

# Bootstrapping factor-augmented regression models

Sílvia Gonçalves and Benoit Perron\*

*Département de sciences économiques, CIREQ and CIRANO, Université de Montréal*

November 9, 2011

## Abstract

The main contribution of this paper is to propose and theoretically justify bootstrap methods for regressions where some of the regressors are factors estimated from a large panel of data. We derive our results under the assumption that  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$  ( $N$  and  $T$  are the cross-sectional and the time series dimensions, respectively), thus allowing for the possibility that factors estimation error enters the limiting distribution of the OLS estimator. We consider general residual-based bootstrap methods and provide a set of high level conditions on the bootstrap residuals and on the idiosyncratic errors such that the bootstrap distribution of the OLS estimator is consistent. We subsequently verify these conditions for a simple wild bootstrap residual-based procedure.

Our main results can be summarized as follows. When  $c = 0$ , as in Bai and Ng (2006), the crucial condition for bootstrap validity is the ability of the bootstrap regression scores to mimic the serial dependence of the original regression scores. Mimicking the cross sectional and/or serial dependence of the idiosyncratic errors in the panel factor model is asymptotically irrelevant in this case since the limiting distribution of the original OLS estimator does not depend on these dependencies. Instead, when  $c > 0$ , a two-step residual-based bootstrap is required to capture the factors estimation uncertainty, which shows up as an asymptotic bias term (as we show here and as was recently discussed by Ludvigson and Ng (2009b)). Because the bias depends on the cross sectional dependence of the idiosyncratic error term, bootstrap validity depends crucially on the ability of the bootstrap panel factor model to capture this cross sectional dependence.

Keywords: factor model, bootstrap, asymptotic bias.

## 1 Introduction

Factor-augmented regressions where some of the regressors, called factors, are estimated from a large set of data are increasingly popular in empirical work. Inference in these models is complicated by the fact that the regressors are estimated and thus measured with error. Recently, Bai and Ng (2006) derived the asymptotic distribution of the OLS estimator in this case under a set of standard regularity conditions. In particular, they show that the asymptotic distribution of the OLS estimator is unaffected

---

\*We would like to thank participants at the Conference in honor of Halbert White (San Diego, May 2011), the Conference in honor of Ron Gallant (Toulouse, May 2011), the Conference in honor of Hashem Pesaran (Cambridge, July 2011), the Fifth CIREQ time series conference (Montreal, May 2011), the 17th International Panel Data conference (Montreal, July 2011), and the NBER/NSF Time Series conference at MSU (East Lansing, September 2011) as well as seminar participants at HEC Montreal, Syracuse University, Georgetown University, Tilburg University, Rochester University, Cornell University, University of British Columbia and University of Victoria. We also thank Valentina Corradi, Nikolay Gospodinov, Guido Kursteiner and Serena Ng for very useful comments and discussions. Gonçalves acknowledges financial support from the NSERC and MITACS whereas Perron acknowledges financial support from the SSHRCC and MITACS.

by the estimation of the factors when  $\sqrt{T}/N \rightarrow 0$ , where  $N$  and  $T$  are the cross-sectional and the time series dimensions, respectively. While their simulation study does not consider inference on the coefficients themselves (they look at the conditional mean), they report noticeable size distortions in some situations.

The main contribution of this paper is to propose and theoretically justify bootstrap methods for inference in the context of the factor-augmented regression model. Shintani and Guo (2011) prove the validity of the bootstrap to carry out inference on the persistence of factors, but not in the context of factor-augmented regressions. Recent empirical applications of the bootstrap in this context include Ludvigson and Ng (2007, 2009a,b) and Gospodinov and Ng (2011), where the bootstrap has been used in the context of predictability tests based on factor-augmented regressions without theoretical justification. Our main contribution is to establish the first order asymptotic validity of the bootstrap for factor-augmented regression models under assumptions similar to those of Bai and Ng (2006) but without the condition that  $\sqrt{T}/N \rightarrow 0$ . Specifically, we assume that  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$ , thus allowing for the possibility that factor estimation error affects the limiting distribution of the OLS estimator. As it turns out, when  $c > 0$ , an asymptotic bias term appears in the distribution, reflecting the contribution of factors estimation uncertainty. This bias problem was recently discussed by Ludvigson and Ng (2009b), who proposed an analytical bias correction procedure. Instead, here we focus on the bootstrap and provide a set of conditions under which it can reproduce the whole limiting distribution of the OLS estimator, including the bias term.

The bootstrap method we propose is made up of two main steps. In a first step, we obtain a bootstrap panel data set from which we estimate the bootstrap factors by principal components. The bootstrap panel observations are generated by adding the estimated common components from the original panel and bootstrap idiosyncratic residuals. In a second step, we generate a bootstrap version of the response variable by again relying on a residual-based bootstrap where the bootstrap observations of the dependent variable are obtained by summing the estimated regression mean and a bootstrap regression residual. To mimic the fact that in the original regression model the true factors are latent and need to be estimated, we regress the bootstrap response variable on the estimated bootstrap factors. This produces a bootstrap OLS estimator whose bootstrap distribution can be used to replicate the distribution of the OLS estimator.

We first provide a set of high level conditions on the bootstrap residuals and idiosyncratic errors that allow us to characterize the limiting distribution of the bootstrap OLS estimator under the assumption that  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$ . These high level conditions essentially require that the bootstrap idiosyncratic errors be weakly dependent across individuals and over time and that the bootstrap regression scores satisfy a central limit theorem. We then verify these high level conditions for a residual-based wild bootstrap scheme, where the wild bootstrap is used to generate the bootstrap idiosyncratic error term in the first step, and also in the second step when generating the regression residuals. The two steps are performed independently of each other.

A crucial result in proving the first order asymptotic validity of the bootstrap in this context is the consistency of the bootstrap principal component estimator. Given our residual-based bootstrap, the “latent” factors underlying the bootstrap data generating process (DGP) are given by the estimated factors. Nevertheless, these are not identified by the bootstrap principal component estimator due to the well-known identification problem of factor models. By relying on results of Bai and Ng (2011) (see also Stock and Watson (2002)), we show that the bootstrap estimated factors identify the estimated factors up to a change of sign. Contrary to the rotation indeterminacy problem that affects the principal component estimator, this sign indetermination is easily resolved in the bootstrap world, where the bootstrap rotation matrix depends on bootstrap population values that are functions of the original data. As a consequence, to bootstrap the distribution of OLS estimator, our proposal is to rotate the bootstrap OLS estimator using the feasible bootstrap rotation matrix. This amounts to sign-adjusting the bootstrap OLS regression estimates asymptotically.

Our results can be summarized as follows. When  $c = 0$ , as in Bai and Ng (2006), the crucial condition for bootstrap validity is the ability of the bootstrap regression scores to mimic the serial dependence and heteroskedasticity of the original regression scores. For instance, a wild bootstrap in the second-step of our residual-based bootstrap is appropriate if we assume the regression errors to be a possibly heteroskedastic martingale difference sequence (as in Bai and Ng (2006)). Under a more general dependence assumption, the wild bootstrap is not appropriate and we should instead consider a block bootstrap. We do not pursue this possibility here but note that our bootstrap high level conditions would be useful in establishing the validity of the block bootstrap in this context as well. Mimicking the cross sectional and/or serial dependence of the idiosyncratic errors in the panel factor model is asymptotically irrelevant when  $c = 0$  since the limiting distribution of the original OLS estimator does not depend on these dependencies under this condition. Thus, a wild bootstrap in the first step of the residual-based bootstrap is asymptotically valid under the general setup of Bai and Ng (2006) that allows for weak time series and cross sectional dependence in the idiosyncratic error term. In fact, a simple one-step residual-based bootstrap method that does not take into account the factor estimation uncertainty in the bootstrap samples (i.e. a bootstrap method based only on the second step of our proposed method) is asymptotically valid when  $c = 0$ .<sup>1</sup> Instead, when  $c > 0$ , a two-step residual-based bootstrap is required to capture the asymptotic bias term that reflects the factors estimation error. Because the bias depends on the cross sectional dependence of the idiosyncratic error term, bootstrap validity depends crucially on the ability of the bootstrap panel factor model to capture this cross sectional dependence. Since the wild bootstrap generates bootstrap idiosyncratic errors that are heteroskedastic but independent along the time and cross sectional dimensions, this method is consistent only under cross sectional independence of the idiosyncratic errors.

The rest of the paper is organized as follows. In Section 2, we first describe the setup and the

---

<sup>1</sup>We do not consider this here because factor estimation uncertainty has an impact in finite samples. For instance, Yamamoto (2011) compares bootstrap methods with and without factor estimation for the factor-augmented vector autoregression (FAVAR) model and concludes that the latter is worse than the first in terms of finite sample accuracy.

assumptions, and then derive the asymptotic theory of the OLS estimator when  $\sqrt{T}/N \rightarrow c$ . In Section 3, we introduce the residual-based bootstrap method and characterize a set of conditions under which the bootstrap distribution consistency follows. Section 4 proposes a wild bootstrap implementation of the residual-based bootstrap and proves its consistency. Section 5 discusses the Monte Carlo results and Section 6 concludes. Three mathematical appendices are included. Appendix A contains the proofs of the results in Section 2, Appendix B the proofs of the results in Section 3, and Appendix C the proofs of the results in Section 4.

A word on notation. As usual in the bootstrap literature, we use  $P^*$  to denote the bootstrap probability measure, conditional on a given sample. For any bootstrap statistic  $T_{NT}^*$ , we write  $T_{NT}^* = o_{P^*}(1)$ , in probability, or  $T_{NT}^* \rightarrow^{P^*} 0$ , in probability, when for any  $\delta > 0$ ,  $P^*(|T_{NT}^*| > \delta) = o_{P^*}(1)$ . We write  $T_{NT}^* = O_{P^*}(1)$ , in probability, when for all  $\delta > 0$  there exists  $M_\delta < \infty$  such that  $\lim_{N,T \rightarrow \infty} P[P^*(|T_{NT}^*| > M_\delta) > \delta] = 0$ . Finally, we write  $T_{NT}^* \rightarrow^{d^*} D$ , in probability, if conditional on a sample with probability that converges to one,  $T_{NT}^*$  weakly converges to the distribution  $D$  under  $P^*$ , i.e.  $E^*(f(T_{NT}^*)) \rightarrow^P E(f(D))$  for all bounded and uniformly continuous functions  $f$ .

## 2 Asymptotic theory when $\sqrt{T}/N \rightarrow c$ , where $0 \leq c < \infty$

We consider the following regression model

$$y_{t+h} = \alpha' F_t + \beta' W_t + \varepsilon_{t+h}, \quad t = 1, \dots, T-h, \quad (1)$$

where  $h \geq 0$ . The  $q$  observed regressors are contained in  $W_t$ . The  $r$  unobserved regressors  $F_t$  are the common factors in the following panel factor model,

$$X_{it} = \lambda_i' F_t + e_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad (2)$$

where the  $r \times 1$  vector  $\lambda_i$  contains the factor loadings and  $e_{it}$  is an idiosyncratic error term. In matrix form, we can write (2) as

$$X = F\Lambda' + e,$$

where  $X$  is a  $T \times N$  matrix of stationary data,  $F = (F_1, \dots, F_T)'$  is  $T \times r$ ,  $r$  is the number of common factors,  $\Lambda = (\lambda_1, \dots, \lambda_N)'$  is  $N \times r$ , and  $e$  is  $T \times N$ .

The factor-augmented regression model described in (1) and (2) has recently attracted a lot of attention in econometrics. One of the first papers to discuss this model in the forecasting context was Stock and Watson (2002). Recent empirical applications include Ludvigson and Ng (2007) who consider predictive regressions of excess stock returns and augment the usual set of predictors by including estimated factors from a large panel of macro and financial variables, Ludvigson and Ng (2009a,b) who consider this approach in the context of predictive regressions of bond excess returns, Gospodinov and Ng (2011) who study predictive regressions for inflation using principal components from a panel of commodity convenience yields, and Eichengreen, Mody, Nedeljkovic, and Sarno (2009)

who use common factors extracted from credit default swap (CDS) spreads during the recent financial crisis to look at spillovers across banks.

Estimation proceeds in two steps. Given  $X$ , we estimate  $F$  and  $\Lambda$  with the method of principal components. In particular,  $F$  is estimated with the  $T \times r$  matrix  $\tilde{F} = (\tilde{F}_1 \dots \tilde{F}_T)'$  composed of  $\sqrt{T}$  times the eigenvectors corresponding to the  $r$  largest eigenvalues of  $XX'/TN$  (arranged in decreasing order), where the normalization  $\frac{\tilde{F}'\tilde{F}}{T} = I_r$  is used. The matrix containing the estimated loadings is then  $\tilde{\Lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_N)' = X'\tilde{F}(\tilde{F}'\tilde{F})^{-1} = X'\tilde{F}/T$ .

In the second step, we run an OLS regression of  $y_{t+h}$  on  $\hat{z}_t = (\tilde{F}_t' \ W_t')'$ , i.e. we compute

$$\hat{\delta} \equiv \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \left( \sum_{t=1}^{T-h} \hat{z}_t \hat{z}_t' \right)^{-1} \sum_{t=1}^{T-h} \hat{z}_t y_{t+h}, \quad (3)$$

where  $\hat{\delta}$  is  $p \times 1$  with  $p = r + q$ .

As is well known in this literature, the principal components  $\tilde{F}_t$  can only consistently estimate a transformation of the true factors  $F_t$ , given by  $HF_t$ , where  $H$  is a rotation matrix defined as

$$H = \tilde{V}^{-1} \frac{\tilde{F}'F}{T} \frac{\Lambda'\Lambda}{N}, \quad (4)$$

where  $\tilde{V}$  is the  $r \times r$  diagonal matrix containing on the main diagonal the  $r$  largest eigenvalues of  $XX'/NT$ , in decreasing order.

One important implication is that  $\hat{\delta}$  consistently estimates  $\delta \equiv (\alpha'H^{-1} \ \beta')'$ , and not  $(\alpha' \ \beta')'$ . In particular, given (1), adding and subtracting appropriately yields

$$y_{t+h} = \underbrace{(\alpha'H^{-1} \ \beta')}'_{=\delta'} \underbrace{\begin{pmatrix} \tilde{F}_t \\ W_t \end{pmatrix}}_{=\hat{z}_t} + \alpha'H^{-1} (HF_t - \tilde{F}_t) + \varepsilon_{t+h},$$

or, equivalently,

$$y_{t+h} = \hat{z}_t' \delta + \alpha'H^{-1} (HF_t - \tilde{F}_t) + \varepsilon_{t+h}, \quad (5)$$

where the second term represents the contribution from estimating the factors.

Recently, Bai and Ng (2006) derived the asymptotic distribution of  $\sqrt{T}(\hat{\delta} - \delta)$  under a set of regularity conditions and the assumption that  $\sqrt{T}/N \rightarrow 0$ . Our goal in this section is to derive the limiting distribution of  $\hat{\delta}$  under the assumption that  $\sqrt{T}/N \rightarrow c$ , where  $c$  is not necessarily zero. We use the following assumptions, which are similar to Bai's (2003) assumptions and slightly weaker than the Bai and Ng (2006) assumptions. We let  $z_t = (F_t' \ W_t')'$ , where  $z_t$  is  $p \times 1$ , with  $p = r + q$ .

### Assumption 1 - Factors and factor loadings

(a)  $E\|F_t\|^4 \leq M$  and  $\frac{1}{T} \sum_{t=1}^T F_t F_t' \rightarrow^P \Sigma_F > 0$ , where  $\Sigma_F$  is a non-random  $r \times r$  matrix.

(b) The loadings  $\lambda_i$  are either deterministic such that  $\|\lambda_i\| \leq M$ , or stochastic such that  $E\|\lambda_i\|^4 \leq M$ .

In either case,  $\Lambda'\Lambda/N \rightarrow^P \Sigma_\Lambda > 0$ , where  $\Sigma_\Lambda$  is a non-random matrix.

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

**Assumption 2 - Time and cross section dependence and heteroskedasticity**

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

(b)  $E(e_{it}e_{js}) = \sigma_{ij,ts}$ ,  $|\sigma_{ij,ts}| \leq \bar{\sigma}_{ij}$  for all  $(t, s)$  and  $|\sigma_{ij,ts}| \leq \tau_{ts}$  for all  $(i, j)$  such that  $\frac{1}{N} \sum_{i,j=1}^N \bar{\sigma}_{ij} \leq M$ ,  $\frac{1}{T} \sum_{t,s=1}^T \tau_{ts} \leq M$ , and  $\frac{1}{NT} \sum_{t,s,i,j} |\sigma_{ij,ts}| \leq M$ .

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

**Assumption 3 - Moments and weak dependence among  $\{z_t\}$ ,  $\{\lambda_i\}$  and  $\{e_{it}\}$ .**

(a)  $E \left( \frac{1}{N} \sum_{i=1}^N \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^T F_t e_{it} \right\|^2 \right) \leq M$ , where  $E(F_t e_{it}) = 0$  for all  $(i, t)$ .

(b) For each  $t$ ,  $E \left\| \frac{1}{\sqrt{TN}} \sum_{s=1}^T \sum_{i=1}^N z_s (e_{it}e_{is} - E(e_{it}e_{is})) \right\|^2 \leq M$ , where  $z_s = (F'_s \ W'_s)'$ .

(c)  $E \left\| \frac{1}{\sqrt{NT}} \sum_{t=1}^T z_t e'_t \Lambda \right\|^2 \leq M$ , where  $E(z_t \lambda'_i e_{it}) = 0$  for all  $(i, t)$ .

(d)  $E \left( \frac{1}{T} \sum_{t=1}^T \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^N \lambda_i e_{it} \right\|^2 \right) \leq M$ , where  $E(\lambda_i e_{it}) = 0$  for all  $(i, t)$ .

(e) As  $N, T \rightarrow \infty$ ,  $\frac{1}{TN} \sum_{t=1}^T \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda'_j e_{it} e_{jt} - \Gamma \rightarrow^P 0$ , where  $\Gamma \equiv \lim_{N,T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \Gamma_t > 0$ , and  $\Gamma_t \equiv \text{Var} \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N \lambda_i e_{it} \right)$ .

**Assumption 4 - weak dependence between idiosyncratic errors and regression errors**

(a) For each  $t$  and  $h \geq 0$ ,  $E \left| \frac{1}{\sqrt{TN}} \sum_{s=1}^{T-h} \sum_{i=1}^N \varepsilon_{s+h} (e_{it}e_{is} - E(e_{it}e_{is})) \right|^2 \leq M$ .

(b)  $E \left\| \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-h} \sum_{i=1}^N \lambda_i e_{it} \varepsilon_{t+h} \right\|^2 \leq M$ , where  $E(\lambda_i e_{it} \varepsilon_{t+h}) = 0$  for all  $(i, t)$ .

**Assumption 5 - moments and CLT for the score vector**

(a)  $E(\varepsilon_{t+h}) = 0$  and  $E|\varepsilon_{t+h}|^2 < M$ .

(b)  $E\|z_t\|^4 \leq M$  and  $\frac{1}{T} \sum_{t=1}^T z_t z'_t \rightarrow^P \Sigma_{zz} > 0$ .

(c) As  $T \rightarrow \infty$ ,  $\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} z_t \varepsilon_{t+h} \rightarrow^d N(0, \Omega)$ , where  $E \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} z_t \varepsilon_{t+h} \right\|^2 < M$ , and  $\Omega \equiv \lim_{T \rightarrow \infty} \text{Var} \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} z_t \varepsilon_{t+h} \right) > 0$ .

Assumption 1(a) imposes the assumption that factors are non-degenerate. Assumption 1(b) ensures that each factor contributes non-trivially to the variance of  $X_t$ , i.e. the factors are pervasive and affect all cross sectional units. These assumptions ensure that there are  $r$  identifiable factors in the model. Recently, Onatski (2011) considers a class of “weak” factor models, where the factor loadings are modeled as local to zero. Under this assumption, the estimated factors are no longer consistent for

the unobserved (rotated) factors. In this paper, we do not consider this possibility. Assumption 1(c) ensures that  $Q \equiv p \lim \left( \tilde{F}'F/T \right)$  is unique. Without this assumption,  $Q$  is only determined up to orthogonal transformations. See Bai (2003, proof of his Proposition 1).

Assumption 2 imposes weak cross-sectional and serial dependence conditions in the idiosyncratic error term  $e_{it}$ . In particular, we allow for the possibility that  $e_{it}$  is dependent across individual units and over time, but we require that the degree of dependence decreases as the time and the cross sectional distance (regardless of how it is defined) between observations increases. This assumption is compatible with the approximate factor model of Chamberlain and Rothschild (1983) and Connor and Korajczyk (1986, 1993), in which cross section units are weakly correlated. Assumption 2 allows for heteroskedasticity in both dimensions and requires the idiosyncratic error term to have finite eighth moments.

Assumption 3 restricts the degree of dependence among the vector of regressors  $\{z_t\}$ , the factor loadings  $\{\lambda_i\}$  and the idiosyncratic error terms  $\{e_{it}\}$ . If we assume that  $\{z_t\}$ ,  $\{\lambda_i\}$  and  $\{e_{it}\}$  are mutually independent (as in Bai and Ng (2006)), then Assumptions 3(a), 3(c) with  $z_t = F_t$  and 3(d) are implied by Assumptions 1 and 2. Similarly, Assumption 3(b) holds if we assume that  $\{z_t\}$  and  $\{e_{it}\}$  are independent and the following weak dependence condition on  $\{e_{it}\}$  holds

$$\frac{1}{NT^2} \sum_{t,s,l,u} \sum_{i,j} |Cov(e_{it}e_{is}, e_{jl}e_{ju})| < \Delta < \infty. \quad (6)$$

Bai (2009) relies on a similar condition (part 1 of his Assumption C.4) to establish the asymptotic properties of the interactive effects estimator. As he explains, this condition is slightly stronger than the assumption that the second moment of  $N^{-1/2} \sum_{i=1}^N \eta_i$  is bounded, where  $\eta_i \equiv \left( T^{-1/2} \sum_{t=1}^T e_{it} \right)^2 - E \left( T^{-1/2} \sum_{t=1}^T e_{it} \right)^2$  (note that  $Var \left( N^{-1/2} \sum_{i=1}^N \eta_i \right)$  equals the left hand side of the above inequality without the absolute value). It holds if  $e_{it}$  is i.i.d. over  $i$  and  $t$  and  $E(e_{it}^4) < M$ . Assumptions 3(a)-3(c) are equivalent to Assumptions D, F1 and F2 of Bai (2003) when  $z_t = F_t$ .

To describe Assumption 3(e), for each  $t$ , let

$$\phi_t \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^N \lambda_i e_{it}, \text{ and } \Gamma_t \equiv Var(\phi_t) = E(\phi_t \phi_t'),$$

since we assume that  $E(\lambda_i e_{it}) = 0$  for all  $(i, t)$ . Assumption 3(e) requires that  $\frac{1}{T} \sum_{t=1}^T [\phi_t \phi_t' - E(\phi_t \phi_t')]$  converges in probability to zero. This follows under weak dependence conditions on  $\{\lambda_i e_{it}\}$  over  $(i, t)$ .  $\Gamma_t$  is related to the asymptotic variance of  $\sqrt{N} \left( \tilde{F}_t - HF_t \right)$ , as shown by Bai (2003, cf. Theorem 1).

Assumption 4 imposes weak dependence between the idiosyncratic errors and the regression errors. Part (a) holds if  $\{e_{it}\}$  is independent of  $\{\varepsilon_t\}$  and the weak dependence condition (6) holds. Similarly, part (b) holds if  $\{\lambda_i\}$ ,  $\{e_{is}\}$  and  $\{\varepsilon_t\}$  are three mutually independent groups of random variables and Assumption 2 holds.

Assumption 5 imposes moment conditions on  $\{\varepsilon_{t+h}\}$ , on  $\{z_t\}$  and on the score vector  $\{z_t \varepsilon_{t+h}\}$ . Part b) requires  $\{z_t z_t'\}$  to satisfy a law of large numbers. Part c) requires the score to satisfy a

central limit theorem, where  $\Omega$  denotes the limiting variance of the scaled average of the scores. The dependence structure of the scores  $\{z_t \varepsilon_{t+h}\}$  dictates the form of the covariance matrix estimator to be used for inference on  $\delta$ . For instance, Bai and Ng (2006) impose the assumption that  $\varepsilon_{t+h}$  is a martingale difference sequence with respect to  $\{z_t, y_t, z_{t-1}, y_{t-1}, \dots\}$  when  $h > 0$ . Under this assumption,  $\Omega = \lim \frac{1}{T} \sum_{t=1}^{T-h} E(z_t z_t' \varepsilon_{t+h}^2)$ , and the appropriate covariance matrix estimator is an heteroskedasticity robust variance estimator. As we will see later, the form of  $\Omega$  will also dictate the type of bootstrap we should use. In particular, under the martingale difference sequence on  $\varepsilon_{t+h}$ , a wild bootstrap method will be appropriate. Section 4 will consider this specific bootstrap scheme.

