data spans over long periods, often several decades, the factor loadings are likely to be exhibit some instability. We propose a simple test for time-varying loadings in factor models.

The test is constructed using principal components estimates of the common factors and residuals. From a regression of the squared residuals on the squared factors we obtain the test statistics as the squared correlation coefficient R^2 times the sample size T. Under the null hypothesis of constant factor loadings, the test statistic has a limiting χ^2 distribution with degrees of freedom equal to the number of factors. The result is based on the observation that a factor model with time-varying factor loadings can be written as $x_{it} = \lambda_{it}' F_t + e_{it} = \lambda_i' F_t + \xi_{it}' F_t + e_{it}$. The observed data for variable i at time t is x_{it}, F_t is the vector of common factors, λ_{it} is the time-varying factor loadings, and e_{it} is the idiosyncratic errors. The variable $\xi_{it} = \lambda_{it} - \lambda_i$ is the stationary variations in the factor loadings around the constant λ_i. The principal components estimator estimates the factors F_t and the constant λ_i, and the residuals are thus an estimate of $\xi_{it}' F_t + e_{it}$, which we denote u_{it}. In the single factor case, the second moment of u_{it} is $E(u_{it}^2) = E(\xi_{it}^2)E(F_t^2) + E(e_{it}^2)$. A regression of the squared residuals on the squared factors will thus give an estimate of the variance of the factor loadings $E(\xi_{it}^2)$. When the loadings are constant, the R^2 from this regression is close to zero, and a large R^2 is evidence of variation in the factor loadings.

Under the condition that the number of variables N satisfies $T/N^2 \to 0$, the estimation error in the factors does not affect the limiting distribution of the test statistics. In the analysis of the test statistic, the idiosyncratic errors are assumed to be white noise. Serial correlation in the errors can be controlled for by basing the test statistic on GLS estimation of the factor model suggested by Breitung and Tenhofen (2011), and we show that the GLS version of the test has the correct size in our simulations.

A related test is the Chow test of Breitung and Eickmeier (2011) for structural breaks in the factor loadings. They consider a factor model in which the factor loadings are λ_1 for $t = 1, \ldots t^*$, and λ_2 for $t > t^*$. They find evidence of structural breaks in a large number of variables for both US and European datasets. Eickmeier et al. (2015) and Del Negro and Otrok (2008) suggest factor models in which the factor loadings are time-varying.

We consider two empirical applications of our testing procedure. We use the dataset of McCracken and Ng (2015) for the US economy, and we apply our testing procedure for different choices for the number of factors. We find evidence of time-varying factor loadings in over half of series, irrespective of the number of included factors. In the second application, we consider excess returns on 100 portfolios sorted on size and book-to-market. We find that around 80% of the portfolios are associated with time-varying factor loadings, indicating that the portfolios have time-varying exposures to the risk factors. Furthermore, the estimated factors are closely related to the three risk factors of Fama and French (1993). The squared canonical correlations between the estimated factors and the Fama French factors are all larger than 0.90.

The rest of the paper is organized as follows. In Section 2 we introduce the factor model and the test statistic. Section 3 states the assumption for the data-generating process and the main result on the limiting distribution of the test statistic. In Section 4, a Monte Carlo study shows the finite samples properties of the test statistic. In Section 5 we report results for the empirical applications. Section 6 concludes.

2 Testing for time-varying loadings

Let X_{it} denote the observed data at time $t = 1, ..., T$ for observation $i = 1, ..., N$. We consider a factor model with r common factors and time-varying factor loadings:

$$X_{it} = \lambda_{it}' F_t + e_{it},$$

where $F_t = (F_{1t}, ..., F_{rt})'$ is the r-dimensional vector of common factors, and e_{it} is the idiosyncratic error. We assume the factor loadings to be stationary and define the variable:

$$\xi_{it} = \lambda_{it} - \lambda_i,$$

where $\lambda_i = E(\lambda_{it})$ is the mean value of the loadings, and ξ_{it} is a mean-zero stationary random process. The variable ξ_{it} separates the time-varying loadings into a constant part and a time-varying part. The factor model can be written as:

$$X_{it} = \lambda_i' F_t + \xi_{it}' F_t + e_{it}. \tag{1}$$

If the factor loadings are constant over time, the variable ξ_{it} will be zero for all t with zero variance, whereas in the time-varying case, ξ_{it} will have a non-zero variance, $E(\xi_{it}^2) \neq 0$. Under the null hypothesis we therefore assume:

$$H_0 : \quad E(\xi_{it}^2) = 0,$$

and we construct a test that can detect non-zero variances of ξ_{it}, which corresponds to time-varying factor loadings. To test the null hypothesis, we form a test statistic using estimated factors and residuals. A regression of the squared residuals on the squared factors gives the test statistic TR_i^2, where R_i^2 is the squared correlation coefficient from the regression. The test statistic has a χ^2 limiting distribution with r degrees of freedom under the null of constant loadings.

To develop some intuition for the test, consider the case of a single observed factor, $r = 1$. Define the vectors $F = (F_1, ..., F_T)'$ and $X_i = (X_{i1}, ..., X_{iT})'$. As we assume in the next section, the sample average of the squared factors converges to a positive definite matrix, $T^{-1} \sum_t F_t^2 \xrightarrow{p} \Sigma_F > 0$. Since we can observe the factors, we can consider the OLS estimator for the loadings $\hat{\lambda}_i = (F'F)^{-1}F'X_i$. From (1), we have:

$$\hat{\lambda}_i \approx \lambda_i + \Sigma_F^{-1} T^{-1} \sum_{t=1}^{T} F_t^2 \xi_{it} + \Sigma_F^{-1} T^{-1} \sum_{t=1}^{T} F_t e_{it}.$$

When ξ_{it} and e_{it} both have limited serial dependence, the last two terms will converge to zero in probability, as both $E(\xi_{it}) = 0$ and $E(e_{it}) = 0$. The residuals from this regression are therefore an estimate of $\xi_{it} F_t + e_{it}$. Define $u_{it} := \xi_{it} F_t + e_{it}$, and consider the second moment of u_{it}:

$$E(u_{it}^2) = E(\xi_{it} F_t + e_{it})^2 = E(\xi_{it}^2) E(F_t^2) + E(e_{it}^2).$$

With time-varying factor loadings, the loadings variance is non-zero, $E(\xi_{it}^2) \neq 0$, and u_{it}^2 and F_t^2 will therefore be correlated. This observation shows that a regression of u_{it}^2 on F_t^2 and a constant can be used to test for time-varying loadings, because the coefficient on F_t^2 will be an estimate of the variance of the loadings. Under the null, $E(\xi_{it}^2) = 0$, and the residuals u_{it} will be equal to e_{it}. A large value of the test statistic is therefore evidence of time-varying factor loadings.

In practice, the test statistic is constructed using estimates of the unobservable quantities F_t, λ_i, and e_{it}. The principal components estimator gives estimates of the factors and the constant part of the loadings $(\tilde{F}_t, \tilde{\lambda}_i)$, as well as estimates of the idiosyncratic components, $\tilde{e}_{it} = X_{it} - \tilde{\lambda}_i' \tilde{F}_t$. The definition of the estimator is stated in the Appendix. The test statistic is obtained as TR_i^2 from the regression of $\tilde{e}_{it}^2$ on the squared principal components and a constant, and we denote this statistic as LM_i. The test statistic can be written as:

$$LM_i = T \tilde{D}_i' \tilde{B}_i^{-1} \tilde{D}_i,$$

with

$$\tilde{D}_i = T^{-1} \sum_{t=1}^{T} (\tilde{e}_{it}^2 - \tilde{\sigma}_i^2) g(\tilde{F}_t \tilde{F}_t' - \tilde{F}'\tilde{F}/T),$$
$$\tilde{B}_i = T^{-1} \sum_{t=1}^{T} [\tilde{e}_{it}^2 - \tilde{\sigma}_i^2]^2 T^{-1} \sum_{t=1}^{T} g(\tilde{F}_t \tilde{F}_t' - \tilde{F}'\tilde{F}/T) g(\tilde{F}_t \tilde{F}_t' - \tilde{F}'\tilde{F}/T)'.$$

where $g(A)$ denotes the column vector of diagonal elements of a square matrix A, and $\tilde{\sigma}_i^2 = T^{-1} \sum_{t=1}^{T} \tilde{e}_{it}^2$ is the estimator for the variance of e_{it}. Theorem 1 in the next section states that the LM_i statistic has a χ^2 limiting distribution with r degrees of freedom under the null hypothesis of constant factor loadings. The LM_i test has power against covariance-stationary forms of variation in the factor loadings, with stationary autoregressions being a leading example. We assume that the number of factor r is known. In practice, the number of factors can be estimated consistently under the null of constant loadings, e.g. by the information criteria of Bai and Ng (2002).

3 Assumptions

To establish the limiting distribution of the test statistic, we make a similar set of assumptions as in Bai (2003). Let $\|A\| = [tr(A'A)]^{1/2}$ denote the norm of matrix A. The constant $M \in (0, \infty)$ is common to all assumptions below.

Assumption A. $E\|F_t\|^4 \leq M < \infty$, and $T^{-1}\sum_{t=1}^{T} F_t F_t' \xrightarrow{p} \Sigma_F$ for some $r \times r$ positive definite matrix Σ_F.

Assumption B. $\|\lambda_i\| \leq \bar{\lambda} < \infty$, and $\|\Lambda'\Lambda/N - \Sigma_\Lambda\| \to 0$ for some positive definite matrix Σ_Λ.

Assumption C. There exists a positive constant $M < \infty$ such that for all N and T:

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

2. $E(e_s'e_t/N) = E(N^{-1}\sum_{i=1}^{N} e_{is}e_{it}) = \gamma_N(s,t)$, $|\gamma_N(s,s)| \leq M$ for all s, and $\sum_{s=1}^{T} |\gamma_N(s,t)| \leq M$ for all t.

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

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

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

Assumption D. F_t is independent of e_{is} for all (i,t,s).

Assumption E. There exists a positive constant $M < \infty$ such that for all N and T:

1. For each t,
$$E\|(NT)^{-1/2}\sum_{s=1}^{T}\sum_{k=1}^{N} F_s[e_{ks}e_{kt} - E(e_{ks}e_{kt})]\|^2 \leq M,$$

2. For each t,
$$E\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} \lambda_i e_{it}\right\|^4 \leq M,$$

3. For each i,
$$E\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} F_t e_{it}\right\|^2 \leq M,$$

4. For each i,

$$E\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_{it}(F_tF_t' \otimes F_t')\right\|^2 \leq M.$$

Assumption F. The eigenvalues of the $r \times r$ matrix $(\Sigma_\Lambda \cdot \Sigma_F)$ are distinct.

Assumptions A and B imply the existence of r common factors. Assumption C allows the idiosyncratic errors to exhibit limited serial correlation and cross-sectional dependence. If e_{it} are i.i.d., Assumption C is satisfied. Assumption D requires the factors and idiosyncratic errors to be independent, but dependence within groups is allowed. In particular, F_t can be serially correlated. Assumptions A-D permit consistent estimation of the factor space $H'F_t$ by the principal components estimator $\tilde{F}_t$, where H is an invertible matrix. The moment conditions in Assumption E are similar to Assumption F in Bai (2003). Assumption F ensures that the rotation matrix H has a unique limit. To study the limit distribution of the test statistic, we impose additional assumptions on the idiosyncratic errors, as well as a central limit theorem.

Assumption G.

1. For all t, $E(e_{it}) = \sigma_i^2$, $E(e_{it}^4) = \mu_{4,i}$, and e_{it} and e_{is} are independent for $t \neq s$.

2. For each i, as $T \to \infty$,

$$B_i^{-1/2}\sqrt{T}D_i \xrightarrow{d} N(0, I_r),$$

where

$$D_i = T^{-1}\sum_{t=1}^{T}(e_{it}^2 - \sigma_i^2)g[H'(F_tF_t' - F'F/T)H],$$

and the asymptotic covariance matrix of $\sqrt{T}D_i$ is:

$$B_i = \text{plim}_{T\to\infty}T^{-1}\sum_{t=1}^{T}E\left[(e_{it}^2 - \sigma_i^2)^2 g[H'(F_tF_t' - F'F/T)H]g[H'(F_tF_t' - F'F/T)H]'\right] > 0.$$

Assumption G implies that the limiting distribution of the infeasible test statistic $TD_iB_i^{-1}D_i'$ is χ^2 with r degrees of freedom. Since the principal components estimator $\tilde{F}_t$ is consistent for a rotation of the factors $H'F_t$, the rotation matrix H appears in the limiting distribution. Under Assumptions A-D and F, Bai (2003) show that $H \xrightarrow{p} Q^{-1}$, where Q^{-1} is an invertible matrix, so H can be replaced by Q^{-1} in Assumption G. However, since the test statistic is based on the R^2 from the regression of the squared residuals on the factors, the test statistic is invariant to the scaling of the factors, and the limiting distribution is therefore not affected by rotations of the factors.

The null distribution of the LM_i test statistic is presented in the following theorem.

Theorem 1. *Under Assumptions A-G and if $N, T \to \infty$ and $\sqrt{T}/N \to 0$, the statistic LM_i has a limiting χ^2 distribution with r degrees of freedom for each i.*

The proof is presented in the Appendix. Theorem 1 states that the LM_i statistic based on the principal components estimates $\tilde{F}_t$ and $\tilde{e}_{it}$ has the same limiting distribution as the statistic based on the population quantities F_t and e_{it}. The result follows since $\tilde{D}_i - D_i = O_p(\delta_{NT}^{-2})$, where $\delta_{NT} = min\{\sqrt{N}, \sqrt{T}\}$. The rate condition $\sqrt{T}/N \to 0$ ensures that the estimation error in $\tilde{F}_t$ does not affect the limiting distribution of the test statistic, and $\sqrt{T}\tilde{D}_i$ therefore has the same limiting distribution as $\sqrt{T}D_i$. Under Assumption G.1, the asymptotic variance B_i can be consistently estimated by $\tilde{B}_i$, and the LM_i statistics therefore has the same limiting distribution as the statistic obtained from the infeasible regression that uses F_t and e_{it} instead of $\tilde{e}_{it}$ and $\tilde{F}_t$.

3.1 Serially correlated errors

The limiting distribution of the LM_i test statistic is derived under the assumption that e_{it} is $i.i.d.$ (Assumption G.1), while F_t is allowed to exhibit serial correlation. In practice, the idiosyncratic errors can be serially correlated if the factors do not adequately capture the serial correlation in the data. When both the idiosyncratic errors e_{it} and the factors F_t are serially correlated, the asymptotic covariance matrix B_i is not valid. If e_{it} exhibits serial correlation, e_{it}^2 will also be serially correlated, and we have:

$$E\left[g[H'(F_sF_s' - F'F/T)H](e_{it}^2 - \sigma_i^2)(e_{is}^2 - \sigma_i^2)g[H'(F_sF_s' - F'F/T)H]'\right] \neq 0 \quad \text{for} \quad t \neq s.$$

Instead the covariance matrix of $\sqrt{T}B_i$ takes the form:

$$B_i^* = \text{plim}_{T\to\infty} T^{-1} \sum_{t=1}^{T}\sum_{s=1}^{T} E[g[H'(F_tF_t' - F'F/T)H](e_{it}^2 - \sigma_i^2)(e_{is}^2 - \sigma_i^2)g[H'(F_sF_s' - F'F/T)H]'].$$

The consequence is that the size of the test will be affected. Assumptions A-E are sufficient to ensure that $\tilde{D}_i$ converges to D_i, and D_i converges to zero even with serially correlated errors. The asymptotic power of the test is unaffected as serial correlation in e_{it} only affects the asymptotic covariance matrix of $\sqrt{T}D_i$.

To improve the size properties of the LM_i test when Assumption G.1 is violated, the test can be based on GLS estimation of the factor model. By fitting an auxiliary model to the residuals, we can capture the idiosyncratic dynamics in the errors. The estimated dynamics can then be used to perform a GLS transformation of the factor model. The GLS residuals will resemble white noise if the auxiliary model for the errors captures the idiosyncratic dynamics, and the LM_i test based on the GLS transformed model will therefore have better size properties.

