Département de sciences économiques

Cahier 2002-18

Testing for a Unit Root in Panels with Dynamic Factors

MOON, H.R.
PERRON, B.

Université de Montréal

Département de sciences économiques
Université de Montréal
Faculté des arts et des sciences
C.P. 6128, succursale Centre-Ville
Montréal (Québec) H3C 3J7
Canada
http://www.sceco.umontreal.ca
SCECO-information@UMontreal.CA
Téléphone : (514) 343-6539
Télécopieur : (514) 343-7221

Ce cahier a également été publié par le Centre interuniversitaire de recherche en économie quantitative (CIREQ) sous le numéro 18-2002.

This working paper was also published by the Center for Interuniversity Research in Quantitative Economics (CIREQ), under number 18-2002.

ISSN 0709-9231

Testing for A Unit Root in Panels with Dynamic Factors

Hyungsik Roger Moon*
Department of Economics
University of Southern California
&
Benoit Perron†
Département de sciences économiques, CIREQ and CIRANO
Université de Montréal

September 2002

Abstract

This paper studies testing for a unit root for large n and T panels in which the cross-sectional units are correlated. To model this cross-sectional correlation, we assume that the data is generated by an unknown number of unobservable common factors. We propose unit root tests in this environment and derive their (Gaussian) asymptotic distribution under the null hypothesis of a unit root and local alternatives. We show that these tests have significant asympotitic power when the model has no incidental trends. However, when there are incidental trends in the model and it is necessary to remove heterogeneous deterministic components, we show that these tests have no power against the same local alternatives. Through Monte Carlo simulations, we provide evidence on the finite sample properties of these new tests.

Résumé:

Cet article analyse des tests de racines unitaires dans les panels où n et T sont tous les deux grands et où les unités d'observations sont corrélées. Cette corrélation transversale est modélisée à l'aide d'un modèle à facteurs dynamiques inobservables. Nous proposons plusieurs tests dans cet environnement et dérivons leurs lois limites sous l'hypothèse nulle d'une racine unitaire et des alternatives locales. Nous montrons aussi que ces tests n'ont pas de puissance contre ces mêmes alternatives locales lorsqu'il est nécessaire d'estimer des composantes déterministes. À l'aide d'une expérience de Monte Carlo, nous comparons les propriétés en échantillons finis de ces tests et de d'autres suggérés dans la littérature.

Keywords: unit roots, panel data, factor models.

Mots clés: racines unitaires, données de panels, mdèles à facteurs.

*Email: moonr@usc.edu, Tel: 213-740-2108

†Email: benoit.perron@umontreal.ca, Tel: 514-343-2126. Financial support from FCAR (Fonds pour la Formation de Chercheurs et l'Aide à la recherche) and MITACS (Mathematics of Information Technology and Complex Systems) is gratefully acknowledged.

JEL classification: C23, C33.

Classification JEL: C23, C33.

1 Introduction

In this paper we propose several unit root test statistics for panels in which cross-sections are correlated. Over the last few years, there has been a lot of research on nonstationary panels with large cross section and time series dimensions in particular in the context of testing for the presence of a unit root. A common feature of these studies is the restriction that the cross-sections are independent. Under this assumption, various central limit theorems can be applied to obtain test statistics with an asymptotic normal distribution. However, this cross sectional independence assumption is quite restrictive in many empirical applications in macroeconomics, finance or international finance. For example, consider a panel of cross-country real exchange rates. As argued by O'Connell (1998), due to the strong links across markets and the use of a numeraire country in defining real exchange rates, real exchange rates should have high cross-correlation both in the short-run and in the long run.

To model this cross sectional dependence, we consider an approximate linear dynamic factor model in which the panel data is generated by both idiosyncratic shocks and unobservable dynamic factors that are common to all the individual units but to which each individual reacts heterogeneously.

When common factors exist in the panel, tests that suppose independence among cross-sectional units will suffer size distortions. To correct this problem, we propose a test statistic that uses de-factored panel data obtained by projecting the panel data to the space orthogonal to the factor loadings. To estimate the matrix of factor loadings, we use a modified version of the principal component method used in Stock and Watson (1998) and Bai and Ng (2001, 2002). A similar orthogonalization procedure is also suggested in Phillips and Sul (2002),

Considering cross sectional dependence in a panel context is quite new. Recently, Chen and Conley (2001) study a semiparametric spatial model for fixed n and large T panels in which the time series component is stationary and mixing. For nonstationary panels, Chang (2001) develops a nonlinear instrument variable unit root test for a panel with large T and fixed n and Choi (2001) proposes a unit root test based on a Fisher-type statistic for panels with large n and T. In independent work, Bai and Ng (2001) and Phillips and Sul (2002) also use a factor structure to model cross-sectional dependence in panels and to construct unit root tests in such a setting.

An important contribution of this paper is the study of the behavior of our test statistics under local alternative hypotheses. We show that our tests have power against alternatives that shrink towards the unit root at rate $1/\sqrt{n}T$ under some circumstances. However, we also show that our tests do not have power in such a neighborhood in the case where heterogeneous deterministic trends exist in the data, the so-called incidental trends problem (cf. Moon and Phillips, 1999). We also provide an upper bound on the rate at which the alternative hypothesis can drift towards the null for nontrivial power to exist.

The paper is organized as follows. Section 2 introduces the model and proposes various test statistics for a unit root and finds their asymptotic properties under local alternatives. In section 3, we show that the proposed tests have no power against the same local alternatives when deterministic components have to be removed. In Section 4 we compare

finite sample properties of our proposed panel unit test statistics, while section 5 concludes All the technical proofs and derivations are in the Appendix.

A word on notation. We use notation M to denote a generic constant that is finite. For an $n \times K$ matrix β, $P_\beta = \beta\left(\beta'\beta\right)^{-1}\beta'$ and $Q_\beta = I - P_\beta$. For a matrix A, $A > 0$ denotes that A is positive definite. For a matrix A, $\|A\|$ denotes the Euclidean norm, $\|A\| = \left(tr\left(A'A\right)\right)^{1/2}$.

2 A Simple Model

The model we will consider is the dynamic panel model:

$$\begin{aligned} z_{it} &= \alpha_i + z_{it}^0 \\ z_{it}^0 &= \rho_i z_{it-1}^0 + y_{it}, \end{aligned} \tag{1}$$

where we set the initial observations $z_{i0}^0 = 0$ for all i. The model allows for fixed effects and is therefore suitable for use with panels of macroeconomic data that do not exhibit deterministic trends such as most real exchange rates, interest rates, or inflation rates. Deterministic trends, relevant for variables such as GDP or industrial production, will be considered in the next section. We are interested in testing the null hypothesis:

$$H_0 : \rho_i = 1 \quad \forall i$$

against the stationary alternative hypothesis:

$$H_1 : \rho_i < 1 \quad \text{for some } i$$

In order to simultaneously handle unit roots and local alternatives, we will nest these two hypotheses using the near unit root model:

$$\rho_i = 1 - \frac{\theta_i}{\sqrt{n}T}, \tag{2}$$

where θ_i is a non-negative random variable.

Assumption 1 *The random variables θ_i are iid with mean μ_θ and a finite fourth moment, and they are defined on $[0, \bar{M}_\theta]$.*[1]

With this assumption, the hypotheses we will consider are

$$\mathbb{H}_0 : \mu_\theta = 0.$$

against the local alternative

$$\mathbb{H}_1' : \mu_\theta > 0.$$

Under Assumption 1, the null hypothesis is equivalent to

$$\mathbb{H}_0 : \theta_i = 0 \text{ for all } i.$$

The near unit root model has been extensively used in the univariate literature starting with Phillips (1987) to study the behavior of test statistics under local alternatives. The rate of approach of the local alternative to the null hypothesis is faster in our case because

[1] Assuming the upper bound $\bar{M}_\theta$ is for convenience. It could be relaxed at the expense of technical complexity.

the use of panel data will entail faster convergence of the estimator of the autoregressive parameter as we will see below.

To model the correlation among the cross-sectional units, we will assume that the error term in (1) follows an approximate factor model:

$$y_{it} = \beta_i^{0\prime} f_t^0 + e_{it}, \tag{3}$$

where f_t^0 are $K-$vectors of unobservable random factors, β_i^0 are nonrandom factor loading coefficient vectors (also $K-$vectors), e_{it} are idiosyncratic shocks, and the number of factors K is possibly unknown. The extent of the correlation is determined by the factor loading coefficients β_i^0, *i.e.*,

$$E\left(y_{it}y_{jt}\right) = \beta_i^{0\prime} E\left(f_t^0 f_t^{0\prime}\right) \beta_j^0.$$

The factor model in (3) is introduced to model cross sectional dependence. The common factors f_t^0 play an important role in reducing the dimensionality of the cross sectional covariance structure of y_{it}.

Assumption 2 *(i)* $e_{it} = \sum_{j=0}^{\infty} d_{ij} v_{it-j}$, *where* v_{it} *are iid*$(0,1)$ *across* i *and over* t, *have a finite eighth moment.*

(ii) $\inf_i \sum_{j=0}^{\infty} d_{ij} > 0$.
(iii) Let $\bar{d}_j = \sup_i |d_{ij}|$ *.Then,* $\sum_{j=0}^{\infty} j^m \bar{d}_j < M$ *for some* $m > 1$.

Assumption 3 *(i)* $f_t^0 = \sum_{j=0}^{\infty} c_j u_{t-j}$, *where* c_j *are* $K \times K$ *matrices of real numbers and the* K*-vectors* u_i *are iid*$(0, I_K)$ *across* i *and over* t.

(ii) $\sum_{j=0}^{\infty} j^m \|c_j\| < M$ *for some* $m > 1$.

Assumption 4 θ_i, u_t, *and* v_{js} *are independent.*

Assumption 5 $1 \leq K \leq \bar{K} < \infty$, *where* $\bar{K}$ *is known.*

Assumption 6 *As* $n \to \infty$, $\frac{1}{n} \sum_{i=1}^{n} \beta_i^0 \beta_i^{0\prime} \to \Sigma_\beta > 0$.

Assumption 7 *As* $T \to \infty$, $\frac{1}{T} \sum_{t=1}^{T} f_t^0 f_t^{0\prime} \to_p \Sigma_f > 0$.

Define $\sigma_{e,i}^2 = \sum_{j=0}^{\infty} d_{ij}^2$, $\omega_{e,i}^2 = \left(\sum_{j=0}^{\infty} d_{ij}\right)^2$, and $\lambda_{e,i} = \sum_{l=1}^{\infty} \sum_{j=0}^{\infty} d_{ij} d_{ij+l}$. In this notation, $\sigma_{e,i}^2$ signifies the variance of e_{it}, $\omega_{e,i}^2$ the long-run variance of e_{it}, and $\lambda_{e,i}$ the one-sided long-run variance of e_{it}.

Assumption 8 *As* $n \to \infty$,
(i) $\omega_e^2 \overset{let}{=} \lim_n \frac{1}{n} \sum_{i=1}^{n} \omega_{e,i}^2 \ (> 0)$ *is well defined.*
(ii) $\phi_e^4 \overset{let}{=} \lim_n \frac{1}{n} \sum_{i=1}^{n} \omega_{e,i}^4 \ (> 0)$ *is well defined.*
(iii) $\sigma_e^2 \overset{let}{=} \lim_n \frac{1}{n} \sum_{i=1}^{n} \sigma_{e,i}^2 \ (> 0)$ *is well defined.*

Assumption 9 $\sup_i E\left(\alpha_i^2\right) < \infty$.

Remarks

(a) In the special case where $\beta_i^0 = \beta_j^0$ for all i, j, our factor model becomes an error component model with time specific effect f_t^0 as studied in Choi (2002). Phillips and Sul (2002) study a similar model with a single factor ($K = 1$). In addition to unit root testing using Hausman-type and Fisher-type tests, they also analyze median unbiased estimation and general homogeneity hypotheses. Finally, Bai and Ng (2001) develop a testing methodology for dynamic panels with factors that allows for stationary and nonstationary factors and idiosyncratic errors. Their methodology allows to test separately the nonstationarity of the factors and the idiosyncratic components.

(b) Under Assumption 3, it is known that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} f_t^0 \Rightarrow B_f(r), \tag{4}$$

$B_f(r)$ is a K-vector Brownian motion with covariance $\Omega_f = c(1) c(1)'$, $c(1) = \sum_{j=1}^{\infty} c_j$. In this paper, we do not require that the long-run covariance matrix of the factors, Ω_f, be of full rank. The positive definiteness restriction is imposed on the variance matrix of f_t, Σ_f (see Assumption 7). This implies that under the null it is possible to have cointegrating relations among the nonstationary factors. Also, note under Assumption 7 that f_t^0 is allowed to contain some lagged variables: *e.g.*, $f_t^0 = (g_t, g_{t-1})'$ for some random variable g_t with finite second moments. Then, under the null hypothesis, we have $F_t^0 = (G_t, G_{t-1})$, where $F_t^0 = \sum_{s=1}^{t} f_t^0$, $G_t = \sum_{s=1}^{t} g_s$. In this case, G_t and G_{t-1} are cointegrated in the sense that $G_t - G_{t-1} = g_t$ is stationary.

(c) However, the restrictions in Assumption 2 exclude the possibility of cointegrating relations in the integrated idiosyncratic shocks $E_{it} = \sum_{s=1}^{t} e_{is}$. The assumed independence across i implies that the covariance matrix of the stacked E_{it} is block-diagonal and Assumption 2 (*ii*) ensures that each element along the diagonal is non-zero.

(d) Assumptions 2 and 3 assume that the random factors f_t^0 and the idiosyncratic shocks e_{it} are stationary linear processes and that they are independent of each other. These assumptions correspond to Assumption 1 of Forni *et al* (2000) but are more restrictive than Assumptions C and D of Bai and Ng (2002), Assumptions C, D, and E of Bai (2001), or Condition M of Stock and Watson (1998). The conditions assumed in Bai and Ng (2002), Bai (2001), or Stock and Watson (1998) do not restrict f_t^0 and e_{it} to be linear processes and allow for weak dependence between e_{it} and f_t^0 and among the cross sectional units of e_{it}. The assumption of linear processes for e_{it} and f_t^0 is convenient yet very general in terms of the temporal dependence allowed and allows avoiding high level assumptions such as Assumption E of Bai (2001). Also, assuming independence between e_{it} and f_t^0 is not likely to be too restrictive considering the nature of the factor model and it could be relaxed as in, for example, Bai and Ng (2002) at the expense of complexity of the proofs.

(e) Assumptions 6 and 7 are standard assumptions in factor models (*e.g.*, Bai and Ng, 2002). An implication of Assumption 6 is that the contribution from each factor to at least one of the y_{it} is significant, and in this context it may correspond to Assumption 4 of Forni *et al* (2000). However, this assumption does not impose that all cross-sections respond to all factors so that some of the factor loadings could be zero.

(f) Assumption 5 assumes that there exists at least a common factor in y_{it} and the number of factors, K, is bounded by a finite number $\bar{K}$ that is assumed to be known. In this paper, we first discuss testing the null hypothesis of a unit root assuming that K is known, and then later we discuss how to estimate consistently the true number of factors.

(g) By definition, $\omega_{e,i}^2 = 2\lambda_{e,i} + \sigma_{e,i}^2$. Thus, under Assumptions 8(a) and (c),

$$\lambda_e = \lim \frac{1}{n} \sum_{i=1}^{n} \lambda_{e,i}$$

is well defined.

(h) Assumption 9 restricts the moment of the incidental parameters α_i. Under this assumption, since the stochastic trend term z_{it}^0 dominates the incidental parameters, the presence of the incidental parameters can be ignored as we will see. In Section 3, we investigate a model that does not assume the restriction in Assumption 9.

We now define our matrix notation: Define

$$\begin{aligned}
y &= \left(\underline{y}_1, ..., \underline{y}_T\right), \ \underline{y}_i = (y_{i1}, ..., y_{iT})', \\
e &= (\underline{e}_1, ..., \underline{e}_n), \ \underline{e}_i = (e_{i1}, ..., e_{iT})', \\
Z &= (\underline{Z}_1, ..., \underline{Z}_n), \ \underline{Z}_i = (z_{i1}, ..., z_{iT})', \\
Z_{-1} &= \left(\underline{Z}_{-1,1}, ..., \underline{Z}_{-1,n}\right), \ \underline{Z}_{-1,i} = (z_{i0}, ..., z_{iT-1})', \\
Z^0 &= \left(\underline{Z}_1^0, ..., \underline{Z}_n^0\right), \ \underline{Z}_i^0 = \left(z_{i1}^0, ..., z_{iT}^0\right)', \\
Z_{-1}^0 &= \left(\underline{Z}_{-1,1}^0, ..., \underline{Z}_{-1,n}^0\right), \ \underline{Z}_{-1,i}^0 = \left(z_{i0}^0, ..., z_{iT-1}^0\right)', \\
f^0 &= \left(f_1^0, ..., f_T^0\right)' \ ; \ \beta^0 = \left(\beta_1^0, ..., \beta_n^0\right)'.
\end{aligned}$$

Define

$$\rho(L) = diag(\rho_1 L, ..., \rho_n L),$$

where L denotes a lag operator. Write $l_T = (1, ..., 1)'$, $T \times 1$ vector of ones. Using our matrix notation, we rewrite the model as

$$\begin{aligned}
Z &= l_T \alpha' + Z^0, \\
Z^0 (I_n - \rho(L)) &= f^0 \beta^{0\prime} + e.
\end{aligned} \tag{5}$$

2.1 Pooled Estimators and Their Asymptotics

Define the pooled autoregressive estimator:

$$\hat{\rho}_{pool} = \frac{tr\left(Z_{-1}' Z\right)}{tr\left(Z_{-1}' Z_{-1}\right)}.$$

Our choice of the pooled estimator rests on three reasons: First, it simplifies the joint limit theory (as $n, T \to \infty$). Secondly, this allows us to analyze our tests under the local alternative $\mathbb{H}_1$. Thirdly, it is clearly appropriate for the linear structure of the model and is an implication of our null hypothesis.

If there is no common factor, *i.e*, $\beta_i^{0\prime} f_t^0 = 0$ for all i, t, then the error term y_{it} contains only idiosyncratic shocks and is thus cross-sectionally independent. It is well known in

this case that it is possible to modify the pooled estimator $\hat{\rho}_{pool}$ by fixing a second-order bias due to the serial correlation in the time series of the panel Z and make the modified pooled $\sqrt{n}T-$ consistent and have a normal limit distribution (see, for example, Levin *et al*, 2001, and Moon and Phillips, 2000).

As mentioned in the previous section, when the panel is generated by the common factors $\beta_i^{0\prime} f_t^0$ satisfying Assumptions 6 and 7, the influence of the common factor f_t^0 and cross-sectional correlation persist in all the individual units. In this case, the conventional central limit theorem cannot be applied to the conventional modified pooled estimator. Hence, this estimator is not $\sqrt{n}T-$ consistent and does not have a normal limit. The limit of the pooled estimator $\hat{\rho}_{pool}$ is derived in the following lemma.

Lemma 1 *Suppose that Assumptions 1 – 9 hold. Then, as* $(n,T \to \infty)$, *under the null of unit root,*

$$T\left(\hat{\rho}_{pool}-1\right) \Rightarrow \frac{\frac{1}{2}tr\left(B_f(1)B_f(1)'\Sigma_\beta\right)+\frac{1}{2}\omega_e^2-\frac{1}{2}tr\left(\Sigma_f\Sigma_\beta\right)-\frac{1}{2}\sigma_e^2}{tr\left(\int_0^1 B_f(r)B_f(r)'dr\right)\Sigma_\beta+\frac{1}{2}\omega_e^2},$$

where $B_f(r)$ *is the Brownian motion in* (4).

In Lemma 1, due to the presence of the nonstationary common factor and since each factor is significant in the sense that $\frac{1}{n}\sum_{i=1}^n \beta_i^0\beta_i^{0\prime}$ converges to a positive definite matrix Σ_β, the pooled estimator $\hat{\rho}_{pool}$ has a limit that is a function of $B_f(r)$, and its convergence rate is determined only by the time dimension. Adding cross-sectional units does not improve the rate of convergence.

In order to achieve $\sqrt{n}T-$ consistency and the conventional normal limit, we need to eliminate the common factors from the panel. To have an intuition, we first consider a simple case where the factor loading matrix β^0 and the coefficients of the dynamics of the idiosyncratic shock e are known. Notice that under the null hypothesis,

$$Z^0 = Z_{-1}^0 + f^0\beta^{0\prime} + e. \tag{6}$$

To eliminate the common factor in the panel, in this case, we multiply equation (6) by the projection matrix Q_{β^0}[2]. Then, under the null hypothesis, we have

$$Z^0Q_{\beta^0} = Z_{-1}^0Q_{\beta^0} + eQ_{\beta^0}. \tag{7}$$

In view of (7), we define

$$\hat{\rho}_{pool}^+ = \frac{tr\left(Z_{-1}Q_{\beta^0}Z'\right)-nT\lambda_e^n}{tr\left(Z_{-1}Q_{\beta^0}Z_{-1}'\right)}, \tag{8}$$

$\lambda_e^n = \frac{1}{n}\sum_{i=1}^n \lambda_{e,i}$. The estimator $\hat{\rho}_{pool}^+$ is a modified pooled OLS estimator using the de-factored panel data. The modification is required because of the serial correlation in eQ_{β^0}.

Define $\Theta = diag(\theta_1,...,\theta_n)$. Now to find the asymptotic distribution of $\hat{\rho}_{pool}^+$, we write by definition that

$$\begin{aligned}
&\sqrt{n}T\left(\hat{\rho}_{pool}^+-1\right)\\
=\;& \frac{\sqrt{n}\left(\frac{1}{nT}tr\left(Z_{-1}Q_{\beta^0}\left(Z-Z_{-1}\right)'\right)-\lambda_e^n\right)}{\frac{1}{nT^2}tr\left(Z_{-1}Q_{\beta^0}Z_{-1}'\right)}\\
=\;& -\frac{\frac{1}{nT^2}tr\left(Z_{-1}Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)}{\frac{1}{nT^2}tr\left(Z_{-1}Q_{\beta^0}Z_{-1}'\right)}+\frac{\sqrt{n}\left(\frac{1}{nT}tr\left(Z_{-1}^0Q_{\beta^0}e'\right)-\lambda_e^n\right)}{\frac{1}{nT^2}tr\left(Z_{-1}Q_{\beta^0}Z_{-1}'\right)}.
\end{aligned}$$

[2]Recently, Phillips and Sul (2002) independently propose a similar orthogonalization procedure.

Lemma 2 provides the asymptotic analysis of each component in this expression.

Lemma 2 *Suppose that Assumptions 1 – 9 hold. Assume that* $(n,T \to \infty)$ *with* $\frac{n}{T} \to 0$. *Then, the following holds.*
(a) $\frac{1}{nT^2} tr\left(Z_{-1} Q_{\beta^0} Z_{-1}'\right) \to_p \frac{1}{2}\omega_e^2$.
(b) $\frac{1}{nT^2} tr\left(Z_{-1} Q_{\beta^0} \Theta Z_{-1}^{0\prime}\right) \to_p \frac{1}{2}\mu_\theta \omega_e^2$.
(c) $\sqrt{n}\left(\frac{1}{nT} tr\left(Z_{-1}^0 Q_{\beta^0} e'\right) - \lambda_e^n\right) \Rightarrow N\left(0, \frac{1}{2}\phi_e^4\right)$.

Using the results in Lemma 2, we can derive the asymptotic distribution of $\sqrt{n}T\left(\hat{\rho}_{pool}^{+} - 1\right)$ as follows.

Theorem 1 *Suppose that Assumptions 1 – 9 hold. Assume that* $(n,T \to \infty)$ *with* $\frac{n}{T} \to 0$. *Then,*

$$\sqrt{n}T\left(\hat{\rho}_{pool}^{+} - 1\right) \Rightarrow N\left(-\mu_\theta, \frac{2\phi_e^4}{\omega_e^4}\right).$$

Theorem 1 shows that under the null, $\hat{\rho}_{pool}^{+}$ is $\sqrt{n}T-$ consistent and asymptotically normal. This result is fully expected because the common factors in the panel are eliminated and the de-factored data ZQ_{β^0} has no cross-sectional dependence. Also, the estimator $\hat{\rho}_{pool}^{+}$ does not depend on the factors in f_t^0. Thus, if a test is constructed using $\hat{\rho}_{pool}^{+}$, then we may expect that the test statistic is robust with respect to the factors in f_t^0.

Another way to eliminate the factors would be to project the panel data on the orthogonal space generated by the nonstationary factors F_t^0. In this paper, we do not consider this approach because handling nonstationary factors would be more complicated than handling non-random factor loading vectors and the primary interest of this paper is to eliminate the common factors $\beta_i^{0\prime} F_t^0$ under the null, and not to estimate them.

In view of Theorem 1 and Lemma 2, we may deduce that

$$\frac{\sqrt{n}T\left(\hat{\rho}_{pool}^{+} - 1\right)}{\sqrt{\frac{2\phi_e^4}{\omega_e^4}}} \Rightarrow N\left(-\mu_\theta \sqrt{\frac{\omega_e^4}{2\phi_e^4}}, 1\right) \tag{9}$$

and

$$\sqrt{n}T\left(\hat{\rho}_{pool}^{+} - 1\right)\sqrt{\frac{1}{nT^2} tr\left(Z_{-1} Q_{\beta_0} Z_{-1}'\right)\frac{\omega_e^2}{\phi_e^4}} \Rightarrow N\left(-\mu_\theta \sqrt{\frac{\omega_e^4}{2\phi_e^4}}, 1\right) \tag{10}$$

as $(n,T \to \infty)$ with $\frac{n}{T} \to 0$. These statistic are not useful in applications, however, because one does not observe the true factor loading coefficient β_i^0 and the long-run variances of the idiosyncratic shock e_{it} in general. Feasible versions of these statistics will be developed in the next section.