Given Assumptions 1-5, we can state our main result as follows. We introduce the following notation:

$$\begin{aligned} H_0 &\equiv p \lim H = V^{-1} Q \Sigma_\Lambda, \quad Q \equiv p \lim \left( \frac{\tilde{F}' F}{T} \right), \quad V \equiv p \lim \tilde{V}, \quad \text{and} \\ \Phi_0 &\equiv \text{diag}(H_0, I_q). \end{aligned}$$

Additionally, we let  $\Sigma_{WF} = E(W_t F_t')$ .

**Theorem 2.1** *Let Assumptions 1-5 hold. If  $\sqrt{T}/N \rightarrow c$ , with  $0 \leq c < \infty$ , then*

$$\sqrt{T} (\hat{\delta} - \delta) \rightarrow^d N(-c \Delta_\delta, \Sigma_\delta),$$

where  $\Sigma_\delta = (\Phi_0')^{-1} \Sigma_{zz}^{-1} \Omega \Sigma_{zz}^{-1} \Phi_0^{-1}$ , and

$$\begin{aligned} \Delta_\delta &\equiv \begin{pmatrix} \Delta_\alpha \\ \Delta_\beta \end{pmatrix} = (\Phi_0 \Sigma_{zz} \Phi_0')^{-1} \begin{pmatrix} \Sigma_{\tilde{F}} + V \Sigma_{\tilde{F}} V^{-1} \\ \Sigma_{W\tilde{F}} V \Sigma_{\tilde{F}} V^{-1} \end{pmatrix} (H_0^{-1})' \alpha, \quad \text{with} \\ \Sigma_{\tilde{F}} &\equiv V^{-1} (Q \Gamma Q') V^{-1}, \\ \Sigma_{W\tilde{F}} &\equiv p \lim \left( \frac{W' \tilde{F}}{T} \right) = \Sigma_{WF} H_0'. \end{aligned}$$

If  $\Sigma_{WF} = 0$ , the asymptotic bias is equal to

$$-c \Delta_{\hat{\delta}} = -c \begin{pmatrix} [\Sigma_{\tilde{F}} + V \Sigma_{\tilde{F}} V^{-1}] (H_0^{-1})' \alpha \\ 0 \end{pmatrix}.$$

Theorem 2.1 gives the asymptotic distribution of  $\sqrt{T} (\hat{\delta} - \delta)$  under the condition that  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$ . When  $c = 0$ , we obtain the same limiting distribution as in Bai and Ng (2006) under a set of assumptions that is weaker than theirs, as we discussed above. As in Bai and Ng (2006), factors estimation error does not contribute to the asymptotic distribution when  $c = 0$ . This is no longer the case when  $c > 0$ . Under this alternative condition, an asymptotic bias appears, as was recently discussed by Ludvigson and Ng (2009b) in the context of a simpler regression model without observed regressors  $W_t$ . Our Theorem 2.1 complements their results by providing an expression of the bias for  $\hat{\delta}$  when the factor-augmented regression model includes also observed regressors in addition to the unobserved factors  $F_t$ .



Several remarks follow. First, the expression for  $\Delta_\delta$  is proportional to  $(H_0^{-1})' \alpha = p \lim \hat{\alpha}$ , implying that when  $\alpha = 0$ , no asymptotic bias exists independently of the value of  $c$ . Second, the asymptotic bias for both  $\hat{\alpha}$  and  $\hat{\beta}$  is a function of

$$\Sigma_{\tilde{F}} \equiv V^{-1} Q \Gamma Q' V^{-1} = \lim_{N, T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T V^{-1} Q \Gamma_t Q' V^{-1},$$

where  $V^{-1} Q \Gamma_t Q' V^{-1}$  is the asymptotic variance of  $\sqrt{N} (\tilde{F}_t - H F_t)$  (see Bai (2003)). Thus, the bias depends on the sampling variance-covariance matrix of the estimation error incurred by the principal components estimator  $\tilde{F}_t$ , averaged over time. This variance matrix depends on the cross sectional dependence of  $\{e_{it}\}$  via  $\Gamma_t \equiv \text{Var} \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N \lambda_i e_i \right)$ . As we will see next, the main implication for the validity of the bootstrap is that it needs to reproduce this cross sectional dependence when  $c \neq 0$  but not otherwise. Third, the existence of measurement error in  $\tilde{F}_t$  contaminates the estimators of the remaining parameters  $\beta$ , i.e.  $\hat{\beta}$  is asymptotically biased due to the measurement error in  $\tilde{F}_t$ . The asymptotic bias associated with  $\hat{\beta}$  will be zero only in the special case when the observed regressors and the factors are not correlated (i.e.  $\Sigma_{WF} = 0$ ) (or when  $\alpha = 0$ ).

### 3 A general residual-based bootstrap

The main contribution of this section is to propose a general residual-based bootstrap method and discuss its consistency for factor-augmented regression models under a set of sufficient high-level conditions on the bootstrap residuals. These high level conditions can be verified for any bootstrap scheme that resamples residuals. We verify these conditions for a two-step wild bootstrap scheme in Section 4.

#### 3.1 Bootstrap data generating process and estimation

Let  $\{e_t^* = (e_{1t}^*, \dots, e_{Nt}^*)'\}$  denote a bootstrap sample from  $\{\tilde{e}_t = X_t - \tilde{\Lambda} \tilde{F}_t\}$  and  $\{\varepsilon_{t+h}^*\}$  a bootstrap sample from  $\{\hat{\varepsilon}_{t+h} = y_{t+h} - \hat{\alpha}' \tilde{F}_t - \hat{\beta}' W_t\}$ . We consider the following bootstrap DGP:

$$y_{t+h}^* = \hat{\alpha}' \tilde{F}_t + \hat{\beta}' W_t + \varepsilon_{t+h}^*, \quad t = 1, \dots, T - h, \quad (7)$$

$$X_t^* = \tilde{\Lambda} \tilde{F}_t + e_t^*, \quad t = 1, \dots, T. \quad (8)$$

Estimation proceeds in two stages. First, we estimate the factors by the method of principal components using the bootstrap panel data set  $\{X_t^*\}$ . Second, we run a regression of  $y_{t+h}^*$  on the bootstrap *estimated* factors and on the fixed observed regressors  $W_t$ .

More specifically, given  $\{X_t^*\}$ , we estimate the bootstrap factor loadings and the bootstrap factors by minimizing the bootstrap objective function

$$V^*(F, \Lambda) = \frac{1}{TN} \sum_{t=1}^T \sum_{i=1}^N (X_{it}^* - \lambda_i' F_t)^2$$

subject to the normalization constraint that  $F'F/T = I_r$ . The  $T \times r$  matrix containing the estimated bootstrap factors is denoted by  $\tilde{F}^* = (\tilde{F}_1^*, \dots, \tilde{F}_T^*)'$  and it is equal to the  $r$  eigenvectors of  $X^*X^*/NT$  (multiplied by  $\sqrt{T}$ ) corresponding to the  $r$  largest eigenvalues. The  $N \times r$  matrix of estimated bootstrap loadings is given by  $\tilde{\Lambda}^* = (\tilde{\lambda}_1^*, \dots, \tilde{\lambda}_N^*)' = X^*\tilde{F}^*/T$ .

Given the estimated bootstrap factors  $\tilde{F}_t^*$ , the second estimation stage is to regress  $y_{t+h}^*$  on  $\hat{z}_t^* = (\tilde{F}_t^{*'} \quad W_t')'$ . Because the bootstrap scheme used to generate  $y_{t+h}^*$  is residual-based, we fix the observed regressors  $W_t$  in the bootstrap regression. We replace  $\tilde{F}_t$  with  $\tilde{F}_t^*$  to mimic the fact that in the original regression model the factors  $F_t$  are latent and need to be estimated with  $\tilde{F}_t$ . This yields the bootstrap OLS estimator

$$\hat{\delta}^* = \left( \frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t^* \hat{z}_t^{*'} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t^* y_{t+h}^*. \quad (9)$$

$\hat{\delta}^*$  is the bootstrap analogue of  $\hat{\delta}$ , the OLS estimator based on the original sample.

### 3.2 Bootstrap high level conditions

In this section, we provide a set of high level conditions on  $\{e_{it}^*\}$  and  $\{\varepsilon_{t+h}^*\}$  that will allow us to characterize the bootstrap distribution of  $\hat{\delta}^*$ .

**Condition A\*** (weak time series and cross section dependence in  $e_{it}^*$ )

- (a)  $E^*(e_{it}^*) = 0$  for all  $(i, t)$ .
- (b)  $\frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T |\gamma_{st}^*|^2 = O_P(1)$ , where  $\gamma_{st}^* = E^* \left( \frac{1}{N} \sum_{i=1}^N e_{it}^* e_{is}^* \right)$ .
- (c)  $\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T E^* \left| \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{it}^* e_{is}^* - E^*(e_{it}^* e_{is}^*)) \right|^2 = O_P(1)$ .

**Condition B\*** (weak dependence among  $\hat{z}_t, \tilde{\lambda}_i$ , and  $e_{it}^*$ )

- (a)  $\frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T \tilde{F}_s \tilde{F}_t' \gamma_{st}^* = O_P(1)$ .
- (b)  $\frac{1}{T} \sum_{t=1}^T E^* \left\| \frac{1}{\sqrt{TN}} \sum_{s=1}^T \sum_{i=1}^N \hat{z}_s (e_{it}^* e_{is}^* - E^*(e_{it}^* e_{is}^*)) \right\|^2 = O_P(1)$ , where  $\hat{z}_s = (\tilde{F}_s' \quad W_s')'$ .
- (c)  $E^* \left\| \frac{1}{\sqrt{TN}} \sum_{t=1}^T \sum_{i=1}^N \hat{z}_t \tilde{\lambda}_i' e_{it}^* \right\|^2 = O_P(1)$ .
- (d)  $\frac{1}{T} \sum_{t=1}^T E^* \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^N \tilde{\lambda}_i e_{it}^* \right\|^2 = O_P(1)$ .
- (e)  $\frac{1}{T} \sum_{t=1}^T \left( \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \right) \left( \frac{e_t^{*\prime} \tilde{\Lambda}}{\sqrt{N}} \right) - \Gamma^* = o_P^*(1)$ , in probability, where  $\Gamma^* \equiv \frac{1}{T} \sum_{t=1}^T \text{Var}^* \left( \frac{1}{\sqrt{N}} \tilde{\Lambda}' e_t^* \right) > 0$ .

**Condition C\*** (weak dependence between  $e_{it}^*$  and  $\varepsilon_{t+h}^*$ )

- (a)  $\frac{1}{T} \sum_{t=1}^T E^* \left| \frac{1}{\sqrt{TN}} \sum_{s=1}^{T-h} \sum_{i=1}^N \varepsilon_{s+h}^* (e_{it}^* e_{is}^* - E^*(e_{it}^* e_{is}^*)) \right|^2 = O_P(1)$ .
- (b)  $E^* \left\| \frac{1}{\sqrt{TN}} \sum_{t=1}^{T-h} \sum_{i=1}^N \tilde{\lambda}_i e_{it}^* \varepsilon_{t+h}^* \right\|^2 = O_P(1)$ , where  $E(e_{it}^* \varepsilon_{t+h}^*) = 0$  for all  $(i, t)$ .

(c)  $\frac{1}{T} \sum_{t=1}^{T-h} \sum_{s=1}^T \tilde{F}_s \varepsilon_{t+h}^* \gamma_{st}^* = O_{P^*}(1)$ , in probability.

**Condition D\* (bootstrap CLT)**

(a)  $E^*(\varepsilon_{t+h}^*) = 0$  and  $\frac{1}{T} \sum_{t=1}^{T-h} E^* |\varepsilon_{t+h}^*|^2 = O_P(1)$ .

(b)  $\Omega^{*-1/2} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \hat{z}_t \varepsilon_{t+h}^* \xrightarrow{d^*} N(0, I_p)$ , in probability, where  $E^* \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \hat{z}_t \varepsilon_{t+h}^* \right\|^2 = O_P(1)$ , and  $\Omega^* \equiv \text{Var}^* \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \hat{z}_t \varepsilon_{t+h}^* \right) > 0$ .

Conditions A\*-D\* are the bootstrap analogue of Assumptions 1 through 5. However, contrary to Assumptions 1-5, which pertain to the data generating process and cannot be verified in practice, Conditions A\*-D\* can be verified for any particular bootstrap algorithm used to generate the bootstrap residuals and idiosyncratic error terms. More importantly, we can devise bootstrap schemes to verify these conditions and hence ensure bootstrap validity. For instance, part (a) of Condition A\* requires the bootstrap mean of  $e_{it}^*$  to be zero for all  $(i, t)$  whereas part (a) of Condition D\* requires that the same is true for  $\varepsilon_t^*$ . The practical implication is that we should make sure to construct bootstrap residuals with mean zero, e.g. to recenter residuals before applying a nonparametric bootstrap method. Parts b) and c) of Condition A\* impose weak dependence conditions on  $\{e_{it}^*\}$  over  $(i, t)$ . For instance, these conditions are satisfied if we resample  $\{e_{it}^*\}$  in an i.i.d. fashion over the two indices  $(i, t)$ . Condition B\* imposes further restrictions on the dependence among  $\hat{z}_t, \tilde{\lambda}_i$  and the idiosyncratic errors  $e_{it}^*$ . Since  $\hat{z}_t$  and  $\tilde{\lambda}_i$  are fixed in the bootstrap world, Condition B\* is implied by appropriately restricting the dependence of  $e_{it}^*$  over  $(i, t)$ . Similarly, Condition C\* restricts the amount of dependence between  $\{e_{is}^*\}$  and  $\{\varepsilon_{t+h}^*\}$ . If these two sets of bootstrap innovations are independent of one another, then weak dependence on  $\{e_{is}^*\}$  suffices for Condition C\* to hold. Finally, Condition D\* requires the bootstrap regression scores  $\hat{z}_t \varepsilon_{t+h}^*$  to obey a central limit theorem in the bootstrap world.

**3.3 Bootstrap results**

Under Conditions A\*-D\*, we can show the consistency of the bootstrap principal component estimator  $\tilde{F}^*$  for a rotated version of the true “latent” bootstrap factors  $\tilde{F}$ , a crucial result in proving the first order asymptotic validity of the bootstrap in this context.

More specifically, according to (7)-(8), the common factors underlying the bootstrap panel data  $\{X_t^*\}$  are given by  $\tilde{F}_t$  (with  $\tilde{\Lambda}$  as factor loadings). Nevertheless, just as  $\tilde{F}_t$  estimates a rotation of  $F_t$ , the estimated bootstrap factors  $\tilde{F}_t^*$  estimate  $H^* \tilde{F}_t$ , where  $H^*$  is the bootstrap analogue of the rotation matrix  $H$  defined in (4), i.e.

$$H^* = \tilde{V}^{*-1} \frac{\tilde{F}^{*'} \tilde{F} \tilde{\Lambda}' \tilde{\Lambda}}{T N}, \tag{10}$$

where  $\tilde{V}^*$  is the  $r \times r$  diagonal matrix containing on the main diagonal the  $r$  largest eigenvalues of  $X^* X^{*'} / NT$ , in decreasing order.

**Lemma 3.1** *Assume Assumptions 1-5 hold and suppose we generate bootstrap data  $\{y_{t+h}^*, X_t^*\}$  according to the residual-based bootstrap DGP (7) and (8) by relying on bootstrap residuals  $\{\varepsilon_{t+h}^*\}$  and  $\{e_{it}^*\}$  such that Conditions A\*-D\* are satisfied. Then, as  $N, T \rightarrow \infty$ ,*

$$\frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t^* - H^* \tilde{F}_t \right\|^2 = O_{P^*}(\delta_{NT}^{-2}),$$

*in probability, where  $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$ .*

According to Lemma 3.1, the time average of the squared deviations between the *estimated* bootstrap factors  $\tilde{F}_t^*$  and a rotation of the “latent” bootstrap factors given by  $H^* \tilde{F}_t$  vanishes in probability under the bootstrap measure  $P^*$  as  $N, T \rightarrow \infty$ , conditional on a sample which lies in a set with probability converging to one. Contrary to  $H$ ,  $H^*$  does not depend on population values and can be computed for any bootstrap sample, given the original sample. Hence, rotation indeterminacy is not a problem in the bootstrap world. Because the bootstrap factor DGP (8) satisfies the constraints that  $\tilde{F}' \tilde{F} / T = I_r$  and  $\tilde{\Lambda}' \tilde{\Lambda}$  is a diagonal matrix, we can actually show that  $H^*$  is asymptotically (as  $N, T \rightarrow \infty$ ) equal to  $H_0^* = \text{diag}(\pm 1)$ , a diagonal matrix with diagonal elements equal to  $\pm 1$ , where the sign of is determined by the sign of  $\tilde{F}' \tilde{F} / T$  (see Lemma B.1; the proof follows by arguments similar to those used in Bai and Ng (2011) and Stock and Watson (2002)). Thus, the bootstrap factors are identified up to a change of sign.

The main implication from Lemma 3.1 is that the bootstrap OLS estimator that one obtains from regressing  $y_{t+h}^*$  on  $\hat{z}_t^*$  estimates a rotated version of  $\hat{\delta}$ , given by  $\delta^* \equiv \begin{pmatrix} \hat{\alpha}' H^{*-1} & \hat{\beta}' \end{pmatrix}' = \Phi^{*-1} \hat{\delta}$ , where  $\Phi^* = \text{diag}(H^*, I_q)$ . Asymptotically,  $\delta^*$  is equal to  $\Phi_0^{*-1} \hat{\delta}$ , where  $\Phi^* = \text{diag}(H_0^*, I_q)$ , which can be interpreted as a sign-adjusted version of  $\hat{\delta}$ .

Our next result characterizes the asymptotic bootstrap distribution of  $\sqrt{T}(\hat{\delta}^* - \delta^*)$  when  $\sqrt{T}/N \rightarrow c$ , with  $0 \leq c < \infty$ . We add the two following conditions.

**Condition E\*.**  $p \lim \Omega^* = \Phi_0 \Omega \Phi_0'$ .

**Condition F\*.**  $p \lim \Gamma^* = Q \Gamma Q'$ .

$\Omega^*$  is the bootstrap variance of the scaled average of the bootstrap regression scores  $\hat{z}_t \varepsilon_{t+h}^*$  (as defined in Condition D\*(b)). Since  $\tilde{F}_t$  estimates a rotated version of the latent factors given by  $H_0 F_t$ ,  $\hat{z}_t$  estimates a rotated version of  $z_t$  given by  $\Phi_0 z_t$ , and therefore  $\Omega^*$  is the sample analog of  $\Phi_0 \Omega \Phi_0'$  provided we choose  $\varepsilon_{t+h}^*$  to mimic the time series dependence of  $\varepsilon_{t+h}$ . Condition E\* imposes formally this condition. Similarly, by Condition B\*(e),  $\Gamma^* \equiv \frac{1}{T} \sum_{t=1}^T \text{Var}^* \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N \tilde{\lambda}_i e_{it}^* \right)$ . Because  $\tilde{\lambda}_i$  estimates  $Q \lambda_i$ ,  $\Gamma^*$  is the sample analogue of  $Q \Gamma Q'$  if we choose  $e_{it}^*$  to mimic the cross sectional dependence of  $e_{it}$  (interestingly enough, mimicking the time series dependence of  $e_{it}$  is not relevant). Condition F\* formalizes this requirement.

**Theorem 3.1** *Let Assumptions 1-5 hold and consider any residual-based bootstrap scheme for which Conditions A\*-D\* are verified. Suppose  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$ . If in addition the two following conditions hold: (1) Condition E\* is verified, and (2)  $c = 0$  or Condition F\* is verified; then as  $N, T \rightarrow \infty$ ,*

$$\sqrt{T} \left( \hat{\delta}^* - \delta^* \right) \rightarrow^{d^*} N \left( -c \left( \Phi_0^{*'} \right)^{-1} \Delta_\delta, \left( \Phi_0^{*'} \right)^{-1} \Sigma_\delta \Phi_0^{*-1} \right),$$

*in probability, where  $\delta^* \equiv \left( \Phi_0^{*-1} \right)' \hat{\delta}$ , with  $\Phi_0^* = p \lim \Phi^* = \text{diag} \left( H_0^*, I_q \right)$  a diagonal matrix with  $\pm 1$  in the main diagonal, and  $\Delta_\delta$  and  $\Sigma_\delta$  are as defined in Theorem 2.1.*

According to Theorem 3.1,  $\sqrt{T} \left( \hat{\delta}^* - \delta^* \right)$  is asymptotically distributed as a normal random vector with mean equal to  $-c \left( \Phi_0^{*'} \right)^{-1} \Delta_\delta$ . Just as the asymptotic bias of  $\sqrt{T} \left( \hat{\alpha} - \left( H^{-1} \right)' \alpha \right)$  is proportional to  $\left( H^{-1} \right)' \alpha$ , the bootstrap asymptotic bias is proportional to  $\left( H_0^{*-1} \right)' \hat{\alpha}$ . Since  $\hat{\alpha}$  converges in probability to  $\left( H^{-1} \right)' \alpha$ , the bootstrap bias of  $\hat{\alpha}^*$  converges to  $-c \left( H_0^{*'} \right)^{-1} \Delta_\alpha$  provided we ensure that Condition F\* is satisfied. The bootstrap bias of  $\hat{\beta}^*$  depends both on the bootstrap analogue of  $\Sigma_{W\tilde{F}} = p \lim \frac{W'\tilde{F}}{T} = \Sigma_{WF} H_0'$  and on  $\left( H_0^{*-1} \right)' \hat{\alpha}$ . Since the bootstrap analogue of  $\Sigma_{W\tilde{F}}$  is  $p \lim \frac{W'\tilde{F}^*}{T}$ , which converges to  $\Sigma_{W\tilde{F}} H_0^{*'}$ , the rotation matrix  $H_0^{*'}$  “cancels out” with  $\left( H_0^{*-1} \right)' \hat{\alpha}$ , leaving the bootstrap bias of  $\hat{\beta}^*$  unaffected by the sign problem that arises for  $\hat{\alpha}^*$  (Lemma B.4 formalizes this argument). Similarly, the asymptotic variance-covariance matrix of  $\hat{\delta}^*$  is equal to  $\left( \Phi_0^{*'} \right)^{-1} \Sigma_\delta \Phi_0^{*-1}$  provided we choose  $\varepsilon_{t+h}^*$  so as to verify Condition E\*.

For bootstrap consistency, we need the bootstrap bias and variance to match the bias and variance of the limiting distribution of the original OLS estimator. Since  $H_0^*$  (hence  $\Phi_0^*$ ) is not necessarily equal to the identity matrix, Theorem 3.1 shows that this is not the case. Hence, the bootstrap distribution of  $\sqrt{T} \left( \hat{\delta}^* - \delta^* \right)$  is not a consistent estimator of the sampling distribution of  $\sqrt{T} \left( \hat{\delta} - \delta \right)$  in general. This is true even if we choose  $\varepsilon_{t+h}^*$  and  $e_{it}^*$  so that Conditions E\* and F\* are satisfied. The reason is that the bootstrap factors are not identified. In particular, because the bootstrap principal components estimator does not necessarily identify the sign of the bootstrap factors, the mean of each element of  $\sqrt{T} \left( \hat{\delta}^* - \delta^* \right)$  corresponding to the coefficients associated with the latent factors may have the “wrong” sign even asymptotically. The same “sign” problem will affect the off-diagonal elements of the bootstrap covariance matrix asymptotically (although not the main diagonal elements). As we mentioned above, the coefficients associated with  $W_t$  are correctly identified in the bootstrap world as well as in the original sample and therefore this sign problem does not affect these coefficients.

In order to obtain a consistent estimator of the distribution of  $\sqrt{T} \left( \hat{\delta} - \delta \right)$ , our proposal is to consider the bootstrap distribution of the rotated version of  $\sqrt{T} \left( \hat{\delta}^* - \delta^* \right)$  given by  $\sqrt{T} \left( \Phi^{*'} \hat{\alpha}^* - \hat{\alpha} \right)$ . This rotation is feasible because  $\Phi^*$  does not depend on any population quantities and can be computed for any bootstrap and original samples. Since  $\Phi^*$  is asymptotically equal to  $\Phi_0^* = \text{diag} \left( \pm 1, I_q \right) = \text{diag} \left( \text{sign} \left( \tilde{F}^{*'} \tilde{F} \right), I_q \right)$ ,  $\Phi^{*'} \hat{\delta}^*$  is asymptotically equal to a sign-adjusted version of  $\hat{\delta}^*$ . The following result is an immediate corollary to Theorems 2.1 and 3.1.

**Corollary 3.1** *Under the conditions of Theorem 3.1, if  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$ , as  $N, T \rightarrow \infty$ , then  $\sup_{x \in \mathbb{R}^p} \left| P^* \left( \sqrt{T} \left( \Phi^{*\prime} \hat{\delta}^* - \hat{\delta} \right) \leq x \right) - P \left( \sqrt{T} \left( \hat{\delta} - \delta \right) \leq x \right) \right| \rightarrow^P 0$ .*

Corollary 3.1 justifies the use of a residual-based bootstrap method for constructing bootstrap percentile confidence intervals for the elements of  $\delta$ . When  $c = 0$ , the crucial condition for bootstrap validity is Condition E\*, which requires  $\{\varepsilon_t^*\}$  to be chosen so as to mimic the dependence structure of the scores  $z_t \varepsilon_{t+h}$ . This condition ensures that the bootstrap variance-covariance matrix of  $\Phi^{*\prime} \hat{\delta}^*$  is correct asymptotically. When  $c \neq 0$ , Condition F\* is also crucial to ensure that the bootstrap distribution correctly captures the bias. When both Conditions E\* and F\* are satisfied, the bootstrap contains a built-in bias correction term that is absent in the Bai and Ng (2006) asymptotic normal distribution, and we might expect it to outperform the normal approximation. A bootstrap method that does not involve factor estimation in the bootstrap world will not contain this bias correction and will not be valid in this context.

