Estimation of Factor-Augmented Panel Regressions with Weakly Influential Factors

Joakim Westerlund
Simon Reese

February 2014
Revised: January 2015
ESTIMATION OF FACTOR-AUGMENTED PANEL REGRESSIONS WITH WEAKLY INFLUENTIAL FACTORS*

Simon Reese
Lund University

Joakim Westerlund†
Lund University
and
Financial Econometrics Group
Centre for Research in Economics and Financial Econometrics
Deakin University

January 27, 2015

Abstract

The use of factor-augmented panel regressions has become very popular in recent years. Existing methods for such regressions require that the common factors are strong, an assumption that is likely to be mistaken in practice. Motivated by this, the current paper offers an analysis of the effect of weak, semi-weak and semi-strong factors on two of the most popular estimators for factor-augmented regressions, namely, principal components (PC) and common correlated effects (CCE).

JEL Classification: C12; C13; C33.

Keywords: Non-strong common factors; factor-augmented panel regressions; common factor models.

1 Introduction

Consider the scalar and $m \times 1$ vector of observable panel data variables $y_{i,t}$ and $x_{i,t}$, where $i = 1, \ldots, N$ and $t = 1, \ldots, T$ index the cross-sectional and time series dimensions respectively.

* A previous version of the paper was presented at a seminar at Lund University. The authors would like to thank seminar participants, and in particular two anonymous referees for many valuable comments and suggestions. Westerlund thanks the Knut and Alice Wallenberg Foundation for financial support through a Wallenberg Academy Fellowship. Both authors thank the Jan Wallander and Tom Hedelius Foundation for financial support under research grant number P2014-0112-1.

† Corresponding author: Department of Economics, Lund University, Box 7082, 220 07 Lund, Sweden. Telephone: +46 46 222 8997. Fax: +46 46 222 4613. E-mail address: joakim.westerlund@nek.lu.se.
The data generating process (DGP) of the $T \times 1$ vector $y_i = (y_{i,1}, ..., y_{i,t})'$ is similar to the DGP of Bai (2009a), Greenaway-McGrevy et al. (2012), Pesaran (2006), and Westerlund and Urbain (2015), and is given by

$$y_i = X_i \beta_i + \eta_i,$$  

(1)

$$\eta_i = F \lambda_i + \nu_i,$$  

(2)

where $X_i = (x_{i,1}, ..., x_{i,t})'$ is $T \times m$, $\beta_i$ is a $m \times 1$ vector of slope coefficients, $F = (f_1, ..., f_T)'$ is a $T \times r$ matrix of unobservable common factors with $\lambda_i$ being the associated $r \times 1$ vector of factor loadings, and $\nu_i = (\nu_{i,1}, ..., \nu_{i,T})'$ is a $T \times 1$ vector of idiosyncratic errors.$^1$ In (1), $\beta_i$ is allowed to vary across the cross-section. Most existing work (see, for example, Bai, 2009a; Greenaway-McGrevy et al., 2012; Westerlund and Urbain, 2015), however, focuses on the case when $\beta_1 = ... = \beta_N = \beta$, and therefore so shall we. Hence, unless otherwise stated, $\beta_i$ is assumed to be homogenous, although we also consider the case when the slopes are not all equal.

The above model is the prototypical pooled panel regression with a factor error structure, in which $\nu_i$ is assumed independent of $X_i$. If $F$ is also independent of $X_i$, then (1) is nothing more than a static panel data regression with exogenous regressors. As is well known, such models can be consistently estimated using least squares (LS), which is true even if $F$ is unobserved and hence omitted in the estimation. If, however, $X_i$ is correlated with $F$, then consistency may be lost.$^2$ To allow for this possibility, we follow Pesaran (2006) and assume that

$$X_i = FA_i' + E_i,$$  

(3)

where $A_i$ is a $m \times r$ loading matrix and $E_i = (e_{i,1}, ..., e_{i,T})'$ is a $T \times m$ matrix of idiosyncratic errors. This specification of $X_i$ makes for a nontrivial estimation problem in the sense that $F$ can no longer be ignored, but has to be estimated and included in (1) as an additional regressor. By combining (1)–(3),

$$Z_i = FC_i + U_i,$$  

(4)

where $Z_i = (y_i, X_i) = (z_{i,1}, ..., z_{i,T})'$ is $T \times (m+1)$, $z_{i,t} = (y_{i,t}, x_{i,t}')'$ is $(m+1) \times 1$, $C_i = (A_i' \beta + \lambda_i, \Lambda_i')$ is $r \times (m+1)$, and $U_i = (u_{i,1}', ..., u_{i,T}')'$ is $(E_i\beta + \nu_i, E_i)$ is $T \times (m+1)$. Thus, (1)–(3) can be rewritten equivalently as a static factor model for $Z_i$, which is convenient because

$^1$If the model includes unit-specific fixed effects, then $y_i, X_i, \eta_i, F$ and $\nu_i$ are simply the correspondingly (time) demeaned variables.

$^2$One occasion when LS will remain consistent even if $X_i$ is correlated with $F$ is if the factor loadings in the equations for $\eta_i$ and $X_i$ are uncorrelated with zero expected values.
it means that \( \mathbf{F} \) can be estimated (up to a matrix rotation) using existing approaches for such models. The two main approaches are principal components (PC) (see Bai, 2009a; Greenaway-McGrevy et al., 2012) and common correlated effects (CCE) (see Pesaran, 2006).

Despite the generated regressor problem caused by the use of estimated rather than known factors, the CCE and PC estimators of \( \beta \) are \( \sqrt{NT} \)-consistent and asymptotically normal. This requires that the factors are strong, however, which of course need not be the case in practice, as when the factors are only weakly influential (see Chudik et al., 2011, page C58; Chudik and Pesaran, 2013, page 15, for some motivating examples). This case is discussed to some extent in a survey of Bai and Ng (2008), who use Monte Carlo simulation to show that the PC factors can be severely impaired when the factors are not strong (see also Boivin and Ng, 2006). They therefore reach the conclusion that “although much work has been accomplished in this research, much more remains to be done” (page 155). These findings recently motivated Chudik et al. (2011) to more formally consider the implications of weak, semi-weak and semi-strong factors on the CCE estimator of \( \beta \). The following quote summarizes their findings: “As predicted by the theory the CCE estimator performs well and show very little size distortions, in contrast with the iterated PC approach of Bai (2009a), which exhibit significant size distortions. The latter is partly due to the fact that in the presence of weak or semi-strong factors the PC estimates of the (rotated) factors need not be consistent. This problem does not affect the CCE estimator because it does not aim at consistent estimation of the factors but deals with error cross-section dependence generally by using cross-section averages to mop up such effects” (page C47).

The present paper is motivated by these findings. Our starting point is the preference of Chudik et al. (2011) to assume that the researcher knows with full certainty which factors that are strong and which factors that are not, and also that the non-strong factor component is uncorrelated with \( \mathbf{X}_i \), which means that, in analogy to the literature on omitted variables, it can be omitted without consequence. A more realistic scenario is that the researcher is unaware of both the strength and correlation of the factors, and therefore that some of the supposedly strong factors may potentially be non-strong and possibly correlated with \( \mathbf{X}_i \). Onatski (2012) considers the case when some of the included factors are semi-strong, but only within the context of PC estimation of a pure common factor model.\(^3\) According to his results, the presence

\(^3\)See Onatski (2012) for a brief review of some papers that have considered the presence of non-strong factors within the context of a pure common factor model.
of such factors causes the PC estimator to become inconsistent.

The purpose of the present study is to provide an analysis of the effects of the presence of non-strong factors on the CCE and PC estimators. In so doing we generalize the present literature in several directions. First, to the best of our knowledge, the present study is the first to consider estimation of factor-augmented regressions when some of the included factors are potentially non-strong and/or correlated with \( X_i \). Second, the study is the first to allow for a dependence between the strength of the factors and the relative expansion rate of \( N \) and \( T \), a feature that is shown to deliver significant insight. Third, the study is also the first to enable a direct comparison between the CCE and PC estimators when the factors are non-strong. Fourth, the study is the first to consider the properties of the PC estimator in case of a violation of the common slope assumption.

2 Assumptions

The conditions under which we will be working are summarized in Assumptions HOM, HET, ERR, LAM, RK–CCE, RK–PC and KAP. Here and throughout this paper \( tr(A) \), \( rk(A) \) and \( |A| = \sqrt{tr(A^tA)} \) denote the trace, rank and Frobenius (Euclidean) norm, respectively, of the matrix \( A \), and \( M < \infty \) is a generic positive number.

**Assumption HOM.** \( \beta_1 = \ldots = \beta_N = \beta \).

As mentioned in the introduction, the assumption of a common slope coefficient \( \beta \) is standard in the literature. One exception is Pesaran (2006) (see also Chudik at al., 2011), who allows some variation in \( \beta_i \) by requiring that it satisfies a random coefficient assumption. Because of this, our main analysis in Section 3.1, which is conducted under Assumption HOM, is in Section 3.2 complemented with an analysis under Assumption HET.

**Assumption HET.**

(i) \( \beta_i = \beta + \xi_i \), where \( \xi_i \) is iid with \( E(\xi_i) = 0, E(\xi_i\xi'_i) = \Sigma_\xi \) positive semi-definite, \( ||\beta|| \leq M \) and \( ||\Sigma_\xi|| \leq M \);

(ii) \( \xi_i \) is independent of \( \lambda_i, A_i, v_i, E_i \) and \( F \).

**Assumption ERR.**
(i) \( \nu_{i,t} \) is independently and identically distributed (iid) across both \( i \) and \( t \) with \( E(\nu_{i,t}) = 0, E(\nu_{i,t}^2) > 0 \) and \( E(\nu_{i,t}^4) \leq M \);

(ii) \( \epsilon_{i,t} \) is iid across both \( i \) and \( t \) with \( E(\epsilon_{i,t}) = 0, E(\epsilon_{i,t}\epsilon_{i,t}') = \Sigma_{\epsilon,i} \) positive definite and \( E(||\epsilon_{i,t}||^4) \leq M \);

(iii) \( f_t \) is covariance stationary such that \( E(||f_t||^4) \leq M \) and \( E(f_t f_t') = \Sigma_f \) is positive definite;

(iv) \( \nu, E \) and \( F \) are mutually and pairwise independent.

Assumption ERR is quite restrictive, but can be relaxed at the expense of added expositional and technical complexity. However, as argued by Chudik et al. (2011, page C53), this seems unnecessary in the present case, because the results regarding the strength of the factors are unlikely to be affected by the removal of nuisance parameters. The assumption that \( \nu_{i,t} \) is iid over time is, for instance, not necessary and can be relaxed by simply replacing \( \sigma_{\nu,i}^2 \) by \( \omega_{\nu,i}^2 = \sum_{s=-\infty}^{\infty} E(\nu_{i,t} \nu_{i,t-s}) \). Standard (parametric or kernel-based) estimators can be used if \( \omega_{\nu,i}^2 \) is unknown. Serial correlation in \( \epsilon_{i,t} \) can be accounted for in the same way.

Assumption LAM.

(i) \( \lambda_i = N^{-\alpha} \lambda_i^0 \) and \( \Lambda_i = N^{-\alpha} \Lambda_i^0 \), where \( \alpha \in [0,1] \);

(ii) \( \lambda_i^0 \) and \( \Lambda_i^0 \) are either random such that \( E(||\lambda_i^0||^4) \leq M \) and \( E(||\Lambda_i^0||^4) \leq M \), or non-random such that \( ||\lambda_i^0|| \leq M \) and \( ||\Lambda_i^0|| \leq M \);

(iii) \( \lambda_i^0 \) and \( \Lambda_i^0 \) are independent of \( \nu, E \) and \( F \).

Assumption LAM allows for very general cross-section dependencies. Indeed, in contrast to Onatski (2012), who only consider the case when \( \alpha = 1/2 \), the factors may be strong (\( \alpha = 0 \)), semi-strong (\( 0 < \alpha < 1/2 \)), semi-weak (\( 1/2 \leq \alpha < 1 \)), or weak (\( \alpha = 1 \)) (see Chudik et al., 2011, Definition 3.1, and the discussion that follows it). The fact that \( \alpha \) is the same for all loadings in \( \lambda_i \) and \( \Lambda_i \) is not very restrictive. If the loadings shrink to zero at different rates, then the results will be dominated by the slowest shrinking loading, and in such a case our results can

\[\text{lim}_{N \to \infty} N^{-\gamma} \sum_{i=1}^{\infty} ||\lambda_i|| < \infty.\] According to Chudik et al. (2011, condition (3.7)), the conditions for \( \lambda_i \) to be strong, semi-strong, semi-weak, and weak are given by \( \gamma = 1, 1/2 \leq \gamma < 1, 0 < \gamma < 1/2, \) and \( \gamma = 0 \). These conditions correspond to our classification via the relation \( \alpha = 1 - \gamma \). For ease of exposition, we deviate from Chudik et al. (2011) by counting \( \alpha = 1/2 \) to semi-weak factors.
be thought of as emanating from an analysis of the effects of the strongest factor. Indeed, as Chudik and Pesaran (2013, page 11) point out, in a multi-factor setup with differing rates of shrinking a can be thought of as representing their maximum.

In addition to Assumption LAM, we require that the fitted number of factors is no less than the true number, \( r \). Let us therefore denote by \( k \) the number of factors used in PC. Let \( \overline{C}^0 = N^{-1} \sum_{i=1}^{N} C_i^0 \) and \( \overline{Q}^0 = N^{-1} \sum_{i=1}^{N} C_i^0 C_i^{0'} = N^{-1} C^0 C^0', \) where \( C_i^0 = [(\Lambda_i^0)' \beta + \lambda_i^0, (\Lambda_i^0)'] \) is such that \( C_i = N^{-\delta} C_i^0 \) and \( C^0 = (C_1^0, \ldots, C_N^0)' \). Also, \( C^* = E(\overline{C}^0) \) and \( Q^* = E(\overline{Q}^0) \). The following assumptions are enough for our purposes.

**Assumption RK–CCE.** \( ||\overline{C}^0 - C^*|| = o_p(1) \), where \( rk(\overline{C}^0) = rk(C^*) = r \leq m + 1 \) and \( ||C^*|| \leq M \).

**Assumption RK–PC.** \( ||\overline{Q}^0 - Q^*|| = o_p(1) \), where \( rk(\overline{Q}^0) = rk(Q^*) = r \leq k \) and \( ||Q^*|| \leq M \).

Assumptions RK–CCE and RK–PC are similar to those employed by Pesaran (2006, equation (21)) and Bai (2009b, Section C.3), respectively. As Assumption RK–CCE makes clear, the number of common factors permitted within the CCE approach is bounded by the number of observables, \( m + 1 \). Specifically, the number of common factors is bounded by the number of observables that depend on \( F \). In the DGP considered here, \( F \) enters the equations of both \( y_i \) and \( X_i \), which means that there can be at most \( m + 1 \) factors. These restrictions make CCE less flexible than PC, in which \( k \) can be set arbitrarily by the researcher. That is, unlike in CCE, in PC the number of common factors is not bounded by the number of observables. Indeed, as Bai (2009a) shows, provided that it enters the equation for \( y_i \), in PC \( F \) need not enter the equation for \( X_i \). Assumption RK–CCE can be relaxed also within the CCE framework (see Pesaran, 2006). However, as Westerlund and Urbain (2013) show, this requires assuming the loadings \( \lambda_i \) and \( \Lambda_i \) to be mutually uncorrelated which, if false, renders the CCE estimator inconsistent. Hence, even if RK–CCE can in principle be relaxed, in most situations of practical relevance this is not the case. PC is therefore more general in this regard.

Chudik et al. (2011) assume the existence of \( r \leq m + 1 \) strong factors and \( n \) non-strong factors. Since the researcher is assumed to know the strength of the factors, the strong factors can be treated exactly as in Pesaran (2006). Moreover, by assuming that the non-strong factors are mean zero and independent of all other random elements of the model (including \( X_i \)), in analogy to the classical literature on omitted variables, the non-strong factors can be ignored,
provided that $n$ is fixed or increases at a slower rate than $N$ (see Chudik and Pesaran, 2013, Example 4, for an illustration of the effects of omitted factors when the factors are correlated with $X_t$).\footnote{In the classical literature the number of omitted variables is usually treated as fixed. A contribution of Chudik et al. (2011) over this literature is therefore the consideration of an increasing number of non-strong factors. However, since these factors are not allowed to be correlated with the explanatory variables, which is highly plausible in practice, this contribution is mainly technical.} Of course, in practice one never knows beforehand which factors that are strong and which factors that are not, and therefore the included factors are likely to have different strengths. Indeed, as Onatski (2012) shows, in some circumstances methods originally designed to estimate $r$ in the strong factor-only case remain valid also in the non-strong factor case. Hence, from a practical point of view the most relevant scenario is when strength of the included factors are unknown.

The assumption that $C^0$ ($Q^0$) has the same rank as $C^*$ ($Q^*$) is not necessary and can be relaxed at the cost of technical complexity. A minimal requirement in case of $C^0$ is that $P[rk(C^0) = rk(C^*)] \to 1$ as $N, T \to \infty$, which can be shown to hold under the conditions of the paper (see Chudik et al., 2011, Proof of Lemma A.2, for a formal argument).

For the sake of simplicity, in what follows we will frequently refer to Assumption RK–CCE/RK–PC with the understanding that Assumption RK–CCE (RK–PC) only applies to CCE (PC).

The types of factors that can be permitted are related to the relative expansion rate of $N$ and $T$. In order to capture relationships of this sort, we make use of the following assumption.

**Assumption KAP.** $T = N^\kappa$, where $\kappa > 0$.

Assumption KAP is less “flexible” than assuming that $T$ is proportional to $N^\kappa$. However, since the conclusions are qualitatively the same, and since assuming $T = N^\kappa$ greatly simplifies both transparency and notation, we opt for the less flexible specification in the present paper. Note in particular how Assumption KAP does not put any restrictions on the values taken by $\kappa$, provided that $\kappa > 0$. This is in contrast to Onatski (2012), who assumes that $N/T \to c > 0$, which under Assumption KAP is equivalent to requiring $\kappa = 1$. 
3 Results

3.1 Results under Assumption HOM

As is well known, since $F$ and $C_i$ are not separately identifiable, $F$ can only be estimated up to a matrix rotation. Depending on the assumed DGP, the CCE and PC estimators of the suitably rotated version of $F$ can be constructed in different ways. For example, while in Bai (2009a) $F$ is estimated by applying PC to the residuals from the LS fit of (1), in Greenaway-McGrevy et al. (2012), Kapetanios and Pesaran (2005), and Westerlund and Urbain (2015) the estimation is carried out by applying PC (and CCE) to $Z_i$. In the current setup, the estimators are applied to $Z_i$, in order to exploit the common factor structure of $X_i$. Hence, while the CCE estimator is just $\hat{F}_{\text{CCE}} = (\hat{f}_{\text{CCE}}^1, ..., \hat{f}_{\text{CCE}}^T)' = \bar{Z} = N^{-1} \sum_{i=1}^N Z_i$, the PC estimator, denoted $\hat{F}_{\text{PC}} = (\hat{f}_{\text{PC}}^1, ..., \hat{f}_{\text{PC}}^T)'$, is $\sqrt{T}$ times the matrix consisting of the eigenvectors corresponding to the $k$ largest eigenvalues of the $T \times T$ matrix $\hat{Z} \hat{Z}'$, where $\hat{Z} = (Z_1, ..., Z_N)$ is $T \times N(m+1)$.

The motivation for using panel estimation has always been, and continues to be, the increased accuracy that becomes available when the slope coefficients can be assumed to be homogenous. Most of the existing literature on factor-augmented regressions is therefore based on assuming that the slopes are indeed homogenous. Moreover, since under homogeneity the most efficient method of combining the data across the cross-section is “pooling”, as opposed to “mean grouping”, most estimators are pooled. In fact, in case of PC the literature has not yet ventured outside the homogeneous slope environment, and to the best of our knowledge all PC estimators in the literature are pooled (see, for example, Bai, 2009a; Greenaway-McGrevy et al., 2012). In this paper we therefore focus on the pooled CCE and PC estimators under Assumption HOM, although in Section 3.2 we also consider briefly the properties of these estimators under Assumption HET. In Appendix D we provide some results for both the individual and the group mean CCE estimators considered by Pesaran (2006).

The pooled factor-augmented CCE and PC estimators of $\beta$ have the following form:

$$\hat{\beta}_n^p = \left( \sum_{i=1}^N X_i' M_{F_n} X_i \right)^{-1} \sum_{i=1}^N X_i' M_{F_n} y_i,$$

where $n \in \{\text{CCE, PC}\}$ and the $p$ superscript signifies that the estimators are pooled. Furthermore, $M_{F_n} = I_T - \hat{F}_n S_{F_n}^{-1} \hat{F}_n'$, where $S_{F_n}$ is a regularized version of $\hat{F}_n' \hat{F}_n$, the definition of which is given in (A8) in Appendix A. Regularization is necessary to enable a formal derivation of the asymptotic results for CCE (PC) when $m+1 > r$ ($k > r$), although in small samples use of
\( S_{\hat{p}} \) instead of \( \hat{F}^p \hat{F}^n \) seem to have little or no effect. Coming back to (5), it is obvious that \( \hat{\beta}_n^p \) is simply the LS estimator of \( \beta \) with \( \hat{F}^n \) in place of \( F \). Theorem 1, which is our main result, gives the asymptotic distribution of \( \sqrt{NT}(\hat{\beta}_n^p - \beta) \). Before we come to the theorem, however, we need to introduce some notation. In particular, let us define

\[
\mathbf{b}_{n,i} = \mathbf{b}_{1n,i} - \mathbf{b}_{2n,i} - \mathbf{b}_{3n,i},
\]

where

\[
\begin{align*}
\mathbf{b}_{1CCE,i} &= \Lambda_i^0[(\mathbf{C}_i^0)']^{-1}\Sigma_i^0(\mathbf{C}_i^0)^{-1}\lambda_i^0, \\
\mathbf{b}_{2CCE,i} &= \Sigma_{e,i}(\beta, \mathbf{I}_m)(\mathbf{C}_i^0)^{-1}\lambda_i^0, \\
\mathbf{b}_{3CCE,i} &= \sigma_{e,i}^2\Lambda_i^0[(\mathbf{C}_i^0)^{-1}](1,0)',
\end{align*}
\]

for \( n = CCE \), and

\[
\begin{align*}
\mathbf{b}_{1PC,i} &= \Lambda_i^0(\mathbf{Q}_i^0)^{-1}\mathbf{S}_i^0(\mathbf{Q}_i^0)^{-1}\lambda_i^0, \\
\mathbf{b}_{2PC,i} &= \Sigma_{e,i}(\beta, \mathbf{I}_m)(\mathbf{Q}_i^0)^{-1}\lambda_i^0, \\
\mathbf{b}_{3PC,i} &= \sigma_{e,i}^2\Lambda_i^0((\mathbf{Q}_i^0)^{-1}1,0)',
\end{align*}
\]

for \( n = PC \). Theorem 1 is stated in terms of \( \mathbf{b}_n = N^{-1}\sum_{i=1}^N \mathbf{b}_{n,i} \quad \mathbf{W} = N^{-1}\sum_{i=1}^N \sigma_{e,i}^2\Sigma_{e,i} \) and \( \Sigma_e = N^{-1}\sum_{i=1}^N \Sigma_{e,i} \).

**Theorem 1.** Suppose that Assumptions HOM, ERR, LAM, RK–CCE/RK–PC and KAP hold, and that \( \alpha < 1/2 \). Suppose also that \( \kappa \in \mathcal{K}_C = (2\alpha, 3 - 4\alpha) \) for \( n = CCE \), and \( \kappa \in \mathcal{K}_P = (\max\{1/2 - \alpha, (4\alpha + 1)/3\}, 2 - \alpha) \) for \( n = PC \). Then, as \( N, T \to \infty \),

\[
\sqrt{NT}(\hat{\beta}_n^p - \beta) \to_d N\left(0, \lim_{N \to \infty} \Sigma^{-1}_e \mathbf{W} \Sigma^{-1}_e \right) + \lim_{N \to \infty} \Sigma^{-1}_e N^{(\kappa - 1)/2} \mathbf{b}_n,
\]

where \( \to_d \) signifies convergence in distribution.

The asymptotic distribution given in Theorem 1 is similar to those reported by Bai (2009a, Theorem 3) and Pesaran (2006, Theorem 4). In particular, we see that the asymptotic distribution consists of two terms; (i) a normal variate with mean zero and covariance matrix \( \lim_{N \to \infty} \Sigma^{-1}_e \mathbf{W} \Sigma^{-1}_e \), which is identical to the asymptotic distribution of the LS estimator of \( \beta \) when \( F \) is known, and (ii) a bias term involving \( N^{(\kappa - 1)/2} \mathbf{b}_n \) that depends on both \( \kappa \) and \( n \in \{CCE, PC\} \). Since the normal variate does not depend on \( n \), from a distributional point of view, the main difference between the estimators is the bias. Westerlund and Urban (2015)
compare the properties of the CCE and PC estimators under the strong factor assumption. They end up with similar bias expressions.\(^6\) Their analysis suggests that, except for some highly specialized cases, it is not possible to work out the relative bias, at least not analytically. The same is true in the present more general case. Of course, in this paper we are not interested in the magnitude of the different bias components per se, but rather in how the distributional results are affected by \(\alpha\) (and \(\kappa\)). The following observations regarding the impact of \(\alpha\) can be made.

- Both estimators require \(\alpha < 1/2\), that is, the factors can be at most semi-strong. If this is not the case, the asymptotic distributions of the CCE and PC estimators become dependent on nuisance parameters reflecting, among other things, the value of \(\alpha \geq 1/2\). This finding is in agreement with the results of Onatski (2012), showing that the PC estimator of \(F\) is inconsistent when \(\alpha = 1/2\).

- The ranges of values for \(\kappa\) that can be allowed under Theorem 1, \(K_{CCE}\) and \(K_{PC}\), depend critically on \(\alpha\). Consider the two extreme cases of \(\alpha = 0\) and \(\alpha = 1/2 - \nu\), where \(\nu > 0\) is a small number. When \(\alpha = 0\), \(K_{CCE} = (0, 3)\) and \(K_{PC} = (1/2, 2)\), whereas when \(\alpha = 1/2 - \nu\), then \(K_{CCE} = (1 - 2\nu, 1 + 4\nu)\) and \(K_{PC} = (1 - 4\nu/3, 3/2 + \nu)\). This result leads to the following two conclusions. First, \(K_{CCE}\) and \(K_{PC}\) get narrower when \(\alpha\) increases. Second, \(\alpha\) determines not only the width of \(K_{CCE}\) and \(K_{PC}\) but also their relative width; when \(\alpha = 0\), \(K_{CCE}\) is relatively wide, but this relation reverses when \(\alpha = 1/2 - \nu\) with \(\nu\) sufficiently small. Hence, when it comes to the allowable values of \(\kappa\), the choice of which estimator to use depends critically on \(\alpha\).

- When \(\alpha \to 1/2\), \(K_{CCE}\) and \(K_{PC}\) shrink towards the value one, suggesting that the weaker the factor the more restrictive the conditions placed on \(N\) and \(T\).

While theoretically a restriction, the assumed relationship between \(N\) and \(T\) is very relevant from an applied point of view. Indeed, since \(\kappa = \ln(T)/\ln(N)\) for any \(N\) and \(T\), as the following examples illustrate, the practical implications of Theorem 1 are very straightforward. Many empirical studies use data sets where \(N \approx T\), suggesting that \(\kappa \approx 1\) (see, for

\(^6\)In case of the PC estimator, the bias also appears in Theorem 3 of Bai (2009a). However, there is no bias in Theorem 4 of Pesaran (2006). The reason is that he assumes that \(T/N = N^{k-1} = o(1)\), in which case the bias in our Theorem 1 is negligible. In this sense, Theorem 1 generalizes Theorem 4 of Pesaran (2006) in two directions; (i) it allows for non-strong factors, and (ii) there is no requirement that \(N/T\) should go to zero.
example, Cavalcanti et al., 2011; Bertoli and Fernández-Huertas Moraga, 2012; Calderon et al., 2014; Herzer et al., 2012). As Theorem 1 makes clear, in terms of the allowable values of \( \alpha \), when \( \kappa = 1 \) there is no difference between the CCE and PC estimators. The choice of which estimator to use in this case is therefore mainly a matter of personal preference. However, there are also studies where \( N \) and \( T \) are quite different, in which case \( \kappa \neq 1 \). In Eberhardt et al. (2013), for example, \( N = 119 \) and the average \( T \) (over the unbalanced cross-section) is 22.2, implying that \( \kappa \approx 0.65 \). Since in this case \( \kappa > 1/2 \), when it comes to the allowable values of \( \alpha \), the PC estimator is clearly the preferred choice. By contrast, in studies such as in Arnold et al. (2013), where \( T > N \), the CCE estimator is preferred, for in this case the set of allowable values of \( \alpha \) is larger for CCE than for PC.

In order to illustrate the implications of Theorem 1 in Table 1 we report some Monte Carlo results for the simulated means and standard deviations (STDs) of \( \sqrt{N/T} (\hat{\beta}_{PC}^\rho - \beta) \) and \( \sqrt{N/T} (\hat{\beta}_{CCE}^\rho - \beta) \), and their theoretical predictions. The DGP is a restricted version of the one given in (4), and sets \( m = 1, r = 2 \) and \( (f', v_{ij}, e_{ij})' \sim N(0, \text{diag}(\Sigma_f, \sigma^2_{v_{ij}}, \Sigma_{e_{ij}})) \), where \( \Sigma_f = 4j_2, \sigma^2_{v_{ij}} = 4 \) and \( \Sigma_{e_{ij}} = I_m \). While the variation in \( \lambda^0_i \) and \( \Lambda^0_i \) is not unimportant, for the purpose of illustration it is enough to consider a single change.\(^7\) We therefore set \( \lambda^0_i = \lambda^* e_1 + [0.5 - 1(i > \lfloor 0.5N \rfloor)] t_2 \), where \( e_1 = (1, 0)' \), \( t_2 = (1, 1)' \), \( 1(A) \) is the indicator function for the event \( A \), and \( \lfloor x \rfloor \) is the integer part of \( x \). Thus, if \( \lambda^* = 0 \), then \( \lambda^0_i \) is 0.5\( t_2 \) for the first \( \lfloor 0.5N \rfloor \) units and \(-0.5t_2 \) for the rest. The corresponding specification of \( (\Lambda^0_i)' \) is given by \( (\Lambda^0_i)' = \Lambda^* e_2 + [0.5 - 1(i > \lfloor 0.5N \rfloor)] t_2 \), where \( e_2 = (0, 1)' \). In order to ensure that Assumptions RK–CCE and RK–PC are satisfied, in this experiment we set \( \lambda^* = 1 \) and \( \Lambda^* = 0.4 \). All results are based on 3,000 replications. As expected, we see that the asymptotic theory provides very accurate approximations of actual behavior when \( \alpha \in \{0, 1/4\} \). This is true regardless of the value of \( T \). However, this is no longer the case when \( \alpha = 1/2 \). Specifically, while the variance predictions are still very accurate, the predicted means are way off target. A similar pattern emerges when looking at the relative performance of the two estimators. Specifically, while the STDs tend to be very close, there is a difference in bias. In the particular DGP considered here CCE is more biased than PC, but it is possible to find other DGPs in which it is the other way around. The results therefore give little or no practical guidance, except that the one should be careful not to assume that all factors are strong when in fact some (or indeed all)

\(^7\)While the results do depend to some extent on the variation in \( \lambda^0_i \) and \( \Lambda^0_i \), in terms of effect of \( \kappa \) and \( \alpha \), the conclusions do not depend on how \( \lambda^0_i \) and \( \Lambda^0_i \) are generated.
of the factors may be weak.

Table 1: Monte Carlo evaluation of the asymptotic distributions of the pooled PC and CCE estimators.