Notice by the Cauchy-Schwarz inequality that $\frac{\omega_e^2}{\phi_e^2} \leq 1$, and $\frac{\omega_e^2}{\phi_e^2} = 1$ only when $\omega_{e,i}^2$ are identical for all i. So, the test statistic will have a better power when the cross section data are homogeneous, more specifically, when the long-run variance of the idiosyncractic shocks e_{it} are identical across i then the panel of heterogeneous cross section.

2.2 Feasible Panel Unit Root Test Statistics

In the previous section, we have defined test statistics in (9) and (10) that are not feasible since they depend on unknown parameters. In this section, we obtain feasible versions of these statistics that will have the same asymptotic behavior. We proceed by first discussing the estimation of the long-run variances of e_{it} and factor loading coefficients assuming that the true number of the factors, K, is known. We show that the estimation of these quantities does not affect the distribution of our statistics in large samples. Finally, we discuss how to obtain a consistent estimator of K.

Estimation of β^0

To estimate β_i^0, we use the principal component method. This approach has been used widely in the literature on factor models of panels with large n and large T, for example Connor and Korajczyk (1986, 1988, 1993), Stock and Watson (1998), Bernanke and Boivin (2002), Brisson, Campbell, and Galbraith (2001) , Bai and Ng (2002), and Bai (2001).

In model (1) , since the error term y_{it} is not observable, we use the residual

$$\hat{y} = Z - \hat{\rho}_{pool} Z_{-1}.$$

To estimate β^0 and f^0, we minimize

$$V_{nT}(f, \beta, K) = \frac{tr\left(\left(\hat{y} - f\beta'\right)\left(\hat{y} - f\beta'\right)'\right)}{nT}$$

with respect to $\frac{\beta'\beta}{n} = I_K$ or $\frac{f'f}{T} = I_K$. With the normalization $\frac{\beta'\beta}{n} = I_K$, we have the estimated factor loading matrix $\bar{\beta}_K$ that is a $(n \times K)$ matrix of $\sqrt{n}$ times the eigenvectors corresponding to the K largest eigenvalues of $\hat{y}'\hat{y}$. Then, we obtain an estimator of the factor, $\bar{f}_K = \frac{1}{n}\hat{y}\bar{\beta}_K$. On the other hand, if we use the normalization $\frac{f'f}{T} = I_K$, we have the estimated factor $\check{f}_K$ that is a $(T \times K)$ matrix of $\sqrt{T}$ times the eigenvectors corresponding to the K largest eigenvalues of $\hat{y}\hat{y}'$, and the estimated factor loading $\check{\beta}_K = \frac{1}{T}\hat{y}'\check{f}_K$. Define

$$\hat{\beta}_K = \check{\beta}_K \left(\frac{\check{\beta}_K'\check{\beta}_K}{n}\right)^{1/2}, \tag{11}$$

a re-scaled estimator of the factor loading[3]. This is the estimator of β^0 that we will use in defining our feasible panel unit root test statistics.

The following lemma shows that the projection matrix $Q_{\hat{\beta}_K}$ is consistent and provides its convergence rate.

Lemma 3 *Suppose that Assumptions 1 – 9 hold. Assume that* $(n,T \to \infty)$ *with* $\frac{n}{T} \to 0$. *Then,*

$$\left\| Q_{\hat{\beta}_K} - Q_{\beta^0} \right\| = O_p\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right)\right).$$

Estimation of the long-run variances

[3]The rescaled estimator studied in Bai (2001) is $\bar{\beta}_K \left(\frac{\bar{f}_K'\bar{f}_K}{T}\right)^{-1/2}$.

In order to implement the t-statistics in (9) and (10) , we also need consistent estimators, say $\hat{\lambda}_e^n$, $\hat{\omega}_e^2$, and $\hat{\phi}_e^4$, for λ_e^n, ω_e^2, and ϕ_e^4, respectively, satisfying

$$\sqrt{n}\left(\hat{\lambda}_e^n - \lambda_e^n\right) = o_p(1), \tag{12}$$

$$\hat{\omega}_e^2 - \omega_e^2 = o_p(1), \tag{13}$$

and

$$\hat{\phi}_e^4 - \phi_e^4 = o_p(1). \tag{14}$$

In this section we propose estimators of λ_e^n, ω_e^2, and ϕ_e^4 that satisfy these conditions.

Let $\hat{e}_{it}$ denote the $(t,i)^{th}$ element of $\hat{e} = \bar{y}Q_{\hat{\beta}_K}$. Define the sample covariances $\hat{\Gamma}_i(j) = \frac{1}{T}\sum_t \hat{e}_{it}\hat{e}_{it+j}$, where the summation $\sum_t$ is defined over $1 \leq t, t+j \leq T$. To define the estimators of the long-run variances λ_e^n, ω_e^2, and ϕ_e^4, we use the following kernel estimators of $\lambda_{e,i}$ and $\omega_{e,i}^2$,

$$\hat{\lambda}_{e,i} = \sum_{j=1}^{T-1} w\left(\frac{j}{h}\right)\hat{\Gamma}_i(j) \tag{15}$$

$$\hat{\omega}_{e,i}^2 = \sum_{j=-T+1}^{T-1} w\left(\frac{j}{h}\right)\hat{\Gamma}_i(j), \tag{16}$$

where $w(\cdot)$ is a kernel function and h is a bandwidth parameter. Define

$$\hat{\lambda}_e^n = \frac{1}{n}\sum_{i=1}^n \hat{\lambda}_{e,i},\ \hat{\omega}_e^2 = \frac{1}{n}\sum_{i=1}^n \hat{\omega}_{e,i}^2,\ \text{and}\ \hat{\phi}_e^4 = \frac{1}{n}\sum_{i=1}^n \hat{\omega}_{e,i}^4. \tag{17}$$

In order for the estimators $\hat{\lambda}_e^n$, $\hat{\omega}_e^2$, and $\hat{\phi}_e^4$ to satisfy the desirable properties in $(12)-(14)$, we need the following assumptions on the kernel function and the bandwidth parameter.

Assumption 10 *(Restriction on the convergence rate of n and T). The size of the panel (n,T) tends to infinity with* $\liminf_{(n,T\to\infty)} \frac{\log T}{\log n} > 1$.

Define $a = \liminf_{(n,T\to\infty)} \frac{\log T}{\log n}$. The parameter a is related to the speed of $\frac{n}{T}$ tending to zero. The restriction $a > 1$ implies that $(n, T \to \infty)$ with $\frac{n}{T} \to 0$ because for n, T large,

$$\frac{n}{T} = e^{\log n - \log T} = e^{\left(1-\frac{\log T}{\log n}\right)\log n} = n^{\left(1-\frac{\log T}{\log n}\right)} \leq n^{(1-a)} \to 0.$$

The above assumption allows the parameter a to be infinity.

Assumption 11 *(Kernel Conditions) The kernel function $w(\cdot) : \mathbb{R} \to [0,1]$ is continuous at zero and all but a finite number of other points, satisfying*

$$\textit{(i)}\ w(0) = 1,\ w(x) = w(-x),\ \int_{-\infty}^{\infty} w(x)^2\,dx < M,$$

$$\textit{(ii)}\ w_q = \lim_{x\to 0}\left[1 - w(x)/|x|^q\right] < \infty$$

for some $0 < q \leq m$, where parameter m is defined in Assumptions 2 and 3.

In some cases, we need to strengthen this assumption to:

Assumption 12 *(Kernel Conditions*) The kernel function $w(\cdot)$ satisfies the kernel conditions in Assumption 11 as well as*

$$\textit{(iii)}\ \max\left\{1, \frac{1}{a-1}\right\} < q.$$

The parameter q is related to the smoothness of the kernel $w(\cdot)$ at zero. It is well known that for the truncated kernel, $w_q = 0$ for all $q < \infty$, for the Bartlett kernel, $w_q < \infty$ only if $q \leq 1$, and for the Parzen, Tukey-Hanning, and Quadratic-Spectral kernel, $w_q < \infty$ only if $q \leq 2$ (*e.g.*, see Andrews, 1991). The requirement $0 < q \leq m$ is related to the smoothness of the spectral densities of e_{it} and f_t^0.

Assumption 13 *(Bandwidth Conditions) The bandwidths h_λ, h_ω, and h_ϕ tend to infinity satisfying the following conditions.*

(a) $h_\lambda \sim n^b$ with $\frac{1}{2q} < b < \min\left\{\frac{a-1}{2}, \frac{a}{q}, \frac{1}{2}\right\}$.

(b) For $0 < q < 1$, $h_\omega \sim n^b$ with $0 < b < \min\left\{1, \frac{a}{2}\right\}$. For $q \geq 1$, $h_\omega \sim n^b$ with $0 < b < \min\left\{1, \frac{a}{2}, \frac{a}{q}\right\}$.

(c) For $0 < q < 1$, $h_\phi \sim n^b$ with $0 < b < \frac{1}{4}$. For $q \geq 1$, $h_\phi \sim n^b$ with $0 < b < \min\left\{\frac{1}{4}, \frac{a}{q}\right\}$.

Lemma 4 *Suppose that Assumptions 1 – 10 hold.*

(a) If the kernel window satisfies Assumption 12 and the bandwidth h_λ satisfies Assumption 13(a), then,

$$\sqrt{n}\left(\hat{\lambda}_e^n - \lambda_e^n\right) = o_p(1).$$

(b) If the kernel window satisfies Assumption 11 and the bandwidth h_ω satisfies Assumption 13(b), then

$$\hat{\omega}_e^2 - \omega_e^2 = o_p(1).$$

(c) If the kernel window satisfies Assumption 11 and the bandwidth h_ϕ satisfies Assumption 13(c), then

$$\hat{\phi}_e^4 - \phi_e^4 = o_p(1).$$

In view of (9) and (10), using (11) and (17), we may define the following feasible t-statistics for $\mathbb{H}_0$:

$$t_a^* = \frac{\sqrt{n}T\left(\hat{\rho}_{pool}^* - 1\right)}{\sqrt{\frac{2\hat{\phi}_e^4}{\hat{\omega}_e^4}}},$$

and

$$t_b^* = \sqrt{n}T\left(\hat{\rho}_{pool}^* - 1\right)\sqrt{\frac{1}{nT^2}tr\left(Z_{-1}Q_{\hat{\beta}_K}Z_{-1}'\right)}\left(\frac{\hat{\omega}_e}{\hat{\phi}_e^2}\right)$$

where

$$\hat{\rho}_{pool}^* = \frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}Z'\right) - nT\hat{\lambda}_e^n}{tr\left(Z_{-1}Q_{\hat{\beta}_K}Z_{-1}'\right)}.$$

Theorem 2 *Suppose that Assumptions 1 – 13 hold. Then,*

$$t_a^*,\ t_b^* \Rightarrow N\left(-\mu_\theta \sqrt{\frac{\omega_e^4}{2\phi_e^4}}, 1\right).$$

Remarks

(a) Theorem 2 indicates that the t-ratio statistics t_a^* and t_b^* have significant asymptotic power in $\frac{1}{\sqrt{nT}}-$ neighborhoods of the null of unit root. As expected, the power increases as μ_θ deviates from zero.

(b) The finite-sample performance of univariate unit root tests suffer from the difficulty of estimating long-run variances such $\omega_{e,i}^2$. Our panel tests should perform better in this regard since what is required is the *average* $\omega_{e,i}^2$. This averaging should remove a lot of the uncertainty in long-run variance estimation. Of course, this averaging will not remove the bias in the estimation of the long-run variances.

2.3 Estimation of the number of factors

As mentioned in the beginning of the previous section, all the results in the previous section are established under the condition that the number of the factors, K, is known. When it is unknown, a natural approach is treat the estimation problem as a model selection issue and to estimate K using an information criterion. In this section, we discuss how to obtain a consistent estimator of the number of factors, K, using this approach. Now for a given $(n \times r)$ matrix β_r, let

$$W_{nT}\left(\beta_r, r\right) = \min_{f_r} \frac{tr\left(\left(\hat{y} - f_r\beta_r'\right)\left(\hat{y} - f_r\beta_r'\right)'\right)}{nT} = \frac{tr\left(\hat{y}'Q_{\beta_r}\hat{y}\right)}{nT}.$$

To estimate the true number of factors, K, Bai and Ng (2002) propose to maximize the following criterion functions,

$$\begin{aligned} PC\left(r\right) &= W_{nT}\left(\hat{\beta}_r, r\right) + rG_{nT}, \\ IC\left(r\right) &= \ln\left(W_{nT}\left(\hat{\beta}_r, r\right)\right) + rG_{nT}, \end{aligned}$$

where the penalty function $G_{n,T}$ satisfies (i) $G_{n,T} \to 0$ and (ii) $\min\left\{n,T\right\} G_{n,T} \to \infty$. as $(n,T \to \infty)$.

Theorem 3 *Suppose that Assumptions 1 – 9 hold and* $(n,T \to \infty)$ *following Assumption 10. Let*

$$\hat{K} = \underset{1 \le r \le \bar{K}}{\arg\min} PC\left(r\right),\ \check{K} = \underset{1 \le r \le \bar{K}}{\arg\min} IC\left(r\right).$$

Then,

$$(a)\ plim1\left\{\hat{K} = K\right\} = 1\ and\ (b)\ plim1\left\{\check{K} = K\right\} = 1.$$

The specific forms of the penalty function proposed by Bai and Ng (2002) are:

$$\begin{aligned} G_{PC,1,nT} &= \hat{\sigma}_e^{2,n}\frac{n+T}{nT}\ln\left(\frac{n+T}{nT}\right), \\ G_{PC,2,nT} &= \hat{\sigma}_e^{2,n}\frac{n+T}{nT}\ln\left(\min\left\{n,T\right\}\right), \\ G_{PC,3,nT} &= \hat{\sigma}_e^{2,n}\frac{n+T}{nT}\left(\frac{\ln\left(\min\left\{n,T\right\}\right)}{\min\left\{n,T\right\}}\right), \end{aligned}$$

and

$$\begin{aligned} G_{IC,1,nT} &= \frac{n+T}{nT} \ln\left(\frac{n+T}{nT}\right), \\ G_{IC,2,nT} &= \frac{n+T}{nT} \ln\left(\min\left\{n,T\right\}\right), \\ G_{IC,3,nT} &= \frac{n+T}{nT} \left(\frac{\ln\left(\min\left\{n,T\right\}\right)}{\min\left\{n,T\right\}}\right). \end{aligned}$$

We also consider the modified BIC criterion that they proposed which they called BIC_3 :

$$BIC_3 = W_{nT}\left(\hat{\beta}_r, r\right) + r\hat{\sigma}_e^{2,n} \frac{n+T}{nT} \ln\left(nT\right)$$

because their evidence suggests that it performs better in selecting the number of factors when $\min\left(n,T\right)$ is small (≤ 20) as is often the case in empirical applications (such as ours). Bai and Ng rejected this criterion because it does not satisfy the required conditions for consistency when either n or T dominates the other one exponentially, but this appears to be a rather unusual case. For small n and T of roughly the same magnitude, this criterion preformed best among those they considered.

3 A Model with Incidental Trends

In this section, we extend our analysis and consider the dynamic panel model that may include incidental trends:

$$\begin{aligned} z_{it} &= \alpha_{ki}' g_{kt} + z_{it}^0 \\ z_{it}^0 &= \rho_i z_{it-1}^0 + y_{it}, \end{aligned} \tag{18}$$

where

$$g_{0t} = 1 \text{ and } g_{1t} = \left(1,t\right)'.$$

We continue to assume the local-to-unity framework (2) for ρ_i and the approximate factor structure (3) for y_{it}. We also assume that $z_{i0}^0 = 0$ for all i. As in the previous section, we want to test for the null hypothesis $\mathbb{H}_0$ against the (local) alternative $\mathbb{H}_1$.

The model (18) is an extension of the model in the previous section as it adds incidental trend components $\alpha_{ki}' g_{kt}$ representing individual effects. When $k = 0$, *i.e.*, $g_{kt} = 1$, the model with incidental trends (18) reduces to our original model (1). However, in this section we do not assume any restriction on the incidental parameters (or trends) such as Assumption 9. To distinguish the two different models of $k = 0$ and $k = 1$, we call them model $k = 0$ and model $k = 1$, respectively.

The main purpose of this section is to study the local power of t-ratio type statistics based on a (bias-modified) pooled estimator (such as $\hat{\rho}^+$) using detrended (or demeaned) panel data. Detrending (or demeaning) is required to eliminate the incidental trends (parameters) in the model and have a test statistic that is independent of the incidental trends (or parameters).

To make our point, we will simplify our analysis as much as possible by making more restrictive assumptions than in the previous section. In particular, we assume that $e_{it} \sim iid\left(0,1\right)$ with finite fourth moments across i and over t, $f_t \sim iid\left(0,1\right)$ over t, and e_{is} and f_t are independent. We also assume for convenience that the factor loading coefficient β_i^0 is observed.

Finally, we replace our local-to-unity framework of (2) by assuming that the autoregressive roots are identical cross i and assume that

$$\rho_i = 1 - \frac{\mu_\theta}{n^\eta T} \text{ for all } i.$$

In what follows, we will investigate the asymptotic powers of the models $k = 0$ and $k = 1$ within a $\frac{1}{n^\eta T}-$ neighborhood of the null hypothesis of a unit root and find that the test has no asymptotic power if $\eta > \frac{1}{4}$ for $k = 0$ and $\eta > \frac{1}{6}$ for $k = 1$. The restrictions made in this section could be relaxed to the more general conditions assumed in the previous section without changing any of the main results.

To distinguish the notation for the panel under the null hypothesis, we denote $Z^0(0)$ for Z^0 in (1) when $\rho_i = 1$ (or equivalently $\theta_i = 0$) for all i. Define $F^0 = \Xi f^0$ and $E = \Xi e$, where Ξ be a $(T \times T)$ lower triangular matrix such that

$$\underset{(T\times T)}{\Xi} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & & 0 \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

So,

$$Z^0(0) = F^0 \beta^{0\prime} + E.$$

Similarly we define $Z^0_{-1}(0)$, F^0_{-1}, and E_{-1} to denote the matrices of lagged panel data of $Z^0(0)$, F^0, and E, respectively. Define $G_{kT} = (g_{k1}, ..., g_{kT})'$, a $T \times (k+1)$ trend matrix (or a vector consisting of ones). Let $Q_G = I_T - G_{kT}(G'_{kT}G_{kT})^{-1} G'_{kT}$. When A is a $(T \times n)$ matrix of panel data, we denote $\tilde{A} = Q_G A$, the matrix of panel data from which the trends (or the means) along the time dimension have been subtracted. Using this notation together with the notation defined in the previous section, we may deduce that

$$\tilde{Z} = \tilde{Z}^0$$

and

$$\tilde{Z}^0(I_n - \rho(L)) = \tilde{f}^0 \beta^{0\prime} + \tilde{e}.$$

Analogous to $\hat{\rho}^+_{pool}$ in (8), we may define

$$\hat{\rho}^{\#}_{pool} = \frac{tr\left(\tilde{Z}_{-1} Q_{\beta^0} \tilde{Z}'\right) - nTb_{k,nT}}{tr\left(\tilde{Z}_{-1} Q_{\beta^0} \tilde{Z}'_{-1}\right)}$$

where the bias term is $b_{k,nT} = \frac{1}{nT} E\left(tr\left(\tilde{E}_{-1}\tilde{e}'\right)\right)$. The limit of the bias $b_{kn,T}$ as $T \to \infty$ is $b_k = -E\left(\int_0^1 \int_0^1 W(r) h_k(r,s) dW(s) dr\right)$[4], where $W(r)$ is a Wiener process, $h_k(r,s) = g_k(r)' \left(\int_0^1 g_k(r) g_k(r)' dr\right)^{-1} g_k(s)$, $g_0(r) = 1$ and $g_1(r) = (1, r)'$. The correction term $b_{k,nT}$ is the mean of the bias generated by the correlation between the detrended data $\tilde{E}_{-1}$ and $\tilde{e}$.

The typical t- ratio statistic is defined as

$$t^{\#} = \sqrt{tr\left(\tilde{Z}_{-1} Q_{\beta^0} \tilde{Z}'_{-1}\right)} \left(\hat{\rho}^{\#}_{pool} - 1\right).$$

[4] A direct calculation shows that $b_k = -\frac{1}{2}$ for $k = 0$ and 1.

Then, we may write

$$
\begin{aligned}
t^{\#} &= \frac{\sqrt{n}\left[\frac{1}{nT}tr\left\{\tilde{Z}_{-1}^{0}Q_{\beta^{0}}\left(-\frac{\mu_{\theta}}{n^{\eta}T}\tilde{Z}_{-1}^{0}+\tilde{y}\right)'\right\}-b_{k,nT}\right]}{\sqrt{\frac{1}{nT^{2}}tr\left(Z_{-1}Q_{\beta^{0}}Z_{-1}'\right)}} \\
&= \frac{\sqrt{n}\left[\frac{1}{nT}tr\left(\tilde{E}_{-1}\tilde{e}'\right)-b_{k,nT}\right]}{\sqrt{\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}Q_{\beta^{0}}\tilde{Z}_{-1}'\right)}}+\frac{\sqrt{n}\left[\frac{1}{nT}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^{0}}\tilde{e}'\right)\right]}{\sqrt{\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}Q_{\beta^{0}}\tilde{Z}_{-1}'\right)}}. \\
&\quad -\frac{\mu_{\theta}}{n^{\eta}}\frac{\sqrt{n}\left[\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}\left(0\right)Q_{\beta^{0}}\tilde{Z}_{-1}'\left(0\right)\right)\right]}{\sqrt{\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}Q_{\beta^{0}}\tilde{Z}_{-1}'\right)}} \\
&\quad +\mu_{\theta}\frac{\sqrt{n}}{n^{\eta}}\frac{\frac{1}{nT^{2}}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^{0}}\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)'\right)}{\sqrt{\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}Q_{\beta^{0}}\tilde{Z}_{-1}'\right)}} \\
&\quad -2\mu_{\theta}\frac{\sqrt{n}}{n^{\eta}}\frac{\frac{1}{nT^{2}}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^{0}}\tilde{Z}_{-1}'\left(0\right)\right)}{\sqrt{\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}Q_{\beta^{0}}\tilde{Z}_{-1}'\right)}}
\end{aligned}
$$

Lemma 5 provides the asymptotic analysis of each component in this expression.

Lemma 5 *Assume that $\eta > \frac{1}{4}$ for model $k = 0$ and $\eta > \frac{1}{6}$ for model $k = 1$. Under the assumptions made in this section, the following hold.*

(a) $\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}^{0}Q_{\beta^{0}}\tilde{Z}_{-1}^{0\prime}\right) \rightarrow_{p} \left(\int_{0}^{1} rdr - \int_{0}^{1}\int_{0}^{1} \min\left(r,s\right) h_{k}\left(r,s\right) dsdr\right)$.

(b) $\sqrt{n}\left[\frac{1}{nT}tr\left(\tilde{E}_{-1}\tilde{e}'\right)-b_{k,nT}\right] \Rightarrow N\left(0, \lim_{n,T} E\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{E}_{it-1}\tilde{e}_{it}-b_{k,nT}\right)^{2}\right)$.

(c) $\sqrt{n}\left[\frac{1}{nT}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^{0}}\tilde{e}'\right)-\frac{\mu_{\theta}}{n^{\eta}}\int_{0}^{1}\int_{0}^{r}\left(r-s\right)h_{k}\left(r,s\right)dsdr\right] = o_{p}\left(1\right)$.

(d) $n^{1/2-\eta}\left[\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}^{0}\left(0\right)Q_{\beta^{0}}\tilde{Z}_{-1}^{0\prime}\left(0\right)\right)-\left(\int_{0}^{1} rdr - \int_{0}^{1}\int_{0}^{1} \min\left(r,s\right) h_{k}\left(r,s\right) dsdr\right)\right] = o_{p}\left(1\right)$.

(e) $\frac{1}{n^{1/2+\eta}T^{2}}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^{0}}\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)'\right) = o_{p}\left(1\right)$.

(f) $\frac{1}{n^{1/2+\eta}T^{2}}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^{0}}\tilde{Z}_{-1}'\left(0\right)\right) = o_{p}\left(1\right)$.