## 4 Wild bootstrap

In this section we propose a particular bootstrap method for generating  $\{\varepsilon_{t+h}^*\}$  and  $\{e_{it}^*\}$  and show its first-order asymptotic validity under a set of primitive conditions.

### Bootstrap algorithm

1. For  $t = 1, \dots, T$ , let

$$X_t^* = \tilde{\Lambda} \tilde{F}_t + e_t^*,$$

where  $\{e_t^* = (e_{1t}^*, \dots, e_{Nt}^*)\}$  is such that

$$e_{it}^* = \tilde{e}_{it} \eta_{it},$$

is a resampled version of  $\{\tilde{e}_{it} = X_{it} - \tilde{\lambda}_i' \tilde{F}_t\}$  obtained with the wild bootstrap. The external random variables  $\eta_{it}$  are i.i.d. across  $(i, t)$  and have mean zero and variance one.

2. Estimate the bootstrap factors  $\tilde{F}^*$  and the bootstrap loadings  $\tilde{\Lambda}^*$  using  $X^*$ .
3. For  $t = 1, \dots, T - h$ , let

$$y_{t+h}^* = \hat{\alpha}' \tilde{F}_t + \hat{\beta}' W_t + \varepsilon_{t+h}^*,$$

where the error term  $\varepsilon_{t+h}^*$  is a wild bootstrap resampled version of  $\hat{\varepsilon}_{t+h}$ , i.e.

$$\varepsilon_{t+h}^* = \hat{\varepsilon}_{t+h} v_{t+h},$$

where the external random variable  $v_{t+h}$  is i.i.d.  $(0, 1)$  and is independent of  $\eta_{it}$ .

4. Regress  $y_{t+h}^*$  generated in 3. on the estimated bootstrap factors and the fixed regressors  $\hat{z}_t^* = (\tilde{F}_t^{*'}, W_t')'$ . This yields the bootstrap OLS estimators

$$\hat{\delta}^* = \left( \sum_{t=1}^{T-h} \hat{z}_t^* \hat{z}_t^{*'} \right)^{-1} \sum_{t=1}^{T-h} \hat{z}_t^* y_{t+h}^*.$$

To prove the validity of this residual-based wild bootstrap, we add the following assumptions.

**Assumption 6.**  $\lambda_i$  are either deterministic such that  $\|\lambda_i\| \leq M < \infty$ , or stochastic such that  $E \|\lambda_i\|^{12} \leq M < \infty$  for all  $i$ ;  $E \|F_t\|^{12} \leq M < \infty$ ;  $E |e_{it}|^{12} \leq M < \infty$ , for all  $(i, t)$ ; and for some  $q > 1$ ,  $E |\varepsilon_{t+h}|^{4q} \leq M < \infty$ , for all  $t, h$ .

**Assumption 7.**  $E(\varepsilon_{t+h} | y_t, F_t, y_{t-1}, F_{t-1}, \dots) = 0$  for any  $h > 0$ , and  $F_t$  and  $\varepsilon_t$  are independent of the idiosyncratic errors  $e_{is}$  for all  $(i, s, t)$ .

**Assumption 8.**  $E(e_{it}e_{js}) = 0$  if  $i \neq j$ .

Assumption 6 strengthens the moment conditions assumed in Assumption 1.b), 2.a), and 5.a), respectively. The moment conditions on  $\lambda_i$ ,  $F_t$  and  $e_{it}$  suffice to show that  $E |\lambda_i' F_s e_{it}|^4 < M$  (while maintaining that  $E |e_{it}|^8 < M$ ). If we assume that the three groups of random variables  $\{F_t\}$ ,  $\{e_{it}\}$  and  $\{\lambda_i\}$  are mutually independent (as in Bai and Ng (2006)), then it suffices that  $E \|\lambda_i\|^4 \leq M < \infty$ ,  $E \|F_t\|^4 \leq M < \infty$  (and  $E |e_{it}|^8 < M$ ).

Assumption 7 was used by Bai and Ng (2006). It imposes a martingale difference condition on the regression errors  $\varepsilon_{t+h}$ , implying that these are serially uncorrelated but possibly heteroskedastic. In addition,  $\varepsilon_t$  is independent of  $e_{is}$  for all  $(i, t, s)$ . Under Assumption 7,

$$\Omega = \lim Var \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} z_t \varepsilon_{t+h} \right) = \lim \frac{1}{T} \sum_{t=1}^{T-h} E (z_t z_t' \varepsilon_{t+h}^2),$$

which motivates using a wild bootstrap to generate  $\varepsilon_{t+h}^*$ . For this bootstrap scheme,

$$\Omega^* = \frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t \hat{z}_t' \hat{\varepsilon}_{t+h}^2,$$

which corresponds to the estimator used by Bai and Ng (2006) (cf. their equation (3)) and is consistent for  $\Phi_0 \Omega \Phi_0'$  under Assumptions 1-7. We assume independence between  $e_{it}$  and  $\varepsilon_{t+h}$  and generate  $\varepsilon_{t+h}^*$  independently of  $e_{it}^*$ , but we conjecture that our results will be valid under weak forms of correlation between the two sets of errors because the limiting distribution of the OLS estimator remains unchanged under Assumptions 1-5, which allow for dependence between  $e_{it}$  and  $\varepsilon_{t+h}$ , as we proved in Theorem 2.1.

Assumption 8 assumes cross section independence in  $\{e_{it}\}$ , but allows for serial correlation and heteroskedasticity in both directions. This assumption motivates the use of a wild bootstrap to

generate  $\{e_{it}^*\}$ . For this bootstrap scheme, we can show that

$$\Gamma^* = \frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}_i \tilde{\lambda}'_i \tilde{e}_{it}^2 \equiv \frac{1}{T} \sum_{t=1}^T \tilde{\Gamma}_t,$$

where  $\tilde{\Gamma}_t$  corresponds to estimator (5a) in Bai and Ng (2006, p. 1140). As shown by Bai and Ng, this estimator is consistent for  $Q\Gamma Q'$  under cross section independence (and potential heteroskedasticity). Assumption 8 assumes this is the case and thus justifies Condition F\* in this context. As we discussed in the previous section, Condition F\* is not needed if  $c = 0$ . Thus, a wild bootstrap is still asymptotically valid if the idiosyncratic errors are cross sectionally (and serially) dependent when  $\sqrt{T}/N$  converges to zero (as assumed in Bai and Ng (2006)).

Our main result is as follows.

**Theorem 4.1** *Suppose that a residual-based wild bootstrap is used to generate  $\{e_{it}^*\}$  and  $\{\varepsilon_{t+h}^*\}$  with  $E^* |\eta_{it}|^4 < C$  for all  $(i, t)$  and  $E^* |v_{t+h}|^{4q} < C$  for all  $t$ , for some  $q > 1$ . Under Assumptions 1-7, if  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$ , and either Assumption 8 holds or  $c = 0$ , the conclusions of Corollary 3.1 follow.*

## 5 Monte Carlo results

In this section, we report results from a simulation experiment that documents the properties of our bootstrap procedure in factor-augmented regressions.

The data-generating process (DGP) is similar to the one used in Ludvigson and Ng (2009b) to analyze bias. We consider the single factor model:

$$y_t = \alpha F_t + \varepsilon_t, \tag{11}$$

where  $F_t$  is drawn from a standard normal distribution independently over time. The regression error  $\varepsilon_t$  will either be standard normal or heteroskedastic over time. The  $(T \times N)$  matrix of panel variables is generated as:

$$X_{it} = \lambda_i F_t + e_{it}, \tag{12}$$

where  $\lambda_i$  is drawn from a  $U[0, 1]$  distribution (independent across  $i$ ) and the properties of  $e_{it}$  will be discussed below. The only difference with Ludvigson and Ng (2009b) is that they draw the loadings from a standard normal distribution. The use of a uniform distribution increases the cross-correlations and leads to larger biases without having to set the idiosyncratic variance to large values (they set it to 16 in one experiment). Note that this DGP satisfies the conditions PC1 in Bai and Ng (2011) which implies that  $H_0 = \pm 1$ .

We consider two values for the coefficient, either  $\alpha = 0$  or 1. We consider six different scenarios outlined in the table below. When  $\alpha = 0$ , the OLS estimator is unbiased, and the properties of the idiosyncratic components do not matter asymptotically. This leads us to consider only one scenario



with  $\alpha = 0$ . The other five scenarios have  $\alpha = 1$  but differ according to the properties of the regression error,  $\varepsilon_t$ , and of the idiosyncratic error,  $e_{it}$ .

DGP	$\alpha$	$\varepsilon_t$	$e_{it}$
1 (homo-homo)	0	$N(0, 1)$	$N(0, 1)$
2 (homo-homo)	1	$N(0, 1)$	$N(0, 1)$
3 (hetero-homo)	1	$N\left(0, \frac{F_t^2}{3}\right)$	$N(0, 1)$
4 (hetero-hetero)	1	$N\left(0, \frac{F_t^2}{3}\right)$	$N(0, \sigma_i^2)$
5 (hetero-AR)	1	$N\left(0, \frac{F_t^2}{3}\right)$	$AR + N(0, \sigma_i^2)$
6 (hetero-CS)	1	$N\left(0, \frac{F_t^2}{3}\right)$	$CS + N(0, 1)$

DGP 1 is the simplest case with  $\alpha = 0$  and both error terms i.i.d. standard normal in both dimensions. DGP 2 is the same but with  $\alpha = 1$ . This will allow us to isolate the effect of a non-zero coefficient on bias and inference while keeping everything else the same. The third experiment introduces conditional heteroskedasticity in the regression error. We do so by making the variance of  $\varepsilon_t$  depend on the factor and scale so that the asymptotic variance of  $\hat{\alpha}$ ,  $\Sigma_\alpha$ , is 1 in all experiments.

The fourth DGP adds heteroskedasticity to the idiosyncratic error. The variance of  $e_{it}$  is drawn from  $U[.5, 1.5]$  so that the average variance is the same as the homoskedastic case. The fifth DGP introduces serial correlation in the idiosyncratic error term with autoregressive parameter equal to 0.5. The innovations are scaled by  $(1 - .5^2)^{1/2}$  to preserve the variance of the idiosyncratic errors. Finally, the last experiment introduces cross-sectional dependence among idiosyncratic errors. The design is similar to the one in Bai and Ng (2006): the correlation between  $e_{it}$  and  $e_{jt}$  is  $0.5^{|i-j|}$  if  $|i - j| \leq 5$ .

We concentrate on inference about the parameter  $\alpha$  in (11). We consider asymptotic and bootstrap symmetric percentile t confidence intervals at a nominal level of 95%. We report experiments based on 1000 replications with  $B = 399$  bootstrap repetitions. We consider three values for  $N$  (50, 100, and 200) and  $T$  (50, 100, and 200).

We tailor our inference procedures to the properties of the error terms. In other words, when  $\varepsilon_t$  is homoskedastic (DGP 1 and 2), we use the variance estimator under homoskedasticity:

$$\hat{\Sigma}_{\hat{\alpha}} = \hat{\sigma}_\varepsilon^2 \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t^2 \right)^{-1} \quad (13)$$

whereas we use the heteroskedastic-robust version for DGPs 3-6:

$$\hat{\Sigma}_\alpha = \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t^2 \right)^{-1} \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t^2 \hat{\varepsilon}_{t+h}^2 \right) \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t^2 \right)^{-1}. \quad (14)$$

Similarly, we use the homoskedastic estimator of  $\Gamma$  for cases 1-3, the heteroskedasticity-robust estimator for cases 4-5, and the CS-HAC estimator of Bai and Ng (2006) in case 6 with the window size  $n$  equal to  $\min(\sqrt{N}, \sqrt{T})$ . We consider the wild residual-based bootstrap described in Section 4 with the two external variables  $\eta_{it}$  and  $v_t$  both drawn from i.i.d.  $N(0, 1)$ .

We report two sets of results. The first set of results is the bias of the OLS estimator. Because the OLS estimator does not converge to  $\alpha$  but to  $H^{-1}\alpha$  (and  $H$  converges to +1 or -1) and because its bias is proportional to this, the bias will be positive for some replications and negative for others. Reporting the average bias over replications is therefore misleading in this situation. To circumvent this, we report the bias of the *rotated* OLS coefficient,  $H'\hat{\alpha}$ . This rotated coefficient converges to  $\alpha$  in all replications. Note that this rotation is not possible in the real world since the matrix  $H$  depends on population parameters. Note also that this rotation is possible in the bootstrap world (and indeed necessary to obtain consistent inference of the entire coefficient vector, see Corollary 3.1). For the bootstrap world, we report the average of  $H'H^*\hat{\alpha}^* - H'\hat{\alpha}$ , again to ensure that the sign of this bias is always the same. Secondly, we present coverage rates of the associated confidence intervals. For comparison, we also include results for the case where factors do not need to be estimated and are treated as observed (row labeled "true factor" in the tables). This quantifies the loss from having to estimate the factors.

Table 1 provides results for the first two DGPs and illustrates our results. For each DGP, the top panel gives the bias associated with the OLS estimator as well as the plug-in and bootstrap estimates. The second panel for each design provides the coverage rate of intervals based on asymptotic theory, either using the OLS estimator or its bias-corrected version, based on OLS using true factors, and based on the wild bootstrap. From table 1, we see that, as expected, bias is nil when  $\alpha = 0$  (case 1). When  $\alpha \neq 0$ , a negative bias appears. DGP 2 shows that the magnitude of this bias is decreasing in  $N$  (and  $T$ ). The sample estimate of this quantity provides a reasonable approximation to it. However, the bootstrap captures the behavior of the bias well as predicted by theory and better than the sample estimate.

Coverage rates are consistent with these findings. When  $\alpha = 0$  (DGP 1), asymptotic theory is nearly perfect and matches closely the results based on the observed factors. In case 2, with  $\alpha = 1$ , OLS inference suffers from noticeable distortions for all sample sizes. This is because the estimator is biased and the associated t-statistic is not centered at 0. Analytical bias correction corrects most of these distortions. The bootstrap provides even better inference and is quite accurate for  $N \geq 100$ . This loss in accuracy in inference is due to the estimation of the factors as illustrated by the results with the true factors.

Tables 2 and 3 provide results for the other DGPs and show the robustness of the results to the presence of heteroskedasticity in both errors (DGPs 3 and 4), and serial correlation in the idiosyncratic errors (DGP 5). The bias results in table 2 are very similar to those in table 1, although coverage rates reported in table 3 deteriorate relative to the simpler homoskedastic case. The presence of cross-sectional dependence (DGP 6) is interesting. Firstly the bias is larger here (because  $\Gamma$  also includes cross-product terms). Secondly, the wild bootstrap is theoretically not valid since it does not replicate the cross-sectional dependence. Indeed, we see that, contrary to the other cases, the sample estimate of the bias is often better than the bootstrap, especially with  $N = 50$ .

## 6 Conclusion

The main contribution of this paper is to give a set of sufficient high-level conditions under which any residual-based bootstrap method is valid in the context of the factor-augmented regression model in cases where  $\sqrt{T}/N \rightarrow c$ ,  $0 \leq c < \infty$ . Our results show that two crucial conditions for bootstrap validity in this context are that the bootstrap regression scores replicate the time series dependence of the true regression scores, and that either  $c = 0$  or the bootstrap replicates also the cross-sectional dependence of the idiosyncratic error terms.

Our high-level conditions can be checked for any implementation of the bootstrap in this context. We verify them for a particular scheme based on a two-step application of the wild bootstrap. Although the wild bootstrap preserves heteroskedasticity, its validity depends on a martingale difference condition on the regression errors and on cross-sectional independence of the idiosyncratic errors when  $c \neq 0$ .

The martingale difference condition on the regression errors was used in Bai and Ng (2006), but it is overly restrictive when the forecasting horizon  $h$  is larger than one. Thus, relaxing this assumption is important. Although our general results in Sections 2 and 3 allow for serial correlation in the scores (see in particular our Assumption 5(b)), the particular implementation of the two-step residual-based wild bootstrap we consider in Section 4 is not robust to serial dependence. A block bootstrap based method is required in this case. We plan on investigating the validity of such a method in future work. Our high-level conditions will be useful in establishing this result.

A second important extension of the results in this paper is to propose a bootstrap scheme that is able to replicate the cross-sectional dependence of the idiosyncratic error term. As our results show, this is crucial for capturing the bias when  $c \neq 0$ . Our wild bootstrap based method does not allow for cross-sectional dependence. Because there is no natural cross-sectional ordering, devising a nonparametric bootstrap method that is robust to cross-sectional dependence of unknown form is a challenging task.

Another important extension of the results in this paper is the construction of interval forecasts, which we are currently investigating. Finally, the extension to factor-augmented vector autoregressions (FAVAR) first suggested by Boivin and Bernanke (2003) is also important. This case has recently been analyzed by Yamamoto (2011), who proposes a bootstrap scheme that exploits the VAR structure in the factors and the panel variables.

## A Appendix A: Proofs of results in Section 2

We rely on the following lemmas to prove Theorem 2.1.

**Lemma A.1** *Let Assumptions 1-5 hold. Then,  $\frac{1}{T} \sum_{t=1}^{T-h} (\tilde{F}_t - HF_t) \varepsilon_{t+h} = O_P\left(\frac{1}{\delta_{NT}\sqrt{T}}\right)$ , where  $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$ .*

**Lemma A.2** *Let Assumptions 1-5 hold. Then, if  $\sqrt{T}/N \rightarrow c$ , where  $0 \leq c < \infty$ , for any fixed  $h \geq 0$ ,*

a)  $\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (\tilde{F}_t - HF_t) (\tilde{F}_t - HF_t)' = cV^{-1}Q\Gamma Q'V^{-1} + o_P(1)$ .

b)  $\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} HF_t (\tilde{F}_t - HF_t)' = cQ\Gamma Q'V^{-2} + o_P(1)$ .

c)  $\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} W_t (\tilde{F}_t - HF_t)' = c\Sigma_{WF}H_0'Q\Gamma Q'V^{-2} + o_P(1)$ .

d) *Letting  $\Sigma_{\tilde{F}} \equiv V^{-1}Q\Gamma Q'V^{-1}$ , we have that*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \tilde{F}_t (\tilde{F}_t - HF_t)' (H^{-1})' \alpha = \underbrace{c(\Sigma_{\tilde{F}} + V\Sigma_{\tilde{F}}V^{-1})}_{\equiv B_\alpha} p \lim(\hat{\alpha}) + o_P(1), \quad (15)$$

and

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} W_t (\tilde{F}_t - HF_t)' (H^{-1})' \alpha = \underbrace{c(\Sigma_{W\tilde{F}}V\Sigma_{\tilde{F}}V^{-1})}_{\equiv B_\beta} p \lim(\hat{\alpha}) + o_P(1), \quad (16)$$

where  $\Sigma_{W\tilde{F}} = p \lim\left(\frac{W'\tilde{F}}{T}\right) = \Sigma_{WF}H_0'$ , with  $\Sigma_{WF} \equiv E(W_t F_t')$ .

**Proof of Theorem 2.1.** Write

$$\sqrt{T}(\hat{\delta} - \delta) = \left(\frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t \hat{z}_t'\right)^{-1} \left\{ \underbrace{\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \begin{pmatrix} HF_t \\ W_t \end{pmatrix} \varepsilon_{t+h}}_{\equiv A} + \underbrace{\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \begin{pmatrix} \tilde{F}_t - HF_t \\ 0 \end{pmatrix} \varepsilon_{t+h}}_{\equiv B} - \underbrace{\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \hat{z}_t (\tilde{F}_t - HF_t)' (H^{-1})' \alpha}_{\equiv C} \right\}. \quad (17)$$

By Assumption 5 and given the definition of  $\Phi_0 = \text{diag}(p \lim H, I_q)$ ,

$$A = \begin{pmatrix} H & 0 \\ 0 & I_q \end{pmatrix} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} z_t \varepsilon_{t+h} \rightarrow^d N(0, \Phi_0 \Omega \Phi_0').$$

By Lemma A.1,  $B \rightarrow^P 0$  and by Lemma A.2d),

$$C = -\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \begin{pmatrix} \tilde{F}_t \\ W_t \end{pmatrix} (\tilde{F}_t - HF_t)' (H^{-1})' \alpha = -c \begin{pmatrix} B_\alpha \\ B_\beta \end{pmatrix} + o_P(1),$$

where  $B_\alpha$  and  $B_\beta$  are as defined in (15) and (16). Given Assumptions 1-5,

$$\frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t \hat{z}'_t = \Phi_0 \left( \frac{1}{T} \sum_{t=1}^{T-h} z_t z'_t \right) \Phi'_0 + o_P(1) = \Phi_0 \Sigma_{zz} \Phi'_0 + o_P(1).$$

This implies that if  $\sqrt{T}/N \rightarrow c$ ,  $\sqrt{T}(\hat{\delta} - \delta) \rightarrow^d N(-c\Delta_\delta, \Sigma_\delta)$ , with  $\Sigma_\delta = \Phi_0'^{-1} \Sigma_{zz}^{-1} \Omega \Sigma_{zz}^{-1} \Phi_0^{-1}$ , and

$$\Delta_\delta \equiv \begin{pmatrix} \Delta_\alpha \\ \Delta_\beta \end{pmatrix} = (\Phi_0 \Sigma_{zz} \Phi'_0)^{-1} \begin{pmatrix} \Sigma_{\tilde{F}} + V \Sigma_{\tilde{F}} V^{-1} \\ \Sigma_{W\tilde{F}} V \Sigma_{\tilde{F}} V^{-1} \end{pmatrix} p \lim(\hat{\alpha}).$$

When  $\Sigma_{WF} \equiv E(W_t F'_t) = 0$ , we have that  $\Sigma_{W\tilde{F}} = 0$ , implying that

$$\begin{aligned} \Delta_\delta &= \begin{pmatrix} H_0'^{-1} \Sigma_F^{-1} H_0^{-1} & 0 \\ 0 & \Sigma_W^{-1} \end{pmatrix} \begin{pmatrix} \Sigma_{\tilde{F}} + V \Sigma_{\tilde{F}} V^{-1} \\ 0 \times V \Sigma_{\tilde{F}} V^{-1} \end{pmatrix} p \lim(\hat{\alpha}) \\ &= \begin{pmatrix} \Sigma_{\tilde{F}} + V \Sigma_{\tilde{F}} V^{-1} \\ 0 \end{pmatrix} p \lim(\hat{\alpha}), \end{aligned}$$

since  $H_0'^{-1} \Sigma_F^{-1} H_0^{-1} = I_r$  given that we can show that  $H_0 \Sigma_F = Q = (H_0')^{-1}$ .