We follow Breitung and Tenhofen (2011) and specify individual-specific $AR(p_i)$ models for the idiosyncratic components. After obtaining the initial principal components estimates $\tilde{e}_{it}$ of the idiosyncratic errors, we estimate an $AR(p_i)$ model for the residuals by least squares:

$$\tilde{e}_{it} = \rho_{i,1}\tilde{e}_{i,t-1} + ... + \rho_{i,p_i}\tilde{e}_{i,t-p_i} + v_{it}.$$

The individual-specific lag lengths p_i can be determined by information criteria. Denote the resulting lag polynomial by $\tilde{\rho}_i(L)$. The GLS transformed model is:

$$\tilde{\rho}_i(L)x_{it} = \lambda_i'[\tilde{\rho}_i(L)\tilde{F}_t] + e_{it}^*,$$

where $\tilde{F}_t$ is the principal components estimator of the common factors. A new estimate $\tilde{e}_{it}^*$ of the residuals is obtained from least squares regression of the GLS transformed model. These residuals are serially uncorrelated if the $AR(p_i)$ model sufficiently approximates the correlation structure in e_{it}. To test the null hypothesis of constant loadings, the LM_i statistic is constructed from the GLS residuals $\tilde{e}_{it}^*$ and the GLS transformed factors $\tilde{\rho}_i(L)\tilde{F}_t$. Using the GLS transformed model to construct the LM_i statistic therefore gives a test statistic that is robust to serial correlation in the errors. The Monte Carlo simulations in the next section show that the GLS based statistic has an actual size very close to the nominal in the presence of serially correlated errors.

4 Small sample properties

We perform a Monte Carlo study to investigate the small sample properties of the test statistic. The simulation design is as follows:

$$\begin{aligned}
X_{it} &= \lambda_{it}'F_t + e_{it}, & \lambda_i &\sim \text{i.i.d. } \mathcal{U}(0,1),\\
(1-b_{ip}L)(\lambda_{itp}-\lambda_{ip}) &= \eta_{itp}, & \eta_{it} &\sim \text{i.i.d. } \mathcal{N}(0,\sigma_{ip}^2(1-b_{ip}^2)),\\
(1-\alpha L)e_{it} &= v_{it}, & v_t &\sim \text{i.i.d. } \mathcal{N}(0,(1-\alpha^2)I_N),\\
F_{tp} &= \rho F_{t-1,p} + u_{tp}, & u_{tp} &\sim \text{i.i.d. } \mathcal{N}(0,1-\rho^2),
\end{aligned}$$

where $i = 1, ..., N$, $t = 1, ..., T$, and $p = 1, ..., r$ are factor and loadings indices. We omit the subscript p when there is no ambiguity. The processes $\{u_{tp}\}$, $\{\eta_{it}\}$, and $\{v_t\}$ are mutually independent. We generate time-variation in the loadings by simulating them as $AR(1)$'s independent over i. The constant part of the loadings is $\lambda_i \sim$ i.i.d. $\mathcal{U}(0,1)$, and the degree of variation is determined by the variance parameter σ_{ip}^2, which is the unconditional variance of the factor loadings, $\sigma_{ip}^2 = E(\xi_{ipt}^2)$. The parameters α and ρ determine the degree of serial correlation in the idiosyncratic errors and factors, respectively. In our baseline simulations we set $\alpha = 0$ and $\rho = 0$. We also consider the effect of having $\alpha \neq 0$ and $\rho \neq 0$, in which case Assumption G.1 is not fulfilled and the asymptotic covariance matrix B_i is invalid. When computing results, we discard the first 200 observations to avoid any dependence on initial values.

Table 1 shows the empirical sizes of the LM_i statistic. The results in (a) and (b) are for a model with one and two factors, respectively. The data is generated with $\alpha = 0$, $\rho = 0$, and $\sigma_{i,p}^2 = 0$, such that the loadings are constant, and the model satisfies Assumptions A-G. The rejection frequencies are similar for all N and T. The empirical sizes are close to the nominal size, but slightly undersized for the majority of the sample sizes. The number of factors does not seem to affect the empirical size of the test. The rejection frequencies are similar in (a) and (b).

Table 2 shows results for a factor model with serial correlation in factors and the idiosyncratic errors, such that Assumption G.1 is violated. We set the parameters to $\alpha = 0.5$ and $\rho = 0.9$. Table 2 (a) reports the empirical sizes of the LM_i test. The test is seen to reject the null hypothesis too frequently when the factors and errors are serially correlated. The empirical size is larger than the nominal for all N and T, and the size distortion generally increases with T. The results in (b) are for the GLS transformed model. We choose the order of the autoregression for the residuals by AIC. The rejection frequencies are now much closer to the nominal size and the empirical sizes are similar to those in Table 1. The GLS transformation thus works well for correcting for serial correlation in the idiosyncratic errors.

To study the empirical power properties of the LM_i test, we simulate the model with time-varying factor loadings and different degrees of loading variance, $E(\xi_{it}^2) = \sigma_i^2$. The AR-parameter of the loadings is set to $b_i = 0.9$, and the variance parameter is $\sigma_i^2 = (0.1, 0.5, 1, 1.5)$. The other parameters are the same as in Table 1. Figure 1 plots the empirical power of the test for the model with one factor and time-varying factor loadings. Table 3 reports the corresponding rejection frequencies. The rejection frequencies increase monotonically with σ_i^2 for all combinations of N and T, so the larger the difference between the null, $E(\xi_{it}^2) = 0$, and the alternative, the higher is the empirical power. The sample size T is also seen to increase power. The rejection frequencies increase with T for any fixed σ_i^2 and N. The rejection rates approach 1 as T increases for all choices of loadings variance except for $\sigma^2 = 0.1$. In the case with $\sigma^2 = 0.1$, the rejection rates are around 0.35 for $T = 400$. This is, however, a very limited amount of variation in the loadings, and the difference between the null hypothesis and the alternative is small. The cross-section size has a smaller impact on the empirical power, but does tend to increase the rejection frequencies for a given σ^2 and T.

Finally, we repeat the simulations in Figure 1, but with serial correlation in errors and factors, $\alpha = 0.5$ and $\rho = 0.9$. Figure 2 shows the empirical power of the LM_i test, and Figure 3 shows the results for the GLS transformed model. The corresponding tables of rejection frequencies are in Tables 4 and 5, respectively. Serial correlation generally leads to lower rejection frequencies. For $T = 50, .., 200$, the empirical power is lower for the LM_i test compared to the results in Figure 1. For the largest sample size, serial correlation has only a minor impact. The rejection frequencies are close to one when N and T are large. When the LM_i test is based on the GLS transformed model, we see a further reduction in the rejection frequencies. The empirical power in Figure 3 is lower for all sample sizes. Otherwise, the same patterns as in Figure 1 are evident: Power increases with σ^2 and T, whereas N has a smaller effect on the rejection frequencies.

5 Empirical application

We apply our test procedure to two settings. The first is a large dataset of macroeconomic variables for the US, and the second is a dataset of portfolio returns.

The macroeconomic dataset is the FRED-MD database of McCracken and Ng (2015). The dataset contains 135 monthly variables and includes measures of real activity, prices, money and credit aggregates, interest rates, stock prices, and exchange rates. We perform the same pre-treatment of the data as in McCracken and Ng (2015). Specifically, we difference the non-stationary series to stationarity, and standardize the series to have zero mean and unit variance before extracting principal components. The reader is referred to their paper for a closer variable description and details of the pre-treatment of the series.

We apply our test procedure for the period 1984:2014. Breitung and Eickmeier (2011) test for structural breaks in the factor loadings associated with the Great Moderation using a similar dataset. They find evidence of structural breaks in over half the series in 1984, and argue that it leads to an inflation in the number of factors. Our testing procedure requires the number of factors to be constant over the sample period, and we therefore restrict the sample to 1984:2014, resulting in $T = 372$ observations for each variable. To determine the number of factors, we use the Bai and Ng (2002) IC_{p1} criterion. The number of factors is estimated to be 10, but the criterion is very flat for $r = 7, ..., 12$.[1] We therefore consider $r = 7, ..., 12$ when testing for time-varying loadings.

Table 6 shows the rejection rates, i.e. the share of the 135 variables for which we reject the null hypothesis. For $r = 10$ factors, the rejection rate is 50%, so for half of the variables we reject the null of constant factor loadings. If we increase or decrease the number of factors, the rejection rates are similar. When we redo the tests based on the GLS transformed model, the rejection rates increase slightly. For $r = 10$ the rejection rate increases to 59%, with a similar increase for the other number of included factors. In Table 7 we report the rejection frequencies for the individual t-statistics for the GLS transformed model with 10 factors. The rejection rates are highest for the first three factors with a rejection rate around 40%. For the remaining factors, the test also rejects for a non-trivial share of the variables. If we include more or fewer factors, the rejection rates for the individual t-statistics are similar. There is thus substantial evidence of time-varying factor loadings in the macroeconomic series.

For the second application we consider returns on portfolios. The dataset is from Kenneth French's website and consists of excess returns on 100 portfolios sorted on size and book-to-market. Data descriptions and details on the sorting of portfolios can be found in Fama and French (1993). The data includes $T = 636$ observations and covers 1963:1 to 2015:12. The IC_{p1} criterion results in 6 factors. We also consider $r = 1, ..., 5$ as these are common choices in the asset pricing literature.[2]

The results of the tests are shown in Table 8. The rejection rates using the LM_i test are larger than 0.80 for $r = 2, ..., 6$ factors, and 0.46 when only a single factor is included. The results based on GLS estimation are similar. The rejection rates are slightly lower, but we still reject for the majority of the variables. The results for the individual t-statistics in Table 9 also show high rejection rates. We get similar results for the t-statistics with fewer factors included. We thus identify time-varying factor loadings for a large share of the asset portfolios.

This implies that some portfolios have time-varying exposure to the underlying risk factors.

[1] We have also tried the method of Alessi et al. (2010) to determine the number of factors. The results are very sensitive to the choice of tuning parameters and do not give any clear indication of the number of factors.

[2] The method of Alessi et al. (2010) tends to pick 2-4 factors, depending on the choice of tuning parameters.

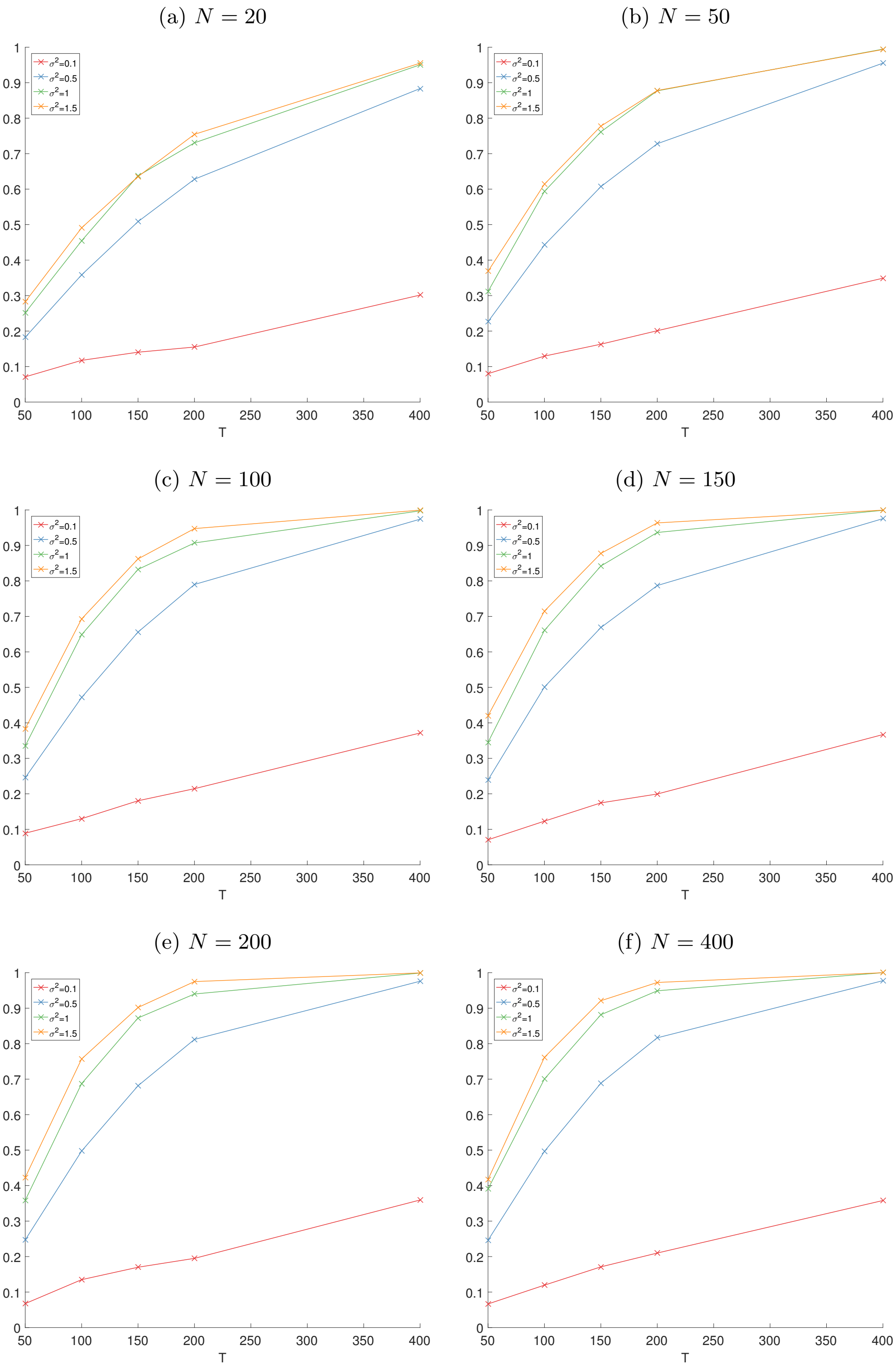
The estimated factors bear a strong resemblance to the three risk factors of Fama and French (1993): the market excess return, small minus big factor (SMB), and high minus low factor (HML). The squared canonical correlations between the three Fama French factors and the 6 factors in our analysis are 0.993, 0.951, and 0.917, respectively. Bai and Ng (2006) also find that the Fama French factors are strong proxies for systematic risk. They find canonical correlations of 0.992, 0.917, 0.832 for the period 1960-1996.

6 Conclusions

In this paper we propose a simple procedure to test for stationary variations in factor loadings. The test is based on principal components estimation of the factors and is constructed as TR^2 from a regression of the squared residuals on the squared factors. We show that under the assumption of an approximate factor model, the limiting distribution of the test statistic is unaffected by the estimation error of the common factors. The test statistic converges to a χ^2 random variable with degrees of freedom equal to the number of factors. The limiting distribution is therefore the same as if the factors could be observed. Furthermore, the test can be based on GLS estimation of the factor model such that serial correlation in the idiosyncratic errors is removed.