Note that according to Remark (c) on page 950 of Moon and Phillips (2000), we have

$$\int_{0}^{1} rdr - \int_{0}^{1}\int_{0}^{1} \min\left(r,s\right) h_{k}\left(r,s\right) dsdr = \int_{0}^{1}\int_{0}^{r}\left(r-s\right)h_{k}\left(r,s\right)dsdr \text{ for } k=0,1.$$

Then, we may deduce that for $\eta > \frac{1}{4}$ in model $k = 0$ and $\eta > \frac{1}{6}$ in model $k = 1$,

$$
\begin{aligned}
t^{\#} &= \frac{\sqrt{n}\left[\frac{1}{nT}tr\left(\tilde{E}_{-1}\tilde{e}'\right)-b_{k,nT}\right]}{\sqrt{\frac{1}{nT^{2}}tr\left(\tilde{Z}_{-1}Q_{\beta^{0}}\tilde{Z}_{-1}'\right)}}+o_{p}\left(1\right) \\
&\Rightarrow N\left(0, \frac{\lim_{n,T} E\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{E}_{it-1}\tilde{e}_{it}-b_{k,nT}\right)^{2}}{\int_{0}^{1} rdr - \int_{0}^{1}\int_{0}^{1} \min\left(r,s\right) h_{k}\left(r,s\right) dsdr}\right),
\end{aligned}
$$

from which we find that the limit distribution does not depend on μ_θ and is identical under $\mathbb{H}_0$ and $\mathbb{H}_1$. Therefore, the $t^{\#}$, the t-ratio statistic based on $\hat{\rho}^{\#}_{pool}$ does not have any asymptotic power in $\frac{1}{n^{\eta}T}-$ neighborhoods of the null of a unit root, where $\eta > \frac{1}{4}$ in model $k = 0$ and $\eta > \frac{1}{6}$ in model $k = 1$. Summarizing this, we have the following theorem.

Theorem 4 *Assume that $\eta > \frac{1}{4}$ for model $k = 0$ and $\eta > \frac{1}{6}$ for model $k = 1$. Under the assumptions made in this section, the $t^{\#}$ statistic does not have any asymptotic power in $\frac{1}{n^{\eta}T}-$ neighborhoods of the null of unit root.*

Remarks

(a) Again, as mentioned earlier, the same result as Theorem 4 could be obtained under more general conditions than those assumed in the previous section at the cost of complexity. However, due to space limitations, we omit this.

(b) Recently, Ploberger and Phillips (2002) have investigated optimal invariant panel unit root tests that maximizes average power, and Moon and Phillips (2002) and Moon, Perron, and Phillips (2002) have proposed a similar test using the OLS detrended data[5]. One of their main findings is that when the panel includes incidental trends (in our case, it corresponds to model $k = 1$), their tests have power within $\frac{1}{n^{1/4}T}$ neighborhoods of the null of unit root. In view of Theorem 4, their tests asymptotically dominates the test based on $t^{\#}$ since the latter does not have any power in such a shrinking neighborhood.

4 Simulation experiment

In this section, we will conduct a small Monte Carlo experiment to assess the finite-sample properties of the tests presented on section 2. We also want to demonstrate the lack of power of the tests when deterministic components are estimated. For this purpose, the data generating process we will use is:

$$\begin{aligned} z_{it} &= \alpha_{i0} + \alpha_{i1} t + z^0_{it} \\ z^0_{it} &= \rho_i z^0_{it-1} + y_{it} \\ z^0_{i0} &= 0 \end{aligned}$$

with a factor structure for the error terms:

$$y_{it} = \tau \sum_{j-1}^{K} \beta_{ij} f_{tj} + \sqrt{K} e_{it}$$

and all stochastic shocks i.i.d. standard normal:

$$\left(f_{tj}, \beta_{ij}, e_{it}\right) \sim iidN\left(0, I_3\right)$$

The DGP for y_{it} is identical to the one used by Bai and Ng (2002) for $\tau = 1$. We add to that an autoregressive structure as well as deterministic components.

We will conduct 2 experiments that will differ according to the specifications of the deterministic components.:

$$\begin{aligned} \text{Experiment 1 (fixed effects, no trend)} &: \quad \alpha_{i0} \sim N\left(0,1\right), \alpha_{i1} = 0 \\ \text{Experiment 2 (deterministic trend)} &: \quad \alpha_{i0} \sim N\left(0,1\right), \alpha_{i1} \sim N\left(0,1\right) \end{aligned}$$

[5] However, these papers do not consider cross sectional dependence in the panels.

For each of these data-generating processes, we will consider two assumptions on the autoregressive parameters:

$$\begin{aligned} \text{Case } A &: \quad \rho_i = 1 \quad \forall i \\ \text{Case } B &: \quad \rho_i \sim U\left[0.98, 1\right] \end{aligned}$$

Case A is used to study the size of the proposed tests. Case B is used to study the power of the tests. Two distinctive features of this case are worth emphasizing. First, we do not impose a common autoregressive parameter under the alternative. As we have shown in section 2, our test is consistent against this type of alternative despite our pooling approach. Of course, tests that do allow for different autoregressive parameters under the alternative hypothesis may prove to be more powerful. Secondly, we consider a fixed alternative regardless of the size of n and T to show the increased power as n and/or T increase. Thus in terms of the theoretical framework of section 2, the random variable representing the local alternative, θ_i, has a uniform distribution over the interval $\left[0, \frac{\sqrt{n}T}{50}\right]$. The chosen specification ensures that the average value of ρ is 0.99.

Finally, we will consider three values of the parameter controlling the relative importance of common versus idiosyncratic shocks τ : 0, 1, and 10. A larger value of τ represents a greater importance of the common shocks relative to the idiosyncratic ones. The base case $\tau = 1$ represents a situation in which the two components have equal importance, and the case $\tau = 0$ corresponds to the absence of common factors (i.e. an independent panel).

We set the true number of factors at $K = 1$. For the estimation of the number of factors we use the IC and BIC_3 criteria and follow Bai and Ng and set the maximum number of factors $\overline{K} = 8$. Since the three IC criteria gave essentially identical results, we only report the results using the IC_3 criterion here. The long-run variances are estimated using the Andrews-Monahan (1992) estimator using the quadratic spectral kernel and prewhitening. All tests are carried out at the 5% significance level, and the number of replications is set at 1000.

We choose two values for each of n and T : $n = 10$ or 20 and $T = 100$ or 300. The larger value of the time dimension T corresponds to panels of about 25 years of monthly data, approximately the length of post Bretton Woods exchange rate data often used for testing of the purchasing power parity (PPP) theory. The smaller value of T is meant to represent similar panels of quarterly data.

Table 1 presents the results for the size of the tests in experiment 1. Except for the last two columns, each entry in the table represents the percentage of replications in which the null hypothesis of a unit root is rejected using the appropriate test statistic. The first column provides results for our t_a^* test with the number of factors either set to the true number (1), estimated using the information criteria suggested by Bai and Ng (2002) and BIC_3, or set to 0 (*i.e.* neglecting to defactor the data). The following column gives the same information for the t_b^* test. The last two columns give the mean estimated number of factors using either the IC_3 or BIC_3 criteria.

In almost all cases, the test based on the t_b^* statistic has better size properties . As pointed out by Bai and Ng (2002) , with at least 20 cross-sections, the number of factors is estimated with extreme precision. This means that for the small number of cross-sections ($n = 10$) , the number of factors is grossly overestimated by their information criteria, and this distorts the size of our tests. The use of the BIC_3 criterion alleviates this problem to some degree. In fact, with such a small number of cross-sectional units, it is often preferable to use tests that assume cross-sectional independence rather than estimate the factor structure.

The value of the parameter τ controls the relative importance of common versus idiosyncratic shocks. Looking down the table, it appears that the more important are the

common shocks (larger value of τ), the more difficult it is to control size with the small value of n. This results stems from the imprecise estimation of the factors allowed by the small number of cross sections. The last part of table 1 shows the price to pay for handling cross-sectional dependence when it is not present (the case where $\tau = 0$). The size distortions in that case are quite mild for the t_b^* test.

Table 2 presents power results for the first experiment. Remember that the autoregressive parameter varies with i but has a mean value of 0.99. Each cell in the table has two entries, the first one is raw power and the second one is size-adjusted power. Our tests have very good power for this very difficult alternative hypothesis. As expected, power goes down as the importance of the factor (characterized by the parameter τ) increases.

Table 3 reports the size results for experiment 2 (with a deterministic trend). The same results as without trend are observed. With a small number of cross-sections, we overestimate the number of factors and obtain large size distortions as a result. Once that number is well estimated (with 20 cross-sections), our tests have good size properties, especially when $T = 300$. The size properties also do not seem to be much affected by the value of τ in this case.

Finally, table 4 reports the power of our test with deterministic trends. As expected, power is almost absent. Our theoretical result that the distribution of the statistics is the same under the null and local alternative hypotheses implies that size-adjusted power should be approximately equal to size or 5%. Our results are quite suggestive of that.

In summary, our tests require at least 20 cross-sectional units for precise estimation of the number of factors. Once this is accomplished, our tests show good size properties and are quite powerful when no deterministic component is estimated. As we have shown theoretically, our tests have no power (beyond size) under local alternative hypotheses when deterministic components have to be estimated.

5 Conclusion

This paper has developed new procedures for testing the null hypothesis of a unit root in cross-sectionally correlated panels of large dimensions. This is quite important since recent use of panels involving long time series of a large number of countries, regions or industries has been rapidly increasing. Assuming independence among cross-sectional units in such cases seems very restrictive as there should be common shocks such as business cycle effects. The approach used here is to model the dependence among cross-sectional units through an approximate factor model. Conditional upon these factors, the cross-sectional units are assumed to be independent, though their idiosyncratic shocks could be serially correlated.

We have shown that with individual fixed effects, we can construct tests based on a bias-modified pooled estimate of the autoregressive parameter that have power in a local neighborhood that shrinks towards the null hypothesis at rate $\frac{1}{\sqrt{n}T}$. The limiting distribution of our test statistics is normal and therefore no special table is required to compute p-values.

Secondly, we have shown that corresponding tests in cases where heterogeneous deterministic components have to be estimated has no power in such neighborhoods. We have given upper bounds on the rate at which local alternative must approach the unit root null hypothesis in order for nontrivial asymptotic power to exist in such cases. This rate is slower than the $\frac{1}{\sqrt{n}T}$ neighborhood obtained when no deteministic components have to be estimated and is $\frac{1}{n^{1/6}T}$ with deterministic linear trends.

Finally, we have provided simulation evidence that supports our theoretical results. In particular, we have shown that we can have tests with good size and excellent power

when no estimation of deterministic components is necessary. When such estimation is necessary, the tests typically have no power beyond their size.

6 Appendix A: Preliminary Results

Suppose that A and B are $(n \times n)$ matrices. The following facts will be used frequently in the following proofs; (a) $tr(AB) \leq \|A\| \|B\|$ by the Cauchy-Schwarz inequality, (b) if A is symmetric and positive semidefinite, then $\|A\| \leq tr(A)$ and $tr(A) \leq \sqrt{n}\|A\|$, and (c) if both of A and B are positive semidefinite, then $tr(AB) \leq tr(A)\|B\|$, and $tr(AB) \leq tr(B)\|A\|$. To distinguish the notation for the panel with $\rho_i = 1 (c_i = 0)$ for all i, we denote $Z(0)$ and $Z^0(0)$ for Z and Z^0 in (5), respectively. Also we define $F^0 = \Xi f^0$ and $E = \Xi e$, where Ξ be a $(T \times T)$ lower triangular matrix such that

$$\underset{(T\times T)}{\Xi} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & & 0 \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

So,

$$\begin{aligned} Z(0) &= l_T \alpha' + Z^0(0), \\ Z^0(0) &= F^0 \beta^{0\prime} + E, \end{aligned}$$

where $l_T = \underbrace{(1, ..., 1)}_{T}'$. Similarly we define $Z_{-1}, Z_{-1}(0), Z^0_{-1}(0)$, F^0_{-1}, and E_{-1} to denote the matrices of lagged panel data of $Z, Z(0), Z^0(0)$, F^0, and E, respectively. We denote

$$\rho = diag(\rho_1, ..., \rho_n).$$

We use M to denote a generic positive constant.

The following lemmas are useful in proving the main results of the paper.

Lemma 6 *Let* $X_{ij,T} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (e_{it}e_{jt} - E(e_{it}e_{jt}))$. *Under Assumption 2,* $sup_{i,j} E\left(X^4_{ij,T}\right) < M$.

Lemma 7 *Under Assumptions 1 – 9, the following hold. Let* $E_{it} = \sum_{s=1}^{t} e_{is}$ *with* $E_{i0} = 0$.

(a) As $(n, T \to \infty)$, $\frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \theta_i E^2_{it-1} \to_p \frac{1}{2}\mu_\theta \omega_e^2$.

(b) As $(n, T \to \infty)$ *with* $\frac{n}{T} \to 0$, $\sqrt{n}\left(\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} E_{it-1}e_{it} - \lambda^n_e\right) \Rightarrow N\left(0, \frac{1}{2}\phi^4_e\right)$.

Lemma 8 *We assume Assumptions 1 – 9. Then, Parts (h) holds as* $(n, T \to \infty)$ *with* $\frac{n}{T} \to 0$, *and the other parts hold as* $(n, T \to \infty)$, *where*

(a) $\frac{1}{nT^2}\|Z_{-1}\|^2 = O_p(1)$,

(b) $\frac{1}{nT}\left\|Z'_{-1}y + y'Z_{-1}\right\| = O_p(1)$,

(c) $\frac{1}{nT^2} tr\left(\beta^{0\prime}E'_{-1}E_{-1}\beta^0\right) = O_p(1)$,

(d) $\frac{1}{nT}\left\|\beta^{0\prime}e'E_{-1}\beta^0 + \beta^{0\prime}E'_{-1}e\beta^0\right\| = O_p(1)$,

(e) $\frac{1}{nT}\|e'e\| = O_p\left(\max\left(\frac{1}{\sqrt{T}}, \frac{1}{\sqrt{n}}\right)\right)$,

(f) $\frac{1}{nT} tr\left(f^{0\prime}ee'f^0\right) = O_p(1)$,

(g) $\frac{1}{nT} tr\left(\beta^{0\prime}e'e\beta^0\right) = O_p(1)$,

(h) $\frac{1}{n\sqrt{n}T\sqrt{T}}\left(\beta^{0\prime}e'eE_T\right) = o_p(1)$, *where* $E_T = e'l_T$,

(i) $\frac{1}{\sqrt{n}\sqrt{T}}\left\|\beta^{0\prime}e'l_T\right\| = O_p(1)$. *(i*)* $\frac{1}{\sqrt{n}\sqrt{T}}\|e'l_T\| = O_p(1)$, *(i**)* $\frac{1}{\sqrt{n}\sqrt{T}}\|y'l_T\| = O_p(1)$.

(j) $\frac{1}{\sqrt{n}T\sqrt{T}}\left\|\beta^{0\prime}E'_{-1}l_T\right\| = O_p(1)$,

(k) $\frac{\|y\|^2}{nT} = O_p(1)$,

(l) $\frac{1}{nT\sqrt{T}}\left|\alpha'E'_{-1}l_T\right| = O_p(1)$.

Lemma 9 *Suppose that Assumptions 1 – 9 hold. Then, the following hold.*

(a) $\frac{1}{nT^2}\sum_{i=1}^{n}\theta_i^2\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\frac{t-s-1}{T}\beta_i^{0\prime}f_s^0\right)^2=O_p(1)$.

(b) $\frac{1}{nT^2}\sum_{i=1}^{n}\theta_i^2\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\frac{t-s-1}{T}e_{is}\right)^2=O_p(1)$.

(c) $\frac{1}{nT}\sum_{i=1}^{n}\theta_i\beta_i^{0\prime}\left(\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{t-s-1}{T}f_s^0f_t^{0\prime}\right)\beta_i^0=O_p(1)$.

(d) $\frac{1}{nT}\sum_{i=1}^{n}\theta_i\beta_i^{0\prime}\left(\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{t-s-1}{T}f_s^0e_{it}\right)=o_p(1)$.

(e) $\frac{1}{nT}\sum_{i=1}^{n}\theta_i\left(\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{t-s-1}{T}e_{is}e_{it}\right)=o_p(1)$.

(f) $\left\|\frac{1}{n^2T}\sum_{i=1}^{n}\sum_{i=1}^{n}\beta_i^0\beta_i^{0\prime}\theta_j\sum_{t=2}^{T}\tilde{e}_{it}\left(\sum_{s=1}^{t-1}\frac{t-s-1}{T}\tilde{f}_s^0\right)\beta_j^{0\prime}\right\|=o_p(1)$

(g) $\left\|\frac{1}{n^2T}\sum_{i=1}^{n}\sum_{i=1}^{n}\beta_i^0\beta_i^{0\prime}\theta_j\sum_{t=2}^{T}\tilde{e}_{it}\left(\sum_{s=1}^{t-1}\frac{t-s-1}{T}\tilde{e}_{js}\right)\beta_j^{0\prime}\right\|=o_p(1)$.

(h) $\frac{1}{n\sqrt{T}}\left\|\sum_{i=1}^{n}\theta_i\beta_i^0\beta_i^{0\prime}\left(\sum_{s=1}^{T}\left(1-\frac{s}{T}\right)f_s^0\right)\right\|=O_p(1)$.

(i) $\frac{1}{n\sqrt{T}}\left\|\sum_{i=1}^{n}\beta_i^0\theta_i\left(\sum_{s=1}^{T}\left(1-\frac{s}{T}\right)e_{is}\right)\right\|=o_p(1)$.

Lemma 10 *Suppose that Assumptions 1 – 9 hold. Then, the following hold.*

(a) $\frac{1}{T}\left\|Z_{-1}-Z_{-1}(0)\right\|=O_p(1)$.

(b) $\frac{1}{\sqrt{nT}}tr\left(\left(Z_{-1}-Z_{-1}(0)\right)y'\right)=O_p(1)$.

(c) $\frac{1}{\sqrt{nT}}tr\left(\left(Z_{-1}-Z_{-1}(0)\right)Q_{\beta_0}e'\right)=o_p(1)$.

(d) $\frac{1}{n\sqrt{T}}\left\|\beta^{0\prime}\left(Z_T-Z_T(0)\right)\right\|$,

where $Z_T=(z_{1T},...,z_{nT})'$ *and* $Z_T(0)=(z_{1T}(0),...,z_{nT}(0))'$.

Let β_r denote an $(n\times r)$ matrix, $r\leq K$. Define

$$\begin{aligned}\mathcal{H}_{1nT}(\beta_r) &= tr\left(\frac{\beta_r'}{\sqrt{n}}\left(\frac{\hat{y}'\hat{y}}{nT}\right)\frac{\beta_r}{\sqrt{n}}\right),\\ \mathcal{H}_{2nT}(\beta_r) &= tr\left(\frac{\beta_r'}{\sqrt{n}}\left(\frac{y'y}{nT}\right)\frac{\beta_r}{\sqrt{n}}\right),\end{aligned}$$

and

$$\mathcal{H}_{3nT}(\beta_r)=tr\left(\frac{\beta_r'}{\sqrt{n}}\left(\frac{\beta^0f^{0\prime}f^0\beta^{0\prime}}{nT}\right)\frac{\beta_r}{\sqrt{n}}\right).$$

The following lemma establishes the uniform convergence of the three functions.

Lemma 11 *Suppose that Assumptions 1 – 9 hold. As* $(n,T\to\infty)$ *with* $\frac{n}{T}\to 0$,

(a) $\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{1nT}(\beta_r)-\mathcal{H}_{2nT}(\beta_r)\right|=o_p(1)$

(b) $\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{2nT}(\beta_r)-\mathcal{H}_{3nT}(\beta_r)\right|=o_p(1)$.

The relationship among the various estimators for β^0 and f^0 are well known. Let $\check{\Lambda}_{nT,K}$ denote the diagonal matrix of the K largest eigenvalues of $\hat{y}\hat{y}'$. Then, by definition,

$$\hat{y}\hat{y}'\frac{\check{f}_K}{\sqrt{T}}=\frac{\check{f}_K}{\sqrt{T}}\check{\Lambda}_{nT,K},$$

and so

$$\hat{y}'\hat{y}\left(\hat{y}'\frac{\check{f}_K}{\sqrt{T}}\check{\Lambda}_{nT,K}^{-1/2}\right)=\left(\hat{y}'\frac{\check{f}_K}{\sqrt{T}}\check{\Lambda}_{nT,K}^{-1/2}\right)\check{\Lambda}_{nT,K}.$$

Since $tr\left(\check{\Lambda}_{nT,K}^{-1/2}\frac{\check{f}_K'}{\sqrt{T}}\hat{y}\hat{y}'\frac{\check{f}_K}{\sqrt{T}}\check{\Lambda}_{nT,K}^{-1/2}\right)=I_K$, we have

$$\bar{\beta}_K=\frac{\sqrt{n}}{\sqrt{T}}\hat{y}'\check{f}_K\check{\Lambda}_{nT,K}^{-1/2},$$

and in consequence,

$$\bar{f}_K=\frac{1}{n}\hat{y}\bar{\beta}_K=\check{f}_K\left(\frac{\check{\Lambda}_{nT,K}}{nT}\right)^{1/2}.$$

Also, using the definition of $\check{\beta}_K=\frac{1}{T}\hat{y}'\check{f}_K$ and the relations above, we deduce that

$$\hat{\beta}_K=\check{\beta}_K\left(\frac{\check{\beta}_K'\check{\beta}_K}{n}\right)^{1/2}=\frac{1}{T}\hat{y}'\check{f}_K\left(\frac{\check{\Lambda}_{nT,K}}{nT}\right)^{1/2}=\frac{1}{T}\hat{y}'\bar{f}_K=\frac{\hat{y}'\hat{y}}{nT}\bar{\beta}_K.$$

This relation between $\hat{\beta}_K$ and $\bar{\beta}_K$ will be used a lot in the proofs of the appendix.

Recall that $\bar{\beta}_K$ is $\sqrt{n}$ times the $(n\times K)$ matrix of the orthonormal eigenvectors of the first K largest eigenvalues of $\frac{\hat{y}'\hat{y}}{nT}$. Let $\Lambda_{nT,K}$ be the $(K\times K)$ diagonal matrix consisting of the first K largest eigenvalues of $\frac{\hat{y}'\hat{y}}{nT}$ (and also of $\hat{y}'\hat{y}$), *i.e.*,

$$\frac{\hat{y}'\hat{y}}{nT}\bar{\beta}_K=\bar{\beta}_K\Lambda_{nT,K}.$$

Define Λ_K to the $(K\times K)$ diagonal matrix consisting of the eigenvalues of $\Sigma_f\Sigma_\beta$. The following lemma shows that the limit of $\Lambda_{nT,K}$ is Λ_K. This lemma corresponds to Lemma A.3 of Bai (2001), which was also implicitly proved by Stock and Watson (1998). The main difference between the two lemmas is that Bai analyzes the relationship between two estimators of the factors f_t^0 using the observable data, while the following lemma characterizes the relationship between two estimators of the factor loadings β_i^0 using the residuals.

Lemma 12 *As $(n,T\to\infty)$ with $\frac{n}{T}\to 0$, under Assumptions 1 – 9, the following hold.*

(a) $\frac{1}{n}\bar{\beta}_K'\frac{\hat{y}'\hat{y}}{nT}\bar{\beta}_K=\Lambda_{nT,K}\to_p\Lambda_K$.

(b) $\left(\frac{\bar{\beta}_K'\beta^0}{n}\right)\left(\frac{f^{0\prime}f^0}{n}\right)\left(\frac{\beta^{0\prime}\bar{\beta}_K}{n}\right)\to_p\Lambda_K$.

Lemma 13 *Suppose that Assumptions 1 – 9 hold. Assume that $(n,T\to\infty)$ with $\frac{n}{T}\to 0$. Then, the limit of*

$$H_K=\left(\frac{f^{0\prime}f^0}{T}\right)\left(\frac{\beta^{0\prime}\bar{\beta}_K}{n}\right)$$

is of full rank, and H_K is asymptotically bounded.

Lemma 14 *Suppose that Assumptions 1 – 9 hold.*

(a) Suppose that $(n,T\to\infty)$. Then,

$$\left\|\frac{\hat{\beta}_K-\beta_K^*}{\sqrt{n}}\right\|=O_p\left(\max\left(\frac{1}{\sqrt{n}},\frac{1}{\sqrt{T}}\right)\right),$$

where $\beta_K^*=\beta^0H_K$, $H_K=\left(\frac{f^{0\prime}f^0}{n}\right)\left(\frac{\beta^{0\prime}\bar{\beta}_K}{n}\right)$.

(b) Suppose that $(n,T\to\infty)$ with $\frac{n}{T}\to 0$. Then,

$$\left\|\hat{\beta}_K-\beta_K^*\right\|=O_p(1).$$

(c) Suppose that $(n, T \to \infty)$ *with* $\frac{n}{T} \to 0$. *Then,*

$$\left\| \hat{\beta}_K' \left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}} \right) \right\| = o_p(1).$$

(d) Suppose that $(n, T \to \infty)$ *with* $\frac{n}{T} \to 0$. *Then,*

$$\left\| \left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}} \right)' \beta_K^* \right\| = o_p(1).$$

Proofs of Lemmas 6, 7, 8, and 9

In this paper, we omit the proofs. The required results for Lemmas 6, 8, and 9 could be obtained by calculating relavent moments directly. The proof of Lemma 7 is similar to the proofs of Theorem 14. All the detailed proofs could be found in Moon and Perron (2002).