**Proof of Lemma A.1.** The proof is based on the following identity:

$$\begin{aligned} \tilde{F}_t - HF_t &= \tilde{V}^{-1} \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \gamma_{st} + \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \zeta_{st} + \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \eta_{st} + \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \xi_{st} \right) \\ &\equiv \tilde{V}^{-1} (A_{1t} + A_{2t} + A_{3t} + A_{4t}), \end{aligned} \tag{18}$$

where

$$\begin{aligned} \gamma_{st} &= E \left( \frac{1}{N} \sum_{i=1}^N e_{is} e_{it} \right), \quad \zeta_{st} = \frac{1}{N} \sum_{i=1}^N (e_{is} e_{it} - E(e_{is} e_{it})), \\ \eta_{st} &= \frac{1}{N} \sum_{i=1}^N \lambda'_i F_s e_{it} = F'_s \frac{\Lambda' e_t}{N} \text{ and } \xi_{st} = F'_t \frac{\Lambda' e_s}{N} = \eta_{ts}. \end{aligned}$$

Using the identity (18), we have that

$$\frac{1}{T} \sum_{t=1}^{T-h} (\tilde{F}_t - HF_t) \varepsilon_{t+h} = \tilde{V}^{-1} (I + II + III + IV),$$

where

$$\begin{aligned} I &= \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T \tilde{F}_s \gamma_{st} \varepsilon_{t+h}, \quad II = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T \tilde{F}_s \zeta_{st} \varepsilon_{t+h}, \\ III &= \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T \tilde{F}_s \eta_{st} \varepsilon_{t+h}, \quad \text{and } IV = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T \tilde{F}_s \xi_{st} \varepsilon_{t+h}. \end{aligned}$$

Since  $\tilde{V}^{-1} = O_P(1)$  (see Lemma A.3 of Bai (2003), which shows that  $\tilde{V} \rightarrow^P V > 0$ ), we can ignore  $\tilde{V}^{-1}$ . Start with  $I$ . We can write

$$I = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T (\tilde{F}_s - HF_s) \gamma_{st} \varepsilon_{t+h} + H \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \gamma_{st} \varepsilon_{t+h} \equiv I_1 + I_2.$$

We can show that  $I_2 = O_P\left(\frac{1}{T}\right)$  by showing that  $E\|I_2\| = O\left(\frac{1}{T}\right)$ . Ignoring  $H$  (which is  $O_P(1)$ ), we have

$$\begin{aligned} E\|I_2\| &\leq \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T E\|F_s \gamma_{st} \varepsilon_{t+h}\| \leq \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T |\gamma_{st}| \left(E\|F_s\|^2\right)^{1/2} \left(E|\varepsilon_{t+h}|^2\right)^{1/2} \\ &\leq \Delta \frac{1}{T} \left(\frac{1}{T} \sum_{t,s} |\gamma_{st}|\right) = O\left(\frac{1}{T}\right), \end{aligned}$$

since  $E\|F_s\|^2 \leq M$  and  $E|\varepsilon_{t+h}|^2 \leq M$  for some constant  $M < \infty$ , and  $\frac{1}{T} \sum_{t,s} |\gamma_{st}| < M$  by assumption. For  $I_1$ , repeated application of the Cauchy-Schwartz inequality implies that

$$\begin{aligned} \|I_1\| &= \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \left( \frac{1}{T} \sum_{t=1}^{T-h} \gamma_{st} \varepsilon_{t+h} \right) \right\| \leq \left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^T \left| \frac{1}{T} \sum_{t=1}^{T-h} \gamma_{st} \varepsilon_{t+h} \right|^2 \right)^{1/2} \\ &\leq \frac{1}{\sqrt{T}} \left( \underbrace{\frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2}_{O_P(\delta_{NT}^{-2}) \text{ by Lemma A.1 of Bai (2003)}} \right)^{1/2} \left( \underbrace{\frac{1}{T} \sum_{s=1}^T \sum_{t=1}^{T-h} |\gamma_{st}|^2}_{O_P(1)} \underbrace{\frac{1}{T} \sum_{t=1}^{T-h} \varepsilon_{t+h}^2}_{O_P(1)} \right)^{1/2} = O_P\left(\frac{1}{\sqrt{T}\delta_{NT}}\right). \end{aligned}$$

Thus,  $I = O_P\left(\frac{1}{\sqrt{T}\delta_{NT}}\right)$ . Next, consider  $II$ . We have that

$$II = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T (\tilde{F}_s - HF_s) \zeta_{st} \varepsilon_{t+h} + H \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \zeta_{st} \varepsilon_{t+h} \equiv II_1 + II_2.$$

We can show that

$$\|II_1\| \leq \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \right)^{1/2}}_{O_P(\delta_{NT}^{-1})} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left| \frac{1}{T} \sum_{t=1}^{T-h} \zeta_{st} \varepsilon_{t+h} \right|^2 \right)^{1/2}}_{O_P\left(\frac{1}{\sqrt{NT}}\right)} = O_P\left(\frac{1}{\sqrt{NT}\delta_{NT}}\right).$$

Indeed,

$$\begin{aligned} \frac{1}{T} \sum_{s=1}^T E \left| \frac{1}{T} \sum_{t=1}^{T-h} \zeta_{st} \varepsilon_{t+h} \right|^2 &= \frac{1}{T} \sum_{s=1}^T E \left| \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{N} \sum_{i=1}^N (e_{is} e_{it} - E(e_{is} e_{it})) \right) \varepsilon_{t+h} \right|^2 \\ &= \frac{1}{TN} \frac{1}{T} \sum_{s=1}^T E \underbrace{\left| \frac{1}{\sqrt{TN}} \sum_{t=1}^{T-h} \sum_{i=1}^N (e_{is} e_{it} - E(e_{is} e_{it})) \varepsilon_{t+h} \right|^2}_{=O(1) \text{ by Assumption 4(a)}} = O\left(\frac{1}{TN}\right). \end{aligned}$$

For  $II_2$ , ignoring  $H$  (which is  $O_P(1)$ ), we have that

$$II_2 = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \zeta_{st} \varepsilon_{t+h} = \frac{1}{\sqrt{TN}} \frac{1}{T} \sum_{t=1}^{T-h} \underbrace{\left( \frac{1}{\sqrt{TN}} \sum_{s=1}^T \sum_{i=1}^N F_s (e_{it} e_{is} - E(e_{it} e_{is})) \right)}_{\equiv m_t} \varepsilon_{t+h} \equiv \frac{1}{\sqrt{TN}} \frac{1}{T} \sum_{t=1}^{T-h} m_t \varepsilon_{t+h}.$$

We can show that  $\frac{1}{T} \sum_{t=1}^{T-h} m_t \varepsilon_{t+h} = O_P(1)$ , implying that  $II_2 = O_P\left(\frac{1}{\sqrt{NT}}\right)$ . By Cauchy-Schwartz inequality, we have that

$$\frac{1}{T} \sum_{t=1}^{T-h} m_t \varepsilon_{t+h} \leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \|m_t\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \varepsilon_{t+h}^2 \right)^{1/2} = O_P(1),$$

given that  $E|\varepsilon_{t+h}|^2 < M < \infty$ , provided  $\frac{1}{T} \sum_{t=1}^{T-h} \|m_t\|^2 = O_P(1)$ , or  $\frac{1}{T} \sum_{t=1}^{T-h} E\|m_t\|^2 = O(1)$ , by Markov's inequality. But

$$\frac{1}{T} \sum_{t=1}^{T-h} E\|m_t\|^2 \leq \frac{1}{T} \sum_{t=1}^T E \left\| \frac{1}{\sqrt{TN}} \sum_{s=1}^T \sum_{i=1}^N F_s (e_{it} e_{is} - E(e_{it} e_{is})) \right\|^2 = O(1)$$

by Assumption 3(b). Thus,  $II = O_P\left(\frac{1}{\sqrt{NT}}\right)$ . Next, consider

$$III = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T \tilde{F}_s \eta_{st} \varepsilon_{t+h} = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T (\tilde{F}_s - HF_s) \eta_{st} \varepsilon_{t+h} + H \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \eta_{st} \varepsilon_{t+h} \equiv III_1 + III_2.$$

Starting with  $III_1$ , we have that

$$\|III_1\| \leq \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \right)^{1/2}}_{O_P(\delta_{NT}^{-1})} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left| \frac{1}{T} \sum_{t=1}^{T-h} \eta_{st} \varepsilon_{t+h} \right|^2 \right)^{1/2}}_{O_P\left(\frac{1}{\sqrt{TN}}\right)} = O_P\left(\frac{1}{\sqrt{TN} \delta_{NT}}\right)$$

since

$$\begin{aligned} \frac{1}{T} \sum_{s=1}^T E \left| \frac{1}{T} \sum_{t=1}^{T-h} \eta_{st} \varepsilon_{t+h} \right|^2 &= \frac{1}{T} \sum_{s=1}^T E \left| \frac{1}{T} \sum_{t=1}^{T-h} \left( F_s' \frac{\Lambda' e_t}{N} \right) \varepsilon_{t+h} \right|^2 = \frac{1}{T} \sum_{s=1}^T E \left| F_s' \frac{1}{TN} \sum_{t=1}^{T-h} \Lambda' e_t \varepsilon_{t+h} \right|^2 \\ &\leq \underbrace{\frac{1}{T} \sum_{s=1}^T \|F_s\|^2}_{=O_P(1)} \cdot \underbrace{\frac{1}{TN} E \left\| \frac{1}{\sqrt{TN}} \sum_{t=1}^{T-h} \Lambda' e_t \varepsilon_{t+h} \right\|^2}_{=O(1) \text{ by Assumption 4(b)}} = O_P\left(\frac{1}{TN}\right). \end{aligned}$$

Ignoring again  $H$  and replacing  $\eta_{st}$  with its definition, we have that

$$III_2 = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \eta_{st} \varepsilon_{t+h} = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \left( F_s' \frac{\Lambda' e_t}{N} \right) \varepsilon_{t+h} = \underbrace{\left( \frac{1}{T} \sum_{s=1}^T F_s F_s' \right)}_{O_P(1)} \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \frac{\Lambda' e_t}{N} \varepsilon_{t+h} \right)}_{O_P\left(\frac{1}{\sqrt{TN}}\right)} = O_P\left(\frac{1}{\sqrt{TN}}\right),$$

given Assumption 4(b). Thus,  $III = O_P\left(\frac{1}{\sqrt{TN}}\right)$ . For  $IV$ , write

$$IV = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T (\tilde{F}_s - HF_s) \xi_{st} \varepsilon_{t+h} + H \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \xi_{st} \varepsilon_{t+h} \equiv IV_1 + IV_2.$$

For  $IV_1$ , we have that

$$\|IV_1\| \leq \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \right)^{1/2}}_{O_P(\delta_{NT}^{-1})} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left( \frac{1}{T} \sum_{t=1}^{T-h} \xi_{st} \varepsilon_{t+h} \right)^2 \right)^{1/2}}_{O_P\left(\frac{1}{\sqrt{NT}}\right)} = O_P\left(\frac{1}{\delta_{NT}\sqrt{NT}}\right).$$

Indeed,

$$\begin{aligned} & \frac{1}{T} \sum_{s=1}^T \left( \frac{1}{T} \sum_{t=1}^{T-h} \xi_{st} \varepsilon_{t+h} \right)^2 = \frac{1}{T} \sum_{s=1}^T \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t' \left( \frac{\Lambda' e_s}{N} \right) \varepsilon_{t+h} \right)^2 \\ &= \frac{1}{T} \sum_{s=1}^T \left( \left( \frac{e_s' \Lambda}{N} \right) \frac{1}{T} \sum_{t=1}^{T-h} F_t \varepsilon_{t+h} \right)^2 \leq \frac{1}{T} \sum_{s=1}^T \left\| \frac{e_s' \Lambda}{N} \right\|^2 \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \varepsilon_{t+h} \right\|^2 \\ &= \frac{1}{TN} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left\| \frac{e_s' \Lambda}{\sqrt{N}} \right\|^2 \right)}_{=O_P(1) \text{ by Assumption 3(d)}} \underbrace{\left( \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} F_t \varepsilon_{t+h} \right\|^2 \right)}_{=O_P(1) \text{ by Assumption 5(c)}} = O_P\left(\frac{1}{TN}\right). \end{aligned}$$

For  $IV_2$ , we have that

$$\begin{aligned} IV_2 &= \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \left( F_t' \frac{\Lambda' e_s}{N} \right) \varepsilon_{t+h} \\ &= \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T F_s \left( \frac{e_s' \Lambda}{N} F_t \right) \varepsilon_{t+h} = \left( \frac{1}{T} \sum_{s=1}^T F_s \frac{e_s' \Lambda}{N} \right) \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t \varepsilon_{t+h} \right) = O_P\left(\frac{1}{\sqrt{NT}}\right) \times O_P\left(\frac{1}{\sqrt{T}}\right), \end{aligned}$$

given Assumptions 3(c) and 5(b). Thus, we have shown that under our assumptions,

$$I+II+III+IV = O_P\left(\frac{1}{\delta_{NT}\sqrt{T}}\right) + O_P\left(\frac{1}{\sqrt{NT}}\right) + O_P\left(\frac{1}{\sqrt{NT}}\right) + O_P\left(\frac{1}{\delta_{NT}\sqrt{NT}}\right) = O_P\left(\frac{1}{\delta_{NT}\sqrt{T}}\right).$$

**Proof of Lemma A.2. Proof of part a)** Using the identity (18), we can write

$$\begin{aligned} B &\equiv \sqrt{T} \frac{1}{T} \sum_{t=1}^{T-h} \left( \tilde{F}_t - HF_t \right) \left( \tilde{F}_t - HF_t \right)' \\ &= \sqrt{T} \tilde{V}^{-1} \frac{1}{T} \sum_{t=1}^{T-h} (A_{1t} + A_{2t} + A_{3t} + A_{4t}) (A_{1t} + A_{2t} + A_{3t} + A_{4t})' \tilde{V}^{-1}, \end{aligned}$$

where  $A_{it}$  ( $i = 1, \dots, 4$ ) are defined in (18). We analyze each term separately (ignoring  $\tilde{V}^{-1}$ ).

- We can show that  $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{1t}' = O_P\left(\frac{1}{T\delta_{NT}^2}\right) + O_P\left(\frac{1}{T}\right)$ . Thus,

$$\sqrt{T} \frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{1t}' = O_P\left(\frac{1}{\sqrt{T}\delta_{NT}^2}\right) + O_P\left(\frac{1}{\sqrt{T}}\right) = o_P(1).$$



Indeed, by Cauchy-Schwartz

$$\begin{aligned} \left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A'_{1t} \right\| &\leq \frac{1}{T} \sum_{t=1}^{T-h} \|A_{1t}\|^2 = \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \gamma_{st} \right\|^2 = \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s + HF_s) \gamma_{st} \right\|^2 \\ &\leq \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \gamma_{st} \right\|^2 + \frac{1}{T} \sum_{t=1}^{T-h} \left\| H \frac{1}{T} \sum_{s=1}^T F_s \gamma_{st} \right\|^2 \equiv a_{11.1} + a_{11.2}. \end{aligned}$$

We have that

$$\begin{aligned} a_{11.1} &\equiv \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \gamma_{st} \right\|^2 \leq \frac{1}{T} \sum_{t=1}^{T-h} \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \frac{1}{T} \sum_{s=1}^T |\gamma_{st}|^2 \\ &= \frac{1}{T} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \right)}_{=O_P(\delta_{NT}^{-2})} \underbrace{\left( \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T |\gamma_{st}|^2 \right)}_{=O_P(1)} = O_P\left(\frac{1}{T\delta_{NT}^2}\right), \end{aligned}$$

where we have use Assumption 2 (see Lemma 1(i) of Bai and Ng (2002)) to bound  $\frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T |\gamma_{st}|^2$ .

Next,

$$\begin{aligned} a_{11.2} &\equiv \frac{1}{T} \sum_{t=1}^{T-h} \left\| H \frac{1}{T} \sum_{s=1}^T F_s \gamma_{st} \right\|^2 \leq \|H\|^2 \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T F_s \gamma_{st} \right\|^2 \\ &\leq \underbrace{\|H\|^2}_{=O_P(1)} \frac{1}{T} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|F_s\|^2 \right)}_{=O_P(1)} \underbrace{\frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T |\gamma_{st}|^2}_{=O_P(1)} = O_P\left(\frac{1}{T}\right). \end{aligned}$$

- We can show that  $\frac{1}{T} \sum_{t=1}^{T-h} A_{2t} A'_{2t} = O_P\left(\frac{1}{TN}\right) + O_P\left(\frac{1}{N\delta_{NT}^2}\right)$ , implying that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} A_{2t} A'_{2t} = O_P\left(\frac{1}{\sqrt{TN}}\right) + O_P\left(\frac{\sqrt{T}}{N} \frac{1}{\delta_{NT}^2}\right) = o_P(1),$$

if  $\sqrt{T}/N \rightarrow c < \infty$ .

Indeed,

$$\begin{aligned} \left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{2t} A'_{2t} \right\| &\leq \frac{1}{T} \sum_{t=1}^{T-h} \|A_{2t}\|^2 = \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s + HF_s) \zeta_{st} \right\|^2 \\ &\leq \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \zeta_{st} \right\|^2 + \frac{1}{T} \sum_{t=1}^{T-h} \left\| H \frac{1}{T} \sum_{s=1}^T F_s \zeta_{st} \right\|^2 \equiv a_{22.1} + a_{22.2}. \end{aligned}$$

We have that

$$a_{22.1} \equiv \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \zeta_{st} \right\|^2 \leq \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \right)}_{=O_P(\delta_{NT}^{-2})} \underbrace{\left( \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T |\zeta_{st}|^2 \right)}_{=O_P(1/N) \text{ by Assumption 2(c)}} = O_P\left(\frac{1}{N\delta_{NT}^2}\right).$$

Indeed,

$$\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T E |\zeta_{st}|^2 = \frac{1}{N} \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T E \left| \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{is} e_{it} - E(e_{is} e_{it})) \right|^2 = O\left(\frac{1}{N}\right)$$

given Assumption 2(c). Next,

$$a_{22.2} \leq \|H\|^2 \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T F_s \zeta_{st} \right\|^2 = \|H\|^2 \frac{1}{TN} \frac{1}{T} \sum_{t=1}^{T-h} \underbrace{\left\| \frac{1}{\sqrt{TN}} \sum_{s=1}^T \sum_{i=1}^N F_s (e_{is} e_{it} - E(e_{is} e_{it})) \right\|^2}_{=O_P(1) \text{ by Assumption 3(b)}} = O_P\left(\frac{1}{TN}\right),$$

implying that

$$\left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{2t} A'_{2t} \right\| = O_P\left(\frac{1}{N\delta_{NT}^2}\right) + O_P\left(\frac{1}{TN}\right) = O_P\left(\frac{1}{N\delta_{NT}^2}\right).$$

- We can show that under Assumptions 1-5, if  $\sqrt{T}/N \rightarrow c$ ,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} A_{3t} A'_{3t} = cQ\Gamma Q' + o_P(1).$$

*Proof.*

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} A_{3t} A'_{3t} &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \eta_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \eta_{st} \right)' \\ &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s + HF_s) \eta_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s + HF_s) \eta_{st} \right)' \\ &\equiv a_{33.1} + a_{33.2} + a'_{33.2} + a_{33.3}. \end{aligned}$$

It follows that

$$\|a_{33.1}\| \leq \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \eta_{st} \right\|^2 \leq \underbrace{\frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2}_{=O_P\left(\frac{1}{\delta_{NT}^2}\right)} \underbrace{\frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T |\eta_{st}|^2}_{=O_P\left(\frac{1}{N}\right)} = O_P\left(\frac{1}{N\delta_{NT}^2}\right),$$

since

$$\frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T |\eta_{st}|^2 = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T \left| F'_s \frac{\Lambda' e_t}{N} \right|^2 \leq \underbrace{\frac{1}{T} \sum_{s=1}^T \|F_s\|^2}_{=O_P(1)} \underbrace{\frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{\Lambda' e_t}{N} \right\|^2}_{=O_P\left(\frac{1}{N}\right) \text{ given Assumption 3(d)}} = O_P\left(\frac{1}{N}\right).$$

Thus,

$$\sqrt{T} a_{33.1} = O_P\left(\frac{\sqrt{T}}{N} \frac{1}{\delta_{NT}^2}\right) = o_P(1),$$

if  $\sqrt{T}/N \rightarrow c$ . Next we show that  $\sqrt{T}a_{33.2} = O_P\left(\frac{\sqrt{T}}{N\delta_{NT}}\right) = o_P(1)$ . Indeed,

$$\|a_{33.2}\| \leq \underbrace{\left(\frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T HF_s \eta_{st} \right\|^2\right)^{1/2}}_{=O_P\left(\frac{1}{\sqrt{N}}\right)} \underbrace{\left(\frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \eta_{st} \right\|^2\right)^{1/2}}_{=O_P\left(\frac{1}{\delta_{NT}\sqrt{N}}\right)} = O_P\left(\frac{1}{N\delta_{NT}}\right),$$

since by Cauchy-Schwartz and Assumption 3(d),

$$\frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T HF_s \eta_{st} \right\|^2 \leq \|H\|^2 \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T F_s \eta_{st} \right\|^2 \leq \|H\|^2 \underbrace{\frac{1}{T} \sum_{s=1}^T \|F_s\|^2}_{=O_P(1)} \underbrace{\frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T |\eta_{st}|^2}_{=O_P\left(\frac{1}{N}\right)} = O_P\left(\frac{1}{N}\right).$$

For  $a_{33.3}$ , we have that

$$\begin{aligned} a_{33.3} &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T HF_s \eta_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T HF_s \eta_{st} \right)' = H \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T F_s \eta_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T \eta_{st} F_s' \right) H' \\ &= H \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T F_s F_s' \frac{\Lambda' e_t}{N} \right) \left( \frac{1}{T} \sum_{s=1}^T \frac{e_t' \Lambda}{N} F_s F_s' \right) H' \\ &= H \left( \frac{F' F}{T} \right) \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\Lambda' e_t}{N} \right) \left( \frac{e_t' \Lambda}{N} \right) \left( \frac{F' F}{T} \right) H' = O_P\left(\frac{1}{N}\right), \end{aligned}$$

given Assumption 3(e). When multiplied by  $\sqrt{T}$ , this will therefore be of order  $O_P\left(\frac{\sqrt{T}}{N}\right)$  and therefore it will not go to zero in probability when  $c \neq 0$ . Its probability limit can be computed as follows. First, notice that

$$\frac{HF'F}{T} = \frac{(\tilde{F} - FH' + \tilde{F})' F}{T} = \frac{(\tilde{F} - FH')' F}{T} + \frac{\tilde{F}' F}{T} = Q + o_P(1)$$

given that by Lemma B.2 of Bai (2003),

$$\frac{(\tilde{F} - FH')' F}{T} = O_P(\delta_{NT}^{-2}) = o_P(1),$$

and  $p \lim \frac{\tilde{F}' F}{T} = Q$ . Second, by Assumption 3(e),

$$\frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\Lambda' e_t}{\sqrt{N}} \right) \left( \frac{e_t' \Lambda}{\sqrt{N}} \right) = \Gamma + o_P(1).$$

Thus, letting  $\sqrt{T}/N = c + o(1)$ ,

$$\begin{aligned} \sqrt{T}a_{33.3} &= \frac{\sqrt{T}}{N} H \left( \frac{F' F}{T} \right) \left\{ \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\Lambda' e_t}{\sqrt{N}} \right) \left( \frac{e_t' \Lambda}{\sqrt{N}} \right) \right\} \left( \frac{F' F}{T} \right) H' \\ &= (c + o(1)) [Q + o_P(1)] (\Gamma + o_P(1)) [Q' + o_P(1)] \\ &= cQ\Gamma Q' + o_P(1). \end{aligned}$$

- We can show that  $\frac{1}{T} \sum_{t=1}^{T-h} A_{4t} A'_{4t} = O_P\left(\frac{1}{N\delta_{NT}^2}\right) + O_P\left(\frac{1}{N\delta_{NT}}\right) + O_P\left(\frac{1}{TN}\right)$ , which implies that if  $\sqrt{T}/N \rightarrow c$ ,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} A_{4t} A'_{4t} = O_P\left(\frac{\sqrt{T}}{N\delta_{NT}^2}\right) + O_P\left(\frac{\sqrt{T}}{N\delta_{NT}}\right) + O_P\left(\frac{1}{\sqrt{T}N}\right) = o_P(1).$$

*Proof.*

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} A_{4t} A'_{4t} &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \xi_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \xi_{st} \right)' \\ &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s + HF_s) \xi_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s + HF_s) \xi_{st} \right)' \\ &\equiv a_{44.1} + a_{44.2} + a'_{44.2} + a_{44.3}. \end{aligned}$$

We have that

$$\|a_{44.1}\| \leq \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \xi_{st} \right\|^2 \leq \underbrace{\frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2}_{=O_P\left(\frac{1}{\delta_{NT}^2}\right)} \underbrace{\frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T |\xi_{st}|^2}_{=O_P\left(\frac{1}{N}\right)} = O_P\left(\frac{1}{N\delta_{NT}^2}\right),$$

since

$$\frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T |\xi_{st}|^2 = \frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T \left| F'_t \frac{\Lambda' e_s}{N} \right|^2 \leq \underbrace{\frac{1}{T} \sum_{t=1}^{T-r} \|F_t\|^2}_{=O_P(1)} \underbrace{\frac{1}{T} \sum_{s=1}^T \left\| \frac{\Lambda' e_s}{N} \right\|^2}_{=O_P\left(\frac{1}{N}\right) \text{ given Assumption 3(d)}} = O_P\left(\frac{1}{N}\right).$$

This implies that  $\sqrt{T}a_{44.1} = O_P\left(\sqrt{T}/N\right) O_P\left(\frac{1}{\delta_{NT}^2}\right) = o_P(1)$ . Next,

$$\|a_{44.2}\| \leq \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T HF_s \xi_{st} \right\|^2 \right)^{1/2}}_{=O_P\left(\frac{1}{\sqrt{N}}\right)} \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \xi_{st} \right\|^2 \right)^{1/2}}_{=O_P\left(\frac{1}{\delta_{NT}\sqrt{N}}\right)} = O_P\left(\frac{1}{N\delta_{NT}}\right),$$

since by Cauchy-Schwartz and Assumption 3(d),

$$\frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T HF_s \xi_{st} \right\|^2 \leq \|H\|^2 \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right\|^2 \leq \|H\|^2 \underbrace{\frac{1}{T} \sum_{s=1}^T \|F_s\|^2}_{=O_P(1)} \underbrace{\frac{1}{T^2} \sum_{t=1}^{T-h} \sum_{s=1}^T |\xi_{st}|^2}_{=O_P\left(\frac{1}{N}\right)} = O_P\left(\frac{1}{N}\right).$$

Thus,

$$\sqrt{T}a_{44.2} = O_P\left(\frac{\sqrt{T}}{N}\right) O_P\left(\frac{1}{\delta_{NT}}\right) = o_P(1),$$

when  $\sqrt{T}/N \rightarrow c$ . For  $a_{44.3}$ , we have that

$$\begin{aligned}
a_{44.3} &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T H F_s \xi_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T H F_s \xi_{st} \right)' = H \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T \xi_{st} F_s' \right) H' \\
&= H \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T F_s F_t' \frac{\Lambda' e_s}{N} \right) \left( \frac{1}{T} \sum_{s=1}^T \frac{e_s' \Lambda}{N} F_t F_s' \right) H' \\
&= H \underbrace{\left( \frac{1}{T} \sum_{s=1}^T F_s \frac{e_s' \Lambda}{N} \right)}_{=O_P\left(\frac{1}{\sqrt{TN}}\right)} \underbrace{\frac{1}{T} \sum_{t=1}^{T-h} F_t F_t'}_{=O_P(1)} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} F_s' \right)}_{=O_P\left(\frac{1}{\sqrt{TN}}\right) \text{ by Assumption 3(c)}} H' = O_P\left(\frac{1}{TN}\right).
\end{aligned}$$

Thus,  $\sqrt{T} a_{44.3} = O_P\left(\frac{1}{\sqrt{TN}}\right) = o_P(1)$ .