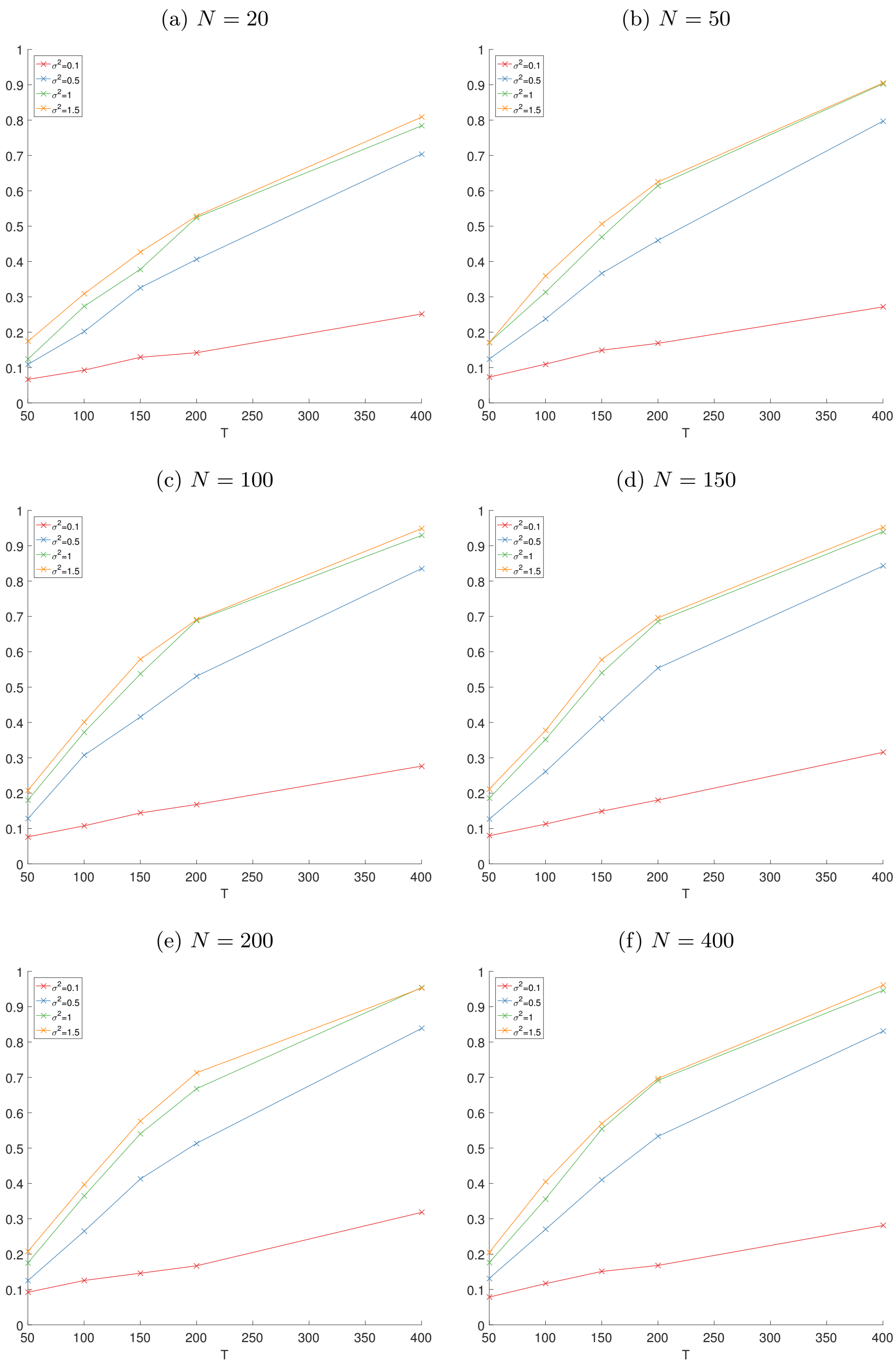
When testing for time-varying factor loadings in a large macroeconomic dataset, we find evidence for time-varying factor loadings in around half of the series. When applying the test to returns on portfolios, we find that the factor loadings are time-varying for 80% of the portfolios. These portfolios therefore have a time-varying exposure to the risk embedded in the underlying factors. Furthermore, the factors have a strong relation with the three Fama French factors with squared canonical correlations all larger than 0.90.

Figure 1: Empirical power (average rejection frequencies).



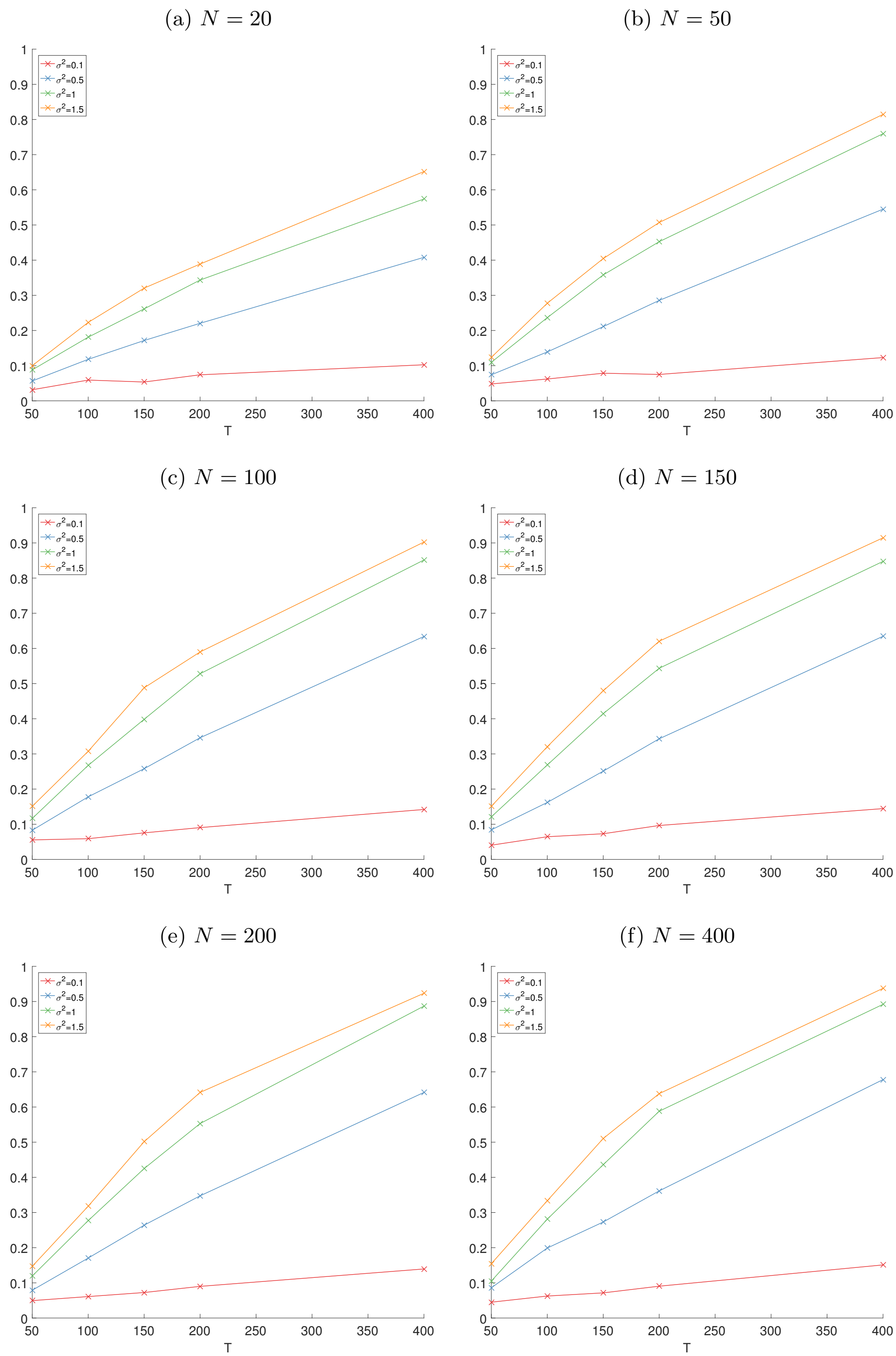
NOTE: Figures (a)-(f) plot the average rejections frequencies of rejection of the LM_i test for the model with factor loadings variances $\sigma^2 = (0.1, 0.5, 1, 1.5)$. Actual observations are marked with an "x". The lines are piecewise linear interpolations.

Figure 2: Empirical power (average rejection frequencies).
Serially correlated errors – LM_i.



NOTE: Figures (a)-(f) plot the average frequencies of rejection of the LM_i test for the model with factor loadings variances $\sigma^2 = (0.1, 0.5, 1, 1.5)$. The errors and factors are serially correlated with AR-parameters, $\alpha = 0.5$ and $\rho = 0.9$, respectively. Actual observations are marked with an "x". The lines are piecewise linear interpolations.

Figure 3: Empirical power (average rejection frequencies).
Serially correlated errors – LM_i–GLS.



Note: Figures (a)-(f) plot the average frequencies of rejection of the LM_i test based on GLS estimation for the model with factor loadings variances $\sigma^2 = (0.1, 0.5, 1, 1.5)$. The errors and factors are serially correlated with AR-parameters, $\alpha = 0.5$ and $\rho = 0.9$, respectively. Actual observations are marked with an "x". The lines are piecewise linear interpolations.

Table 1: Empirical sizes (average rejection frequencies).

(a) 1 factor

N	$T = 50$	$T = 100$	$T = 150$	$T = 200$	$T = 400$
20	0.037	0.040	0.041	0.057	0.043
50	0.039	0.037	0.048	0.041	0.050
100	0.036	0.044	0.047	0.044	0.039
150	0.035	0.043	0.042	0.043	0.044
200	0.042	0.047	0.039	0.036	0.043
400	0.035	0.038	0.045	0.040	0.055

(b) 2 factors

N	$T = 50$	$T = 100$	$T = 150$	$T = 200$	$T = 400$
20	0.049	0.044	0.051	0.048	0.058
50	0.043	0.048	0.045	0.039	0.051
100	0.044	0.039	0.049	0.044	0.053
150	0.035	0.047	0.044	0.056	0.050
200	0.048	0.036	0.047	0.044	0.050
400	0.039	0.040	0.051	0.057	0.055

NOTE: The table reports the average rejection frequencies of the LM_i test for the factor model with 1 factor (a) and 2 factors (b), and constant factor loadings, $E(\xi_{it}^2) = 0$. The nominal size is 0.05 and the results are based on 2000 replications.

Table 2: Empirical sizes (average rejection frequencies).
Serially correlated errors.

(a) LM_i

N	$T = 50$	$T = 100$	$T = 150$	$T = 200$	$T = 400$
20	0.051	0.073	0.076	0.076	0.095
50	0.065	0.076	0.077	0.092	0.098
100	0.068	0.072	0.079	0.076	0.094
150	0.067	0.075	0.086	0.076	0.092
200	0.065	0.070	0.100	0.091	0.089
400	0.070	0.077	0.079	0.082	0.105

(b) $LM_i - GLS$

N	$T = 50$	$T = 100$	$T = 150$	$T = 200$	$T = 400$
20	0.042	0.043	0.033	0.049	0.044
50	0.035	0.037	0.045	0.045	0.049
100	0.035	0.040	0.043	0.040	0.052
150	0.032	0.052	0.042	0.050	0.049
200	0.033	0.046	0.042	0.046	0.045
400	0.038	0.036	0.037	0.048	0.050

NOTE: The table reports the average rejection frequencies of the LM_i test for the factor model with 1 factor and constant factor loadings, $E(\xi_{it}^2) = 0$. The errors and factors are serially correlated with AR-parameter $\alpha = 0.5$ and $\rho = 0.9$, respectively. Results in (a) for the LM_i test, and results in (b) are for the GLS transformed model. The nominal size is 0.05 and the results are based on 2000 replications.

Table 3: Empirical power (average rejection frequencies).

(a) $N = 20$

T	50	100	150	200	400
σ^2					
0.1	0.071	0.117	0.141	0.155	0.301
0.5	0.183	0.358	0.508	0.628	0.883
1.0	0.252	0.455	0.637	0.731	0.951
1.5	0.282	0.491	0.635	0.754	0.955

(b) $N = 50$

T	50	100	150	200	400
σ^2					
0.1	0.080	0.130	0.163	0.201	0.349
0.5	0.227	0.443	0.607	0.728	0.955
1.0	0.312	0.594	0.761	0.877	0.995
1.5	0.368	0.614	0.777	0.878	0.994

(c) $N = 100$

T	50	100	150	200	400
σ^2					
0.1	0.089	0.130	0.180	0.215	0.372
0.5	0.246	0.471	0.656	0.790	0.974
1.0	0.335	0.648	0.833	0.907	0.998
1.5	0.384	0.693	0.862	0.948	1.000

(d) $N = 150$

T	50	100	150	200	400
σ^2					
0.1	0.070	0.123	0.174	0.200	0.367
0.5	0.239	0.501	0.668	0.787	0.976
1.0	0.345	0.660	0.842	0.937	0.999
1.5	0.420	0.715	0.877	0.964	1.000

(e) $N = 200$

T	50	100	150	200	400
σ^2					
0.1	0.068	0.135	0.170	0.195	0.359
0.5	0.247	0.498	0.681	0.812	0.976
1.0	0.358	0.688	0.873	0.940	0.999
1.5	0.423	0.757	0.902	0.975	1.000

(f) $N = 400$

T	50	100	150	200	400
σ^2					
0.1	0.067	0.120	0.171	0.210	0.358
0.5	0.246	0.497	0.689	0.817	0.978
1.0	0.391	0.701	0.881	0.949	1.000
1.5	0.417	0.761	0.921	0.973	1.000

NOTE: The table reports the average frequencies of rejection of the LM_i test for the model with factor loadings variance $E(\xi_{it}^2) = \sigma^2$. The results are based on 2000 replications.

Table 4: Empirical power (average rejection frequencies)
Serially correlated errors – LM_i.

(a) $N = 20$

T	50	100	150	200	400
σ^2					
0.1	0.067	0.092	0.129	0.142	0.252
0.5	0.109	0.202	0.326	0.406	0.704
1.0	0.124	0.273	0.378	0.524	0.784
1.5	0.174	0.309	0.426	0.528	0.809

(b) $N = 50$

T	50	100	150	200	400
σ^2					
0.1	0.073	0.110	0.149	0.169	0.272
0.5	0.125	0.237	0.366	0.460	0.797
1.0	0.171	0.313	0.469	0.615	0.902
1.5	0.171	0.359	0.506	0.625	0.905

(c) $N = 100$

T	50	100	150	200	400
σ^2					
0.1	0.076	0.108	0.144	0.168	0.277
0.5	0.128	0.307	0.415	0.531	0.836
1.0	0.180	0.372	0.537	0.689	0.929
1.5	0.208	0.401	0.579	0.691	0.949

(d) $N = 150$

T	50	100	150	200	400
σ^2					
0.1	0.080	0.113	0.149	0.180	0.316
0.5	0.127	0.261	0.410	0.554	0.843
1.0	0.186	0.352	0.540	0.686	0.940
1.5	0.211	0.377	0.578	0.696	0.952

(e) $N = 200$

T	50	100	150	200	400
σ^2					
0.1	0.092	0.126	0.146	0.167	0.318
0.5	0.125	0.265	0.412	0.513	0.839
1.0	0.175	0.364	0.540	0.668	0.954
1.5	0.207	0.396	0.577	0.713	0.953

(f) $N = 400$

T	50	100	150	200	400
σ^2					
0.1	0.079	0.117	0.151	0.168	0.281
0.5	0.132	0.270	0.410	0.533	0.831
1.0	0.176	0.356	0.553	0.692	0.946
1.5	0.204	0.405	0.570	0.697	0.961

NOTE: The table reports the average frequencies of rejection of the LM_i test for the model with factor loadings variance $E(\xi_{it}^2) = \sigma^2$. The errors and factors are serially correlated with AR-parameters, $\alpha = 0.5$ and $\rho = 0.9$, respectively. The results are based on 2000 replications.

Table 5: Empirical power (average rejection frequencies)
Serially correlated errors – LM_i–GLS.

(a) $N = 20$

T	50	100	150	200	400
σ^2					
0.1	0.032	0.059	0.053	0.074	0.102
0.5	0.057	0.118	0.172	0.221	0.408
1.0	0.088	0.181	0.262	0.344	0.574
1.5	0.100	0.223	0.321	0.389	0.652

(b) $N = 50$

T	50	100	150	200	400
σ^2					
0.1	0.048	0.062	0.079	0.075	0.123
0.5	0.074	0.139	0.211	0.285	0.545
1.0	0.110	0.236	0.358	0.453	0.759
1.5	0.124	0.278	0.405	0.507	0.815

(c) $N = 100$

T	50	100	150	200	400
σ^2					
0.1	0.056	0.059	0.075	0.090	0.142
0.5	0.083	0.178	0.258	0.346	0.634
1.0	0.117	0.268	0.399	0.528	0.852
1.5	0.151	0.308	0.488	0.590	0.902

(d) $N = 150$

T	50	100	150	200	400
σ^2					
0.1	0.041	0.065	0.073	0.097	0.144
0.5	0.084	0.162	0.251	0.343	0.634
1.0	0.121	0.269	0.414	0.543	0.848
1.5	0.151	0.321	0.479	0.621	0.914

(e) $N = 200$

T	50	100	150	200	400
σ^2					
0.1	0.050	0.061	0.072	0.090	0.140
0.5	0.079	0.171	0.264	0.347	0.641
1.0	0.120	0.277	0.425	0.553	0.887
1.5	0.147	0.318	0.501	0.641	0.923

(f) $N = 400$

T	50	100	150	200	400
σ^2					
0.1	0.045	0.063	0.071	0.090	0.151
0.5	0.086	0.199	0.273	0.361	0.678
1.0	0.105	0.281	0.436	0.588	0.892
1.5	0.155	0.334	0.510	0.638	0.938

NOTE: The table reports the average frequencies of rejection of the LM_i test based on GLS estimation for the model with factor loadings variance $E(\xi_{it}^2) = \sigma^2$. The errors and factors are serially correlated with AR-parameters, $\alpha = 0.5$ and $\rho = 0.9$, respectively. The results are based on 2000 replications.

Table 6: Rejection rates for LM_i-statistics – macro data.

Number of factors:	$r = 7$	$r = 8$	$r = 9$	$r = 10$	$r = 11$	$r = 12$
rej % LM	0.50	0.50	0.49	0.50	0.58	0.54
rej % LM–GLS	0.55	0.54	0.56	0.59	0.64	0.64

NOTE: 'rej % LM' is the rejection rate of the N individual LM statistics, and 'rej % LM–GLS' is the rejection rates for the GLS transformed model. The significance level is 5%.