Proof of Lemma 10

Part (a). The required result follows by Lemma 9(a) and (b) because

$$\begin{aligned} & \frac{1}{T^2} \|Z_{-1} - Z_{-1}(0)\|^2 = \frac{1}{T^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(z_{it-1}^0 - z_{it-1}^0(0) \right)^2 \\ = & \frac{1}{T^2} \sum_{i=1}^{n} \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \left(1 - \rho_i^{t-s-1}\right) y_{is} \right)^2 \sim \frac{1}{nT^2} \sum_{i=1}^{n} \theta_i^2 \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \frac{(t-s-1)}{T} y_{is} \right)^2 \\ \leq & \frac{2}{nT^2} \sum_{i=1}^{n} \theta_i^2 \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \frac{(t-s-1)}{T} \beta_i^{0\prime} f_s^0 \right)^2 + \frac{2}{nT^2} \sum_{i=1}^{n} \theta_i^2 \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \frac{(t-s-1)}{T} e_{is} \right)^2. \blacksquare \end{aligned}$$

Part (b). By Lemma 9(c), (d), and (e),

$$\begin{aligned} & \frac{1}{\sqrt{n}T} tr\left(\left(Z_{-1} - Z_{-1}(0) \right) y' \right) \\ = & \frac{1}{\sqrt{n}T} \sum_{i=1}^{n} \sum_{t=2}^{T} \left(z_{it-1}^0 - z_{it-1}^0(0) \right) y_{it} \sim -\frac{1}{nT} \sum_{i=1}^{n} \theta_i \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) y_{is} \right) y_{it} = O_p(1), \end{aligned}$$

as required. ■

Part (c). Write

$$\begin{aligned} & \frac{1}{\sqrt{n}T} tr\left(\left(Z_{-1}^0 - Z_{-1}^0(0) \right) Q_{\beta^0} e' \right) \\ = & \frac{1}{\sqrt{n}T} tr\left(\left(Z_{-1}^0 - Z_{-1}^0(0) \right) e' \right) - \frac{1}{\sqrt{n}T} tr\left(\left(Z_{-1}^0 - Z_{-1}^0(0) \right) \beta^0 \left(\beta^{0\prime} \beta^0 \right)^{-1} \beta^{0\prime} e' \right) \\ = & I_c - II_c, \text{ say.} \end{aligned}$$

By definition,

$$\begin{aligned} I_c = & \frac{1}{\sqrt{n}T} \sum_{i=1}^{n} \sum_{t=2}^{T} \left(z_{it-1}^0 - z_{it-1}^0(0) \right) e_{it} \sim \frac{1}{nT} \sum_{i=1}^{n} \theta_i \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) y_{is} \right) e_{it} \\ = & \frac{1}{nT} \sum_{i=1}^{n} \theta_i \beta_i^{0\prime} \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) f_s \right) e_{it} + \frac{1}{nT} \sum_{i=1}^{n} \theta_i \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) e_{is} \right) e_{it} \\ = & o_p(1), \end{aligned}$$

where the last equality holds by Lemma 9(d) and (e). So we have

$$I_c = o_p(1).$$

Also,

$$\begin{aligned} II_c &= \frac{1}{\sqrt{n}T} tr\left(\left(Z_{-1}^0 - Z_{-1}^0(0)\right)\beta^0\left(\beta^{0\prime}\beta^0\right)^{-1}\beta^{0\prime}e'\right) \\ &= \frac{1}{n\sqrt{n}T} tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\beta^{0\prime}e'\left(Z_{-1}^0 - Z_{-1}^0(0)\right)\beta^0\right) \\ &\leq \left\|\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\right\| \frac{1}{n\sqrt{n}T}\left\|\beta^{0\prime}e'\left(Z_{-1}^0 - Z_{-1}^0(0)\right)\beta^0\right\|. \end{aligned}$$

Since $\left\|\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\right\| = O_p(1)$, the required result follows if we show that $\frac{1}{n\sqrt{n}T}\left\|\beta^{0\prime}e'\left(Z_{-1}^0 - Z_{-1}^0(0)\right)\beta^0\right\| = o_p(1)$, which follows because

$$\begin{aligned} &\frac{1}{n\sqrt{n}T}\left\|\beta^{0\prime}e'\left(Z_{-1}^0 - Z_{-1}^0(0)\right)\beta^0\right\| \\ = &\frac{1}{n\sqrt{n}T}\left\|\sum_{i=1}^n \beta_i^0 \sum_{j=1}^n\sum_{t=2}^T e_{it}\left(z_{jt-1}^0 - z_{jt-1}^0(0)\right)\beta_j^{0\prime}\right\| \sim \frac{1}{n^2T^2}\left\|\sum_{i=1}^n \beta_i^0 \sum_{j=1}^n \theta_j \sum_{t=2}^T e_{it}\left(\sum_{s=1}^{t-1}(t-s-1)\,y_{js}\right)\beta_j^{0\prime}\right\| \\ \leq &\frac{1}{n^2T^2}\left(\begin{array}{l} \left\|\sum_{i=1}^n \beta_i^0 \sum_{j=1}^n \beta_j^{0\prime}\beta_j^0\theta_j \sum_{t=2}^T e_{it}\left(\sum_{s=1}^{t-1}(t-s-1)\,f_s\right)\right\| \\ +\left\|\sum_{i=1}^n \beta_i^0 \sum_{j=1}^n \beta_j^{0\prime}\theta_j \sum_{t=2}^T e_{it}\left(\sum_{s=1}^{t-1}(t-s-1)\,e_{js}\right)\right\| \end{array}\right) = o_p(1) \end{aligned}$$

by Lemma 9(f) and (g). ■

Part (d). The required result follows because

$$\begin{aligned} &\frac{1}{\sqrt{nT}}\left\|\beta^{0\prime}\left(Z_T - Z_T(0)\right)\right\| \\ = &\frac{1}{\sqrt{nT}}\left\|\sum_{i=1}^n \beta_i^0\left(z_{iT}^0 - z_{iT}^0(0)\right)\right\| = \frac{1}{\sqrt{nT}}\left\|\sum_{i=1}^n \beta_i^0\left(\sum_{s=1}^T\left(1-\rho_i^{T-s-1}\right)y_{is}\right)\right\| \\ \sim &\frac{1}{n\sqrt{T}}\left\|\sum_{i=1}^n \beta_i^0\theta_i\left(\sum_{s=1}^T\left(1-\frac{s-1}{T}\right)y_{is}\right)\right\| \\ \leq &\frac{1}{n\sqrt{T}}\left\|\sum_{i=1}^n \theta_i\beta_i^0\beta_i^{0\prime}\left(\sum_{s=1}^T\left(1-\frac{s-1}{T}\right)f_s^0\right)\right\| + \frac{1}{n\sqrt{T}}\left\|\sum_{i=1}^n \beta_i^0\theta_i\left(\sum_{s=1}^T\left(1-\frac{s-1}{T}\right)e_{is}\right)\right\| = O_p(1) \end{aligned}$$

by Lemma 9(g) and (i). ■

Proof of Lemma 11

Part (a).

By definition,

$$\hat{y} = Z - \hat{\rho}_{pool}Z_{-1} = y + l_T\alpha'\left(1-\hat{\rho}_{pool}\right) + Z_{-1}^0\left(\rho - \hat{\rho}_{pool}I_n\right).$$

Then, we may write

$$\frac{\hat{y}'\hat{y}}{nT} - \frac{y'y}{nT} = \sum_{k=1}^7 \mathcal{R}_k,$$

where

$$\begin{aligned}
\mathcal{R}_1 &= T^2\left(1-\hat{\rho}_{pool}\right)^2\frac{\alpha\alpha'}{nT^2},\ \mathcal{R}_2 = T\left(1-\hat{\rho}_{pool}\right)\left(\frac{y'l_T\alpha'+\alpha l_T'y}{nT^2}\right)\\
\mathcal{R}_3 &= \frac{(\rho-I_n)Z_{-1}^{0\prime}y+y'Z_{-1}^0(\rho-I_n)}{nT},\ \mathcal{R}_4 = T\left(1-\hat{\rho}_{pool}\right)\left(\frac{Z_{-1}^{0\prime}y+y'Z_{-1}^0}{nT^2}\right)\\
\mathcal{R}_5 &= T\left(1-\hat{\rho}_{pool}\right)\left(\frac{\alpha l_T'Z_{-1}^0(\rho-I_n)+(\rho-I_n)Z_{-1}^{0\prime}l_T\alpha'}{nT^2}\right)\\
\mathcal{R}_6 &= T^2\left(1-\hat{\rho}_{pool}\right)^2\left(\frac{Z_{-1}^{0\prime}l_T\alpha'+\alpha l_T'Z_{-1}^0}{nT^3}\right),\ \mathcal{R}_7 = \frac{\left(\rho-\hat{\rho}_{pool}I_n\right)Z_{-1}^{0\prime}Z_{-1}^0\left(\rho-\hat{\rho}_{pool}I_n\right)}{nT}.
\end{aligned}$$

By the triangle inequality,

$$\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{1nT}(\beta_r)-\mathcal{H}_{2nT}(\beta_r)\right| \le \sum_{k=1}^{7}\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|tr\left(\frac{\beta_r'}{\sqrt{n}}\mathcal{R}_k\frac{\beta_r}{\sqrt{n}}\right)\right| \le r\sum_{k=1}^{7}\left\|\mathcal{R}_k\right\|.$$

Notice that $T\left(1-\hat{\rho}_{pool}\right)=O_p(1)$ by Lemma 1, $\frac{\|\alpha\|^2}{n}=O_p(1)$ by Assumption 9, and $T\left\|\rho-I_n\right\|=\sqrt{\frac{1}{n}\sum_{i=1}^n\theta_i^2}=O_p(1)$. Using the results in Lemmas 8 and 9, we can show that

$$\begin{aligned}
\left\|\mathcal{R}_1\right\| &= \frac{1}{T^2}T^2\left(1-\hat{\rho}_{pool}\right)^2\frac{\|\alpha\|^2}{n}=O_p\left(\frac{1}{T^2}\right),\\
\left\|\mathcal{R}_2\right\| &\le \frac{1}{T\sqrt{T}}\left|T\left(1-\hat{\rho}_{pool}\right)\right|\frac{\|y'l_T\|}{\sqrt{nT}}\frac{\|\alpha\|}{\sqrt{n}}=O_p\left(\frac{1}{T\sqrt{T}}\right),\\
\left\|\mathcal{R}_3\right\| &\le \frac{2}{\sqrt{nT}}\frac{\left\|Z_{-1}^0\Theta\right\|}{\sqrt{n}T}\frac{\|y\|}{\sqrt{nT}}=O_p\left(\frac{1}{\sqrt{nT}}\right),\\
\left\|\mathcal{R}_4\right\| &= \frac{1}{T}\left|T\left(1-\hat{\rho}_{pool}\right)\right|\frac{\left\|Z_{-1}^{0\prime}y+y'Z_{-1}^0\right\|}{nT}=O_p\left(\frac{1}{T}\right),\\
\left\|\mathcal{R}_5\right\| &\le \frac{2}{\sqrt{n}T\sqrt{T}}\left|T\left(1-\hat{\rho}_{pool}\right)\right|\frac{\|\alpha\|}{\sqrt{n}}\frac{\|l_T\|}{\sqrt{T}}\frac{\left\|Z_{-1}^0\Theta\right\|}{\sqrt{n}T}=O_p\left(\frac{1}{\sqrt{n}T\sqrt{T}}\right),\\
\left\|\mathcal{R}_6\right\| &\le \frac{2}{T\sqrt{T}}\left|T^2\left(1-\hat{\rho}_{pool}\right)^2\right|\frac{\|\alpha\|}{\sqrt{n}}\frac{\|l_T\|}{\sqrt{T}}\frac{\left\|Z_{-1}^0\right\|}{\sqrt{n}T}=O_p\left(\frac{1}{T\sqrt{T}}\right),\\
\left\|\mathcal{R}_7\right\| &\le \frac{2}{T}\left[\frac{1}{n}\frac{\left\|Z_{-1}^0\Theta\right\|^2}{nT^2}+T^2\left(1-\hat{\rho}_{pool}\right)^2\frac{\left\|Z_{-1}^0\right\|^2}{nT^2}\right]=O_p\left(\frac{1}{T}\right), \qquad (19)
\end{aligned}$$

where $\frac{\left\|Z_{-1}^0\Theta\right\|}{\sqrt{n}T}=O_p(1)$ holds by similar arguments in the proof of Lemma 8(a). Then,

$$\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{1nT}(\beta_r)-\mathcal{H}_{2nT}(\beta_r)\right| \le r\sum_{k=1}^{7}\left\|\mathcal{R}_k\right\| = O_p\left(\frac{1}{\sqrt{T}}\max\left\{\frac{1}{\sqrt{n}},\frac{1}{\sqrt{T}}\right\}\right)=o_p(1),$$

and this completes the proof of Part (a). ■

Part (b).

Write

$$\begin{aligned}
&\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{2nT}(\beta_r)-\mathcal{H}_{3nT}(\beta_r)\right|\\
=\ &\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|tr\left(\frac{\beta_r'}{\sqrt{n}}\left(\frac{\beta^0f^{0\prime}e}{nT}+\frac{ef^0\beta^{0\prime}}{nT}+\frac{e'e}{nT}\right)\frac{\beta_r}{\sqrt{n}}\right)\right| \le I+II+III,
\end{aligned}$$

where

$$\begin{aligned} I &= \sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left| tr\left(\frac{\beta_r'}{\sqrt{n}}\left(\frac{\beta^0 f^{0\prime} e}{nT}\right)\frac{\beta_r}{\sqrt{n}}\right)\right| \\ II &= \sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left| tr\left(\frac{\beta_r'}{\sqrt{n}}\left(\frac{e f^0 \beta^{0\prime}}{nT}\right)\frac{\beta_r}{\sqrt{n}}\right)\right| \\ III &= \sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left| tr\left(\frac{\beta_r'}{\sqrt{n}}\left(\frac{e'e}{nT}\right)\frac{\beta_r}{\sqrt{n}}\right)\right|. \end{aligned}$$

First,

$$I \leq \sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left\|\frac{\beta^0 f^{0\prime} e}{nT}\right\| \left\|\frac{\beta_r}{\sqrt{n}}\right\|^2 = r\left\|\frac{\beta^0 f^{0\prime} e}{nT}\right\| \leq \frac{r}{\sqrt{T}}\left\|\frac{\beta^0}{\sqrt{n}}\right\|\left\|\frac{f^{0\prime} e}{\sqrt{nT}}\right\| = O_p\left(\frac{1}{\sqrt{T}}\right)$$

since $\left\|\frac{\beta^0}{\sqrt{n}}\right\| = O(1)$ by Assumption 6, $\left\|\frac{f^{0\prime} e}{\sqrt{nT}}\right\|^2 = tr\left(\frac{f^{0\prime} e e' f^0}{nT}\right) = O_p(1)$ by Lemma 8(f). Thus,

$$I = o_p(1).$$

Similarly, we can show that

$$II = O_p\left(\frac{1}{\sqrt{T}}\right) = o_p(1).$$

Consider the last term, III. Let $\beta_r = \left(\beta_{1,r}, ..., \beta_{i,r}, ..., \beta_{n,r}\right)'$. Notice that

$$\begin{aligned} III &= \sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left| \frac{1}{n^2 T}\sum_{i=1}^n\sum_{j=1}^n\sum_{t=1}^T e_{it}e_{jt}\beta_{i,r}'\beta_{j,r}\right| \\ &\leq \sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left| \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{T}\sum_{t=1}^T\left(e_{it}e_{jt} - E\left(e_{it}e_{jt}\right)\right)\right)\beta_{i,r}'\beta_{j,r}\right| \\ &\quad + \sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left| \frac{1}{n^2}\sum_{i=1}^n \beta_{i,r}'\beta_{i,r}\left(\frac{1}{T}\sum_{t=1}^T E\left(e_{it}^2\right)\right)\right| \\ &= III_a + III_b, \text{ say.} \end{aligned}$$

Notice by Lemma 7(a) that $\sup_i \frac{1}{T}\sum_{t=1}^T E\left(e_{it}^2\right) < M$. Then, since

$$\begin{aligned} &\sup_{\frac{\beta_r'\beta_r}{n}=I_r} \left| \frac{1}{n^2}\sum_{i=1}^n \beta_{i,r}'\beta_{i,r}\left(\frac{1}{T}\sum_{t=1}^T E\left(e_{it}^2\right)\right)\right| \\ \leq\ & \frac{1}{n}\sup_i\left(\frac{1}{T}\sum_{t=1}^T E\left(e_{it}^2\right)\right)\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\frac{1}{n}\sum_{i=1}^n \beta_{i,r}'\beta_{i,r}\right| \leq \frac{rM}{n}, \end{aligned}$$

we have

$$III_b = O\left(\frac{1}{n}\right) = o(1).$$

Next, we consider III_a. Recalling the definition of

$$X_{ij,T} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(e_{it} e_{jt} - E\left(e_{it} e_{jt} \right) \right),$$

we write

$$\begin{aligned} III_a &= \sup_{\frac{\beta_r' \beta_r}{n} = I_r} \left| \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{X_{ij,T}}{\sqrt{T}} \beta_{i,r}' \beta_{j,r} \right| \leq \sup_{i,j} \left| \frac{X_{ij,T}}{\sqrt{T}} \right| \sup_{\frac{\beta_r' \beta_r}{n} = I_r} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| \beta_{i,r}' \beta_{j,r} \right| \\ &\leq \sup_{\frac{\beta_r' \beta_r}{n} = I_r} \left(\frac{1}{n} \sum_{i=1}^{n} \left\| \beta_{i,r} \right\| \right)^2 \sup_{i,j} \left| \frac{X_{ij,T}}{\sqrt{T}} \right| \leq r \left(\sup_{i,j} \left| \frac{X_{ij,T}}{\sqrt{T}} \right| \right). \end{aligned}$$

Notice by Lemma 6,

$$\sup_{i,j} E X_{ij,T}^4 < M.$$

So, for any $\varepsilon > 0$,

$$P\left\{ \sup_{i,j} \left| \frac{X_{ij,T}}{\sqrt{T}} \right| > \varepsilon \right\} \leq \frac{E\left(\sup_{i,j} \left| \frac{X_{ij,T}}{\sqrt{T}} \right| \right)^4}{\varepsilon^4} \leq \frac{\frac{1}{T^2} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(X_{ij,T}^4 \right)}{\varepsilon^4} \leq \frac{n^2}{T^2} \frac{\sup_{i,j} E X_{ij,T}^4}{\varepsilon^4} \to 0$$

as $(n, T \to \infty)$ with $\frac{n}{T} \to 0$. Thus,

$$\sup_{i,j} \left| \frac{X_{ij,T}}{\sqrt{T}} \right| = O_p\left(\sqrt{\frac{n}{T}} \right),$$

from which we deduce that as $(n, T \to \infty)$ with $\frac{n}{T} \to 0$,

$$III_a = O_p\left(\sqrt{\frac{n}{T}} \right) = o_p\left(1 \right).$$

In view of III_a and III_b, we have

$$III = O_p\left(\max\left(\frac{1}{n}, \sqrt{\frac{n}{T}} \right) \right) = o_p\left(1 \right), \tag{20}$$

as required for Part (b). ■

Proof of Lemma 12

Let $\bar{\beta}_r$ be $\sqrt{n}$ times the orthonormal eigenvectors corresponding to the first r $(\leq K)$ largest eigenvalues of $\frac{\hat{y}'\hat{y}}{nT}$. Similarly, let $\tilde{\beta}_r$ be $\sqrt{n}$ times the orthonormal eigenvectors of the first r $(\leq K)$ largest eigenvalues of $\frac{\beta^0 f^{0\prime} f^0 \beta^{0'}}{nT}$. Then, $\bar{\beta}_r$ and $\tilde{\beta}_r$ maximize $\mathcal{H}_{1nT}\left(\beta_r \right)$ and $\mathcal{H}_{3nT}\left(\beta_r \right)$, respectively, among the $\beta_r' s$ satisfying $\frac{\beta_r' \beta_r}{n} = I_r$. Define

$$\Lambda_{nT,r} = \frac{\bar{\beta}_r'}{\sqrt{n}} \left(\frac{\hat{y}'\hat{y}}{nT} \right) \frac{\bar{\beta}_r}{\sqrt{n}}$$

$$\tilde{\Lambda}_{nT,r} = \frac{\tilde{\beta}_r'}{\sqrt{n}} \left(\frac{\beta^0 f^{0\prime} f^0 \beta^{0'}}{nT} \right) \frac{\tilde{\beta}_r}{\sqrt{n}}. \tag{21}$$

We first show that $\left\|\tilde{\Lambda}_{nT,K}-\Lambda_K\right\|\rightarrow_p 0$. Let

$$\tilde{\beta}_K^*=\left(\frac{f^{0\prime}f^0}{T}\right)\left(\frac{\beta^{0\prime}\tilde{\beta}_K}{n}\right) \text{ and } \tilde{\beta}_K^{**}=\frac{\tilde{\beta}_K^*}{\left\|\tilde{\beta}_K^*\right\|}.$$

Then,

$$\tilde{\Lambda}_{nT,K}=\tilde{\beta}_K^{**\prime}\left(\frac{f^{0\prime}f^0}{T}\right)\left(\frac{\beta^{0'}\beta^0}{n}\right)\tilde{\beta}_K^{**}.$$

Notice that $\tilde{\Lambda}_{nT,K}$ is the $(K\times K)$ diagonal matrix consisting of the first K eigenvalues of $\frac{\beta^0 f^{0\prime}f^0\beta^{0'}}{nT}$, and it is also the $(K\times K)$ diagonal matrix consisting of the eigenvalues of the $(K\times K)$ matrix $\left(\frac{f^{0\prime}f^0}{T}\right)\left(\frac{\beta^{0'}\beta^0}{n}\right)$.

Now since $\left(\frac{f^{0\prime}f^0}{T}\right)\left(\frac{\beta^{0'}\beta^0}{n}\right)\rightarrow_p \Sigma_f\Sigma_\beta$, a full rank matrix, by Assumptions 7 and 6, we have

$$\left\|\tilde{\Lambda}_{nT,K}-\Lambda_K\right\|\rightarrow_p 0 \tag{22}$$

as $(n,T\rightarrow\infty)$ (for further details on this, refer to Anderson, 1963).

Next, we establish that $\left\|\Lambda_{nT,K}-\tilde{\Lambda}_{nT,K}\right\|\rightarrow_p 0$. By Lemma 11, as $(n,T\rightarrow\infty)$ with $\frac{n}{T}\rightarrow 0$,

$$\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{1nT}\left(\beta_r\right)-\mathcal{H}_{2nT}\left(\beta_r\right)\right|=o_p\left(1\right)$$

and

$$\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{2nT}\left(\beta_r\right)-\mathcal{H}_{3nT}\left(\beta_r\right)\right|=o_p\left(1\right)$$

for $1\leq r\leq K$.

Then, by the triangle inequality, it follows that

$$\sup_{\frac{\beta_r'\beta_r}{n}=I_r}\left|\mathcal{H}_{1nT}\left(\beta_r\right)-\mathcal{H}_{3nT}\left(\beta_r\right)\right|=o_p\left(1\right) \tag{23}$$

for $1\leq r\leq K$. Since $\bar{\beta}_r$ and $\tilde{\beta}_r$ are maximizers of $\mathcal{H}_{1nT}\left(\beta_r\right)$ and $\mathcal{H}_{3nT}\left(\beta_r\right)$, respectively, it follows that

$$\left|\mathcal{H}_{1nT}\left(\bar{\beta}_r\right)-\mathcal{H}_{3nT}\left(\tilde{\beta}_r\right)\right|=o_p\left(1\right) \tag{24}$$

for $1\leq r\leq K$. Also, by (23) and (24), we have

$$\left|\mathcal{H}_{3nT}\left(\tilde{\beta}_r\right)-\mathcal{H}_{3nT}\left(\bar{\beta}_r\right)\right|\leq\left|\mathcal{H}_{1nT}\left(\bar{\beta}_r\right)-\mathcal{H}_{3nT}\left(\bar{\beta}_r\right)\right|+\left|\mathcal{H}_{1nT}\left(\bar{\beta}_r\right)-\mathcal{H}_{3nT}\left(\tilde{\beta}_r\right)\right|=o_p\left(1\right) \tag{25}$$

for $1\leq r\leq K$.

>From (24), it follows that

$$\begin{aligned}\left|tr\left(\Lambda_{nT,r}\right)-tr\left(\tilde{\Lambda}_{nT,r}\right)\right| &= \left|tr\left(\frac{\bar{\beta}_r'}{\sqrt{n}}\left(\frac{\hat{y}'\hat{y}}{nT}\right)\frac{\bar{\beta}_r}{\sqrt{n}}\right)-tr\left(\frac{\tilde{\beta}_r'}{\sqrt{n}}\left(\frac{\beta^0f^{0\prime}f^0\beta^{0'}}{nT}\right)\frac{\tilde{\beta}_r}{\sqrt{n}}\right)\right| \\ &= \left|\mathcal{H}_{1nT}\left(\bar{\beta}_r\right)-\mathcal{H}_{3nT}\left(\tilde{\beta}_r\right)\right|\rightarrow_p 0,\end{aligned} \tag{26}$$

for $1\leq r\leq K$. Since the sequences $\{\Lambda_{nT,r}\}_r$ and $\left\{\tilde{\Lambda}_{nT,r}\right\}_r$ are nested, (26) implies that

$$\left\|\Lambda_{nT,K}-\tilde{\Lambda}_{nT,K}\right\|\rightarrow_p 0. \tag{27}$$

Then, Part (a) follows by the triangle inequality and (22) and (27), so

$$\|\Lambda_{nT,K} - \Lambda_K\| \leq \left\|\Lambda_{nT,K} - \tilde{\Lambda}_{nT,K}\right\| + \left\|\tilde{\Lambda}_{nT,K} - \Lambda_K\right\| \to_p 0.$$

Similarly, Part (b) follows by (25) and (22). ■

Proof of Lemma 13

>From Lemma 12(b), as $(n, T \to \infty)$ with $\frac{n}{T} \to 0$, we have

$$\left(\frac{\bar{\beta}_K' \beta^0}{n}\right)\left(\frac{f^{0\prime} f^0}{T}\right)\left(\frac{\beta^{0\prime} \bar{\beta}_K}{n}\right) \to_p \Lambda_K, \tag{28}$$

where Λ_K a $(K \times K)$ full rank matrix. Since $\frac{f^{0\prime} f^0}{n} \to_p \Sigma_f$, a full rank matrix, $\frac{\bar{\beta}_K' \beta^0}{n}$ is asymptotically of full rank. Then, we have that $H_K = \left(\frac{f^{0\prime} f^0}{T}\right)\left(\frac{\beta^{0\prime} \bar{\beta}_K}{n}\right)$ is asymptotically of full rank.