- We can show that  $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{2t}' = O_P\left(\frac{1}{\delta_{NT} \sqrt{TN}}\right)$ , which implies that  $\sqrt{T} \frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{2t}' = o_P(1)$ .

*Proof.* Given the bounds we found above for the terms that depend on  $A_{1t}$  and  $A_{2t}$ , it is immediate to see that

$$\left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{2t}' \right\| \leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{1t}\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{2t}\|^2 \right)^{1/2} = O_P\left(\frac{1}{\sqrt{T}}\right) O_P\left(\frac{1}{\delta_{NT} \sqrt{N}}\right) = O_P\left(\frac{1}{\delta_{NT} \sqrt{TN}}\right)$$

- We can show that  $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{3t}' = O_P\left(\frac{1}{\sqrt{TN}}\right)$ , implying that  $\frac{\sqrt{T}}{T} \sum_{t=1}^{T-h} A_{1t} A_{3t}' = O_P\left(\frac{\sqrt{T}}{\sqrt{TN}}\right) = O_P\left(\frac{1}{\sqrt{N}}\right) = o_P(1)$ . Indeed,

$$\left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{3t}' \right\| \leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{1t}\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{3t}\|^2 \right)^{1/2} = O_P\left(\frac{1}{\sqrt{T}}\right) O_P\left(\frac{1}{\sqrt{N}}\right) = O_P\left(\frac{1}{\sqrt{TN}}\right).$$

- Similarly,  $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{4t}' = O_P\left(\frac{1}{\sqrt{TN}}\right)$ , which implies that  $\sqrt{T} \frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{4t}' = o_P(1)$ .

$$\begin{aligned}
\left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{1t} A_{4t}' \right\| &\leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{1t}\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{4t}\|^2 \right)^{1/2} = O_P\left(\frac{1}{\sqrt{T}}\right) O_P\left(\frac{1}{\sqrt{N}}\right) \\
&= O_P\left(\frac{1}{\sqrt{TN}}\right),
\end{aligned}$$

where we can prove that  $\frac{1}{T} \sum_{t=1}^{T-h} \|A_{4t}\|^2 = O_P\left(\frac{1}{N}\right)$  by an application of Cauchy-Schwartz.

*Proof.* Given the definition of  $A_{4t}$ , we have that

$$\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T-h} \|A_{4t}\|^2 &= \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \xi_{st} \right\|^2 \leq \frac{1}{T} \sum_{t=1}^{T-h} \underbrace{\left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - H F_s) \xi_{st} \right\|^2}_{O_P\left(\frac{1}{N \delta_{NT}^2}\right)} + \frac{1}{T} \sum_{t=1}^{T-h} \underbrace{\left\| H \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right\|^2}_{O_P\left(\frac{1}{N}\right)} \\
&= O_P\left(\frac{1}{N}\right).
\end{aligned}$$

By Cauchy-Schwartz,

$$\frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \xi_{st} \right\|^2 \leq \underbrace{\frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2}_{=O_P(\delta_{NT}^{-2})} \underbrace{\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \xi_{st}^2}_{=O_P(\frac{1}{N})} = O_P\left(\frac{1}{\delta_{NT}^2 N}\right)$$

since

$$\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \xi_{st}^2 = \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \left\| F_t' \frac{\Lambda' e_s}{N} \right\|^2 \leq \frac{1}{T} \sum_{t=1}^T \|F_t\|^2 \frac{1}{T} \sum_{s=1}^T \left\| \frac{\Lambda' e_s}{N} \right\|^2 = O_P(1) O_P\left(\frac{1}{N}\right),$$

by Assumptions 1(a) and 3(d). For the second term,

$$\frac{1}{T} \sum_{t=1}^{T-h} \left\| H \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right\|^2 \leq \|H\|^2 \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right\|^2 \leq \|H\|^2 \frac{1}{T} \sum_{t=1}^{T-h} \frac{1}{T} \sum_{s=1}^T \|F_s\|^2 \frac{1}{T} \sum_{s=1}^T |\xi_{st}|^2 = O_P\left(\frac{1}{N}\right).$$

- Similarly, we can show that  $\frac{1}{T} \sum_{t=1}^{T-h} A_{2t} A'_{3t} = O_P\left(\frac{1}{N\delta_{NT}}\right)$ , implying that  $\frac{\sqrt{T}}{T} \sum_{t=1}^{T-h} A_{2t} A'_{3t} = o_P(1)$ .

$$\left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{2t} A'_{3t} \right\| \leq \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{2t}\|^2 \right)^{1/2}}_{O_P\left(\frac{1}{N\delta_{NT}^2}\right)} \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{3t}\|^2 \right)^{1/2}}_{O_P\left(\frac{1}{N}\right)} = O_P\left(\frac{1}{N\delta_{NT}}\right).$$

- $\frac{1}{T} \sum_{t=1}^{T-h} A_{2t} A'_{4t} = O_P\left(\frac{1}{N\delta_{NT}}\right)$ , implying that  $\frac{\sqrt{T}}{T} \sum_{t=1}^{T-h} A_{2t} A'_{4t} = o_P(1)$ .

$$\left\| \frac{1}{T} \sum_{t=1}^{T-h} A_{2t} A'_{4t} \right\| \leq \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{2t}\|^2 \right)^{1/2}}_{O_P\left(\frac{1}{N\delta_{NT}^2}\right)} \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{4t}\|^2 \right)^{1/2}}_{O_P\left(\frac{1}{N}\right)} = O_P\left(\frac{1}{N\delta_{NT}}\right).$$

- $\frac{1}{T} \sum_{t=1}^{T-h} A_{3t} A'_{4t} = O_P\left(\frac{1}{N\delta_{NT}}\right)$ , implying that  $\frac{\sqrt{T}}{T} \sum_{t=1}^{T-h} A_{3t} A'_{4t} = o_P(1)$ .

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} A_{3t} A'_{4t} &= \frac{1}{T} \sum_{t=1}^{T-h} A_{3t} \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \xi_{st} \right)' = \frac{1}{T} \sum_{t=1}^{T-h} A_{3t} \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \xi_{st} + H \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right)' \\ &= \frac{1}{T} \sum_{t=1}^{T-h} A_{3t} \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s)' \xi_{st} + H \frac{1}{T^2} \sum_{t=1}^{T-h} A_{3t} \left( \sum_{s=1}^T F_s \xi_{st} \right)' \end{aligned}$$

The first term can be bounded by

$$\left( \frac{1}{T} \sum_{t=1}^{T-h} \|A_{3t}\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \xi_{st} \right\|^2 \right)^{1/2} = O_P \left( \frac{1}{\sqrt{N}} \right) O_P \left( \frac{1}{\delta_{NT}\sqrt{N}} \right) = O_P \left( \frac{1}{N\delta_{NT}} \right).$$

For the second term, note that

$$A_{3t} = \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \eta_{st} = \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \eta_{st} + H \frac{1}{T} \sum_{s=1}^T F_s \eta_{st}.$$

Thus, we get that

$$\begin{aligned} \frac{1}{T^2} \sum_{t=1}^{T-h} A_{3t} \sum_{s=1}^T F_s \xi_{st} &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \eta_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right)' \\ &\quad + H \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T F_s \eta_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right)'. \end{aligned}$$

The first term is bounded by

$$\begin{aligned} &\left( \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \eta_{st} \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right\|^2 \right)^{1/2} \\ &= O_P \left( \frac{1}{\delta_{NT}\sqrt{N}} \right) O_P \left( \frac{1}{\sqrt{N}} \right) = O_P \left( \frac{1}{N\delta_{NT}} \right). \end{aligned}$$

For the second term, we get that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T F_s \eta_{st} \right) \left( \frac{1}{T} \sum_{s=1}^T F_s \xi_{st} \right)' &= \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{T} \sum_{s=1}^T F_s \left( F_s' \frac{\Lambda' e_t}{N} \right) \right) \left( \frac{1}{T} \sum_{s=1}^T F_s F_t' \frac{\Lambda' e_s}{N} \right)' \\ &= \frac{1}{T} \sum_{t=1}^{T-h} \frac{F' F}{T} \left( \frac{\Lambda' e_t}{N} \right) \left( \frac{1}{T} \sum_{s=1}^T F_s F_t' \frac{\Lambda' e_s}{N} \right)' = \frac{F' F}{T} \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\Lambda' e_t}{N} \right) \left( \frac{1}{T} \sum_{s=1}^T F_s \frac{e'_s \Lambda}{N} F_t \right)' \\ &= \frac{F' F}{T} \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\Lambda' e_t}{N} F_t' \right) \left( \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} F_s' \right) = O_P(1) O_P \left( \frac{1}{\sqrt{TN}} \right) O_P \left( \frac{1}{\sqrt{TN}} \right) = O_P \left( \frac{1}{TN} \right), \end{aligned}$$

given Assumption 3(c).

Thus, it follows that the only term that contributes in a non-negligible way to

$$\sqrt{T}B = \tilde{V}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (A_{1t} + A_{2t} + A_{3t} + A_{4t}) (A_{1t} + A_{2t} + A_{3t} + A_{4t})' \tilde{V}^{-1}$$

is the term that depends on  $\frac{1}{T} \sum_{t=1}^{T-h} A_{3t} A_{3t}'$ . More precisely, since  $\tilde{V} \xrightarrow{P} V$  (by Lemma A.3 of Bai (2003)), we have that

$$\sqrt{T}B = cV^{-1}Q\Gamma Q'V^{-1} + o_P(1),$$

which proves the result.

**Proof of part b).** Consider now

$$C \equiv H \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} F_t \left( \tilde{F}_t - HF_t \right)'.$$

Replacing  $\tilde{F}_t - HF_t = \tilde{V}^{-1} (A_{1t} + A_{2t} + A_{3t} + A_{4t})$  yields

$$C = H \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} F_t (A_{1t} + A_{2t} + A_{3t} + A_{4t})' \tilde{V}^{-1} \equiv \sqrt{T} H (b_{f1} + b_{f2} + b_{f3} + b_{f4}) \tilde{V}^{-1},$$

where  $A_{it}$  are as defined previously. Again, we consider each term separately.

- We can show that  $\frac{1}{T} \sum_{t=1}^{T-h} F_t A'_{1t} = O_P \left( \frac{1}{\delta_{NT} \sqrt{T}} \right)$ , which implies that  $\sqrt{T} H b_{f1} \tilde{V}^{-1} = o_P(1)$  under our assumptions.

*Proof.* Write

$$\begin{aligned} b_{f1} &\equiv \frac{1}{T} \sum_{t=1}^{T-h} F_t A'_{1t} = \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T \left( \tilde{F}_s - HF_s \right)' \gamma_{st} + \frac{1}{T} \sum_{s=1}^T F'_s \gamma_{st} H' \right) \\ &= \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T \left( \tilde{F}_s - HF_s \right)' \gamma_{st} \right) + \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T F'_s \gamma_{st} \right) H' \\ &\equiv b_{f1.1} + b_{f1.2}. \end{aligned}$$

By Cauchy-Schwartz,

$$\|b_{f1.1}\| \leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \|F_t\|^2 \right)^{1/2} \left( \underbrace{\frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T \left( \tilde{F}_s - HF_s \right)' \gamma_{st} \right\|^2}_{O_P \left( \frac{1}{\delta_{NT}^2 T} \right)} \right)^{1/2} = O_P(1) O_P \left( \frac{1}{\delta_{NT} \sqrt{T}} \right),$$

where the  $O_P \left( \frac{1}{\delta_{NT}^2 T} \right)$  term is equal to  $a_{11.1}$  in part a). Next, consider  $b_{f1.2}$ :

$$\begin{aligned} b_{f1.2} &\equiv \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T F'_s \gamma_{st} \right) H' \\ &= \frac{1}{T} \underbrace{\left( \frac{1}{T} \sum_{t=1}^{T-h} \sum_{s=1}^T F_t F'_s \gamma_{st} \right)}_{O_P(1) \text{ given Assumptions 1 and 2}} H' = O_P \left( \frac{1}{T} \right). \end{aligned}$$

Indeed,

$$E \left\| \frac{1}{T} \sum_{t=1}^{T-h} \sum_{s=1}^T F_t F'_s \gamma_{st} \right\| \leq \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T E \|F_t F'_s \gamma_{st}\| \leq \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T |\gamma_{st}| \left( E \|F_t\|^2 \right)^{1/2} \left( E \|F_s\|^2 \right)^{1/2} = O(1),$$



given the moment conditions on  $F_t$  and the weak dependence assumptions on  $e_{it}$ . Thus,

$$b_{f1} \equiv b_{f1.1} + b_{f1.2} = O_P\left(\frac{1}{\delta_{NT}\sqrt{T}}\right) + O_P\left(\frac{1}{T}\right) = O_P\left(\frac{1}{\delta_{NT}\sqrt{T}}\right),$$

as we wanted to prove.

- We can show that  $b_{f2} \equiv \frac{1}{T} \sum_{t=1}^{T-h} F_t A'_{2t} = O_P\left(\frac{1}{\delta_{NT}\sqrt{TN}}\right) + O_P\left(\frac{1}{\sqrt{TN}}\right) = O_P\left(\frac{1}{\sqrt{TN}}\right)$ , which implies that  $\sqrt{T}$  times this object is  $o_P(1)$ .

*Proof.*

$$\begin{aligned} b_{f2} &\equiv \frac{1}{T} \sum_{t=1}^{T-h} F_t A'_{2t} = \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \zeta_{st} \right)' \\ &= \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s)' \zeta_{st} \right) + \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{1}{T} \sum_{s=1}^T F'_s \zeta_{st} H' \equiv b_{f2.1} + b_{f2.2}. \end{aligned}$$

We can write

$$\begin{aligned} \|b_{f2.1}\| &= \left\| \frac{1}{T} \sum_{s=1}^T \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t \zeta_{st} \right) (\tilde{F}_s - HF_s)' \right\| \leq \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \zeta_{st} \right\|^2 \right)^{1/2}}_{O_P\left(\frac{1}{TN}\right)} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left\| \tilde{F}_s - HF_s \right\|^2 \right)^{1/2}}_{O_P\left(\frac{1}{\delta_{NT}^2}\right)} \\ &= O_P\left(\frac{1}{\delta_{NT}\sqrt{TN}}\right), \end{aligned}$$

given that

$$\begin{aligned} \frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \zeta_{st} \right\|^2 &= \frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{1}{N} \sum_{i=1}^N (e_{is} e_{it} - E(e_{is} e_{it})) \right\|^2 \\ &= \frac{1}{TN} \frac{1}{T} \sum_{s=1}^T \underbrace{\left\| \frac{1}{\sqrt{TN}} \sum_{t=1}^{T-h} \sum_{i=1}^N F_t (e_{is} e_{it} - E(e_{is} e_{it})) \right\|^2}_{O_P(1) \text{ by Assumption 3(b)}} = O_P\left(\frac{1}{TN}\right). \end{aligned}$$

For  $b_{f2.2}$ , we have that (ignoring  $H$ ),

$$\|b_{f2.2}\| = \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T F'_s \zeta_{st} \right) \right\| \leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \|F_t\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \left\| \frac{1}{T} \sum_{s=1}^T F'_s \zeta_{st} \right\|^2 \right)^{1/2} = O_P\left(\frac{1}{\sqrt{TN}}\right),$$

using again Assumption 3(b) to bound the second term.

- We can show that  $b_{f3} \equiv \frac{1}{T} \sum_{t=1}^{T-h} F_t A'_{3t} = O_P\left(\frac{1}{\sqrt{TN}}\right)$ , implying that  $\sqrt{T}b_{f3} = O_P\left(\frac{\sqrt{T}}{\sqrt{TN}}\right) = o_P(1)$ .

*Proof.*

$$\begin{aligned}
b_{f3} &\equiv \frac{1}{T} \sum_{t=1}^{T-h} F_t A'_{3t} = \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) \eta_{st} + H \frac{1}{T} \sum_{s=1}^T F_s \eta_{st} \right)' \\
&= \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s)' \eta_{st} + \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( H \frac{1}{T} \sum_{s=1}^T F_s \eta_{st} \right)' \\
&\equiv b_{f3.1} + b_{f3.2}.
\end{aligned}$$

Rewrite

$$b_{f3.1} = \frac{1}{T} \sum_{s=1}^T \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t \eta_{st} \right) (\tilde{F}_s - HF_s)',$$

so that by Cauchy-Schwartz

$$\|b_{f3.1}\| \leq \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \eta_{st} \right\|^2 \right)^{1/2}}_{O_P\left(\frac{1}{\sqrt{TN}}\right)} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s - HF_s\|^2 \right)^{1/2}}_{=O_P\left(\frac{1}{\delta_{NT}}\right)} = O_P\left(\frac{1}{\delta_{NT}\sqrt{TN}}\right).$$

Note that

$$\begin{aligned}
&\frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \eta_{st} \right\|^2 = \frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t F_s' \frac{\Lambda' e_t}{N} \right\|^2 = \frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{e_t' \Lambda}{N} F_s \right\|^2 \\
&\leq \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{e_t' \Lambda}{N} \right\|^2 \frac{1}{T} \sum_{s=1}^T \|F_s\|^2 = \frac{1}{TN} \underbrace{\left\| \frac{1}{\sqrt{TN}} \sum_{t=1}^{T-h} F_t e_t' \Lambda \right\|^2}_{O_P(1) \text{ by Assumption 3(c)}} \frac{1}{T} \sum_{s=1}^T \|F_s\|^2 = O_P\left(\frac{1}{TN}\right)
\end{aligned}$$

Next consider (ignoring  $H$ ),

$$b_{f3.2} = \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{1}{T} \sum_{s=1}^T F_s' \eta_{st} = \frac{1}{T} \sum_{s=1}^T \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t \eta_{st} \right) F_s',$$

so that by Cauchy-Schwartz inequality and Assumption 3(c), we get that

$$\|b_{f3.2}\| \leq \left( \frac{1}{T} \sum_{s=1}^T \left\| \frac{1}{T} \sum_{t=1}^{T-h} F_t \eta_{st} \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^T \|F_s\|^2 \right)^{1/2} = O_P\left(\frac{1}{\sqrt{TN}}\right).$$

This concludes the proof that

$$b_{f3} \equiv b_{f3.1} + b_{f3.2} = O_P\left(\frac{1}{\sqrt{TN}}\right).$$

- We can show that  $\sqrt{T}b_{f4} \equiv \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} F_t A'_{4t} = c\Sigma_F \Gamma Q' V^{-1} + o_P(1)$ .

*Proof.* Replacing  $A_{4t}$  with its definition yields

$$\begin{aligned} b_{f4} &\equiv \frac{1}{T} \sum_{t=1}^{T-h} F_t A'_{4t} = \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}'_s \xi_{st} \right) \\ &= \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s)' \xi_{st} \right) + \frac{1}{T} \sum_{t=1}^{T-h} F_t \left( \frac{1}{T} \sum_{s=1}^T (HF_s)' \xi_{st} \right) \equiv b_{f4.1} + b_{f4.2}. \end{aligned}$$

Starting with  $b_{f4.2}$ , and given that  $\xi_{st} = F'_t \frac{\Lambda' e_s}{N}$ , we have that

$$\begin{aligned} b_{f4.2} &= \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{1}{T} \sum_{s=1}^T F'_t \frac{\Lambda' e_s}{N} F'_s H' = \frac{1}{\sqrt{TN}} \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t F'_t \right) \left( \frac{1}{\sqrt{TN}} \sum_{s=1}^T \Lambda' e_s F'_s \right) H' \\ &= O\left(\frac{1}{\sqrt{TN}}\right) O_P(1) O_P(1) O_P(1) = O_P\left(\frac{1}{\sqrt{TN}}\right), \end{aligned}$$

given Assumptions 1, 3(c) and the fact that  $H = O_P(1)$ . Thus,  $\sqrt{T} b_{f4.2} = o_P(1)$ .

Next, consider  $b_{f4.1}$ . We have that

$$b_{f4.1} = \frac{1}{T} \sum_{t=1}^{T-h} F_t \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s)' \left( F'_t \frac{\Lambda' e_s}{N} \right) = \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t F'_t \right) \left( \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} (\tilde{F}_s - HF_s)' \right),$$

where the first term is  $O_P(1)$  and the second term can be shown to be  $O_P\left(\frac{1}{N}\right)$ . Thus, we will get a non negligible contribution from  $b_{f4.1}$  when multiplying by  $\sqrt{T}$ . Specifically, using the usual decomposition for  $\tilde{F}_s - HF_s$ , we have that

$$\frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} (\tilde{F}_s - HF_s)' = \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} (A_{1s} + A_{2s} + A_{3s} + A_{4s})' \tilde{V}^{-1}. \quad (19)$$

We first show that the first, second and last terms are  $o_P(1)$  when multiplied by  $\sqrt{T}$ . To end the proof we will study the third term. Starting with the first term in (19) (and ignoring  $\tilde{V}^{-1}$ ),

$$\left\| \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} A'_{1s} \right\| \leq \left( \frac{1}{T} \sum_{s=1}^T \left\| \frac{\Lambda' e_s}{N} \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^T \|A_{1s}\|^2 \right)^{1/2} = O_P\left(\frac{1}{\sqrt{N}}\right) O_P\left(\frac{1}{\sqrt{T}}\right),$$

so that this term is  $o_P(1)$  when multiplied by  $\sqrt{T}$ . Similarly,

$$\left\| \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} A'_{2s} \right\| \leq \left( \frac{1}{T} \sum_{s=1}^T \left\| \frac{\Lambda' e_s}{N} \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^T \|A_{2s}\|^2 \right)^{1/2} = O_P\left(\frac{1}{\sqrt{N}}\right) O_P\left(\frac{1}{\sqrt{N}\delta_{NT}}\right) = O_P\left(\frac{1}{N\delta_{NT}}\right),$$

so that when we multiply by  $\sqrt{T}$  we get  $O_P\left(\frac{\sqrt{T}}{N}\right) O_P\left(\frac{1}{\delta_{NT}}\right) = o_P(1)$ .

For the last term in (19), a more careful analysis is required. We have that

$$\begin{aligned} \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} A'_{4s} &= \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \left( \frac{1}{T} \sum_{t=1}^T (\tilde{F}_t - HF_t) \xi_{ts} + H \frac{1}{T} \sum_{t=1}^T F_t \xi_{ts} \right)' \\ &= \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \frac{1}{T} \sum_{t=1}^T (\tilde{F}_t - HF_t)' \xi_{ts} + H \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \frac{1}{T} \sum_{t=1}^T F'_t \xi_{ts}. \quad (20) \end{aligned}$$

For the first term in (20),

$$\begin{aligned} & \left\| \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \frac{1}{T} \sum_{t=1}^T (\tilde{F}_t - HF_t) \xi_{ts} \right\| = \left\| \frac{1}{T} \sum_{t=1}^T (\tilde{F}_t - HF_t) \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \xi_{ts} \right\| \\ & \leq \left( \frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t - HF_t \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^T \left\| \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \xi_{ts} \right\|^2 \right)^{1/2} = O_P \left( \frac{1}{\delta_{NT}} \right) O_P \left( \frac{1}{N} \right) = O_P \left( \frac{1}{N\delta_{NT}} \right), \end{aligned}$$

since

$$\frac{1}{T} \sum_{t=1}^T \left\| \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \xi_{ts} \right\|^2 = \frac{1}{T} \sum_{t=1}^T \left\| \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \eta_{st} \right\|^2 \leq \frac{1}{T} \sum_{s=1}^T \left\| \frac{\Lambda' e_s}{N} \right\|^2 \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T |\eta_{st}|^2 = O_P \left( \frac{1}{N^2} \right).$$

For the second term in (20), given Assumption 3(c),

$$\begin{aligned} \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \frac{1}{T} \sum_{t=1}^T F_t' \xi_{ts} &= \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \frac{1}{T} \sum_{t=1}^T \left( F_s' \frac{\Lambda' e_t}{N} \right) F_t' \\ &= \frac{1}{TN} \left( \frac{1}{\sqrt{TN}} \sum_{s=1}^T \Lambda' e_s F_s' \right) \left( \frac{1}{\sqrt{TN}} \sum_{t=1}^T \Lambda' e_t F_t' \right) = O_P \left( \frac{1}{TN} \right). \end{aligned}$$

Thus, (20) is  $O_P \left( \frac{1}{N\delta_{NT}} \right) + O_P \left( \frac{1}{NT} \right) = O_P \left( \frac{1}{N\delta_{NT}} \right)$ , which when multiplied by  $\sqrt{T}$  is  $o_P(1)$ .