<table>
<thead>
<tr>
<th>κ</th>
<th>α</th>
<th>PC</th>
<th>CCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Theory</td>
</tr>
<tr>
<td>T = 100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/4</td>
<td>0</td>
<td>-1.0957</td>
<td>-1.1231</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>-0.6389</td>
<td>-0.6589</td>
</tr>
<tr>
<td>3/4</td>
<td>1/4</td>
<td>-1.1591</td>
<td>-1.1231</td>
</tr>
<tr>
<td>1</td>
<td>1/4</td>
<td>-0.6850</td>
<td>-0.6589</td>
</tr>
<tr>
<td>3/4</td>
<td>1/2</td>
<td>-1.4381</td>
<td>-1.1231</td>
</tr>
<tr>
<td>1</td>
<td>1/2</td>
<td>-1.0813</td>
<td>-0.6589</td>
</tr>
</tbody>
</table>

| T = 200 |
| 3/4 | 0   | -1.2607 | -1.2746 | 1.9765 | 2.0306 | 5.0853 | 5.2779 | 2.0475 | 2.0306 |
| 1   | 0   | -0.6088 | -0.6749 | 2.0100 | 2.0286 | 2.5494 | 2.5138 | 2.0302 | 2.0286 |
| 3/4 | 1/4 | -1.3176 | -1.2746 | 1.9731 | 2.0306 | 4.8848 | 5.2779 | 2.0457 | 2.0306 |
| 1   | 1/4 | -0.6444 | -0.6749 | 2.0081 | 2.0286 | 2.4964 | 2.5138 | 2.0304 | 2.0286 |
| 3/4 | 1/2 | -1.6379 | -1.2746 | 1.9545 | 2.0306 | 3.8112 | 5.2779 | 2.0349 | 2.0306 |
| 1   | 1/2 | -1.0835 | -0.6749 | 2.0327 | 2.0286 | 1.9321 | 2.5138 | 2.0291 | 2.0286 |

Notes: “Mean” and “STD” refer to the simulated mean and standard deviation of $\sqrt{NT(\hat{\beta}_n^p - \beta)}$, respectively. The theoretical predictions, reported as “Theory”, are obtained by simulating the asymptotic distributions given in Theorem 1.

Theorem 1 implies that the two estimators are consistent and that the rate of consistency is $(NT)^{-1/2}N^{(\kappa - 1)/2} = N^{-1}$. However, this requires that $\alpha < 1/2$ and $\kappa \in K_{CCE}$ and/or $\kappa \in K_{PC}$, conditions that are necessary for deriving the asymptotic distributions, but that might not be required for consistency. In the following corollary to Theorem 1 we report necessary conditions for consistency.

Corollary 1. Suppose that Assumptions HOM, ERR, LAM, RK–CCE/RK–PC and KAP hold. Suppose also that $\kappa > 2\alpha - 1$ with $\alpha < 1$ for $n = CCE$, and $\kappa > \max\{2\alpha, 4\alpha - 1\}$ for $n = PC$. Then, as $N, T \to \infty$,

$$||\hat{\beta}_n^p - \beta|| = o_p(1).$$

Corollary 1 suggests that there is a trade-off between the allowable values of $\alpha$ and the restrictions placed on $\kappa$. On the one hand, since $2\alpha - 1 \leq \max\{2\alpha, 4\alpha - 1\}$ for all $\alpha \in [0, 1)$, PC is relatively demanding when it comes to the values of $\kappa$ needed for consistency. On the other
hand, if $\kappa$ is just large enough, PC is consistent even when $\alpha = 1$, which is not the case for CCE. Of course, since the relevant condition on $\kappa$ in this case is given by $\kappa > 3$, allowing $\alpha = 1$ comes at a high cost in terms of both bias (as $N^{(\kappa-1)/2}$ is increasing in $\kappa$) and the required size of $T$ (in the sense that $T \gg N$).

The finding that the PC estimator is consistent even when $\alpha \geq 1/2$ would seem to go against the results of Onatski (2012, Theorem 1) showing how the PC estimator of $F$ is rendered inconsistent when $\alpha = 1/2$. However, this is actually not the case. The difference lies with the current more flexible specification of $\alpha$ and $\kappa$. Indeed, if we assume that $\kappa = 1$, which under Assumption KAP is equivalent to Onatski’s (2012) requirement that $N/T \rightarrow c > 0$, then, as Corollary 1 makes clear, the PC estimator is inconsistent for all $\alpha \leq 1/2$ (including $\alpha = 1/2$). The results reported herein are therefore consistent with those of Onatski (2012).

As an illustration of Corollary 1, we collect the mean and STD of $(\hat{\beta}_{CCE}^p - \beta)$ over 3,000 simulated samples and report them in Table 2. We use the same DGP as before but now with the following parameter values: $\Sigma_f = 10j_{2r}$, $\sigma_{e,j}^2 = 10$, $\Sigma_{e,i} = 10^{-4}I_m$ and $\kappa = 1$. The factor loadings are obtained by setting $\lambda^* = 1$, $\Lambda^* = 10$ and by increasing the variation around $\lambda^*e_1$ and $\Lambda^*e_2$ by a factor of 20 relative to the previous DGP. For the purpose of demonstrating the requirements for convergence, it is convenient to look at the STDs as a function of the sample size. For $\alpha < 1$, the values decrease constantly, but the closer $\alpha$ is to one, the slower this decrease becomes. As expected, the STDs stagnate when $\alpha = 1$ and they even begin to increase if the value of $\alpha$ is pushed above one.

### 3.2 Results under Assumption HET

As mentioned in Section 3.1, in the strong factor case, while the CCE strand of the literature has considered both Assumptions HOM and HET (see, for example, Pesaran, 2006; Chudik et al., 2011), the PC strand has only considered Assumption HOM (see, for example, Bai, 2009a; Greenaway-McGrevy et al., 2012; Westerlund and Urbain, 2015). In this section we study the pooled CCE and PC estimators under Assumption HET. Hence, even in the strong factor case, the results reported here represents an extension of the existing PC results under HOM.

**Theorem 2.** Suppose that Assumptions HET, ERR, LAM, RK–CCE and KAP hold. Suppose also that $\kappa > \max\{\alpha, 2\alpha - 1/2\}$ with $\alpha < 3/4$ for $n = \text{CCE}$, and $\kappa > \max\{4\alpha - 2, \alpha + 1/4, 1/2 - 2\alpha\}$ with
<table>
<thead>
<tr>
<th>N/α</th>
<th>Mean</th>
<th>STD</th>
<th>Relative STD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5</td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>1.0005</td>
<td>1.0004</td>
<td>1.0003</td>
</tr>
<tr>
<td>100</td>
<td>1.0001</td>
<td>1.0003</td>
<td>1.0003</td>
</tr>
<tr>
<td>150</td>
<td>1.0003</td>
<td>1.0006</td>
<td>1.0007</td>
</tr>
<tr>
<td>200</td>
<td>1.0002</td>
<td>1.0004</td>
<td>1.0004</td>
</tr>
<tr>
<td>250</td>
<td>1.0000</td>
<td>0.9993</td>
<td>0.9989</td>
</tr>
<tr>
<td>300</td>
<td>0.9999</td>
<td>0.9994</td>
<td>0.9990</td>
</tr>
<tr>
<td>350</td>
<td>0.9999</td>
<td>0.9988</td>
<td>0.9981</td>
</tr>
<tr>
<td>400</td>
<td>0.9998</td>
<td>0.9993</td>
<td>0.9987</td>
</tr>
<tr>
<td>450</td>
<td>0.9998</td>
<td>0.9986</td>
<td>0.9976</td>
</tr>
<tr>
<td>500</td>
<td>0.9999</td>
<td>0.9987</td>
<td>0.9976</td>
</tr>
<tr>
<td>550</td>
<td>1.0001</td>
<td>0.9986</td>
<td>0.9973</td>
</tr>
<tr>
<td>600</td>
<td>1.0000</td>
<td>0.9987</td>
<td>0.9972</td>
</tr>
</tbody>
</table>

Notes: “Mean” and “STD” refer to the simulated mean and standard deviation of $\hat{\beta}_{CCE} - \beta$, respectively. “Relative STD” refers the ratio of the standard deviation corresponding to the given value of $N$ and that of a sample with $N = 50$. 

Table 2: Monte Carlo evaluation of the consistency of the pooled CCE estimator.
\( \alpha < 7/4 \) for \( n = \text{PC} \). Then, as \( N, T \to \infty \),

\[
\sqrt{N}(\hat{\beta}_n^p - \beta) \to_d N \left( 0, \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^N \sum_{e,i} \sum_{e,i} \right) \Sigma_e^{-1}.
\]

Some remarks are in order.

- According to Theorems 1 and 2, the heterogeneity of \( \beta_1, \ldots, \beta_N \) leads to a slower rate of convergence relative to the case when \( \beta_1 = \ldots = \beta_N \), from \( \sqrt{NT} \) to \( \sqrt{N} \). This is consistent with the results of Pesaran (2006, Theorems 2 and 3).

- Unlike what happens under Assumption HOM, under HET \( \hat{\beta}_C^p \) and \( \hat{\beta}_P^p \) are asymptotically unbiased even if \( \kappa > 1 \). Note in particular how in the strong factor case asymptotic unbiasedness holds for all (permissible) \( \kappa > 0 \). This finding is consistent with those of Pesaran (2006, Theorems 2 and 3) and Chudik et al. (2011, Section 4), where in the latter study all non-strong factors are omitted from the estimation process.

- Interestingly, the set of admissible combinations of \( \alpha \) and \( \kappa \) is actually larger under Assumption HET than under Assumption HOM. That is, relaxing the homogeneity assumption also means relaxing the restrictions placed \( \alpha \) and \( \kappa \). For example, in terms of \( \alpha \), while under HOM the relevant requirement for asymptotic normality is that \( \alpha < 1/2 \), under HET it is \( \alpha < 3/4 \) for \( \hat{\beta}_C^p \) and \( \alpha < 7/4 \) for \( \hat{\beta}_P^p \). That is, while under HOM the factors have to be at most semi-strong, under HET we can also allow for semi-weak factors. The “price” of the increased generality under HET is of course the reduced rate of consistency mentioned above.

Corollary 2 provides the conditions on \( \alpha \) and \( \kappa \) required to ensure consistency under Assumption HET.

**Corollary 2.** Suppose that Assumptions HET, ERR, LAM, RK–CCE and KAP hold. Suppose also that \( \kappa > 2\alpha - 1 \) with \( \alpha < 1 \), and \( \kappa > \max\{2\alpha, 4\alpha - 1\} \) for \( n = \text{PC} \). Then, as \( N, T \to \infty \),

\[
||\hat{\beta}_n^p - \beta|| = o_p(1).
\]

The conditions placed on \( \kappa \) and \( \alpha \) are therefore the same as in Corollary 1 (under Assumption HOM).
4 Conclusion

The present study considers a factor-augmented regression model in which the factor loadings go to zero at the rate $N^{-\alpha}$, where $\alpha \in [0, 1]$ and $N$ is related to $T$ via $T = N^\kappa$. The purpose is to study the effect of $\alpha$ on two of the most popular estimators for factor-augmented regressions, namely, PC and CCE. A standard assumption in the literature is that the slopes are homogenous. Our findings in this case can be summarized as follows. First, in order to ensure $\sqrt{NT}$-consistency and limiting normal distributions, both estimators require strong or semi-strong factors ($\alpha < 1/2$). The sets of allowable values of $\kappa$ generally differ between the estimators. However, both sets shrink toward the value one as the factors become weaker ($\alpha \to 0$). Second, unless $\kappa < 1$, both estimators are asymptotically biased. Third, both estimators can be consistent when the factors are semi-weak ($\alpha < 1$), but only PC allows consistent estimation in the weak factor case ($\alpha = 1$). Additional requirements on $\kappa$ have to be fulfilled for consistency to be possible. CCE is very general in this regard, having binding restrictions on $\kappa$ only for semi-weak factors. By contrast, PC restricts $\kappa$ from below already for semi-strong factors and requires that $\kappa > 3$ when the factors are weak, which is clearly very restrictive.

While the PC strand of the literature on factor-augmented regressions assumes that the slopes are homogenous, in the CCE strand there has been some attempts to relax this assumption. Motivated by this development, we also consider the properties of the CCE and PC estimators when the heterogeneity of the slopes can be given a random coefficient representation. From a qualitative point of view, our results for the heterogenous slope case are very similar to those obtained under homogenous slopes. The main difference is threefold. First, the estimators are asymptotically unbiased for all (permissible) values of $\kappa$, including $\kappa \geq 1$. Second, quite unexpectedly, the required restrictions on $\kappa$ and $\alpha$ to ensure asymptotic normality (and also consistency in case of the PC estimator) are less restrictive when the slopes are allowed to be heterogeneous than when they are restricted to be homogenous. Third, the increased generality in terms of $\kappa$ and $\alpha$ under heterogeneous slopes has a “price” in terms of a relatively slow rate of consistency.
References


A Notation

The model for \( Z_{i,t} = (y_{i,t}, x'_{i,t})' \) can be written in matrix notation as

\[
Z_i = FC_i + U_i, \tag{A1}
\]

where \( Z_i = (y_i, X_i) = (z_{i,1}, ..., z_{i,T})' \) is \( T \times (m + 1) \), \( F = (f_1, ..., f_T)' \) is \( T \times r \), \( C_i = (\Lambda_i' \beta + \lambda_i, \Lambda_i') \) is \( r \times (m + 1) \) and \( U_i = (u_{i,1}, ..., u_{i,T})' = (E_i \beta + v_i, E_i) \) is \( T \times (m + 1) \). Alternatively, the model for \( z_{i,t} \) can be written as the following \( N \)-dimensional system:

\[
z_t = C f_t + u_t, \tag{A2}
\]

where \( z_t = (z'_{1,t}, ..., z'_{N,t})' \) and \( u_t = (u_{1,t}, ..., u_{N,t})' \) are \( N(m + 1) \times 1 \), and \( C = (C_1, ..., C_N)' \) is \( N \times (m + 1) \times r \). The matrix notation

\[
Z = FC' + U \tag{A3}
\]

will also be used, where \( Z = (Z_1, ..., Z_N) \) and \( U = (U_1, ..., U_N) \) are \( T \times N(m + 1) \). In what follows the representations in (A1)–(A3) will be used interchangeably.

Many of the results can be expressed in terms of \( D^{CCE} = \hat{F}^{CCE} - FC \) and \( D^{PC} = \hat{F}^{PC} - F \), where \( \bar{C} = N^{-1} \sum_{i=1}^{N} C_i = N^{-a} \bar{C}^0 \), \( \bar{H} = \bar{Q}(T^{-1}F'\hat{F}^{PC}) V_T^{-1} = N^{-2a} \bar{H}^0 \), \( \bar{Q} = N^{-1} \sum_{i=1}^{N} C_i C_i' = N^{-1} \bar{C}' \bar{C} \) and \( \bar{H}^0 = \bar{Q} (T^{-1}F'\hat{F}^{PC}) V_T^{-1} \). It will therefore be convenient to introduce some special notation to simplify such expressions. Consider the PC estimator. As in Bai (2003, page 158), if we denote by \( V_T \) the \( k \times k \) diagonal matrix consisting of the first \( k \) eigenvalues of \((NT)^{-1}ZZ'\) in descending order, then, by the definition of eigenvalues and eigenvectors, \( \hat{F}^{PC} = (NT)^{-1}ZZ' \hat{F}^{PC} V_T^{-1} \). In this notation,

\[
D^{PC} = \hat{F}^{PC} - F \bar{H} = (NT)^{-1}ZZ' \hat{F}^{PC} V_T^{-1} - (NT)^{-1}FC'CF' \hat{F}^{PC} V_T^{-1} = (NT)^{-1}(ZZ' - FC'CF') \hat{F}^{PC} V_T^{-1} = (NT)^{-1}(UU' + UCF' + FC'U') \hat{F}^{PC} V_T^{-1} = (NT)^{-1}G \hat{F}^{PC} V_T^{-1}, \tag{A4}
\]

where \( G = (UU' + UCF' + FC'U') = (g_1, ..., g_T)' \) is \( T \times T \) and \( g_t = (U u_t + FC' u_t + UC f_t) \) is \( T \times 1 \). Note that the dimension of \( D^{PC} \) is \( T \times (m + 1) \). It is further convenient to write \( D^{PC} = (d^{PC}_1, ..., d^{PC}_T)' \), where

\[
d^{PC}_t = \hat{f}^{PC}_t - \bar{H}' f_t = (NT)^{-1}V_T^{-1}(\hat{F}^{PC})'(U u_t + FC' u_t + UC f_t) = (NT)^{-1}V_T^{-1}(\hat{F}^{PC})' g_t \tag{A5}
\]
is \((m + 1) \times 1\). The corresponding quantities in case CCE are given by

\[
\begin{align*}
D_{\text{CCE}} &= \hat{f}_{\text{CCE}} - \hat{C} = U, \\
d_i^{\text{CCE}} &= \hat{f}_i^{\text{CCE}} - \hat{C} f_i = \hat{u}_i.
\end{align*}
\]

(A6)  
(A7)

It is important to note that \(\hat{C}^0\) is not square, but of dimension \(r \times (m + 1)\). Thus, since \(rk(\hat{C}^0) = r \leq (m + 1)\), the \(r \times r\) matrix \(\hat{C}^0(\hat{C}^0)'\) is nonsingular. Let us therefore denote by \(A^-\) the Moore–Penrose inverse of the matrix \(A\). Note in particular that if \(A\) has full row rank, then \(A^- = A'(AA')^{-1}\), whereas if \(A\) has column row rank, then \(A^- = (A'A)^{-1}A'.\) Thus, since \(\hat{C}^0\) has full row rank, we have \((\hat{C}^0)^- = (\hat{C}^0)'(\hat{C}^0(\hat{C}^0)')^{-1},\) such that \(\hat{C}^0(\hat{C}^0)^- = I_r\), and \(\hat{C}\hat{C}^- = I_r\).

A similar definition applies to \((H^0)^-\).

Consider \(M_{\hat{F}_n} = I_T - \hat{F}_n(S_{\hat{F}_n})^{-1}\hat{F}_n'.\) As explained in the main text, \(S_{\hat{F}_n}\) is a regularized version of \(\hat{F}_n'\hat{F}_n\). We now characterize this matrix. By the singular value decomposition, \(\hat{F}_n'\hat{F}_n = V_n\Delta_n U_n'\), where in case of CCE (PC) \(\Delta_n\) is a diagonal matrix containing the \(m + 1\) (\(k\)) singular values of \(\hat{F}_\text{CCE} (\hat{F}_\text{PC})\) in decreasing order. The regularized version of \(\hat{F}_n'\hat{F}_n\) is given by

\[
S_{\hat{F}_n} = V_n \hat{\Delta}_n U_n',
\]

(A8)

where

\[
\hat{\Delta}_n = \begin{bmatrix} \Delta_n & 0 \\ 0 & 0 \end{bmatrix},
\]

with \(\Delta_n\) being the \(r \times r\) diagonal matrix containing the \(r\) largest values of \(\Delta_n\). In case of CCE (PC), \(S_{\hat{F}_n} = \hat{F}_n'\hat{F}_n\) if \(m + 1 = r\) (\(k = r\)), whereas if \(m + 1 > r\) (\(k > r\)), then \(S_{\hat{F}_n}\) satisfies the rank condition of Andrews (1987, Theorem 2). This means that \(S_{\hat{F}_n}\) is “well-behaved” in the sense that \((T^{-1}S_{\hat{F}_n})^- \to_p (\lim_{N,T \to \infty} T^{-1}F_n'F_n)^-\).

For notational simplicity, \(R\) will henceforth be used to indicate the order of lengthy remainder terms. For example, if \(A = O_p(N^n) + O_p(T^b)\), then we write \(A = O_p(R),\) where \(R = N^n + T^b\) (with the dependence on \(N\) and \(T\) in \(R\) omitted).
B Auxiliary lemmas

In Lemma PC1 we show how Assumption RK–PC can be used to show that $||\bar{H}^0 - H^*||^2 = o_p(1)$, where $\bar{H}^0$ is the relevant rotation matrix in case of PC. The corresponding rotation matrix in case of CCE is given by $\bar{C}^0$, for which we already know from Assumption RK–CCE that $||\bar{C}^0 - C^*||^2 = o_p(1)$. There is therefore no “Lemma CCE1”. All other lemmas come in pairs; for every CCE lemma there is a corresponding PC lemma.

**Lemma PC1.** Under Assumptions ERR, LAM, RK–PC and KAP, with $\kappa > \max\{2\alpha, 4\alpha - 1\}$,

$$||\bar{H}^0 - H^*||^2 = o_p(1),$$

where $H^* = (Q^*)^{1/2}SV^{-1/2}$, $V = \text{diag}(v_1, ..., v_r)$, $v_1 > ... > v_r > 0$ are the eigenvalues of $\Sigma^0 = (Q^*)^{1/2}\Sigma_F(Q^*)^{1/2}$, and $S$ is the associated matrix of eigenvectors.

**Proof of Lemma PC1**

The proof of Lemma PC1 is similar to that of Theorem 2 in Stock and Watson (1999). Here we focus on the case when $k = r$; the proof for the case when $k > r$ is simpler and follows by the same arguments as in Stock and Watson (1999, page 47).

We begin by evaluating $T^{-1}\bar{F}^{PC}F$. As will soon become clear this can be done indirectly by evaluating the limit of the PC objective function, which we write as $tr[(Z - F^*C')(Z - F^*C')']$, where $F^*$ is the chosen value for $F$. Using the normalization $T^{-1}F^*F^* = I_r$, we can concentrate the objective function with respect to $C = ZF^*(F^*F^*)^{-1} = T^{-1}Z'F^*$, leading to the concentrated objective function

$$tr(Z'M_{F^*}Z) = tr(Z'Z) - tr(Z'P_{F^*}Z),$$

where $P_A = A(A'A)^{-1}A'$ for any matrix $A$. Hence, minimizing $(NT)^{-1}tr[(Z - F^*C')(Z - F^*C')']$ is equivalent to maximizing $Q(F^*) = (NT)^{-1}tr(Z'P_{F^*}Z)$. Substitution of $Z = FC' + U$ into $Q(F^*)$ gives

$$Q(F^*) = (NT)^{-1}tr(Z'P_{F^*}Z) = (NT)^{-1}tr((FC' + U)'P_{F^*}(FC' + U)) = (NT)^{-1}[tr(CF'P_{F^*}FC') + tr(CF'P_{F^*}U) + tr(U'P_{F^*}FC') + tr(U'P_{F^*}U)],$$

21
where

\[(NT)^{-1} \text{tr}(U'P_F \cdot FC') = \frac{1}{NT} \sum_{i=1}^{N} \text{tr}(U'_i P_F \cdot FC_i)\]

\[= \frac{1}{\sqrt{NT}} \text{tr} \left( \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} C_i U'_i F^* (T^{-1} F^* F^*)^{-1} T^{-1} F^* F \right) = O_p((NT)^{-1/2}),\]

with \((NT)^{-1} \text{tr}(CF'P_F \cdot U)\) having the same order. Similarly,

\[(NT)^{-1} \text{tr}(U'P_F \cdot U) = \frac{1}{NT} \sum_{i=1}^{N} \text{tr}(U'_i P_F \cdot U_i)\]

\[= \frac{1}{T} \sum_{i=1}^{N} \text{tr}[T^{-1/2} U'_i F^* (T^{-1} F^* F^*)^{-1} T^{-1/2} F^* U_i] = O_p(T^{-1}).\]

By using this and the fact that \(N^{-1} C' C = \overline{Q} = N^{-2a} \overline{Q}^0\), where \(|| \overline{Q}^0 - Q^* || = o_p(1)\) (Assumption RK–PC), we obtain

\[Q(F^*) = (NT)^{-1} \text{tr}(CF'P_F \cdot FC') + O_p((NT)^{-1/2}) + O_p(T^{-1})\]

\[= T^{-1} \text{tr}(F'P_F \cdot F\overline{Q}) + O_p((NT)^{-1/2}) + O_p(T^{-1})\]

\[= N^{-2a} T^{-1} \text{tr}(F'P_F \cdot FQ^*) + O_p((NT)^{-1/2}) + O_p(T^{-1}).\]

Moreover, because \(Q^*\) is positive definite we may write \(Q^* = (Q^*)^{1/2} (Q^*)^{1/2}\), suggesting that,

\[Q(F^*) = N^{-2a} Q_F(F^*) + O_p((NT)^{-1/2}) + O_p(T^{-1}),\]  \hspace{1cm} (A9)

where \(Q_F(F^*) = T^{-1} \text{tr}(F^0P_F \cdot F^0)\) and \(F^0 = F(Q^*)^{1/2}\). Consider \(Q_F(F^*)\). Denote by \(S_T\) the eigenvectors of the \(r \times r\) matrix \(T^{-1} F^0 F^0\) and let \(V_T\) be the associated diagonal matrix of eigenvalues. It follows that \(T^{-1} F^0 F^0 S_T = S_T V_T\), and therefore \(T^{-1} F^0 F^0 (F^0 S_T) = (F^0 S_T) V_T\), where now \(F^0 S_T\) contains the eigenvectors of the \(T \times T\) matrix \(T^{-1} F^0 F^0\). By using this and \(F^0 F^* = T I_r\), we obtain

\[Q_F(F^*) = T^{-1} \text{tr}(F^0P_F \cdot F^0) = T^{-2} \text{tr}(F^0 F' F^* F^0) = T^{-1} \text{tr}[F^* (T^{-1} F^0 F^0) F^*],\]

which means that \(Q_F(F^*)\) is maximized for \(F^* = \sqrt{T} F^0 S_T\). While infeasible, \(F^*\) lend itself to simple asymptotics. Specifically, letting \(S = \lim_{T \to \infty} S_T\) and \(\Sigma^0 = \lim_{T \to \infty} T^{-1} F^0 F^0 = \lim_{T \to \infty} (Q^*)^{1/2} T^{-1/2} F^0 F(Q^*)^{1/2}\), we can show that

\[Q_F(F^*) = T^{-1} \text{tr}(F^0P_F \cdot F^0) = T^{-1} \text{tr}[F^0 F^* (F^0 F^0 F^0 F^0)]\]

\[= tr[T^{-1} F^0 F^0 S_T (S_T' T^{-1} F^0 F^0 S_T)^{-1} S_T' T^{-1} F^0 F^0]\]

\[= tr[\Sigma^0 S (\Sigma^0)^{-1} S^* + o_p(1)].\]  \hspace{1cm} (A10)
In order to appreciate the relevance of (A10) for the evaluation of $T^{-1}\hat{F}^{PC}F$, note that using $T^{-1}\hat{F}^{PC}F = I_r$, we obtain

$$Q_F(\hat{F}^{PC}) = T^{-1}F^0P_{FPC}F^0 = T^{-2}F^0\hat{F}^{PC}F^0 = (Q^*)^{1/2}(T^{-1}F^0\hat{F}^{PC})(T^{-1}F^{PC}F)(Q^*)^{1/2}. $$

Hence, if we can show that $Q_F(\hat{F}^{PC}) - Q_F(\hat{F}^*) = o_p(1)$, then the limit of $T^{-1}\hat{F}^{PC}F$ can be worked out from that of $Q_F(\hat{F}^*)$ above. We start with the following expansion:

$$N^{-2a}[Q_F(\hat{F}^*) - Q_F(\hat{F}^{PC})] = N^{-2a}[Q_F(\hat{F}^*) - Q_F(F)] - N^{-2a}[Q_F(\hat{F}^{PC}) - Q_F(F)], \quad (A11)$$

where

$$N^{-2a}[Q_F(\hat{F}^{PC}) - Q_F(F)] = [N^{-2a}Q_F(\hat{F}^{PC}) - Q(\hat{F}^{PC})] + [Q(\hat{F}^{PC}) - N^{-2a}Q_F(\hat{F}^*)] + N^{-2a}[Q_F(\hat{F}^*) - Q_F(F)].$$

By (A9), the first term on the right hand side is $O_p((NT)^{-1/2}) + O_p(T^{-1})$. Also, since $\hat{F}^{PC}$ and $\hat{F}^*$ are the maximizers of $Q(F^*)$ and $Q_F(F^*)$, respectively, and $|Q(F^*) - N^{-2a}Q_F(F^*)| = o_p(1)$, the second term is of the same order. It remains to consider the third term. Note that $\hat{F}^* = \sqrt{T}F^0S_T = \sqrt{T}F(Q^*)^{1/2}S_T$. The eigenvalue interpretation of $F^0S_T = F(Q^*)^{1/2}S_T$ means that $S_T(Q^*)^{1/2}(T^{-1}F^0F)(Q^*)^{1/2}S_T = I_r$ and $T^{-1}F^0F = [(Q^*)^{1/2}S_T]\alpha(Q^*)^{1/2}].$ Making use of these relationships, we obtain

$$P_F = \hat{F}^*F\hat{F}^*^{-1} = T^{-1}F(Q^*)^{-1/2}S_T\alpha(Q^*)^{-1/2}(T^{-1}F^0F)(Q^*)^{1/2}S_T^{-1}\alpha(Q^*)^{1/2}F' = T^{-1}F(Q^*)^{1/2}S_T\alpha(Q^*)^{-1/2}F' = T^{-1}F(F^0F)^{-1}F' = P_F.$$

Therefore,

$$Q_F(\hat{F}^*) - Q_F(F) = T^{-1}tr[F^0(\hat{P}_F - P_F)F^0] = 0,$$

which in turn implies

$$N^{-2a}[Q_F(\hat{F}^{PC}) - Q_F(F)] = O_p((NT)^{-1/2}) + O_p(T^{-1}).$$

and

$$N^{-2a}[Q_F(\hat{F}^*) - Q_F(\hat{F}^{PC})] = O_p((NT)^{-1/2}) + O_p(T^{-1}), \quad (A12)$$

or, by imposing $T = N^x$,

$$Q_F(\hat{F}^*) - Q_F(\hat{F}^{PC}) = O_p(N^{2a-1/2}T^{-1/2}) + O_p(N^{2a}T^{-1})
= O_p(N^{2a-(1+x)/2}) + O_p(N^{2a-x}). \quad (A13)$$

23
Provided that $\kappa > \max\{2\alpha, 4\alpha - 1\}$ this is $o_p(1)$. By using this and (A10),

$$T^{-1}F_{\hat{PC}}^0F^0 = \left(Q^*\right)^{1/2}(T^{-1}F_{\hat{PC}}^0)(T^{-1}F_{\hat{PC}}^0F^0)^{1/2} = \Sigma^0S(S^0\Sigma^0)^{-1}S^0\Sigma^0 + o_p(1),$$

and therefore

$$T^{-1}\hat{F}_{PC}^0F = \left(S^0\Sigma^0S^0\right)^{-1/2}S^0\Sigma^0\left(Q^*\right)^{-1/2} + o_p(1). \quad (A14)$$

Note that $S$ is the eigenvector matrix of $\Sigma^0$ with $V = \lim_{T \to \infty} V_T$ being the associated diagonal matrix of eigenvalues. It follows that $\Sigma^0S = SV$, and since $S = S' = S^{-1}$, we also have that $S^0\Sigma^0S = V$. Hence, since $V$ is symmetric,

$$(S^0\Sigma^0S)^{-1/2}S^0\Sigma^0 = V^{-1/2}VS' = V^{1/2}S',$$

giving

$$T^{-1}\hat{F}_{PC}^0F = \left(S^0\Sigma^0S^0\right)^{-1/2}S^0\Sigma^0\left(Q^*\right)^{-1/2} + o_p(1) = V^{1/2}S'(Q^*)^{-1/2} + o_p(1), \quad (A15)$$

and so we obtain

$$N^{2\alpha}H = N^{2\alpha}Q(T^{-1}F_{\hat{PC}}^0V^{-1} = Q'(V^{1/2}S'(Q^*)^{-1/2})'V^{-1} + o_p(1)$$

$$= H^* + o_p(1), \quad \text{(A16)}$$

where $H^* = (Q^*)^{1/2}SV^{-1/2}$.