Also, from (28),

$$\left\|\left(\frac{f^{0\prime} f^0}{T}\right)^{1/2}\left(\frac{\beta^{0\prime} \bar{\beta}_K}{n}\right)\right\|^2 = tr\left(\left(\frac{\bar{\beta}_K' \beta^0}{n}\right)\left(\frac{f^{0\prime} f^0}{T}\right)\left(\frac{\beta^{0\prime} \bar{\beta}_K}{n}\right)\right) = O_p(1).$$

Therefore,

$$\begin{aligned} H_K &= \left(\frac{f^{0\prime} f^0}{T}\right)^{1/2}\left(\frac{f^{0\prime} f^0}{T}\right)^{1/2}\left(\frac{\beta^{0\prime} \bar{\beta}_K}{n}\right) \\ &= O_p(1) O_p(1) = O_p(1), \end{aligned}$$

as required. ■

Proof of Lemma 14.

Parts (a) and (b).

By the definitions of $\hat{\beta}_K$ and $\bar{\beta}_K$,

$$\hat{\beta}_K = \frac{\hat{y}'\hat{y}}{nT}\bar{\beta}_K, \tag{29}$$

where

$$\begin{aligned} \hat{y} &= Z - \hat{\rho}_{pool} Z_{-1} \\ &= f^0 \beta^{0\prime} + e - Z_{-1}^0\left(\hat{\rho}_{pool} I_n - \rho\right) + l_T \alpha'\left(1 - \hat{\rho}_{pool}\right). \end{aligned} \tag{30}$$

>From (29) and (30) and by applying the triangle inequality, we may have

$$\begin{aligned} \left\|\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}}\right\| &\leq \left\|\frac{\beta^0 f^{0\prime} e \bar{\beta}_K}{\sqrt{n}nT}\right\| + \left\|\frac{e' f^0 \beta^{0\prime} \bar{\beta}_K}{\sqrt{n}nT}\right\| + \left\|\frac{e' e \bar{\beta}_K}{\sqrt{n}nT}\right\| + \sum_{k=1}^{7} \|\mathcal{R}_k\| \left\|\frac{\bar{\beta}_K}{\sqrt{n}}\right\| \qquad (31) \\ &= \frac{2}{\sqrt{T}}\left\|\frac{\beta^0}{\sqrt{n}}\right\|\left\|\frac{e' f^0}{\sqrt{nT}}\right\|\left\|\frac{\bar{\beta}_K}{\sqrt{n}}\right\| + \left\|\frac{e' e}{nT}\right\|\left\|\frac{\bar{\beta}_K}{\sqrt{n}}\right\| + \sum_{k=1}^{7} \|\mathcal{R}_k\| \left\|\frac{\bar{\beta}_K}{\sqrt{n}}\right\| \\ &= \frac{1}{\sqrt{T}} O(1) O_p(1) O_p(1) + O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) O_p(1) + O_p\left(\frac{1}{\sqrt{T}} \max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) O_p(1) \\ &= O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right). \qquad (32) \end{aligned}$$

where the last equality holds by Lemma 8(e) and (f), the results in (19) together with $\left\|\frac{\bar{\beta}_K}{\sqrt{n}}\right\| = \sqrt{K}$ (by the definition of $\bar{\beta}_K$) and $\left\|\frac{\beta^0}{\sqrt{n}}\right\| = O(1)$ (by Assumption 6). From this we deduce both Parts (a) and (b). ■

Part (c).

Similar to (31), we may write

$$\begin{aligned}
&\left\|\hat{\beta}_K'\left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}}\right)\right\| \\
\leq\ & \sqrt{n}\left\|\frac{\hat{\beta}_K}{\sqrt{n}}\right\|\left\|\frac{\beta^0 f^{0\prime} e\bar{\beta}_K}{\sqrt{n}nT}\right\| + \sqrt{n}\left\|\frac{\hat{\beta}_K}{\sqrt{n}}\right\|\left\|\frac{e' f^0 \beta^{0\prime}\bar{\beta}_K}{\sqrt{n}nT}\right\| + \left\|\frac{\hat{\beta}_K' e' e\bar{\beta}_K}{\sqrt{n}nT}\right\| + \left\|\frac{\hat{\beta}_K}{\sqrt{n}}\right\|\sum_{k=1}^{7}\sqrt{n}\left\|\mathcal{R}_k\right\|\left\|\frac{\bar{\beta}_K}{\sqrt{n}}\right\|.
\end{aligned}$$

>From (32) and $\left\|\frac{\hat{\beta}_K}{\sqrt{n}}\right\| = O_p(1)$, we have

$$\left\|\hat{\beta}_K'\left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}}\right)\right\| \leq \frac{\sqrt{n}}{\sqrt{T}}O_p(1)O(1)O_p(1) + \left\|\frac{\hat{\beta}_K' e' e\bar{\beta}_K}{\sqrt{n}nT}\right\| + \sqrt{n}O_p\left(\frac{1}{\sqrt{T}}\max\left\{\frac{1}{\sqrt{n}},\frac{1}{\sqrt{T}}\right\}\right).$$

Thus, as $(n,T\to\infty)$ with $\frac{n}{T}\to 0$, we have

$$\left\|\hat{\beta}_K'\left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}}\right)\right\| = \left\|\frac{\hat{\beta}_K' e' e\bar{\beta}_K}{\sqrt{n}nT}\right\| + o_p(1). \tag{33}$$

Since $\hat{\beta}_K = \frac{\hat{y}'\hat{y}}{nT}\bar{\beta}_K = \bar{\beta}_K\Lambda_{nT,K}$, we have

$$\left\|\frac{\hat{\beta}_K' e' e\bar{\beta}_K}{\sqrt{n}nT}\right\| = \left\|\frac{\hat{\beta}_K' e' e\hat{\beta}_K\Lambda_{nT}^{-1}}{\sqrt{n}nT}\right\| \leq \frac{\left\|e\hat{\beta}_K\right\|^2}{\sqrt{n}nT}\left\|\Lambda_{nT,K}^{-1}\right\| = \frac{\left\|e\hat{\beta}_K\right\|^2}{\sqrt{n}nT}O_p(1),$$

where the last equality holds by Lemma 12. Notice that

$$\begin{aligned}
\left\|e\hat{\beta}_K\right\|^2 &= \left\|e\left(\hat{\beta}_K - \beta_K^*\right) + e\beta_K^*\right\|^2 \leq 2\left\|e\left(\hat{\beta}_K - \beta_K^*\right)\right\|^2 + 2\left\|e\beta_K^*\right\|^2 \\
&\leq 2\left\|e\right\|^2\left\|\hat{\beta}_K - \beta_K^*\right\|^2 + 2\left\|e\beta^0\right\|^2\left\|H_K\right\|^2.
\end{aligned}$$

So,

$$\begin{aligned}
\frac{\left\|e\hat{\beta}_K\right\|^2}{\sqrt{n}nT} &\leq \frac{2}{\sqrt{n}}\frac{\left\|e\right\|^2}{nT}\left\|\hat{\beta}_K - \beta_K^*\right\|^2 + \frac{2}{\sqrt{n}}\frac{\left\|e\beta^0\right\|^2}{nT}\left\|H_K\right\|^2 \\
&= \frac{1}{\sqrt{n}}O_p(1)O_p(1) + \frac{1}{\sqrt{n}}O_p(1)O_p(1),
\end{aligned} \tag{34}$$

where the last equality holds since $\sup_{it}E(e_{it}^2) < M$, by Lemma 14(b), Lemma 8(g), and Lemma 13. So we have

$$\left\|\frac{\hat{\beta}_K' e' e\bar{\beta}_K}{\sqrt{n}nT}\right\| = o_p(1). \tag{35}$$

Thus, in view of (33) and (35), we have the required result,

$$\left\|\hat{\beta}_K'\left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}}\right)\right\| = o_p(1)$$

as $(n, T \to \infty)$ with $\frac{n}{T} \to 0$. ■

Part (d).

Similarly to the proof of Part (c), the required result follows if we show that

$$\left\| \frac{\beta_K^{*\prime} e^\prime e \bar{\beta}_K}{\sqrt{n} nT} \right\| = o_p(1). \tag{36}$$

Notice that

$$\begin{aligned}
\left\| \frac{\beta_K^{*\prime} e^\prime e \bar{\beta}_K}{\sqrt{n} nT} \right\| &\leq \left\| \frac{\beta_K^{*\prime} e^\prime e \hat{\beta}_K}{\sqrt{n} nT} \right\| \left\| \Lambda_{nT,K}^{-1} \right\| \\
&\leq \|H_K\|^2 \left\| \frac{\beta^{0\prime} e^\prime e \beta^0}{\sqrt{n} nT} \right\| \left\| \Lambda_{nT,K}^{-1} \right\| + \|H_K\| \left\| \frac{\beta^{0\prime} e^\prime e \left(\hat{\beta}_K - \beta_K^* \right)}{\sqrt{n} nT} \right\| \left\| \Lambda_{nT,K}^{-1} \right\| \\
&\leq \frac{1}{\sqrt{n}} \|H_K\|^2 \left\| \frac{e\beta^0}{\sqrt{n}\sqrt{T}} \right\|^2 \left\| \Lambda_{nT,K}^{-1} \right\| + \|H_K\| \left\| \frac{\beta^0}{\sqrt{n}} \right\| \left\| \frac{e^\prime e}{nT} \right\| \left\| \hat{\beta}_K - \beta_K^* \right\| \left\| \Lambda_{nT,K}^{-1} \right\| \\
&= \frac{1}{\sqrt{n}} O_p(1) O_p(1) O_p(1) + O_p(1) O(1) O_p\left(\max\left\{ \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}} \right\} \right) O_p(1) O_p(1) \\
&= o_p(1), \qquad (37)
\end{aligned}$$

where the last line holds by Lemma 13, Lemma 8(g), Lemma 8(e), Lemma 14(b), and Lemma 12. Therefore, we have the required result. ■

7 Appendix B: Proofs of the Main Results

Proof of Lemma 1.

The proof is omitted. The proof could be found in Moon and Perron (2002). ■

Proof of Lemma 2.

Part (a).

Part (a) is a special case of Part (b) with $\Theta = I_n$ ($\mu_\theta = 1$) and we omit the proof. ■

Part (b).

By definition,

$$\begin{aligned}\frac{1}{nT^2}tr\left(Z_{-1}Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right) &= \frac{1}{nT^2}tr\left(l_T\alpha' Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)+\frac{1}{nT^2}tr\left(Z_{-1}^0 Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)\\ &= \frac{1}{nT^2}tr\left(Z_{-1}^0 Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)+\frac{1}{\sqrt{T}}\frac{\|l_T\|}{\sqrt{T}}\frac{\|\alpha\|}{\sqrt{n}}\frac{\left\|\Theta Z_{-1}^0\right\|}{\sqrt{n}T}\\ &= \frac{1}{nT^2}tr\left(Z_{-1}^0 Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)+o_p(1).\end{aligned}$$

Also, by Lemma 8(a) and Lemma 9(a)[6], we have

$$\begin{aligned}&\frac{1}{nT^2}\left|tr\left(Z_{-1}^0 Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)-tr\left(Z_{-1}^0(0)\,Q_{\beta^0}\Theta Z_{-1}^0(0)'\right)\right|\\ =\;&\frac{1}{nT^2}\left|tr\left(\left(Z_{-1}^0-Z_{-1}^0(0)\right)Q_{\beta^0}\Theta Z_{-1}^0(0)'+Z_{-1}^0 Q_{\beta^0}\Theta\left(Z_{-1}^0-Z_{-1}^0(0)\right)'\right)\right|\\ \le\;&\frac{1}{\sqrt{n}}\frac{\left\|Z_{-1}^0-Z_{-1}^0(0)\right\|}{T}\left(1+\left(\frac{\left\|\beta^0\right\|}{\sqrt{n}}\left\|\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\right\|\frac{\left\|\beta^{0\prime}\Theta\right\|}{\sqrt{n}}\right)\right)\frac{\left(\left\|Z_{-1}^0(0)\right\|+\left\|Z_{-1}^0\right\|\right)}{\sqrt{n}T}\\ =\;&O_p\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

So

$$\frac{1}{nT^2}tr\left(Z_{-1}Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)=\frac{1}{nT^2}tr\left(Z_{-1}^0(0)\,Q_{\beta^0}\Theta Z_{-1}^0(0)'\right)+o_p(1).$$

Since $Z_{-1}^0(0)=F_{-1}^0\beta^{0\prime}+E_{-1}$,

$$\begin{aligned}\frac{1}{nT^2}tr\left(Z_{-1}^0(0)\,Q_{\beta^0}\Theta Z_{-1}^0(0)'\right) &= \frac{1}{nT^2}tr\left(E_{-1}Q_{\beta^0}\Theta\beta^0F_{-1}'\right)+\frac{1}{nT^2}tr\left(E_{-1}Q_{\beta^0}\Theta E_{-1}'\right)\\ &= I_b+II_b,\ \text{say.}\end{aligned}$$

First,

$$I_b=\frac{1}{nT^2}tr\left(E_{-1}\Theta\beta^0F_{-1}'\right)-\frac{1}{nT^2}tr\left(E_{-1}\beta^0\left(\beta^{0\prime}\beta^0\right)^{-1}\beta^{0\prime}\Theta\beta^0F_{-1}'\right)=I_{ba}+I_{bb},\ \text{say.}$$

Notice that

$$I_{ba}=\frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{i=1}^n\theta_i\beta_i^{0\prime}\frac{1}{T^2}\sum_{t=1}^T F_{t-1}E_{i,t-1}.$$

Then, it is possible to show that

$$E\left(I_{ba}^2\right)=O\left(\frac{1}{n}\right),$$

[6] Notice that $Z_{-1}-Z_{-1}(0)=Z_{-1}^0-Z_{-1}^0(0)$.

which yields

$$I_{ba} = O_p\left(\frac{1}{\sqrt{n}}\right).$$

Also,

$$\begin{aligned} I_{bb} &\leq \left|\frac{1}{nT^2} tr\left(\left(\beta^{0\prime}\beta^0\right)^{-1}\beta^{0\prime}\Theta\beta^0 F'_{-1}E_{-1}\beta^0\right)\right| \leq \frac{1}{\sqrt{n}}\left\|\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\right\| \frac{\left\|\beta^{0\prime}\Theta\beta^0\right\|}{n}\frac{\left\|F'_{-1}E_{-1}\beta^0\right\|}{\sqrt{n}T^2} \\ &= \frac{1}{\sqrt{n}} O(1) O_p(1) O_p(1) = O_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

where $\frac{\left\|F'_{-1}E_{-1}\beta^0\right\|}{\sqrt{n}T^2} = O_p(1)$ holds by similar arguments used in I_{ba}. Combining I_{ba} and I_{bb}, we have

$$I_b = O_p\left(\frac{1}{\sqrt{n}}\right).$$

Next,

$$II_b = \frac{1}{nT^2} tr\left(E_{-1}\Theta E'_{-1}\right) - \frac{1}{n^2T^2} tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\beta^{0\prime}\Theta E'_{-1}E_{-1}\beta^0\right) = II_{ba} - II_{bb}, \text{ say.}$$

Using similar arguments in the proof of Lemma 8(c)[7], we have

$$|II_{bb}| \leq \frac{1}{n}\left\|\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\right\| \frac{\left\|\beta^{0\prime}\Theta E'_{-1}\right\|}{\sqrt{n}T}\frac{\left\|E_{-1}\beta^0\right\|}{\sqrt{n}T} = O_p\left(\frac{1}{n}\right).$$

For II_{ba}, by Lemma 7(a),

$$II_{ba} = \frac{1}{nT^2} tr\left(E_{-1}\Theta E'_{-1}\right) = \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\theta_i^2 E_{it-1}^2 \to_p \frac{1}{2}\mu_\theta\omega_e^2.$$

Combining II_{ba} and II_{bb}, we have

$$II_b \to_p \frac{1}{2}\mu_\theta\omega_e^2.$$

>From I_b, and II_b, we have the required result. ■

Part (c).

Notice that

$$\begin{aligned} &\sqrt{n}\left(\frac{1}{nT} tr\left(Z_{-1}Q_{\beta^0}e'\right) - \lambda_e^n\right) = \frac{1}{\sqrt{n}T} tr\left(l_T\alpha' Q_{\beta^0}e'\right) + \sqrt{n}\left(\frac{1}{nT} tr\left(Z_{-1}^0 Q_{\beta^0}e'\right) - \lambda_e^n\right) \\ = \;& \frac{1}{\sqrt{n}T} tr\left(l_T\alpha' Q_{\beta^0}e'\right) + \frac{1}{\sqrt{n}T} tr\left(\left(Z_{-1}^0 - Z_{-1}^0(0)\right)Q_{\beta^0}e'\right) + \sqrt{n}\left(\frac{1}{nT} tr\left(Z_{-1}^0(0)\, Q_{\beta^0}e'\right) - \lambda_e^n\right) \\ = \;& I_c + II_c + III_c, \text{ say.} \end{aligned}$$

By Lemma 9(c), $II_c = o_p(1)$. In what follows, we will show that as $(n, T \to \infty)$ with $\frac{n}{T} \to 0$, $I_c = o_p(1)$ and $III_c \Rightarrow N\left(0, \frac{1}{2}\phi_e^4\right)$.

[7]For the proof of Lemma 8(c), see Moon and Perron (2002).

First, by Lemma 8 (i) and (i*), Assumptions 6 and 9, we have

$$\begin{aligned} |I_c| &\leq \frac{|tr(\alpha' e' l_T)|}{\sqrt{n}T} + \frac{\left|tr\left(\alpha' P_{\beta^0} e' l_T\right)\right|}{\sqrt{n}T} \\ &\leq \sqrt{\frac{n}{T}} \frac{\|\alpha\|}{\sqrt{n}} \frac{\|e' l_T\|}{\sqrt{nT}} + \frac{1}{\sqrt{T}} \frac{\|\alpha\|}{\sqrt{n}} \frac{\left\|\beta^0\right\|}{\sqrt{n}} \left\| \left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1} \right\| \frac{\left\|\beta^{0\prime} e' l_T\right\|}{\sqrt{nT}} \\ &= O_p\left(\sqrt{\frac{n}{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\right) = o_p(1). \end{aligned}$$

Next, for III_c, write

$$\begin{aligned} III_c &= \sqrt{n}\left(\frac{1}{nT} tr\left(E_{-1} Q_{\beta^0} e'\right) - \lambda_e^n\right) \\ &= \sqrt{n}\left(\frac{1}{nT} tr\left(E_{-1} e'\right) - \lambda_e^n\right) - \frac{1}{\sqrt{n}T} tr\left(E_{-1} P_{\beta^0} e'\right) = III_{ca} - III_{cb}, \text{ say.} \end{aligned}$$

First, we have

$$\begin{aligned} III_{cb} &= \frac{1}{2}\left(\frac{1}{\sqrt{n}T} tr\left(E_{-1} P_{\beta^0} e' + e P_{\beta^0} E'_{-1}\right)\right) = \frac{1}{2\sqrt{n}} tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\left(\frac{\beta^{0\prime} e' E_{-1}\beta^0 + \beta^{0\prime} E'_{-1} e \beta^0}{nT}\right)\right) \\ &\leq \frac{1}{2\sqrt{n}} \left\| \left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1} \right\| \left\| \frac{\beta^{0\prime} e' E_{-1}\beta^0 + \beta^{0\prime} E'_{-1} e \beta^0}{nT} \right\| = \frac{1}{\sqrt{n}} O(1) O_p(1) = O_p\left(\frac{1}{\sqrt{n}}\right) = o_p(1), \end{aligned}$$

where the third equality holds by Assumption 6 and Lemma 8(d).

Next, for III_{ca}, by Lemma 7(b), we have

$$\sqrt{n}\left(\frac{1}{nT} tr\left(E_{-1} e'\right) - \lambda_e^n\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{1}{T} \sum_{t=1}^{T} E_{it-1} e_{it} - \lambda_{e,i}\right) \Rightarrow N\left(0, \frac{1}{2}\phi_e^4\right)$$

as $(n, T \to \infty)$ with $\frac{n}{T} \to \infty$. Therefore,

$$III_c \Rightarrow N\left(0, \frac{1}{2}\phi_e^4\right)$$

as $(n, T \to \infty)$ with $\frac{n}{T} \to \infty$, and we have all the required results. ■

Proof of Lemma 3

Recall that $\beta_K^* = \beta^0 H_K$. Notice by Lemma 13 that, for n and T large, the matrix H_K is invertible. Then, $P_{\beta_K^*} = P_{\beta^0}$. So, the required result follows if we show that

$$\left\| P_{\hat{\beta}_K} - P_{\beta_K^*} \right\| = O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) = o_p(1).$$

By definition,

$$
\begin{aligned}
& \left\| P_{\hat{\beta}_K} - P_{\beta_K^*} \right\| \\
= & \left\| \frac{\hat{\beta}_K}{\sqrt{n}} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \frac{\hat{\beta}_K'}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \frac{\beta_K^{*\prime}}{\sqrt{n}} \right\| \\
\leq & \left\| \frac{\hat{\beta}_K}{\sqrt{n}} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \frac{\hat{\beta}_K'}{\sqrt{n}} - \frac{\hat{\beta}_K}{\sqrt{n}} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \frac{\beta_K^{*\prime}}{\sqrt{n}} \right\| + \left\| \frac{\hat{\beta}_K}{\sqrt{n}} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \frac{\beta_K^{*\prime}}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \frac{\beta_K^{*\prime}}{\sqrt{n}} \right\| \\
& + \left\| \frac{\beta_K^*}{\sqrt{n}} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \frac{\beta_K^{*\prime}}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \frac{\beta_K^{*\prime}}{\sqrt{n}} \right\| \\
\leq & \left\| \frac{\hat{\beta}_K}{\sqrt{n}} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \frac{\hat{\beta}_K}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \right\| + \left\| \frac{\hat{\beta}_K}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \right\| \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \frac{\beta_K^*}{\sqrt{n}} \right\| \\
& + \left\| \frac{\beta_K^*}{\sqrt{n}} \right\|^2 \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} - \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \right\| \\
\leq & \left\| \frac{\hat{\beta}_K}{\sqrt{n}} \right\| \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \frac{\hat{\beta}_K}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \right\| + \left\| \frac{\hat{\beta}_K}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \right\| \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \frac{\beta_K^*}{\sqrt{n}} \right\| \\
& + \left\| \frac{\beta_K^*}{\sqrt{n}} \right\|^2 \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \right\| \left\| \frac{\hat{\beta}_K' \hat{\beta}_K}{n} - \frac{\beta_K^{*\prime} \beta_K^*}{n} \right\| \\
\leq & \left\| \frac{\hat{\beta}_K}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}} \right\| \left\{ \begin{array}{c} \left\| \frac{\hat{\beta}_K}{\sqrt{n}} \right\| \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| + \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \frac{\beta_K^*}{\sqrt{n}} \right\| \\ + \left\| \frac{\beta_K^*}{\sqrt{n}} \right\|^2 \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \right\| \left(\left\| \frac{\hat{\beta}_K}{\sqrt{n}} \right\| + \left\| \frac{\beta_K^*}{\sqrt{n}} \right\| \right) \end{array} \right\}. \qquad (38)
\end{aligned}
$$

First, using the definitions of $\hat{\beta}_K$ and $\bar{\beta}$, we have

$$\hat{\beta}_K = \left(\frac{\tilde{y}'\tilde{y}}{nT} \right) \bar{\beta}_K = \bar{\beta}_K \Lambda_{nT,K}.$$

By Lemma 12(a), as $(n, T \to \infty)$ with $\frac{n}{T} \to 0$,

$$\frac{\hat{\beta}_K' \hat{\beta}_K}{n} = \Lambda_{nT,K} \to_p \Lambda_K,$$

where Λ_K is a full rank matrix. Thus,

$$\left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} = O_p(1). \qquad (39)$$

Next,

$$\left\| \frac{\hat{\beta}_K}{\sqrt{n}} \right\|^2 = O_p(1) \qquad (40)$$

because

$$\left\| \frac{\hat{\beta}_K}{\sqrt{n}} \right\|^2 = tr \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right) = tr\left(\Lambda_{nT,K} \right) \to_p tr(\Lambda_K).$$

Recall that $\beta_K^* = \beta^0 H_K$. By Lemma 13 and Assumption 6,

$$\left(\frac{\beta_K^{*\prime}\beta_K^*}{n}\right)^{-1} = O_p(1).$$

Similarly,

$$\left\|\frac{\beta_K^*}{\sqrt{n}}\right\| = O_p(1).$$

Finally, by Lemma 14(a), $\left\|\frac{\hat{\beta}_K}{\sqrt{n}} - \frac{\beta_K^*}{\sqrt{n}}\right\| = O_p\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right)\right)$.