Next, we analyze the dominant term in  $b_{f4.1}$  which comes from the contribution involving  $A_{3s}$ .

For this term, we have that

$$\begin{aligned} \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} A_{3s}' \tilde{V}^{-1} &= \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \left( \frac{1}{T} \sum_{t=1}^T \tilde{F}_t \eta_{ts} \right)' \tilde{V}^{-1} = \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \left( \frac{1}{T} \sum_{t=1}^T \tilde{F}_t F_t' \frac{\Lambda' e_s}{N} \right)' \tilde{V}^{-1} \\ &= \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \left( \frac{\tilde{F}' F \Lambda' e_s}{T N} \right)' = \frac{1}{N} \frac{1}{T} \sum_{s=1}^T \left( \frac{\Lambda' e_s}{\sqrt{N}} \right) \left( \frac{e_s' \Lambda}{\sqrt{N}} \right) \frac{F' \tilde{F}}{T} \tilde{V}^{-1}, \\ &= \frac{1}{N} (\Gamma + o_P(1)) Q' V^{-1}, \end{aligned}$$

given that  $\frac{F' \tilde{F}}{T} = Q' + o_P(1)$  by Proposition 1 of Bai (2003) and that by Assumption 3(e),

$$\frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\Lambda' e_t}{\sqrt{N}} \right) \left( \frac{e_t' \Lambda}{\sqrt{N}} \right) = \Gamma + o_P(1).$$

Thus,

$$\sqrt{T} b_{f4.1} = \left( \frac{1}{T} \sum_{t=1}^{T-h} F_t F_t' \right) \left( \frac{\sqrt{T}}{N} (\Gamma + o_P(1)) Q' V^{-1} \right) = c_{\Sigma_F} \Gamma Q' V^{-1} + o_P(1).$$

Since

$$C = H \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} F_t (A_{1t} + A_{2t} + A_{3t} + A_{4t})' \tilde{V}^{-1} = \sqrt{T} H (b_{f1} + b_{f2} + b_{f3} + b_{f4}) \tilde{V}^{-1},$$

it follows that

$$C = cH_0\Sigma_F\Gamma Q'V^{-1}V^{-1} + o_P(1) = cQ\Gamma Q'V^{-2} + o_P(1),$$

because  $H_0 = p\lim H$  is such that  $H_0\Sigma_F = Q$ . Indeed, note that  $H_0 = p\lim H = V^{-1}Q\Sigma_\Lambda$ , which implies that

$$\begin{aligned} H_0\Sigma_F &= V^{-1} \left( V^{1/2}\Upsilon'\Sigma_\Lambda^{-1/2} \right) \Sigma_\Lambda\Sigma_F = V^{-1}V^{1/2} \left( \Upsilon'\Sigma_\Lambda^{1/2}\Sigma_F\Sigma_\Lambda^{1/2} \right) \Sigma_\Lambda^{-1/2} \\ &= V^{-1}V^{1/2} (V\Upsilon') \Sigma_\Lambda^{-1/2} = V^{1/2}\Upsilon'\Sigma_\Lambda^{-1/2} = Q, \end{aligned}$$

given the definition of  $Q$  and the fact that  $\Upsilon$  is such that  $\Sigma_\Lambda^{1/2}\Sigma_F\Sigma_\Lambda^{1/2}\Upsilon = \Upsilon V$ .

**Proof of part c).** The proof follows closely the proof of part b) by relying on moment and dependence conditions involving the extra regressors  $W_t$ . In particular, writing

$$\begin{aligned} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} W_t \left( \tilde{F}_t - HF_t \right)' &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} W_t (A_{1t} + A_{2t} + A_{3t} + A_{4t})' \tilde{V}^{-1} \\ &\equiv \sqrt{T} (d_1 + d_2 + d_3 + d_4) \tilde{V}^{-1}, \end{aligned}$$

we verify  $\sqrt{T}d_i = o_P(1)$  for  $i = 1, 2, 3$  by using the same arguments as for  $b_{f1}$ ,  $b_{f2}$  and  $b_{f3}$ . The only term that has a nonzero contribution is  $d_4$ . Following the same arguments as for  $b_{f4}$ , we have that

$$\begin{aligned} \sqrt{T}d_4\tilde{V}^{-1} &= \sqrt{T} \left( \frac{1}{T} \sum_{t=1}^{T-h} W_t F_t' \right) \left( \frac{1}{T} \sum_{s=1}^T \frac{\Lambda' e_s}{N} \left( \tilde{F}_s - HF_s \right)' \right) \tilde{V}^{-1} \\ &= \frac{\sqrt{T}}{N} \left( \frac{1}{T} \sum_{t=1}^{T-h} W_t F_t' \right) \left( \Gamma Q' \tilde{V}^{-1} + o_P(1) \right) \tilde{V}^{-1} \\ &= c\Sigma_{WF}\Gamma Q'V^{-2} + o_P(1) = c\Sigma_{WF}H_0'Q\Gamma Q'V^{-2} + o_P(1), \text{ since } H_0'Q = I_r \\ &= c\Sigma_{WF}H_0'V (V^{-1}Q\Gamma Q'V^{-1}) V^{-1} + o_P(1) \\ &= c\Sigma_{W\tilde{F}}V\Sigma_{\tilde{F}}V^{-1} + o_P(1), \text{ since } \Sigma_{W\tilde{F}} = \Sigma_{WF}H_0' \text{ and } \Sigma_{\tilde{F}} = V^{-1}Q\Gamma Q'V^{-1}. \end{aligned}$$

This ends the proof.

**Proof of part d).** This follows immediately from parts a), b) and c) of this Lemma.

## B Appendix B: Proofs of results in Section 3

This Appendix is organized as follows. First, we provide some auxiliary lemmas, then we prove the results in Section 3, and finally we prove the auxiliary lemmas.

**Lemma B.1** *Let  $H^* = \tilde{V}^{*-1} \frac{\tilde{F}'\tilde{F}}{T} \frac{\tilde{\Lambda}'\tilde{\Lambda}}{N}$ . Under Conditions A\*-D\*, we have that if  $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$ ,*

a)  $H^*H^{*'} = I_r + O_{P^*}(\delta_{NT}^{-2})$ , in probability, i.e.  $H^*$  is asymptotically an orthogonal matrix.

b)  $H^* = H_0^* + O_{P^*}(\delta_{NT}^{-2})$ , in probability, where  $H_0^*$  is a diagonal matrix with  $\pm 1$  on the main diagonal.

c)  $\tilde{V}^* = H^* \tilde{V} H^{*'} + O_{P^*}(\delta_{NT}^{-2}) = \tilde{V} + O_{P^*}(\delta_{NT}^{-2})$ , in probability.

**Lemma B.2** Assume Assumptions 1-5 hold and suppose we generate bootstrap data  $\{y_{t+h}^*, X_t^*\}$  according to the residual-based bootstrap DGP (7) and (8) by relying on bootstrap residuals  $\{\varepsilon_{t+h}^*\}$  and  $\{e_t^*\}$  such that Conditions A\*-D\* are satisfied. Then, as  $N, T \rightarrow \infty$ ,

$$\frac{1}{T} \sum_{t=1}^{T-h} \left( \tilde{F}_t^* - H^* \tilde{F}_t \right) \varepsilon_{t+h}^* = O_{P^*} \left( \frac{1}{\delta_{NT} \sqrt{T}} \right),$$

in probability, for  $h \geq 0$ .

**Lemma B.3** Suppose conditions A\*-D\* hold. Then, as  $N, T \rightarrow \infty$ ,

a)

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} \left( \tilde{F}_t^* - H^* \tilde{F}_t \right) \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' &= \frac{1}{N} \tilde{V}^{*-1} H^* \left[ \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \right) \left( \frac{e_t^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] H^{*'} \tilde{V}^{*-1} \\ &+ O_{P^*} \left( \frac{1}{T} \right) + O_{P^*} \left( \frac{1}{N \delta_{NT}} \right) + O_{P^*} \left( \frac{1}{\sqrt{NT}} \right). \end{aligned}$$

b)

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} H^* \tilde{F}_t \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' &= H^* \frac{1}{N} \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t \tilde{F}_t' \right) \left[ \frac{1}{T} \sum_{s=1}^T \left( \frac{\tilde{\Lambda}' e_s^*}{\sqrt{N}} \right) \left( \frac{e_s^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-2} \\ &+ O_{P^*} \left( \frac{1}{\delta_{NT} \sqrt{T}} \right) + O_{P^*} \left( \frac{1}{N \delta_{NT}} \right). \end{aligned}$$

c)

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} W_t \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' &= \frac{1}{N} \left( \frac{1}{T} \sum_{t=1}^{T-h} W_t \tilde{F}_t' \right) \left[ \frac{1}{T} \sum_{s=1}^T \left( \frac{\tilde{\Lambda}' e_s^*}{\sqrt{N}} \right) \left( \frac{e_s^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-2} \\ &+ O_{P^*} \left( \frac{1}{\delta_{NT} \sqrt{T}} \right) + O_{P^*} \left( \frac{1}{N \delta_{NT}} \right). \end{aligned}$$

**Lemma B.4** Suppose conditions A\*-D\* hold. For any  $h \geq 0$ , if  $\sqrt{T}/N \rightarrow c$ ,  $0 \leq c < \infty$ , then

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \tilde{F}_t^* \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' (H^{*'})^{-1} \hat{\alpha} = c (H_0^{*'})^{-1} \underbrace{\left[ \tilde{V}^{-1} \Gamma^* \tilde{V}^{-1} + \Gamma^* \tilde{V}^{-2} \right]}_{\equiv B_\alpha^*} \hat{\alpha} + o_{P^*}(1),$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} W_t \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' (H^{*'})^{-1} \hat{\alpha} = c \underbrace{\left[ \tilde{\Sigma}_{W\tilde{F}} \Gamma^* \tilde{V}^{-2} \right]}_{\equiv B_\beta^*} \hat{\alpha} + o_{P^*}(1),$$

in probability, where  $\tilde{\Sigma}_{W\tilde{F}} = \frac{1}{T} \sum_{t=1}^{T-h} W_t \tilde{F}_t'$ .

**Proof of Lemma 3.1.** The proof is based on the following identity:

$$\tilde{F}_t^* - H^* \tilde{F}_t = \tilde{V}^{*-1} \left( \underbrace{\frac{1}{T} \sum_{s=1}^T \tilde{F}_s^* \gamma_{st}^*}_{\equiv A_{1t}^*} + \underbrace{\frac{1}{T} \sum_{s=1}^T \tilde{F}_s^* \zeta_{st}^*}_{\equiv A_{2t}^*} + \underbrace{\frac{1}{T} \sum_{s=1}^T \tilde{F}_s^* \eta_{st}^*}_{\equiv A_{3t}^*} + \underbrace{\frac{1}{T} \sum_{s=1}^T \tilde{F}_s^* \xi_{st}^*}_{\equiv A_{4t}^*} \right),$$

where

$$\begin{aligned} \gamma_{st}^* &= E^* \left( \frac{1}{N} \sum_{i=1}^N e_{is}^* e_{it}^* \right), \quad \zeta_{st}^* = \frac{1}{N} \sum_{i=1}^N (e_{is}^* e_{it}^* - E^*(e_{is}^* e_{it}^*)), \\ \eta_{st}^* &= \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}'_i \tilde{F}_s e_{it}^* = \tilde{F}'_s \frac{\tilde{\Lambda}' e_t^*}{N} \quad \text{and} \quad \xi_{st}^* = \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}'_i \tilde{F}_t e_{is}^* = \eta_{ts}^*. \end{aligned}$$

Ignoring  $\tilde{V}^{*-1}$  (which is  $O_{P^*}(1)$ ), it follows that

$$\frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t^* - H^* \tilde{F}_t \right\|^2 \leq \frac{1}{T} \sum_{t=1}^T \left( \|A_{1t}^*\|^2 + \|A_{2t}^*\|^2 + \|A_{3t}^*\|^2 + \|A_{4t}^*\|^2 \right),$$

By the Cauchy-Schwartz inequality,  $\left\| \sum_{s=1}^T \tilde{F}_s^* \gamma_{st}^* \right\|^2 \leq \left( \sum_{s=1}^T \left\| \tilde{F}_s^* \right\|^2 \right) \left( \sum_{s=1}^T \gamma_{st}^{*2} \right)$ , implying that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \|A_{1t}^*\|^2 &\leq \frac{1}{T} \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left\| \tilde{F}_s^* \right\|^2 \right)}_{=\frac{\|\tilde{F}^*\|^2}{T}=r \text{ because } \frac{\tilde{F}^* \tilde{F}^*}{T} = I_r} \underbrace{\left( \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T \gamma_{st}^{*2} \right)}_{=O_P(1) \text{ by Condition A}^*(b)} = O_P \left( \frac{1}{T} \right). \end{aligned}$$

For the second term, we have that

$$\frac{1}{T} \sum_{t=1}^T \|A_{2t}^*\|^2 \leq \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \left\| \tilde{F}_s^* \right\|^2 \right)}_{=r} \underbrace{\left( \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T |\zeta_{st}^*|^2 \right)}_{O_{P^*} \left( \frac{1}{N} \right)} = O_{P^*} \left( \frac{1}{N} \right).$$

In particular, by Condition A\*(c), we can show that

$$\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T E^* |\zeta_{st}^*|^2 = \frac{1}{N} \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T E^* \left| \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{is}^* e_{it}^* - E^*(e_{is}^* e_{it}^*)) \right|^2 = O_P \left( \frac{1}{N} \right),$$

which explains why the second term is  $O_{P^*} \left( \frac{1}{N} \right)$ . For the third term,

$$\frac{1}{T} \sum_{t=1}^T \|A_{3t}^*\|^2 = \frac{1}{T} \sum_{t=1}^T T^{-2} \left\| \sum_{s=1}^T \tilde{F}_s^* \tilde{F}'_s \frac{\tilde{\Lambda}' e_t^*}{N} \right\|^2 \leq \frac{1}{T} \sum_{t=1}^T \left\| \frac{\tilde{\Lambda}' e_t^*}{N} \right\|^2 \left\| T^{-1} \sum_{s=1}^T \tilde{F}_s^* \tilde{F}'_s \right\|^2 = O_{P^*} \left( \frac{1}{N} \right),$$

since  $\left\| T^{-1} \sum_{s=1}^T \tilde{F}_s^* \tilde{F}'_s \right\|^2 \leq r^2$ , whereas by Condition B\*(d) and Markov's inequality,  $\frac{1}{T} \sum_{t=1}^T \left\| \frac{\tilde{\Lambda}' e_t^*}{N} \right\|^2 = O_{P^*} \left( \frac{1}{N} \right)$ . The fourth term in  $d_t$  follows by the same arguments.

**Proof of Theorem 3.1.** The proof follows the proof of Theorem 2.1. In particular, we can write the bootstrap analogue of (17), viz

$$\sqrt{T} (\hat{\delta}^* - \delta^*) = \left( \frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t^* \hat{z}_t^{*'} \right)^{-1} (A^* + B^* + C^*),$$

where conditional on the original data, with probability converging to one, we have that

$$A^* = \Phi^* \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \hat{z}_t \varepsilon_{t+h}^* \xrightarrow{d^*} N(0, \Phi_0^* \Omega \Phi_0^{*'}),$$

given Conditions D\*(b) and E\*, and given that  $\Phi_0^* \equiv p \lim \Phi^*$ ;  $B^* = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (\tilde{F}_t^* - H^* \tilde{F}_t) \varepsilon_{t+h}^* = o_{P^*}(1)$  (given Lemma B.2);

$$C^* = -\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \hat{z}_t^* (\tilde{F}_t^* - H^* \tilde{F}_t)' (H^{*-1})' \hat{\alpha} \xrightarrow{P^*} -c (\Phi_0^{*'})^{-1} \Delta_\delta^*,$$

where  $\Delta_\delta^* \equiv (B_\alpha^{*'}, B_\beta^{*'})'$  is defined in Lemma B.4. Under Assumptions 1-5,  $p \lim \tilde{V} = V$  and  $p \lim \hat{\alpha} = (H_0')^{-1} \alpha$ ,  $p \lim \tilde{\Sigma}_{W\tilde{F}} = \Sigma_{W\tilde{F}}$ , and  $p \lim \Gamma^* = Q\Gamma Q'$  by Condition F\*, which implies that  $\Delta_\delta^* \xrightarrow{P} \Delta_\delta$ ; and finally  $\frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t^* \hat{z}_t^{*'} = \Phi_0^* \Phi_0 \Sigma_{zz} \Phi_0' \Phi_0^{*'} + o_{P^*}(1)$ . This implies that  $\sqrt{T} (\hat{\delta}^* - \delta^*) \xrightarrow{d^*} N(-c (\Phi_0^{*'})^{-1} \Delta_\delta, (\Phi_0^{*'})^{-1} \Sigma_\delta (\Phi_0^*)^{-1})$ , in probability.

**Proof of Corollary 3.1.** By Theorem 2.1, under Assumptions 1-5, and if  $\sqrt{T}/N \rightarrow c$ ,  $0 \leq c < \infty$ ,  $\sqrt{T} (\hat{\delta} - \delta) \xrightarrow{d} Z \sim N(-c\Delta_\delta, \Sigma_\delta)$ . Thus, from a multivariate version of Polya's Theorem (cf. Battacharya and Rao (1986)), it follows that  $\sup_x \left| P(\sqrt{T} (\hat{\delta} - \delta) \leq x) - \Phi(x; -c\Delta_\delta, \Sigma_\delta) \right| = o(1)$ , where  $\Phi(x; -c\Delta_\delta, \Sigma_\delta)$  denotes the distribution function of  $Z$ . Then, the result follows if we show that

$$\sup_x \left| P(\sqrt{T} (\Phi^{*'} \hat{\delta}^* - \hat{\delta}) \leq x) - \Phi(x; -c\Delta_\delta, \Sigma_\delta) \right| = o_{P^*}(1). \quad (21)$$

Under the stated assumptions, from Theorem 3.1 we have that  $\sqrt{T} (\Phi^{*'} \hat{\delta}^* - \hat{\delta}) \xrightarrow{d^*} N(-c\Delta_\delta, \Sigma_\delta)$ , in probability. By arguing along subsequences and applying Polya's Theorem, this suffices to show (21).

**Proof of Lemma B.1.** For part a), writing

$$\frac{\tilde{F}^{*'} \tilde{F}}{T} = \frac{1}{T} (\tilde{F}^* - \tilde{F} H^{*'})' \tilde{F} + \frac{1}{T} H^* \tilde{F}' \tilde{F},$$

and noting that  $\frac{1}{T} \tilde{F}' \tilde{F} = I_r$ , yields

$$H^* = \frac{\tilde{F}^{*'} \tilde{F}}{T} + O_{P^*}(\delta_{NT}^{-2}), \quad (22)$$

where we used Lemma B.3 to bound the second term. Now right multiply by  $H^{*'}$  to get

$$\begin{aligned} H^* H^{*'} &= \frac{\tilde{F}^{*'} \tilde{F}}{T} H^{*'} + O_{P^*}(\delta_{NT}^{-2}) = \frac{\tilde{F}^{*'} (\tilde{F} H^{*'} - \tilde{F}^* + \tilde{F}^*)}{T} + O_{P^*}(\delta_{NT}^{-2}) = \frac{\tilde{F}^{*'} \tilde{F}^*}{T} + O_{P^*}(\delta_{NT}^{*-2}) \\ &= I_r + O_{P^*}(\delta_{NT}^{*-2}), \end{aligned}$$



where again we used Lemma B.3 and the fact that by construction  $\frac{\tilde{F}^{*'}\tilde{F}^*}{T} = I_r$ . This shows that  $H^{*'} = H^{*-1} + O_{P^*}(\delta_{NT}^{-2})$ , i.e.  $H^*$  is asymptotically an orthogonal matrix. The eigenvalues of  $H^*$  are therefore +1 or -1, for large  $N$  and  $T$ . For part b), by definition,

$$H^* = \tilde{V}^{*-1} \left( \frac{\tilde{F}^{*'}\tilde{F}^*}{T} \right) \frac{\tilde{\Lambda}'\tilde{\Lambda}}{N} = \tilde{V}^{*-1} H^* \frac{\tilde{\Lambda}'\tilde{\Lambda}}{N} + O_{P^*}(\delta_{NT}^{-2}),$$

given (22). Left multiplying both sides by  $\tilde{V}^*$  yields

$$\tilde{V}^* H^* = H^* \frac{\tilde{\Lambda}'\tilde{\Lambda}}{N} + O_{P^*}(\delta_{NT}^{-2}) = H^* \tilde{V} + O_{P^*}(\delta_{NT}^{-2}), \quad (23)$$

since  $\tilde{V} = \frac{\tilde{\Lambda}'\tilde{\Lambda}}{N}$  by construction of the principal components. By Lemma A.3 of Bai (2003),  $\tilde{V} \xrightarrow{P} V > 0$  and therefore we can write that

$$\tilde{V}^* H^* = H^* V + o_{P^*}(1),$$

or, transposing, that

$$V H^{*'} = H^{*'} \tilde{V}^* + o_{P^*}(1).$$

Thus,  $H^{*'}$  is (for large  $N$  and  $T$ ) the matrix of eigenvectors of  $V$ . Since  $V$  is a diagonal matrix,  $H^{*'}$  is also diagonal, asymptotically. Moreover, because  $V$  has distinct eigenvalues (by assumption), it follows that its eigenvectors have only one nonzero value and this is +1 or -1 (because  $H^*$  is orthogonal). Therefore  $H^{*'}$  is for large  $N$  and  $T$  a diagonal matrix with  $\pm 1$  in the main diagonal (in particular,  $H^* = \text{diag}(\text{sign}(\tilde{F}^{*'}\tilde{F}^*))$ ). Part c) follows from (23) by right multiplying by  $H^{*'}$  and using parts a) and b).

**Proof of Lemma B.2.** The proof follows exactly as the proof of Lemma A.1, thus we only highlight the differences. Throughout, we denote with a star the bootstrap analogue of a given formula in that proof. First, to bound  $I_2^*$ , the bootstrap analogue of  $I_2$ , we use directly Condition C\*(c) to conclude that  $I_2^* = O_{P^*}(\frac{1}{T})$ , in probability. Second, to bound  $I_1^*$ , we rely on Lemma 3.1 and on Conditions A\*(b) and D\*(a). For  $II_1^*$ , we use Condition C\*(a) instead of Assumption 4(a). For  $II_2^*$ , we use Condition B\*(b) instead of Assumption 3(b) to show that  $\frac{1}{T} \sum_{t=1}^{T-h} E^* \|m_t^*\|^2 = O_P(1)$ , and Condition D\*(a) instead of Assumption 5(a) to show that  $\frac{1}{T} \sum_{t=1}^{T-h} \varepsilon_{t+h}^{*2} = O_{P^*}(1)$ , in probability. To bound  $III_1^*$  and  $III_2^*$ , we rely on Condition C\*(b) instead of Assumption 4(b). Finally, to bound  $IV^*$ , we rely on Conditions B\*(d) and D\*(b) instead of Assumptions 3(d) and 5(b), respectively.

**Proof of Lemma B.3.** *Proof of part a).* We follow closely the proof of Lemma A.2. Specifically, we analyze each of the terms in

$$\frac{1}{T} \sum_{t=1}^{T-h} \left( \tilde{F}_t^* - H^* \tilde{F}_t \right) \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' = \tilde{V}^{*-1} \frac{1}{T} \sum_{t=1}^{T-h} (A_{1t}^* + A_{2t}^* + A_{3t}^* + A_{4t}^*) (A_{1t}^* + A_{2t}^* + A_{3t}^* + A_{4t}^*)' \tilde{V}^{*-1}.$$

In particular, by Lemma 3.1, and the appropriate bootstrap high level conditions, we can show that:

- $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t}^* A_{1t}^{*'} = O_{P^*}(\frac{1}{T})$ , in probability, using Condition A\*(b).

- $\frac{1}{T} \sum_{t=1}^{T-h} A_{2t}^* A_{2t}^{*'} = O_{P^*} \left( \frac{1}{N\delta_{NT}^2} \right)$ , in probability, using Conditions A\*(c) and B\*(b).
- $\frac{1}{T} \sum_{t=1}^{T-h} A_{3t}^* A_{3t}^{*'} = \frac{1}{N} H^* \left( \frac{\tilde{F}' \tilde{F}}{T} \right) \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \right) \left( \frac{e_t^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \left( \frac{\tilde{F}' \tilde{F}}{T} \right) H^{*'} + O_{P^*} \left( \frac{1}{N\delta_{NT}} \right)$ , in probability, using Condition B\*(d).
- Each of  $\frac{1}{T} \sum_{t=1}^{T-h} A_{4t}^* A_{4t}^{*'}$ ,  $\frac{1}{T} \sum_{t=1}^{T-h} A_{2t}^* A_{3t}^{*'}$ ,  $\frac{1}{T} \sum_{t=1}^{T-h} A_{2t}^* A_{4t}^{*'}$  and  $\frac{1}{T} \sum_{t=1}^{T-h} A_{3t}^* A_{4t}^{*'}$  is  $O_{P^*} \left( \frac{1}{N\delta_{NT}} \right)$ , in probability, using condition B\*(d).
- $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t}^* A_{2t}^{*'} = O_{P^*} \left( \frac{1}{\delta_{NT}\sqrt{NT}} \right)$ , in probability.
- $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t}^* A_{3t}^{*'}$  and  $\frac{1}{T} \sum_{t=1}^{T-h} A_{1t}^* A_{4t}^{*'}$  are  $O_{P^*} \left( \frac{1}{\sqrt{NT}} \right)$ , in probability.