Table 7: Rejection rates for t-statistics – macro data.

	F_1	F_2	F_3	F_4	F_5
rej % t–GLS	0.43	0.39	0.38	0.13	0.25
	F_6	F_7	F_8	F_9	F_{10}
rej % t–GLS	0.28	0.12	0.33	0.26	0.24

NOTE: 'rej % t–GLS' is the rejection rate of the N individual t– statistics on the 10 factors for the GLS transformed model. The significance level is 5%.

Table 8: Rejection rates for LM_i-statistics – portfolio data.

Number of factors:	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 5$	$r = 6$
rej % LM	0.46	0.82	0.86	0.86	0.83	0.87
rej % LM–GLS	0.46	0.81	0.81	0.80	0.79	0.78

NOTE: 'rej % LM' is the rejection rate of the N individual LM statistics, and 'rej % LM–GLS' is the rejection rates for the GLS transformed model. The significance level is 5%.

Table 9: Rejection rates for t-statistics – portfolio data.

	F_1	F_2	F_3	F_4	F_5	F_6
rej % t–GLS	0.42	0.56	0.40	0.19	0.32	0.29

NOTE: 'rej % t–GLS' is the rejection rate of the N individual t- statistics on the 6 factors for the GLS transformed model. The significance level is 5%.

Appendix

Let $X = (X_1, ..., X_T)'$ be the $T \times N$ matrix of observations, and let V_{NT} be the $r \times r$ diagonal matrix of the r largest eigenvalues of $(NT)^{-1}XX'$ in decreasing order. The principal components estimator $\tilde{F}$ is obtained as $\sqrt{T}$ times the eigenvectors corresponding to the largest r eigenvalues of the matrix XX'. By the definition of eigenvalues and eigenvectors, we have $(NT)^{-1}XX'\tilde{F} = \tilde{F}V_{NT}$ or $(NT)^{-1}XX'\tilde{F}V_{NT}^{-1} = \tilde{F}$, where $\tilde{F}'\tilde{F}/T = I_r$, and $H = (\Lambda^{0\prime}\Lambda^0/N)(F'\tilde{F}/T)V_{NT}^{-1}$ is the $r \times r$ rotation matrix. The following results are based on the identity (see Bai (2003)):

$$\tilde{F}_t - H'F_t = V_{NT}^{-1}\Big(T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t) + T^{-1}\sum_s \tilde{F}_s\zeta_{st} + T^{-1}\sum_s \tilde{F}_s\eta_{st} + T^{-1}\sum_s \tilde{F}_s\xi_{st}\Big), \quad \text{(A.1)}$$

where

- $\zeta_{st} = \frac{e_s'e_t}{N} - \gamma_N(s,t)$,
- $\eta_{st} = F_s^{0\prime}\Lambda^{0\prime}e_t/N$,
- $\xi_{st} = F_t^{0\prime}\Lambda^{0\prime}e_s/N$.

Lemma A.3 in Bai (2003) implies that $\|V_{NT}^{-1}\| = O_p(1)$ and $\|H\| = O_p(1)$. We also have $T^{-1}\sum_t \|\tilde{F}_t - H'F_t\|^2 = O_p(\delta_{NT}^{-2})$ from Lemma A.1 in Bai (2003) where $\delta_{NT} = min\{\sqrt{N}, \sqrt{T}\}$. As stated in Bai and Ng (2002), p 198, $\|\tilde{F}_t - H'F_t\|^2 = O_p(\delta_{NT}^{-2})$ if $\sum_s \gamma_N(s,t)^2 \leq M$ for all t, which we show in Lemma 1 below. These results will be used extensively in the following. It should be noted that our Assumptions D and E differ slightly from the corresponding assumtions in Bai (2003), but are still sufficient to arrive at the same results. The proof of Theorem 1 requires the following lemmas.

Lemma 1. *Under Assumption C, we have for all t and some $M \leq \infty$:*

$$\sum_s \gamma_N(s,t)^2 \leq M.$$

Proof: As in Bai and Ng (2002), let $\rho(s,t) = \gamma_N(s,t)/[\gamma_N(s,s)\gamma_N(t,t)]^{1/2}$. Then $|\rho(s,t)| \leq 1$. We can write:

$$\begin{aligned}\sum_s \gamma_N(s,t)^2 &= \sum_s \rho(s,t)^2\gamma_N(s,s)\gamma_N(t,t) \\ &\leq \sum_s |\rho(s,t)||\gamma_N(s,s)\gamma_N(t,t)|^{1/2}|\gamma_N(s,s)|^{1/2}|\gamma_N(t,t)|^{1/2} \\ &\leq M\sum_s |\rho(s,t)||\gamma_N(s,s)\gamma_N(t,t)|^{1/2} = M\sum_s |\gamma_N(s,t)| \leq M^2,\end{aligned}$$

for all t by Assumption C.2.

□

Lemma 2. *Under Assumptions A-D, the $r \times r$ matrix satisfies*

$$E\left\| \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} \sum_{k=1}^{N} F_t \lambda_k' e_{kt} \right\|^2 \leq M.$$

Proof: We can write:

$$\begin{aligned} E\left\| \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} \sum_{k=1}^{N} F_t \lambda_k' e_{kt} \right\|^2 &= (NT)^{-1} \sum_{t,s} \sum_{k,l} trE(e_{ls}\lambda_l F_s' F_t \lambda_k' e_{kt}) \\ &\leq (NT)^{-1} \sum_{t,s} \sum_{k,l} E(e_{ls}e_{kt}) E(F_s' F_t) \lambda_k' \lambda_l \\ &\leq \bar{\lambda}^2 (NT)^{-1} \sum_{t,s} \sum_{k,l} |E(e_{ls}e_{kt})| (E\|F_s\|^2)^{1/2} (\|F_t\|^2)^{1/2} \\ &\leq \bar{\lambda}^2 M (NT)^{-1} \sum_{t,s} \sum_{k,l} \tau_{kl,ts} \leq M, \end{aligned}$$

which follows from Assumptions C.4 and D.

□

Lemma 3. *Under Assumptions A-E, we have:*

$$T^{-1} \sum_{t} \|(\tilde{F}_s - H' F_t) e_{it}\|^2 = O_p(\delta_{NT}^{-2}).$$

Proof: From the identity (A.1) we have:

$$\begin{aligned} (\tilde{F}_t - H'F_t) e_{it} = V_{NT}^{-1} \Big(& T^{-1} \sum_s \tilde{F}_s \gamma_N(s,t) e_{it} + T^{-1} \sum_s \tilde{F}_s \zeta_{st} e_{it} \\ & + T^{-1} \sum_s \tilde{F}_s \eta_{st} e_{it} + T^{-1} \sum_s \tilde{F}_s \xi_{st} e_{it} \Big). \end{aligned}$$

Using Loève's inequality gives:

$$\begin{aligned} T^{-1} \sum_t \|(\tilde{F}_t - H'F_t) e_{it}\|^2 \leq 4 \|V_{NT}^{-1}\|^2 \Big(& T^{-1} \sum_t \|T^{-1} \sum_s \tilde{F}_s \gamma_N(s,t) e_{it}\|^2 \\ & + T^{-1} \sum_t \|T^{-1} \sum_s \tilde{F}_s \zeta_{st} e_{it}\|^2 + T^{-1} \sum_t \|T^{-1} \sum_s \tilde{F}_s \eta_{st} e_{it}\|^2 \\ & + T^{-1} \sum_t \|T^{-1} \sum_s \tilde{F}_s \xi_{st} e_{it}\|^2 \Big) = 4 \|V_{NT}^{-1}\|^2 (I + II + III + IV). \end{aligned}$$

Consider I:

$$
\begin{aligned}
I &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t)e_{it}\|^2 \le T^{-3}\sum_t\left(\sum_s \|\tilde{F}_s\||\gamma_N(s,t)||e_{it}|\right)^2\\
&\le T^{-2}\sum_t e_{it}^2\Big(T^{-1}\sum_s\|\tilde{F}_s\|^2\Big)\Big(\sum_s \gamma_N(s,t)^2\Big) = rT^{-2}\sum_t e_{it}^2\Big(\sum_s \gamma_N(s,t)^2\Big)\\
&\le rMT^{-2}\sum_t e_{it}^2 = O_p(T^{-1}),
\end{aligned}
$$

by Lemma 1, and the fact that $T^{-1}\sum_s \|\tilde{F}_s\|^2 = r$.

For II we have:

$$
\begin{aligned}
II &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\zeta_{st}e_{it}\|^2 \le T^{-3}\sum_t\left(\sum_s \|\tilde{F}_s\||\zeta_{st}||e_{it}|\right)^2\\
&\le T^{-2}\sum_t\left(T^{-1}\sum_s\|\tilde{F}_s\|^2\right)\left(\sum_s \zeta_{s,t}^2e_{it}^2\right) = rT^{-2}\sum_t\sum_s \zeta_{s,t}^2e_{it}^2\\
&= rN^{-1}T^{-2}\sum_t\sum_s |N^{-1/2}\sum_i [e_{it}e_{is}-E(e_{is}e_{it})]|^2e_{it}^2.
\end{aligned}
$$

The last term can be bounded in expectation:

$$
E\bigg(|N^{-1/2}\sum_i [e_{it}e_{is}-E(e_{is}e_{it})]|^2e_{it}^2\bigg) \le E|N^{-1/2}\sum_i [e_{it}e_{is}-E(e_{is}e_{it})]|^4E(e_{it}^4) \le M^2,
$$

for all s,t by Assumptions C.1 and C.5. Thus $II = O_p(N^{-1})$.

For III we have:

$$
\begin{aligned}
III &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\eta_{st}e_{it}\|^2 = N^{-1}T^{-1}\sum_t\left\|T^{-1}\sum_s \tilde{F}_sF_s'\Big(\frac{\Lambda' e_t}{\sqrt{N}}\Big)e_{it}\right\|^2\\
&\le N^{-1}T^{-1}\sum_t\left\|\frac{\Lambda' e_t}{\sqrt{N}}e_{it}\right\|^2\|T^{-1}\sum_s \tilde{F}_sF_s'\|^2\\
&\le N^{-1}T^{-1}\sum_t\left\|\frac{\Lambda' e_t}{\sqrt{N}}e_{it}\right\|^2\Big(T^{-1}\sum_s\|\tilde{F}_s\|^2\Big)\Big(T^{-1}\sum_s\|F_s\|^2\Big)\\
&\le rO_p(1)N^{-1}\Bigg(T^{-1}\sum_t\left\|\frac{\Lambda' e_t}{\sqrt{N}}\right\|^4\Bigg)^{1/2}\Bigg(T^{-1}\sum_t e_{it}^4\Bigg)^{1/2} = O_p(N^{-1}),
\end{aligned}
$$

by Assumption E.2.

For IV we have:

$$
\begin{aligned}
IV &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\xi_{st}e_{it}\|^2 = N^{-1}T^{-1}\sum_t \left\| T^{-1}\sum_s \tilde{F}_s F_t'\left(\frac{\Lambda' e_s}{\sqrt{N}}\right)e_{it}\right\|^2 \\
&\leq N^{-1}T^{-1}\sum_t \left(T^{-1}\sum_s \|\tilde{F}_s\|\|F_t\|\left\|\frac{\Lambda' e_s}{\sqrt{N}}\right\||e_{it}|\right)^2 = N^{-1}T^{-1}\sum_t \|F_t\|^2 e_{it}^2\left(T^{-1}\sum_s \|\tilde{F}_s\|\left\|\frac{\Lambda' e_s}{\sqrt{N}}\right\|\right)^2 \\
&N^{-1}\left(T^{-1}\sum_t \|F_t\|^2 e_{it}^2\right)\left(T^{-1}\sum_s \|\tilde{F}_s\|^2\right)\left(T^{-1}\sum_s \left\|\frac{\Lambda' e_s}{\sqrt{N}}\right\|^2\right) = rN^{-1}O_p(1) = O_p(N^{-1}),
\end{aligned}
$$

as $T^{-1}\sum_t \|F_t\|^2 e_{it}^2$ is bounded in expectation. Thus $I + II + III + IV = O_p(T^{-1}) + O_p(N^{-1}) = O_p(\delta_{NT}^{-2})$.

$\square$

Lemma 4. *Under Assumptions A-E, we have:*

$$
T^{-1}\sum_t \|(\tilde{F}_s - H'F_t)F_t'\|^2 = O_p(\delta_{NT}^{-2}).
$$

Proof: From the identity (A.1) we have:

$$
\begin{aligned}
(\tilde{F}_t - H'F_t)F_t' = V_{NT}^{-1}\Bigg(&T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t)F_t' + T^{-1}\sum_s \tilde{F}_s\zeta_{st}F_t' \\
&+ T^{-1}\sum_s \tilde{F}_s\eta_{st}F_t' + T^{-1}\sum_s \tilde{F}_s\xi_{st}F_t'\Bigg).
\end{aligned}
$$

Using Loève's inequality gives:

$$
\begin{aligned}
T^{-1}\sum_t \|(\tilde{F}_t - H'F_t)F_t'\|^2 \leq 4\|V_{NT}^{-1}\|^2\Bigg(&T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t)F_t'\|^2 \\
&+ T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\zeta_{st}F_t'\|^2 + T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\eta_{st}F_t'\|^2 \\
&+ T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\xi_{st}F_t'\|^2\Bigg) = 4\|V_{NT}^{-1}\|^2(I + II + III + IV).
\end{aligned}
$$

Consider I:

$$
\begin{aligned}
I &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t)F_t'\|^2 \leq T^{-3}\sum_t \left(\sum_s \|\tilde{F}_s\||\gamma_N(s,t)|\|F_t\|\right)^2 \\
&\leq T^{-2}\sum_t \|F_t\|^2\left(T^{-1}\sum_s \|\tilde{F}_s\|^2\right)\left(\sum_s \gamma_N(s,t)^2\right) \leq rMT^{-2}\sum_t \|F_t\|^2 = O_p(T^{-1}),
\end{aligned}
$$

by Lemma 1 and Assumption A.

For II we have:

$$
\begin{aligned}
II &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\zeta_{st}F_t'\|^2 \le T^{-3}\sum_t\left(\sum_s \|\tilde{F}_s\||\zeta_{st}|\|F_t\|\right)^2 \\
&\le T^{-1}\sum_t \|F_t\|^2\left(T^{-1}\sum_s\|\tilde{F}_s\|^2\right)\left(T^{-1}\sum_s \zeta_{s,t}^2\right) \\
&= rN^{-1}T^{-1}\sum_t \|F_t\|^2\left(T^{-1}\sum_s |N^{-1/2}\sum_i[e_{it}e_{is}-E(e_{it}e_{is})|^2\right) = O_p(N^{-1}),
\end{aligned}
$$

by Assumptions A and C.5.

For III we have:

$$
\begin{aligned}
III &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\eta_{st}F_t'\|^2 = N^{-1}T^{-1}\sum_t\left\|T^{-1}\sum_s \tilde{F}_sF_s'\left(N^{-1/2}\Lambda' e_tF_t'\right)\right\|^2 \\
&\le N^{-1}T^{-1}\sum_t\left\|N^{-1/2}\Lambda' e_tF_t'\right\|^2\|T^{-1}\sum_s\tilde{F}_sF_s'\|^2 \\
&\le N^{-1}T^{-1}\sum_t\left\|N^{-1/2}\Lambda' e_tF_t'\right\|^2\left(T^{-1}\sum_s\|\tilde{F}_s\|^2\right)\left(T^{-1}\sum_s\|F_s\|^2\right) \\
&= rO_p(1)N^{-1}T^{-1}\sum_t\left\|N^{-1/2}\Lambda' e_tF_t'\right\|^2.
\end{aligned}
$$

The last term can be bounded in expectation:

$$
\begin{aligned}
&T^{-1}\sum_t E\left\|N^{-1/2}\Lambda' e_tF_t'\right\|^2 = T^{-1}N^{-1}\sum_t\sum_i\sum_k E(e_{it}e_{kt})\lambda_i'\lambda_kE(F_t'F_t) \\
&\le \bar{\lambda}^2T^{-1}\sum_t E\|F_t\|^2N^{-1}\sum_i\sum_k|E(e_{it}e_{kt})| \le \bar{\lambda}^2T^{-1}\sum_t E\|F_t\|^2N^{-1}\sum_i\sum_k|\tau_{ik}| \\
&\le \bar{\lambda}^2MT^{-1}\sum_t E\|F_t\|^2 = O_p(1),
\end{aligned}
$$

by Assumptions A, B, and C.3. Thus, $III = O_p(N^{-1})$.