**Lemma CCE2.** Under Assumptions ERR and LAM,

$$\frac{1}{T} \sum_{t=1}^{T} ||d_{t}^{CCE}||^2 = O_p(N^{-1}).$$

**Proof of Lemma CCE2**

The proof of Lemma CCE2 is a simple consequence of the fact that $||\sqrt{N}\bar{u}_t|| = O_p(1)$, as seen by writing

$$\frac{N}{T} \sum_{t=1}^{T} ||d_{t}^{CCE}||^2 \leq \frac{1}{T} \sum_{t=1}^{T} ||\sqrt{N}\bar{u}_t||^2 = O_p(1).$$

**Lemma PC2.** Under Assumptions ERR and LAM,

$$\frac{1}{T} \sum_{t=1}^{T} ||d_{t}^{PC}||^2 = O_p(T^{-1}) + O_p(N^{-1}).$$

24
**Proof of Lemma PC2**

By the definition of \(d_t^{PC}\),

\[
\frac{1}{T} \sum_{t=1}^{T} ||d_t^{PC}||^2 \leq ||V_T^{-1}||^2 \frac{3}{N^2T^3} \sum_{t=1}^{T} \left(||(\hat{F}^{PC})'Uu_t||^2 + ||(\hat{F}^{PC})'FC'u_t||^2 + ||(\hat{F}^{PC})'UCf_t||^2\right).
\]

Here

\[
\frac{1}{N^2T^3} \sum_{t=1}^{T} ||(\hat{F}^{PC})'Uu_t||^2 = \frac{1}{N^2T^3} \sum_{t=1}^{T} \left(\sum_{s=1}^{T} \hat{f}_s^{PC}u'_tu_t\right)^2 \\
\leq \frac{1}{T} \left(\frac{1}{T} \sum_{s=1}^{T} ||\hat{f}_s^{PC}||^2\right) \left(\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} (N^{-1}u'_tu_t)^2\right) \\
= O_p(T^{-1}) + O_p(N^{-1}),
\]

where the last equality holds because of the normalization \(T^{-1}\hat{F}^{PC}\hat{F}^{PC} = I_r\), suggesting that \(T^{-1}||\hat{F}^{PC}||^2 = T^{-1}\sum_{s=1}^{T} ||\hat{f}_s^{PC}||^2 = r\). Also, since \(N^{-1}u'_tu_t = O_p(1)\) and \(N^{-1/2}u'_tu_t = O_p(1)\) whenever \(t \neq s\),

\[
\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} (N^{-1}u'_su_t)^2 = \frac{1}{T} \sum_{t=1}^{T} (N^{-1}u'_tu_t)^2 + \frac{T}{N} \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} (N^{-1/2}u'_su_t)^2 = O_p(1) + O_p(TN^{-1}).
\]

Further use of \(||N^{-1/2}C'u_t|| = N^{-\alpha}||N^{-1/2}C^0u_t|| = O_p(N^{-\alpha})\) gives

\[
\frac{1}{N^2T^3} \sum_{t=1}^{T} ||(\hat{F}^{PC})'FC'u_t||^2 \\
= \frac{1}{N^2T^3} \sum_{t=1}^{T} \left(\sum_{s=1}^{T} \hat{f}_s^{PC}C'u_t\right)^2 \\
\leq N^{-(2a+1)} \left(\frac{1}{T} \sum_{s=1}^{T} ||\hat{f}_s^{PC}||^2\right) \left(\frac{1}{T} \sum_{s=1}^{T} ||f_s||^2\right) \left(\frac{1}{T} \sum_{t=1}^{T} ||N^{-1/2}C^0u_t||^2\right) \\
= O_p(N^{-2a+1}),
\]

with \(N^{-2T^3} \sum_{t=1}^{T} ||(\hat{F}^{PC})'UCf_t||^2\) being of the same order. Hence, since \(||V_T^{-1}|| = O_p(1),\)

\[
\frac{1}{T} \sum_{t=1}^{T} ||d_t^{PC}||^2 \leq ||V_T^{-1}||^2 \frac{3}{N^2T^3} \sum_{t=1}^{T} \left(||(\hat{F}^{PC})'Uu_t||^2 + ||(\hat{F}^{PC})'FC'u_t||^2 + ||(\hat{F}^{PC})'UCf_t||^2\right) \\
= O_p(T^{-1}) + O_p(N^{-1}) + O_p(N^{-(2a+1)}) = O_p(T^{-1}) + O_p(N^{-1}),
\]

and so the proof is complete. ■

**Lemma CCE3.** Under Assumptions ERR and LAM,

\[
||\sqrt{NT^{-1/2}}F'D^{CCE}|| = O_p(1).
\]
Proof of Lemma CCE3

The proof is completed by noting that
\[
\sqrt{NT^{-1/2}}F'D_{CCE}^C = \frac{\sqrt{N}}{\sqrt{T}} \sum_{t=1}^{T} f_{1}u_{t}^C = O_p(1). \tag{A17}
\]

Lemma PC3. Under Assumptions ERR and LAM,

\[||\sqrt{NT^{-1/2}}F'D_{PC}^C|| = O_p(R_1),\]

where

\[R_1 = N^{-3\alpha} + N^{-1/2} + (1 + N^{1/2-2\alpha})T^{-1/2} + \sqrt{NT}^{-1}.\]

Proof of Lemma PC3

Write
\[
\sqrt{NT^{-1/2}}F'D_{PC}^C = \frac{\sqrt{N}}{\sqrt{T}} \sum_{t=1}^{T} f_{1}d_{t}^PC = \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_{1}g_{t}^PC \hat{V}_T^{-1}
\]
\[
= \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_{1}(u_{t}U' + u_{t}CF' + f_{t}C'U') \hat{F}_p V_T^{-1}. \tag{A18}
\]

The first term on the right involves
\[
\frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_{1}u_{t}U' \hat{F}_p^C = \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_{1}u_{t}U'D_{PC} + \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_{1}u_{t}U' \hat{F}_p^C V_T^{-1},
\]

where
\[
\left| \left| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_{1}u_{t}U'D_{PC} \right| \right| = \left| \left| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} \sum_{s=1}^{T} f_{t}u_{t}u_{s}d_{s}^{PC}\right| \right|
\]
\[
\leq \left( \frac{1}{T} \sum_{s=1}^{T} \left| \left| \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} f_{t}u_{t}u_{s}\right| \right| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} \left| \left| d_{s}^{PC}\right| \right| \right)^{1/2}
\]
\[
= [O_p(\sqrt{NT}^{-1/2}) + O_p(1)][O_p(N^{-1/2}) + O_p(T^{-1/2})]
\]
\[
= O_p(\sqrt{NT}^{-1}) + O_p(N^{-1/2}) + O_p(T^{-1/2}),
\]

as \(T^{-1} \sum_{s=1}^{T} \left| \left| d_{s}^{PC}\right| \right| ^{2} = O_p(N^{-1}) + O_p(T^{-1})\) by Lemma PC2 and
\[
\frac{1}{\sqrt{NT}} \sum_{t=1}^{T} f_{t}u_{t}u_{s} = \sqrt{NT}^{-1/2}N^{-1}f_{s}u_{t}u_{s} + \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} f_{t}u_{t}u_{s} = O_p(\sqrt{NT}^{-1/2}) + O_p(1).
\]
Also,
\[
\frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' F = \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} \sum_{s=1}^{T} f_t u_t' f_s' \\
= \sqrt{N} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} f_t u_t' f_t' + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \sum_{s \neq t}^{T} f_t u_t' f_s' \\
= O_p(\sqrt{NT^{-1/2}}) + O_p(T^{-1/2}) = O_p(\sqrt{NT^{-1/2}}),
\]
suggesting that, with \( ||H|| = O_p(N^{-2\alpha}) \),
\[
\left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' U' f_{PC} \right\| \\
\leq \left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' U' f_{PC} \right\| + \left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' U' F \right\| ||H|| \\
= O_p(\sqrt{NT^{-1}}) + O_p(N^{-1/2}) + O_p(T^{-1/2}) + O_p(N^{-2\alpha})O_p(\sqrt{NT^{-1/2}}) \\
= O_p(\sqrt{NT^{-1}}) + O_p(N^{-1/2}) + O_p(T^{-1/2}) + O_p(N^{1/2-2\alpha}T^{-1/2}).
\]

(A19)

From \( ||(NT)^{-1/2} \sum_{t=1}^{T} f_t u_t' C|| = N^{-\alpha} ||(NT)^{-1/2} \sum_{t=1}^{T} f_t u_t' C|| = O_p(N^{-\alpha}) \), we further have
\[
\left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} \sum_{s=1}^{T} f_t u_t' C f_s f_{PC} \right\| \\
= \left( \frac{1}{T} \sum_{t=1}^{T} \left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{s=1}^{T} f_t u_t' C \right\| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} \left\| f_s \right\|^{2} \right)^{1/2} \\
= O_p(N^{-\alpha}) \left( O_p(N^{-1/2}) + O_p(T^{-1/2}) \right) \\
= O_p(N^{-(\alpha+1/2)}) + O_p(N^{-\alpha}T^{-1/2}),
\]
and
\[
\left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' C F' \right\| \leq N^{-\alpha} \left\| \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} f_t u_t' C \right\| \left\| T^{-1/2} F' \right\| = O_p(N^{-\alpha}),
\]
giving
\[
\left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' C F' f_{PC} \right\| \\
\leq \left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' C F' f_{PC} \right\| = \left\| \frac{1}{\sqrt{NT^{3/2}}} \sum_{t=1}^{T} f_t u_t' C F' F \right\| \left\| H \right\| \\
= O_p(N^{-\alpha}T^{-1/2}) + O_p(N^{-\alpha}T^{-1/2}) + O_p(N^{-3\alpha}),
\]
(A20)

with \( ||N^{-1/2}T^{-3/2} \sum_{t=1}^{T} f_t f_t' C' U' f_{PC}|| \) being of the same order. Hence, letting
\[
R_1 = N^{-3\alpha} + N^{-1/2} + (1 + N^{1/2-2\alpha})T^{-1/2} + \sqrt{NT^{-1}},
\]
27
with \(||V_T^{-1}|| = O_p(1),
\[
||\sqrt{NT^{-1/2}}F' D_{PC}|| \\
\leq \left|\frac{1}{\sqrt{NT^3/2}} \sum_{t=1}^{T} f_t(u_t'U' + u_t'C F' + f_t'C'U') \hat{F}_{PC}\right| ||V_T^{-1}|| = O_p(R_1),
\]

\( (A21) \)
as required.

\textbf{Lemma CCE4.} Under Assumptions ERR and LAM,
\[
NT^{-1}D_{CCE}'D_{CCE} = \Sigma_u + O_p(T^{-1/2}),
\]
where
\[
\Sigma_u = \frac{1}{N} \sum_{i=1}^{N} \Sigma_{u,i},
\]
\[
\Sigma_{u,i} = E(u_{i,t}u_{i,t}') = \left[ \begin{array}{c} \beta' \Sigma_{e,i} \beta + \sigma^2_{e,i} \beta' \Sigma_{e,i} \\ \Sigma_{e,i} \beta' \Sigma_{e,i} \end{array} \right].
\]

\textbf{Proof of Lemma CCE4}

A direct calculation reveals that
\[
NT^{-1}D_{CCE}'D_{CCE} = \frac{N}{T} \sum_{t=1}^{T} d_t^{CCE}'d_t^{CCE} = \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} u_{i,t}u_{j,t}'
\]
\[
= \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} u_{i,t}u_{i,t}' + \frac{1}{\sqrt{T}} \frac{1}{N\sqrt{T}} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j \neq i} u_{i,t}u_{j,t}'
\]
\[
= \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} u_{i,t}u_{i,t}' + O_p(T^{-1/2})
\]
\[
= \Sigma_u + \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} (u_{i,t}u_{i,t}' - \Sigma_u) + O_p(T^{-1/2})
\]
\[
= \Sigma_u + O_p(T^{-1/2}),
\]
\( (A22) \)
as was to be shown.

\textbf{Lemma PC4.} Under Assumptions ERR, LAM and RK–PC,
\[
NT^{-1}D_{PC}'D_{PC} = N^{-2a}H^0(\overrightarrow{Q})^{-1}S^0(\overrightarrow{Q})^{-1}H^0 + O_p(R_4),
\]
\( (A23) \)
where
\[
S^0 = \frac{1}{N} \sum_{i=1}^{N} C_0^i \Sigma_{u,i} C_0^i
\]
\[
R_4 = N^{-3a-1} + N^{-4a-1/2} + (N^{-2a-1/2} + N^{-4a} + N^{-1})T^{-1/2} + (N^{-2a} + N^{-1/2}T^{-1}
\]
\[
+ (1 + N^{1/2-3a})T^{-3/2} + (N^{-4a+1} + N^{-2a+1/2})T^{-2} + N^{1-2a}T^{-5/2}.
\]

28
Proof of Lemma PC4

By definition,

\[ NT^{-1}D^{PC}_TD^{PC} \]

\[ = N^{-1}T^{-3}V^{-1}F^{PC}_TG^{PC}T^{-1} \]

\[ = N^{-1}T^{-3}V^{-1}[\hat{H}'F'G'\hat{G} \hat{F} + \overline{H}'F'G'D^{PC} + D^{PC}_T G' \hat{G} + D^{PC}_T G'GD^{PC}]V^{-1}. \]  
(A24)

Consider the second term on the right-hand side. Clearly,

\[ ||N^{-1}T^{-3}F'G'D^{PC}|| = \left| \left| \frac{1}{NT^3} \sum_{s=1}^{T} F'G'_s D^{PC}_s \right| \right| \]

\[ \leq \left( \frac{1}{T} \sum_{s=1}^{T} ||N^{-1}T^{-2}F'G'_s||^2 \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} ||d^{PC}_s||^2 \right)^{1/2}, \]  
(A25)

where the order of the second term in the product is given by Lemma PC2. As for the first term,

\[ N^{-1}T^{-2}F'G'_s \]

\[ = N^{-1}T^{-2}F'(UU' + FC'U' + UCF') (UU_s + FC'u_s + UCF_s) \]

\[ = N^{-1}T^{-2}(F'UU'U_s + F'FC'U_s + F'UCF_s + F'UU'FC'_s + F'FC'UFC'_s) \]

\[ = j_1 + \ldots + j_9, \]  
(A26)

where the dependence on \( s \) in \( j_1, \ldots, j_9 \) has been suppressed.

Consider \( j_1 \), which can be expanded as

\[ ||j_1|| = ||N^{-1}T^{-2}F'UU'U_s|| \]

\[ = \left| \left| \frac{1}{NT^2} \sum_{k=1}^{T} f_k u'_k u_s u'_s \right| \right| \]

\[ \leq \left| \left| \frac{1}{NT^2} \sum_{k=1}^{T} f_k u'_k u_s u'_s \right| \right| + \frac{1}{NT^2} \sum_{t \neq s}^{T} \left| \left| f_t u'_t u_s u'_s \right| \right| \]

\[ \leq NT^{-2} \left| \left| N^{-2} f_s u'_s u_s u'_s \right| \right| + \frac{\sqrt{N}}{T^{3/2}} \left| \left| \frac{1}{N^3/2 \sqrt{T}} \sum_{k \neq s}^{T} f_k u'_k u_s u'_s \right| \right| \]

\[ + \frac{\sqrt{N}}{T^{3/2}} \left| \left| \frac{1}{N^3/2 \sqrt{T}} \sum_{t \neq s}^{T} \sum_{k \neq t}^{T} f_k u'_k u'_t u'_s \right| \right| \]

\[ = O_p(NT^{-2}) + O_p(\sqrt{N}T^{-3/2}) + O_p(T^{-1/2}), \]  
(A27)
where last equality follows from direct evaluation of the individual terms. The order of the second term, for example, is obtained as

\[
\left\| \frac{1}{NT} \sum_{k \neq s}^T f_k u_k' u_k u_s' \right\| \leq \frac{\sqrt{N}}{T^{3/2}} \left\| \frac{1}{\sqrt{NT}} \sum_{k \neq s}^T \sum_{i=1}^N f_k u_k' u_{i,s} \right\| \left\| \frac{1}{N} \sum_{j=1}^N u_j' u_{j,s} \right\| = O_p(\sqrt{NT^{-3/2}}).
\]

Consider \( j_2 \). Let \( \Sigma_u = \text{diag}(\Sigma_{u,1}, \ldots, \Sigma_{u,N}) \), such that \( E(u_i, u_j') = \Sigma_{ij} \), an \( N(m+1) \times N(m+1) \) matrix. Note that

\[
C^0_t(u_i' - \Sigma_u)u_s = \sum_{i=1}^N \sum_{j=1}^N C^0_i[1/(u_i' u_{i,t}) - E(u_i' u_{i,t})]u_{i,s}
\]

\[
= \sum_{i=1}^N C^0_i(u_{i,t} - \Sigma_{u,t})u_{i,s} + \sum_{i=1}^N \sum_{j \neq i}^N C^0_i u_{i,j} u_{i,s}
\]

where both terms on the right are mean zero and independent across \( s \neq t \), suggesting

\[
\frac{1}{NT} \sum_{t \neq s}^T C^0_t u_t u_t' u_s
\]

\[
= \frac{1}{NT} \sum_{t \neq s}^T C^0_t (u_t' - \Sigma_u)u_s + \frac{(T-1)}{NT} C^0_t \Sigma_u u_s
\]

\[
= \frac{1}{\sqrt{NT}} \left( \sum_{t \neq s} \sum_{i=1}^N C^0_i (u_{i,t} - \Sigma_{u,t})u_{i,s} + \frac{1}{\sqrt{T}} \sum_{t \neq s} \sum_{i=1}^N \sum_{j \neq i}^N C^0_i u_{i,j} u_{i,s} \right)
\]

\[
+ \frac{1}{\sqrt{N}} (T-1) \sum_{i=1}^N C^0_i \Sigma_u u_{i,s}
\]

\[
= O_p((NT)^{-1/2}) + O_p(T^{-1/2}) + O_p(N^{-1/2}) = O_p(T^{-1/2}) + O_p(N^{-1/2}).
\]

It follows that

\[
||j_2|| = ||N^{-1}T^{-2}F'FC'U'U_s||
\]

\[
\leq ||T^{-1}F'F|| \left\| \frac{1}{NT} \sum_{t=1}^T C_t u_t u_t' u_s \right\|
\]

\[
\leq N^{1/2-a}T^{-1} ||T^{-1}F'F|| ||N^{-1/2}C^0_t u_s|| ||N^{-1}u_t' u_s||
\]

\[
+ N^{-a} ||T^{-1}F'F|| \left\| \frac{1}{NT} \sum_{t \neq s}^T C^0_t u_t u_t' u_s \right\|
\]

\[
= O_p(N^{1/2-a}T^{-1}) + O_p(N^{-a}T^{-1/2}) + O_p(N^{-(a+1/2)}). \quad (A28)
\]
For \( j_3 \), we use the fact that

\[
\frac{1}{\sqrt{NT}} \sum_{k=1}^{T} f_k u_k' u_s = \frac{1}{\sqrt{NT}} \sum_{t=1}^{N} \sum_{k=1}^{T} f_k u_k' u_{i,s}
\]

\[
= \frac{\sqrt{N}}{\sqrt{T}} \frac{1}{N} \sum_{t=1}^{N} f_{t} u_{i,s} + \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} \sum_{k=1, k \neq t}^{T} f_k u_k' u_{i,s}
\]

\[
= O_p(\sqrt{NT}^{-1/2}) + O_p(1),
\]

giving

\[
||j_3|| = ||N^{-1}T^{-2}F'UUF'Uu_s|| \leq \frac{1}{N^{aT}} || \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} f_{t} u_{i}C' || \frac{1}{\sqrt{NT}} \sum_{k=1}^{T} f_k u_k' u_s || \]

\[
= N^{-a}T^{-1}[O_p(\sqrt{NT}^{-1/2}) + O_p(1)] = O_p(N^{1/2-a}T^{-3/2}) + O_p(N^{-a}T^{-1}). \tag{A29}
\]

Also, since

\[
N^{-1/2}T^{-2}F'UU'F
\]

\[
= \frac{1}{\sqrt{NT}^2} \sum_{t=1}^{T} \sum_{k=1}^{T} f_{t} u_{i}u_k f_k^{'},
\]

\[
= \frac{\sqrt{N}}{T} \frac{1}{NT} \sum_{t=1}^{T} f_{t} u_{i}f_{t} + \frac{1}{T} \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} \sum_{k=1, k \neq t}^{T} f_{t} u_{i}u_k f_k^{'},
\]

\[
= O_p(\sqrt{NT}^{-1}) + O_p(T^{-1}) = O_p(\sqrt{NT}^{-1}),
\]

we can show that

\[
||j_4|| = ||N^{-1}T^{-2}F'UU'FC'u_s|| \leq N^{-a}||N^{-1/2}T^{-2}F'UU'F||||N^{-1/2}C^0u_s||
\]

\[
= O_p(N^{1/2-a}T^{-1}). \tag{A30}
\]

As for \( j_5 \), by using arguments similar to those used in the above,

\[
||j_5|| = ||N^{-1}T^{-2}F'FC'U'FC'u_s||
\]

\[
\leq N^{-2a}T^{-1/2}||T^{-1}F'F||||(NT)^{-1/2}C^0U'F||||N^{-1/2}C^0u_s||
\]

\[
= O_p(N^{-2a}T^{-1/2}), \tag{A31}
\]

with \( ||j_6|| \) being of the same order.

For \( j_7 \),

\[
||j_7|| = ||N^{-1}T^{-2}F'UU'UCf_s||
\]

\[
= \left| \frac{1}{NT^2} \sum_{t=1}^{T} \sum_{k=1}^{T} f_{t} u_{i}u_k C f_s \right|
\]

\[
\leq \frac{1}{N^{a-1/2}T} \left| \frac{1}{NT^2} \sum_{t=1}^{T} f_{t} u_{i}u_{i}C f_s \right| + \frac{1}{N^{a} \sqrt{T}} \left| \frac{1}{NT^{3/2}} \sum_{t=1}^{T} \sum_{k=1, k \neq t}^{T} f_{t} u_{i}u_k C f_s \right|
\]

\[
= O_p(N^{-a+1/2}T^{-1}) + O_p(N^{-a}T^{-1/2}), \tag{A32}
\]

\]

31
where the last equality holds because
\[
\left\| \frac{1}{NT} \sum_{i=1}^{T} f_i u'_i u_i C^0 f_i \right\| \\
\leq \left( \frac{1}{T} \sum_{i=1}^{T} \left| \left| N^{-1} f_i u'_i \right| \right|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{i=1}^{T} \left| \left| N^{-1/2} u'_i C^0 \right| \right|^2 \right)^{1/2} \left\| f_i \right\| = O_p(1),
\]
and
\[
\left\| \frac{1}{NT^{3/2}} \sum_{i=1}^{T} \sum_{k \neq i} f_i u'_i u'_i u^0 f_i \right\| \\
\leq \left( \frac{1}{T} \sum_{k \neq i} \left\| \frac{1}{NT} \sum_{i=1}^{T} f_i u'_i u_k \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{k \neq l} \left| \left| N^{-1/2} u'_i C^0 \right| \right|^2 \right)^{1/2} \left\| f_i \right\| = O_p(1).
\]
The order of $j_8$ can be obtained from
\[
\frac{1}{NT} \sum_{i=1}^{T} C^0 u_i u'_i C^0 \\
= \frac{1}{\sqrt{NT}} \sum_{i=1}^{T} C^0 (u_i u'_i - \Sigma u_i) C^0 + N^{-1} C^0 \Sigma u_i C^0 \\
= \frac{1}{\sqrt{NT}} \sum_{i=1}^{T} \sum_{j=1}^{N} C^0 (u_i, u'_i, \Sigma u_i, j) C^0 + \frac{1}{\sqrt{T}} \sum_{i=1}^{T} \sum_{j=1}^{N} C^0 u_i, u'_i, \Sigma C^0 \\
+ \frac{1}{N} \sum_{i=1}^{T} \sum_{j=1}^{N} C_i \Sigma u_i, j C^0 \\
= O_p((NT)^{-1/2}) + O_p(T^{-1}) + O_p(1) = O_p(1),
\]
where
\[
\frac{1}{NT} \sum_{i=1}^{T} \sum_{j=1}^{N} C^0 u_i, u'_i, C^0 = \frac{1}{NT} \sum_{i=1}^{T} \sum_{j=1}^{N} C^0 u_i, u'_i, C^0 + \frac{1}{NT} \sum_{i=1}^{T} \sum_{j=1}^{N} \sum_{j \neq i} C^0 u_i, u'_i, C^0 \\
= O_p(1) + O_p(T^{-1/2}),
\]
implicating
\[
\left\| j_8 \right\| = \left\| N^{-1} T^{-2} F^t F C^t U^t U C f \right\| \\
\leq N^{-2} \left\| T^{-1} F^t F \right\| \left\| \frac{1}{NT} \sum_{i=1}^{T} C^0 u_i u'_i C^0 \right\| \left\| f_i \right\| = O_p(N^{-2}). \quad \text{(A33)}
\]
It remains to consider $j_9$, whose order is given by
\[
\left\| j_9 \right\| = \left\| N^{-1} T^{-2} F^t U C F U C f \right\| = \frac{1}{N^{2a} T} \left( \frac{1}{\sqrt{NT}} \sum_{i=1}^{T} f_i u'_i C^0 \right)^2 \left\| f_i \right\| = O_p(N^{-2a} T^{-1}). \quad \text{(A34)}
\]
Hence, by putting everything together,

\[
||N^{-1}T^{-2}F'G'g|| \leq ||j_1|| + ... + ||j_9|| = O_p(R_2),
\]  

(A35)

where

\[
R_2 = N^{-(\alpha+1/2)} + N^{-2\alpha} + T^{-1/2} + N^{1/2-\alpha}T^{-1} + \sqrt{NT^{-3/2}} + NT^{-2}.
\]

and therefore

\[
||N^{-1}T^{-3}F'G'GD^{PC}|| \leq \left(\frac{1}{T} \sum_{s=1}^{T} ||N^{-1}T^{-2}F'G'g||^2\right)^{1/2} \left(\frac{1}{T} \sum_{s=1}^{T} ||d_s^{PC}||^2\right)^{1/2} = O_p(N^{-1/2}R_2) + O_p(T^{-1/2}R_2).
\]

(A36)

The order of \( ||N^{-1}T^{-3}D^{PC}G'G'F|| \) is the same. Thus, since \( ||V_T^{-1}|| = O_p(1) \), \( ||H|| = O_p(N^{-2\alpha}) \) and the order of \( ||N^{-1}T^{-3}D^{PC}G'G'GD^{PC}|| \) is dominated by that of \( ||N^{-1}T^{-3}F'G'GD^{PC}|| \), we obtain

\[
NT^{-1}D^{PC}D^{PC}
\]

\[
= N^{-1}T^{-3}V_T^{-1}||H F'G'GFH + H'F'G'GD^{PC} + D^{PC}G'GFH + D^{PC}G'GD^{PC}||V_T^{-1}
\]

\[
= N^{-1}T^{-3}V_T^{-1}H F'G'GFH + O_p(N^{-(4\alpha+1)/2}R_2) + O_p(N^{-2\alpha}T^{1/2}R_2).
\]

(A37)

Consider \( N^{-1}T^{-3}F'G'GF \), which we expand in the following obvious fashion:

\[
N^{-1}T^{-3}F'G'GF
\]

\[
= \frac{1}{NT^3} \sum_{t=1}^{T} F'g_t g'_t F
\]

\[
= \frac{1}{NT^3} \sum_{t=1}^{T} F'(U_{t+} + FC'u_{t+} + UCf_t) (u_t'U' + u_t'CF' + f_t'C'U')F
\]

\[
= \frac{1}{NT^3} \sum_{t=1}^{T} (F'U_{t+}u_t'U'F + F'FC'u_t'u_t'U'F + F'UCf_tu_t'U'F + F'Uu_tu_t'CF'F
\]

\[
+ F'FC'u_t'u_t'CF'F + F'UCf_tu_t'CF'F + F'Uu_tf_t'C'U'F + F'FC'u_t'f_t'C'U'F + F'UCf_tf_t'C'U'F)
\]

\[
= J_1 + ... + J_9.
\]

(A38)
For $J_1$,

\[
J_1 = \frac{1}{NT^3} \sum_{t=1}^{T} F'Uu_t'u_t'F
\]

\[
= \frac{1}{NT^3} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{k=1}^{T} f_\nu u'_s u_t' u_k' f_k'
\]

\[
= \frac{1}{NT^3} \sum_{t=1}^{T} f_\nu u'_t u_t' u_t' f_t' + \frac{1}{NT^3} \sum_{s \neq t}^{T} \sum_{k \neq t}^{T} f_\nu u'_s u_t' u_k' f_k'
\]

\[
+ \frac{1}{NT^3} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \sum_{k \neq t}^{T} f_\nu u'_s u_t' u_k' f_k',
\]

where

\[
\left\| \frac{1}{NT^3} \sum_{t=1}^{T} f_\nu u'_t u_t' u_t' f_t' \right\| \leq \frac{N}{T^2} \frac{1}{T} \sum_{t=1}^{T} \left\| N^{-1} f_\nu u'_t \right\|^2 = O_p(NT^{-2}),
\]

\[
\left\| \frac{1}{NT^3} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \sum_{k \neq t}^{T} f_\nu u'_s u_t' u_k' f_k' \right\| \leq N^{1/2} \frac{1}{T^{3/2}} \left( \frac{1}{T} \sum_{t=1}^{T} \left\| \frac{1}{\sqrt{NT}} \sum_{s \neq t} f_\nu u'_t \right\| \right)^2 \left( \frac{1}{T} \sum_{t=1}^{T} \left\| N^{-1} u'_t f_t' \right\|^2 \right)^{1/2} = O_p(N^{1/2}T^{-3/2}),
\]

and

\[
\left\| \frac{1}{NT^3} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \sum_{k \neq t}^{T} f_\nu u'_s u_t' u_k' f_k' \right\| \leq \frac{1}{NT^3} \frac{1}{T} \sum_{t=1}^{T} \left\| \frac{1}{\sqrt{NT}} \sum_{s \neq t} f_\nu u'_t \right\| \leq \frac{1}{NT^3} \frac{1}{T} \sum_{t=1}^{T} \left\| N^{-1} u'_t f_t' \right\|^2 = O_p(T^{-1}).
\]

$\sum_{t=1}^{T} \sum_{s \neq t}^{T} \sum_{k \neq t}^{T} f_\nu u'_s u_t' u_k' f_k' / NT^3$ is of the same order as $\sum_{t=1}^{T} \sum_{s \neq t}^{T} f_\nu u'_s u_t' u_k' f_k' / NT^3$. Hence,

\[
\|J_1\| = O_p(NT^{-2}) + O_p(T^{-1}) + O_p(N^{1/2}T^{-3/2}). \quad (A39)
\]

For $J_2$, we use

\[
\frac{1}{NT^2} \sum_{t=1}^{T} \sum_{s=1}^{T} C^0 u_t' u_t' f_t' = \frac{1}{NT^2} \sum_{t=1}^{T} C^0 u_t' u_t' f_t' + \frac{1}{NT^2} \sum_{t=1}^{T} \sum_{s \neq t}^{T} C^0 u_t' u_t' f_t',
\]

where

\[
\frac{1}{NT^2} \sum_{t=1}^{T} C^0 u_t' u_t' f_t' = \frac{1}{NT^2} \sum_{t=1}^{T} C^0 (u_t' - \Sigma_u) u_t' f_t' + \frac{1}{NT^2} \sum_{t=1}^{T} C^0 \Sigma_u u_t' f_t'
\]

\[
= \frac{1}{T} \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} C^0 (u_{i,t} u_{i,t} - \Sigma_{u,i}) u_{i,t} f_t' \right\| + \frac{1}{\sqrt{NT^3/2}} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j \neq i}^{N} C^0 u_{i,t} u_{j,t} f_t'
\]

\[
+ \frac{1}{\sqrt{NT^3/2}} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \sum_{i=1}^{N} C^0 \Sigma_{u,i} u_{i,t} f_t'
\]

\[
= O_p(N^{-1/2}T^{-3/2}) + O_p(\sqrt{NT^{-3/2}}) = O_p(\sqrt{NT^{-3/2}}),
\]

34
and, by exactly the same argument,

\[
\frac{1}{NT^2} \sum_{t=1}^{T} \sum_{s=1}^{T} C^0 u_i u_j f_s' = \frac{1}{\sqrt{NT}} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{i=1}^{N} C^0_i u_i u_j f_s' + \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{i=1}^{N} C^0_i u_i u_j f_s' + \frac{1}{\sqrt{NT^2}} \sum_{s \neq t=1}^{T} \sum_{i=1}^{N} C^0_i u_i u_j f_s' = O_p(N^{-1/2}T^{-1}) + O_p(T^{-1}) + O_p(N^{-1/2}T^{-3/2}) = O_p(T^{-1}).
\]

Therefore,

\[
\left\| \frac{1}{NT^2} \sum_{t=1}^{T} \sum_{s=1}^{T} C^0 u_i u_j f_s' \right\| = O_p(T^{-1}) + O_p(\sqrt{NT}T^{-3/2}).
\]

giving

\[
||J_2|| \leq N^{-a} \left|\left| T^{-1} F' F \right|\right| = O_p(N^{-a}T^{-1}) + O_p(N^{1/2-a}T^{-3/2}). \quad (A40)
\]

We know from before that \( ||N^{-1/2}T^{-2}F'UU'F|| = O_p(\sqrt{NT}T^{-1}) \) and \( ||(NT)^{-1/2}F'UC^0|| = O_p(1) \). This implies

\[
||J_3|| \leq \frac{1}{N^a} \left|\left| (NT)^{-1/2}F'UC^0 \right|\right| \left|\left| \frac{1}{\sqrt{NT^2}} \sum_{t=1}^{T} \sum_{s=1}^{T} f_t u_i u_j f_s' \right|\right| = O_p(N^{1/2-a}T^{-3/2}). \quad (A41)
\]

Also, since \( J_4 = J_2 \), \( ||J_4|| \) is of the same order of magnitude as \( ||J_2|| \).