Thus, we conclude that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} \left( \tilde{F}_t^* - H^* \tilde{F}_t \right) \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' &= \tilde{V}^{*-1} \frac{1}{N} H^* \left( \frac{\tilde{F}' \tilde{F}}{T} \right) \left[ \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \right) \left( \frac{e_t^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] \left( \frac{\tilde{F}' \tilde{F}}{T} \right) H^{*'} \tilde{V}^{*-1} \\ &\quad + O_{P^*} \left( \frac{1}{T} \right) + O_{P^*} \left( \frac{1}{N\delta_{NT}} \right) + O_{P^*} \left( \frac{1}{\sqrt{NT}} \right). \end{aligned}$$

*Proof of part b).* We analyze each of the terms in

$$H^* \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t (A_{1t}^* + A_{2t}^* + A_{3t}^* + A_{4t}^*)' \tilde{V}^{*-1} = H^* (b_{f1}^* + b_{f2}^* + b_{f3}^* + b_{f4}^*) \tilde{V}^{*-1},$$

where we let  $b_{fj}^* \equiv \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t A_{jt}^{*'}$ , for  $j = 1, \dots, 4$ . Following exactly the same steps as in the proof of part b) of Lemma A.2, we can show that:

- $b_{f1}^* = O_{P^*} \left( \frac{1}{\delta_{NT}\sqrt{T}} \right)$ , in probability, given Condition B\*(a) and Lemma 3.1.
- $b_{f2}^* = O_{P^*} \left( \frac{1}{\sqrt{TN}} \right)$ , in probability, given Condition B\*(b) and Lemma 3.1.
- $b_{f3}^* = O_{P^*} \left( \frac{1}{\sqrt{TN}} \right)$ , in probability, given Condition B\*(c) and Lemma 3.1.
- $b_{f4}^* = \frac{1}{N} \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t \tilde{F}_t' \right) \left[ \frac{1}{T} \sum_{s=1}^T \left( \frac{\tilde{\Lambda}' e_s^*}{\sqrt{N}} \right) \left( \frac{e_s^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-1} + O_{P^*} \left( \frac{1}{\sqrt{TN}} \right) + O_{P^*} \left( \frac{1}{N\delta_{NT}} \right)$ , in probability, given in particular Condition B\*(c) and Lemma 3.1. Thus, we can conclude that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} H^* \tilde{F}_t \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' &= H^* \frac{1}{N} \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t \tilde{F}_t' \right) \left[ \frac{1}{T} \sum_{s=1}^T \left( \frac{\tilde{\Lambda}' e_s^*}{\sqrt{N}} \right) \left( \frac{e_s^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-2} \\ &\quad + O_{P^*} \left( \frac{1}{\delta_{NT}\sqrt{T}} \right) + O_{P^*} \left( \frac{1}{N\delta_{NT}} \right). \end{aligned}$$

*Proof of part c).* This follows by the same arguments used in the proof of part c) of Lemma A.1.

**Proof of Lemma B.4.** From prove part a) of Lemma B.3, and noting that  $\frac{\tilde{F}'\tilde{F}}{T} = I_r$ , we have that

$$\begin{aligned}
& \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \left( \tilde{F}_t^* - H^* \tilde{F}_t \right) \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' (H^{*'})^{-1} \hat{\alpha} \\
&= \frac{\sqrt{T}}{N} \tilde{V}^{*-1} H^* \left[ \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \right) \left( \frac{e_t^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] H^{*'} \tilde{V}^{*-1} (H^{*'})^{-1} \hat{\alpha} \\
&\quad + O_{P^*} \left( \frac{1}{\sqrt{T}} \right) + O_{P^*} \left( \frac{\sqrt{T}}{N} \frac{1}{\delta_{NT}} \right) + O_{P^*} \left( \frac{1}{\sqrt{N}} \right) \\
&= c \tilde{V}^{*-1} H^* \Gamma^* H^{*'} \tilde{V}^{*-1} (H^{*'})^{-1} \hat{\alpha} + o_{P^*}(1) \\
&= c (H^{*'})^{-1} \tilde{V}^{-1} H^{*-1} H^* \Gamma^* H^{*'} (H^{*'})^{-1} \tilde{V}^{-1} H^{*-1} (H^{*'})^{-1} \hat{\alpha} + o_{P^*}(1) \\
&= c (H_0^{*'})^{-1} \tilde{V}^{-1} \Gamma^* \tilde{V}^{-1} \hat{\alpha} + o_{P^*}(1)
\end{aligned}$$

where the second equality uses Condition B\*(e), and the third equality follows from Lemma B.1.c). The last equality uses the fact that  $H_0^* = \text{diag}(\pm 1)$ , which implies that  $H^{*-1} (H^{*'})^{-1} \hat{\alpha} = \hat{\alpha} + o_{P^*}(1)$ . This proves part a). Similarly, from part b) of Lemma B.3 and given Condition B\*(e), we get that

$$\begin{aligned}
& \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} H^* \tilde{F}_t \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' (H^{*'})^{-1} \hat{\alpha} \\
&= \frac{\sqrt{T}}{N} H^* \left( \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t \tilde{F}_t' \right) \left[ \frac{1}{T} \sum_{s=1}^T \left( \frac{\tilde{\Lambda}' e_s^*}{\sqrt{N}} \right) \left( \frac{e_s^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-2} (H^{*'})^{-1} \hat{\alpha} + o_{P^*}(1) \\
&= c H^* \Gamma^* \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-2} (H^{*'})^{-1} \hat{\alpha} + o_{P^*}(1) \\
&= c H^* \Gamma^* \tilde{V}^{-1} H^{*'} (H^{*'})^{-1} \tilde{V}^{-1} H^{*-1} (H^{*'})^{-1} \hat{\alpha} + o_{P^*}(1) = c H_0^* \Gamma^* \tilde{V}^{-2} \hat{\alpha} + o_{P^*}(1),
\end{aligned}$$

where the second equality uses the fact that  $\frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t \tilde{F}_t' = I_r + o_P(1)$ , the third equality uses Lemma B.1.c) and the fact that  $\frac{\tilde{F}' \tilde{F}^*}{T} \tilde{V}^{*-1} = \tilde{V}^{-1} H^{*'}$ , and the last equality holds because  $H^{*-1} (H^{*'})^{-1} = I_r + o_{P^*}(1)$ , in probability.

To prove part c), we note that from part c) of Lemma B.3, we have that

$$\begin{aligned}
& \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} W_t \left( \tilde{F}_t^* - H^* \tilde{F}_t \right)' (H^{*'})^{-1} \hat{\alpha} \\
&= \frac{\sqrt{T}}{N} \left( \frac{1}{T} \sum_{t=1}^{T-h} W_t \tilde{F}_t' \right) \left[ \frac{1}{T} \sum_{s=1}^T \left( \frac{\tilde{\Lambda}' e_s^*}{\sqrt{N}} \right) \left( \frac{e_s^{*'} \tilde{\Lambda}}{\sqrt{N}} \right) \right] \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-2} (H^{*'})^{-1} \hat{\alpha} + o_{P^*}(1) \\
&= c \tilde{\Sigma}_{W\tilde{F}} \Gamma^* \left( \frac{\tilde{F}' \tilde{F}^*}{T} \right) \tilde{V}^{*-2} (H^{*'})^{-1} \hat{\alpha} + o_{P^*}(1) \\
&= c \tilde{\Sigma}_{W\tilde{F}} \Gamma^* \tilde{V}^{-2} \hat{\alpha} + o_{P^*}(1),
\end{aligned}$$

by relying on the same arguments as those used to prove part b).

## C Appendix C: Proofs of results in Section 4

First, we state an auxiliary result and its proof. Then we prove Theorem 4.1.

**Lemma C.1** *Suppose Assumptions 1-5 hold. If in addition either: (1)  $\{F_s\}$ ,  $\{\lambda_i\}$  and  $\{e_{it}\}$  are mutually independent and for some  $p \geq 2$ ,  $E|e_{it}|^{2p} \leq M < \infty$ ,  $E\|\lambda_i\|^p \leq M < \infty$  and  $E\|F_t\|^p \leq M < \infty$ , or (2) for some  $p \geq 2$ ,  $E|e_{it}|^{3p} \leq M < \infty$ ,  $E\|\lambda_i\|^{3p} \leq M < \infty$  and  $E\|F_t\|^{3p} \leq M < \infty$ , it follows that*

- (i)  $\frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t - HF_t \right\|^p = O_P(1)$ ;
- (ii)  $\frac{1}{N} \sum_{i=1}^N \left\| \tilde{\lambda}_i - H^{-1'} \lambda_i \right\|^p = O_P(1)$ ;
- (iii)  $\frac{1}{TN} \sum_{t=1}^T \sum_{i=1}^N \tilde{e}_{it}^p = O_P(1)$ .

**Proof of Lemma C.1. Proof of (i).** We rely on the following identity (see Bai and Ng (2002), proof of Theorem 1):

$$\tilde{F}_t - HF_t = \tilde{V}^{-1} \left( \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \psi_{st} + \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \eta_{st} + \frac{1}{T} \sum_{s=1}^T \tilde{F}_s \xi_{st} \right),$$

where  $\psi_{st} = \frac{1}{N} \sum_{i=1}^N e_{is} e_{it}$ ;  $\eta_{st} = \frac{1}{N} \sum_{i=1}^N \lambda_i' F_s e_{it}$ ; and  $\xi_{st} = \eta_{ts}$ . It follows that by the  $c-r$  inequality,

$$\frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t - HF_t \right\|^p \leq 3^{p-1} \left\| \tilde{V}^{-1} \right\|^p \left( \frac{1}{T} \sum_{t=1}^T a_t + \frac{1}{T} \sum_{t=1}^T b_t + \frac{1}{T} \sum_{t=1}^T c_t \right),$$

where

$$a_t = \frac{1}{T^p} \left\| \sum_{s=1}^T \tilde{F}_s \psi_{st} \right\|^p; \quad b_t = \frac{1}{T^p} \left\| \sum_{s=1}^T \tilde{F}_s \eta_{st} \right\|^p; \quad \text{and} \quad c_t = \frac{1}{T^p} \left\| \sum_{s=1}^T \tilde{F}_s \xi_{st} \right\|^p.$$

Let  $\chi_{st}$  denote either  $\psi_{st}$ ,  $\eta_{st}$  or  $\xi_{st}$ . We can write

$$\left\| \sum_{s=1}^T \tilde{F}_s \chi_{st} \right\|^p = \left( \left\| \sum_{s=1}^T \tilde{F}_s \chi_{st} \right\|^2 \right)^{p/2} \leq \left( \sum_{s=1}^T \|\tilde{F}_s\|^2 \sum_{s=1}^T |\chi_{st}|^2 \right)^{p/2},$$

where the inequality follows by Cauchy-Schwartz. It follows that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \frac{1}{T^p} \left\| \sum_{s=1}^T \tilde{F}_s \chi_{st} \right\|^p &\leq \frac{1}{T} \sum_{t=1}^T \underbrace{\left( \frac{1}{T} \sum_{s=1}^T \|\tilde{F}_s\|^2 \right)^{p/2}}_{=r} \left( \frac{1}{T} \sum_{s=1}^T |\chi_{st}|^2 \right)^{p/2} \leq r^{p/2} \frac{1}{T} \sum_{t=1}^T \left( \frac{1}{T} \sum_{s=1}^T |\chi_{st}|^2 \right)^{p/2} \\ &\leq r^{p/2} \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T |\chi_{st}|^p, \end{aligned}$$

where the last inequality follows again by the  $c-r$  inequality. Thus it suffices to show that  $E|\chi_{st}|^p \leq M < \infty$  to prove that the above term is  $O_P(1)$ . Starting with  $\chi_{st} = \psi_{st}$ ,

$$E|\psi_{st}|^p = E\left|\frac{1}{N}\sum_{i=1}^N e_{it}e_{is}\right|^p \leq \frac{1}{N}\sum_{i=1}^N E|e_{it}e_{is}|^p \leq \frac{1}{N}\sum_{i=1}^N \left(E|e_{it}|^{2p}\right)^{1/2} \left(E|e_{is}|^{2p}\right)^{1/2} \leq M < \infty,$$

given that we assume  $E|e_{it}|^{2p} \leq M < \infty$ . When  $\chi_{st} = \eta_{st}$ , we have that

$$E|\eta_{st}|^p = E\left|\frac{1}{N}\sum_{i=1}^N \lambda'_i F_s e_{it}\right|^p \leq \frac{1}{N}\sum_{i=1}^N E|\lambda'_i F_s e_{it}|^p < M,$$

since

$$E|\lambda'_i F_s e_{it}|^p \leq \left(E\|\lambda_i e_{it}\|^{\frac{3p}{2}}\right)^{2/3} \left(E\|F_s\|^{3p}\right)^{1/3} \leq \left(E\|\lambda_i\|^{3p} E|e_{it}|^{3p}\right)^{1/3} \left(E\|F_s\|^{3p}\right)^{1/3} \leq M$$

given the assumptions that  $E\|\lambda_i\|^{3p} \leq M$ ,  $E\|F_s\|^{3p} \leq M$ , and  $E|e_{it}|^{3p} \leq M$  in obtaining the last inequality. Note that if we assume that  $\{\lambda_i\}$ ,  $\{F_s\}$  and  $\{e_{it}\}$  are three mutually independent groups of random variables, then it suffices that  $E\|\lambda_i\|^p \leq M$ ,  $E\|F_s\|^p \leq M$ , and  $E|e_{it}|^p \leq M$  to bound  $E|\lambda'_i F_s e_{it}|^p$ . The term that depends on  $\chi_{st} = \xi_{st}$  can be dealt with similarly.

**Proof of (ii).** Note that  $\tilde{\Lambda} = \frac{X'\tilde{F}}{T}$ , which implies that  $\tilde{\Lambda}' = \frac{\tilde{F}'X}{T}$ . Since  $X = F\Lambda' + e$ , it follows that

$$\tilde{\Lambda}' = \frac{\tilde{F}'F}{T}\Lambda' + \frac{\tilde{F}'e}{T},$$

thus implying that  $\tilde{\lambda}_i = \frac{\tilde{F}'F}{T}\lambda_i + \frac{\tilde{F}'e_i}{T}$ , where  $e_i = (e_{i1}, \dots, e_{iT})'$ . We can write

$$\frac{\tilde{F}'FH'}{T} = \frac{\tilde{F}'\tilde{F}}{T} - \frac{\tilde{F}'(\tilde{F} - FH')}{T} = I_r - \frac{\tilde{F}'(\tilde{F} - FH')}{T},$$

from which it follows that

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

Thus,

$$\frac{1}{N}\sum_{i=1}^N \left\|\tilde{\lambda}_i - H'^{-1}\lambda_i\right\|^p \leq 3^{p-1} \left( \begin{array}{l} \frac{1}{N}\sum_{i=1}^N \left\|T^{-1}\tilde{F}'(\tilde{F} - FH')H'^{-1}\lambda_i\right\|^p \\ + \frac{1}{N}\sum_{i=1}^N \left\|T^{-1}(\tilde{F} - FH')'e_i\right\|^p + \frac{1}{N}\sum_{i=1}^N \left\|T^{-1}(FH')'e_i\right\|^p \end{array} \right).$$

For the first term, we have that

$$\frac{1}{N}\sum_{i=1}^N \left\|T^{-1}\tilde{F}'(\tilde{F} - FH')H'^{-1}\lambda_i\right\|^p \leq \left\|T^{-1/2}\tilde{F}\right\|^p \left\|T^{-1/2}(\tilde{F} - FH')\right\|^p \|H'^{-1}\|^p \frac{1}{N}\sum_{i=1}^N \|\lambda_i\|^p.$$

Now,

$$\left\|T^{-1/2}\tilde{F}\right\|^p = \left(T^{-1}\|\tilde{F}\|^2\right)^{p/2} = \left(T^{-1}\sum_{t=1}^T \|\tilde{F}_t\|^2\right)^{p/2} = r^{p/2},$$

given that  $\tilde{F}'\tilde{F}/T = I_r$ . Similarly, under Assumptions 1-5,

$$\left\| T^{-1/2} \left( \tilde{F} - FH' \right) \right\|^p = \left( T^{-1} \sum_{t=1}^T \left\| \tilde{F}_t - HF_t \right\|^2 \right)^{p/2} = O_P \left( \delta_{NT}^{-p} \right) = O_P(1).$$

Since  $\|H'^{-1}\|^p = O_P(1)$ , it follows that the first term is  $O_P(1)$  provided  $E\|\lambda_i\|^p \leq M < \infty$ , which holds by assumption.

For the second term,

$$\frac{1}{N} \sum_{t=1}^N \left\| T^{-1} \left( \tilde{F} - FH' \right)' e_i \right\|^p = \left\| T^{-1/2} \left( \tilde{F} - FH' \right) \right\|^p \frac{1}{N} \sum_{t=1}^N \left\| T^{-1/2} e_i \right\|^p,$$

where the first factor is  $O_P(1)$  and the second factor is dominated by

$$\frac{1}{N} \sum_{i=1}^N \left( \left\| T^{-1/2} e_i \right\|^2 \right)^{p/2} = \frac{1}{N} \sum_{t=1}^N \left( T^{-1} e_i' e_i \right)^{p/2} = \frac{1}{N} \sum_{t=1}^N \left( T^{-1} \sum_{i=1}^N e_{it}^2 \right)^{p/2} \leq \frac{1}{NT} \sum_{t=1}^N \sum_{i=1}^N e_{it}^p,$$

which is  $O_P(1)$  given the assumption that  $E|e_{it}|^p \leq M$ . The third term can be bounded similarly using in particular the fact that  $E\|F_t\|^2 \leq M < \infty$ .

**Proof of (iii).** We can write

$$\tilde{c}_{it} = e_{it} - \lambda_i' H^{-1} \left( \tilde{F}_t - HF_t \right) - \left( \tilde{\lambda}_i - H^{-1'} \lambda_i \right)' \tilde{F}_t,$$

which implies that

$$\begin{aligned} \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N |\tilde{c}_{it}|^p &\leq 3^{p-1} \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N |e_{it}|^p + 3^{p-1} \frac{1}{N} \sum_{i=1}^N \|\lambda_i\|^p \|H^{-1}\|^p \frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t - HF_t \right\|^p \\ &\quad + 3^{p-1} \frac{1}{N} \sum_{i=1}^N \left\| \tilde{\lambda}_i - H^{-1'} \lambda_i \right\|^p \frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t \right\|^p. \end{aligned}$$

The first term is  $O_P(1)$  given that  $E|e_{it}|^p = O(1)$ ; the second term is  $O_P(1)$  since  $E\|\lambda_i\|^p = O(1)$  and given part (i); and the third term is  $O_P(1)$  given parts (ii) and (iii), since in particular

$$\frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t \right\|^p \leq \frac{1}{T} \sum_{t=1}^T \left\| HF_t + \left( \tilde{F}_t - HF_t \right) \right\|^p \leq 2^{p-1} \left( \|H\|^p \frac{1}{T} \sum_{t=1}^T \|F_t\|^p + \frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t - HF_t \right\|^p \right) = O_P(1).$$

**Proof of Theorem 4.1.** We verify Condition A\*-F\*. We start with Condition A\*. Since  $e_{it}^* = \tilde{e}_{it}\eta_{it}$  where i.i.d.  $(0,1)$  across  $(i,t)$ , part a) follows immediately. For part b), note that  $\gamma_{st}^* = \frac{1}{N} \sum_{i=1}^N \tilde{e}_{it}\tilde{e}_{is}1(t=s)$ , which implies that  $\frac{1}{T} \sum_{t,s} \gamma_{st}^{*2} = \frac{1}{T} \sum_{t=1}^T \left( \frac{1}{N} \sum_{i=1}^N \tilde{e}_{it}^2 \right)^2$ . This expression is bounded by  $\frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \tilde{e}_{it}^4$ , which is  $O_P(1)$  under our assumptions by an application of Lemma C.1 (iii) with  $p=4$ . For c), note that for any  $(t,s)$ ,

$$E^* \left| \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{it}^* e_{is}^* - E^*(e_{it}^* e_{is}^*)) \right|^2 = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N Cov^*(e_{it}^* e_{is}^*, e_{jt}^* e_{js}^*).$$

For the wild bootstrap where  $e_{it}^* = \tilde{e}_{it}\eta_{it}$ , with  $\eta_{it}$  i.i.d. across  $(i, t)$ ,

$$Cov^*(e_{it}^*e_{is}^*, e_{jt}^*e_{js}^*) = \tilde{e}_{it}\tilde{e}_{is}\tilde{e}_{jt}\tilde{e}_{js}Cov(\eta_{it}\eta_{is}, \eta_{jt}\eta_{js}) = \begin{cases} \tilde{e}_{it}^2\tilde{e}_{is}^2Var(\eta_{it}\eta_{is}) & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases},$$

which implies that

$$E^* \left| \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{it}^*e_{is}^* - E^*(e_{it}^*e_{is}^*)) \right|^2 = \frac{1}{N} \sum_{i=1}^N \tilde{e}_{it}^2\tilde{e}_{is}^2Var(\eta_{it}\eta_{is}).$$

Thus, condition A\*(c) becomes

$$\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \frac{1}{N} \sum_{i=1}^N \tilde{e}_{it}^2\tilde{e}_{is}^2 \underbrace{Var(\eta_{it}\eta_{is})}_{\leq \bar{\eta}} \leq \bar{\eta} \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T} \sum_{t=1}^T \tilde{e}_{it}^2 \right)^2 \leq \bar{\eta} C \frac{1}{N} \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T \tilde{e}_{it}^4 = O_P(1),$$

for some constants  $\bar{\eta}$  and  $C$ , which holds given Lemma C.1 (iii) with  $p = 4$ . For Condition B\*(a), under the wild bootstrap, in particular the bootstrap time series independence, we have that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T \tilde{F}_s \tilde{F}_t' \gamma_{st}^* &= \frac{1}{T} \sum_{t=1}^T \tilde{F}_t \tilde{F}_t' \gamma_{tt}^* = \frac{1}{T} \sum_{t=1}^T \tilde{F}_t \tilde{F}_t' \left( \frac{1}{N} \sum_{i=1}^N \tilde{e}_{it}^2 \right) \\ &\leq \left( \frac{1}{T} \sum_{t=1}^T \|\tilde{F}_t \tilde{F}_t'\|^2 \right)^{1/2} \left[ \frac{1}{T} \sum_{t=1}^T \left( \frac{1}{N} \sum_{i=1}^N \tilde{e}_{it}^2 \right)^2 \right]^{1/2} \leq \left( \frac{1}{T} \sum_{t=1}^T \|\tilde{F}_t\|^4 \right)^{1/2} \left[ \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \tilde{e}_{it}^4 \right]^{1/2}, \end{aligned}$$

which is  $O_P(1)$  under our assumptions. For Condition B\*(b), we have that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T E^* \left\| \frac{1}{\sqrt{TN}} \sum_{s=1}^T \sum_{i=1}^N \hat{z}_s (e_{it}^*e_{is}^* - E^*(e_{it}^*e_{is}^*)) \right\|^2 &= \frac{1}{T} \sum_{t=1}^T \frac{1}{T} E^* \left\| \sum_{s=1}^T \hat{z}_s \left[ \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{it}^*e_{is}^* - E^*(e_{it}^*e_{is}^*)) \right] \right\|^2 \\ &= \frac{1}{T} \sum_{t=1}^T \frac{1}{T} \sum_{s=1}^T \sum_{l=1}^T \hat{z}_s' \hat{z}_l E^* \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{it}^*e_{is}^* - E^*(e_{it}^*e_{is}^*)) \frac{1}{\sqrt{N}} \sum_{j=1}^N (e_{jt}^*e_{jl}^* - E^*(e_{jt}^*e_{jl}^*)) \right), \end{aligned}$$

where

$$E^* \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N (e_{it}^*e_{is}^* - E^*(e_{it}^*e_{is}^*)) \frac{1}{\sqrt{N}} \sum_{j=1}^N (e_{jt}^*e_{jl}^* - E^*(e_{jt}^*e_{jl}^*)) \right) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N Cov^*(e_{it}^*e_{is}^*, e_{jt}^*e_{jl}^*).$$

Using the properties of the wild bootstrap, in particular the bootstrap cross sectional independence, we can show that

$$Cov^*(e_{it}^*e_{is}^*, e_{jt}^*e_{jl}^*) = 0 \text{ when } i \neq j \text{ for any } t, s, l.$$

When  $i = j$ ,

$$Cov^*(e_{it}^*e_{is}^*, e_{jt}^*e_{jl}^*) = Cov^*(e_{it}^*e_{is}^*, e_{it}^*e_{il}^*) = \begin{cases} 0 & \text{if } s \neq l \\ Var^*(e_{it}^*e_{is}^*) = \tilde{e}_{it}^2\tilde{e}_{is}^2Var^*(\eta_{it}\eta_{is}) & \text{if } s = l. \end{cases}$$