For IV we have:

$$
\begin{aligned}
IV &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\xi_{st}F_t'\|^2 = N^{-1}T^{-1}\sum_t\left\|T^{-1}\sum_s \tilde{F}_sF_t'\left(\frac{\Lambda' e_s}{\sqrt{N}}\right)F_t'\right\|^2 \\
&\le N^{-1}T^{-1}\sum_t\left(T^{-1}\sum_s\|\tilde{F}_s\|\|F_t\|\left\|\frac{\Lambda' e_s}{\sqrt{N}}\right\|\|F_t\|\right)^2 = N^{-1}T^{-1}\sum_t\|F_t\|^4\left(T^{-1}\sum_s\|\tilde{F}_s\|\left\|\frac{\Lambda' e_s}{\sqrt{N}}\right\|\right)^2 \\
&\le N^{-1}\left(T^{-1}\sum_t\|F_t\|^4\right)\left(T^{-1}\sum_s\|\tilde{F}_s\|^2\right)\left(T^{-1}\sum_s\left\|\frac{\Lambda' e_s}{\sqrt{N}}\right\|^2\right) = rN^{-1}O_p(1) = O_p(N^{-1}),
\end{aligned}
$$

by Assumptions A and E.2. Thus $I + II + III + IV = O_p(T^{-1}) + O_p(N^{-1}) = O_p(\delta_{NT}^{-2})$.

□

Lemma 5. *Under Assumptions A-E, we have:*

$$T^{-1}\sum_t \|(\tilde{F}_s - H'F_t)\tilde{F}_t'\|^2 = O_p(\delta_{NT}^{-2}).$$

Proof: From the identity (A.1) we have:

$$\begin{aligned}(\tilde{F}_t - H'F_t)\tilde{F}_t' = V_{NT}^{-1}\Big(&T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t)\tilde{F}_t' + T^{-1}\sum_s \tilde{F}_s\zeta_{st}\tilde{F}_t' \\ &+ T^{-1}\sum_s \tilde{F}_s\eta_{st}\tilde{F}_t' + T^{-1}\sum_s \tilde{F}_s\xi_{st}\tilde{F}_t'\Big).\end{aligned}$$

Using Loève's inequality gives:

$$\begin{aligned}T^{-1}\sum_t \|(\tilde{F}_t - H'F_t)\tilde{F}_t'\|^2 \leq 4\|V_{NT}^{-1}\|^2\Big(&T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t)\tilde{F}_t'\|^2 \\ &+ T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\zeta_{st}\tilde{F}_t'\|^2 + T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\eta_{st}\tilde{F}_t'\|^2 \\ &+ T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\xi_{st}\tilde{F}_t'\|^2\Big) = 4\|V_{NT}^{-1}\|^2(I + II + III + IV).\end{aligned}$$

Consider I:

$$\begin{aligned}I = T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\gamma_N(s,t)\tilde{F}_t'\|^2 &\leq T^{-3}\sum_t \left(\sum_s \|\tilde{F}_s\||\gamma_N(s,t)|\|\tilde{F}_t\|\right)^2 \\ \leq T^{-2}\sum_t \|\tilde{F}_t\|^2\Big(T^{-1}\sum_s \|\tilde{F}_s\|^2\Big)\Big(\sum_s \gamma_N(s,t)^2\Big) &\leq r^2MT^{-1} = O_p(T^{-1}),\end{aligned}$$

by Lemma 1.

For II we have:

$$\begin{aligned}II &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\zeta_{st}\tilde{F}_t'\|^2 \leq T^{-3}\sum_t \left(\sum_s \|\tilde{F}_s\||\zeta_{st}|\|\tilde{F}_t\|\right)^2 \\ &\leq T^{-1}\sum_t \|\tilde{F}_t\|^2\Big(T^{-1}\sum_s \|\tilde{F}_s\|^2\Big)\Big(T^{-1}\sum_s \zeta_{s,t}^2\Big) \\ &= r^2N^{-1}\Big(T^{-1}\sum_s |N^{-1/2}\sum_i [e_{it}e_{is} - E(e_{it}e_{is})|^2\Big) = O_p(N^{-1}),\end{aligned}$$

by Assumption C.5.

For III we have:

$$\begin{aligned}
III &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\eta_{st}\tilde{F}_t'\|^2 \leq T^{-1}\sum_t \|\tilde{F}_t\|^2\left(T^{-1}\sum_s \|\tilde{F}_s\||\eta_{st}|\right)^2 \\
&\leq \left(T^{-1}\sum_t \|\tilde{F}_t\|^2\right)\left(T^{-1}\sum_s \|\tilde{F}_s\|^2\right)\left(T^{-1}\sum_s \eta_{st}^2\right) \\
&= r^2 T^{-1}\sum_s \eta_{st}^2.
\end{aligned}$$

The last term is bounded in expectation:

$$T^{-1}\sum_s E(N^{-1}F_s'\Lambda' e_t)^2 \leq N^{-1}E\|N^{-1/2}\Lambda' e_t\|^2 T^{-1}\sum_s E\|F_s\|^2 = N^{-1}O_p(1),$$

by Assumption E.2, so III is $O_p(N^{-1})$.

For IV we have:

$$\begin{aligned}
IV &= T^{-1}\sum_t \|T^{-1}\sum_s \tilde{F}_s\xi_{st}\tilde{F}_t'\|^2 = N^{-1}T^{-1}\sum_t \left\|T^{-1}\sum_s \tilde{F}_s\left(N^{-1/2}F_t'\Lambda' e_s\right)\tilde{F}_t'\right\|^2 \\
&= N^{-1}T^{-1}\sum_t \left\|T^{-1}\sum_s \tilde{F}_s\left(N^{-1/2}e_s'\Lambda F_t\right)\tilde{F}_t'\right\|^2 \\
&\leq N^{-1}\left(T^{-1}\sum_t \|F_t\tilde{F}_t\|^2\right)\left\|T^{-1}\sum_s \tilde{F}_s\left(N^{-1/2}e_s'\Lambda\right)\right\|^2 \\
&\leq N^{-1}\left(T^{-1}\sum_t \|F_t\tilde{F}_t'\|^2\right)\left(T^{-1}\sum_s \|\tilde{F}_s\|^2\right)\left(T^{-1}\sum_s \|N^{-1/2}e_s'\Lambda\|^2\right) \\
&\leq 2rO_p(1)N^{-1}\left(T^{-1}\sum_t \|F_t(\tilde{F}_t - H'F_t)'\|^2\right) + 2rO_p(1)N^{-1}\left(T^{-1}\sum_t \|F_tF_t'H\|^2\right) \\
&= O_p(N^{-1}\delta_{NT}^{-2}) + O_p(N^{-1}) = O_p(N^{-1}),
\end{aligned}$$

from Lemma 4 above. Thus $I + II + III + IV = O_p(T^{-1}) + O_p(N^{-1}) = O_p(\delta_{NT}^{-2})$.

□

Lemma 6. *Under Assumptions A-E we have:*

$$T^{-1}\sum_t(\tilde{F}_t - H'F_t)e_{it}F_tF_t' = O_p(\delta_{NT}^{-2}).$$

Proof: From the identity (A.1) we have:

$$T^{-1}\sum_t(\tilde{F}_t - H'F_t)e_{it}F_tF_t' = V_{NT}^{-1}\Big(T^{-2}\sum_t\sum_s \tilde{F}_s\gamma_N(s,t)e_{it}F_tF_t' + T^{-2}\sum_t\sum_s \tilde{F}_s\zeta_{st}e_{it}F_tF_t'$$
$$+ T^{-2}\sum_t\sum_s \tilde{F}_s\eta_{st}e_{it}F_tF_t' + T^{-2}\sum_t\sum_s \tilde{F}_s\xi_{st}e_{it}F_tF_t'\Big)$$
$$= V_{NT}^{-1}(I + II + III + IV).$$

For I we can write:

$$I = T^{-2}\sum_t\sum_s(\tilde{F}_s - H'F_s)\gamma_N(s,t)e_{it}F_tF_t' + T^{-2}\sum_t\sum_s H'F_s\gamma_N(s,t)e_{it}F_tF_t'.$$

The first term is bounded by:

$$T^{-2}\sum_s\|\tilde{F}_s - H'F_s\|\Big(\sum_t\gamma_N(s,t)^2\Big)^{1/2}\Big(\sum_t e_{it}^2\|F_t\|^4\Big)^{1/2}$$
$$\le T^{-1/2}\Big(T^{-1}\sum_s\|\tilde{F}_s - H'F_s\|^2\Big)^{1/2}\Big(T^{-1}\sum_s\sum_t\gamma_N(s,t)^2T^{-1}\sum_t e_{it}^2\|F_t\|^4\Big)^{1/2}$$
$$= T^{-1/2}O_p(\delta_{NT}^{-1})O_p(1),$$

where $O_p(1)$ follows as $E(e_{it}^2)E\|F_t\|^4 = O_p(1)$ by Assumptions A and C.1, and from $\sum_s\gamma_N(s,t)^2 \le M$ for all t by Lemma 1.

The second term can be bounded in expectation (ignore H):

$$T^{-2}\sum_t\sum_s|\gamma_N(s,t)|E|e_{it}|(E\|F_s\|^2)^{1/2}(E\|F_t\|^4)^{1/2} \le MT^{-2}\sum_t\sum_s|\gamma_N(s,t)| = O_p(T^{-1}),$$

from Assumption C.2. Thus I is $O_p(T^{-1/2}\delta_{NT}^{-1})$.

For II we write:

$$T^{-2}\sum_t\sum_s(\tilde{F}_s - H'F_s)\zeta_{st}e_{it}F_tF_t' + T^{-2}\sum_t\sum_s H'F_s\zeta_{st}e_{it}F_tF_t'.$$

The first term is bounded by:

$$T^{-2}\sum_s\|\tilde{F}_s - H'F_s\|\|\sum_t\zeta_{st}e_{it}F_tF_t'\| \le \Big(T^{-1}\sum_s\|\tilde{F}_s - H'F_t\|^2\Big)^{1/2}\Big(T^{-3}\sum_s\|\sum_t\zeta_{st}e_{it}F_tF_t'\|^2\Big)^{1/2}$$
$$\le N^{-1/2}O_p(\delta_{NT}^{-1})\Big(T^{-2}\sum_s\sum_t|N^{-1/2}\sum_k[e_{ks}e_{kt} - E(e_{ks}e_{kt})]|^2T^{-1}\sum_t e_{it}^2\|F_t'\|^4\Big)^{1/2}$$
$$= N^{-1/2}O_p(\delta_{NT}^{-1})O_p(1),$$

by Assumption C.5.

For the second term we have:

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}T^{-1}\sum_t\left(\frac{1}{\sqrt{NT}}\sum_s\sum_k F_s[e_{ks}e_{kt}-E(e_{ks}e_{kt})]\right)e_{it}F_tF_t'\\
&\leq \frac{1}{\sqrt{NT}}\left(T^{-1}\sum_t\left\|\frac{1}{\sqrt{NT}}\sum_s\sum_k F_s[e_{ks}e_{kt}-E(e_{ks}e_{kt})]\right\|^2\right)^{1/2}\left(T^{-1}\sum_t e_{it}^2\|F_t\|^4\right)^{1/2}\\
&= O_p\left(\frac{1}{\sqrt{NT}}\right)O_p(1),
\end{aligned}
$$

from Assumption E.1. Thus II is $O_p(N^{-1/2}\delta_{NT}^{-1})$.

We rewrite III as:

$$
T^{-2}\sum_t\sum_s(\tilde{F}_s-H'F_s)\eta_{st}e_{it}F_tF_t'+H'T^{-2}\sum_t\sum_s F_s\eta_{st}e_{it}F_tF_t'.
$$

For the first term, we write:

$$
\begin{aligned}
&N^{-1/2}T^{-2}\sum_s(\tilde{F}_s-H'F_s)F_s'\sum_t\left(\frac{\Lambda' e_t}{\sqrt{N}}\right)e_{it}F_tF_t'\\
&\leq N^{-1/2}\left(T^{-1}\sum_s\|(\tilde{F}_s-H'F_s)F_s'\|^2\right)^{1/2}\left(T^{-3}\sum_s\|\sum_t\left(\frac{\Lambda' e_t}{\sqrt{N}}\right)e_{it}F_tF_t'\|^2\right)^{1/2}.
\end{aligned}
$$

The first parenthesis is $O_p(\delta_{NT}^{-1})$ by Lemma 4. The term inside the second parenthesis is bounded by:

$$
T^{-1}\sum_s\left(T^{-1}\sum_t\left\|\frac{\Lambda' e_t}{\sqrt{N}}\right\|^2\right)\left(T^{-1}\sum_t e_{it}^2\|F_t\|^4\right)=O_p(1),
$$

by Assumption E.2. The first term is thus $N^{-1/2}O_p(\delta_{NT}^{-1})$.

The second term can be written as:

$$
\left(T^{-1}\sum_s F_sF_s'\right)\left(N^{-1}T^{-1}\sum_t\sum_k\lambda_k e_{kt}e_{it}F_tF_t'\right).
$$

The first parenthesis is $O_p(1)$, and the second is bounded in expectation by:

$$
N^{-1}T^{-1}\sum_t\sum_k\lambda_k E(e_{kt}e_{it})E\|F_t\|^2\leq\bar{\lambda}N^{-1}T^{-1}\sum_t\sum_k|\tau_{ik}|=O_p(N^{-1}),
$$

by Assumption C.3, and III is thus $O_p(N^{-1/2}\delta_{NT}^{-1})$.

We rewrite IV as:

$$T^{-2}\sum_t\sum_s \tilde{F}_s\xi_{st}e_{it}F_tF_t' = T^{-2}\sum_t\sum_s \tilde{F}_s(F_t'\Lambda'e_s/N)e_{it}F_tF_t' = T^{-2}\sum_t\sum_s \tilde{F}_s(e_s'\Lambda/N)F_te_{it}F_tF_t'$$
$$= T^{-2}\sum_t\sum_s(\tilde{F}_s - H'F_s)(e_s'\Lambda/N)F_te_{it}F_tF_t' + H'T^{-2}\sum_t\sum_s F_s(e_s'\Lambda/N)F_te_{it}F_tF_t'.$$

The first term is bounded by:

$$N^{-1/2}T^{-2}\sum_s\|\tilde{F}_s - H'F_s\|\left\|N^{-1/2}e_s'\Lambda\right\|\sum_t|e_{it}|\|F_t\|^3$$
$$\leq N^{-1/2}\left(T^{-1}\sum_s\|\tilde{F}_s - H'F_s\|^2\right)^{1/2}\left(T^{-1}\sum_s\left\|N^{-1/2}e_s'\Lambda\right\|^2\right)^{1/2}T^{-1}\sum_t|e_{it}|\|F_t\|^3,$$

which is $N^{-1/2}O_p(\delta_{NT}^{-1})$.
For the second term we write:

$$\frac{1}{\sqrt{NT}}\left(\frac{1}{\sqrt{NT}}\sum_s\sum_k F_se_{ks}\lambda_k'\right)\left(T^{-1}\sum_t F_te_{it}F_tF_t'\right) = O_p\left(\frac{1}{\sqrt{NT}}\right),$$

by Lemma 2. Thus IV is $O_p(N^{-1/2}\delta_{NT}^{-1})$. We therefore have that $I+II+III+IV = O_p(N^{-1/2}\delta_{NT}^{-1})$ $+\, O_p(T^{-1/2}\delta_{NT}^{-1}) = O_p(\delta_{NT}^{-2})$.