Consider \( J_5 \), where

\[
\frac{1}{NT} \sum_{i=1}^{T} C^0_i u_i u_j C^0_j
\]

\[
= \frac{1}{NT} \sum_{i=1}^{T} \sum_{j=1}^{N} \sum_{j=1}^{N} C^0_i u_i u_j C^0_j + \frac{1}{\sqrt{NT}} \sum_{i=1}^{T} \sum_{j=1}^{N} \sum_{j=1}^{N} C^0_i u_i u_j C^0_j = O_p(T^{-1/2})
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} C^0_i \Sigma_{u_i} C^0_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{T} \sum_{j=1}^{N} C^0_i (u_i u_j - \Sigma_{u_i}) C^0_j + O_p(T^{-1/2})
\]

\[
= 1 \sum_{i=1}^{N} C^0_i \Sigma_{u_i} C^0_i + O_p((NT)^{-1/2}) + O_p(T^{-1/2})
\]

\[
= 0 + O_p(T^{-1/2})
\]
suggesting that

\[ J_5 = \frac{1}{NT^3} \sum_{i=1}^{T} \mathbf{F}' \mathbf{C}' \mathbf{u}_i \mathbf{u}_i' \mathbf{C} \mathbf{F} \]

\[ = N^{-2a} (T^{-1} \mathbf{F}' \mathbf{F}) \frac{1}{NT} \sum_{i=1}^{T} \mathbf{C}^{0i} \mathbf{u}_i \mathbf{u}_i' \mathbf{C}^{0i} (T^{-1} \mathbf{F}' \mathbf{F}) \]

\[ = N^{-2a} (T^{-1} \mathbf{F}' \mathbf{F}) \mathbf{S}^{0} (T^{-1} \mathbf{F}' \mathbf{F}) + O_p(N^{-2a} T^{-1/2}). \quad (A42) \]

The order of \(||\mathbf{J}_6|||\) is

\[ ||\mathbf{J}_6|| = \left| \frac{1}{NT^3} \sum_{i=1}^{T} \mathbf{F}' \mathbf{U} \mathbf{C} \mathbf{f}_i \mathbf{u}_i' \mathbf{C} \mathbf{F} \right| \]

\[ \leq \frac{1}{N^{2a} T} \left| \frac{1}{\sqrt{NT}} \sum_{i=1}^{T} \mathbf{f}_i \mathbf{u}_i' \mathbf{C}^{0} \right|^2 \left| T^{-1} \mathbf{F}' \mathbf{F} \right| = O_p(N^{-2a} T^{-1}), \quad (A43) \]

while that of \(||\mathbf{J}_7|||\) is \(O_p(N^{1/2 - a} T^{-3/2})\), as follows from the fact that \(\mathbf{J}_7 = \mathbf{J}_5\). Similarly, since \(\mathbf{J}_8 = \mathbf{J}_6\), we have \(||\mathbf{J}_8||| = O_p(N^{-2a} T^{-1})\).

Finally, consider \(\mathbf{J}_9\), which can be evaluated in the following fashion:

\[ ||\mathbf{J}_9|| = \left| \frac{1}{NT^3} \sum_{i=1}^{T} \mathbf{F}' \mathbf{U} \mathbf{C} \mathbf{f}_i \mathbf{f}_i' \mathbf{C}' \mathbf{U}' \mathbf{F} \right| \]

\[ \leq \frac{1}{N^{2a} T} \frac{1}{T} \sum_{i=1}^{T} \left| (NT)^{-1/2} \mathbf{F}' \mathbf{U} \mathbf{C}^{0} \right|^2 \left| \mathbf{f}_i \right|^2 = O_p(N^{-2a} T^{-1}). \quad (A44) \]

Hence, by adding the results,

\[ N^{-1} T^{-3} \mathbf{F}' \mathbf{G}' \mathbf{G} \mathbf{F} = \mathbf{J}_1 + \ldots + \mathbf{J}_9 \]

\[ = N^{-2a} (T^{-1} \mathbf{F}' \mathbf{F}) \mathbf{S}^{0} (T^{-1} \mathbf{F}' \mathbf{F}) + O_p(NT^{-2}) + O_p(T^{-1}) \]

\[ + O_p(N^{1/2} T^{-3/2}) + O_p(N^{-2a} T^{-1/2}), \quad (A45) \]

which in turn implies

\[ N^{-1} \mathbf{D}^{PC} \mathbf{D}^{PC} \]

\[ = N^{-1} T^{-3} \mathbf{V}_T^{-1} \mathbf{H}' \mathbf{F}' \mathbf{G}' \mathbf{G} \mathbf{F} \mathbf{H} \mathbf{V}_T^{-1} + O_p(N^{-(4a+1)/2} R_2) + O_p(N^{-2a} T^{-1/2} R_2) \]

\[ = N^{-2a} \mathbf{V}_T^{-1} \mathbf{H}' (T^{-1} \mathbf{F}' \mathbf{F}) \mathbf{S}^{0} (T^{-1} \mathbf{F}' \mathbf{F}) \mathbf{H} \mathbf{V}_T^{-1} + O_p(R_3), \quad (A46) \]

where

\[ R_3 = N^{-(4a+1)/2} R_2 + N^{-2a} T^{-1/2} R_2 + N^{-4a} (NT^{-2} + T^{-1} + N^{1/2} T^{-3/2} + N^{-2a} T^{-1/2}) \]

\[ = N^{-3a-1} + N^{-4a-1/2} + (N^{-2a-1/2} + N^{-4a}) T^{-1/2} \]

\[ + N^{-2a-1} T^{-1} + (N^{1/2 - 3a}) T^{-3/2} + N^{1-2a} T^{-5/2} + (N^{-2a+1/2} + N^{4a+1}) T^{-2} \]
is the order of $O_p(N^{-(4a+1)/2}R_2)$ plus $N^{-4a}$ times the four reminder terms in the above expression for $N^{-1}T^{-3}F'G'GF$.

Consider $\mathbf{H}$. According to Lemma PC3, $\|T^{-1}F'D_{PC}\| = O_p((NT)^{-1/2}R_1)$, from which it follows that
\[
\mathbf{H} = \mathbf{Q}(T^{-1}F'\hat{F}_{PC})V_T^{-1}
\]
\[
= N^{-2a}\mathbf{Q}^0(T^{-1}F'\mathbf{H})V_T^{-1} + N^{-2a}\mathbf{Q}^0(T^{-1}F'D_{PC})V_T^{-1}
\]
\[
= N^{-2a}\mathbf{Q}^0T^{-1}F'\mathbf{HV}_T^{-1} + O_p(N^{-2a+1/2}T^{-1/2}R_1),
\]
(A47)

or, since $\mathbf{Q}^0$ is invertible,
\[
T^{-1}F'\mathbf{HV}_T^{-1} = N^{2a}(\mathbf{Q}^0)^{-1}\mathbf{H} + O_p((NT)^{-1/2}R_1) = (\mathbf{Q}^0)^{-1}\mathbf{H} + O_p((NT)^{-1/2}R_1),
\]
with which we obtain
\[
NT^{-1}D_{PC}'D_{PC} = N^{-2a}V_T^{-1}\mathbf{H}'(T^{-1}F'\mathbf{F})\mathbf{S}^0(T^{-1}F'\mathbf{F})\mathbf{HV}_T^{-1} + O_p(R_3)
\]
\[
= N^{-2a}\mathbf{H}'\mathbf{S}^0(\mathbf{Q}^0)^{-1}\mathbf{Q}^0(\mathbf{Q}^0)^{-1}\mathbf{H} + O_p(R_4),
\]
(A48)

where
\[
R_4 = R_3 + (NT)^{-1}R_1
\]
\[
= N^{-3a-1} + N^{-4a-1/2} + (N^{-2a-1/2} + N^{-4a} + N^{-1})T^{-1/2} + (N^{-2a} + N^{-1})T^{-1}
\]
\[
+ (1 + N^{1/2-3a})T^{-3/2} + (N^{-4a+1} + N^{-2a+1/2})T^{-2} + N^{1-2a}T^{-5/2}
\]

This completes the proof. ■

**Lemma CCE5.** Under Assumptions ERR,
\[
NT^{-1}E_i'D_{CCE} = \Sigma e_{i,j}(\beta, I_m) + O_p(\sqrt{NT}^{-1/2}),
\]
\[
\frac{1}{T} \sum_{i=1}^{N} E_i'D_{CCE} = \frac{1}{N} \sum_{i=1}^{N} \Sigma e_{i,j}(\beta, I_m) + O_p(T^{-1/2}).
\]

**Proof of Lemma CCE5**

For the first result, note that $E(\epsilon_{i,t}u_{i,t}') = E[\epsilon_{i,t}(u_{i,t} + \epsilon_{i,t}'\beta, \epsilon_{i,t}')] = (\Sigma e_{i,j}\beta, \Sigma e_{i}) = \Sigma e_{i}(\beta, I_m)$. This suggests
\[
NT^{-1}E_i'D_{CCE} = \frac{N}{T} \sum_{i=1}^{T} \epsilon_{i,t}D_{iCCE}' = \frac{1}{T} \sum_{i=1}^{T} \epsilon_{i,t} \sum_{j=1}^{N} u_{i,t}'
\]
\[
= \Sigma e_{i,j}(\beta, I_m) + \frac{1}{\sqrt{T}} \frac{1}{\sqrt{T}} \sum_{i=1}^{T} (\epsilon_{i,t}u_{i,t}' - \Sigma e_{i,j}(\beta, I_m)) + \frac{\sqrt{N}}{\sqrt{T} \sqrt{NT}} \sum_{i=1}^{T} \sum_{j \neq i}^{N} \epsilon_{i,t}u_{i,t}'
\]
\[
= \Sigma e_{i}(\beta, I_m) + O_p(\sqrt{NT}^{-1/2}).
\]
For the second result, we can use the calculations above to obtain
\[
\frac{1}{T} \sum_{t=1}^{N} E_t^D = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{T} \epsilon_{i,j} \beta_i I_m + \frac{1}{\sqrt{NT}} \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} [\epsilon_{i,t} u_{i,t} - \Sigma_{e,i} \beta_i I_m] \\
+ \frac{1}{\sqrt{T}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j \neq i} \epsilon_{i,t} u_{j,t} \\
= \frac{1}{N} \sum_{i=1}^{N} \Sigma_{e,i} \beta_i I_m + O_p(T^{-1/2}),
\]
and so the proof of the lemma is complete.

Lemma PC5. Under Assumptions ERR, LAM and RK–PC,
\[
||NT^{-1}E[D^P]|| = O_p(N^{-a/2}) + O_p(\sqrt{NT^{-1/2}}) + O_p(N^{-3/2}),
\]
\[
\frac{1}{T} \sum_{i=1}^{N} E_t^D = \frac{1}{N^{a+1}} \sum_{i=1}^{N} \Sigma_{e,i} \beta_i I_m C_{i}^0 (Q_0)^{-1} H_1 + O_p(R_5),
\]
where
\[
R_5 = N^{-(a+1)/2} + T^{-1/2} + \sqrt{NT^{-3/2}}.
\]

Proof of Lemma PC5

Consider the first result. By the definition of $D^P$,
\[
NT^{-1} E[D^P] = \frac{1}{T^2} \sum_{i=1}^{T} \epsilon_{i,t} (u_i U' + u_i' C F' + f_i'C' U') \hat{H}^P V_T^{-1}.
\]

(A49)

Ignoring $V_T^{-1}$, the first term on the right is
\[
\frac{1}{T^2} \sum_{i=1}^{T} \epsilon_{i,t} u_i u_i' \hat{H}^P = \frac{1}{T^2} \sum_{i=1}^{T} \epsilon_{i,t} u_i U' D^P + \frac{1}{T^2} \sum_{i=1}^{T} \epsilon_{i,t} u_i U' F H.
\]

Clearly,
\[
\left| \left| \frac{1}{T} \sum_{t \neq s} \epsilon_{i,t} u_i u_s \right| \right| \leq \left( \frac{1}{T} \sum_{t \neq s} || \epsilon_{i,t} ||^2 \right)^{1/2} \sqrt{N} \left( \frac{1}{T} \sum_{t \neq s} || N^{-1/2} u_i u_s ||^2 \right)^{1/2} = O_p(\sqrt{N}),
\]
giving
\[
\left| \left| \frac{1}{T^2} \sum_{i=1}^{T} \sum_{t=1}^{T} \epsilon_{i,t} u_i u_i' D^P \right| \right| = \left| \left| \frac{1}{T^2} \sum_{i=1}^{T} \sum_{t=1}^{T} \epsilon_{i,t} u_i u_i' d_s^{PC} \right| \right| \\
\leq \left( \frac{1}{T} \sum_{s=1}^{T} \left| \left| \frac{1}{T} \sum_{t=1}^{T} \epsilon_{i,t} u_i u_s \right| \right| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} || d_s^{PC} ||^2 \right)^{1/2} \\
= O_p(\sqrt{N})[O_p(N^{-1/2}) + O_p(T^{-1/2})] \\
= O_p(1) + O_p(\sqrt{NT^{-1/2}}).
Furthermore,
\[
\left|\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| = \left|\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| \\
\leq \left|\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| + \left|\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| \\
\leq \frac{1}{T} \left|\frac{1}{T^3/2} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| + \left|\frac{N}{T} \sqrt{T} \sum_{j \neq t}^{N} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| \\
+ \left|\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| + \left|\frac{T}{\sqrt{T} N \sqrt{T}} \sum_{j \neq t}^{N} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| \\
= O_p(T^{-1}) + O_p(NT^{-3/2}) + O_p(T^{-1/2}) + O_p(\sqrt{NT^{-1}}) \\
= O_p(NT^{-3/2}) + O_p(T^{-1/2}) + O_p(\sqrt{NT^{-1}}),
\]
and so, with \( ||H|| = O_p(N^{-2a}) \),
\[
\left|\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| \leq \left|\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| + \left|\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| ||H|| \\
= O_p(1) + O_p(\sqrt{NT^{-1/2}}) + O_p(N^{1-2a}T^{-3/2}).
\]  

(A50)

Let us now consider the second term in the expansion of \( NT^{-1}E^pD^p \). By definition, 
\[
H = T^{-1}Q F^p D^p V_T^{-1} = N^{-2a} T^{-1} Q^0 F^p D^p V_T^{-1},
\]
or, since \( Q^0 \) is invertible, 
\[
T^{-1}F^p D^p V_T^{-1} = N^{2a} (Q^0)^{-1} H = (Q^0)^{-1} H. \]
Hence, \( ||T^{-1}F^p D^p|| = ||(Q^0)^{-1}H^T V_T|| = O_p(1) \). Moreover,
\[
\left|\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{N} \epsilon_{i,t} u'_i u'_s C_s' \right| = \left|\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{N} \epsilon_{i,t} u'_i u'_s C_s' \right| \\
\leq N^{-a} \left|\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{N} \epsilon_{i,t} u'_i u'_s C_s' \right| + \frac{N^{1/2-a}}{\sqrt{T}} \left|\frac{1}{T} \sum_{t=1}^{T} \sum_{s \neq t}^{N} \epsilon_{i,t} u'_i u'_s C_s' \right| \\
= O_p(N^{-a}) + O_p(N^{1/2-a} T^{-1/2}).
\]
Finally, since
\[
\left|\frac{1}{T} \sum_{t=1}^{T} \epsilon_{i,t} u'_i u'_s f_s \right| \leq \sqrt{N} \left|\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \epsilon_{i,t} u'_i f_s \right| ||(NT)^{-1/2} C^p U' F|| = O_p(\sqrt{N}),
\]
and
\[\left\| \frac{1}{T^{3/2}} \sum_{i=1}^{T} e_{i,t_i} f_{i} C^{0} U^{'} D^{PC} \right\| \]
\[= \left\| \frac{1}{T^{3/2}} \sum_{i=1}^{T} \sum_{s=1}^{T} e_{i,t_i} f_{i} C^{0} u_s d_{i,s}^{PC} \right\| \]
\[\leq \sqrt{N} \left( \frac{1}{T} \sum_{s=1}^{T} \left\| \frac{1}{\sqrt{T}} \sum_{i=1}^{T} e_{i,t_i} f_{i} (N^{-1/2} C^{0} u_s) \right\| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} ||d_{i,s}^{PC}||^2 \right)^{1/2} \]
\[= \sqrt{N}[O_p(N^{-1/2}) + O_p(T^{-1/2})] = O_p(1) + O_p(\sqrt{NT^{-1/2}}),\]

the order of the third term becomes
\[\left\| \frac{1}{T^2} \sum_{i=1}^{T} e_{i,t_i} C^{0} U^{'} \hat{f}^{PC} \right\| \]
\[\leq \frac{1}{N^{3a} T} \left\| \frac{1}{T} \sum_{t=1}^{T} e_{i,t} f_{i} C^{0} U^{'} F \right\| \|H^{0}\| + \frac{1}{N^{a} \sqrt{T}} \left\| \frac{1}{T^{3/2}} \sum_{i=1}^{T} e_{i,t_i} f_{i} C^{0} U^{'} D^{PC} \right\| \]
\[= N^{-3a} T^{-1} O_p(\sqrt{N}) + N^{-a} T^{-1/2}[O_p(1) + O_p(\sqrt{NT^{-1/2}})] \]
\[= O_p(N^{1/2-3a} T^{-1}) + O_p(N^{-a} T^{-1/2}) + O_p(N^{1/2-a} T^{-1}) \]
\[= O_p(N^{-a} T^{-1/2}) + O_p(N^{1/2-a} T^{-1}). \]

(A51)

By adding the above results,
\[||NT^{-1} E^{D^{PC}}|| \leq \left\| \frac{1}{T^2} \sum_{i=1}^{T} e_{i,t_i} (u_i U^{'} + u_i CF^{'} + f_{i} C^{0} U^{'}) \hat{f}^{PC} V_{T}^{-1} \right\| \]
\[= O_p(1) + O_p(\sqrt{NT^{-1/2}}) + O_p(N^{1/2-a} T^{-3/2}). \]
(A52)

The proof of the second result is very similar to that of the first. We begin by noting that
\[\frac{1}{T} \sum_{i=1}^{N} E^{D^{PC}} = \frac{1}{NT^2} \sum_{i=1}^{N} e_{i,t_i} (u_i U^{'} + u_i CF^{'} + f_{i} C^{0} U^{'}) \hat{f}^{PC} V_{T}^{-1}, \]
(A53)

where
\[\frac{1}{NT^2} \sum_{i=1}^{N} e_{i,t_i} u_i^{PC} = \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t_i} u_i^{D^{PC}} + \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t_i} u_i^{U^{'} \hat{F} \hat{H}}.\]

Clearly,
\[\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t_i} u_i^{u_s} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t_i} u_i^{u_t} u_j^{s} + \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t_i} u_i^{u_t} u_j^{s} = O_p(1) + O_p(T^{-1}) = O_p(1),\]

40
Thus, and therefore

\[
\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i u_s = \frac{\sqrt{N}}{T} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_{i,t} (N^{-1} u'_i u_s) + \frac{1}{NT} \sum_{i=1}^{N} \sum_{t \neq s} e_{i,t} u'_i u_s = O_p(\sqrt{NT}^{-1}) + O_p(1).
\]

Thus,

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i U'D_{PC} \right\| = \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i u_s d_s \right\| \leq \left( \frac{1}{T} \sum_{s=1}^{T} \left\| \frac{1}{NT} \sum_{i=1}^{N} e_{i,t} u'_i u_s \right\|^2 \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} \left\| d_s \right\|^2 \right)^{1/2} = [O_p(\sqrt{NT}^{-1}) + O_p(1)] [O_p(N^{-1/2}) + O_p(T^{-1/2})] = O_p(N^{-1/2}) + O_p(\sqrt{NT}^{-3/2}) + O_p(T^{-1/2}).
\]

Moreover,

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i U'F \right\| = \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i u_s f_s \right\| \leq \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i u_s f_s \right\| + \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t \neq s} e_{i,t} u'_i u_s f_s \right\| \leq \frac{1}{T} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i u_s f_s \right\| + \left\| \frac{\sqrt{N}}{T^{3/2}} \frac{1}{\sqrt{NT^{3/2}}} \sum_{i=1}^{N} \sum_{t \neq s} e_{i,t} u'_i u_s f_s \right\| = O_p(T^{-1}) + O_p(\sqrt{NT}^{-3/2}) + O_p((NT)^{-1/2}) + O_p(T^{-1}) = O_p(\sqrt{NT}^{-3/2}) + O_p((NT)^{-1/2}) + O_p(T^{-1}),
\]

from which we deduce that, with \( \|\overline{H}\| = O_p(N^{-2a}) \),

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i U'F_{PC} \right\| \leq \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i U'D_{PC} \right\| + \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u'_i U'F \right\| \|H\| = O_p(N^{-1/2}) + O_p(\sqrt{NT}^{-3/2}) + O_p(T^{-1/2}) + O_p(N^{-2a}T^{-1}).
\]
Consider the second term in the expansion of \( T^{-1} \sum_{i=1}^{N} E[D]'\mathbf{C} \mathbf{F}' \mathbf{V}' \). We have

\[
\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i C T^{-1} \mathbf{F}' \mathbf{V}^{-1} = \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i C \mathbf{0}'(\mathbf{Q}^0)^{-1} \mathbf{H}^0
\]

\[
= \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i C' \mathbf{0}'(\mathbf{Q}^0)^{-1} \mathbf{H}^0.
\]

where \( E(e_{i,t} u_i') = E(e_{i,t}(v_{i,t} + e_{i,t}' \beta, e_{i,t}')) = (\Sigma e_{i,t} \beta, \Sigma e_{i,t}) = \Sigma e_{i,t} (\beta, I_m)\), suggesting that

\[
\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i c_{i,t} C' \mathbf{0}'
\]

\[
= \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i c_{i,t} C' \mathbf{0}' + \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i C' \mathbf{0}'
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} \Sigma e_{i,t}(\beta, I_m) C' \mathbf{0}' + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left| e_{i,t} u_i - C e_{i,t}(\beta, I_m) \right| C' \mathbf{0}'
\]

\[
+ \frac{1}{\sqrt{T}} \frac{1}{N \sqrt{T}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j \neq i} e_{i,t} u_i C' \mathbf{0}'
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} \Sigma e_{i,t}(\beta, I_m) C' \mathbf{0}' + O_p((NT)^{-1/2}) + O_p(T^{-1/2})
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} \Sigma e_{i,t}(\beta, I_m) C' \mathbf{0}' + O_p(T^{-1/2}).
\]

Therefore,

\[
\frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i' C \mathbf{F}' \mathbf{V}' \mathbf{V}^{-1}
\]

\[
= \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} u_i C' \mathbf{0}'(\mathbf{Q}^0)^{-1} \mathbf{H}^0
\]

\[
= \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \Sigma e_{i,t}(\beta, I_m) C' \mathbf{0}'(\mathbf{Q}^0)^{-1} \mathbf{H}^0 + O_p(N^{-a}T^{-1/2}).
\]

One term remains, the order of which is given by

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t} f_i C' \mathbf{U}' \mathbf{F} \right\|
\]

\[
\leq \frac{1}{N^a T} \left\| \frac{1}{NT} \sum_{i=1}^{N} e_{i,t} C' \mathbf{0}' \mathbf{U}' \right\| \left\| \mathbf{H} \right\| + \frac{1}{N^a \sqrt{T}} \left\| \frac{1}{NT^{3/2}} \sum_{i=1}^{N} e_{i,t} C' \mathbf{0}' \mathbf{U}' \mathbf{D}' \right\|
\]

\[
= N^{-a} T^{-1} O_p(N^{-2a}) + N^{-a} T^{-1/2} [O_p(N^{-1/2}) + O_p(T^{-1/2})]
\]

\[
= O_p(N^{-(a+1)/2} T^{-1/2}) + O_p(N^{-a} T^{-1}),
\]

(A56)
as follows from noting that
\[
\left\| \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{i,t} f_i' C^0 u_t' F \right\| \leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{i,t} f_i' \right\| \left\| (NT)^{-1/2} C^0 u_t' F \right\| = O_p(1),
\]
and
\[
\left\| \frac{1}{NT^{3/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{i,t} f_i' C^0 u_t' D^p \right\| = \left\| \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{i,t} f_i' (N^{-1/2} C^0 u_t') \right\|^{1/2} \left( \frac{1}{T} \sum_{i=1}^{T} \| d^p_i \|^2 \right)^{1/2} = O_p(N^{-1/2}) + O_p(T^{-1/2}).
\]

Direct substitution now yields
\[
\frac{1}{T} \sum_{i=1}^{N} E[D^p] = \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{i,t} (u_t' U' + u_t' CF + f_i' C' U') \hat{F}^p V_t^{-1} = \frac{1}{Na+1} \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{i,t} 1_i (\beta, I_i) \tilde{C}_i^0 (\tilde{Q}^0)^{-1} \tilde{H}^0 + O_p(R_5), \tag{A57}
\]
where
\[
R_5 = N^{-1/2} + T^{-1/2} + \sqrt{NT^{-3/2}}.
\]
This establishes the second result, and hence the proof of the lemma is complete. ■

**Lemma CCE6.** Under Assumptions ERR, LAM and RK–CCE,
\[
\| NT^{-1} v_i' D_{CCE} \| = O_p(1) + O_p(\sqrt{NT}^{-1/2}),
\]
\[
\frac{1}{N} \sum_{i=1}^{N} NT^{-1} v_i' D_{CCE} (\tilde{C}^0) - (\Lambda_0') = \frac{1}{N} \sum_{i=1}^{N} \sigma_{v,i}^2 (1, 0)(\tilde{C}^0) - (\Lambda_0') + O_p(T^{-1/2}).
\]

**Proof of Lemma CCE6**

Consider the first result. Note that \( E(v_{i,t} u_{j,t}') = E[v_{i,t} (v_{i,t} + e'_{i,t} \beta, e'_{i,t})] = (\sigma_{v,i}^2, 0) \). Making use of this and
\[
NT^{-1} v_i' D_{CCE} = \frac{N}{T} \sum_{i=1}^{T} v_{i,t} \left( \frac{1}{N} \sum_{j=1}^{N} u_{j,t}' = \frac{1}{T} \sum_{i=1}^{T} \sum_{j=1}^{N} v_{i,t} u_{j,t}' = \frac{1}{T} \sum_{i=1}^{T} v_{i,t} u_{j,t}' + \frac{1}{T} \sum_{i=1}^{T} \sum_{j \neq i} v_{i,t} u_{j,t}' \right) \]
\[
= \sigma_{2,i}^2 (1, 0) + \frac{1}{T} \sum_{i=1}^{T} [v_{i,t} u_{j,t}' - \sigma_{v,i}^2 (1, 0)] + \frac{1}{T} \sum_{i=1}^{T} \sum_{j \neq i} v_{i,t} u_{j,t}'.
\]
we obtain
\[
\left| |NT^{-1}v'_i D^{\text{CCE}}| \right| \leq \sigma^2_{\epsilon,i} + \frac{1}{\sqrt{T}} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} [v_{i,t}u'_{i,t} - \sigma^2_{\epsilon,i}(1,0)] \right| + \frac{\sqrt{N}}{\sqrt{T}} \left| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} v_{i,t}u'_{i,t} \right|
\]

\[= O_p(1) + O_p(T^{-1/2}) + O_p(\sqrt{NT^{-1/2}}) = O_p(1) + O_p(\sqrt{NT^{-1/2}}),\]
as required for the first result.