Thus, in view of (38) , we have the required result,

$$\left\|P_{\hat{\beta}_K} - P_{\beta_K^*}\right\| = O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) = o_p(1). \blacksquare \tag{41}$$

Before we start the proof of Lemma 4, we introduce the following Lemmas. Define $\tilde{\Gamma}_i(j) = \frac{1}{T}\sum_t e_{it}e_{it+j}$, where the summation $\sum_t$ runs over $1 \leq t, t+j \leq T$. Define $\tilde{\lambda}_{e,i}$ and $\tilde{\omega}_{e,i}^2$

$$\tilde{\lambda}_{e,i} = \sum_{j=1}^{T-1} w\left(\frac{j}{h_\lambda}\right)\tilde{\Gamma}_i(j)$$

and

$$\tilde{\omega}_{e,i}^2 = \sum_{j=-T+1}^{T-1} w\left(\frac{j}{h}\right)\tilde{\Gamma}_i(j).$$

We denote

$$\tilde{\lambda}_e^n = \frac{1}{n}\sum_{i=1}^n \tilde{\lambda}_{e,i},\ \tilde{\omega}_e^{n,2} = \frac{1}{n}\sum_{i=1}^n \tilde{\omega}_{e,i}^2,\ \text{and}\ \tilde{\phi}_e^{n,4} = \frac{1}{n}\sum_{i=1}^n \tilde{\omega}_{e,i}^4.$$

Lemma 15 *(a) Under the conditions of Lemma 4(a),* $\sqrt{n}\left(\tilde{\lambda}_e^n - \tilde{\lambda}_e^n\right) = o_p(1)$.

(b) Under the conditions of Lemma 4(b), $\hat{\omega}_e^2 - \tilde{\omega}_e^2 = o_p(1)$.

(c) Under the conditions of Lemma 4(c), $\hat{\phi}_e^4 - \tilde{\phi}_e^4 = o_p(1)$.

Proof of Lemma 15.

We omit the proof. The proof could be found in Moon and Perron (2002). $\blacksquare$

Proof of Lemma 4.

Part (a).

In order to have the required result, by Lemma 15(a), it is enough to show that

$$\sqrt{n}\left(\tilde{\lambda}_e^n - \lambda_e^n\right) = o_p(1).$$

Along the same lines as the proofs of Theorems 9 and 10 of Hannan (1970), it is possible to show under Assumption 2 that

$$E\left(\sqrt{n}\left(\tilde{\lambda}_e^n - \lambda_e^n\right)\right)^2 \leq \sup_i var\left(\tilde{\lambda}_{e,i}\right) + n\left(\sup_i bias\left(\tilde{\lambda}_{e,i}\right)\right)^2 = O\left(\frac{h_\lambda}{T}\right) + O\left(\frac{n}{h_\lambda^{2q}}\right),$$

if $\frac{h_\lambda^q}{T} \to 0$ in the case of $q \geq 1$[8]. So, we have

$$\sqrt{n}\left(\tilde{\lambda}_e^n - \lambda_e^n\right) = O_p\left(\max\left\{\sqrt{\frac{h_\lambda}{T}}, \sqrt{\frac{n}{h_\lambda^{2q}}}\right\}\right),$$

if $\frac{h_\lambda^q}{T} \to 0$ with $q \geq 1$

First, since $\frac{1}{2q} < b$,

$$\frac{n}{h_\lambda^{2q}} \sim n^{(1-2qb)} \to 0.$$

Next, for n, T large,

$$\frac{h_\lambda^q}{T} \sim \frac{n^{bq}}{T} = e^{\log\left(\frac{n^{bq}}{T}\right)} = e^{\left(bq - \frac{\log T}{\log n}\right)\log n} = n^{bq - \frac{\log T}{\log n}} \leq n^{bq - \liminf \frac{\log T}{\log n}} \to 0$$

since $b < \frac{a}{q}$ if $q \geq 1$. Similarly, $\frac{h_\lambda}{T} \to 0$. Therefore,

$$\sqrt{n}\left(\tilde{\lambda}_e^n - \lambda_e^n\right) = O_p\left(\max\left\{\sqrt{\frac{h_\lambda}{T}}, \sqrt{\frac{n}{h_\lambda^{2q}}}\right\}\right) = o_p(1),$$

as required for Part (a). ■

Part (b).

Similar to Part (a), by Lemma 15(b), it is sufficient to show that

$$\tilde{\omega}_e^{2,n} - \omega_e^{2,n} = o_p(1).$$

Along the same lines as the proofs of Theorems 9 and 10 of Hannan (1970), it is possible to show under Assumption 2 that

$$\begin{aligned} E\left(\tilde{\omega}_e^{2,n} - \omega_e^{2,n}\right)^2 &= E\left(\frac{1}{n}\sum_{i=1}^n \left(\tilde{\omega}_{e,i}^2 - E\tilde{\omega}_{e,i}^2\right)\right)^2 + \left(\frac{1}{n}\sum_{i=1}^n \left(E\tilde{\omega}_{e,i}^2 - \omega_{e,i}^2\right)\right)^2 \\ &\leq \frac{1}{n}\sup_i var\left(\tilde{\omega}_{e,i}^2\right) + \left(\sup_i bias\left(\tilde{\omega}_{e,i}^2\right)\right)^2 = O\left(\frac{h_\omega}{nT}\right) + O\left(\frac{1}{h_\omega^{2q}}\right), \end{aligned}$$

if $\frac{h_\omega^q}{T} \to 0$ in the case of $q \geq 1$. So, we have

$$\tilde{\omega}_e^{2,n} - \omega_e^{2,n} = O_p\left(\max\left\{\sqrt{\frac{h_\omega}{nT}}, \frac{1}{h_\omega^q}\right\}\right).$$

Under the conditions in the lemma,

$$\frac{1}{h_\omega^q} \to 0$$

and for n, T large,

$$\frac{h_\omega}{nT} \sim \frac{n^{b-1}}{T} = e^{\log\left(\frac{n^{b-1}}{T}\right)} = e^{\left(b-1-\frac{\log T}{\log n}\right)\log n} = n^{b-1-\frac{\log T}{\log n}} \leq n^{b-1-\liminf \frac{\log T}{\log n}} \to 0$$

[8] For details on this, see equation (B.27) on page 985 of Moon and Phillips (2000).

since $b < 1 < a + 1$. Also, if $q \geq 1$

$$\frac{h_\omega^q}{T} \sim \frac{n^{bq}}{T} = e^{\log\left(\frac{n^{bq}}{T}\right)} = e^{\left(bq - \frac{\log T}{\log n}\right)\log n} = n^{bq - \frac{\log T}{\log n}} \leq n^{bq - \liminf \frac{\log T}{\log n}}$$

since $q < \min\left\{1, \frac{a}{2}, \frac{a}{q}\right\}$, $a = \liminf \frac{\log T}{\log n}$. Therefore,

$$\tilde{\omega}_e^{2,n} - \omega_e^{2,n} = o_p(1),$$

and we have all the required results to complete the proof. ■

Part (c)

Part (c) follows from Lemma 15(c) and the arguments used in the proof of Lemma 15(c). For details, see Moon and Perron (2002). ■

Proof of Theorem 2

In view of (9) and (10) and Lemma 4, the required result follows if we show

$$\text{(a)} \ \frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}Z_{-1}' - Z_{-1}Q_{\beta^0}Z_{-1}'\right)}{nT^2} = o_p(1)$$

and

$$\text{(b)} \ \frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}(Z - Z_{-1})' - Z_{-1}Q_{\beta^0}(Z - Z_{-1})'\right)}{\sqrt{n}T} = o_p(1).$$

Part (a)

Part (a) follows because

$$\frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}Z_{-1}' - Z_{-1}Q_{\beta^0}Z_{-1}'\right)}{nT^2} = \frac{tr\left(Z_{-1}P_{\beta_K^*}Z_{-1}' - Z_{-1}P_{\hat{\beta}_K}Z_{-1}'\right)}{nT^2} = \frac{tr\left(\left(P_{\beta_K^*} - P_{\hat{\beta}_K}\right)Z_{-1}'Z_{-1}\right)}{nT^2}$$

$$\leq \left\|P_{\hat{\beta}_K} - P_{\beta_K^*}\right\| \frac{\|Z_{-1}\|^2}{nT^2} = o_p(1)\,O_p(1) = o_p(1),$$

where the second line holds by Lemma 13 and the last line holds by Lemma 8(a) and Lemma 3, recalling that $\left\|P_{\hat{\beta}_K} - P_{\beta_K^*}\right\| = \left\|Q_{\hat{\beta}_K} - Q_{\beta^0}\right\|$ because H_K is asymptotically invertible. .

Part (b)

By definition $Z - Z_{-1} = Z^0 - Z_{-1}^0 = -\frac{1}{\sqrt{n}T}Z_{-1}^0\Theta + y$, and we write

$$\frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}(Z - Z_{-1})' - Z_{-1}Q_{\beta^0}(Z - Z_{-1})'\right)}{\sqrt{n}T}$$

$$= -\frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}\Theta Z_{-1}^{0\prime} - Z_{-1}Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)}{nT^2} + \frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}y' - Z_{-1}Q_{\beta^0}y'\right)}{\sqrt{n}T} = -I_b + II_b, \text{ say.}$$

First for I_b,

$$I_b = I_{ba} + I_{bb},$$

where

$$I_{ba} = \frac{tr\left(l_T\alpha' Q_{\hat{\beta}_K}\Theta Z_{-1}^{0\prime} - l_T\alpha' Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)}{nT^2} \text{ and } I_{bb} = \frac{tr\left(Z_{-1}^0 Q_{\hat{\beta}_K}\Theta Z_{-1}^{0\prime} - Z_{-1}^0 Q_{\beta^0}\Theta Z_{-1}^{0\prime}\right)}{nT^2}.$$

Then, by modifying the proof of Lemma 8(a)[9] and noting that $\left\|P_{\hat{\beta}_K}\right\| = O_p(1)$, we have

$$I_{ba} \leq \frac{1}{\sqrt{T}}\frac{\|l_T\|}{\sqrt{T}}\frac{\|\alpha\|}{\sqrt{n}}\frac{\left\|Z_{-1}^0\Theta\right\|}{\sqrt{n}T}\left(\left\|P_{\hat{\beta}_K}\right\| + \left\|P_{\beta^0}\right\|\right) = O_p\left(\frac{1}{\sqrt{T}}\right)$$

The proof of $I_{bb} = o_p(1)$ is similar to the proof of Part (a) and we omit it. Combining I_{ba} and I_{bb}, we have

$$I_b = o_p(1).$$

Next, for II_b, notice that

$$\frac{tr\left(Z_{-1}P_{\hat{\beta}_K}y'\right)}{\sqrt{n}T} = \frac{1}{2}\frac{tr\left(Z_{-1}P_{\hat{\beta}_K}y' + yP_{\hat{\beta}_K}Z_{-1}'\right)}{\sqrt{n}T} = \frac{1}{2n\sqrt{n}T}tr\left(\left(\frac{\hat{\beta}_K'\hat{\beta}_K}{n}\right)^{-1}\left[\hat{\beta}_K'\left(y'Z_{-1} + Z_{-1}'y\right)\hat{\beta}_K\right]\right).$$

Then,

$$\begin{aligned}
&\left|\frac{tr\left(Z_{-1}Q_{\hat{\beta}_K}y' - Z_{-1}Q_{\beta^0}y'\right)}{\sqrt{n}T}\right| = \left|\frac{tr\left(Z_{-1}P_{\hat{\beta}_K}y' - Z_{-1}P_{\beta_K^*}y'\right)}{\sqrt{n}T}\right| \\
= &\left|\frac{1}{2n\sqrt{n}T}tr\left(\left(\frac{\hat{\beta}_K'\hat{\beta}_K}{n}\right)^{-1}\left[\hat{\beta}_K'\left(y'Z_{-1} + Z_{-1}'y\right)\hat{\beta}_K\right] - \left(\frac{\beta_K^{*\prime}\beta_K^*}{n}\right)^{-1}\left[\beta_K^{*\prime}\left(y'Z_{-1} + Z_{-1}'y\right)\beta_K^*\right]\right)\right| \\
\leq &\left|\frac{1}{2n\sqrt{n}T}tr\left(\left(\frac{\hat{\beta}_K'\hat{\beta}_K}{n}\right)^{-1}\left[\hat{\beta}_K'\left(y'Z_{-1} + Z_{-1}'y\right)\hat{\beta}_K - \beta_K^{*\prime}\left(y'Z_{-1} + Z_{-1}'y\right)\beta_K^*\right]\right)\right| \\
&+\left|\frac{1}{2n\sqrt{n}T}tr\left(\left[\left(\frac{\hat{\beta}_K'\hat{\beta}_K}{n}\right)^{-1} - \left(\frac{\beta_K^{*\prime}\beta_K^*}{n}\right)^{-1}\right]\beta_K^{*\prime}\left(y'Z_{-1} + Z_{-1}'y\right)\beta_K^*\right)\right| \\
\leq &\frac{1}{2}\left\|\left(\frac{\hat{\beta}_K'\hat{\beta}_K}{n}\right)^{-1}\right\|\left\|\frac{\hat{\beta}_K'\left(y'Z_{-1} + Z_{-1}'y\right)\hat{\beta}_K - \beta_K^{*\prime}\left(y'Z_{-1} + Z_{-1}'y\right)\beta_K^*}{n\sqrt{n}T}\right\| \\
&+\frac{1}{2}\left\|\sqrt{n}\left\{\left(\frac{\hat{\beta}_K'\hat{\beta}_K}{n}\right)^{-1} - \left(\frac{\beta_K^{*\prime}\beta_K^*}{n}\right)^{-1}\right\}\right\|\left\|\frac{\beta_K^*}{\sqrt{n}}\right\|^2\left\|\frac{y'Z_{-1} + Z_{-1}'y}{nT}\right\| \\
= &III_b + IV_b, \text{ say.}
\end{aligned}$$

In what follows, we will show that $III_b, IV_b = o_p(1)$.

First, for IV_b, since $\left\|\frac{\beta_K^*}{\sqrt{n}}\right\|^2 \leq \left\|\frac{\beta^0}{\sqrt{n}}\right\|^2 \|H_K\|^2 = O_p(1)$ by Assumption 6 and Lemma 13. Also, $\left\|\frac{y'Z_{-1} + Z_{-1}'y}{nT}\right\| = O_p(1)$ by Lemma 8(b). Thus, the required result $IV_b = o_p(1)$ follows if we show that

$$\sqrt{n}\left\{\left(\frac{\hat{\beta}_K'\hat{\beta}_K}{n}\right)^{-1} - \left(\frac{\beta_K^{*\prime}\beta_K^*}{n}\right)^{-1}\right\} = o_p(1),$$

[9] See Moon and Perron (2002) for the proof of Lemma 8(a).

which holds because

$$\begin{aligned}
& \left\| \sqrt{n} \left\{ \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} - \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \right\} \right\| \\
= & \left\| \sqrt{n} \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} - \frac{\beta_K^{*\prime} \beta_K^*}{n} \right) \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \\
\leq & \left\| \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \right\| \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\| \sqrt{n} \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} - \frac{\beta_K^{*\prime} \beta_K^*}{n} \right) \right\| \\
\leq & \left\| \left(\frac{\beta_K^{*\prime} \beta_K^*}{n} \right)^{-1} \right\| \left\| \left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} \right\| \left\{ \left\| \hat{\beta}_K' \left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}} \right) \right\| + \left\| \left(\frac{\hat{\beta}_K - \beta_K^*}{\sqrt{n}} \right)' \beta_K^* \right\| \right\} \\
= & O_p(1) O_p(1) (o_p(1) + o_p(1)) = o_p(1),
\end{aligned}$$

where the last line holds by Lemma 14(c) and (d).

For III_b, since $\left(\frac{\hat{\beta}_K' \hat{\beta}_K}{n} \right)^{-1} = O_p(1)$, for the required result, it is enough to show that $\left\| \frac{\hat{\beta}_K' (y' Z_{-1} + Z_{-1}' y) \hat{\beta}_K - \beta_K^{*\prime} (y' Z_{-1} + Z_{-1}' y) \beta_K^*}{n \sqrt{n} T} \right\| = o_p(1)$. Since

$$\begin{aligned}
& y' Z_{-1} + Z_{-1}' y \\
= & y' l_T \alpha' + \alpha l_T' y + Z_T^0 Z_T^{0\prime} + (\rho - I_n) Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \\
& + Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) + (\rho - I_n) Z_{-1}^{0\prime} Z_{-1}^0 + (\rho - I_n) Z_{-1}^{0\prime} y + y' Z_{-1}^0 (\rho - I_n) + y' y,
\end{aligned}$$

the required result follows by the triangle inequality if we show that

$$\text{(i)} \quad \left\| \frac{\hat{\beta}_K' y' l_T \alpha' \hat{\beta}_K - \beta_K^{*\prime} y' l_T \alpha' \beta_K^*}{n \sqrt{n} T} \right\| = o_p(1),$$

$$\text{(ii)} \quad \left\| \frac{\hat{\beta}_K' (\rho - I_n) Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \hat{\beta}_K - \beta_K^{*\prime} (\rho - I_n) Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \beta_K^*}{n \sqrt{n} T} \right\| = o_p(1),$$

$$\text{(iii)} \quad \left\| \frac{\hat{\beta}_K' Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \hat{\beta}_K - \beta_K^{*\prime} Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \beta_K^*}{n \sqrt{n} T} \right\| = o_p(1)$$

$$\text{(iv)} \quad \left\| \frac{\hat{\beta}_K' (\rho - I_n) Z_{-1}^{0\prime} y \hat{\beta}_K - \beta_K^{*\prime} (\rho - I_n) Z_{-1}^{0\prime} y \beta_K^*}{n \sqrt{n} T} \right\| = o_p(1),$$

$$\text{(v)} \quad \left\| \frac{\hat{\beta}_K' Z_T^0 Z_T^{0\prime} \hat{\beta}_K - \beta_K^{*\prime} Z_T^0 Z_T^{0\prime} \beta_K^*}{n \sqrt{n} T} \right\| = o_p(1),$$

$$\text{and (vi)} \quad \left\| \frac{\hat{\beta}_K' y' y \hat{\beta}_K - \beta_K^{*\prime} y' y \beta_K^*}{n \sqrt{n} T} \right\| = o_p(1).$$

Part (i) holds because

$$\begin{aligned}
&\left\| \frac{\hat{\beta}_K' y' l_T \alpha' \hat{\beta}_K - \beta_K^{*\prime} y' l_T \alpha' \beta_K^*}{n\sqrt{n}T} \right\| \\
\leq\ & \left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' l_T \alpha' \hat{\beta}_K}{n\sqrt{n}T} \right\| + \left\| \frac{\beta_K^{*\prime} y' l_T \alpha' \left(\hat{\beta}_K - \beta_K^*\right)}{n\sqrt{n}T} \right\| \\
\leq\ & \frac{2}{\sqrt{T}} \left\| \hat{\beta}_K - \beta_K^* \right\| \frac{\|y' l_T\|}{\sqrt{nT}} \frac{\|\alpha\|}{\sqrt{n}} \left(\frac{\left\| \hat{\beta}_K \right\|}{\sqrt{n}} + \frac{\|\beta_K^*\|}{\sqrt{n}} \right) = O_p\left(\frac{1}{\sqrt{T}}\right) = o_p(1),
\end{aligned}$$

where the first equality holds by Lemma 14(b) and Lemma 8(i**).

Part (ii) holds because

$$\begin{aligned}
&\left\| \frac{\hat{\beta}_K' (\rho - I_n) Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \hat{\beta}_K - \beta_K^{*\prime} (\rho - I_n) Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \beta_K^*}{n\sqrt{n}T} \right\| \\
\leq\ & \frac{\sqrt{n}}{T} \|T(\rho - I_n)\|^2 \frac{\left\| Z_{-1}^0 \right\|^2}{nT^2} \left\{ \frac{\left\| \hat{\beta}_K \right\|^2 + \|\beta_K^*\|^2}{n} \right\} = O_p\left(\frac{\sqrt{n}}{T}\right) = o_p(1).
\end{aligned}$$

For Part (iii), by the triangle inequality and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\left\| \frac{\hat{\beta}_K' Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \hat{\beta}_K - \beta_K^{*\prime} Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \beta_K^*}{n\sqrt{n}T} \right\| \\
\leq\ & \left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)' Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \left(\hat{\beta}_K - \beta_K^*\right)}{n\sqrt{n}T} \right\| + \left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)' Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \beta_K^*}{n\sqrt{n}T} \right\| \\
&+ \left\| \frac{\beta_K^{*\prime} Z_{-1}^{0\prime} Z_{-1}^0 (\rho - I_n) \left(\hat{\beta}_K - \beta_K^*\right)}{n\sqrt{n}T} \right\| \\
\leq\ & \frac{\left\| Z_{-1}^0 \right\|^2}{nT^2} \left\{ \begin{array}{c} \frac{1}{n} \left\| \hat{\beta}_K - \beta_K^* \right\| \left\| \Theta \left(\hat{\beta}_K - \beta_K^*\right) \right\| \\ + \frac{1}{\sqrt{n}} \left(\left\| \hat{\beta}_K - \beta_K^* \right\| \frac{\|\Theta \beta_K^*\|}{\sqrt{n}} + \frac{\|\beta_K^*\|}{\sqrt{n}} \left\| \Theta \left(\hat{\beta}_K - \beta_K^*\right) \right\| \right) \end{array} \right\}. \qquad (42)
\end{aligned}$$

By modifying the proof of Lemma 14(a) and (b), it is possible to show that

$$\left\| \Theta \left(\hat{\beta}_K - \beta_K^*\right) \right\| = O_p(1).$$

Also,

$$\frac{\|\Theta \beta_K^*\|}{\sqrt{n}} \leq \frac{\left\| \Theta \beta^0 \right\|}{\sqrt{n}} \|H_K\| = O_p(1).$$

Thus, together with Lemma 8(a), we have

$$(42) = O_p(1) \left\{ \frac{1}{n} O_p(1) + \frac{1}{\sqrt{n}} O_p(1) \right\} = o_p(1),$$

as required for Part (iii).

Part (iv) holds similarly since

$$\begin{aligned}
&\left\| \frac{\hat{\beta}_K^{'}\left(\rho - I_n\right) Z_{-1}^{0\prime} y \hat{\beta}_K - \beta_K^{*\prime}\left(\rho - I_n\right) Z_{-1}^{0\prime} y \beta_K^*}{n\sqrt{n}T} \right\| \\
\leq & \left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)^{'}\left(\rho - I_n\right) Z_{-1}^{0\prime} y \hat{\beta}_K}{n\sqrt{n}T} \right\| + \left\| \frac{\beta_K^{*\prime}\left(\rho - I_n\right) Z_{-1}^{0\prime} y \left(\hat{\beta}_K - \beta_K^*\right)}{n\sqrt{n}T} \right\| \\
\leq & \frac{1}{\sqrt{nT}} \frac{\left\| Z_{-1}^0 \right\|}{\sqrt{n}T} \frac{\|y\|}{\sqrt{nT}} \left(\left\| \left(\hat{\beta}_K - \beta_K^*\right)\Theta \right\| \frac{\left\|\hat{\beta}_K\right\|}{\sqrt{n}} - \left\| \hat{\beta}_K - \beta_K^* \right\| \frac{\|\beta_K^*\Theta\|}{\sqrt{n}} \right) \\
= & O_p\left(\frac{1}{\sqrt{nT}}\right) = o_p\left(1\right).
\end{aligned}$$

For Part (v), notice that

$$\begin{aligned}
&\left\| \frac{\hat{\beta}_K^{'} Z_T^0 Z_T^{0\prime} \hat{\beta}_K - \beta_K^{*\prime} Z_T^0 Z_T^{0\prime} \beta_K^*}{n\sqrt{n}T} \right\| \\
\leq & \frac{\left\| \left(\hat{\beta}_K - \beta_K^*\right)^{'} Z_T^0 \right\| \left\| Z_T^{0\prime} \hat{\beta}_K \right\|}{n\sqrt{n}T} + \frac{\left\| \beta_K^{*\prime} Z_T^0 \right\| \left\| Z_T^{0\prime}\left(\hat{\beta}_K - \beta_K^*\right) \right\|}{n\sqrt{n}T} \\
\leq & \frac{\left\| \left(\hat{\beta}_K - \beta_K^*\right)^{'} Z_T^0 \right\|^2}{n\sqrt{n}T} + 2\frac{\left\| \beta_K^{*\prime} Z_T^0 \right\| \left\| Z_T^{0\prime}\left(\hat{\beta}_K - \beta_K^*\right) \right\|}{n\sqrt{n}T} \\
= & \left\| \frac{Z_T^{0\prime}\left(\hat{\beta}_K - \beta_K^*\right)}{\sqrt{n}\sqrt{T}} \right\| \left\{ \frac{1}{\sqrt{n}} \left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)^{'} Z_T^0}{\sqrt{n}\sqrt{T}} \right\| + 2\left\|H_K\right\| \left\| \frac{\beta^{0\prime} Z_T^0}{n\sqrt{T}} \right\| \right\}. \qquad (43)
\end{aligned}$$

In what follows we show that

$$(\mathrm{v}_1) \quad \left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)^{'} Z_T^0}{\sqrt{n}\sqrt{T}} \right\| = o_p\left(1\right) \text{ and } (\mathrm{v}_2) \quad \left\| \frac{\beta^{0\prime} Z_T^0}{n\sqrt{T}} \right\| = O_p\left(1\right).$$

Then, the result of Part (v) follows.