□

Lemma 7. *Under Assumptions A-E we have:*

$$T^{-1}\sum_t(\tilde{F}_t - H'F_t)e_{it}^2F_t' = O_p(\delta_{NT}^{-2}).$$

Proof: From the identity (A.1) we have:

$$T^{-1}\sum_t(\tilde{F}_t - H'F_t)e_{it}^2F_t' = V_{NT}^{-1}\Bigg(T^{-2}\sum_t\sum_s\tilde{F}_s\gamma_N(s,t)e_{it}^2F_t' + T^{-2}\sum_t\sum_s\tilde{F}_s\zeta_{st}e_{it}^2F_t'$$
$$+ T^{-2}\sum_t\sum_s\tilde{F}_s\eta_{st}e_{it}^2F_t' + T^{-2}\sum_t\sum_s\tilde{F}_s\xi_{st}e_{it}^2F_t'\Bigg)$$
$$= V_{NT}^{-1}(I + II + III + IV).$$

We rewrite I as:

$$T^{-2}\sum_t\sum_s\tilde{F}_s\gamma_N(s,t)e_{it}^2F_t' = T^{-2}\sum_t\sum_s(\tilde{F}_s - H'F_s)\gamma_N(s,t)e_{it}^2F_t' + H'T^{-2}\sum_t\sum_s F_s\gamma_N(s,t)e_{it}^2F_t'.$$

The first term is bounded by:

$$
\begin{aligned}
&T^{-2}\sum_s \|\tilde{F}_s - H'F_s\| \sum_t |\gamma_N(s,t)| e_{it}^2 \|F_t\| \\
&\leq T^{-2}\sum_s \|\tilde{F}_s - H'F_s\| \Big(\sum_t |\gamma_N(s,t)|^2\Big)^{1/2}\Big(\sum_t e_{it}^4 \|F_t\|^2\Big)^{1/2} \\
&\leq T^{-1/2}\Big(T^{-1}\sum_s \|\tilde{F}_s - H'F_s\|^2\Big)^{1/2}\Big(T^{-1}\sum_s\sum_t |\gamma_N(s,t)|^2 T^{-1}\sum_t e_{it}^4\|F_t\|^2\Big)^{1/2} \\
&= T^{-1/2} O_p(\delta_{NT}^{-1}),
\end{aligned}
$$

by Assumptions A and C.1 and Lemma 1.

The second term is bounded in expectation by:

$$
\begin{aligned}
&T^{-2}\sum_t\sum_s |\gamma_N(s,t)| E(e_{it}^2)\Big(E\|F_t\|^2\Big)^{1/2}\Big(E\|F_s\|^2\Big)^{1/2} \\
&\leq MT^{-1}\Big(T^{-1}\sum_t\sum_s |\gamma_N(s,t)|\Big) = O_p(T^{-1}),
\end{aligned}
$$

by Assumption C.2. Thus I is $T^{-1/2}O_p(\delta_{NT}^{-1})$.

For II we have:

$$
T^{-2}\sum_t\sum_s \tilde{F}_s\zeta_{st}e_{it}^2 F_t' = T^{-2}\sum_t\sum_s (\tilde{F}_s - H'F_s)\zeta_{st}e_{it}^2F_t' + H'T^{-2}\sum_t\sum_s F_s\zeta_{st}e_{it}^2F_t'.
$$

The first term is bounded by:

$$
\begin{aligned}
&T^{-2}\sum_s \|\tilde{F}_s - H'F_s\| \|\sum_t \zeta_{st}e_{it}^2F_t'\| \leq \Big(T^{-1}\sum_s \|\tilde{F}_s - H'F_t\|^2\Big)^{1/2}\Big(T^{-3}\sum_s \|\sum_t \zeta_{st}e_{it}^2F_t'\|^2\Big)^{1/2} \\
&\leq N^{-1/2}O_p(\delta_{NT}^{-1})\Big(T^{-2}\sum_s\sum_t |N^{-1/2}\sum_k [e_{ks}e_{kt} - E(e_{ks}e_{kt})]|^2 T^{-1}\sum_t e_{it}^4\|F_t'\|^2\Big)^{1/2} \\
&= N^{-1/2}O_p(\delta_{NT}^{-1}),
\end{aligned}
$$

by Assumption C.5.

The second term is:

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}T^{-1}\sum_t \left(\frac{1}{\sqrt{NT}}\sum_s\sum_k F_s[e_{ks}e_{kt}-E(e_{ks}e_{kt})]\right)e_{it}^2F_t' \\
&\leq \frac{1}{\sqrt{NT}}\left(T^{-1}\sum_t\left\|\frac{1}{\sqrt{NT}}\sum_s\sum_k F_s[e_{ks}e_{kt}-E(e_{ks}e_{kt})]\right\|^2\right)^{1/2}\left(T^{-1}\sum_t e_{it}^4\|F_t\|^2\right)^{1/2} \\
&= O_p\left(\frac{1}{\sqrt{NT}}\right)O_p(1),
\end{aligned}
$$

by Assumption E.1, so II is $O_p(N^{-1/2}\delta_{NT}^{-1})$.

For III we have:

$$
T^{-2}\sum_t\sum_s \tilde{F}_s\eta_{st}e_{it}^2F_t' = T^{-2}\sum_t\sum_s(\tilde{F}_s-H'F_s)\eta_{st}e_{it}^2F_t' + H'T^{-2}\sum_t\sum_s F_s\eta_{st}e_{it}^2F_t'.
$$

We write the first term as:

$$
N^{-1/2}\left(T^{-1}\sum_s(\tilde{F}_s-H'F_s)F_s'\right)\left(T^{-1}\sum_t\left(\frac{\Lambda' e_t}{\sqrt{N}}\right)e_{it}^2F_t'\right).
$$

By Lemma B.2 in Bai (2003), the first parenthesis is $O_p(\delta_{NT}^{-2})$. The second parenthesis is bounded by:

$$
\left(T^{-1}\sum_t\left\|\frac{\Lambda' e_t}{\sqrt{N}}\right\|^2\right)^{1/2}\left(T^{-1}\sum_t e_{it}^4\|F_t\|^2\right)^{1/2} = O_p(1),
$$

by Assumption E.2. The first term of III is thus $N^{-1/2}O_p(\delta_{NT}^{-2})$.

The second term can be written as:

$$
\left(T^{-1}\sum_s F_sF_s'\right)\left(T^{-1}N^{-1}\sum_t\sum_k \lambda_k e_{kt}F_t'e_{it}^2\right).
$$

The first parenthesis is $O_p(1)$. We can bound the second parenthesis in expectation:

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}E\left(\frac{1}{\sqrt{NT}}\sum_t\sum_k \lambda_k e_{kt}F_t'\right)e_{it}^2 \\
&\leq \frac{1}{\sqrt{NT}}\left(E\left\|\frac{1}{\sqrt{NT}}\sum_t\sum_k \lambda_k e_{kt}F_t'\right\|^2 E(e_{it}^4)\right)^{1/2} = O_p\left(\frac{1}{\sqrt{NT}}\right),
\end{aligned}
$$

by Lemma 2. Thus III is $O_p(N^{-1/2}\delta_{NT}^{-1})$.

For the IV we can write:

$$T^{-2}\sum_t\sum_s \tilde{F}_s\xi_{st}e_{it}^2F_t' = T^{-2}\sum_t\sum_s \tilde{F}_s(F_t'\Lambda' e_s/N)e_{it}^2F_t' = T^{-2}\sum_t\sum_s \tilde{F}_s(e_s'\Lambda/N)F_te_{it}^2F_t'$$
$$= T^{-2}\sum_t\sum_s(\tilde{F}_s - H'F_s)(e_s'\Lambda/N)F_te_{it}^2F_t' + H'T^{-2}\sum_t\sum_s F_s(e_s'\Lambda/N)F_te_{it}^2F_t'.$$

The first term is bounded by:

$$N^{-1/2}T^{-2}\sum_s \|\tilde{F}_s - H'F_s\|\left\|N^{-1/2}e_s'\Lambda\right\|\sum_t e_{it}^2\|F_t\|^2$$
$$\leq N^{-1/2}\left(T^{-1}\sum_s\|\tilde{F}_s - H'F_s\|^2\right)^{1/2}\left(T^{-1}\sum_s\left\|N^{-1/2}e_s'\Lambda\right\|^2\right)^{1/2}T^{-1}\sum_t e_{it}^2\|F_t\|^2,$$

which is $N^{-1/2}O_p(\delta_{NT}^{-1})$.

For the second term we write:

$$\frac{1}{\sqrt{NT}}\left(\frac{1}{\sqrt{NT}}\sum_s\sum_k F_se_{ks}\lambda_k'\right)\left(T^{-1}\sum_t F_te_{it}^2F_t'\right) = O_p\left(\frac{1}{\sqrt{NT}}\right),$$

by Lemma 2, and IV is therefore $O_p(N^{-1/2}\delta_{NT}^{-1})$. Collecting results thus gives that $I + II + III + IV = O_p(N^{-1/2}\delta_{NT}^{-1}) + O_p(T^{-1/2}\delta_{NT}^{-1}) = O_p(\delta_{NT}^{-2})$.

□

Lemma 8. *Under Assumptions A-D, we have:*

a. $\|T^{-1}(\tilde{F} - FH)'e_i\|^2 = O_p(\delta_{NT}^{-2})$,

b. $\|T^{-1}(\tilde{F} - FH)'F\|^2 = O_p(\delta_{NT}^{-2})$,

c. $\|T^{-1}(\tilde{F} - FH)'\tilde{F}\|^2 = O_p(\delta_{NT}^{-2})$.

Proof: For a we have:

$$\|T^{-1}(\tilde{F} - FH)'e_i\|^2 = \|T^{-1}\sum_t(\tilde{F}_t - H'F_t)e_{it}\|^2$$
$$\leq \left(T^{-1}\sum_t\|\tilde{F}_t - H'F_t\|^2\right)\left(T^{-1}\sum_t e_{it}^2\right) = O_p(\delta_{NT}^{-2})O_p(1),$$

as $T^{-1}\sum_t e_{it}^2 = O_p(1)$ by Assumption C.1. The proof of b and c follows in the same way by using $T^{-1}\sum_t\|F_t\|^2 = O_p(1)$ and $T^{-1}\sum_t\|\tilde{F}_t\|^2 = r$.

□

Lemma 9. *Under Assumptions A-E, we have:*

$$\|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2 = O_p(\delta_{NT}^{-2}).$$

Proof: Following in Bai (2003), p 165, we can write $\tilde{\lambda}_i - H^{-1}\lambda_i$ as:

$$\tilde{\lambda}_i - H^{-1}\lambda_i = T^{-1}\tilde{F}'(FH - \tilde{F})H^{-1}\lambda_i + T^{-1}(\tilde{F} - FH)'e_i + T^{-1}H'F'e_i.$$

The norm is thus bounded by:

$$\begin{aligned}
\|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2 &\leq 3\|T^{-1}\tilde{F}'(FH - \tilde{F})H^{-1}\lambda_i\|^2 + 3\|T^{-1}(\tilde{F} - FH)'e_i\|^2 + 3\|T^{-1}H'F'e_i\|^2 \\
&\leq 3\|T^{-1}\tilde{F}'(FH - \tilde{F})\|^2\|H^{-1}\|^2\|\lambda_i\|^2 + 3\|T^{-1}(\tilde{F} - FH)'e_i\|^2 + 3\|H\|^2\|T^{-1}F'e_i\|^2 \\
&= O_p(\delta_{NT}^{-2}),
\end{aligned}$$

which follows from Lemmas 8.a and 8.c, and since $\|T^{-1}F'e_i\|^2 = T^{-1}\|T^{-1/2}F'e_i\|^2 = T^{-1}O_p(1)$ by Assumption E.3.

□

Lemma 10. *Under Assumptions A-E we have for all t:*

a. $\tilde{e}_{it} = e_{it} + O_p(\delta_{NT}^{-1})$,

b. $\tilde{C}_{it} = C_{it} + O_p(\delta_{NT}^{-1})$.

Proof: We start with *a*. We can write the principal components residual as:

$$\begin{aligned}
\tilde{e}_{it} &= x_{it} - (\tilde{\lambda}_i - H^{-1}\lambda_i)'\tilde{F}_t - \lambda_i'(H^{-1})'\tilde{F}_t \\
&= x_{it} - (\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t) - (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t - \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t) - \lambda_i'(H^{-1})'H'F_t \\
&= x_{it} - \lambda_i'F_t - (\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t) - (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t - \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t) \\
&= e_{it} - (\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t) - (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t - \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t).
\end{aligned} \tag{A.2}$$

For the second term we have:

$$|(\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)| \leq \|\tilde{\lambda}_i - H^{-1}\lambda_i\|\|\tilde{F}_t - H'F_t\| = O_p(\delta_{NT}^{-1})O_p(\delta_{NT}^{-1}),$$

by Lemma 9. The third term is bounded by:

$$|(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t| \leq \|\tilde{\lambda}_i - H^{-1}\lambda_i\|\|H\|\|F_t\| = O_p(\delta_{NT}^{-1})O_p(1)O_p(1),$$

by Lemma 9 and Assumption A. The last term is bounded by:

$$|\lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)| \leq \|\lambda_i\|\|H^{-1}\|\|\tilde{F}_t - H'F_t\| = O_p(1)O_p(1)O_p(\delta_{NT}^{-1}),$$

by Assumption B. We thus have that:

$$\tilde{e}_{it} = e_{it} + O_p(\delta_{NT}^{-1}).$$

The proof of b follows immediately by noting:

$$\tilde{C}_{it} - C_{it} = \tilde{\lambda}_i'\tilde{F}_t - \lambda_i'F_t = (x_{it} - \lambda_i'F_t) - (x_{it} - \tilde{\lambda}_i'\tilde{F}_t) = e_{it} - \tilde{e}_{it} = O_p(\delta_{NT}^{-1}).$$

□

Lemma 11. *Under Assumtions A-E we have*

$$T^{-1}\sum_t \tilde{e}_{it}^2 = T^{-1}\sum_t e_{it}^2 + O_p(\delta_{NT}^{-2}).$$

Proof: We can write $T^{-1}\sum_t \tilde{e}_{it}^2$ as:

$$\begin{aligned} T^{-1}\sum_t \tilde{e}_{it}^2 &= T^{-1}\sum_t (x_{it} - \tilde{C}_{it})^2 = T^{-1}\sum_t (x_{it} - C_{it} + C_{it} - \tilde{C}_{it})^2 \\ &= T^{-1}\sum_t (x_{it} - C_{it})^2 + T^{-1}\sum_t (C_{it} - \tilde{C}_{it})^2 + 2T^{-1}\sum_t e_{it}(C_{it} - \tilde{C}_{it})^2 \\ &= T^{-1}\sum_t e_{it}^2 + T^{-1}\sum_t (C_{it} - \tilde{C}_{it})^2 + 2T^{-1}\sum_t (C_{it} - \tilde{C}_{it})e_{it}. \end{aligned}$$

We start with $T^{-1}\sum_t (C_{it} - \tilde{C}_{it})e_{it}$. Using (A.2) we can write:

$$\begin{aligned} T^{-1}\sum_t (C_{it} - \tilde{C}_{it})e_{it} = &-T^{-1}\sum_t (\tilde{\lambda}_i + H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)e_{it} - T^{-1}\sum_t (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t e_{it} \\ &- T^{-1}\sum_t \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)e_{it}. \end{aligned}$$

For the first term we write:

$$-T^{-1}\sum_t (\tilde{\lambda}_i + H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)e_{it} = -(\tilde{\lambda}_i + H^{-1}\lambda_i)'T^{-1}\sum_t (\tilde{F}_t - H'F_t)e_{it}.$$

Bai (2003), p 165, shows that $(\tilde{\lambda}_i + H^{-1}\lambda_i) = O_p(T^{-1/2})$, and Lemma B.1, also in Bai (2003), states that $T^{-1}\sum_t (\tilde{F}_t - H'F_t)e_{it} = O_p(\delta_{NT}^{-2})$. The first term is thus $O_p(T^{-1/2})O_p(\delta_{NT}^{-2})$.

The second term is

$$-T^{-1}\sum_t (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t e_{it} = -(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'T^{-1}\sum_t F_t e_{it} = O_p(T^{-1/2})O_p(T^{-1/2}),$$

again from Bai (2003), p 165 and Assumption E.3.

The last term is:

$$-T^{-1}\sum_t \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)e_{it} = -\lambda_i'(H^{-1})'T^{-1}\sum_t(\tilde{F}_t - H'F_t)e_{it} = O_p(\delta_{NT}^{-2}),$$

by Assumption B and Lemma B.1 in Bai (2003). Thus $T^{-1}\sum_t(C_{it}-\tilde{C}_{it})e_{it} = O_p(T^{-1/2})O_p(\delta_{NT}^{-2}) + O_p(T^{-1/2})O_p(T^{-1/2}) + O_p(\delta_{NT}^{-2}) = O_p(\delta_{NT}^{-2})$.

For $T^{-1}\sum_t(C_{it} - \tilde{C}_{it})^2$ we again use (A.2). By Loève's inequality we get:

$$\begin{aligned}
T^{-1}\sum_t(C_{it} - \tilde{C}_{it})^2 = T^{-1}\sum_t &((\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t) \\
&+ (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t + \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t))^2 \\
&\leq 3T^{-1}\sum_t \|(\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)\|^2 + 3T^{-1}\sum_t \|(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t\|^2 \\
&+ 3T^{-1}\sum_t \|\lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)\|^2.
\end{aligned}$$

The first term is:

$$T^{-1}\sum_t \|(\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)\|^2 \leq \|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2 T^{-1}\sum_t \|\tilde{F}_t - H'F_t\|^2 = O_p(\delta_{NT}^{-2})O_p(\delta_{NT}^{-2}),$$

by Lemma 9.