In order to show the second result we again make use of the above expansion of $NT^{-1}v'_i D^{\text{CCE}}$, which, together with
\[
\left| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} [v_{i,t}u'_{i,t} - \sigma^2_{\epsilon,i}(1,0)](C^0)^{-1}(A^i_0)' \right| = O_p(1),
\]
\[
\left| \frac{1}{N\sqrt{T}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j \neq i}^{N} v_{i,t}u'_{j,t}(C^0)^{-1}(A^i_0)' \right| = O_p(1),
\]
can be used to show that
\[
\frac{1}{N} \sum_{i=1}^{N} NT^{-1}v'_i D^{\text{CCE}}(C^0)^{-1}(A^i_0)' = \frac{1}{N} \sum_{i=1}^{N} \sigma^2_{\epsilon,i}(1,0)(C^0)^{-1}(A^i_0)' + O_p(T^{-1/2}).
\]

**Lemma PC6.** Under Assumptions ERR, LAM and RK–PC,
\[
\frac{1}{N} \sum_{i=1}^{N} NT^{-1}v'_i D^{\text{PC}}(H^0)^{-1}(A^i_0)' = \frac{1}{N^{1+a}} \sum_{i=1}^{N} \sigma^2_{\epsilon,i}(1,0)C^0_{i,i}^{-1}(A^i_0)' + O_p(N^{-a} T^{-1/2})
\]
\[+ O_p((NT)^{-1/2}) + O_p(T^{-1}) + O_p(N^{1/2}T^{-3/2}),\]
where
\[
R_6 = N^{-1/2} + (1 + N^{1/2-a})T^{-1/2} + \sqrt{NT^{-1}} + NT^{-3/2}.
\]

**Proof of Lemma PC6**

This proof is very similar to Proof of Lemma PC5. Consider the first result. We have
\[
NT^{-1}v'_i D^{\text{PC}} = \frac{1}{T^2} \sum_{i=1}^{T} v_{i,t}(u'_i U + u'_i CF + f'_i C'U') \hat{F}^{\text{PC}}V^{-1}.
\]
(A58)

Here
\[
\frac{1}{T^2} \sum_{i=1}^{T} v_{i,t}u'_i U' \hat{F}^{\text{PC}} = \frac{1}{T^2} \sum_{i=1}^{T} v_{i,t}u'_i U' D^{\text{PC}} + \frac{1}{T^2} \sum_{i=1}^{T} v_{i,t}u'_i U' \hat{F}H,
\]

44
with
\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{T} v_{i,t} u'_{t} U^F \right\| \\
= \left\| \frac{1}{NT^2} \sum_{i=1}^{T} \sum_{s=1}^{T} v_{i,t} u'_{t} u_s d_{s}^{PC} \right\| \\
\leq \left( \frac{1}{T} \sum_{s=1}^{T} \left\| \frac{1}{NT} \sum_{i=1}^{T} v_{i,t} u'_{t} u_s \right\| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} \left\| d_{s}^{PC} \right\| ^{2} \right)^{1/2} \\
= [O_p(N^{-1}) + O_p(T^{-1}) + O_p((NT)^{-1/2})][O_p(N^{-1/2}) + O_p(T^{-1/2})] \\
= O_p(N^{-3/2}) + O_p(N^{-1}T^{-1/2}) + O_p(T^{-3/2}) + O_p(N^{-1/2}T^{-1}), \quad (A59)
\]
where we have made use of the fact that
\[
\left\| \frac{1}{NT} \sum_{i=1}^{T} v_{i,t} u'_{t} u_s \right\| \\
\leq \left\| \frac{1}{NT} \sum_{i=1}^{T} v_{i,t} u'_{t} u_s \right\| + \left\| \frac{1}{NT} \sum_{j \neq i}^{T} v_{j,t} u'_{t} u_j \right\| \\
\leq \frac{1}{N} \left\| \frac{1}{T} \sum_{i=1}^{T} v_{i,t} u'_{t} u_s \right\| + \frac{1}{N} \left\| \frac{1}{T} \sum_{j \neq i}^{T} v_{j,s} u'_{s} u_j \right\| + \frac{1}{\sqrt{NT}} \left\| \sum_{j \neq i \neq j}^{T} v_{j,t} u'_{t} u_j \right\| \\
= O_p(N^{-1}) + O_p(T^{-1}) + O_p((NT)^{-1/2}).
\]
By using the same steps as in the proof of Lemma PC5, we can further show that
\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{T} v_{i,t} u'_{t} U^F \right\| \\
\leq \frac{1}{NT} \left\| \frac{1}{T} \sum_{i=1}^{T} v_{i,t} u'_{t} u_i f_i \right\| + \frac{1}{\sqrt{N}} \left\| \sum_{j \neq i}^{T} v_{j,t} u'_{t} u_j f_j \right\| \\
+ \frac{1}{\sqrt{NT}} \left\| \sum_{j \neq i}^{T} v_{j,t} u'_{t} u_j f_j \right\| \\
= O_p(N^{-1}T^{-1}) + O_p(T^{-3/2}) + O_p(N^{-1}T^{-1/2}) + O_p(N^{-1/2}T^{-1}) \\
= O_p(T^{-3/2}) + O_p(N^{-1}T^{-1/2}) + O_p(N^{-1/2}T^{-1}),
\]
and so we obtain
\[
\left\| \frac{1}{T^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t} u'_{t} U^F \right\| \\
\leq N \left\| \frac{1}{NT^2} \sum_{i=1}^{T} v_{i,t} u'_{t} U^D \right\| + N^{1-2a} \left\| \frac{1}{NT^2} \sum_{i=1}^{T} v_{i,t} u'_{t} U^F \right\| \left\| \mathbf{H}^0 \right\| \\\n= O_p(N^{-1/2}) + O_p(T^{-1/2}) + O_p(NT^{-3/2}) + O_p(\sqrt{NT^{-1}}), \quad (A60)
\]
Again, in analogy to the proof of Lemma PC5, we may write
\[ \frac{1}{T} \sum_{t=1}^{T} v_{t,i} u_i'C T^{-1} F \hat{\Phi}^p C T^{-1} = N^{-\alpha} \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{N} v_{t,i} u_i' C_j^0 ( \hat{Q}^0 )^{-1} \hat{H}^0, \]
where \( E(v_{t,i} u_i') = E[v_{t,i} ( v_{t,i} + e_i^t \beta, e_i^t )] = ( \sigma_{v,i}^2 0 ) \) and
\[ \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{N} v_{t,i} u_i' C_j^0 \]
\[ = \frac{1}{T} \sum_{t=1}^{T} v_{t,i} u_i' C_i^0 + \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{N} v_{t,i} u_i' C_j^0 \]
\[ = \sigma_{v,i}^2 (1,0) C_i^0 + \frac{1}{\sqrt{T}} \frac{1}{\sqrt{N}} \sum_{t=1}^{T} ( v_{t,i} u_i' - \sigma_{v,i}^2 (1,0) ) C_i^0 + \frac{\sqrt{N}}{\sqrt{N T}} \sum_{t=1}^{T} \sum_{j=1, j \neq i}^{N} v_{t,i} u_i' C_j^0 \]
\[ = \sigma_{v,i}^2 (1,0) C_i^0 + O_p(\sqrt{N T}^{-1/2}), \]
yielding
\[ \frac{1}{T^2} \sum_{t=1}^{T} v_{t,i} u_i'C F \hat{\Phi}^p C T^{-1} = N^{-\alpha} \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{N} v_{t,i} u_i' C_j^0 ( \hat{Q}^0 )^{-1} \hat{H}^0 \]
\[ = N^{-\alpha} \sigma_{v,i}^2 (1,0) C_i^0 ( \hat{Q}^0 )^{-1} \hat{H}^0 + O_p(N^{1/2-\alpha} T^{-1/2}). \quad (A61) \]

Finally, since
\[ \left\| \frac{1}{N T^2} \sum_{t=1}^{T} v_{t,i} f_i' C U' F \right\| = \frac{1}{N^{a+1/2} T} \left\| \frac{1}{\sqrt{T}} \sum_{i=1}^{T} v_{t,i} f_i' \right\| \left\| (NT)^{-1/2} C^0 U' F \right\| \]
\[ = O_p(N^{-(a+1/2)} T^{-1}), \]
and
\[ \left\| \frac{1}{N T^2} \sum_{t=1}^{T} v_{t,i} f_i' C U' D^p \right\| \]
\[ \leq \left( \frac{1}{N^{a+1/2} T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} v_{t,i} f_i' (N^{-1/2} C^0 u_s) \right\| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} \| d_s^p \|^2 \right)^{1/2} \]
\[ = N^{-a+1/2} T^{-1/2} [O_p(N^{-1/2}) + O_p(T^{-1/2})] \]
\[ = O_p(N^{-(a+1/2)} T^{-1/2}) + O_p(N^{-(a+1/2)} T^{-1}), \]
we have
\[ \left\| \frac{1}{T^2} \sum_{t=1}^{T} v_{t,i} f_i' C U' F \hat{\Phi}^p \right\| \leq \left\| \frac{1}{T^2} \sum_{t=1}^{T} v_{t,i} f_i' C U' F \right\| \left\| \hat{H} \right\| + \left\| \frac{1}{T^2} \sum_{t=1}^{T} v_{t,i} f_i' C U' D^p \right\| \]
\[ = O_p(N^{-(a+1)} T^{-1/2}) + O_p(N^{-(a+1/2)} T^{-1}), \quad (A62) \]
which can be combined with the above results to obtain

\[
NT^{-1}v'_{i}D^{PC} = \frac{1}{T^2} \sum_{i=1}^{T} v_{i,t}(u'_{i}U' + u'_{i}CF' + f_{i}'C'U')\hat{F}^{PC}V_{T}^{-1}
\]

\[
= N^{-a}a_{ii}'^{2}(1,0)C_{i}'(\mathbf{Q})^{-1}\mathbf{H}^{0} + O_p(R_{6}),
\]

(A63)

where

\[
R_{6} = N^{-(a+1)/2} + (1 + N^{1/2-a})T^{-1/2} + \sqrt{NT}^{-1} + NT^{-3/2}.
\]

Hence,

\[
||NT^{-1}v'_{i}D^{PC}|| = O_p(N^{-a}) + O_p(R_{6}),
\]

(A64)

as required.

As in the proof of CCE6, the second result follows from a simple manipulation of the proof of the first. Therefore, only essential details will be given. Note first that

\[
\frac{1}{N} \sum_{i=1}^{N} NT^{-1}v'_{i}D^{PC}(\mathbf{H}^{0})^{-1}(\Lambda_{i}^{0})'
\]

\[
= \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t}(u'_{i}U' + u'_{i}CF' + f_{i}'C'U')\hat{F}^{PC}V_{T}^{-1}(\mathbf{H}^{0})^{-1}(\Lambda_{i}^{0})'.
\]

(A65)

We begin by considering the first and third terms on the right. These are negligible and therefore the analysis is unaffected by the scaling of \(V_{T}^{-1}(\mathbf{H}^{0})^{-1}(\Lambda_{i}^{0})'\). Ignoring this matrix, the order of the first term is given by

\[
\left|\frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t}u'_{i}U'\hat{F}^{PC}\right| \leq \left|\frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t}u'_{i}U'D^{PC}\right| + \left|\frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t}u'_{i}F\right| ||\mathbf{H}||
\]

\[
= O_p(N^{-(a/2+1)}) + O_p((NT)^{-1/2}) + O_p(T^{-1}) + [O_p((NT)^{-1/2}) + O_p(T^{-1})]O_p(N^{-2a})
\]

\[
= O_p(N^{-(a/2+1)}) + O_p((NT)^{-1/2}) + O_p(T^{-1}).
\]

(A66)

In order to appreciate that this must be so, note that

\[
\left|\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t}u'_{i}\right| \leq \left|\frac{1}{\sqrt{N}} \left|\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t}u'_{i}\right|\right| + \left|\frac{1}{\sqrt{T}} \left|\frac{1}{\sqrt{T}} \sum_{i=j}^{N} \sum_{t=1}^{T} v_{i,t}u'_{i}\right|\right|\right| \leq O_p(N^{-1/2}) + O_p(T^{-1/2}).
\]

47
and therefore

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T U D^{PC} \right\| = \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T u_s d^{PC}_s \right\| \\
\leq \left( \frac{1}{T} \sum_{s=1}^{T} \left\| \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T u_s \right\| \right)^{1/2} \left( \frac{1}{T} \sum_{s=1}^{T} \left\| d^{PC}_s \right\|^2 \right)^{1/2} \\
= O_p(\frac{N^{-1/2}}{T^{1/2}}) + O_p(\frac{T^{-1/2}}{N^{-1/2}} + O_p(\frac{T^{-1/2}}{N^{-1/2}})) \\
= O_p(\frac{N^{-1}}{T^{1/2}}) + O_p(\frac{(NT)^{-1/2}}{T^{1/2}}) + O_p(T^{-1}).
\]

Moreover,

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T U F \right\| \\
\leq \left\| \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T u_j^T f_i \right\| + \frac{\sqrt{N}}{T^{3/2}} \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} v_{ij} u_i^T u_j f_i \right\| \\
= O_p(T^{-1}) + O_p(\frac{N^{1/2}}{T^{3/2}}) + O_p(\frac{(NT)^{-1/2}}{T^{1/2}}) + O_p(T^{-1}) \\
= O_p(\frac{(NT)^{-1/2}}{T^{1/2}}) + O_p(T^{-1}) + O_p(\frac{N^{1/2}}{T^{3/2}}).
\]

The second term is

\[
\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T C T^{-1} F \hat{P}^{PC} \hat{V}^{-1}(\bar{H})^0 - (\Lambda_i^0)^0 \]

\[
= \frac{1}{N^{a+1}T} \sum_{j=1}^{N} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T C_i^0 (\bar{Q})^{-1} (\Lambda_i^0)^0 \]

\[
= \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \sum_{t=1}^{T} v_{ij} u_i^T C_i^0 (\bar{Q})^{-1} (\Lambda_i^0)^0 + \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \sum_{t=1}^{T} v_{ij} u_i^T C_i^0 (\bar{Q})^{-1} (\Lambda_i^0)^0 \]

\[
= \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sigma^2_{e_i}(1,0) C_i^0 (\bar{Q})^{-1} (\Lambda_i^0)^0 \]

\[
+ \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{ij} u_i^T d_i^0 (\bar{Q})^{-1} (\Lambda_i^0)^0 \]

\[
+ \frac{1}{N^{a}T} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \sum_{t=1}^{T} v_{ij} u_i^T C_i^0 (\bar{Q})^{-1} (\Lambda_i^0)^0 \]

\[
= \frac{1}{N^{a+1}T} \sum_{i=1}^{N} \sigma^2_{e_i}(1,0) C_i^0 (\bar{Q})^{-1} (\Lambda_i^0)^0 + O_p(N^{-a}T^{-1/2}). \quad (A67)
\]
The third and final term is, again ignoring the scaling by $V_T^{-1}(\mathbf{H}_0^0 - (\Lambda_i^0)'$, 

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t} f'_i C' \mathbf{U}' \mathbf{F}^PC \right\| \leq \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t} f'_i C' \mathbf{U}' \mathbf{F} \right\| \left\| \mathbf{H} \right\| + \left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t} f'_i C' \mathbf{U}' \mathbf{D}^PC \right\| = O_p(N^{-(3\alpha+1)/2}T^{-1/2}) + O_p(N^{-\alpha}T^{-1}),
\]

which uses

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t} f'_i C' \mathbf{U}' \mathbf{F} \right\| = \frac{1}{N^a T} \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t} f'_i \right\| \left\| (NT)^{-1/2} C^0 \mathbf{U}' \mathbf{F} \right\| = O_p(N^{-\alpha}T^{-1}),
\]

and

\[
\left\| \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} v_{i,t} f'_i C' \mathbf{U}' \mathbf{D}^PC \right\| = O_p(N^{-3\alpha-1/2}T^{-1/2}) + O_p(N^{-\alpha}T^{-1}).
\]

The above results imply

\[
\frac{1}{N} \sum_{i=1}^{N} NT^{-1} v'_i \mathbf{D}^PC (\mathbf{H}_0^0 - (\Lambda_i^0)'). = \frac{1}{N^{1+\alpha}} \sum_{i=1}^{N} \sigma_{v,i}^2 (1, 0) C^0 (\mathbf{Q}_i^0)' - (\Lambda_i^0)' + O_p(N^{-\alpha}T^{-1/2}) + O_p((NT)^{-1/2}) + O_p(T^{-1}) + O_p(N^{1/2}T^{-3/2}). \]  \hspace{1cm} (A69)

This establishes the second result and hence the proof of the lemma is complete. \hfill \blacksquare

**Lemma CCE7** Under Assumptions LAM and RK–CCE,

\[
T \left\| \mathbf{S}_{\text{CCE}}^{-} - (\mathbf{C}' \mathbf{F}' \mathbf{F})^{-} \right\| = O_p(N^{\alpha-1/2}T^{-1/2}) + O_p(N^{2\alpha-1}),
\]

\[
\left\| (T^{-1} \mathbf{S}_{\text{CCE}})^{-} \right\| = O_p(N^{2\alpha}).
\]

**Proof of Lemma CCE7**
From
\[
\bar{\mathbf{f}}_{\text{CCE}} = (\mathbf{F}\mathbf{C} + \mathbf{D}_{\text{CCE}})'(\mathbf{F}\mathbf{C} + \mathbf{D}_{\text{CCE}}) = \mathbf{C}'\mathbf{F}\mathbf{C} + \mathbf{D}_{\text{CCE}}'\mathbf{F}\mathbf{C} + \mathbf{C}'\mathbf{F}'\mathbf{D}_{\text{CCE}} + \mathbf{D}_{\text{CCE}}'\mathbf{D}_{\text{CCE}},
\]
and
\[
||\mathbf{S}_{\text{CCE}} - \mathbf{C}'\mathbf{F}\mathbf{C}|| \leq ||\bar{\mathbf{f}}_{\text{CCE}} - \mathbf{C}'\mathbf{F}\mathbf{C}||,
\]
we obtain
\[
T^{-1}||\mathbf{S}_{\text{CCE}} - \mathbf{C}'\mathbf{F}\mathbf{C}|| \leq 2N^{-(a+1/2)}T^{-1/2}||\sqrt{N}T^{-1/2}\mathbf{D}_{\text{CCE}}'\mathbf{F}||||\mathbf{C}'|| + N^{-1}||NT^{-1}\mathbf{D}_{\text{CCE}}'\mathbf{D}_{\text{CCE}}|| = O_p(N^{-(a+1/2)}T^{-1/2}) + O_p(N^{-1}). \tag{A70}
\]

By the properties of the Moore–Penrose inverse, if \(\mathbf{A} = \mathbf{B}\mathbf{C}\), where \(\mathbf{B}\) and \(\mathbf{C}\) have full column and row rank, respectively, then \(\mathbf{A}^{-} = \mathbf{C}^{-}\mathbf{B}^{-}\). Applying this (twice) to \((\mathbf{C}'\mathbf{F}\mathbf{C})^{-}\) we obtain \((\mathbf{C}'\mathbf{F}\mathbf{C})^{-} = \mathbf{C}^{-}(\mathbf{F}'\mathbf{F})^{-}\mathbf{C}' = \mathbf{C}^{-}(\mathbf{F}'\mathbf{F})^{-1}(\mathbf{C}')^{-}\), where \(\mathbf{C}^{-} = \mathbf{C}'(\mathbf{C}'\mathbf{C})^{-1}\), \((\mathbf{C}')^{-} = (\mathbf{C}'\mathbf{C}')^{-1}\), and the last equality holds because \(\mathbf{C}'\) has full row (column) rank and \(\mathbf{F}'\mathbf{F}\) is nonsingular. Hence, since \(\mathbf{F}'\mathbf{F}\mathbf{C}\mathbf{C}^{-} = (\mathbf{F}'\mathbf{F})^{-1}(\mathbf{C}')^{-}\mathbf{C}' = \mathbf{I}_r\),
\[
\mathbf{C}[\mathbf{S}_{\text{CCE}} - (\mathbf{C}'\mathbf{F}\mathbf{C})^{-}]\mathbf{C}' = \mathbf{C}[\mathbf{S}_{\text{CCE}} - \mathbf{C}^{-}(\mathbf{F}'\mathbf{F})^{-1}(\mathbf{C}')^{-}]\mathbf{C}'
\]
and therefore
\[
T||\mathbf{S}_{\text{CCE}} - (\mathbf{C}'\mathbf{F}\mathbf{C})^{-}|| \leq N^{2\alpha}||\mathbf{(T}^{-1}\mathbf{S}_{\text{CCE}})^{-}||T^{-1}||\mathbf{S}_{\text{CCE}} - \mathbf{C}'\mathbf{F}\mathbf{C}||||\mathbf{(C)}^{-}||^2||\mathbf{(T}^{-1}\mathbf{F}'\mathbf{F}^{-1})|| = O_p(N^{a-1/2}T^{-1/2}) + O_p(N^{2\alpha-1}). \tag{A71}
\]

or
\[
||\mathbf{(T}^{-1}\mathbf{S}_{\text{CCE}})^{-}|| = O_p(N^{2\alpha}) + O_p(N^{a-1/2}T^{-1/2}) + O_p(N^{2\alpha-1}) = O_p(N^{2\alpha}). \tag{A72}
\]

\[\text{Lemma PC7} \quad \text{Under Assumptions LAM and RK–PC,}\]
\[
T||\mathbf{S}_{\text{PC}} - (\mathbf{H}'\mathbf{F}'\mathbf{H})^{-}|| = O_p(N^{a-1/2}T^{-1/2}R_{10}),
\]
where
\[ R_{10} = N^{-1/2} \sqrt{T} + N^{-5a} + \sqrt{N} T^{-1/2}. \]

**Proof of Lemma PC7**

By using the same trick as in Proof of CCE7 for expanding \( C[S_{PC} - (F'F)^{-1}]C' \), \( T^{-1} S_{PC} = I_p, O_p(R_{10}) + O_p(N^{-2a} R_1) = O_p(R_{10}) \) and \( ||H|| = O_p(N^{-2a}) \), we obtain
\[
T|S_{PC} - (H'F'H)^{-1}|| \\
\leq (NT)^{-1/2}||T^{-1} S_{PC} - H'F'H||||(T^{-1} H'F'H)^{-1}|| \\
= (NT)^{-1/2}||T^{-1} S_{PC} - H'F'H||\sqrt{N} T^{-1/2}||D^{PC} C^{PC} + H'F'D^{PC}||||T^{-1} H'F'H||^{-1} \\
\leq (NT)^{-1/2}||T^{-1} S_{PC} - H'F'H||\sqrt{N} T^{-1/2}||D^{PC} C^{PC} + H'F'D^{PC}||||H||^{-2}||(T^{-1} F')||^{-1} \\
= (NT)^{-1/2}[O_p(R_{10}) + O_p(N^{-2a} R_1)]O_p(N^{4a}) = O_p(N^{4a - 1/2} T^{-1/2} R_{10}). \] (A73)

**Lemma CCE8.** Under Assumptions ERR and LAM,
\[ ||NT^{-1}X'D^{CCE}|| = O_p(1) + O_p(\sqrt{NT}^{-1/2}). \]

**Proof of Lemma CCE8**

From Lemmas CCE3 and CCE5,
\[
||NT^{-1}X'D^{CCE}|| = ||NT^{-1}(FA_i' + E_i)'D^{CCE}|| \\
\leq N^{1/2 - a} T^{-1/2}||A_i'||||\sqrt{NT}^{-1/2} F'D^{CCE}|| + ||NT^{-1}E_i'D^{CCE}|| \\
= O_p(N^{1/2 - a} T^{-1/2}) + O_p(1) + O_p(\sqrt{NT}^{-1/2}) = O_p(1) + O_p(\sqrt{NT}^{-1/2}),
\]
as was to be shown.

**Lemma PC8.** Under Assumptions ERR, LAM and RK–PC,
\[ ||NT^{-1}X'D^{PC}|| = O_p(N^{1/2 - a} T^{-1/2} R_1) + O_p(N^{-a/2}) + O_p(\sqrt{NT}^{-1/2}) + O_p(N T^{-3/2}). \]

**Proof of Lemma PC8**
Lemmas PC3 and PC5 imply

\[ ||NT^{-1}X'D^PC|| = ||NT^{-1}(FA_i' + E_i')D^PC|| \]
\[ \leq N^{1/2 - \alpha}T^{-1/2}||A^0||\sqrt{NT^{-1/2}F'D^PC|| + ||NT^{-1}E'D^PC||} \]
\[ = O_p(N^{1/2 - \alpha}T^{-1/2}R_1) + O_p(N^{-\alpha/2}) + O_p(\sqrt{NT^{-1/2}}) + O_p(NT^{-3/2}). \]

\[ \blacksquare \]

**Lemma CCE9.** Under Assumptions ERR, LAM, RK–CCE, KAP and \( \kappa > \max\{0, 2\alpha - 1\} \),

\[ T^{-1}X'M_{F_{CCE}}X_i = \Sigma_e + o_p(1), \]
\[ \frac{1}{NT} \sum_{i=1}^{N} X'M_{F_{CCE}}X_i = \Sigma_e + o_p(1). \]

**Proof of Lemma CCE9**

We can expand the expression as

\[ T^{-1}X'M_{F_{CCE}}X_i = T^{-1}X'M_{F_{C}}X_i - T^{-1}X'(M_{F_{C}} - M_{F_{CCE}})X_i. \]  (A74)

Now, from the definitions of \( M_{F_{CCE}} \) and \( M_{F_{C}} \) in (A8), we have

\[ M_{F_{C}} - M_{F_{CCE}} = D^{CCE}S_{F_{CCE}}D^{CCE'} + D^{CCE}S_{F_{CCE}}\Sigma'C'F' \]
\[ + fCS_{F_{CCE}}D^{CCE'} + fC(S_{F_{CCE}} - (C'F'C)^{-1}C'F'). \]  (A75)

Note furthermore that

\[ ||T^{-1}X'F|| \leq N^{-\alpha} ||A^0|| ||T^{-1/2}F'|| + T^{-1/2}||T^{-1/2}E_i'F|| = O_p(N^{-\alpha}) + O_p(T^{-1/2}). \]  (A76)
These two results and Lemmas CCE7 and CCE8 imply

\[
\|T^{-1}\mathbf{X}'_i (\mathbf{M}_{\text{FC}} - \mathbf{M}_{\text{FCE}}) \mathbf{X}_i\|
\]
\[
= N^{-2} \|NT^{-1}\mathbf{X}'_i \mathbf{D}^{\text{CCE}} \|_2^2 \|T^{-1} \mathbf{S}_{\text{FCE}}\| - \|\mathbf{C}\|_2 \|T^{-1} \mathbf{F}' \mathbf{X}_i\|
\]
\[
+ 2N^{-(a+1)} \|NT^{-1}\mathbf{X}'_i \mathbf{D}^{\text{CCE}} \| \|T^{-1} \mathbf{S}_{\text{FCE}}\| - \|\mathbf{C}\|_2 \|T^{-1} \mathbf{F}' \mathbf{X}_i\|
\]
\[
+ N^{-2a} \|T^{-1}\mathbf{X}'_i \mathbf{F}\|_2^2 \|\mathbf{C}\|_2 \|T \mathbf{S}_{\text{FCE}} - (\mathbf{C}' \mathbf{F}' \mathbf{C})^{-1}\|
\]
\[
= N^{-2} \mathbb{O}_p(1) + \mathbb{O}_p(\sqrt{N} T^{-1/2}) \|\mathbf{F}\|_2^2 N^{2a} + \mathbb{O}_p((\sqrt{N} T^{-1/2}) \mathbb{O}_p(N^{-a}) \mathbb{O}_p(NT^{-1/2}) + \mathbb{O}_p(T^{-1/2})
\]
\[
+ N^{-2a} \mathbb{O}_p(\mathbb{O}_p(\mathbb{O}_p(N^{-a}) + \mathbb{O}_p(T^{-1/2})^2 [\mathbb{O}_p(N^{a-1/2} T^{-1/2}) + \mathbb{O}_p(N^{2a-1})]
\]
\[
= \mathbb{O}_p(N^{2a-2}) + \mathbb{O}_p(N^{2a-3/2} T^{-1/2}) + \mathbb{O}_p(N^{2a-1} T^{-1}) + \mathbb{O}_p(N^{-a} T^{-1/2}) + \mathbb{O}_p(N^{a-1/2} T^{-1})
\]
\[
+ \mathbb{O}_p(N^{-a+1/2} T^{-3/2}) + \mathbb{O}_p((\sqrt{N} T^{-1/2}) \mathbb{O}_p(N^{a-1} T^{-1/2}) + \mathbb{O}_p(N^{a-1/2} T^{-1})
\]
\[
+ \mathbb{O}_p(N^{2a-2}) + \mathbb{O}_p(N^{-a+1/2} T^{-3/2}) + \mathbb{O}_p((\sqrt{N} T^{-1/2}) \mathbb{O}_p(N^{a-1} T^{-1/2}) + \mathbb{O}_p(N^{a-1/2} T^{-1})
\]
\[
\frac{1}{\sqrt{N}} \sum_{i=1}^{N} X'_i M_{\text{FCE}} X_i = \Sigma_{e,i} + \mathbb{O}_p(1),
\]
(A77)

which is clearly \(\mathbb{O}_p(1)\) under \(T = N^\kappa\) and \(\kappa > \max\{0, 2a - 1\}\). Also, the first term on the right-hand side of (A74) is given by

\[
T^{-1}\mathbf{X}'_i \mathbf{M}_{\text{FC}} \mathbf{E}_i = T^{-1} \mathbf{E}'_i \mathbf{M}_{\text{FC}} \mathbf{E}_i
\]
\[
= T^{-1} \mathbf{E}'_i \mathbf{E}_i - T^{-1} (T^{-1/2} \mathbf{E}'_i \mathbf{F}) \mathbf{C}^0 (T^{-1} \mathbf{C}^0)' \mathbf{F}' \mathbf{F} \mathbf{E}_i - \mathbf{C}^0 (T^{-1/2} \mathbf{F}' \mathbf{E}_i)
\]
\[
= \Sigma_{e,i} + \mathbb{O}_p(T^{-1}),
\]
(A78)

When taken together these two results imply that if \(\kappa > \max\{0, 2a - 1\}\), we have

\[
T^{-1}\mathbf{X}'_i \mathbf{M}_{\text{FCE}} \mathbf{X}_i = \Sigma_{e,i} + \mathbb{O}_p(1).
\]
(A79)

The second result of the lemma follows from

\[
\left\| \frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}'_i \mathbf{M}_{\text{FCE}} \mathbf{X}_i \right\| \leq \frac{1}{N} \sum_{i=1}^{N} \|T^{-1}\mathbf{X}'_i \mathbf{M}_{\text{FCE}} \mathbf{X}_i\|,
\]

and so we are done. ■
Lemma PC9. Under Assumptions ERR, LAM, RK–PC and KAP,
\[
\frac{1}{NT} \sum_{i=1}^{N} X'_i (M_{\text{FPC}} - M_{\text{FCE}}) X_i = \Sigma_c + o_p(1).
\]

Proof of Lemma PC9

\((NT)^{-1} \sum_{i=1}^{N} X'_i M_{\text{FPC}} X_i\) can be expanded as follows:
\[
\frac{1}{NT} \sum_{i=1}^{N} X'_i M_{\text{FPC}} X_i = \frac{1}{NT} \sum_{i=1}^{N} X'_i M_{\text{FCE}} X_i - \frac{1}{NT} \sum_{i=1}^{N} X'_i (M_{\text{FCE}} - M_{\text{FPC}}) X_i.
\]

From (A75) and Lemmas PC7 and PC8, we obtain
\[
\alpha \text{ which does not impose any restrictions on the admissible values of } \alpha \text{ and } \kappa.
\]

Hence, by using the fact that \((NT)^{-1} \sum_{i=1}^{N} X'_i | M_{\text{FPC}} X_i| \leq N^{-1} \sum_{i=1}^{N} ||T^{-1} X'_i | M_{\text{FPC}} X_i||\), we can proceed analogously to Proof of Lemma CCE9 to obtain
\[
\frac{1}{NT} \sum_{i=1}^{N} X'_i M_{\text{FPC}} X_i = \Sigma_c + o_p(1),
\]

which holds irrespectively of the values of \(\alpha\) and \(\kappa\). \(\blacksquare\)

Lemma CCE10. Under Assumptions ERR, LAM, RK–CCE and KAP,
\[
-T^{-1} X'_i M_{\text{FCE}} D_{\text{CCE}} C^{-1} \lambda_i = N^{-1} (b_{1\text{CCE},i} - b_{2\text{CCE},i}) + O_p(R_8),
\]
\[
-\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i M_{\text{FCE}} D_{\text{CCE}} C^{-1} \lambda_i = \sqrt{TN^{-1/2}} (b_{1\text{CCE}} - b_{2\text{CCE}}) + \sqrt{NT}O_p(R_8),
\]
where

\[
\begin{align*}
\mathbf{b}_{1\text{CCE},i} &= \Lambda_i^0[\mathbf{c}_{i} - \mathbf{\Sigma}_{u}(\mathbf{c}_0^0)^\prime \Lambda_i^0], \\
\mathbf{b}_{2\text{CCE},i} &= \mathbf{\Sigma}_{c,i}(\beta, \mathbf{I}_m)(\mathbf{c}_0^0)^\prime \Lambda_i^0, \\
B_{1\text{CCE}} &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{b}_{1\text{CCE},i}, \\
B_{2\text{CCE}} &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{b}_{2\text{CCE},i}, \\
R_8 &= N^{2\alpha-2} + N^{\alpha-(3+\kappa)/2} + N^{-(1+\kappa)/2} + N^{-(1/2+\kappa)}.
\end{align*}
\]

Proof of Lemma CCE10

Clearly, \( \mathbf{M}_{\text{FCE}} \mathbf{D}^{\text{CCE}} \mathbf{C}^{-} = \mathbf{M}_{\text{FCE}}(\hat{\mathbf{F}}^{\text{CCE}} - \mathbf{F})\mathbf{C}^{-} = -\mathbf{M}_{\text{FCE}} \mathbf{F} \), and therefore

\[
\begin{align*}
-T^{-1} \mathbf{X}_i^\prime \mathbf{M}_{\text{FCE}} \mathbf{D}^{\text{CCE}} \mathbf{C}^{-} \lambda_i &= T^{-1} \mathbf{X}_i^\prime \mathbf{M}_{\text{FCE}} \mathbf{F} \lambda_i \\
&= T^{-1} \Lambda_i \mathbf{F}' \mathbf{M}_{\text{FCE}} \mathbf{F} \lambda_i + T^{-1} \mathbf{E}_i^\prime \mathbf{M}_{\text{FCE}} \mathbf{F} \lambda_i \\
&= \mathbf{k}_1 + \mathbf{k}_2. \tag{A83}
\end{align*}
\]

Consider \( \mathbf{k}_1 \). We can write

\[
\mathbf{C}^\prime \mathbf{F}' \mathbf{M}_{\text{FCE}} \mathbf{F} = \mathbf{D}^{\text{CCE}} \mathbf{M}_{\text{FCE}} \mathbf{D}^{\text{CCE}} = \mathbf{D}^{\text{CCE}} \mathbf{M}_{\text{FCE}} \mathbf{D}^{\text{CCE}} - \mathbf{D}^{\text{CCE}}(\mathbf{M}_{\text{FCE}} - \mathbf{M}_{\text{FCE}}) \mathbf{D}^{\text{CCE}},
\]

from which we obtain

\[
\begin{align*}
\mathbf{k}_1 &= T^{-1} \Lambda_i \mathbf{F}' \mathbf{M}_{\text{FCE}} \mathbf{F} \lambda_i \\
&= T^{-1} \Lambda_i (\mathbf{c}_0^0)^\prime \mathbf{C}^\prime \mathbf{F}' \mathbf{M}_{\text{FCE}} \mathbf{F} \mathbf{C}^{-} \lambda_i \\
&= T^{-1} \Lambda_i (\mathbf{c}_0^0)^\prime \mathbf{D}^{\text{CCE}} \mathbf{M}_{\text{FCE}} \mathbf{D}^{\text{CCE}} \mathbf{C}^{-} \lambda_i + T^{-1} \Lambda_i (\mathbf{c}_0^0)^\prime \mathbf{D}^{\text{CCE}}(\mathbf{M}_{\text{FCE}} - \mathbf{M}_{\text{FCE}}) \mathbf{D}^{\text{CCE}} \mathbf{C}^{-} \lambda_i \\
&= \mathbf{k}_{11} + \mathbf{k}_{12}. \tag{A84}
\end{align*}
\]