First, Part (v_2) follows because

$$\begin{aligned}
\left\| \frac{\beta^{0\prime} Z_T^0}{n\sqrt{T}} \right\| &\leq \left\| \frac{\beta^{0\prime} Z_T^0\left(0\right)}{n\sqrt{T}} \right\| + \left\| \frac{\beta^{0\prime}\left(Z_T^0 - Z_T^0\left(0\right)\right)}{n\sqrt{T}} \right\| \\
&\leq \left\| \frac{\beta^{0\prime} Z_T^0\left(0\right)}{n\sqrt{T}} \right\| + \frac{1}{\sqrt{n}} \frac{\left\|\beta^0\right\|}{\sqrt{n}} \frac{\left\| Z_T^0 - Z_T^0\left(0\right) \right\|}{\sqrt{T}} \\
&= O_p\left(1\right) + O_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}$$

where the $O_p\left(\frac{1}{\sqrt{n}}\right)$ term holds by similar arguments used in Lemma 10(a), and the $O_p\left(1\right)$

holds since

$$\begin{aligned}\left\|\frac{\beta^{0\prime}Z_T^0(0)}{n\sqrt{T}}\right\| &\leq \left\|\frac{\beta^{0\prime}\beta^0 F_T}{n\sqrt{T}}\right\| + \left\|\frac{\beta^{0\prime}E_T}{n\sqrt{T}}\right\| \\ &= \left\|\frac{\beta^{0\prime}\beta^0}{n}\right\|\left\|\frac{F_T}{\sqrt{T}}\right\| + \frac{1}{\sqrt{n}}\left\|\frac{\beta^{0\prime}e'l_T}{\sqrt{n}\sqrt{T}}\right\| \\ &= O(1)\,O_p(1) + \frac{1}{\sqrt{n}}O_p(1) = O_p(1) \qquad (44)\end{aligned}$$

by Assumptions 6, 3, and Lemma 8(i). Next, for Part (v_2), by the triangle inequality,

$$\begin{aligned}&\frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)' Z_T^0\right\|}{\sqrt{nT}} \\ \leq\ & \frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)' Z_T^0(0)\right\|}{\sqrt{nT}} + \frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)'\left(Z_T^0 - Z_T^0(0)\right)\right\|}{\sqrt{nT}} \\ \leq\ & \frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)'\beta^0 F_T\right\|}{\sqrt{nT}} + \frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)' E_T\right\|}{\sqrt{nT}} + \frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)'\left(Z_T^0 - Z_T^0(0)\right)\right\|}{\sqrt{nT}} \\ =\ & V_a + V_b + V_c, \text{ say.}\end{aligned}$$

In view of (31) , (36) , and by the triangle inequality, we have

$$\begin{aligned}V_a &\leq \frac{\left\|\bar{\beta}_K' e' f^0 \beta^{0\prime}\beta^0 F_T\right\|}{n\sqrt{n}T\sqrt{T}} + \frac{\left\|\bar{\beta}_K'\beta^0 f^{0\prime} e\beta^0 F_T\right\|}{n\sqrt{n}T\sqrt{T}} + \frac{\left\|\bar{\beta}_K' e' e\beta^0 F_T\right\|}{n\sqrt{n}T\sqrt{T}} \\ &\quad + \sum_{k=1}^{7}\sqrt{n}\left\|\mathcal{R}_k\right\|\frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}}\frac{\left\|\beta^0\right\|}{\sqrt{n}}\frac{\|F_T\|}{\sqrt{T}} \\ &= 2\sqrt{\frac{n}{T}}\left\|\frac{\bar{\beta}_K}{\sqrt{n}}\right\|\left\|\frac{e' f^0}{\sqrt{nT}}\right\|\left\|\frac{\beta^0}{\sqrt{n}}\right\|^2\left\|\frac{F_T}{\sqrt{T}}\right\| + \left\|\frac{\bar{\beta}_K' e' e\beta_K^*}{n\sqrt{n}T}\right\|\left\|H_K^{-1}\right\|\left\|\frac{F_T}{\sqrt{T}}\right\| \\ &\quad + \sum_{k=1}^{7}\sqrt{n}\left\|\mathcal{R}_k\right\|\frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}}\frac{\left\|\beta^0\right\|}{\sqrt{n}}\frac{\|F_T\|}{\sqrt{T}} \\ &= O_p\left(\sqrt{\frac{n}{T}}\right) + o_p(1) + O_p\left(\sqrt{\frac{n}{T}}\max\left\{\frac{1}{\sqrt{n}},\frac{1}{\sqrt{T}}\right\}\right) = o_p(1).\end{aligned}$$

Similarly,

$$\begin{aligned}V_b &\leq \left\|\frac{\bar{\beta}_K' e' f^0\beta^{0\prime}E_T}{n\sqrt{n}T\sqrt{T}}\right\| + \left\|\frac{\bar{\beta}_K'\beta^0 f^{0\prime}eE_T}{n\sqrt{n}T\sqrt{T}}\right\| + \left\|\frac{\bar{\beta}_K' e' eE_T}{n\sqrt{n}T\sqrt{T}}\right\| \\ &\quad + \sum_{k=1}^{7}\sqrt{n}\left\|\mathcal{R}_k\right\|\frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}}\frac{\|E_T\|}{\sqrt{nT}} \\ &\leq 2\sqrt{\frac{n}{T}}\frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}}\frac{\left\|e' f^0\right\|}{\sqrt{nT}}\frac{\left\|\beta^0\right\|}{\sqrt{n}}\frac{\|E_T\|}{\sqrt{nT}} + \left\|\frac{\bar{\beta}_K' e' eE_T}{n\sqrt{n}T\sqrt{T}}\right\| + \sum_{k=1}^{7}\sqrt{n}\left\|\mathcal{R}_k\right\|\frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}}\frac{\|E_T\|}{\sqrt{nT}} \\ &= \sqrt{\frac{n}{T}}O_p(1) + \frac{\left\|\bar{\beta}_K' e' eE_T\right\|}{n\sqrt{n}T\sqrt{T}} + \sqrt{\frac{n}{T}}O_p\left(\max\left\{\frac{1}{\sqrt{n}},\frac{1}{\sqrt{T}}\right\}\right),\end{aligned}$$

where the last line holds by the same principle of the last line of Part V_1 with $\frac{\|E_T\|}{\sqrt{nT}} = \frac{\|e'l_T\|}{\sqrt{nT}} = O_p(1)$ by Lemma 8 (i) and (i*). For the remaining term $\frac{\left\|\bar{\beta}_K' e' e E_T\right\|}{n\sqrt{n}T\sqrt{T}}$, using $\hat{\beta}_K = \bar{\beta}_K \Lambda_{nT,K}$ and $\beta_K^* = \beta^0 H_K$ and by the triangle inequality,

$$\frac{\left\|\bar{\beta}_K' e' e E_T\right\|}{n\sqrt{n}T\sqrt{T}} \leq \left\|\Lambda_{nT,K}^{-1}\right\| \frac{\left\|(\hat{\beta}_K - \beta_K^*)' e' e E_T\right\|}{n\sqrt{n}T\sqrt{T}} + \left\|\Lambda_{nT,K}^{-1}\right\| \left\|H_K^{-1}\right\| \frac{\left\|\beta^{0\prime} e' e E_T\right\|}{n\sqrt{n}T\sqrt{T}}.$$

The first term is

$$\begin{aligned} \left\|\Lambda_{nT,K}^{-1}\right\| \frac{\left\|(\hat{\beta}_K - \beta_K^*)' e' e E_T\right\|}{n\sqrt{n}T\sqrt{T}} &\leq \left\|\Lambda_{nT,K}^{-1}\right\| \left\|\hat{\beta}_K - \beta_K^*\right\| \frac{\|e'e\|}{nT} \frac{\|E_T\|}{\sqrt{nT}} \\ &= O_p(1) O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) O_p(1) = o_p(1), \end{aligned}$$

where the last line holds by Lemma 14(b), Lemma 8(e). Also the second term is

$$\left\|\Lambda_{nT,K}^{-1}\right\| \left\|H_K^{-1}\right\| \frac{\left\|\beta^{0\prime} e' e E_T\right\|}{n\sqrt{n}T\sqrt{T}} = O_p(1) O_p(1) o_p(1) = o_p(1)$$

by Lemma 8(h). Therefore, $\frac{\left\|\bar{\beta}_K' e' e E_T\right\|}{n\sqrt{n}T\sqrt{T}} = o_p(1)$ and so in consequence,

$$V_b = o_p(1). \tag{45}$$

Before we start the proof of $V_c = o_p(1)$, we define

$$F_T^* = \sum_{t=1}^{T} \left(1 - \frac{t}{T}\right) f_t$$

and

$$E_T^* = \sum_{t=1}^{T} \left(1 - \frac{t}{T}\right) e_t,$$

where $e_t = (e_{1t}, ..., e_{nt})'$. Using this notation, we write

$$Z_T^0 - Z_T^0(0) \sim -\frac{1}{\sqrt{n}} \Theta \beta^0 F_T^* - \frac{1}{\sqrt{n}} \Theta E_T^*.$$

Since $\frac{E\|F_T^*\|^2}{T}, \frac{E\|\Theta E_T^*\|^2}{nT} < M$, we have

$$\frac{\|F_T^*\|}{\sqrt{T}} = O_p(1)$$

and

$$\frac{\|\Theta E_T^*\|}{\sqrt{nT}} = O_p(1).$$

By the triangle inequality,

$$V_c \leq \frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)' \Theta \beta^0 F_T^*\right\|}{n\sqrt{T}} + \frac{\left\|\left(\hat{\beta}_K - \beta_K^*\right)' \Theta E_T^*\right\|}{n\sqrt{T}} = V_{ca} + V_{cb}, \text{ say.}$$

Similar to V_a, we have

$$\begin{aligned}
V_{ca} &\leq \frac{\left\|\bar{\beta}_K' e' f^0 \beta^{0\prime} \Theta \beta^0 F_T^*\right\|}{n^2 T\sqrt{T}} + \frac{\left\|\bar{\beta}_K' \beta^0 f^{0\prime} e \Theta \beta^0 F_T^*\right\|}{n^2 T\sqrt{T}} + \frac{\left\|\bar{\beta}_K' e' e \Theta \beta^0 F_T^*\right\|}{n^2 T\sqrt{T}} \\
&\quad + \sum_{k=1}^{7} \|\mathcal{R}_k\| \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\left\|\Theta\beta^0\right\|}{\sqrt{n}} \frac{\|F_T^*\|}{\sqrt{T}} \\
&\leq \frac{2}{\sqrt{T}} \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\left\|e' f^0\right\|}{\sqrt{nT}} \frac{\left\|\beta^0\right\|}{\sqrt{n}} \frac{\left\|\Theta\beta^0\right\|}{\sqrt{n}} \frac{\|F_T^*\|}{\sqrt{T}} + \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\|e'e\|}{nT} \frac{\left\|\Theta\beta^0\right\|}{\sqrt{n}} \frac{\|F_T^*\|}{\sqrt{T}} \\
&\quad + \sum_{k=1}^{7} \|\mathcal{R}_k\| \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\left\|\Theta\beta^0\right\|}{\sqrt{n}} \frac{\|F_T^*\|}{\sqrt{T}} \\
&= O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) + \frac{1}{\sqrt{T}} O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) = o_p(1).
\end{aligned}$$

Similar to V_b, we have

$$\begin{aligned}
V_{cb} &\leq \left\|\frac{\bar{\beta}_K' e' f^0 \beta^{0\prime} \Theta E_T^*}{n^2 T\sqrt{T}}\right\| + \left\|\frac{\bar{\beta}_K' \beta^0 f^{0\prime} e \Theta E_T^*}{n^2 T\sqrt{T}}\right\| + \left\|\frac{\bar{\beta}_K' e' e \Theta E_T^*}{n^2 T\sqrt{T}}\right\| \\
&\quad + \sum_{k=1}^{7} \|\mathcal{R}_k\| \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\|\Theta E_T^*\|}{\sqrt{nT}} \\
&\leq \frac{2}{\sqrt{T}} \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\left\|e' f^0\right\|}{\sqrt{nT}} \frac{\left\|\beta^0\right\|}{\sqrt{n}} \frac{\|\Theta E_T^*\|}{\sqrt{nT}} + \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\|e'e\|}{nT} \frac{\|\Theta E_T^*\|}{\sqrt{nT}} + \sum_{k=1}^{7} \|\mathcal{R}_k\| \frac{\left\|\bar{\beta}_K\right\|}{\sqrt{n}} \frac{\|\Theta E_T^*\|}{\sqrt{nT}} \\
&= \frac{1}{\sqrt{T}} O_p(1) + O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) + \sqrt{\frac{1}{T}} O_p\left(\max\left\{\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}}\right\}\right) = o_p(1).
\end{aligned}$$

>From V_{ca} and V_{cb}, we have the required result that $V_c = o_p(1)$. Combining V_a, V_b, and V_c we have the required result for Part (v_2), and in consequence, we complete the proof of Part (v).

Finally, for Part (vi), notice by the triangle inequality that

$$\begin{aligned}
&\left\|\frac{\hat{\beta}_K' y' y \hat{\beta} - \beta_K^{*\prime} y' y \beta_K^*}{n\sqrt{n}T}\right\| = \left\|\frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' y \hat{\beta}_K + \beta_K^{*\prime} y' y \left(\hat{\beta}_K - \beta_K^*\right)}{n\sqrt{n}T}\right\| \\
&= \left\|\frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' y \left(\hat{\beta}_K - \beta_K^*\right) + \left(\hat{\beta}_K - \beta_K^*\right)' y' y \beta_K^* + \beta_K^{*\prime} y' y \left(\hat{\beta}_K - \beta_K^*\right)}{n\sqrt{n}T}\right\| \\
&\leq \left\|\frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' y \left(\hat{\beta}_K - \beta_K^*\right)}{n\sqrt{n}T}\right\| + 2\left\|\frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' y \beta_K^*}{n\sqrt{n}T}\right\| \\
&\leq \frac{1}{\sqrt{n}} \left\|\hat{\beta}_K - \beta_K^*\right\|^2 \frac{\|y\|^2}{nT} + 2\left\|\frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' y \beta^0}{n\sqrt{n}T}\right\| \|H_K\| \\
&= VI_a + VI_b, \text{ say.}
\end{aligned}$$

By Lemma 14(b) and Lemma 8(k), the first term

$$VI_a = \frac{1}{\sqrt{n}} O_p(1) = o_p(1).$$

Next we consider the term VI_b. Using (31), we expand $\left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' y \beta^0}{n\sqrt{nT}} \right\|$. Using $\frac{\|y'y\|}{nT} \leq \frac{\|y\|^2}{nT} = O_p(1)$ by Lemma 8(k), $\left\| \frac{\beta^0}{\sqrt{n}} \right\| = O(1)$, and (31), it is not difficult to find that all the expanded terms of $\left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)' y' y \beta^0}{n\sqrt{nT}} \right\|$ are of order $O_p\left(\sqrt{\frac{n}{T}}\right) = o_p(1)$ (under Assumption 10) except for the term $\left\| \frac{\bar{\beta}_K' e' e y' y \beta^0}{n^2\sqrt{n}T^2} \right\|$. Expand $\frac{\bar{\beta}_K' e' e y' y \beta^0}{n^2\sqrt{n}T^2}$ by substituting $f^0\beta^{0\prime} + e$ for y and apply the triangle inequality. Then,

$$\begin{aligned}
& \left\| \frac{\bar{\beta}_K' e' e y' y \beta^0}{n^2\sqrt{n}T^2} \right\| \\
\leq\ & \left\| \frac{\bar{\beta}_K' e' e \beta^0 f^{0\prime} f^0 \beta^{0\prime} \beta^0}{n^2\sqrt{n}T^2} \right\| + \left\| \frac{\bar{\beta}_K' e' e \beta^0 f^{0\prime} e \beta^0}{n^2\sqrt{n}T^2} \right\| + \left\| \frac{\bar{\beta}_K' e' e e' f^0 \beta^{0\prime} \beta^0}{n^2\sqrt{n}T^2} \right\| + \left\| \frac{\bar{\beta}_K' e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\| \\
\leq\ & \frac{\left\| \bar{\beta}_K' e' e \beta_K^* \right\|}{n\sqrt{n}T} \left\| H_K^{-1} \right\| \left\| \frac{f^{0\prime} f^0}{T} \right\| \left\| \frac{\beta^{0\prime} \beta^0}{n} \right\| \\
& + \frac{2}{\sqrt{T}} \left\| \frac{\bar{\beta}_K}{\sqrt{n}} \right\| \left\| \frac{\beta^0}{\sqrt{n}} \right\|^2 \left\| \frac{e'e}{nT} \right\| \left\| \frac{f^{0\prime} e}{\sqrt{n}\sqrt{T}} \right\| + \left\| \frac{\bar{\beta}_K' e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\| \\
=\ & o_p(1) O_p(1) O_p(1) + \frac{1}{\sqrt{T}} O_p(1) O(1) O_p\left(\max\left\{ \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{T}} \right\} \right) O_p(1) + \left\| \frac{\bar{\beta}_K' e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\|,
\end{aligned}$$

where the last line holds by Equation (36), Lemma 13, Assumption 7, Lemmas 8(e) and (f). For the remaining term $\left\| \frac{\bar{\beta}_K' e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\|$, using $\hat{\beta}_K = \bar{\beta}_K \Lambda_{nT,K}$ and $\beta_K^* = \beta^0 H_K$ and the triangle inequality, we have

$$\begin{aligned}
& \left\| \frac{\bar{\beta}_K' e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\| = \left\| \Lambda_{nT,K}^{-1} \frac{\hat{\beta}_K' e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\| \\
\leq\ & \left\| \Lambda_{nT,K}^{-1} \right\| \left\| \frac{\left(\hat{\beta}_K - \beta_K^*\right)' e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\| + \left\| \Lambda_{nT,K}^{-1} \right\| \|H_K\| \left\| \frac{\beta^{0\prime} e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right\| \\
\leq\ & \left\| \Lambda_{nT,K}^{-1} \right\| \left\| \hat{\beta}_K - \beta_K^* \right\| \left\| \frac{e'e}{nT} \right\|^2 \left\| \frac{\beta^0}{\sqrt{n}} \right\| + \left\| \Lambda_{nT,K}^{-1} \right\| \|H_K\| tr\left(\frac{\beta^{0\prime} e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right) \\
=\ & O_p(1) O_p(1) O_p\left(\max\left\{ \frac{1}{n}, \frac{1}{T} \right\} \right) O(1) + O_p(1) O_p(1) O_p\left(\frac{1}{\sqrt{n}} \right) = o_p(1),
\end{aligned}$$

where the last line holds by Lemmas 12, 14(b), 8(e), and 13, by Assumption 6, and by

$$tr\left(\frac{\beta^{0\prime} e' e e' e \beta^0}{n^2\sqrt{n}T^2} \right) = tr\left(\frac{e e' e \beta^0 \beta^{0\prime} e'}{n^2\sqrt{n}T^2} \right) \leq \frac{1}{\sqrt{n}} \frac{\|e\|^2}{nT} \frac{tr\left(\beta^{0\prime} e' e \beta^0\right)}{nT} = \frac{1}{\sqrt{n}} O_p(1) O_p(1)$$

due to $\sup_{it} E\left(e_{it}^2\right) < M$ and Lemma 8(g). From this, we deduce that $\left\| \frac{\bar{\beta}_K' e' e y' y \beta^0}{n^2\sqrt{n}T^2} \right\| = o_p(1)$, which leads to

$$VI_b = o_p(1).$$

Therefore,

$$\left\| \frac{\hat{\beta}_K^{'} y'y \hat{\beta}_K - \beta_K^{*\prime} y'y \beta_K^*}{n\sqrt{n}T} \right\| = o_p(1),$$

as required, and this completes the proof part (vi). ■

Proof of Lemma 5.
The proofs of Parts (a) and (b) are similar to those of Lemma 2(a) - (c), and we omit it.[10]

Part (c). Let $P_G = I_T - Q_G$. For a $T \times n$ matrix A, A_t' denotes the transpose of the t^{th} row of A. Then,

$$\begin{aligned} & \frac{1}{nT} tr\left(\left(\tilde{Z}_{-1}^0 - \tilde{Z}_{-1}^0(0) \right) Q_{\beta^o} \tilde{e}' \right) \\ = & \frac{1}{nT} tr\left(e' \left(Z_{-1}^0 - Z_{-1}^0(0) \right) Q_{\beta^o} \right) - \frac{1}{nT} tr\left(e' P_G \left(Z_{-1}^0 - Z_{-1}^0(0) \right) Q_{\beta^o} \right). \end{aligned}$$

In what follows we show that

$$\frac{1}{\sqrt{n}T} tr\left(e' \left(Z_{-1}^0 - Z_{-1}^0(0) \right) Q_{\beta^o} \right) = o_p(1), \tag{46}$$

and

$$\sqrt{n}\left[\frac{1}{nT} tr\left(e' P_G \left(Z_{-1}^0 - Z_{-1}^0(0) \right) Q_{\beta^o} \right) + \left(\frac{\mu_\theta}{n^\eta} \frac{1}{T^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) h_{kT}(s,t) \right) \right] = o_p(1), \tag{47}$$

Then, since $\frac{1}{T^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) h_{kT}(s,t) - \int_0^1 \int_0^r (r-s) h_k(r,s) \, dsdr = O\left(\frac{1}{T} \right)$, we have the required result.

First, for (46),

$$\begin{aligned} & \frac{1}{\sqrt{n}T} tr\left(e' \left(Z_{-1}^0 - Z_{-1}^0(0) \right) Q_{\beta^o} \right) \\ = & \frac{1}{\sqrt{n}T} \sum_{t=2}^{T} \left(Z_{t-1}^0 - Z_{t-1}^0(0) \right)' Q_{\beta^o} e_t \sim - \frac{\mu_\theta}{n^{1/2+\eta}T} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) y_s' Q_{\beta^o} e_t \\ = & - \frac{\mu_\theta}{n^{1/2+\eta}T} \sum_{i=1}^{n} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) e_{is} e_{it} \\ & + \frac{\mu_\theta}{n^{1/2+\eta}T} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right) \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} e_{is} \beta_i^{0\prime} \right) \left(\frac{1}{n} \sum_{i=1}^{n} \beta_i^{0\prime} \beta_i^0 \right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \beta_i^0 e_{it} \right) \\ = & I_a + II_a, \text{ say.} \end{aligned}$$

Since

$$\begin{aligned} E\left(I_a^2 \right) &= \frac{\mu_\theta^2}{n^{2\eta}T^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \sum_{p=2}^{T} \sum_{q=1}^{t-1} \left(\frac{t-s-1}{T} \right) \left(\frac{p-q-1}{T} \right) E\left(e_{is} e_{it} e_{ip} e_{iq} \right) \\ &= \frac{\mu_\theta^2}{n^{2\eta}T^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \left(\frac{t-s-1}{T} \right)^2 = O\left(\frac{1}{n^{2\eta}} \right) = o(1), \end{aligned}$$

[10] Upon requestion, the details of the proofs could be obtained from the authors.

we have

$$I_a = o_p(1).$$

Also, similarly, it is possible to show that

$$\begin{aligned} E\left(II_a^2\right) &= E\frac{\mu_\theta^2}{n^{1+2\eta}T^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{p=2}^{T}\sum_{q=1}^{t-1}\left(\frac{t-s-1}{T}\right)\left(\frac{p-q-1}{T}\right) \\ &\quad\times\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}e_{is}\beta_i^{0\prime}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i^{0\prime}\beta_i^0\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\beta_j^0 e_{jt}\right) \\ &\quad\times\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}e_{kq}\beta_k^{0\prime}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i^{0\prime}\beta_i^0\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{l=1}^{n}\beta_l^0 e_{lp}\right) \\ &= O\left(\frac{1}{n^{1+2\eta}}\right), \end{aligned}$$

so that

$$II_a = o_p(1),$$

which leads (46).