The second term is:

$$T^{-1}\sum_t \|(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t\|^2 \leq \|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2\|H\|^2 T^{-1}\sum_t \|F_t\|^2 = O_p(\delta_{NT}^{-2}),$$

by Lemma 9 and Assumption A.

The last term is:

$$T^{-1}\sum_t \|\lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)\|^2 \leq \|\lambda_i\|^2\|H^{-1}\|^2 T^{-1}\sum_t \|\tilde{F}_t - H'F_t\|^2 = O_p(\delta_{NT}^{-2}),$$

by Assumption B. We thus have $T^{-1}\sum_t(C_{it} - \tilde{C}_{it})^2 = O_p(\delta_{NT}^{-2})$.

□

Lemma 12. *Under Assumtions A-E and if $T/N^2 \to 0$, we have:*

$$T^{-1}\sum_t \tilde{e}_{it}^4 = T^{-1}\sum_t e_{it}^4 + O_p(T/N^2) + O_p(\delta_{NT}^{-1}).$$

Proof: We can write $T^{-1}\sum_t \tilde{e}_{it}^4$ as:

$$
\begin{aligned}
T^{-1}\sum_t \tilde{e}_{it}^4 &= T^{-1}\sum_t (x_{it}-\tilde{C}_{it})^4 = T^{-1}\sum_t (x_{it}-C_{it}+C_{it}-\tilde{C}_{it})^4 \\
&= T^{-1}\sum_t (e_{it}+C_{it}-\tilde{C}_{it})^4 = T^{-1}\sum_t e_{it}^4 + T^{-1}\sum_t (C_{it}-\tilde{C}_{it})^4 \\
&+ 6T^{-1}\sum_t e_{it}^2(C_{it}-\tilde{C}_{it})^2 + 4T^{-1}\sum_t e_{it}^3(C_{it}-\tilde{C}_{it}) + 4T^{-1}\sum_t e_{it}(C_{it}-\tilde{C}_{it})^3 \\
&= T^{-1}\sum_t e_{it}^4 + I + II + III + IV.
\end{aligned}
$$

Using Loève's inequality, I can be written as:

$$
\begin{aligned}
T^{-1}\sum_t (C_{it}-\tilde{C}_{it})^4 = T^{-1}\sum_t \Big[&(\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t) \\
&+ (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t + \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)\Big]^4 \\
&\leq 27T^{-1}\sum_t [(\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)]^4 + 27T^{-1}\sum_t [(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t]^4 \\
&+ 27T^{-1}\sum_t [\lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)]^4.
\end{aligned}
$$

The first term is bounded by:

$$
\begin{aligned}
T^{-1}\sum_t [(\tilde{\lambda}_i - H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)]^4 &\leq \|\tilde{\lambda}_i - H^{-1}\lambda_i\|^4 \max_t \|\tilde{F}_t - H'F_t\|^2 T^{-1}\sum_t \|\tilde{F}_t - H'F_t\|^2 \\
&= O_p(\delta_{NT}^{-4})O_p(\delta_{NT}^{-2})\max_t \|\tilde{F}_t - H'F_t\|^2.
\end{aligned}
$$

By Proposition 2 in Bai (2003), $\max_t \|\tilde{F}_t - H'F_t\|^2 = O_p(T^{-1}) + O_p(T/N)$, so the term is $O_p(\delta_{NT}^{-6})(O_p(T^{-1}) + O_p(T/N))$.

The second term is:

$$
T^{-1}\sum_t [(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t]^4 \leq \|\tilde{\lambda}_i - H^{-1}\lambda_i\|^4 \|H\|^4 T^{-1}\sum_t \|F_t\|^4 = O_p(\delta_{NT}^{-4}),
$$

by Lemma 9 and Assumption A.

The third term is:

$$
\begin{aligned}
T^{-1}\sum_t [\lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)]^4 &\leq \|\lambda_i\|^4 \|H^{-1}\|^4 T^{-1}\sum_t \|\tilde{F}_t - H'F_t\|^4 \\
&\leq O_p(1)\max_t \|\tilde{F}_t - H'F_t\|^2 T^{-1}\sum_t \|\tilde{F}_t - H'F_t\|^2 \\
&= [O_p(T^{-1}) + O_p(T/N)]O_p(\delta_{NT}^{-2}).
\end{aligned}
$$

Now $[O_p(T^{-1})+O_p(T/N)]O_p(\delta_{NT}^{-2})=O_p(T^{-1})+O_p(T/N^2)$, so $I=O_p(T^{-1})+O_p(T/N^2)$.

For II we have:

$$\begin{aligned}T^{-1}\sum_t e_{it}^2(C_{it}-\tilde{C}_{it})^2 &= T^{-1}\sum_t e_{it}^2\Big[(\tilde{\lambda}_i-H^{-1}\lambda_i)'(\tilde{F}_t-H'F_t)+(\tilde{\lambda}_i-H^{-1}\lambda_i)'H'F_t\\ &\quad+\lambda_i'(H^{-1})'(\tilde{F}_t-H'F_t)\Big]^2 = T^{-1}\sum_t\Big[(\tilde{\lambda}_i-H^{-1}\lambda_i)'(\tilde{F}_t-H'F_t)e_{it}\\ &\quad+(\tilde{\lambda}_i-H^{-1}\lambda_i)'H'F_te_{it}+\lambda_i'(H^{-1})'(\tilde{F}_t-H'F_t)e_{it}\Big]^2\\ &\le 3T^{-1}\sum_t\|(\tilde{\lambda}_i-H^{-1}\lambda_i)'(\tilde{F}_t-H'F_t)e_{it}\|^2\\ &\quad+3T^{-1}\sum_t\|(\tilde{\lambda}_i-H^{-1}\lambda_i)'H'F_te_{it}\|^2+3T^{-1}\sum_t\|\lambda_i'(H^{-1})'(\tilde{F}_t-H'F_t)e_{it}\|^2.\end{aligned}$$

The first term is bounded by:

$$\begin{aligned}&T^{-1}\sum_t\|(\tilde{\lambda}_i-H^{-1}\lambda_i)'(\tilde{F}_t-H'F_t)e_{it}\|^2\le\|\tilde{\lambda}_i-H^{-1}\lambda_i\|^2T^{-1}\sum_t\|(\tilde{F}_t-H'F_t)e_{it}\|^2\\ &=O_p(\delta_{NT}^{-2})O_p(\delta_{NT}^{-2}),\end{aligned}$$

by Lemmas 3 and 9.

The second term is:

$$T^{-1}\sum_t\|(\tilde{\lambda}_i-H^{-1}\lambda_i)'H'F_te_{it}\|^2\le\|\tilde{\lambda}_i-H^{-1}\lambda_i\|^2\|H\|^2T^{-1}\sum_t\|F_te_{it}\|^2=O_p(\delta_{NT}^{-2})O_p(1),$$

as the last term is bounded in expectation $T^{-1}\sum_t E\|F_te_{it}\|^2\le T^{-1}\sum_t E\|F_t\|^2E(e_{it}^2)$
$\le T^{-1}\sum_t M=O_p(1)$.

The third term is:

$$T^{-1}\sum_t\|\lambda_i'(H^{-1})'(\tilde{F}_t-H'F_t)e_{it}\|^2\le\|\lambda_i\|^2\|H^{-1}\|^2T^{-1}\sum_t\|(\tilde{F}_t-H'F_t)e_{it}\|^2=O_p(1)O_p(\delta_{NT}^{-2}),$$

by Assumption B and Lemma 3. So $II=O_p(\delta_{NT}^{-2})$.

The term III can be written as:

$$T^{-1}\sum_t e_{it}^3(C_{it}-\tilde{C}_{it})\le\Big(T^{-1}\sum_t e_{it}^6\Big)^{1/2}\Big(T^{-1}\sum_t(C_{it}-\tilde{C}_{it})^2\Big)^{1/2}=O_p(\delta_{NT}^{-1}),$$

since $E(e_{it}^8)\le M$, and the second term was show to be $O_p(\delta_{NT}^{-2})$ in the proof of Lemma 11.

For IV we have:

$$T^{-1}\sum_t e_{it}(C_{it}-\tilde{C}_{it})^3 \leq \left(T^{-1}\sum_t e_{it}^2(C_{it}-\tilde{C}_{it})^2\right)^{1/2}\left(T^{-1}\sum_t(C_{it}-\tilde{C}_{it})^4\right)^{1/2}$$
$$= \left(O_p(\delta_{NT}^{-2})\right)^{1/2}\left(O_p(T^{-1})+O_p(T/N^2)\right)^{1/2} = O_p(\delta_{NT}^{-2}),$$

which follows from I and II. We now have that $I+II+III+IV = O_p(T^{-1})+O_p(T/N^2)+O_p(\delta_{NT}^{-2})+O_p(\delta_{NT}^{-1})+O_p(\delta_{NT}^{-2}) = O_p(T/N^2)+O_p(\delta_{NT}^{-1})$.

□

Lemma 13. *Under Assumptions A-E:*

a. $T^{-1}\sum_t \tilde{F}_t\tilde{F}_t'\tilde{e}_{it}^2 - T^{-1}\sum_t \tilde{F}_t\tilde{F}_t'e_{it}^2 = O_p(\delta_{NT}^{-2})$,

b. $T^{-1}\sum_t \tilde{F}_t\tilde{F}_t'e_{it}^2 - T^{-1}\sum_t H'F_tF_t'He_{it}^2 = O_p(\delta_{NT}^{-2})$.

Proof: As in the proof of 11, we use that $\tilde{e}_{it}^2 = e_{it}^2 + (C_{it}-\tilde{C}_{it})^2 + 2(C_{it}-\tilde{C}_{it})e_{it}$. We can therefore write a as:

$$T^{-1}\sum_t \tilde{F}_t(\tilde{e}_{it}^2-e_{it}^2)\tilde{F}_t' = T^{-1}\sum_t \tilde{F}_t(C_{it}-\tilde{C}_{it})^2\tilde{F}_t' + 2T^{-1}\sum_t \tilde{F}_t(C_{it}-\tilde{C}_{it})e_{it}\tilde{F}_t'.$$

For the first term we have:

$$\begin{aligned}
&T^{-1}\sum_t \tilde{F}_t(C_{it}-\tilde{C}_{it})^2\tilde{F}_t' \leq T^{-1}\sum_t \|(C_{it}-\tilde{C}_{it})\tilde{F}_t'\|^2 = \\
&T^{-1}\sum_t \|(\tilde{\lambda}_i-H^{-1}\lambda_i)'(\tilde{F}_t-H'F_t)\tilde{F}_t' + (\tilde{\lambda}_i-H^{-1}\lambda_i)'H'F_t\tilde{F}_t' + \lambda_i'(H^{-1})'(\tilde{F}_t-H'F_t)\tilde{F}_t'\|^2 \\
&\leq 3T^{-1}\sum_t \|(\tilde{\lambda}_i-H^{-1}\lambda_i)'(\tilde{F}_t-H'F_t)\tilde{F}_t'\|^2 + 3T^{-1}\sum_t \|(\tilde{\lambda}_i-H^{-1}\lambda_i)'H'F_t\tilde{F}_t'\|^2 \\
&+ 3T^{-1}\sum_t \|\lambda_i'(H^{-1})'(\tilde{F}_t-H'F_t)\tilde{F}_t'\|^2 = I+II+III.
\end{aligned}$$

We can bound I by:

$$\|\tilde{\lambda}_i - H^{-1}\lambda_i'\|^2 T^{-1}\sum_t \|(\tilde{F}_t-H'F_t)\tilde{F}_t'\|^2 = O_p(\delta_{NT}^{-2})O_p(\delta_{NT}^{-2}),$$

by Lemmas 5 and 9 .

For II we can write:

$$T^{-1}\sum_t \|(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t\tilde{F}_t'\|^2 \le \|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2\|H\|^2T^{-1}\sum_t \|F_t\tilde{F}_t'\|^2$$
$$\le \|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2\|H\|^2T^{-1}\sum_t \|F_t(\tilde{F}_t - H'F_t)' + F_tF_t'H\|^2$$
$$\le 2\|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2\|H\|^2T^{-1}\sum_t \|F_t(\tilde{F}_t - H'F_t)'\|^2$$
$$+ 2\|\tilde{\lambda}_i - H^{-1}\lambda_i\|^2\|H\|^4T^{-1}\sum_t \|F_t\|^4 = O_p(\delta_{NT}^{-2})O_p(\delta_{NT}^{-2}) + O_p(\delta_{NT}^{-2})O_p(1) = O_p(\delta_{NT}^{-2}),$$

from Lemmas 4 and 9, and Assumption A.

For III we have:

$$T^{-1}\sum_t \|\lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)\tilde{F}_t'\|^2 \le \|\lambda_i\|^2\|H^{-1}\|^2T^{-1}\sum_t \|(\tilde{F}_t - H'F_t)\tilde{F}_t'\|^2 = O_p(\delta_{NT}^{-2}),$$

by Lemma 5, so we have $T^{-1}\sum_t \tilde{F}_t(C_{it} - \tilde{C}_{it})^2\tilde{F}_t' = O_p(\delta_{NT}^{-4}) + O_p(\delta_{NT}^{-2}) = O_p(\delta_{NT}^{-2})$.

For the second term we write:

$$T^{-1}\sum_t \tilde{F}_t(C_{it} - \tilde{C}_{it})e_{it}\tilde{F}_t' = T^{-1}\sum_t (\tilde{F}_t - H'F_t)(C_{it} - \tilde{C}_{it})e_{it}\tilde{F}_t'$$
$$+ T^{-1}\sum_t H'F_t(C_{it} - \tilde{C}_{it})e_{it}(\tilde{F}_t - H'F_t)' + T^{-1}\sum_t H'F_t(C_{it} - \tilde{C}_{it})e_{it}F_t'H.$$

We can bound $T^{-1}\sum_t (\tilde{F}_t - H'F_t)(C_{it} - \tilde{C}_{it})e_{it}\tilde{F}_t'$ by:

$$\left(T^{-1}\sum_t (C_{it} - \tilde{C}_{it})^2e_{it}^2\right)^{1/2}\left(T^{-1}\sum_t \|(\tilde{F}_t - H'F_t)\tilde{F}_t'\|^2\right)^{1/2} = O_p(\delta_{NT}^{-1})O_p(\delta_{NT}^{-1}),$$

which follows as the term in the first parenthesis was shown to be $O_p(\delta_{NT}^{-2})$ in the proof of Lemma 12, and the term in the second parenthesis is also $O_p(\delta_{NT}^{-2})$ by Lemma 5. By similar calculations we also get $T^{-1}\sum_t H'F_t(C_{it} - \tilde{C}_{it})e_{it}(\tilde{F}_t - H'F_t)' = O_p(\delta_{NT}^{-2})$.

For $T^{-1}\sum_t H'F_t(C_{it} - \tilde{C}_{it})e_{it}F_t'H$ we can write (ignore H as it is $O_p(1)$):

$$T^{-1}\sum_t F_t(C_{it} - \tilde{C}_{it})e_{it}F_t' = T^{-1}\sum_t F_t\Big[(\tilde{\lambda}_i + H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t) + (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t$$
$$+ \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)\Big]e_{it}F_t' = T^{-1}\sum_t F_t(\tilde{\lambda}_i + H^{-1}\lambda_i)'(\tilde{F}_t - H'F_t)e_{it}F_t'$$
$$+ T^{-1}\sum_t F_t(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_te_{it}F_t' + T^{-1}\sum_t F_t\lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)e_{it}F_t'.$$

The first term is bounded by:

$$\|\tilde{\lambda}_i + H^{-1}\lambda_i\| \left(T^{-1} \sum_t \|(\tilde{F}_t - H'F_t)e_{it}\|^2 \right)^{1/2} \left(T^{-1} \sum_t \|F_t\|^4 \right)^{1/2}$$
$$= O_p(T^{-1/2})(\delta_{NT}^{-1})O_p(1) = O_p(\delta_{NT}^{-2}),$$

by Lemmas 3 and 9.