First, consider \( \mathbf{k}_{12} \). The decomposition in (A75) suggests that

\[
\begin{align*}
\mathbf{D}^{\text{CCE}}(\mathbf{M}_{\text{FCE}} - \mathbf{M}_{\text{FCE}}) \mathbf{D}^{\text{CCE}} \\
&= \mathbf{D}^{\text{CCE}} \mathbf{D}^{\text{CCE}} \mathbf{S}_{\text{FCE}}^{\prime} \mathbf{D}^{\text{CCE}} \mathbf{D}^{\text{CCE}} + \mathbf{D}^{\text{CCE}} \mathbf{D}^{\text{CCE}} \mathbf{S}_{\text{FCE}}^{\prime} \mathbf{C}^\prime \mathbf{F}' \mathbf{D}^{\text{CCE}} \\
&+ \mathbf{D}^{\text{CCE}} \mathbf{F} \mathbf{C}^{\prime} \mathbf{S}_{\text{FCE}}^{\prime} \mathbf{D}^{\text{CCE}} \mathbf{D}^{\text{CCE}} + \mathbf{D}^{\text{CCE}} \mathbf{F} \mathbf{C}^{\prime} \mathbf{S}_{\text{FCE}}^{\prime} (\mathbf{c}_0^0)^\prime (\mathbf{c}_0^0)^\prime \mathbf{F} \mathbf{D}^{\text{CCE}}. \tag{A85}
\end{align*}
\]
Application of Lemmas CCE3, CCE4 and CCE7 to this expression yields

\[
\| T^{-1} D^{CCE} (M_F\mathbf{C} - M_{\mathbf{CCE}}) D^{CCE} \| \\
\leq N^{-2} \| |NT^{-1} D^{CCE} D^{CCE} \| \| (T^{-1} S_{\mathbf{CCE}})^{-} \| \\
+ 2N^{-(a+3/2)} T^{-1/2} \| \mathbf{C}^0 \| \| NT^{-1} D^{CCE} D^{CCE} \| \| (T^{-1} S_{\mathbf{CCE}})^{-} \| \| \sqrt{NT^{-1/2} F'D^{CCE}} \| \\
+ N^{-(2a+1)} T^{-1} \| \mathbf{C}^0 \| \| NT^{-1/2} D^{CCE} F \| \| S_{\mathbf{CCE}} - (\mathbf{C}' \mathbf{F}' \mathbf{F})^{-} \| \\
= N^{-2} O_p(N^{2a}) + N^{-(a+3/2)} T^{-1/2} O_p(N^{2a}) \\
+ N^{-(2a+1)} T^{-1} [O_p(N^{2a-1}) + O_p(N^{a-1/2} T^{-1/2})] = O_p(R_7),
\]

where

\[ R_7 = N^{2a-2} + N^{a-3/2} T^{-1/2}. \]

This result can in turn be used to show that

\[
\| k_{12} \| = \| T^{-1} \Lambda_i(\mathbf{C}^-)' D^{CCE} (M_F\mathbf{C} - M_{\mathbf{CCE}}) D^{CCE} \mathbf{C}^0 \| \\
\leq \| \Lambda^0 \| \| (\mathbf{C}^0)^- \| \| T^{-1} D^{CCE} (M_F\mathbf{C} - M_{\mathbf{CCE}}) D^{CCE} \| \| \Lambda^0 \| \\
= O_p(R_7). \tag{A86}
\]

Let us now consider \( k_{11} \). Write \( M_F\mathbf{C} = I_T - P_{\mathbf{C}}, \) such that \( D^{CCE} M_F D^{CCE} = D^{CCE} D^{CCE} - D^{CCE} P_{\mathbf{C}} D^{CCE} \). Here, given Lemma CCE3 and the fact that \( F \) and \( \mathbf{C} \) span the same vector space,

\[
\| T^{-1} \Lambda_i(\mathbf{C}^-)' D^{CCE} P_{\mathbf{C}} D^{CCE} \mathbf{C}^0 \| \\
= \| T^{-1} \Lambda_i(\mathbf{C}^-)' D^{CCE} P_{\mathbf{C}} (\mathbf{C}' \mathbf{F}' \mathbf{F}) - (\mathbf{C}' \mathbf{F}' \mathbf{F}) D^{CCE} \mathbf{C}^0 \| \\
= \| T^{-1} \Lambda_i(\mathbf{C}^-)' D^{CCE} (\mathbf{F}' \mathbf{F}^-1)' D^{CCE} \mathbf{C}^0 \| \\
\leq (NT)^{-1} \| \Lambda^0 \| \| (\mathbf{C}^0)^- \| \| \sqrt{NT^{-1/2} D^{CCE} \mathbf{F} \| \| (T^{-1} \mathbf{F}' \mathbf{F})^{-1} \| \| \Lambda^0 \| \\
= O_p(N^{-1} T^{-1}),
\]

which implies

\[
k_{11} = T^{-1} \Lambda_i(\mathbf{C}^-)' D^{CCE} M_F D^{CCE} \mathbf{C}^0 \lambda_i \\
= T^{-1} \Lambda_i(\mathbf{C}^-)' D^{CCE} D^{CCE} \mathbf{C}^0 \lambda_i + O_p(N^{-1} T^{-1}). \tag{A87}
\]

By combining (A86) and (A87),

\[
k_1 = k_{11} + k_{12} \\
= N^{-1} \Lambda^0 \| (\mathbf{C}^0)^- \| (NT^{-1} D^{CCE} D^{CCE})(\mathbf{C}^0) - \lambda^0 \| + O_p(R_7) + O_p(N^{-1} T^{-1}).
\]
Application of Lemma CCE4 now yields

\[ k_1 = \sqrt{T}N^{-1}b_{1CCE,i} + O_p(R_7) + O_p(N^{-1/2}) \],

where \( b_{1CCE,i} = \Lambda_i^0(\Theta_i^0) - \Xi_i(\Theta_i^0) - \lambda_i^0 \).

Next, consider \( k_2 \). By using \( M_{fC}F\lambda_i = M_{fC}\hat{\bar{F}}\bar{C}\bar{F}\lambda_i = 0 \), and then substitution for \( (M_{fC} - M_{fCCE}) \),

\[
k_2 = T^{-1}E_i^{M_{fCCE}}F\lambda_i
\]

\[
= -T^{-1}E_i(M_{fC} - M_{fCCE})F\lambda_i
\]

\[
= -T^{-1}E_i^{D_{CCE}S_{fCCE}^*F}F\lambda_i - T^{-1}E_i^{D_{CCE}^*F}S_{fCCE}^{-1}F\lambda_i
\]

\[
\quad - T^{-1}E_i^{SFCS_{fCCE}^*D_{CCE}^*F\lambda_i} - T^{-1}E_i^{FCS_{fCCE}^*D_{CCE}^*F\lambda_i} - (\hat{\bar{F}}\bar{C}^T\bar{F})^{-1}\hat{\bar{F}}\bar{C}^T\lambda_i
\]

\[ = -k_{21} - \ldots - k_{24}. \]

Consider \( k_{21}, k_{23} \) and \( k_{24} \). Here we may make use of Lemmas CCE3, CCE5 and CCE7 to show that

\[
||k_{21}|| = ||T^{-1}E_i^{D_{CCE}S_{fCCE}^*D_{CCE}^*F\lambda_i}||
\]

\[
\leq N^{-\alpha+3/2}T^{-1/2}||NT^{-1}E_i^{D_{CCE}^*F}||||(T^{-1}S_{fCCE}^{-1})^{-1}||\sqrt{N}T^{-1/2}D_{CCE}^*F||\lambda_i^0||
\]

\[
= N^{-\alpha+3/2}T^{-1/2}[O_p(1) + O_p(\sqrt{NT^{-1/2}})]O_p(N^\alpha)
\]

\[ = O_p(N^{-1/2}T^{-1}), \quad (A89) \]

\[
||k_{23}|| = ||T^{-1}E_i^{FCS_{fCCE}^*D_{CCE}^*F\lambda_i}||
\]

\[
\leq N^{-2\alpha+1/2}T^{-1/2}||T^{-1/2}E_i^F||\lambda_i^0||T^{-1}S_{fCCE}^{-1}||\sqrt{N}T^{-1/2}D_{CCE}^*F||\lambda_i^0||
\]

\[
= N^{-2\alpha+1/2}T^{-1/2}O_p(N^\alpha) = O_p(N^{-1/2}T^{-1}), \quad (A90) \]

and

\[
||k_{24}|| = ||T^{-1}E_i^{F[C_{fCCE}^* - (\hat{\bar{F}}\bar{C}^T\bar{F})^{-1}\bar{C}^T\bar{F}\lambda_i]}||
\]

\[
\leq N^{-3\alpha}T^{-1/2}||T^{-1/2}E_i^F||\lambda_i^0||T^{-1}S_{fCCE}^{-1} - (\hat{\bar{F}}\bar{C}^T\bar{F})^{-1}||T^{-1}F||\lambda_i^0||
\]

\[ = N^{-3\alpha}T^{-1/2}[O_p(N^{2\alpha-1}) + O_p(N^{\alpha-1/2}T^{-1/2})]
\]

\[ = O_p(N^{-\alpha+1/2}T^{-1/2}) + O_p(N^{-2\alpha+1/2}T^{-1}). \quad (A91) \]

\( k_{22} \) can be expanded in the following obvious fashion:

\[
k_{22} = T^{-1}E_i^{D_{CCE}S_{fCCE}^*F\lambda_i}
\]

\[
= T^{-1}E_i^{D_{CCE}^*F\lambda_i} + T^{-1}E_i^{D_{CCE}S_{fCCE}^*F\lambda_i - (\hat{\bar{F}}\bar{C}^T\bar{F})^{-1}\bar{C}^T\bar{F}\lambda_i},
\]

57
where, via Lemmas CCE5 and CCE7,
\[
\| T^{-1} E_i^T D^{CCE} [S_{\text{FCE}}^{-1} - (\bar{C}^T F \bar{C})^{-1}] \bar{C}^T F \lambda_i \| \leq N^{-(2a+1)} \| N T^{-1} E_i^T D^{CCE} \| \| T \| \| S_{\text{FCE}}^{-1} - (\bar{C}^T F \bar{C})^{-1} \| \| \bar{C}^0 \| \| T^{-1} F \| \| \lambda_i^0 \|
\]
\[
= N^{-(2a+1)} |O_p(1) + O_p(\sqrt{N T^{-1/2}})| |O_p(N^{2a-1}) + O_p(N^{a-1/2} T^{-1/2})|
\]
\[
= O_p(N^{-2}) + O_p(N^{-3/2} T^{-1/2}) + O_p(N^{-(a+3/2) T^{-1/2}}) + O_p(N^{-(a+1) T^{-1}}).
\]
Moreover,
\[
T^{-1} E_i^T D^{CCE} \bar{C}^{-} \lambda_i = N^{-1} (N T^{-1}) E_i^T D^{CCE} \bar{C}^{-} \lambda_i = N^{-1} \Sigma_{e,i} (\beta, I_m) (\bar{C}^0)\lambda_i^0 + O_p(N^{-1} T^{-1/2}).
\]
Hence,
\[
k_{22} = N^{-1} b_{2CCE,i} + O_p(N^{-5/2}) + O_p(N^{-1} T^{-1/2}) + O_p(N^{-a-3/2} T^{-1}),
\]
(A92)
where \( b_{2CCE,i} = \Sigma_{e,i} (\beta, I_m) (\bar{C}^0)\lambda_i^0 \). Taking together (A89)–(A92), it is clear that
\[
k_2 = -k_{21} - ... - k_{24}
\]
\[
= -N^{-1} b_{2CCE,i} + O_p(N^{a-3/2} T^{-1/2}) + O_p(N^{-1} T^{-1/2}) + O_p(N^{-5/2})
\]
\[
+ O_p(N^{a-3/2} T^{-1}) + O_p(N^{-1/2} T^{-1}).
\]
(A93)
By adding (A88) and (A93),
\[
-T^{-1} X_i^T M_{\text{FCE}} D^{CCE} \bar{C}^{-} \lambda_i = k_1 + k_2 = N^{-1} (b_{1CCE,i} - b_{2CCE,i}) + O_p(R_8),
\]
(A94)
where, under \( T = N^x \),
\[
R_8 = N^{2a-2} + N^{a-3/2} T^{-1/2} + N^{-1} T^{-1/2} + N^{-1/2} T^{-1}
\]
\[
= N^{2a-2} + N^{a-(3+\kappa)/2} + N^{-(1+\kappa)/2} + N^{-(1/2+\kappa)}.
\]
The second result follows directly from the first;
\[
\frac{1}{\sqrt{N T}} \sum_{i=1}^{N} X_i^T M_{\text{FCE}} D^{CCE} \bar{C}^{-} \lambda_i = \sqrt{N T^{-1}} \frac{1}{N} \sum_{i=1}^{N} T^{-1} X_i^T M_{\text{FCE}} D^{CCE} \bar{C}^{-} \lambda_i
\]
\[
= \sqrt{T N^{-1/2}} (\bar{b}_{1CCE} - \bar{b}_{2CCE}) + \sqrt{N T O_p(R_8)},
\]
where \( \bar{b}_{1CCE} \) and \( \bar{b}_{2CCE} \) are simply the average \( b_{1CCE,i} \) and \( b_{2CCE,i} \), respectively.

Lemma PC10. Under Assumptions ERR, LAM, RK–PC and KAP,
\[
- \frac{1}{\sqrt{N T}} \sum_{i=1}^{N} X_i^T M_{\text{FPC}} D^{PC} H^{-} \lambda_i = \sqrt{T N^{-1/2}} (\bar{b}_{1PC} - \bar{b}_{2PC}) + O_p(R_{12}),
\]
(A95)
where

$$\mathbf{b}_{1PC} = \frac{1}{N} \sum_{i=1}^{N} \Lambda_i^0 (\mathbf{Q}^0) - \mathbf{S}^0 (\mathbf{Q}^0) - \mathbf{\alpha}_i^0,$$

$$\mathbf{b}_{2PC} = \frac{1}{N} \sum_{i=1}^{N} \Sigma_{E,i}(\beta, I_m) \mathbf{C}_i^0 (\mathbf{Q}^0) - \mathbf{\alpha}_i^0,$$

$$R_{12} = (N^{2a+1/2} R_{11} + N^{2a-1/2} R_4 + N^{a-1/2} R_5) \sqrt{T} + N^{-(3a+1)/2} R_1 + N^{a-1/2} R_5 R_{10}$$

$$+ (N^{-3a} R_1 + N^{-(a+1)} R_1 + N^{-a} R_{10} + N^{2a-1/2} R_1^2) T^{-1/2} + N^{-a} R_1 T^{-3/2},$$

$$R_{11} = N^{-4} + (N^{-5a+3/2} + N^{-(2a+2)} T^{-1/2} + (N^{-3} + N^{-(2a+3/2)} + N^{-4a+1}) T^{-1}$$

$$+ (N^{-(5a+1/2)} + N^{-(2a+1)}) T^{-3/2} + (N^{-2} + N^{-4a} + N^{-(2a+1)}) T^{-2}$$

$$+ N^{-2a} T^{-5/2} + N^{-1} T^{-3} + T^{-4}.$$ 

**Proof of Lemma PC10**

Analogous to Proof of Lemma CCE10, we can write

$$- \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} X_i^\prime M_{FC} D_{FC}^T \mathbf{H} - \mathbf{\lambda}_i = \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} \Lambda_i F^\prime M_{FC} F \mathbf{\lambda}_i + \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} E_i^\prime M_{FC} F \mathbf{\lambda}_i$$

$$= k_1 + k_2.$$  \hspace{1cm} \text{(A96)}$$

where

$$k_1 = \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} \Lambda_i (\mathbf{H}^\prime)^\prime D_{FC}^T M_{FC} D_{FC}^T \mathbf{H} - \mathbf{\lambda}_i$$

$$+ \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} \Lambda_i (\mathbf{H}^\prime)^\prime D_{FC}^T (M_{FC} - M_{FC}) D_{FC}^T \mathbf{H} - \mathbf{\lambda}_i$$

$$= k_{11} + k_{12}. $$ \hspace{1cm} \text{(A97)}$$

and

$$k_2 = - \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} E_i^\prime D_{FC}^T S_{FC}^\prime D_{FC}^T F \mathbf{\lambda}_i - \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} E_i^\prime D_{FC}^T S_{FC}^\prime \mathbf{H}^\prime F^\prime F \mathbf{\lambda}_i$$

$$- \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} E_i^\prime \mathbf{H}^\prime S_{FC}^\prime D_{FC}^T F \mathbf{\lambda}_i$$

$$- \frac{1}{\sqrt{N}T} \sum_{i=1}^{N} E_i^\prime \mathbf{H}^\prime (S_{FC} - (\mathbf{H}^\prime F^\prime F^\prime) F^\prime F \mathbf{\lambda}_i$$

$$= -k_{21} - \ldots - k_{24}.$$
Consider \( k_{12} \). Application of Lemmas PC2 and PC3 yields

\[
\| \sqrt{N} T^{-1/2} D^{PC} f^{PC} \| \\
\leq \sqrt{N} T \| T^{-1} D^{PC} D^{PC} \| + \| \sqrt{N} T^{-1/2} D^{PC} F \| \| H \| \\
= \sqrt{N} T \left[ O_p(T^{-1}) + O_p(N^{-1}) \right] + O_p(R_1) O_p(N^{-2a}) = O_p(R_{10}),
\]

(A98)

where

\[ R_{10} = \sqrt{N} T^{-1/2} + N^{-1/2} \sqrt{T} + N^{-2a} R_1 = N^{-1/2} \sqrt{T} + N^{-5a} + \sqrt{N} T^{-1/2}. \]

and therefore

\[
\| T^{-1} D^{PC} (M_{PC} - M_{PC}| F) D^{PC} \| \\
\leq N^{-1} \| T^{-1} D^{PC} D^{PC} \|^2 \| (T^{-1} S_{PC})^{-1} \| \\
+ 2N^{-1/2} T^{-1/2} \| H \| \| T^{-1} D^{PC} D^{PC} \| \| (T^{-1} S_{PC})^{-1} \| \| \sqrt{N} T^{-1/2} F T D^{PC} \| \\
+ N^{-1} T^{-1} \| H \|^2 \| \sqrt{N} T^{-1/2} D^{PC} F \|^2 T \| S_{PC}^{-1} \| (H' F H)^{-1} \| \\
= N^{-2} \left[ O_p(N^{-1}) + O_p(T^{-1}) \right]^2 \\
+ N^{-2a+1/2} T^{-1/2} \left[ O_p(N^{-1}) + O_p(T^{-1}) \right] O_p(R_1) \\
+ N^{-2a+1/2} T^{-1} O_p(R_1^2) O_p(N^{4a-1/2} T^{-1/2} R_{10}).
\]

(A99)

The order of the first term on the right is \( N^{-2} (N^{-1} + T^{-1})^2 \leq N^{-4} + (NT)^{-2} \), while the order of the second is

\[
N^{-2} (N^{-1} + T^{-1}) R_1 \\
= (N^{-2a+1/2} + N^{-2a+2}) T^{-1/2} + (N^{-2a+3/2} + N^{-2a+4a+1}) T^{-1} \\
+ (N^{-2a+1} + N^{-5a+1/2}) T^{-3/2} + (N^{-4a} + N^{-2a+1/2}) T^{-2} + N^{-2a} T^{-5/2}.
\]

The order of the third term can be written as follows, via the definition of \( R_{10} \):

\[
N^{-3/2} T^{-3/2} R_1^2 R_{10} = N^{-3/2} T^{-3/2} R_1^2 \left[ \sqrt{N} T^{-1/2} + \sqrt{T} N^{-1/2} \right] + N^{-2a+3/2} T^{-3/2} R_1^3,
\]

where, after considerable simplification,

\[
N^{-3/2} T^{-3/2} R_1^2 \left[ \sqrt{N} T^{-1/2} + \sqrt{T} N^{-1/2} \right] \\
\leq (N^{-3} + N^{-6a+2}) T^{-1} + (N^{-2} + N^{-4a+1}) T^{-2} + (N^{-4a} + N^{-1}) T^{-3} + T^{-4},
\]

60
and
\[ N^{-2a+3/2} T^{-3/2} R_3 \leq (N^{-2a+3} + N^{-11a+3/2}) T^{-3/2} + (N^{-2a+3/2} + N^{-8a}) T^{-3} + N^{-2a} T^{-9/2}, \]

implying
\[ N^{-3/2} T^{-3/2} R_3^2 R_{10} \]
\[ \leq (N^{-3} + N^{-6a+2}) T^{-1} + N^{-11a+3/2} T^{-3/2} + (N^{-4a+1} + N^{-2}) T^{-2} \]
\[ + (N^{-4a} + N^{-1}) T^{-3} + T^{-4}. \]

Thus, letting
\[ R_{11} = N^{-4} (N^{-5a+3/2} + N^{-2a+2}) T^{-1/2} + (N^{-3} + N^{-2a+3/2} + N^{-4a+1}) T^{-1} \]
\[ + (N^{-5a+1/2} + N^{-2a+1}) T^{-3/2} + (N^{-2} + N^{-4a} + N^{-2a+1/2}) T^{-2} \]
\[ + N^{-2a} T^{-5/2} + N^{-1} T^{-3} + T^{-4}, \]

we have
\[ ||T^{-1} D^{\text{PC}} (M_{\text{FPT}} - M_{\text{FRC}}) D^{\text{PC}}|| = O_p(R_{11}). \quad \text{(A100)} \]

It follows that, with \( (\bar{H}^0_0)^- = (\bar{H}^0)^{-1}(\bar{H}^0_0)^{-1} \),
\[ ||k_{12}|| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} A_i (\bar{H}^I) \cdot D^{\text{PC}} (M_{\text{FPT}} - M_{\text{FRC}}) D^{\text{PC}} H^T \lambda_i \right\| \]
\[ \leq N^{2a+1/2} \sqrt{T} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} ||A_i^0||^2 ||(\bar{H}^0)^-||^2 ||T^{-1} D^{\text{PC}} (M_{\text{FPT}} - M_{\text{FRC}}) D^{\text{PC}}||^2 ||\lambda_i^0|| \]
\[ = O_p(N^{2a+1/2} \sqrt{T} R_{11}). \quad \text{(A101)} \]

\( k_{11} \) can be rewritten as follows:
\[ k_{11} = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \Lambda_i (\bar{H}^I) \cdot D^{\text{PC}} P_{\text{FHT}} D^{\text{PC}} H^T \lambda_i - \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \Lambda_i (\bar{H}^I) \cdot D^{\text{PC}} P_{\text{FHT}} D^{\text{PC}} H^T \lambda_i \quad \text{(A102)} \]

where, by Lemma PC3,
\[ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \Lambda_i (\bar{H}^I) \cdot D^{\text{PC}} P_{\text{FHT}} (\bar{H}^0 F \bar{H}^I)^{-1} F^T D^{\text{PC}} H^T \lambda_i \right\| \]
\[ \leq \frac{1}{N^{1/2-2a} \sqrt{T} N} \sum_{i=1}^{N} ||A_i^0||^2 ||(\bar{H}^I)^2||^2 \sqrt{NT}^{-1/2} ||D^{\text{PC}} F||^2 ||(T^{-1} \bar{H}^0 F \bar{H}^I)^-||^2 ||\lambda_i^0|| \]
\[ = O_p(N^{2a-1/2} T^{-1/2} R_1^2). \]

61
From (A97), (A101) and (A102), we therefore obtain

\[ k_1 = \sqrt{T}N^{2a-1/2} \frac{1}{N} \sum_{i=1}^{N} \Lambda_i^0 (\mathbf{H}^0)^{-t} N T^{-1} \mathbf{D}^{PC} \mathbf{D}^{PC} (\mathbf{H}^0)^{-1} + O_p(N^{2a+1/2} \sqrt{T} R_{11}) \]

and so, by further use of Lemma PC4,

\[ k_1 = \sqrt{T}N^{-1/2} \bar{b}_{1PC} + O_p(\sqrt{T}N^{2a-1/2} R_4) + \sqrt{T}N^{2a+1/2} R_1, \]

where \( \bar{b}_{1PC} = N^{-1} \sum_{i=1}^{N} \Lambda_i^0 (\mathbf{Q}^0)^{-1} \mathbf{S} (\mathbf{Q}^0)^{-1} \Lambda_i^0. \)

Next, consider \( k_2. \) Making use of Lemmas PC3, PC5 and PC7 it is possible to show that

\[ ||k_{21}|| = \left| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i \mathbf{D}^{PC} \mathbf{S}_{E_{PC}} \mathbf{D}^{PC} \mathbf{F} \Lambda_i \right| \]

\[ \leq N^{-(a+1)} \frac{1}{N} \sum_{i=1}^{N} \left| (NT^{-1} E_i \mathbf{D}^{PC}) (I - T^{-1} S_{E_{PC}})^{-1} \right| \left| \sqrt{NT^{-1/2} D^{PC} F} \right| \left| \Lambda_i^0 \right| \]

\[ = N^{-(a+1)} \left| O_p(N^{-a/2} \sqrt{T}) + O_p(\sqrt{NT^{-1/2}}) + O_p(NT^{-3/2}) \right| O_p(R_1) \]

\[ = O_p(N^{-3a/2+1} R_1) + O_p(N^{-a/2} T^{-1/2} R_1) + O_p(N^{-a} T^{-3/2} R_1), \]  

(A103)

\[ ||k_{22}|| = \left| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i \Lambda_i \mathbf{F} \mathbf{H} \mathbf{S}_{E_{PC}} \mathbf{D}^{PC} \mathbf{F} \Lambda_i \right| \]

\[ \leq N^{-3a} T^{-1/2} \frac{1}{N} \sum_{i=1}^{N} \left| T^{-1/2} E_i \mathbf{F} \right| \left| \mathbf{H}^0 \right| \left| (T^{-1} S_{E_{PC}})^{-1} \right| \left| \sqrt{NT^{-1/2} D^{PC} F} \right| \left| \Lambda_i^0 \right| \]

\[ = O_p(N^{-3a} T^{-1/2} R_1), \]  

(A104)

and

\[ ||k_{24}|| \]

\[ = \left| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i \Lambda_i \mathbf{F} \mathbf{H} \mathbf{S}_{E_{PC}} - \left( \mathbf{H} \mathbf{F} \mathbf{H} \right)^{-1} \mathbf{H} \mathbf{F} \Lambda_i \right| \]

\[ \leq N^{1/2-5a} \frac{1}{N} \sum_{i=1}^{N} \left| T^{-1/2} E_i \mathbf{F} \right| \left| \mathbf{H}^0 \right| 2T \left| (T^{-1} S_{E_{PC}}) - \left( \mathbf{H} \mathbf{F} \mathbf{H} \right)^{-1} \right| \left| T^{-1} \mathbf{F} \right| \left| \Lambda_i^0 \right| \]

\[ = N^{1/2-5a} O_p(N^{-a} T^{-1/2} R_1) = O_p(N^{-a} T^{-1/2} R_1). \]  

(A105)

For \( k_{22}, \)

\[ k_{22} = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i \mathbf{D}^{PC} \mathbf{S}_{E_{PC}} \mathbf{H} \mathbf{F} \Lambda_i \]

\[ = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i \mathbf{D}^{PC} \mathbf{H} \mathbf{F} \Lambda_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i \mathbf{D}^{PC} \mathbf{S}_{E_{PC}} - \left( \mathbf{H} \mathbf{F} \mathbf{H} \right)^{-1} \mathbf{H} \mathbf{F} \Lambda_i, \]

62
where, by Lemmas PC5 and PC7,

\[
\left| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i^D PC \left[S_{FC}^{-} - (\mathbf{H}' \mathbf{F} \mathbf{F}^{-})^{-1} \mathbf{H}' \mathbf{F} \mathbf{F} \lambda_i^0 \right] \right|
\]

\[
\leq \sqrt{T} N^{-\left(3 \alpha + 1/2\right)} \frac{1}{N} \sum_{i=1}^{N} \left| \right| |NT^{-1} E_i^D PC \left| |T|| |S_{FC}^{-} - (\mathbf{H}' \mathbf{F} \mathbf{F}^{-})^{-1} \mathbf{H}' \mathbf{F} \mathbf{F} \lambda_i^0 \right| ||
\]

\[
= \sqrt{T} N^{-\left(3 \alpha + 1/2\right)} \left[ O_p(N^{-\alpha/2}) + O_p(\sqrt{N} \sqrt{T}^{-1/2}) + O_p(N \sqrt{T}^{-3/2}) + O_p(N^{3\alpha-1/2} \sqrt{T}^{-1/2} R_{10}) \right]
\]

\[
= O_p(N^{\alpha-1/2} R_{10}) + O_p(N^{\alpha-1/2} T^{-1/2} R_{10}) + O_p(N^{\alpha T^{-1/2}} R_{10}).
\]

Another application of Lemma PC5 yields

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i^D PC \mathbf{H}^{-1} \lambda_i
\]

\[
= \sqrt{T} N^{-1/2} \frac{1}{N} \sum_{i=1}^{N} E_i^D PC \mathbf{H}^{-1} \lambda_i
\]

\[
= \sqrt{T} N^{-1/2} \frac{1}{N} \sum_{i=1}^{N} \Sigma_{e,i}(\beta, \mathbf{I}_m) \mathbf{C}^0_j(\mathbf{Q}^0) \lambda_i^0 + O_p(\sqrt{T} N^{\alpha - 1/2} R_5),
\]

and so, with \( \mathbf{b}_{2PC} = N^{-1} \sum_{i=1}^{N} \Sigma_{e,i}(\beta, \mathbf{I}_m) \mathbf{C}^0_j(\mathbf{Q}^0) \lambda_i^0 \),

\[
k_2 = \sqrt{T} N^{-1/2} \mathbf{b}_{2PC} + O_p(\sqrt{T} N^{\alpha - 1/2} R_5) + O_p(N^{\alpha - 1/2} R_5 R_{10}).
\]  

(A107)

By adding (A104)–(A107), we obtain

\[
k_2 = \sqrt{T} N^{-1/2} \mathbf{b}_{2PC} + O_p(\sqrt{T} N^{\alpha - 1/2} R_5) + O_p(N^{\alpha - 1/2} R_5 R_{10}) + O_p(N^{-3\alpha} T^{-1/2} R_1)
\]

\[
+ O_p(N^{-(3\alpha + 2/1)} R_1) + O_p(N^{-(\alpha + 1/2)} T^{-1/2} R_1) + O_p(N^{\alpha T^{-3/2}} R_1)
\]

\[
+ O_p(N^{\alpha T^{-1/2}} R_{10}).
\]  

(A108)

The above results for \( k_1 \) and \( k_2 \) imply

\[- \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i^D M_{FC}^P \mathbf{D}^P \mathbf{H}^{-1} \lambda_i = k_1 + k_2 = \sqrt{T} N^{-1/2} (\mathbf{b}_{1PC} - \mathbf{b}_{2PC}) + O_p(R_{12}),
\]

where

\[
R_{12} = (N^{2\alpha + 1/2} R_{11} + N^{\alpha - 1/2} R_4 + N^{\alpha - 1/2} R_5) \sqrt{T} + N^{-(3\alpha + 2/1)} R_1 + N^{\alpha - 1/2} R_3 R_{10}
\]

\[
+ (N^{-3\alpha} R_1 + N^{-(\alpha + 1/2)} R_1 + N^{\alpha} R_{10} + N^{2\alpha - 1/2} R_1^2) T^{-1/2} + N^{-\alpha} R_1 T^{-3/2}.
\]
Lemma CCE11. Under Assumptions ERR, LAM, RK–CCE and KAP,

\[
T^{-1}X_i' M_{\text{FCCE}} v_i = T^{-1}E_i' v_i - N^{-1}b_{3\text{CCE},i} + O_p(R_9) + O_p(T^{-1}),
\]

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' M_{\text{FCCE}} v_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i' v_i - \sqrt{T}N^{-1/2}b_{3\text{CCE}} + \sqrt{NT}O_p(R_9) + O_p(T^{-1/2}),
\]

where

\[
b_{3\text{CCE},i} = \sigma_{v,i}^2 \mathbf{A}_0^0 (\mathbf{C}^0)'^{1}(1,0)',
\]

\[
\mathbf{b}_{3\text{CCE}} = \frac{1}{N} \sum_{i=1}^{N} b_{3\text{CCE},i},
\]

\[
R_9 = N^{2a-2} + N^{a-1-\kappa/2} + N^{2a-3/2-\kappa/2} + N^{a-(1/2+\kappa)} + N^{2a-(1+\kappa)}.
\]