Next, for (47) we write

$$\begin{aligned} &\frac{1}{nT}tr\left(e'P_G\left(Z_{-1}^0 - Z_{-1}^0(0)\right)Q_{\beta^0}\right) \\ =\ &\frac{1}{nT}tr\left(e'G_{kT}\left(G_{kT}'G_{kT}\right)^{-1}G_{kT}'\left(Z_{-1}^0 - Z_{-1}^0(0)\right)Q_{\beta^0}\right) \\ =\ &\frac{1}{nT}tr\left(\sum_{t=1}^{T}e_t g_{kt}'\right)\left(\sum_{t=1}^{T}g_{kt}g_{kt}'\right)^{-1}\left(\sum_{t=1}^{T}g_{kt}\left(Z_{t-1}^0 - Z_{t-1}^0(0)\right)'Q_{\beta^0}\right) \\ \sim\ &-\frac{\mu_\theta}{n^{1+\eta}T}tr\left(\sum_{t=1}^{T}e_t g_{kt}'\right)\left(\sum_{t=1}^{T}g_{kt}g_{kt}'\right)^{-1}\left(\sum_{t=2}^{T}g_{kt}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)e_s'Q_{\beta^0}\right) \\ =\ &-\frac{\mu_\theta}{n^{1+\eta}T^2}\sum_{p=1}^{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)h_{kT}(p,t)\,e_s'Q_{\beta^0}e_p \\ =\ &-\frac{\mu_\theta}{n^{1+\eta}T^2}\sum_{p=1}^{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)h_{kT}(p,t)\left(\sum_{i=1}^{n}e_{is}e_{ip}\right) \\ &+\frac{\mu_\theta}{n^{1+\eta}T^2}\sum_{p=1}^{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)h_{kT}(p,t)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}e_{is}\beta_i^{0\prime}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i^{0\prime}\beta_i^0\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\beta_j^0 e_{jp}\right) \\ =\ &I_b + II_b, \text{ say.} \end{aligned}$$

A direct calculation shows that

$$E\left(\sqrt{n}II_b\right)^2 = O\left(\frac{1}{n^{1+2\eta}}\right),$$

which yields

$$\sqrt{n}II_b = o_p(1).$$

Also,

$$
\begin{aligned}
& E\left[\sqrt{n}\left\{I_b+\left(\frac{\mu_\theta}{n^\eta}\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)h_{kT}(s,t)\right)\right\}\right]^2 \\
= \; & E\left[\frac{\mu_\theta}{n^\eta T^2}\sum_{p=1}^{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)h_{kT}(p,t)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}e_{is}e_{ip}-1\right)\right]^2 = O\left(\frac{1}{n^{2\eta}}\right),
\end{aligned}
$$

so

$$
\sqrt{n}\left\{I_b+\left(\frac{\mu_\theta}{n^\eta}\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)h_{kT}(s,t)\right)\right\}=o_p(1),
$$

as required for Part (c). ■

Part (d). Here we show that

$$
\frac{\sqrt{n}}{n^\eta}\left[\frac{tr\left(\tilde{Z}_{-1}(0)\,Q_{\beta^0}\tilde{Z}'_{-1}(0)\right)}{nT^2}-\left(\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}-\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{T}\min\left(\frac{t-1}{T},\frac{s-1}{T}\right)h_{kT}(t,s)\right)\right]=o_p(1). \tag{48}
$$

Then, since

$$
\begin{aligned}
& \left(\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}-\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{T}\min\left(\frac{t-1}{T},\frac{s-1}{T}\right)h_{kT}(t,s)\right) \\
= \; & \int_0^1 r\,dr-\int_0^1\int_0^1\min(r,s)\,h(r,s)\,ds\,dr+O\left(\frac{1}{T}\right),
\end{aligned}
$$

the proof of Part (d) is done.

Since $\tilde{Z}_{-1}(0)=\tilde{F}^0_{-1}\beta^{0\prime}+\tilde{E}_{-1}$, we have

$$
\begin{aligned}
& \frac{1}{n^\eta\sqrt{n}T^2}tr\left(\tilde{Z}^0_{-1}(0)\,Q_{\beta^0}\tilde{Z}^0_{-1}(0)'\right) \\
= \; & \frac{1}{n^\eta\sqrt{n}T^2}tr\left(\tilde{E}_{-1}Q_{\beta^0}\tilde{E}'_{-1}\right) \\
= \; & \frac{1}{n^\eta\sqrt{n}T^2}tr\left(\tilde{E}_{-1}\tilde{E}'_{-1}\right)-\frac{1}{n^{1+\eta}\sqrt{n}T^2}tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\beta^{0\prime}\tilde{E}'_{-1}\tilde{E}_{-1}\beta^0\right)
\end{aligned}
$$

Using similar arguments in the proof of Lemma 8(c), we can show that

$$
\begin{aligned}
\frac{1}{n^{1+\eta}\sqrt{n}T^2}tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\beta^{0\prime}\tilde{E}'_{-1}\tilde{E}_{-1}\beta^0\right) &\le \frac{1}{n^{1/2+\eta}}\left\|\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\right\|\frac{\left\|\beta^{0\prime}\tilde{E}'_{-1}\right\|}{\sqrt{n}T}\frac{\left\|\tilde{E}_{-1}\beta^0\right\|}{\sqrt{n}T} \\
&= O_p\left(\frac{1}{n^{1/2+\eta}}\right).
\end{aligned}
$$

Therefore, to have the required result (48), it is enough to show

$$
\frac{\sqrt{n}}{n^\eta}\left[\frac{tr\left(\tilde{E}_{-1}\tilde{E}'_{-1}\right)}{nT^2}-\left(\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}-\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{T}\min\left(\frac{t-1}{T},\frac{s-1}{T}\right)h_{kT}(t,s)\right)\right]=o_p(1),
$$

which follows by a direct calculation that leads to

$$\begin{aligned} & E\left[\frac{\sqrt{n}}{n^{\eta}}\left\{\frac{tr\left(\tilde{E}_{-1}\tilde{E}_{-1}'\right)}{nT^2}-\left(\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}-\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{T}\min\left(\frac{t-1}{T},\frac{s-1}{T}\right)h_{kT}\left(t,s\right)\right)\right\}\right]^2 \\ = \quad & O\left(\frac{1}{n^{2\eta}}\right)=o\left(1\right), \end{aligned}$$

and we have all the required result.

Part (e). Notice that

$$\left\|Z_{-1}^{0}-Z_{-1}^{0}\left(0\right)\right\|=O_p\left(Tn^{1/2-\eta}\right)$$

because

$$\begin{aligned} & \frac{1}{n^{1-2\eta}T^2}\left\|Z_{-1}^{0}-Z_{-1}^{0}\left(0\right)\right\|^2=\frac{1}{n^{1-2\eta}T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(z_{it-1}^{0}-z_{it-1}^{0}\left(0\right)\right)^2 \\ \sim \quad & \frac{\mu_{\theta}^2}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)y_{is}\right)^2=O_p\left(1\right) \end{aligned}$$

by Lemma 9(a) and (b). Then, the required result follows because for some constant $M>0$,

$$\frac{1}{n^{1/2+\eta}T^2}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^0}\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)'\right)\leq n^{1/2-3\eta}M\frac{\left\|Z_{-1}^{0}-Z_{-1}^{0}\left(0\right)\right\|^2}{n^{1-2\eta}T^2}=O_p\left(n^{1/2-3\eta}\right)$$

and since $1/6<\eta$ in both model $k=0$ and $k=1$, $n^{1/2-3\eta}=o\left(1\right)$, as required. ■

Part (f). Define

$$E_{it}\left(\mu_{\theta}\right)=\sum_{s=1}^{t}\left(1-\frac{\mu_{\theta}}{n^{\eta}T}\right)^{t-s}e_{is}.$$

Let

$$\begin{aligned} E_t\left(\mu_{\theta}\right)' &= \left(E_{1t}\left(\mu_{\theta}\right),...,E_{nt}\left(\mu_{\theta}\right)\right), \\ E_{-1}\left(\mu_{\theta}\right) &= \left(0',E_1\left(\mu_{\theta}\right)',...,E_{T-1}\left(\mu_{\theta}\right)'\right)'. \end{aligned}$$

By definition,

$$E_{-1}\left(0\right)=E_{-1}.$$

Using this notation, write

$$\begin{aligned} & \frac{1}{n^{1/2+\eta}T^2}tr\left(\left(\tilde{Z}_{-1}^{0}-\tilde{Z}_{-1}^{0}\left(0\right)\right)Q_{\beta^0}\tilde{Z}_{-1}'\left(0\right)\right) \\ = \quad & \frac{1}{n^{1/2+\eta}T^2}tr\left(Q_G\left(E_{-1}\left(\mu_{\theta}\right)-E_{-1}\right)Q_{\beta^0}E_{-1}'Q_G\right) \\ = \quad & \frac{1}{n^{1/2+\eta}T^2}tr\left(E_{-1}'Q_G\left(E_{-1}\left(\mu_{\theta}\right)-E_{-1}\right)\right)-\frac{1}{n^{1/2+\eta}T^2}tr\left(Q_G\left(E_{-1}\left(\mu_{\theta}\right)-E_{-1}\right)P_{\beta^0}E_{-1}'Q_G\right) \\ = \quad & I_f-II_f, \text{ say.} \end{aligned}$$

Then,

$$\begin{aligned}
II_f &= \frac{1}{n^{3/2+\eta}T^2}tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\beta^{0\prime}E'_{-1}Q_G\left(E_{-1}\left(\mu_\theta\right)-E_{-1}\right)\beta^0\right)\\
&= \frac{1}{n^{3/2+\eta}T^2}tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\beta^{0\prime}E'_{-1}\left(E_{-1}\left(\mu_\theta\right)-E_{-1}\right)\beta^0\right)\\
&\quad -\frac{1}{n^{3/2+\eta}T^2}tr\left(\left(\frac{\beta^{0\prime}\beta^0}{n}\right)^{-1}\beta^{0\prime}E'_{-1}G_{kT}\left(G'_{kT}G_{kT}\right)^{-1}G'_{kT}\left(E_{-1}\left(\mu_\theta\right)-E_{-1}\right)\beta^0\right)\\
&= -\frac{\mu_\theta}{n^{3/2+2\eta}T^2}tr\left(\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i^0\beta_i^{0\prime}\right)^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_i^0\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)E_{it-1}e_{js}\beta_j^{0\prime}\right)\\
&\quad +\frac{\mu_\theta}{n^{3/2+2\eta}T^3}tr\left(\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i^0\beta_i^{0\prime}\right)^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=2}^{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\beta_i^0E_{ip-1}h_{k,T}\left(p,t\right)\left(\frac{t-s-1}{T}\right)e_{js}\beta_j^{0\prime}\right)\\
&= II_{fa}-II_{fb},\text{ say.}
\end{aligned}$$

A direct calculation shows that

$$II_{fa}=O_p\left(\frac{1}{n^{1+2\eta}}\right)=o_p\left(1\right)\text{ and }II_{fb}=O_p\left(\frac{1}{n^{1+2\eta}}\right)=o_p\left(1\right).$$

Also, for I_f,

$$\begin{aligned}
I_f &= -\frac{\mu_\theta}{n^{1/2+2\eta}T^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)e_{is}\right)E_{it-1}\\
&\quad +\frac{\mu_\theta}{n^{1/2+2\eta}T^3}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)e_{is}\right)\sum_{p=2}^{T}E_{ip-1}h_{k,T}\left(t,p\right)
\end{aligned}$$

Notice that

$$\begin{aligned}
&E\left(\frac{1}{T^2}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)e_{is}\right)E_{it-1}-\frac{1}{T^3}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)e_{is}\right)\sum_{p=2}^{T}E_{ip-1}h_{k,T}\left(t,p\right)\right)\\
&= \frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left(\frac{t-s-1}{T}\right)-\frac{1}{T^3}\sum_{t=2}^{T}\sum_{p=2}^{T}h_{k,T}\left(t,p\right)\sum_{s=1}^{\min(t,p)-1}\left(\frac{t-s-1}{T}\right)\\
&= \int_0^1\int_0^r\left(r-s\right)dsdr-\int_0^1\int_0^1h_k\left(r,p\right)\int_0^{\min(r,p)}\left(r-s\right)dsdpdr+O\left(\frac{1}{T}\right).
\end{aligned}$$

Notice that a direct calculation shows that

$$\int_0^1\int_0^r\left(r-s\right)dsdr-\int_0^1\int_0^1h_k\left(r,p\right)\int_0^{\min(r,p)}\left(r-s\right)dsdpdr=\frac{1}{8}\text{ for }k=0\text{ (so that }h_k\left(r,p\right)=1)$$

and

$$\int_0^1\int_0^r\left(r-s\right)dsdr-\int_0^1\int_0^1h_k\left(r,p\right)\int_0^{\min(r,p)}\left(r-s\right)dsdpdr=0\text{ for }k=1.$$

Then, for model $k=0$, it is possible to show that

$$I_f = -\mu_\theta n^{1/2-2\eta} \begin{bmatrix} \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\left(\frac{t-s}{T}\right)e_{is}\right)E_{it-1} \\ -\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\left(\frac{t-s}{T}\right)e_{is}\right)\sum_{p=2}^{T}E_{ip-1}h_{k,T}(t,p) \end{bmatrix} = O_p\left(n^{1/2-2\eta}\right) = o_p(1)$$

because $\eta > \frac{1}{4}$, and for model $k=1$, it is possible to show that

$$E\left(I_f^2\right) = O\left(\frac{1}{n^{4\eta}}\right) = o(1),$$

so

$$I_f = o_p(1),$$

as required. ■

Proof of Theorem 3

The proof is quite similar to the proof of Theorem 2 in Bai and Ng (2002), and we omit the proof. The details of the proof could be found in Moon and Perron (2002). ■

References

[1] Anderson, T. (1963): Asymptotic Theory for Principal Component Analysis, *Annals of Mathematical Statistics*, 34, 122–148.

[2] Andrews, D. (1991): Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, 59, 817–858.

[3] Andrews, D. and C. Monahan (1992): An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator, *Econometrica*, 60, 953-966.

[4] Bai, J. (2001): Inference on Factor Models of Large Dimensions, *Mimeo.*

[5] Bai, J. and S. Ng (2001): A *PANIC* Attack on Unit Roots and Cointegration, *Mimeo,*

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

[7] Bernanke, B. and J. Boivin (2002): Monetary Policy in a Data Rich Environment, *Journal of Monetary Economics*, forthcoming..

[8] Brisson, M., B. Campbell, and J. W. Galbraith (2001): Forecasting Some Low-Predictability Time Series Using Diffusion Indices, CIRANO working paper 2001s-46.

[9] Chang, Y. (2001): Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency, *forthcoming in Journal of Econometrics.*

[10] Chen, X. and T. Conley (2001): A New Semiparametric Spatial Model for Panel Time Series, *Journal of Econometrics*, 105, 59–84.

[11] Choi, I. (2001a): Unit Root Tests for Panel Data, *Journal of International Money and Finance*, 20, 249-272.

[12] Choi, I. (2001b): Unit Root Tests for Cross-Sectionally Correlated Panels, *Mimeo.*

[13] Conley, T. (1999): GMM Estimation with Cross Sectional Dependence, *Journal of Econometrics*, 92, 1–45.

[14] Connor, G. and R. Korajczyk (1986): Performance Measurement with the Arbitrage Pricing Theory: A New Test Methodology, *Journal of Financial Economics,* 15, 373–394.

[15] Connor, G. and R. Korajczyk (1988): Risk and Return in an Equilibrium APT: Application to a New Test Methodology, *Journal of Financial Economics*, 21, 255–289.

[16] Connor, G. and R. Korajczyk (1993): A Test for the Number of Factors in an Approximate Factor Model, *Journal of Finance*, XLVIII:4, 1263–1291.

[17] Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): The Generalized Dynamic Factor Model: Identification and Estimation, *Review of Economics and Statistics*, 82, 540–554.

[18] Hannan, E. (1970): *Multiple Time Series*, New York, Wiley.

[19] Im, K.S., M.H. Pesaran, and Y. Shin (1997): Testing for Unit Roots in Heterogeneous Panels, *Mimeo.*

[20] Levin, A., F. Lin, and C. Chu (2002): Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties, *Journal of Econometrics*, 108, 1-24.

[21] Maddala, G.S. and S. Wu (1999): A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

[22] Moon, H.R. and B. Perron (2002): Appendix: Omitted Proofs of "Testing for A Unit Root in Panels with Dynamic Factors", *CLEO Discussion Papers*, University of Southern California.

[23] Moon, H.R., B. Perron, and P.C.B. Phillips (2002): Power Comparison of Panel Unit Root Tests, *Mimeo*.

[24] Moon, H.R. and P.C.B. Phillips (1999): Maximum Likelihood Estimation in Panels with Incidental Trends, *Oxford Bulletin of Economics and Statistics*, 61, 711-747.

[25] Moon, H.R. and P.C.B. Phillips (2000): Estimation of Autoregressive Roots Near Unity Using Panel Data, *Econometric Theory*, 16, 927–992.

[26] Moon, H.R. and P.C.B. Phillips (2002): GMM Estimation of Autoregressive Roots Near Unity with Panel Data, *Mimeo*.

[27] O'Connell, P. (1998): The Overvaluation of Purchasing Power Parity, *Journal of International Economics*, 44, 1–19.

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

[29] Phillips, P.C.B. and V. Solo (1992): Asymptotics for Linear Processes, *Annals of Statistics*, 20, 971-1001.

[30] Phillips, P.C.B. and D. Sul (2002): Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Mimeo*.

[31] Ploberger, W. and P.C.B. Phillips (2002): Optimal Testing for Unit Roots in Panel Data, *mimeo*.

[32] Quah, D. (1994): Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Panels, *Economics Letters*, 44, 9-19.

[33] Stock, J. and M. Watson (1998): Diffusion Indexes, *NBER Working Paper 6702*.

Table 1. Size of tests for experiment 1

DGP: $z_{it} = \alpha_{i0} + z_{it}^0$

$z_{it}^0 = z_{it-1}^0 + \tau \sum_{j=1}^{K} \beta_{ij} f_{tj} + \sqrt{K} e_{it}$

$\alpha_{i0}, \beta_{ij}, f_{tj}, e_{it} \sim iidN(0,1)$

$K = 1$

			t_a^*				t_b^*				$\widehat{K}$	
	n	T	True K	IC_i	BIC_3	$\widehat{K}=0$	True K	IC_i	BIC_3	$\widehat{K}=0$	IC_i	BIC_3
	10	100	13.6	40.6	34.6	22.4	8.0	34.0	27.7	14.1	8.00	6.75
$\tau=1$	20	100	10.4	10.4	10.4	26.3	7.0	7.0	7.0	18.2	1.00	1.00
	10	300	14.1	38.5	21.2	24.5	8.0	29.6	13.7	14.1	8.00	3.25
	20	300	11.1	11.1	11.1	28.5	7.3	7.3	7.3	20.3	1.00	1.00
	10	100	20.7	59.7	57.0	46.3	11.5	51.6	49.5	37.1	8.00	6.78
$\tau=10$	20	100	20.5	30.3	26.2	49.1	14.2	26.3	21.5	44.4	1.18	1.12
	10	300	16.5	58.9	43.1	46.0	8.6	50.5	36.0	36.0	8.00	3.30
	20	300	12.6	15.7	13.6	49.9	8.1	12.1	9.7	44.9	1.05	1.02
	10	100	13.9	36.1	30.4	13.9	7.6	27.1	20.5	7.6	8.00	6.71
$\tau=0$	20	100	9.7	12.7	12.7	9.7	6.8	9.1	9.1	6.8	1.00	1.00
	10	300	12.5	35.9	17.0	12.5	6.6	28.2	11.4	6.6	8.00	2.97
	20	300	8.5	10.9	10.9	8.5	5.0	7.2	7.2	5.0	1.00	1.00

Note: Each entry represents the percentage of replications in which the null hypothesis of a unit root is rejected for the appropriate test with the number of factors either set to the true number (1), estimated using the information criteria suggested by Bai and Ng (2002) , or set to 0 The last two columns provide the mean number of estimated factors with $\overline{K} = 8$ for both criteria considered The number of replications is 1000.

Table 2. Power of tests for experiment 1

DGP: $z_{it} = \alpha_{i0} + z_{it}^0$

$z_{it}^0 = \rho_i z_{it-1}^0 + \tau \sum_{j=1}^{K} \beta_{ij} f_{tj} + \sqrt{K} e_{it}$

$\alpha_{i0}, \beta_{ij}, f_{tj}, e_{it} \sim iidN(0,1)$

$\rho_i \sim U[0.98, 1]$

$K = 1$

			t_a^*						t_b^*					
	n	T	True K		IC_i		BIC_3		True K		IC_i		BIC_3	
	10	100	56.6	33.3	64.4	6.9	63.3	10.8	44.5	34.2	53.7	6.8	52.4	11.2
$\tau=1$	20	100	74.4	57.6	74.4	57.6	74.4	57.6	64.8	57.7	64.8	57.7	64.8	57.7
	10	300	89.5	77.8	83.7	25.3	88.2	67.8	82.9	77.6	75.7	24.6	83.6	64.7
	20	300	96.6	93.8	96.6	93.8	96.6	93.8	94.5	93.8	94.5	93.8	94.5	93.8
	10	100	26.3	10.0	74.8	6.1	73.2	7.1	18.0	9.7	68.5	6.4	66.6	6.3
$\tau=10$	20	100	26.4	13.9	63.6	6.6	46.5	5.8	23.0	14.4	61.0	5.9	43.7	6.2
	10	300	25.4	15.5	84.5	15.2	82.1	16.2	21.7	16.9	78.4	15.0	77.5	16.1
	20	300	23.4	16.5	73.0	23.4	40.2	31.5	21.9	17.1	70.4	23.7	38.3	31.2
	10	100	62.4	38.1	59.4	10.6	59.5	18.8	47.7	40.5	51.0	10.8	50.0	17.8
$\tau=0$	20	100	80.0	65.5	77.2	56.8	77.2	56.8	71.7	64.5	69.6	57.2	69.6	57.2
	10	300	96.1	90.2	84.5	25.2	90.9	75.5	92.0	89.8	77.3	25.9	87.0	74.7
	20	300	99.2	98.9	99.4	98.5	99.4	98.5	98.8	98.7	98.7	98.1	98.7	98.1

Note: The first entry in each cell represents the percentage of replications in which the null hypothesis of a unit root is rejected for the appropriate test with the number of factors either set to the true number (1), estimated using the information criteria suggested by Bai and Ng (2002), or set to 0 using the asymptotic critical values The second entry in each cell is the corresponding size-adjusted power. The number of replications is 1000.

Table 3. Size of tests for experiment 2

DGP: $z_{it} = \alpha_{i0} + \alpha_{i1} t + z_{it}^0$

$z_{it}^0 = z_{it-1}^0 + \tau \sum_{j=1}^{K} \beta_{ij} f_{tj} + \sqrt{K} e_{it}$

$\alpha_{i0}, \alpha_{i1}, \beta_{ij}, f_{tj}, e_{it} \sim iidN(0,1)$

$K = 1$

			t_a^*				t_b^*				$\widehat{K}$	
	n	T	True K	IC_i	BIC_3	$\widehat{K}=0$	True K	IC_i	BIC_3	$\widehat{K}=0$	IC_i	BIC_3
	10	100	13.2	45.3	45.5	18.7	11.5	38.5	39.5	18.2	8.00	6.80
$\tau = 1$	20	100	18.5	19.0	18.5	26.5	16.7	17.2	16.7	24.9	1.01	1.00
	10	300	8.7	27.2	18.3	11.4	8.6	24.2	16.0	11.2	8.00	3.28
	20	300	6.9	6.9	6.9	16.0	6.8	6.8	6.8	15.7	1.00	1.00
	10	100	11.0	51.6	55.2	26.4	10.0	44.1	47.1	23.4	8.00	6.78
$\tau = 10$	20	100	14.0	23.5	17.1	32.7	12.7	20.2	14.9	29.2	1.15	1.06
	10	300	7.6	30.1	21.1	16.3	6.8	24.9	17.2	14.8	8.00	3.36
	20	300	6.4	6.5	6.4	21.8	6.4	6.5	6.4	19.2	1.01	1.00
	10	100	11.6	47.5	46.9	11.6	10.6	41.0	41.9	10.6	8.00	6.70
$\tau = 0$	20	100	15.6	38.2	38.2	15.6	14.0	34.5	34.5	14.0	1.00	1.00
	10	300	5.8	25.4	15.4	5.8	5.5	22.7	13.6	5.5	8.00	2.97
	20	300	8.1	13.0	13.0	8.1	8.5	12.0	12.0	8.5	1.00	1.00

Note: see table 1.

Table 4. Power of tests for experiment 2

DGP: $z_{it} = \alpha_{i0} + \alpha_{i1} t + z_{it}^0$

$z_{it}^0 = \rho_i z_{it-1}^0 + \tau \sum_{j=1}^{K} \beta_{ij} f_{tj} + \sqrt{K} e_{it}$

$\alpha_{i0}, \alpha_{i1}, \beta_{ij}, f_{tj}, e_{it} \sim iidN(0,1)$

$\rho_i \sim U[0.98, 1]$

$K = 1$

			t_a^*						t_b^*					
	n	T	True K		IC_i		BIC_3		True K		IC_i		BIC_3	
	10	100	12.3	4,5	43.1	5.9	46.6	5.3	9.8	4.3	35.1	4.9	38.5	5.7
$\tau = 1$	20	100	18.1	5.0	18.3	4.4	18.1	5.0	15.9	4.5	16.1	4.2	15.9	4.5
	10	300	4.0	2.1	23.1	3.6	12.1	3.7	2.7	1.8	18.2	3.8	8.3	2.5
	20	300	5.4	2.8	5.4	2.8	5.4	2.8	3.9	2.2	3.9	2.2	3.9	2.2
	10	100	12.9	7.5	57.0	4.6	58.7	5.0	13.4	8.4	47.4	6.0	51.1	5.9
$\tau = 10$	20	100	17.3	10.7	50.1	15.5	34.2	21.7	17.8	12.8	46.4	20.4	32.3	22.0
	10	300	9.5	8.5	31.1	5.5	23.2	6.1	14.9	13.9	24.9	3.9	18.3	6.2
	20	300	16.9	14.6	28.0	23.6	22.4	19.7	22.1	21.1	25.6	23.3	25.2	23.4
	10	100	12.0	7.7	42.7	4.1	46.5	5.2	10.6	7.1	37.1	4.5	39.0	4.0
$\tau = 0$	20	100	17.9	6.0	38.4	5.8	38.4	5.8	15.9	5.1	33.6	5.3	33.6	5.3
	10	300	4.1	3.9	24.0	3.6	11.7	2.9	2.7	2.6	18.6	2.4	8.0	1.9
	20	300	4.9	2.5	8.8	3.1	8.8	3.1	3.5	1.5	6.7	1.7	6.7	1.7

Note: see table 2.