For the second term we apply the *vec* operator:

$$T^{-1}\sum_t \left(vec[F_t(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'F_t e_{it} F_t'] \right)' = T^{-1}\sum_t \left(e_{it}(F_tF_t' \otimes F_t) vec[(\tilde{\lambda}_i - H^{-1}\lambda_i)'H'] \right)'$$
$$= T^{-1}\sum_t (\tilde{\lambda}_i - H^{-1}\lambda_i)'H'(F_tF_t' \otimes F_t')e_{it} = T^{-1/2}(\tilde{\lambda}_i - H^{-1}\lambda_i)'H' \left(T^{-1/2}\sum_t (F_tF_t' \otimes F_t')e_{it} \right)$$
$$= O_p(T^{-1/2})O_p(T^{-1/2}) = O_p(T^{-1}),$$

by Assumption E.4.

For the third term we can write:

$$T^{-1}\sum_t \lambda_i'(H^{-1})'(\tilde{F}_t - H'F_t)e_{it}F_tF_t' = \lambda_i'(H^{-1})'T^{-1}\sum_t (\tilde{F}_t - H'F_t)e_{it}F_tF_t' = O_p(1)O_p(\delta_{NT}^{-2}),$$

by Assumption B and Lemma 6. We therefore have that a is $O_p(\delta_{NT}^{-4}) + O_p(T^{-1}) + O_p(\delta_{NT}^{-2}) = O_p(\delta_{NT}^{-2})$.

For b we can write:

$$T^{-1}\sum_t \tilde{F}_t\tilde{F}_t' e_{it}^2 - T^{-1}\sum_t H'F_tF_t'He_{it}^2 = T^{-1}\sum_t (\tilde{F}_t\tilde{F}_t' - H'F_tF_t'H)e_{it}^2$$
$$= T^{-1}\sum_t \left((\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t) \right) e_{it}^2$$
$$= T^{-1}\sum_t (\tilde{F}_t - H'F_t)\tilde{F}_t' e_{it}^2 + T^{-1}\sum_t H'F_t(\tilde{F}_t - H'F_t)e_{it}^2$$
$$= T^{-1}\sum_t (\tilde{F}_t - H'F_t)(\tilde{F}_t - H'F_t)'e_{it}^2 + T^{-1}\sum_t (\tilde{F}_t - H'F_t)F_t'He_{it}^2 + T^{-1}\sum_t H'F_t(\tilde{F}_t - F_t'H)'e_{it}^2.$$

From Lemma 7, the last two terms are $O_p(\delta_{NT}^{-2})$. The first term is bounded by:

$$\|T^{-1}\sum_t (\tilde{F}_t - H'F_t)e_{it}e_{it}(\tilde{F}_t - H'F_t)\| \leq T^{-1}\sum_t \|(\tilde{F}_t - H'F_t)e_{it}\|^2 = O_p(\delta_{NT}^{-2}),$$

from Lemma 3. Thus b is $O_p(\delta_{NT}^{-2})$.

□

Proof of Theorem 1: The proof consists of two steps. First we show that $\tilde{D}_i = D_i + O_p(\delta_{NT}^{-2})$. This implies that $\sqrt{T}\tilde{D}_i$ has the same limiting distribution as $\sqrt{T}D_i$ if $\sqrt{T}/N \to 0$. The second step shows that $\tilde{B}_i$ is a consistent estimator for B_i, and this implies that $T\tilde{D}_i\tilde{B}_i^{-1}\tilde{D}_i'$ has a χ^2 distribution with r degrees of freedom from Assumption G.

First consider $\tilde{D}_i = T^{-1}\sum_{t=1}^T(\tilde{e}_{it}^2 - \tilde{\sigma}_i^2)g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)$. We can ignore $\tilde{\sigma}_i^2$ as $T^{-1}\sum_{t=1}^T g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T) = 0$. We will show that the $r\times r$ matrix $T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)$ matrix converges to $T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2(H'F_tF_t'H - H'F'FH/T)$, and this will imply that $\tilde{D}_i$ converges to D_i as the $g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)$ is the vector of diagonal elements of $\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T$. We can write:

$$T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T) = T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2\tilde{F}_t\tilde{F}_t' - T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2\tilde{F}'\tilde{F}/T.$$

From Lemma 13 we have:

$$T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2\tilde{F}_t\tilde{F}_t' = T^{-1}\sum_{t=1}^T e_{it}^2 H'F_tF_t'H + O_p(\delta_{NT}^{-2}).$$

From Lemma 11 we have that $T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2 = T^{-1}\sum_{t=1}^T e_{it}^2 + O_p(\delta_{NT}^{-2})$, and for $\tilde{F}'\tilde{F}/T$ we have:

$$\begin{aligned}\tilde{F}'\tilde{F}/T &= \tilde{F}'(\tilde{F} - FH)/T + (\tilde{F} - FH)'FH/T + H'F'FH/T \\ &= H'F'FH/T + O_p(\delta_{NT}^{-2}),\end{aligned}$$

from Lemmas B.2 and B.3 in Bai (2003), so $T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2\tilde{F}'\tilde{F}/T = T^{-1}\sum_{t=1}^T e_{it}^2 H'F'FH/T + O_p(\delta_{NT}^{-2})$, and we therefore have that $\tilde{D}_i = D_i + O_p(\delta_{NT}^{-2})$.

Next we need to show that $\tilde{B}_i$ is a consistent estimator for B_i. Under Assumption G.1 and as F_t and e_{it} are assumed to be independent, we have that:

$$\begin{aligned}B_i &= \text{plim}_{T\to\infty}T^{-1}\sum_{t=1}^T E\Big[(e_{it}^2 - \sigma_i^2)^2 g[H'(F_tF_t' - F'F/T)H]g[H'(F_tF_t' - F'F/T)H]'\Big] \\ &= \text{plim}_{T\to\infty}T^{-1}\sum_{t=1}^T E(e_{it}^2 - \sigma_i^2)^2 E\Big[g[H'(F_tF_t' - F'F/T)H]g[H'(F_tF_t' - F'F/T)H]'\Big] \\ &= (\mu_{4,i} - \sigma_i^4)\text{plim}_{T\to\infty}T^{-1}\sum_{t=1}^T E\Big[g[H'(F_tF_t' - F'F/T)H]g[H'(F_tF_t' - F'F/T)H]'\Big].\end{aligned}$$

From Lemma 11 we have that $\tilde{\sigma}_i^2 = T^{-1}\sum_{t=1}^T \tilde{e}_{it}^2 = T^{-1}\sum_{t=1}^T e_{it}^2 + O_p(\delta_{NT}^{-2})$, so σ_i^4 can be consistently estimated by $\tilde{\sigma}_i^4$. Lemma 12 states that $T^{-1}\sum_t \tilde{e}_{it}^4 = T^{-1}\sum_t e_{it}^4 + O_p(T/N^2) + O_p(\delta_{NT}^{-1})$, so

$T^{-1}\sum_t \tilde{e}_{it}^4$ is consistent for $\mu_{4,i}$. As the final step we show that

$$
\begin{aligned}
&T^{-1}\sum_t g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)' \\
&= T^{-1}\sum_{t=1}^{T} g[H'(F_tF_t' - F'F/T)H]g[H'(F_tF_t' - F'F/T)H]' + O_p(\delta_{NT}^{-1})
\end{aligned}
$$

which will imply that $\tilde{B}_i$ is consistent for B_i.

We can write $g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)$ as:

$$
\begin{aligned}
g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T) =& g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)' + H'F_tF_t'H - \tilde{F}'\tilde{F}/T] \\
=& g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)'] + g[H'F_tF_t'H - \tilde{F}'\tilde{F}/T] \\
=& g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)'] \\
&+ g[H'F_tF_t'H - H'F'FH/T] \\
&+ g[H'F'FH/T - \tilde{F}'\tilde{F}/T].
\end{aligned}
$$

We therefore get:

$$
\begin{aligned}
&T^{-1}\sum_t g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)g(\tilde{F}_t\tilde{F}_t' - \tilde{F}'\tilde{F}/T)' \\
&= T^{-1}\sum_t g[H'F_tF_t'H - H'F'FH/T]g[H'F_tF_t'H - H'F'FH/T]' \\
&+ T^{-1}\sum_t g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']' \\
&+ T^{-1}\sum_t g[H'F'FH/T - \tilde{F}'\tilde{F}/T]g[H'F'FH/T - \tilde{F}'\tilde{F}/T]' \\
&+ 2T^{-1}\sum_t g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']g[H'F_tF_t'H - H'F'FH/T]' \\
&+ 2T^{-1}\sum_t g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']g[H'F'FH/T - \tilde{F}'\tilde{F}/T] \\
&+ 2T^{-1}\sum_t g[H'F_tF_t'H - H'F'FH/T]g[H'F'FH/T - \tilde{F}'\tilde{F}/T] \\
&= T^{-1}\sum_t g[H'F_tF_t'H - H'F'FH/T]g[H'F_tF_t'H - H'F'FH/T]' + I + II + III + IV + IV.
\end{aligned}
$$

We start with I. Recalling that $g(A)$ is the vector of diagonal elements of A, we have for any square matrix A:

$$\|g(A)\|^2 = g(A)'g(A) \le vec(A)'vec(A) = tr(A'A) = \|A\|^2.$$

We can therefore write:

$$\begin{aligned}
&T^{-1}\sum_t g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']' \\
&\leq T^{-1}\sum_t \left\| g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)'] \right\|^2 \\
&\leq 2T^{-1}\sum_t \|(\tilde{F}_t - H'F_t)\tilde{F}_t'\|^2 + 2T^{-1}\sum_t \|H'F_t(\tilde{F}_t - H'F_t)'\|^2 = O_p(\delta_{NT}^{-2}),
\end{aligned}$$

by Lemmas 4 and 5.

For II we have:

$$\begin{aligned}
&T^{-1}\sum_t g[H'F'FH/T - \tilde{F}'\tilde{F}/T]g[H'F'FH - \tilde{F}'\tilde{F}/T]' \\
&\leq \|H'F'FH/T - \tilde{F}'\tilde{F}/T\|^2 = \|T^{-1}\tilde{F}'(\tilde{F} - FH) + T^{-1}(\tilde{F} - FH)'FH\|^2 = O_p(\delta_{NT}^{-2}),
\end{aligned}$$

by Lemmas 8.b and 8.c.

For III we have:

$$\begin{aligned}
&T^{-1}\sum_t g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']g[H'F_tF_t'H - H'F'FH/T]' \\
&\leq \left(T^{-1}\sum_t \|(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)'\|^2\right)^{1/2}\left(T^{-1}\sum_t \|H'F_tF_t'H - H'F'FH/T\|^2\right)^{1/2} \\
&\leq O_p(\delta_{NT}^{-1})\left(2\|H'F'FH/T\|^2 T^{-1}\sum_t \|H'F_tF_t'H\|^2\right)^{1/2} = O_p(\delta_{NT}^{-1})O_p(1),
\end{aligned}$$

from I above and Assumption A.

For IV we have:

$$\begin{aligned}
&T^{-1}\sum_t g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)']g[H'F'FH/T - \tilde{F}'\tilde{F}/T] \\
&= g[H'F'FH/T - \tilde{F}'\tilde{F}/T]T^{-1}\sum_t g[(\tilde{F}_t - H'F_t)\tilde{F}_t' + H'F_t(\tilde{F}_t - H'F_t)'].
\end{aligned}$$

We can write $H'F'FH/T - \tilde{F}'\tilde{F}/T = T^{-1}\tilde{F}'(\tilde{F} - FH) + T^{-1}(\tilde{F} - FH)'FH$. From Lemmas B.2 and B.3 in Bai (2003), these terms are $O_p(\delta_{NT}^{-2})$, so $IV = O_p(\delta_{NT}^{-2})O_p(\delta_{NT}^{-2})$.

Finally, for V we have:

$$T^{-1}\sum_t g[H'F_tF_t'H - H'F'FH/T]g[H'F'FH/T - \tilde{F}'\tilde{F}/T]$$
$$= g[H'F'FH/T - \tilde{F}'\tilde{F}/T]T^{-1}\sum_t g[H'F_tF_t'H - H'F'FH/T] = O_p(\delta_{NT}^{-2}),$$

again from Lemmas B.2 and B.3 in Bai (2003). Thus $I+II+III+IV+V = O_p(\delta_{NT}^{-2}) + O_p(\delta_{NT}^{-4}) + O_p(\delta_{NT}^{-1}) = O_p(\delta_{NT}^{-1})$, and the Theorem 1 follows.

□

References

Alessi, L., M. Barigozzi, and M. Capasso (2010): "Improved penalization for determining the number of factors in approximate factor models," *Statistics & Probability Letters*, 80, 1806–1813.

Bai, J. (2003): "Inferential theory for factor models of large dimensions," *Econometrica*, 71, 135–171.

Bai, J. and S. Ng (2002): "Determining the number of factors in approximate factor models," *Econometrica*, 70, 191–221.

——— (2006): "Evaluating latent and observed factors in macroeconomics and finance," *Journal of Econometrics*, 131, 507–537.

Bernanke, B., J. Boivin, and P. Eliasz (2005): "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach," *Quarterly Journal of Economics*, February, 387–422.

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

Breitung, J. and J. Tenhofen (2011): "GLS estimation of dynamic factor models," *Journal of the American Statistical Association*, 106, 1150–1166.

Del Negro, M. and C. Otrok (2008): "Dynamic factor models with time-varying parameters: measuring changes in international business cycles," Staff report 326, Federal Reserve Bank of New York, New York.

Eickmeier, S., W. Lemke, and M. Marcellino (2015): "Classical time varying factor-augmented vector auto-regressive models – estimation, forecasting and structural analysis," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178, 493–533.

Fama, E. F. and K. R. French (1993): "Common risk factors in the returns on stocks and bonds," *Journal of financial economics*, 33, 3–56.

Giannone, D., L. Reichlin, and L. Sala (2006): "VARs, common factors and the empirical validation of equilibrium business cycle models," *Journal of Econometrics*, 132, 257–279.

McCracken, M. and S. Ng (2015): "FRED-MD: A Monthly Database for Macroeconomic Research," *forthcoming, Journal of Business and Economic Statistics.*

Stock, J. and M. Watson (2002): "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20, 147–162.

Research Papers 2016

- 2017-05: Matthew T. Holt and Timo Teräsvirta: Global Hemispheric Temperatures and Co–Shifting: A Vector Shifting–Mean Autoregressive Analysis
- 2017-06: Tobias Basse, Robinson Kruse and Christoph Wegener: The Walking Debt Crisis
- 2017-07: Oskar Knapik: Modeling and forecasting electricity price jumps in the Nord Pool power market
- 2017-08: Malene Kallestrup-Lamb and Carsten P.T. Rosenskjold: Insight into the Female Longevity Puzzle: Using Register Data to Analyse Mortality and Cause of Death Behaviour Across Socio-economic Groups
- 2017-09: Thomas Quistgaard Pedersen and Erik Christian Montes Schütte: Testing for Explosive Bubbles in the Presence of Autocorrelated Innovations
- 2017-10: Jeroen V.K. Rombouts, Lars Stentoft and Francesco Violante: Dynamics of Variance Risk Premia, Investors' Sentiment and Return Predictability
- 2017-11: Søren Johansen and Morten Nyboe Tabor: Cointegration between trends and their estimators in state space models and CVAR models
- 2017-12: Lukasz Gatarek and Søren Johansen: The role of cointegration for optimal hedging with heteroscedastic error term
- 2017-13: Niels S. Grønborg, Asger Lunde, Allan Timmermann and Russ Wermers: Picking Funds with Confidence
- 2017-14: Martin M. Andreasen and Anders Kronborg: The Extended Perturbation Method: New Insights on the New Keynesian Model
- 2017-15: Andrea Barletta, Paolo Santucci de Magistris and Francesco Violante: A Non-Structural Investigation of VIX Risk Neutral Density
- 2017-16: Davide Delle Monache, Stefano Grassi and Paolo Santucci de Magistris: Does the ARFIMA really shift?
- 2017-17: Massimo Franchi and Søren Johansen: Improved inference on cointegrating vectors in the presence of a near unit root using adjusted quantiles
- 2017-18: Matias D. Cattaneo, Michael Jansson and Kenichi Nagasawa: Bootstrap-Based Inference for Cube Root Consistent Estimators
- 2017-19: Daniel Borup and Martin Thyrsgaard: Statistical tests for equal predictive ability across multiple forecasting methods
- 2017-20: Tommaso Proietti and Alessandro Giovannelli: A Durbin-Levinson Regularized Estimator of High Dimensional Autocovariance Matrices
- 2017-21: Jeroen V.K. Rombouts, Lars Stentoft and Francesco Violante: Variance swap payoffs, risk premia and extreme market conditions
- 2017-22: Jakob Guldbæk Mikkelsen: Testing for time-varying loadings in dynamic factor models