It follows that

$$\begin{aligned}
& \frac{1}{T} \sum_{t=1}^T E^* \left\| \frac{1}{\sqrt{TN}} \sum_{s=1}^T \sum_{i=1}^N \hat{z}_s (e_{it}^* e_{is}^* - E^*(e_{it}^* e_{is}^*)) \right\|^2 \\
&= \frac{1}{T} \sum_{t=1}^T \frac{1}{T} \sum_{s=1}^T \hat{z}'_s \hat{z}_s \frac{1}{N} \sum_{i=1}^N \hat{e}_{it}^2 \hat{e}_{is}^2 \text{Var}^*(\eta_{it} \eta_{is}) \leq \bar{\eta} \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T} \sum_{t=1}^T \hat{e}_{it}^2 \right) \left( \frac{1}{T} \sum_{s=1}^T \hat{z}'_s \hat{z}_s \hat{e}_{is}^2 \right) \\
&= \bar{\eta} \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T} \sum_{t=1}^T \hat{e}_{it}^2 \right) \left( \frac{1}{T} \sum_{s=1}^T \|\hat{z}_s\|^2 \hat{e}_{is}^2 \right) \leq \bar{\eta} \left[ \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T} \sum_{t=1}^T \hat{e}_{it}^2 \right)^2 \right]^{1/2} \left[ \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T} \sum_{s=1}^T \|\hat{z}_s\|^2 \hat{e}_{is}^2 \right)^2 \right]^{1/2} \\
&\leq \bar{\eta} \left[ \frac{1}{N} \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T \hat{e}_{it}^4 \right]^{1/2} \left[ \frac{1}{T} \sum_{s=1}^T \|\hat{z}_s\|^4 \frac{1}{N} \frac{1}{T} \sum_{i=1}^N \sum_{s=1}^T \hat{e}_{is}^4 \right]^{1/2}.
\end{aligned}$$

This expression is  $O_P(1)$  under our assumptions. Next consider Condition B\*(c). We have that

$$\begin{aligned}
E^* \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \hat{z}'_t \right\|^2 &= \frac{1}{TN} E^* \left\| \sum_{t=1}^T \left( \sum_{i=1}^N \tilde{\lambda}_i e_{it}^* \right) \hat{z}'_t \right\|^2 \\
&= \frac{1}{TN} \text{tr} \left( \sum_{t=1}^T \hat{z}_t \left( \sum_{i=1}^N \tilde{\lambda}'_i \tilde{\lambda}_i E^*(e_{it}^{*2}) \right) \hat{z}'_t \right) \\
&= \frac{1}{TN} \sum_{t=1}^T \hat{z}'_t \hat{z}_t \left( \sum_{i=1}^N \tilde{\lambda}'_i \tilde{\lambda}_i \hat{e}_{it}^2 \right) = \frac{1}{T} \sum_{t=1}^T \|\hat{z}_t\|^2 \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^2 \hat{e}_{it}^2 \\
&\leq \underbrace{\left( \frac{1}{T} \sum_{t=1}^T \|\hat{z}_t\|^4 \right)}_{=O_P(1)}^{1/2} \left( \frac{1}{T} \sum_{t=1}^T \left( \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^2 \hat{e}_{it}^2 \right)^2 \right)^{1/2}.
\end{aligned}$$

But by Cauchy-Schwartz,

$$\frac{1}{T} \sum_{t=1}^T \left( \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^2 \hat{e}_{it}^2 \right)^2 \leq \underbrace{\frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^4}_{O_P(1)} \underbrace{\frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \hat{e}_{it}^4}_{=O_P(1)} = O_P(1),$$

under our conditions. In particular, note that

$$\frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^4 \leq 2^3 \left( \frac{1}{N} \sum_{i=1}^N \|H^{-1'} \lambda_i\|^4 + \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i - H^{-1'} \lambda_i\|^4 \right)$$

and use Lemma C.1 (ii) with  $p = 4$  to bound the second term. The first term is bounded by the assumptions on  $\{\lambda_i\}$ . For Condition B\*(d), we have that

$$E^* \left\| \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \right\|^2 = E^* \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^N \tilde{\lambda}_i e_{it}^* \right\|^2 = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \tilde{\lambda}'_i \tilde{\lambda}_j \underbrace{E^*(e_{it}^* e_{jt}^*)}_{=0 \text{ for } i \neq j} = \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}'_i \tilde{\lambda}_i \hat{e}_{it}^2,$$



since  $Cov(\eta_{it}, \eta_{jt}) = 0$  and  $Var(\eta_{it}) = 1$ . Thus, Condition B\*(d) becomes

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T E^* \left\| \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \right\|^2 &= \frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}_i' \tilde{\lambda}_i \tilde{e}_{it}^2 = \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^2 \left( \frac{1}{T} \sum_{t=1}^T \tilde{e}_{it}^2 \right) \\ &\leq \left( \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^4 \right)^{1/2} \left( \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{T} \sum_{t=1}^T \tilde{e}_{it}^2 \right)^2 \right)^{1/2} \leq \left( \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^4 \right)^{1/2} \left( \frac{1}{N} \sum_{i=1}^N \frac{1}{T} \sum_{t=1}^T \tilde{e}_{it}^4 \right)^{1/2} = O_P(1), \end{aligned}$$

under our assumptions, by an application of Lemma C.1. For Condition B\*(e), we need to show that

$$A^* \equiv \frac{1}{T} \sum_{t=1}^T (\phi_t^* \phi_t^{*'} - E^*(\phi_t^* \phi_t^{*'})) = \frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \tilde{\lambda}_i \tilde{\lambda}_j' (e_{it}^* e_{jt}^* - E^*(e_{it}^* e_{jt}^*)) = o_P(1).$$

This expression has mean zero under the bootstrap measure by construction. So, it suffices to show that its variance tends to zero in probability. Take the case where the number of factors  $r$  is equal to 1, for simplicity. Then,

$$Var^*(A^*) = Var^* \left( \frac{1}{T} \sum_{t=1}^T (\phi_t^{*2} - E^*(\phi_t^{*2})) \right) = \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \frac{1}{N^2} \sum_{i,j,k,l} \tilde{\lambda}_i \tilde{\lambda}_j \tilde{\lambda}_k \tilde{\lambda}_l Cov^*(e_{it}^* e_{jt}^*, e_{is}^* e_{ks}^*).$$

Using the properties of the wild bootstrap, we can show that  $Cov^*(e_{it}^* e_{jt}^*, e_{is}^* e_{ks}^*) = 0$  if  $t \neq s$ , for any  $(i, j, k, l)$ , and

$$Cov^*(e_{it}^* e_{jt}^*, e_{it}^* e_{kt}^*) = \begin{cases} \tilde{e}_{it}^4 Var^*(\eta_{it}^2) & \text{if } i = j = k = l \\ \tilde{e}_{it}^2 \tilde{e}_{jt}^2, & \text{if } i = k \neq j = l \\ 0, & \text{otherwise} \end{cases}.$$

This implies that for some finite constant  $\bar{\eta}$ ,

$$\begin{aligned} Var^*(A^*) &= \frac{1}{T^2} \sum_{t=1}^T \frac{1}{N^2} \left( \sum_{i=1}^N \tilde{\lambda}_i^4 \tilde{e}_{it}^4 Var^*(\eta_{it}^2) + \sum_{i \neq j} \tilde{\lambda}_i \tilde{\lambda}_j \tilde{e}_{it}^2 \tilde{e}_{jt}^2 \right) \\ &\leq \bar{\eta} \frac{1}{T^2} \sum_{t=1}^T \left( \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}_i^2 \tilde{e}_{it}^2 \right)^2 \leq \bar{\eta} \frac{1}{T} \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}_i^4 \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \tilde{e}_{it}^4 \\ &= O_P \left( \frac{1}{T} \right) = o_P(1), \end{aligned}$$

provided  $\frac{1}{N} \sum_{i=1}^N \tilde{\lambda}_i^4 = O_P(1)$  and  $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \tilde{e}_{it}^4 = O_P(1)$ , which holds under our moment assumptions by an application of lemma C.1 with  $p = 4$ . Thus,  $\Gamma^* = \frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \tilde{\lambda}_i^2 \tilde{e}_{it}^2$  for the wild bootstrap. Condition F\* is satisfied because by Bai and Ng (2006),  $\Gamma^* \xrightarrow{P} Q\Gamma Q'$ .

Next, we verify Condition C\*. For  $t = 1, \dots, T - h$ , let  $\varepsilon_{t+h}^* = \hat{\varepsilon}_{t+h} v_{t+h}$ , where  $v_{t+h} \sim \text{i.i.d.}(0, 1)$ . Part a) follows as Condition B\*(b), using the independence between  $\varepsilon_{t+h}^*$  and  $e_{it}^*$ . For part b), we have

that

$$\begin{aligned} E^* \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}} \varepsilon_{t+h}^* \right\|^2 &= \frac{1}{TN} E^* \left\| \sum_{t=1}^{T-h} \left( \sum_{i=1}^N \tilde{\lambda}_i e_{it}^* \right) \varepsilon_{t+h}^* \right\|^2 \\ &= \frac{1}{TN} E^* \left\{ \left( \sum_{t=1}^{T-h} \sum_{s=1}^{T-h} \varepsilon_{t+h}^* \varepsilon_{s+h}^* \left( \sum_{i=1}^N \tilde{\lambda}'_i e_{it}^* \right) \left( \sum_{j=1}^N \tilde{\lambda}_j e_{js}^* \right) \right) \right\}. \end{aligned}$$

By the independence between  $\{e_{it}^*\}$  and  $\{\varepsilon_s^*\}$ , we have that

$$\begin{aligned} &= \frac{1}{TN} \left( \sum_{t=1}^{T-h} \sum_{s=1}^{T-h} \underbrace{E^* (\varepsilon_{t+h}^* \varepsilon_{s+h}^*)}_{=0 \text{ if } t \neq s} \sum_{i=1}^N \sum_{j=1}^N \tilde{\lambda}'_i \tilde{\lambda}_j \underbrace{E^* (e_{it}^* e_{js}^*)}_{=0 \text{ if } i \neq j \text{ or } t \neq s} \right) = \frac{1}{TN} \left( \sum_{t=1}^{T-h} E^* (\varepsilon_{t+h}^{*2}) \sum_{i=1}^N \tilde{\lambda}'_i \tilde{\lambda}_i E^* (e_{it}^{*2}) \right) \\ &= \frac{1}{T} \sum_{t=1}^{T-h} \hat{\varepsilon}_{t+h}^2 \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^2 \hat{e}_{it}^2 \leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \hat{\varepsilon}_{t+h}^4 \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{N} \sum_{i=1}^N \|\tilde{\lambda}_i\|^2 \hat{e}_{it}^2 \right)^2 \right)^{1/2} = O_P(1) \times O_P(1), \end{aligned}$$

under our assumptions and using the same arguments used above to show condition B\*(b). Part c) follows because  $\gamma_{st}^* = 0$  for  $t \neq s$  and by repeated application of Cauchy-Schwartz inequality, we have that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-h} \sum_{s=1}^T \tilde{F}_s \varepsilon_{t+h}^* \gamma_{st}^* &= \frac{1}{T} \sum_{t=1}^T \tilde{F}_t \varepsilon_{t+h}^* \gamma_{tt}^* = \frac{1}{T} \sum_{t=1}^{T-h} \tilde{F}_t \varepsilon_{t+h}^* \left( \frac{1}{N} \sum_{i=1}^N \hat{e}_{it}^2 \right) \\ &\leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \|\tilde{F}_t \varepsilon_{t+h}^*\|^2 \right)^{1/2} \left[ \frac{1}{T} \sum_{t=1}^{T-h} \left( \frac{1}{N} \sum_{i=1}^N \hat{e}_{it}^2 \right)^2 \right]^{1/2} \leq \underbrace{\left( \frac{1}{T} \sum_{t=1}^T \|\tilde{F}_t\|^4 \frac{1}{T} \sum_{t=1}^{T-h} \|\varepsilon_{t+h}^*\|^4 \right)^{1/4}}_{O_P(1)} \underbrace{\left[ \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \hat{e}_{it}^4 \right]^{1/2}}_{=O_P(1)}, \end{aligned}$$

under our assumptions, provided in particular that  $\frac{1}{T} \sum_{t=1}^{T-h} \varepsilon_{t+h}^{*4} = O_{P^*}(1)$  in probability. For this, it suffices that  $\frac{1}{T} \sum_{t=1}^{T-h} E^* (\varepsilon_{t+h}^{*4}) = O_P(1)$ . But by the properties of the wild bootstrap on  $\varepsilon_{t+h}^*$ , we have that

$$\frac{1}{T} \sum_{t=1}^{T-h} E^* (\varepsilon_{t+h}^{*4}) = \frac{1}{T} \sum_{t=1}^{T-h} \hat{\varepsilon}_{t+h}^4 \underbrace{E^* (v_{t+h}^4)}_{\leq C < \infty \text{ by assumption}} \leq \frac{1}{T} \sum_{t=1}^{T-h} \hat{\varepsilon}_{t+h}^4,$$

which is verified under our conditions. Finally, we verify Condition D\*.  $E^* (\varepsilon_{t+h}^*) = 0$  by construction. Moreover, we have that  $\frac{1}{T} \sum_{t=1}^{T-h} E^* |\varepsilon_{t+h}^*|^2 = \frac{1}{T} \sum_{t=1}^{T-h} \hat{\varepsilon}_{t+h}^2 \leq \left( \frac{1}{T} \sum_{t=1}^{T-h} \hat{\varepsilon}_{t+h}^4 \right)^{1/2} = O_P(1)$ , under our assumptions. So, part a) is verified. For part b) of Condition D\*, we need to verify that the bootstrap CLT result holds for the wild bootstrap. By the Cramer-Wold device, it suffices to show that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \ell' \underbrace{(\Omega^*)^{-1/2} \hat{z}_t \varepsilon_{t+h}^*}_{\equiv w_t^*} \rightarrow^{d^*} N(0, 1),$$

in probability for any  $\ell$  such that  $\ell' \ell = 1$ . Note that  $w_t^*$  is an heterogeneous array of independent random variables (given that  $\varepsilon_{t+h}^*$  is conditionally independent but heteroskedastic). Thus, we apply

a CLT for heterogeneous independent arrays. Note that  $E^*(w_t^*) = 0$  and

$$\begin{aligned} \text{Var}^* \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} w_t^* \right) &= \ell' (\Omega^*)^{-1/2} \text{Var}^* \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \hat{z}_t \varepsilon_{t+h}^* \right) (\Omega^*)^{-1/2} \ell \\ &= \ell' (\Omega^*)^{-1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t \hat{z}_t' \hat{\varepsilon}_{t+h}^2 \right) (\Omega^*)^{-1/2} \ell = 1. \end{aligned}$$

Thus, it suffices to verify Lyapunov's condition, i.e. for some  $r > 1$ ,  $\frac{1}{T^r} \sum_{t=1}^{T-h} E^* |w_t^*|^{2r} \rightarrow^P 0$ . Noting that  $w_t^* = \ell' (\Omega^*)^{-1/2} \hat{z}_t \varepsilon_{t+h}^*$ , we have that

$$\begin{aligned} \frac{1}{T^r} \sum_{t=1}^{T-h} E^* |w_t^*|^{2r} &= \frac{1}{T^{r-1}} \frac{1}{T} \sum_{t=1}^{T-h} \left| \ell' (\Omega^*)^{-1/2} \hat{z}_t \right|^{2r} |\hat{\varepsilon}_{t+h}|^{2r} \\ &\leq \frac{1}{T^{r-1}} \|\ell\|^{2r} \left\| (\Omega^*)^{-1/2} \right\|^{2r} \frac{1}{T} \sum_{t=1}^{T-h} \|\hat{z}_t\|^{2r} |\hat{\varepsilon}_{t+h}|^{2r} \underbrace{E^* |v_{t+h}|^{2r}}_{\leq M < \infty} \\ &\leq C \frac{1}{T^{r-1}} \left\| (\Omega^*)^{-1/2} \right\|^{2r} \left( \frac{1}{T} \sum_{t=1}^{T-h} \|\hat{z}_t\|^{4r} \right)^{1/2} \left( \frac{1}{T} \sum_{t=1}^{T-h} |\hat{\varepsilon}_{t+h}|^{4r} \right)^{1/2} = O_P \left( \frac{1}{T^{r-1}} \right) = o_P(1), \end{aligned}$$

by an application of Lemma C.1. Since  $\Omega^* = \frac{1}{T} \sum_{t=1}^{T-h} \hat{z}_t \hat{z}_t' \hat{\varepsilon}_{t+h}^2$ , under our assumptions  $\Omega^*$  converges to  $\Phi_0 \Omega \Phi_0' > 0$ , by Bai and Ng (2006). Thus, Condition E\* is satisfied. Condition F\* was verified in the main text.

## References

- [1] Bai, J., 2003. "Inferential theory for factor models of large dimensions," *Econometrica*, 71, 135-172.
- [2] Bai, J., 2009. "Panel data models with interactive fixed effects," *Econometrica*, 77, 1229-1279.
- [3] Bai, J. and S. Ng, 2002. "Determining the number of factors in approximate factor models," *Econometrica*, 70, 191-221.
- [4] Bai, J. and S. Ng, 2006. "Confidence intervals for diffusion index forecasts and inference with factor-augmented regressions," *Econometrica*, 74, 1133-1150.
- [5] Bai, J. and S. Ng, 2011. "Principal Components Estimation and Identification of the Factors", manuscript, Columbia University.
- [6] Boivin, J. and B. S. Bernanke, 2003, "Monetary policy in a data-rich environment", *Journal of Monetary Economics*, 50, 525-546.
- [7] Chamberlain, G. and M. Rothschild, 1983. "Arbitrage, factor structure and mean-variance analysis in large asset markets," *Econometrica*, 51, 1305-1324.
- [8] Connor, G. and R. A. Korajczyk, 1986. "Performance measurement with the arbitrage pricing theory," *Journal of Financial Economics*, 15, 373-394.
- [9] Connor, G. and R. A. Korajczyk, 1993. "A test for the number of factors in an approximate factor model," *Journal of Finance*, XLVIII, 1263-1291.
- [10] Eichengreen, B. A. Mody, M. Nedeljkovic, and L. Sarno, 2009. "How the Subprime Crisis Went Global: Evidence from Bank Credit Default Swap Spreads", NBER working paper 14904.
- [11] Gospodinov, N. and S. Ng, 2011. "Commodity prices, convenience yields and inflation," *Review of Economics and Statistics*, forthcoming.
- [12] Ludvigson, S. and S. Ng, 2007. "The empirical risk return relation: a factor analysis approach," *Journal of Financial Economics*, 83, 171-222.
- [13] Ludvigson, S. and S. Ng, 2009a. "Macro factors in bond risk premia," *Review of Financial Studies*, 22, 5027-5067.
- [14] Ludvigson, S. and S. Ng, 2009b. "A factor analysis of bond risk premia," *Handbook of Applied Econometrics*, forthcoming.
- [15] Onatski, A., 2011. "Asymptotics of the principal components estimator of large factor models with weakly influential factors," manuscript, University of Cambridge.

- [16] Shintani, M. and Z-Y Guo, 2011. “Finite sample performance of principal components estimators for dynamic factor models: asymptotic vs. bootstrap approximations,” manuscript, Vanderbilt University.
- [17] Stock, J. H. and M. Watson, 2002. “Forecasting using principal components from a large number of predictors,” *Journal of the American Statistical Association*, 97, 1167-1179.
- [18] Yamamoto, Y., 2011. “Bootstrap inference for impulse response functions in factor-augmented vector autoregressions,” manuscript, University of Alberta.

Table 1: Bias and coverage rate of 95% CIs for delta - Homoskedastic cases

		N = 50			N = 100			N = 200			
		T = 50	T = 100	T = 200	T = 50	T = 100	T = 200	T = 50	T = 100	T = 200	
<b>DGP 1</b> alpha = 0 homo, homo		Bias									
		bias	-0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
		estimate	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		WB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Coverage rate									
		OLS	94.0	95.4	95.0	94.5	95.7	94.2	95.7	94.4	94.1
		BC	91.0	93.1	91.9	93.1	94.0	93.3	95.1	93.8	93.7
		True factor	93.8	95.5	94.3	94.3	95.4	93.7	95.0	94.6	94.1
		WB	96.9	96.2	95.8	96.0	97.0	95.1	96.7	94.7	94.3
		<b>DGP 2</b> alpha = 1 homo, homo		Bias							
bias	-0.17			-0.14	-0.13	-0.11	-0.09	-0.08	-0.09	-0.06	-0.05
estimate	-0.09			-0.09	-0.10	-0.05	-0.05	-0.05	-0.03	-0.03	-0.03
WB	-0.12			-0.11	-0.10	-0.09	-0.07	-0.06	-0.07	-0.05	-0.04
Coverage rate											
OLS	71.1			66.0	50.7	84.7	83.1	79.3	88.3	89.8	88.1
BC	83.0			88.1	86.5	88.6	90.4	90.2	90.1	92.1	92.7
True factor	93.8			95.5	94.3	94.3	95.4	93.7	94.6	94.6	94.1
WB	90.9			92.7	90.7	93.8	93.7	92.0	93.8	94.5	94.3

Table 2. Bias in estimation of alpha - More general cases

		N = 50			N = 100			N = 200		
		T = 50	T = 100	T = 200	T = 50	T = 100	T = 200	T = 50	T = 100	T = 200
<b>DGP 3</b> alpha = 1 hetero, homo	bias	-0.16	-0.14	-0.13	-0.11	-0.09	-0.07	-0.09	-0.06	-0.05
	estimate	-0.09	-0.09	-0.10	-0.05	-0.05	-0.05	-0.03	-0.03	-0.03
	WB	-0.12	-0.11	-0.10	-0.09	-0.07	-0.06	-0.07	-0.05	-0.04
<b>DGP 4</b> alpha = 1 hetero, hetero	bias	-0.17	-0.15	-0.14	-0.12	-0.10	-0.08	-0.10	-0.06	-0.05
	estimate	-0.10	-0.10	-0.10	-0.05	-0.06	-0.06	-0.03	-0.03	-0.03
	WB	-0.13	-0.12	-0.11	-0.10	-0.08	-0.07	-0.08	-0.06	-0.04
<b>DGP 5</b> alpha = 1 hetero, AR+hetero	bias	-0.19	-0.16	-0.15	-0.13	-0.10	-0.08	-0.10	-0.07	-0.05
	estimate	-0.09	-0.10	-0.10	-0.05	-0.06	-0.06	-0.03	-0.03	-0.03
	WB	-0.13	-0.11	-0.11	-0.09	-0.08	-0.07	-0.08	-0.06	-0.04
<b>DGP 6</b> alpha = 1 hetero, CS+homo	bias	-0.23	-0.21	-0.20	-0.15	-0.13	-0.11	-0.11	-0.08	-0.07
	estimate	-0.11	-0.12	-0.12	-0.07	-0.07	-0.08	-0.04	-0.04	-0.04
	WB	-0.10	-0.09	-0.09	-0.08	-0.07	-0.06	-0.07	-0.05	-0.04

Table 3: Coverage rate of 95% CIs for delta - More general cases

		N = 50			N = 100			N = 200		
		T = 50	T = 100	T = 200	T = 50	T = 100	T = 200	T = 50	T = 100	T = 200
<b>DGP 3</b> alpha = 1 hetero, homo	OLS	63.5	57.3	45.6	76.5	78.4	76.2	80.6	85.0	86.0
	BC	81.5	84.3	84.0	83.1	87.6	88.0	83.5	89.6	89.6
	True factor	93.3	93.6	93.5	90.2	93.1	92.9	88.9	94.3	92.8
	WB	93.1	92.3	91.1	93.5	93.4	92.1	92.6	94.0	92.5
<b>DGP 4</b> alpha = 1 hetero, hetero	OLS	58.8	51.8	38.7	73.0	75.3	73.1	78.7	84.5	84.6
	BC	78.5	81.1	82.1	82.5	87.1	86.6	83.2	89.3	89.3
	True factor	93.3	93.6	93.5	90.2	93.1	92.9	88.9	94.3	92.8
	WB	93.7	92.8	91.8	92.9	93.8	92.3	92.6	93.6	93.1
<b>DGP 5</b> alpha = 1 hetero, AR + hetero	OLS	54.7	49.0	36.7	72.4	72.9	72.6	76.9	83.7	83.4
	BC	74.6	79.7	81.1	80.8	86.3	87.8	81.1	88.5	88.1
	True factor	93.3	93.6	93.5	90.2	93.1	92.9	88.9	94.3	92.8
	WB	91.2	90.8	91.1	91.8	92.8	92.4	91.3	93.1	92.3
<b>DGP 6</b> alpha = 1 hetero, CS + homo	OLS	42.2	30.6	13.7	65.8	65.2	54.7	76.1	79.5	78.3
	BC	68.3	67.2	61.4	77.5	81.7	81.1	80.0	87.9	87.4
	True factor	93.3	93.6	93.5	90.2	93.1	92.9	88.9	94.3	92.8
	WB	81.5	75.1	59.4	87.5	84.5	80.3	88.6	90.2	87.3