Proof of Lemma CCE11

Let us rewrite \(T^{-1}X_i' M_{\text{FCCE}} v_i\) as

\[
T^{-1}X_i' M_{\text{FCCE}} v_i = T^{-1}X_i' M_{\text{FC}} v_i - T^{-1}X_i' (M_{\text{FC}} - M_{\text{FCCE}}) v_i,
\]

where, via (A75),

\[
T^{-1}X_i' (M_{\text{FC}} - M_{\text{FCCE}}) v_i = T^{-1}X_i' D^{\text{CCE}} S_{\text{FCCE}} D^{\text{CCE}} v_i + T^{-1}X_i' D^{\text{CCE}} S_{\text{FCCE}} \mathbf{C}' F v_i
\]

\[
+ T^{-1}X_i' F S_{\text{FCCE}} D^{\text{CCE}} v_i + T^{-1}X_i' F (S_{\text{FCCE}} - (\mathbf{C}' F \mathbf{C})^-) \mathbf{C}' F v_i
\]

\[
= I_1 + ... + I_4.
\]

The order of \(I_1, ..., I_4\) can be obtained by using the same steps as when analyzing \(k_2\). Lemmas CCE6–CCE8 imply

\[
||I_1|| = ||T^{-1}X_i' D^{\text{CCE}} S_{\text{FCCE}} D^{\text{CCE}} v_i||
\]

\[
\leq N^{-2}||NT^{-1}X_i' D^{\text{CCE}}|| ||T^{-1}S_{\text{FCCE}}|| ||NT^{-1}D^{\text{CCE}} v_i||
\]

\[
= N^{-2}[O_p(1) + O_p(\sqrt{NT^{-1/2}})]O_p(N^{2a})
\]

\[
= O_p(N^{2a-2}) + O_p(N^{2a-3/2-1/2}) + O_p(N^{2a-1-1}),
\]

(A111)

\[
||I_2|| = ||T^{-1}X_i' D^{\text{CCE}} S_{\text{FCCE}} \mathbf{C}' F v_i||
\]

\[
\leq N^{-(a+1)} T^{-1/2} ||NT^{-1}X_i' D^{\text{CCE}}|| ||T^{-1}S_{\text{FCCE}}|| ||\mathbf{C}'|| ||T^{-1/2} F v_i||
\]

\[
= N^{-(a+1)} T^{-1/2}[O_p(1) + O_p(\sqrt{NT^{-1/2}})]O_p(N^{2a})
\]

\[
= O_p(N^{a-1-1/2}) + O_p(N^{a-1/2-1})
\]

(A112)
and, additionally using (B),
\[
||I_3|| = ||T^{-1}X'_i\bar{F}\bar{C}S_{\text{CCE}}^{-} - \widetilde{\bar{C}'F'\bar{C}}^{-}\bar{C}'F'v_i||
\leq N^{-a}T^{-1/2}||T^{-1}X'_i||\bar{C}^0||T||\bar{C}^{-}||S_{\text{CCE}}^{-} - \widetilde{\bar{C}'F'\bar{C}}^{-}||||T^{-1/2}F'v_i||
= N^{-a}T^{-1/2}[O_p(N^{-a}) + O_p(T^{-1/2})][O_p(N^{2a-1}) + O_p(N^{a-1/2}T^{-1/2})]
= O_p(N^{-(a+1)}T^{-1/2}) + O_p(N^{-1}T^{-1}) + O_p(N^{-(2a+1/2)}T^{-1})
+ O_p(N^{-a+1})T^{-3/2}).
\tag{A113}
\]

\(I_3\) can be decomposed as follows:
\[
I_3 = T^{-1}X'_i\bar{F}\bar{C}S_{\text{CCE}}^{-}D^{CCE}v_i
= T^{-1}X'_i\bar{F}\bar{C}(\bar{C}'\bar{F}'\bar{C}) - D^{CCE}v_i + T^{-1}X'_i\bar{F}\bar{C}S_{\text{CCE}}^{-} - \widetilde{\bar{C}'F'\bar{C}}^{-}D^{CCE}v_i.
\tag{A114}
\]

Via Lemma CCE7, we obtain
\[
||T^{-1}X'_i\bar{F}\bar{C}S_{\text{CCE}}^{-} - \widetilde{\bar{C}'F'\bar{C}}^{-}D^{CCE}v_i||
\leq N^{-(a+1)}||T^{-1}X'_i||\bar{C}^0||T||\bar{C}^{-}||S_{\text{CCE}}^{-} - \widetilde{\bar{C}'F'\bar{C}}^{-}||||NT^{-1}D^{CCE}v_i||
= N^{-(a+1)}[O_p(N^{-a}) + O_p(T^{-1/2})][O_p(N^{2a-1}) + O_p(N^{a-1/2}T^{-1/2})]
\times [O_p(1) + O_p(\sqrt{NT^{-1/2}})]
= O_p(N^{-2}) + O_p(N^{a-2}T^{-1/2}) + O_p(N^{-3/2}T^{-1/2}) + O_p(N^{-(a+1)}T^{-1})
+ O_p(N^{a-3/2}T^{-1}) + O_p(N^{-1}T^{-3/2}).
\tag{A115}
\]

Note also that
\[
||T^{-1}E'_i\bar{F}\bar{C}(\bar{C}'\bar{F}'\bar{C}) - D^{CCE}v_i||
= ||T^{-1}E'_i\bar{F}\bar{C}(\bar{C}'\bar{F}'\bar{C}) - D^{CCE}v_i||
\leq N^{a-1}T^{-1/2}||T^{-1/2}E'_i||\bar{C}^0||((\bar{C}'\bar{F}'\bar{C})^{-1})||NT^{-1}D^{CCE}v_i||
= N^{a-1}T^{-1/2}[O_p(1) + O_p(\sqrt{NT^{-1/2}})] = O_p(N^{a-1}T^{-1/2}) + O_p(N^{a-1/2}T^{-1}).
\]

which, together with Lemmas CCE6 and CCE7, implies that the first term in (A114) can be
written as

\[ T^{-1}X'F \tilde{C}(\tilde{C}'\tilde{F}\tilde{C})^{-1}D^{CCE}v_i \]

\[ = T^{-1}\Lambda_i F' \tilde{C}(\tilde{C}'\tilde{F}\tilde{C})^{-1}D^{CCE}v_i + T^{-1}E_i' \tilde{F}(\tilde{C}'\tilde{F}\tilde{C})^{-1}D^{CCE}v_i \]

\[ = T^{-1}\Lambda_i (\tilde{C})' F' \tilde{C}(\tilde{C}'\tilde{F}\tilde{C})^{-1}D^{CCE}v_i + O_p(N^{a-1}T^{-1/2}) + O_p(N^{a-1/2}T^{-1}) \]

\[ = \sqrt{T}N^{-1}\Lambda_i [(\tilde{C})']^{-1}NT^{-1}D^{CCE}v_i + O_p(N^{a-1}) + O_p(N^{a-1/2}T^{-1/2}) \]

\[ = \sqrt{T}N^{-1}\sigma_{v_i}^2 A_0^0([\tilde{C}']^{-1})'N, (A117) \]

where the third equality is due to \((\tilde{C}^{-1})\tilde{F}' \tilde{C}(\tilde{C}'\tilde{F}\tilde{C})^{-1} = (\tilde{C}^{-1})\tilde{F}' \tilde{F}\tilde{C}(\tilde{F}'\tilde{F})^{-1} = (\tilde{C}^{-1})\tilde{C}(\tilde{C}')^{-1} = (\tilde{C}^{-1})' = (\tilde{C}')^{-1}\), while the last equality holds because the effect of the remainder in Lemma CCE6 is \(N^{-1/2}O_p(T^{-1/2}) = O_p((NT)^{-1/2}) \leq O_p(N^{a-1/2}T^{-1/2})\). It follows from (A115) and (A116) that

\[ b_{3\text{CCE},i} = \sigma_{v_i}^2 A_0^0([\tilde{C}']^{-1})'N. \]

The results in (A111)–(A113) and (A117) imply

\[ T^{-1}X'_{i}(M_{FC} - M_{F\text{CCE}})v_i = I_1 + \ldots + I_4 = TN^{-1}b_{3\text{CCE}} + O_p(R_9), \]  

where

\[ R_9 = N^{2a-2} + N^{a-1}T^{-1/2} + N^{2a-3/2}T^{-1/2} + (N^{a-1/2}2\alpha - 1)T^{-1} \]

\[ = N^{2a-2} + N^{a-1-\kappa/2} + N^{2a-3/2-\kappa/2} + N^{a-1/2+\kappa} + N^{2a-1+\kappa}. \]  

Equation (A109) can be rewritten as

\[ T^{-1}X'_{i}M_{F\text{CCE}}v_i = T^{-1}X'_{i}M_{F\text{C}}v_i - T^{-1}X'_{i}(M_{F\text{C}} - M_{F\text{CCE}})v_i \]

\[ = T^{-1}X'_{i}M_{F\text{C}}v_i - N^{-1}b_{3\text{CCE},i} + O_p(R_9) \]

\[ = T^{-1}E_i'v_i - T^{-1}E_i'P_{F\text{C}}v_i - N^{-1}b_{3\text{CCE},i} + O_p(R_9), \]  

where the last step arises from \(M_{F\text{C}}X_i = M_{F\text{C}}E_i\). Consider \(T^{-1}X_i'P_{F\text{C}}v_i\). From \(E(v_i'v_i') = \sigma_{v_i}^2 I_T\)
and \( P_{FC} P_{FC} = P_{FC} \),
\[
E[(T^{-1} E'_i [P_{FC} v_i]) (T^{-1} E'_i [P_{FC} v_i])'] = T^{-2} E[E_i P_{FC} E(v_i | E_i, E_p, \overline{FC}) P_{FC} E_i] = T^{-2} \sigma^2_{v,i} E(E_i P_{FC} E_i)
\]
\[
= T^{-2} \sigma^2_{v,i} E(E_i P_{FC} (\overline{C} F' FC) - \overline{C} F f_i]
\]
\[
= T^{-2} \sigma^2_{v,i} \sum_{i=1}^{T} E(E_i P_{FC} (\overline{C} F' FC) - C' f_i] = (m+1) T^{-2} \sigma^2_{v,i} \Sigma_{e,i} = O_p(T^{-2}),
\]

where last equality holds because
\[
\sum_{t=1}^{T} E_i \overline{C} (\overline{C} F' FC) - \overline{C} f_i] = \sum_{t=1}^{T} tr \left( \sum_{i=1}^{T} (C_i)' f_i C_i (C_i)' f_i \right) = tr(1_{m+1}) = m+1.
\]

Thus, since the variance is \( O_p(T^{-2}) \), we have
\[
\|T^{-1} E'_i [P_{FC} v_i]\| = O_p(T^{-1}).
\] (A121)

Taking together the results in (A120) and (A121), we obtain
\[
T^{-1} X'_i M_{FCCE} v_i = T^{-1} E'_i v_i - N^{-1} b_{3CCCE,i} + O_p(R_9) + O_p(T^{-1}).
\] (A122)

As for the second result of the lemma, note that
\[
E \left[ \left( \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E'_i P_{FC} v_i \right)' \left( \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E'_i P_{FC} v_i \right) \right] = \frac{1}{NT} \sum_{i=1}^{N} \sigma^2_{v,i} \Sigma_{e,i} \sum_{i=1}^{T} E[E_i \overline{C} (\overline{C} F' FC) - C' f_i] = O_p(T^{-1})
\]

Hence, \((NT)^{-1/2} \sum_{i=1}^{N} E'_i P_{FC} v_i = O_p(T^{-1/2})\), and we can therefore conclude that
\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i M_{FCCE} v_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E'_i v_i - \sqrt{T} N^{-1/2} \overline{B}_{3CCCE} + \sqrt{T} N^{-1/2} O_p(R_9)
\]
\[
+ O_p(T^{-1/2}),
\]

where \( \overline{B}_{3CCCE} = N^{-1} \sum_{i=1}^{N} b_{3CCCE,i} \). \( \square \)

**Lemma PC11.** Under Assumptions ERR, LAM, RK–PC and KAP,
\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i M_{PC} v_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E'_i v_i - \sqrt{T} N^{-1/2} \overline{B}_{3PC} + O_p(R_{13}) + O_p(T^{-1/2}),
\]
where

\[ \mathcal{B}_{3PC} = \frac{1}{N} \sum_{i=1}^{N} \alpha_{vi}^2 \Lambda_i^0 (\overline{Q}_0^i)^{-1} C_i^0 (1, 0)', \]

\[ \mathcal{R}_{13} = (N^{-(a+3)/2} R_6 + N^{-(3a+3)/2}) \sqrt{T} + N^{a-1/2} + N^{-1} R_{10} + (N^{2a-1/2} + N^{a-1} R_{10}) R_6 \]

\[ + (N^{-2a} + N^{-3a} R_1 + (N^{-a} + N^{a-1}) R_{10} + N^{2a-1} R_6 R_{10}) T^{-1/2} \]

\[ + (N^a + N^{-1/2} R_6 + R_{10}) T^{-1} + N^{-2a+1/2} T^{-3/2}. \]

**Proof of Lemma PC11**

Analogous to Proof of Lemma CCE11, we write

\[ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i^\prime (M_{DPC} - M_{DP}) v_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i^\prime (M_{DPC} - M_{DP}) v_i - \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i^\prime (M_{DPC} - M_{DP}) v_i = \]

(A124)

where

\[ T^{-1} X_i^\prime (M_{DPC} - M_{DP}) v_i \]

\[ = T^{-1} X_i^\prime D_{PC} S_{PC}^{-1} D_{PC}^\prime v_i + T^{-1} X_i^\prime D_{PC} S_{PC}^{-1} \overline{H} F' v_i \]

\[ + T^{-1} X_i^\prime \overline{F} \overline{H} S_{PC}^{-1} D_{PC}^\prime v_i + T^{-1} X_i^\prime \overline{F} \overline{H} S_{PC}^{-1} (\overline{H} F' F \overline{F} - \overline{H} F' F) \overline{H} F' v_i \]

\[ = \bar{I}_1 + ... + \bar{I}_4. \]

(A125)

Using Lemmas PC6, PC7 and PC8, we can work out the following orders for \( \bar{I}_1, \bar{I}_2 \) and \( \bar{I}_4 \):

\[ || \bar{I}_1 || \]

\[ \leq \frac{\sqrt{T}}{N^{3/2}} \left[ \frac{1}{N} \sum_{i=1}^{N} || NT^{-1} X_i^\prime D_{PC} || || (T^{-1} S_{PC}^{-1})^{-1} || NT^{-1} D_{PC} v_i || \right] \]

\[ = \sqrt{T} N^{-3/2} [O_p(N^{1/2-a} T^{-1/2} R_1) + O_p(N^{-a/2}) + O_p(\sqrt{T} N^{-1/2}) + O_p(NT^{-3/2})] \]

\[ \times [O_p(N^{-a}) + O_p(R_6)] \]

\[ = O_p(N^{-2a+1} R_1) + O_p(N^{-a+1} R_1 R_6) + O_p(\sqrt{T} N^{-3a+3/2}) + O_p(\sqrt{T} N^{-a+3/2} R_6) \]

\[ + O_p(N^{-a+1}) + O_p(N^{-a+1} R_6) + O_p(N^{-a+1} T^{-1}) + O_p(N^{-a+1} R_6) \]

(A126)
\[
||l_2|| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i D^{PC} S_{\bar{F}\bar{PC}} \bar{H}' F' v_i \right\|
\]
\[
\leq N^{-(2\alpha+1/2)} \left\| \sum_{i=1}^{N} (T^{-1} X'_i D^{PC}) (T^{-1} S_{\bar{F}\bar{PC}})^{-1} \bar{H}^0 \right\| \left\| T^{-1/2} F' v_i \right\|
\]
\[
= N^{-(2\alpha+1/2)} [O_p(N^{1-2\alpha} T^{-1/2} R_1) + O_p(N^{-\alpha/2}) + O_p(\sqrt{NT^{-1/2}}) + O_p(NT^{-3/2})]
\]
\[
= O_p(N^{-3\alpha} T^{-1/2} R_1) + O_p(N^{-(5\alpha+1)/2}) + O_p(N^{-2\alpha} T^{-1/2})
\]
\[
+ O_p(N^{-2\alpha+1/2} T^{-3/2}),
\]
(A127)

and, recalling (B),

\[
||l_4|| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i F H S_{\bar{F}\bar{PC}} - (\bar{H}' F' F \bar{H})^- \bar{H}' F' v_i \right\|
\]
\[
\leq N^{1/2-4\alpha} \left\| \sum_{i=1}^{N} (T^{-1} X'_i F) \left\| \bar{H}^0 \right\|^2 T \left\| (S_{\bar{F}\bar{PC}} - (\bar{H}' F' F \bar{H})^-) \right\| \left\| T^{-1/2} F' v_i \right\|
\]
\[
= N^{1/2-4\alpha} [O_p(N^{-\alpha}) + O_p(T^{-1/2})] O_p(N^{4\alpha-1/2} T^{-1/2} R_1)
\]
\[
= O_p(N^{-\alpha} T^{-1/2} R_1) + O_p(T^{-1} R_1).
\]
(A128)

For \(l_3\),

\[
l_3 = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i F H S_{\bar{F}\bar{PC}} D^{PC} v_i
\]
\[
= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i F H (\bar{H}' F' F \bar{H})^- D^{PC} v_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i F H [S_{\bar{F}\bar{PC}} - (\bar{H}' F' F \bar{H})^-] D^{PC} v_i,
\]

where, by Lemmas PC6 and PC7,

\[
\left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i F H [S_{\bar{F}\bar{PC}} - (\bar{H}' F' F \bar{H})^-] D^{PC} v_i \right\|
\]
\[
\leq \sqrt{TN} N^{-(2\alpha+1/2)} \frac{1}{N} \left\| T^{-1} X'_i F \right\| \left\| \bar{H}^0 \right\| T \left\| (S_{\bar{F}\bar{PC}} - (\bar{H}' F' F \bar{H})^-) \right\| \left\| NT^{-1} D^{PC} v_i \right\|
\]
\[
= \sqrt{TN} N^{-(2\alpha+1/2)} [O_p(N^{-\alpha}) + O_p(T^{-1/2})] O_p(N^{4\alpha-1/2} T^{-1/2} R_1 + O_p(N^{-\alpha}) + O_p(R_6)]
\]
\[
= O_p(N^{-1} R_1) + O_p(N^{a-1} R_6 R_1) + O_p(N^{a-1} T^{-1/2} R_1) + O_p(N^{a-1} T^{-1/2} R_6 R_1).
\]

Also,

\[
\left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E'_i F H (\bar{H}' F' F \bar{H})^- D^{PC} v_i \right\|
\]
\[
= N^{2\alpha-1/2} \left\| \frac{1}{N} \sum_{i=1}^{N} T^{-1/2} E'_i F H [T^{-1} (\bar{H}^0)' F' F \bar{H}^0]^- NT^{-1} D^{PC} v_i \right\|
\]
\[
= N^{2\alpha-1/2} [O_p(N^{-\alpha}) + O_p(R_6)] = O_p(N^{a-1/2}) + O_p(N^{2\alpha-1/2} R_6),
\]

69
which, via \((\hat{H}^{-})'\hat{H}'\hat{F}\hat{F}\hat{H}(\hat{H}'\hat{F}\hat{F}\hat{H})^{-} = (\hat{H}^{-})'(\hat{H}')^{-} = (\hat{H}^{-})'\) and Lemma PC6, yields

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' \hat{F}\hat{H}(\hat{H}'\hat{F}\hat{F}\hat{H})^{-} D^{PC_i} v_i
\]

\[
= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \Lambda_i(\hat{H}^{-})'\hat{F}\hat{F}\hat{H}(\hat{H}'\hat{F}\hat{F}\hat{H})^{-} D^{PC_i} v_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i' \hat{F}\hat{H}(\hat{H}'\hat{F}\hat{F}\hat{H})^{-} D^{PC_i} v_i
\]

\[
= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \Lambda_i(\hat{H}^{-})'\hat{F}\hat{F}\hat{H}(\hat{H}'\hat{F}\hat{F}\hat{H})^{-} D^{PC_i} v_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i' \hat{F}\hat{H}(\hat{H}'\hat{F}\hat{F}\hat{H})^{-} D^{PC_i} v_i
\]

\[
= \sqrt{T N^{a-1/2}} \frac{1}{N} \sum_{i=1}^{N} \Lambda_i(\hat{H}^{-})'\hat{F}\hat{F}\hat{H}(\hat{H}'\hat{F}\hat{F}\hat{H})^{-} D^{PC_i} v_i + O_p(N^{a-1/2}) + O_p(N^{2a-1/2} R_6)
\]

\[
= \sqrt{T N^{a-1/2}} \bar{b}_{3PC} + O_p(N^{a-1/2}) + O_p(N^{2a-1/2} R_6 + O_p(N^{a-1} R_6 R_{10}) + O_p(N^{a-1} T^{-1/2} R_{10} + O_p(N^{2a-1} T^{-1/2} R_6 R_{10}). \tag{A129}
\]

Taking together equations (A126)–(A129), we obtain

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' (\hat{M}_{FH} - M_{FPC}) v_i = I_1 + \ldots + I_4 = \sqrt{T N^{a-1/2}} \bar{b}_{3PC} + O_p(R_{13}), \tag{A130}
\]

where

\[
R_{13} = \left( N^{-(a+3)/2} R_6 + N^{-(3a+3)/2} \right) \sqrt{T} + N^{a-1/2} + N^{-1} R_6 + (N^{2a-1/2} + N^{a-1} R_6) R_6
\]

\[
+ \left( N^{-2a} + N^{-3a} R_1 + (N^{-a} + N^{a-1}) R_6 + N^{2a-1} R_6 R_{10} \right) T^{-1/2}
\]

\[
+ \left( N^a + N^{-1/2} R_6 + R_{10} \right) T^{-1} + N^{-2a+1/2} T^{-3/2}.
\]

Therefore, since in the case of CCE, \((NT)^{-1/2} \sum_{i=1}^{N} X_i' M_{FH} v_i = (NT)^{-1/2} \sum_{i=1}^{N} E_i' v_i + O_p(T^{-1/2}),\)

we can show that

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' M_{FPC} v_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' M_{FH} v_i - \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' (M_{FH} - M_{FPC}) v_i
\]

\[
= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i' v_i - \sqrt{T N^{a-1/2}} \bar{b}_{3PC} + O_p(R_{13}) + O_p(T^{-1/2}). \tag{A131}
\]

\[\blacksquare\]
C Proofs of main results

Proof of Theorem 1

We begin by considering the CCE estimator. The equation for $y_i$ can now be written as

$$y_i = X_i \beta + \hat{F}_{CCE} \lambda_i - D_{CCE} \lambda_i + \nu_i,$$  \hspace{1cm} (A132)

where $D_{CCE} = \hat{F}_{CCE} - F_C$ is as before. The CCE estimator of $\beta$ is given by

$$\hat{\beta}_{CCE} = \left( \sum_{i=1}^{N} X_i' M_{CCE} X_i \right)^{-1} \sum_{i=1}^{N} X_i' M_{CCE} y_i,$$

implying that

$$\sqrt{NT}(\hat{\beta}_{CCE} - \beta) = \left( \frac{1}{NT} \sum_{i=1}^{N} X_i' M_{CCE} X_i \right)^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' M_{CCE} (\nu_i - D_{CCE} \lambda_i).$$  \hspace{1cm} (A133)

Now let $\tilde{\beta}_{CCE} = \tilde{b}_{1CCE} - \tilde{b}_{2CCE} - \tilde{b}_{3CCE}$. From Lemmas CCE10 and CCE11, we know that

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' M_{CCE} (\nu_i - D_{CCE} \lambda_i) = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i' \nu_i + N^{(\kappa - 1)/2} \tilde{b}_{CCE}$$

$$+ \sqrt{NTO_p(R_8)} + \sqrt{NTO_p(R_9)} + O_p(T^{-1/2}).$$  \hspace{1cm} (A134)

The asymptotic covariance matrix of $(NT)^{-1/2} \sum_{i=1}^{N} E_i' \nu_i$ is given by $\tilde{W} = N^{-1} \sum_{i=1}^{N} \sigma_{i,t} \Sigma_{e,i}$. Moreover, since the fourth-order moments of $\epsilon_{i,t}$ and $\nu_{i,t}$ are bounded by assumption, by a central limit law,

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i' \nu_i \rightarrow_d N \left( 0, \lim_{N \to \infty} \tilde{W} \right)$$  \hspace{1cm} (A135)

as $N, T \to \infty$. For $\sqrt{NTO_p(R_8)}$ and $\sqrt{NTO_p(R_9)}$ to be negligible, according to the definitions of $R_8$ and $R_9$, the following conditions have to be satisfied:

(i) $\kappa < 3 - 4\alpha$,  
(ii) $\alpha < 1$,  
(iii) $\kappa > 0$,  
(iv) $\alpha < 1/2$,  
(v) $\kappa > 2\alpha$,  
(vi) $\kappa > 4\alpha - 1$.

If (iv) is satisfied, only (i) and (v) are binding. These restrictions are tighter than those implied by Lemma CCE9 for the denominator of (A133). Hence, under the condition that $\alpha < 1/2$ and
\[ \kappa \in K_{\text{CCE}} = (2\alpha, 3 - 4\alpha), \] (A133) reduces to

\[ \sqrt{NT}(\hat{\beta}_{\text{CCE}}^p - \beta) = \sum_{e}^{-1} \left( \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i'v_i + N^{(k-1)/2}b_{\text{CCE}} \right) + o_p(1) \]

\[ \rightarrow_d N \left( 0, \lim_{N \to \infty} \sum_{e}^{-1} \mathbf{W} \sum_{e}^{-1} \right) + \lim_{N \to \infty} N^{(k-1)/2} \sum_{e}^{-1} b_{\text{CCE}}, \]

which requires that \( N, T \to \infty. \)

It remains to consider the PC estimator. Using the same steps as above, we obtain

\[ \sqrt{NT}(\hat{\beta}_{\text{PC}}^p - \beta) = \left( \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_{\text{PC}}X_i \right)^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_{\text{PC}} (v_i - D_{\text{PC}}^p \mathbf{H} \lambda_i). \] (A136)

Define \( b_{\text{PC}} = b_{1\text{PC}} - b_{2\text{PC}} - b_{3\text{PC}}. \) Using Lemmas PC10 and PC11, we obtain

\[ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_{\text{PC}} (v_i - D_{\text{PC}}^p \mathbf{H} \lambda_i) = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i'v_i + \sqrt{T} N^{-1/2} b_{\text{PC}} + O_p(R_{12}) + O_p(R_{13}) + O_p(T^{-1/2}). \] (A137)

Let us now consider the order of the remainder term in the above expression, which can be expanded in the following fashion:

\[ R_{12} + R_{13} + T^{-1/2} \]

\[ = (N^{-3\alpha+3)/2} + N^{2\alpha-1/2} R_4 + N^a-1/2 R_5 + N^{-(\alpha+3)/2} R_6 + N^{2\alpha+1/2} R_{11}) \sqrt{T} \]

\[ + \ N^a-1/2 + (N^{-3\alpha+2+1)} R_1 + N^{2\alpha-1/2} R_6 + N^{-1} R_{10} + N^a-1/2 R_5 R_{10} + N^a-1 R_6 R_{10} \]

\[ + \ (1 + N^{-2\alpha} + (N^{-(\alpha+1/2)} + N^{-3\alpha}) R_1 + N^{2\alpha-1/2} R_2 \]

\[ + \ N^{2\alpha-1} R_6 R_{10} + (N^{a-1} + N^{-\alpha}) R_{10}) T^{-1/2} \]

\[ + \ (N^a + N^{-1/2} R_6 + R_{10}) T^{-1} + (N^{-2\alpha+1/2} + N^{-\alpha} R_1) T^{-3/2}. \] (A138)

Consider the coefficient of \( \sqrt{T}, \) the terms of which are given by

\[ N^{2\alpha+1/2} R_{11} = N^{2\alpha-7/2} + (N^{-3/2} + N^{-(3\alpha+1)}) T^{-1/2} + (N^{-1} + N^{-(2\alpha+1)} + N^{2\alpha-5)/2}) T^{-1} \]

\[ + \ (N^{-3\alpha} + N^{-1/2}) T^{-3/2} + (1 + N^{-2\alpha+1/2} + N^{2\alpha-5/2}) T^{-2} + \sqrt{NT}^{-5/2} \]

\[ + \ N^{2\alpha-1/2} T^{-3} + N^{2\alpha+1/2} T^{-4}, \]

\[ N^{2\alpha-1/2} R_4 = N^{-(\alpha+3)/2} + N^{-(2\alpha+1)} + (N^{-1} + N^{-2\alpha-1/2} + N^{2\alpha-3/2}) T^{-1/2} \]

\[ + \ (N^{2\alpha-1} + N^{-1/2}) T^{-1} + (N^{2\alpha-1/2} + N^{-\alpha}) T^{-3/2} + (N^{1/2-2\alpha + 1}) T^{-2} \]

\[ + \ \sqrt{NT}^{-5/2}, \]

72
\[
N^{a-1/2}R_5 = N^{a/2-1} + N^{a-1/2}T^{-1/2} + N^aT^{-3/2},
\]

and
\[
N^{-(a+3)/2}R_6 = N^{-(a/2+2)} + (N^{-(a+3)/2} + N^{-(3a+2)/2})T^{-1/2} + N^{-(a/2+1)}T^{-1} + N^{-(a+1)/2}.
\]

Insertion and simplification yield
\[
N^{-(3a+3)/2} + N^{2a+1/2}R_{11} + N^{2a-1/2}R_4 + N^{a-1/2}R_5 + N^{-(a+3)/2}R_6
\]
\[
= N^{-1} + N^{2a-7/2} + (N^{a-1/2} + N^{2a-3/2})T^{-1/2} + N^{2a-1}T^{-1} +
\]
\[
+ (N^a + N^{2a-1/2})T^{-3/2} + N^{1/2-2a}T^{-2} + \sqrt{NT^{-5/2}} + N^{2a+1/2}T^{-4}.
\]

Consider next the part that is constant in \( T \). Here
\[
N^{-(3a/2+1)}R_1 = N^{-(9a/2+1)} + N^{-(3a/2+3)/2} + (N^{-(3a/2+1)} + N^{-(7a+1)/2})T^{-1/2}
\]
\[
+ N^{-(3a+1)/2}T^{-1},
\]
\[
N^{2a-1/2}R_6 = N^{2a-1} + (N^{2a-1/2} + N^a)T^{-1/2} + N^aT^{-1} + N^{2a+1/2}T^{-3/2},
\]
\[
N^{-1}R_{10} = \sqrt{T}N^{-3/2} + N^{-(5a+1)} + N^{-1/2}T^{-1/2},
\]
\[
N^{a-1/2}R_5R_{10} = N^{-3/2}\sqrt{T} + N^aT^{-1} + N^{-1/2}T^{-1/2} + N^aT^{-1} + N^{a+1/2}T^{-2},
\]
\[
N^{a-1}R_6R_{10} = \sqrt{T}N^{-2} + N^{-1} + N^{-(a+3/2)} + (N^{-a-1} + N^{-(5a+1/2)})T^{-1/2}
\]
\[
+ (1 + N^{a-1})T^{-1} + N^aT^{-3/2} + N^{a+1/2}T^{-2},
\]
giving
\[
N^{a-1/2} + N^{-(3a/2+1)}R_1 + N^{2a-1/2}R_6 + N^{-1}R_{10} + N^{a-1/2}R_5R_{10} + N^{a-1}R_6R_{10}
\]
\[
= (N^{-3/2} + N^{a-2})\sqrt{T} + N^aT^{-1/2} + N^{2a-1} + (N^{2a-1/2} + N^a)T^{-1/2}
\]
\[
+ N^{2a}T^{-1} + N^{2a+1/2}T^{-3/2}.
\]

The coefficient of \( T^{-1/2} \) is
\[
1 + N^{-2a} + (N^{-(a+1/2)} + N^{-3a})R_1 + N^{2a-1/2}R_1^2 + N^{2a-1}R_6R_{10} + (N^{a-1} + N^{-a})R_{10},
\]

where
\[
(N^{-(a+1/2)} + N^{-3a})R_1
\]
\[
= N^{6a} + N^{-(3a+1/2)} + N^{-(a+1)} + (N^{-3a} + N^{-(a+1/2)})T^{-1/2} + (N^{-(3a+1/2)} + N^{-a})T^{-1},
\]
\[
73
\]
\[ N^{2a-1/2} R_1^2 \leq N^{-(4a+1/2)} + N^{2a-3/2} + (N^{2a-1/2} + N^{1/2-2a})T^{-1} + N^{2a+1/2}T^{-2}, \]
\[ N^{2a-1} R_6 R_{10} = N^{2a-2} \sqrt{T} + N^{2a-3/2} + N^{a-1} + (N^{-(4a+1/2)} + N^{2a-1})T^{-1/2} \]
\[ + \ (N^{2a-1/2} + N^{a})T^{-1} + N^{2a}T^{-3/2} + N^{2a+1/2}T^{-2}, \]
and
\[ (N^{a-1} + N^{-a})R_{10} \]
\[ = (N^{-(a+1/2)} + N^{a-3/2}) \sqrt{T} + N^{-(4a+1)} + N^{-6a} + (N^{a-1/2} + N^{1/2-a})T^{-1/2}. \]

Insertion and simplification yield
\[ 1 + N^{-2a} + (N^{-(a+1/2)} + N^{-3a})R_1 + N^{2a-1/2} R_1^2 + N^{2a-1} R_6 R_{10} + (N^{a-1} + N^{-a}) R_{10} \]
\[ = (N^{2a-2} + N^{a-3/2} + N^{-(a+1/2)}) \sqrt{T} + N^{2a-3/2} + N^{a-1} + 1 + N^{-2a} \]
\[ + \ (N^{2a-1} + N^{a-1/2} + N^{1/2-a})T^{-1/2} + (N^{2a-1/2} + N^{a})T^{-1} + N^{2a}T^{-3/2} + N^{2a+1/2}T^{-2}. \]

The coefficient of \( T^{-1} \) is simple;
\[ N^{a} + N^{-1/2} R_6 + R_{10} = N^{-1/2} \sqrt{T} + N^{a} + \sqrt{NT^{-1/2}}, \]
and so is that of \( T^{-3/2}; \)
\[ N^{-2a+1/2} + N^{-a} R_1 = N^{-2a+1/2} + N^{-(a+1/2)} + N^{-a}T^{-1/2} + N^{1/2-a}T^{-1}. \]

Putting everything together, the remainder term in (A137) becomes
\[ R_{12} + R_{13} + T^{-1/2} = (N^{2a-7/2} + N^{a-2} + N^{a/2-1}) \sqrt{T} + N^{2a-1} + N^{a-1/2} \]
\[ + \ (N^{2a-1/2} + N^{a} + 1)T^{-1/2} + (N^{2a} + N^{1/2-a})T^{-1} \]
\[ + \ N^{2a+1/2}T^{-3/2} \]
\[ = (N^{2a-7/2} + N^{a-2} + N^{a/2-1})N^{\kappa/2} + N^{2a-1} + N^{a-1/2} \]
\[ + \ (N^{2a-1/2} + N^{a} + 1)N^{-\kappa/2} + (N^{2a} + N^{1/2-a})N^{-\kappa} \]
\[ + \ N^{2a+1/2-3\kappa/2}. \quad (A139) \]

For this term to converge to go to zero, the following restrictions have to be satisfied:

(i) \( \kappa < 7 - 4a, \) (iii) \( \kappa < 2 - a, \) (v) \( \kappa > 4a - 1, \)

(ii) \( \kappa < 4 - 2a, \) (iv) \( \alpha < 1/2, \) (vi) \( \kappa > 2a, \)

74
(vii) \( \kappa > 0 \), \quad (viii) \( \kappa > 1/2 - \alpha \), \quad (ix) \( \kappa > (4\alpha + 1)/3 \).

Obviously, (iv) provides an upper bound on \( \alpha \). Given the admissible values of \( \alpha \), only (iii), (viii) and (ix) are binding. Consequently, we have that if \( \alpha < 1/2 \) and \( \kappa \in K_{PC} = (\max\{1/2 - \alpha, (4\alpha + 1)/3\}, 2 - \alpha) \), then

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i^T M_{PC}(v_i - D_{PC}^T \hat{H} \lambda_i) = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i' v_i + \sqrt{T} N^{-1/2} \hat{b}_{PC} + o_p(1)
\]

\[
\rightarrow_d N \left( \mathbf{0}, \lim_{N \to \infty} \mathbf{W} \right)
\]  

(A140)

as \( N, T \to \infty \). The limit of the denominator of the estimator is given by Lemma PC9, a result that holds irrespectively of the value of \( \alpha \) and \( \kappa \). Hence,

\[
\sqrt{NT} (\hat{\beta}_{PC}^p - \beta) \rightarrow_d N \left( \mathbf{0}, \lim_{N \to \infty} \mathbf{W} \right) \left[ \sum_{e}^{-1} \mathbf{W} \sum_{e}^{-1} \right] + \lim_{N \to \infty} N^{(k-1)/2} \sum_{e}^{-1} \mathbf{b}_{PC},
\]

(A141)

which again requires \( \alpha < 1/2 \) and \( \kappa \in K_{PC} \). This establishes the required result for the PC estimator and hence the proof of Theorem 1 is complete. \( \blacksquare \)

**Proof of Corollary 1**

Consider the CCE estimator. By using (A133), and Lemmas CCE10 and CCE11, it is not difficult to show that

\[
\hat{\beta}_{CCE}^p - \beta = O_p \left( (NT)^{-1/2} \right) + O_p(R_8) + O_p(R_9).
\]

(A142)

For the right-hand side of this expression to be \( o_p(1) \), \( R_8 \) and \( R_9 \) has to be negligible, which in turn impose the following restrictions:

(i) \( \alpha < 1 \), \quad (iv) \( \kappa > 2\alpha - 2 \), \quad (vii) \( \kappa > 2\alpha - 1 \).

(ii) \( \kappa > 2\alpha - 3 \), \quad (v) \( \kappa > 4\alpha - 3 \),

(iii) \( \kappa > -1/2 \), \quad (vi) \( \kappa > \alpha - 1/2 \),

Given that (i) is satisfied, only (v), (vi) and (vii) are binding. Hence, for (A142) to converge to zero, we need \( \alpha < 1 \) and \( \kappa > \max\{4\alpha - 3, \alpha - 1/2, 2\alpha - 1\} \), which for the relevant range of values for \( \kappa > 0 \) and \( \alpha \in [0, 1] \) reduces to \( \kappa > 2\alpha - 1 \).

As for the PC estimator, if we assume that \( \kappa > \max\{2\alpha, 4\alpha - 1\} \), such that Lemma PC1 holds, then by (A137),

\[
\hat{\beta}_{PC}^p - \beta = N^{-(1+\kappa)/2} O_p(R_{12} + R_{13}),
\]

(A143)
where, using the definition of $R_{12}$ and $R_{13}$, the remainder term is
\[
N^{-(1+\kappa)/2}(R_{12} + R_{13})
\]
\[
= (N^{2a-7/2} + N^{\alpha-2} + N^{-1})N^{-1/2} + N^{2a-(3+\kappa)/2} + N^{\alpha-(1+\kappa)/2}
\]
\[
+ (N^{2a-1/2} + N^{\alpha} + 1)N^{-(1/2+\kappa)} + (N^{2a} + N^{1/2-\alpha})N^{-(1+3\kappa)/2} + N^{2(\alpha-\kappa)}
\]
\[
= N^{2a-(3+\kappa)/2} + N^{2a-(1+\kappa)} + N^{\alpha-(1/2+\kappa)} + N^{2a-(1+3\kappa)/2} + N^{2(\alpha-\kappa)}.
\]
which converges to zero in probability if

(i) $\kappa > 4\alpha - 3,$

(ii) $\kappa > 2\alpha - 1,$

(iii) $\kappa > \alpha - 1/2,$

(iv) $\kappa > (2\alpha - 1/2)/3,$

(v) $\kappa > \alpha.$

However, none of these conditions are stricter than those implied by Lemma PC1. Hence, the values for $\kappa$ that imply $\hat{\beta}_{pc}^p - \beta = o_p(1)$ are bounded by $\kappa > \max\{2\alpha, 4\alpha - 1\}.$

Proof of Theorem 2

Consider the CCE estimator, which is given by
\[
\hat{\beta}_{CCE}^p = \left( \sum_{i=1}^N X_i^\prime M_{CCE} X_i \right)^{-1} \sum_{i=1}^N X_i^\prime M_{CCE} y_i.
\]
Under Assumption HET, this implies
\[
\sqrt{N}(\hat{\beta}_{CCE}^p - \beta) = \left( \frac{1}{NT} \sum_{i=1}^N X_i^\prime M_{CCE} X_i \right)^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^N X_i^\prime M_{CCE} (X_i \zeta_i + v_i - D_{CCE}^C C_{\lambda_i}).
\]
(A144)
The first term in the numerator is
\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^N X_i^\prime M_{FCE} X_i \zeta_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N X_i^\prime M_{FCE} X_i \zeta_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^N X_i^\prime (M_{CCE} - M_{FCE}) X_i \zeta_i.
\]
(A145)
Analogously to (A78) in Proof of Lemma CCE9, we obtain
\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^N X_i^\prime M_{FCE} X_i \zeta_i
\]
\[
= \frac{1}{\sqrt{NT}} \sum_{i=1}^N E_i^\prime M_{FCE} E_i \zeta_i
\]
\[
= \frac{1}{\sqrt{NT}} \sum_{i=1}^N E_i^\prime \xi_i - T^{-1/2} \frac{1}{\sqrt{N}} \sum_{i=1}^N (T^{-1/2} E_i^\prime F) \bar{C}^0 (T^{-1/2} C^0) E_i^\prime \xi_i
\]
\[
- (C^0) (T^{-1/2} F^\prime E_i) \xi_i
\]
\[
= \frac{1}{\sqrt{N}} \sum_{i=1}^N \Sigma_{\epsilon,i} \xi_i + O_p(T^{-1}),
\]
(A146)
It is easily seen that the first term on the right-hand side of the above expression is zero in expectation and also that the variance is given by

\[
E \left[ \left( \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Sigma_{e,i} \xi_i \right) \left( \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Sigma_{e,i} \xi_i \right)' \right] = \frac{1}{N} \sum_{i=1}^{N} \Sigma_{e,i} E( \xi_i | \xi_i ) \Sigma_{e,i} \\
= \frac{1}{N} \sum_{i=1}^{N} \Sigma_{e,i} \Sigma_{\xi} \Sigma_{e,i}.
\]

(A147)

The order of the second term on the right-hand side of (A145) can be obtained by using (A77);

\[
\left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i (M_{\text{FCE}} - M_{\text{FC}}) X_i \xi_i \right\| \\
\leq \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \| X'_i (M_{\text{FCE}} - M_{\text{FC}}) X_i \| \| \xi_i \| \\
= O_p (N^{2\alpha-3/2}) + O_p (N^{-1/2}) + O_p (N^{2\alpha-(2+\kappa)/2}) + O_p (N^{\alpha-(1+\kappa)/2}) \\
+ O_p (N^{-\kappa/2}) + O_p (N^{\alpha-\kappa}) + O_p (N^{2\alpha-\kappa-1/2}).
\]

(A148)

Hence, from (A146) and (A148), we have that (A145) reduces to

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X'_i M_{\text{FCE}} X_i \xi_i = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Sigma_{e,i} \xi_i + O_p (R_{14})
\]

(A149)

where

\[
R_{14} = N^{2\alpha-3/2} + N^{-1/2} + N^{2\alpha-(2+\kappa)/2} + N^{\alpha-(1+\kappa)/2} \\
+ N^{-\kappa/2} + N^{\alpha-\kappa} + N^{2\alpha-\kappa-1/2} + N^{-\kappa}.
\]

The remaining terms in (A144) are obtained by following the same steps as in the proof of
Theorem 1. Hence, applying (A134) and the result above, we obtain the following:

\[
\sqrt{N}(\hat{\beta}_{CCE}^p - \beta) = \left( \frac{1}{NT} \sum_{i=1}^{N} X_i'M_{CCE}X_i \right)^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_{CCE}(v_i - D_{CCE}^C \xi_i - \lambda_i) \\
+ \left( \frac{1}{NT} \sum_{i=1}^{N} X_i'M_{CCE}X_i \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Sigma_{e,i} \xi_i \\
+ \left( \frac{1}{NT} \sum_{i=1}^{N} X_i'M_{CCE}X_i \right)^{-1} \left( T^{-1/2} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E_i^Tv_i + N^{-1/2} \Sigma_{CCE} \right) \\
+ O_p(R_{14}) + \sqrt{N}O_p(R_8) + \sqrt{N}O_p(R_9) + O_p(T^{-1/2}) \\
= \left( \frac{1}{NT} \sum_{i=1}^{N} X_i'M_{CCE}X_i \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Sigma_{e,i} \xi_i + O_p(R_{14}) \\
+ \sqrt{N}O_p(R_8) + \sqrt{N}O_p(R_9) + O_p(N^{-1/2}) + O_p(T^{-1/2}). \tag{A150}
\]

Consider \( \sqrt{N}R_8 \) and \( \sqrt{N}R_9 \). To ensure that these are negligible, we again put restrictions \( \alpha \) and \( \kappa \). In particular, the following restrictions have to be met:

(i) \( \alpha < 3/4 \),

(ii) \( \kappa > 2\alpha - 2 \),

(iii) \( \kappa > -1 \),

(iv) \( \kappa > 0 \),

(v) \( \kappa > 2\alpha - 1 \),

(vi) \( \kappa > 4\alpha - 2 \),

(vii) \( \kappa > \alpha \),

(viii) \( \kappa > 2\alpha - 1/2 \).

Clearly, given that (i) is satisfied, only (vii) and (viii) are binding. These conditions are, however, more stringent than those implied by Lemma CCE9. We can hence conclude that if \( \alpha < 3/4 \) and \( \kappa > \max\{\alpha, 2\alpha - 1/2\} \), then \( \sqrt{N}R_8 \) and \( \sqrt{N}R_9 \) are negligible. Additional conditions are needed for \( R_{14} \) to go to zero. These can, however, be shown either to be less stringent or to coincide with the above conditions. We can therefore conclude that if \( \alpha < 3/4 \) and \( \kappa > \max\{\alpha, 2\alpha - 1/2\} \), as \( N, T \to \infty \),

\[
\sqrt{N}(\hat{\beta}_{CCE}^p - \beta) = \Sigma_e^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Sigma_{e,i} \xi_i + o_p(1) \\
\to_d N\left(0, \Sigma_e^{-1} \left( \frac{1}{N} \sum_{i=1}^{N} \Sigma_{e,i} \Sigma_{e,i}^T \right) \Sigma_e^{-1} \right).
\]

The proof for the PC estimator is very similar to that provided above for the CCE estimator.
The main difference is that instead of (A77) we use (A81):
\[
\left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' (M_{\text{FFH}} - M_{\text{PC}}) X_i \zeta_i \right\| = \frac{1}{\sqrt{N}} \left( \sum_{i=1}^{N} X_i' M_{\text{PC}} X_i \right) \zeta_i = O_p(R_{15}),
\]
where
\[
R_{15} = N^{-2\alpha - 3/2} + N^{-\alpha} + N^{-(\alpha + 2\alpha + 1)} + N^{-1/2 - \kappa} + N^{-2\alpha + 1/2} + N^{-\alpha + 1/2 - 3/2} + N^{-2\kappa + 1/2},
\]
By using this, (A137), and the same steps as when considering the CCE estimator, we can show that
\[
\sqrt{N}(\hat{\beta}_{\text{PC}} - \beta) = \left( \frac{1}{NT} \sum_{i=1}^{N} X_i' M_{\text{PC}} X_i \right)^{-1} \left( \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' M_{\text{PC}} X_i \zeta_i \right) + \left( \frac{1}{NT} \sum_{i=1}^{N} X_i' M_{\text{PC}} X_i \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \zeta_i + O_p(R_{15}) + T^{-1/2} [O_p(R_{12}) + O_p(R_{13}) + O_p(T^{-1/2}) + O_p(N^{-1/2}) + O_p(T^{-1/2})].
\]
By using the results provided in the PC part of Proof of Theorem 1,
\[
T^{-1/2}(R_{12} + R_{13} + T^{-1/2}) = N^{2\alpha - 7/2} + N^{-\alpha} + N^{\alpha/2 - 1} + N^{2\alpha - 1/2} + N^{-1/2 - \kappa} + (N^{2\alpha - 1/2} + N^{\alpha} + 1)N^{-\kappa} + (N^{2\alpha} + N^{1/2 - \alpha}) N^{-3\kappa/2} + N^{2\alpha + 1/2 - 2\kappa}.
\]
The following conditions have to be met for this term to go to zero:

(i) $\alpha < 7/4$
(ii) $\kappa > 4\alpha - 2$
(iii) $\kappa > 2\alpha - 1/2$
(iv) $\kappa > 4\alpha / 3$
(v) $\kappa > (1 - 2\alpha) / 3$
(vi) $\kappa > \alpha + 1/4$.

While condition (i) binds $\alpha$, (ii), (vi) and (v) bind for $\kappa$, depending on the value of $\alpha$. The relevant restrictions to put on $\kappa$ and $\alpha$ are therefore given by $\alpha < 1$ and $\kappa > \max\{4\alpha - 2, \alpha + 1/4, (1 - 2\alpha) / 3\}$, respectively. Similarly, in order to ensure that $R_{15} = o(1)$, we need
Among these conditions, only (i) matters since it is stricter than \((1 - 2\alpha)/3\) for \(\alpha\) close to zero. Hence, provided that \(\alpha < 7/4\) and \(\kappa > \max\{4\alpha - 2, \alpha + 1/4, 1/2 - 2\alpha\}\),

\[
\sqrt{N}(\hat{\beta}_{PC}^p - \beta) \rightarrow_d N\left(0, \Sigma_c^{-1}\left(\frac{1}{N} \sum_{i=1}^{N} \Sigma_{c,i} \Sigma_{c,i} \Sigma_{c,i}^{-1}\right) \Sigma_c^{-1}\right)
\]
as \(N, T \rightarrow \infty\).

**Proof of Corollary 2**

It follows from Proof of Theorem 2 that

\[
\hat{\beta}_{CCE}^p - \beta = O_p(N^{-1/2}) + N^{-1/2}O_p(R_{14}) + O_p(R_8) + O_p(R_9) + O_p(N^{-1}) + O_p((NT)^{-1/2}).
\]

As noted in Proof of Theorem 2, the requirements for \(R_{14}\) to go to zero are never more stringent than those required for \(\sqrt{NR_8}\) and \(\sqrt{NR_9}\) to go to zero. Thus, it is enough to consider \(R_8\) and \(R_9\), for which the conditions are the same as in Proof of Corollary 1. Hence, provided that \(\alpha < 1\) and \(\kappa > 2\alpha - 1\),

\[
\hat{\beta}_{CCE}^p - \beta = o_p(1).
\]

The corresponding requirement for the PC estimator is that

\[
\hat{\beta}_{PC}^p - \beta = O_p(N^{-1/2}) + N^{-1/2}O_p(R_{15}) + (NT)^{-1/2}[O_p(R_{12}) + O_p(R_{13}) + O_p(T^{-1/2})] + O_p(N^{-1}) + O_p((NT)^{-1/2}).
\]

should be \(o_p(1)\). Now, by using the results provided in Proof of Theorem 2, we have that \(N^{-1/2}R_{15} = o(1)\) for all \(\alpha \in [0, 1]\) and \(\kappa > 0\). As shown in Proof of Corollary 1, \(\kappa > \alpha\) is required for the remaining terms to converge to zero. However, this condition is not as strict as the one required for Lemma PC1 to hold. Hence, as long as \(\kappa > \max\{2\alpha, 4\alpha - 1\}\), we have

\[
\hat{\beta}_{PC}^p - \beta = o_p(1).
\]
D Some results for the individual and mean group CCE estimators

As mentioned in the main text, Pesaran (2006) does not only consider the pooled CCE estimator, but also an individual estimator, henceforth denoted $\hat{\beta}_{CCE,i}$, and a mean group-type CCE estimator, here denoted $\hat{\beta}_{mg}^{CCE}$. In this section report some results for the latter two estimators under Assumption HET.

The individual CCE estimator is given simply by

$$\hat{\beta}_{CCE,i} = (X_i' M_{FCE}^{-1} X_i')^{-1} X_i' M_{FCE}^{-1} y_i,$$  \hspace{1cm} (A154)

the asymptotic distribution of which is given in Theorem D1.

**Theorem D1.** Suppose that Assumptions HET, ERR, LAM, RK–CCE and KAP hold, and that $\alpha < 3/4$ and $\kappa \in K_{CCE} = (\max\{0, 4\alpha - 2\}, 4 - 4\alpha)$. Then, as $N, T \to \infty$,

$$\sqrt{T}(\hat{\beta}_{CCE,i} - \beta_i) \rightarrow_d N(0, \sigma_{\nu,i}^2 \Sigma_{\epsilon,i}^{-1}) + \lim_{N \to \infty} N^{(x-2)/2} \Sigma_{\epsilon,i}^{-1} b_{CCE,i}. \hspace{1cm} \text{(A155)}$$

**Proof of Theorem D1**

Analogous to Proof of Theorem 1, we have

$$\sqrt{T}(\hat{\beta}_{CCE,i} - \beta_i) = (T^{-1} X_i' M_{FCE}^{-1} X_i')^{-1} T^{-1/2} X_i' M_{FCE}^{-1} (v_i - D_{CCE}^{-1} \lambda_i).$$

From Lemmas CCE10 and CCE11, we know that

$$T^{-1/2} X_i' M_{FCE}^{-1} (v_i - D_{CCE}^{-1} \lambda_i) = T^{-1/2} E_i' v_i + N^{-1} \sqrt{T} b_{CCE,i} + \sqrt{T} O_p(R_8) + \sqrt{T} O_p(R_9) + O_p(T^{-1}), \hspace{1cm} \text{(A155)}$$

where $b_{CCE,i} = b_{1CCE,i} - b_{2CCE,i} - b_{3CCE,i}$. The asymptotic variance of $T^{-1/2} E_i' v_i$ is $\sigma_{v,i}^2 \Sigma_{\epsilon,i}$, which in conjunction with the assumptions placed on $E_i$ and $v_i$ implies

$$T^{-1/2} E_i' v_i \rightarrow_d N(0, \sigma_{v,i}^2 \Sigma_{\epsilon,i}) \hspace{1cm} \text{(A156)}$$

as $N, T \to \infty$. As for the remainder terms in (A155), the restrictions for $\sqrt{T} R_8$ and $\sqrt{T} R_9$ to converge to zero are

(i) $\kappa < 4 - 4\alpha$,   (ii) $\alpha < 3/2$,   (iii) $\kappa > -1$,   (iv) $\alpha < 1$,   (v) $\alpha < 3/4$,   (vi) $\alpha < 4 \alpha - 2$,   (vii) $\kappa > 2 \alpha - 1$.
While the binding conditions for $\kappa$ are given by (i) and (vii), (v) binds $\alpha$. Note that condition (vi) is the same as in Lemma CCE9. Hence, provided that $\alpha < 3/4$ and $\kappa \in (\max\{0, 4\alpha - 2\}, 4 - 4\alpha)$, as $N, T \to \infty$,

$$
\sqrt{T}(\hat{\beta}_{CCE,i} - \beta_i) = \Sigma_{\nu,i}^{-1}T^{1/2}\epsilon'_{\nu,i} + \Sigma_{\nu,i}^{-1}\sqrt{T}\gamma_{CCE,i,i} + o_p(1)
$$

$$
\rightarrow_d N(0, \sigma^2_{\nu,i} \Sigma_{\nu,i}^{-1}N^{-1} \sqrt{T}\gamma_{CCE,i,i},
$$

(A157)

as required.

Pesaran (2006, Theorem 1) derives the asymptotic distribution of $\sqrt{T}(\hat{\beta}_{CCE,i} - \beta_i)$ in the strong factor case. He assumes that $\sqrt{T}/N = N^{(\kappa-2)/2} \to 0$ as $N, T \to \infty$, which is more restrictive than the condition required by Theorem D1. However, we see that if $\kappa < 2$ such that $N^{(\kappa-2)/2} \to 0$, then according to Theorem D1, $\sqrt{T}(\hat{\beta}_{CCE,i} - \beta_i)$ is asymptotically unbiased, which is in line with the results of Pesaran (2006). Hence, even in the strong factor case Theorem D1 represents an extension of the existing work in the field; it extends the work of Pesaran (2006) to the case when $\kappa \in (0, 4)$ such that $\sqrt{T}/N$ is not necessarily zero. In the non-strong factor case, there is no previous research to which we can refer. However, we see that the impact of $\alpha$ on the set of allowable values of $\kappa$, $K_{CCE}$, is very similar to that found in Theorem 1 for the pooled CCE estimator under Assumption HOM. In particular, the smaller is $\alpha$ (the weaker the factors), the narrower is $K_{CCE}$. In the extreme case when $\alpha \to 3/4$, $K_{CCE}$ collapses to a single value equal to one.

**Corollary D1.** Suppose that Assumptions HET, ERR, LAM, RK–CCE and KAP hold. Suppose also that $\kappa > 2\alpha - 1$ with $\alpha < 1$. Then, as $N, T \to \infty$,

$$
||\hat{\beta}_{CCE,i} - \beta_i|| = o_p(1).
$$

**Proof of Corollary D1**

It follows from (A155) and (A157) that

$$
\hat{\beta}_{CCE,i} - \beta_i = O_p(T^{-1/2}) + O_p(N^{-1}) + O_p(R_8) + O_p(R_9) + O_p(T^{-3/2}).
$$

It remains to establish conditions under which $R_8$ and $R_9$ are $o(1)$. These are given in Proof of Corollary 1, and hence we can conclude that if $\alpha < 1$ and $\kappa > 2\alpha - 1$, then

$$
\hat{\beta}_{CCE,i} - \beta_i = o_p(1).
$$
According to Corollaries 1 and D1, the required conditions on $\alpha$ and $\kappa$ to ensure consistency of the individual CCE estimator is the same as for the pooled CCE estimator.

The mean group CCE estimator is simply the average $\hat{\beta}_{CCE,i}$:

$$\hat{\beta}_{CCE}^mg = \frac{1}{N} \sum_{i=1}^{N} \hat{\beta}_{CCE,i}.$$ 

**Theorem D2.** Suppose that Assumptions HET, ERR, LAM, RK–CCE and KAP hold. Suppose also that $\alpha < 3/4$ and $\kappa > \max\{\alpha, 2\alpha - 1/2\}$. Then, as $N, T \to \infty$,

$$\sqrt{N}(\hat{\beta}_{CCE}^mg - \beta) \to_d N(0, \Sigma_\xi).$$

**Proof of Theorem D2**

Under Assumption HET, $\sqrt{N}(\hat{\beta}_{CCE}^mg - \beta)$ can be written as

$$\sqrt{N}(\hat{\beta}_{CCE}^mg - \beta) = N^{-1/2} \sum_{i=1}^{N} (\hat{\beta}_{CCE,i} - \beta)$$

$$= N^{-1/2} \sum_{i=1}^{N} \xi_i + \sqrt{N} \sum_{i=1}^{N} \left( T^{-1}X'_iM_{\text{CCE}}X_i \right)^{-1}T^{-1}X'_iM_{\text{CCE}}\nu_i$$

$$- \sqrt{N} \sum_{i=1}^{N} \left( T^{-1}X'_iM_{\text{CCE}}X_i \right)^{-1}T^{-1}X'_iM_{\text{CCE}}D_{\text{CCE}}C^{-1}x_i.$$ 

Since $T^{-1}X'_iM_{\text{CCE}}X_i = O_p(1)$, we can make use of Lemmas CCE10 and CCE11 to obtain

$$\sqrt{N}(\hat{\beta}_{CCE}^mg - \beta) = N^{-1/2} \sum_{i=1}^{N} \xi_i + T^{-1/2} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \left( T^{-1}X'_iM_{\text{CCE}}X_i \right)^{-1}E'_i\nu_i + \sqrt{NO_p}(R_9)$$

$$+ O_p(T^{-1}) + N^{-1} \sum_{i=1}^{N} N^{-1/2} \left( T^{-1}X'_iM_{\text{CCE}}X_i \right)^{-1}b_{\text{CCE},i}$$

$$+ \sqrt{NO_p}(R_9). \quad (A158)$$

Now let $\Delta = N^{-1/2} \sum_{i=1}^{N} \xi_i + T^{-1/2}(NT)^{-1} \sum_{i=1}^{N} \Sigma_{\xi,i}^{-1}E'_i\nu_i$. Given the assumptions on $E_i, \nu_i$ and $\xi_i$, the asymptotic variance of this term is given by

$$E(\Delta') = \frac{1}{N} \sum_{i=1}^{N} E(\xi_i\xi'_i) + T^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} E \left[ \left( \frac{1}{T}X'_iM_{\text{CCE}}X_i \right)^{-1}E'_i\nu_i\nu'_iE_i \left( \frac{1}{T}X'_iM_{\text{CCE}}X_i \right)^{-1} \right]$$

$$= \Sigma_\xi + O(T^{-1})$$

Also, in order to avoid that the remainder terms in (A158) dominate the asymptotic behavior of $\sqrt{N}(\hat{\beta}_{CCE}^mg - \beta)$, conditions have to be placed on $\alpha$ and $\kappa$. Specifically, to ensure that $\sqrt{NR_8}$ and $\sqrt{NR_9}$ are $o(1)$, we need
It is easy to see that given condition (i) for $\alpha$, only (vii) and (viii) are binding for $\kappa$. These conditions are, however, more stringent than those implied by Lemma CCE9. We can therefore conclude that if $\alpha < 3/4$ and $\kappa < \max\{\alpha, 2\alpha - 1/2\}$, then

$$
\sqrt{N}(\hat{\beta}_{CCE} - \beta) = N^{-1/2} \sum_{i=1}^{N} \xi_i + T^{-1/2} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \xi_i \epsilon_i + N^{-1/2} \frac{1}{N} \sum_{i=1}^{N} \xi_i b_{CCE,i} + o_p(1) \to_d N(0, \Sigma_\xi)
$$

as $N, T \to \infty$.

According to Theorems 2 and D2, under Assumption HET, the asymptotic distributions of $\hat{\beta}_{CCE}^p$ and $\hat{\beta}_{CCE}^{mg}$ differ only in terms of the asymptotic variance. Note in particular that by the Cauchy–Schwarz inequality, $N^{-1} \sum_{i=1}^{N} \xi_i \epsilon_i \epsilon_i - \Sigma_\xi \Sigma_\epsilon$ is positive semi-definite, implying that $\hat{\beta}_{CCE}^p$ is more efficient than $\hat{\beta}_{CCE}^{mg}$.

The conditions on $\alpha$ and $\kappa$ that ensure consistency under Assumption HET are the same for $\hat{\beta}_{CCE}^p$ and $\hat{\beta}_{CCE}^{mg}$. These conditions are in turn identical to those of $\hat{\beta}_{CCE,i}$ and $\hat{\beta}_{CCE}$ under Assumption HOM. Corollary D2 formalizes this.

**Corollary D2.** Suppose that Assumptions HET, ERR, LAM, RK–CCE and KAP hold. Suppose also that $\kappa > 2\alpha - 1$ with $\alpha < 1$. Then, as $N, T \to \infty$,

$$
||\hat{\beta}_{CCE}^{mg} - \beta|| = o_p(1).
$$

**Proof of Corollary D2**

According to Proof of Theorem D2,

$$
\hat{\beta}_{CCE}^{mg} - \beta = O_p(N^{-1/2}) + O_p((NT)^{-1/2}) + O_p(N^{-1}) + O_p(R_8) + O_p(R_9).
$$

Given that $R_8$ and $R_9$ enter this expression without being multiplied by either $N$ or $T$, the conditions for convergence to zero are the same as in Corollary D1.

According to Corollaries 2, D1 and D2, if consistency is the only concern, in terms of the allowable values of $\alpha$ and $\kappa$, the choice of which CCE estimator to use is just a matter of personal
preference. However, if we are also concerned about the rate at which consistency is achieved, if \( \kappa < 1 \) such that \( \sqrt{T} = N^{\kappa/2} < \sqrt{N} \), then \( \hat{\beta}_{CCE}^p \) and \( \hat{\beta}_{CCE}^m \) are preferred, whereas if \( \kappa \in (1, 2) \), then \( \hat{\beta}_{CCE,i} \) is the preferred choice (if \( \kappa \geq 2 \), then \( \sqrt{T}(\hat{\beta}_{CCE,i} - \beta_i) \) is no longer asymptotically unbiased).