ISSN: 2281-1346

UNIVERSITÀ
DI PAVIA

Department of Economics and Management

DEM Working Paper Series

Testing for no factor structures: on the use of average-type and Hausman-type statistics

Carolina Castagnetti
(Università di Pavia)

Eduardo Rossi
(Università di Pavia)

Lorenzo Trapani
(City University, London)

 # 92 (10-14)

Via San Felice, 5
I-27100 Pavia
http://epmq.unipv.eu/site/home.html

October 2014

Testing for no factor structures: on the use of average-type and Hausman-type statistics.

Carolina Castagnetti
University of Pavia

Eduardo Rossi
University of Pavia

Lorenzo Trapani
Cass Business School, City University London

September 30, 2014

Abstract

Castagnetti, Rossi and Trapani (2014) propose two max-type statistics to test for the presence of a factor structure in a large stationary panel data model. We investigate the use of alternative approaches as average-type and Hausman-type statistics. We show that both approaches can not be used. The average-type statistics diverge under the null, while the tests based on the Hausman principle are inconsistent.

JEL codes: C12, C33.

Keywords: Large Panels, Testing for Factor Structure, Hausman-type test.

1 Introduction

Consider the following panel data model:

$$y_{it} = \beta_i' x_{it} + \gamma_i' f_t + \epsilon_{it}, \tag{1}$$

$$x_{it} = \Lambda_i f_t + \epsilon_{it}^x, \tag{2}$$

where $i = 1, ..., n$, $t = 1, ..., T$, x_{it} is an m-dimensional vector of observable explanatory variables and f_t is an r-dimensional vector of unobservable common factors; in equation (2), Λ_i is a matrix of coefficients of dimension $m \times r$. The specification above is based on Pesaran (2006) and represents a very general setup which includes many panel data models as special cases. The inferential theory on the slope coefficients β_i has been developed in several contributions. Pesaran (2006) proposes the class of Common Correlated

Effects (CCE) estimators. Bai (2009) introduces the Interactive Effect (IE) estimator for the case of homogeneous slope ($\beta_i = \beta$) which has been extended by Song (2013) to the case of heterogenous slopes.

The inferential procedures developed by Castagnetti, Rossi and Trapani (2014) (CRT) are built upon a two stage estimation procedure for f_t and γ_i based on a consistent (first stage) estimator for β_i and the Principal Component estimator. As a first stage estimator for β_i, either the Pesaran (2006) CCE estimator, or the Song (2013) estimator, may be used. We refer the reader to CRT for the two-stage estimation procedure and the asymptotics of $\hat{\gamma}_i$ and $\hat{f}_t$. In particular, throughout the paper, we keep Assumptions 1-5 and we rely on Theorem 1 and Theorem 2 in CRT.

Formally, CRT propose two tests for the null hypotheses:

$$H_0^a \quad : \quad \gamma_i = \gamma \text{ for all } i; \tag{3}$$

$$H_0^b \quad : \quad f_t = f \text{ for all } t. \tag{4}$$

Both tests can be used to verify whether a factor structure in (1)-(2) indeed exists, or whether simpler specifications nested in (1)-(2) should be employed. Under H_0^a, model (1) becomes the panel data model with time effect while under H_0^b, model (1) is equivalent to the familiar panel data model with a unit specific effect. From a methodological point of view, CRT propose two max-type statistics. When H_0^a is considered, the maximum is taken over the deviation of the individual estimate of γ_i with respect to their cross-sectional average. When H_0^b is considered, the maximum is taken over the deviation of the estimate of f_t with respect to their time series average.

In this paper we show that the average-type and Hausman-type statistics to test for (3) and (4) in (1)-(2) cannot be used. The average-type statistics diverge under the null as $(n, T) \to \infty$, while the tests based on the Hausman principle are inconsistent.

The tests based on average-type statistics are presented in Section 2 while Section 3 shows the results for the Hausman-type statistics. We refer to the Appendix for all the proofs.

2 Tests based on average-type statistics

Pesaran and Yamagata (2008) suggest using averages of F-statistics in order to test for the null of slope homogeneity in a model with observable regressors, viz.

$$\tilde{S}_{\gamma,nT} = \sqrt{\frac{n}{2r}}\frac{1}{n}\sum_{i=1}^{n}\left[T\left(\hat{\gamma}_i-\widehat{\overline{\gamma}}\right)'\hat{\Sigma}_{\gamma i}^{-1}\left(\hat{\gamma}_i-\widehat{\overline{\gamma}}\right)-r\right], \tag{5}$$

$$\tilde{S}_{f,nT} = \sqrt{\frac{T}{2r}}\frac{1}{T}\sum_{t=1}^{T}\left[n\left(\hat{f}_t-\widehat{\overline{f}}\right)'\hat{\Sigma}_{ft}^{-1}\left(\hat{f}_t-\widehat{\overline{f}}\right)-r\right]. \tag{6}$$

where $\widehat{\overline{\gamma}} = n^{-1}\sum_{i=1}^{n}\hat{\gamma}_i$ and $\widehat{\overline{f}} = T^{-1}\sum_{t=1}^{T}\hat{f}_t$. $\hat{\Sigma}_{\gamma i}$ and $\hat{\Sigma}_{ft}$ are the estimator for the asymptotic variance of $\hat{\gamma}_i$ and $\hat{f}_t$ given by equations (7) and (10) in CRT, respectively.

We show that $\tilde{S}_{\gamma,nT}$ and $\tilde{S}_{f,nT}$ cannot be employed in our context: in essence, this is because $\tilde{S}_{\gamma,nT}$ and $\tilde{S}_{f,nT}$ diverge under the null as $(n,T)\to\infty$, so that tests based on (5) and (6) always reject the null of no factor structure. Results are summarized in the following Theorem:

Theorem 1 *Let Assumptions 1-4 in CRT hold.*

1. *If, in addition, as $(n,T)\to\infty$*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\epsilon_i' M_{xi}F\Sigma_{fMe,i}^{-1}F'M_{Xi}\epsilon_i - r\right] = O_p(1), \tag{7}$$

then, under H_0^a it holds that $\tilde{S}_{\gamma,nT} = O_p(1) + O_p\left(\sqrt{\frac{T}{n}}\right) + O_p\left(\sqrt{\frac{n}{T}}\right)$.

2. *If, in addition, as $(n,T)\to\infty$*

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_i'\left(\frac{\hat{F}'F}{T}\right)\left(\frac{F'\hat{F}}{T}\right)\gamma_j\epsilon_{it}\epsilon_{jt} - r\right] = O_p(1), \tag{8}$$

then, under H_0^b it holds that $\tilde{S}_{f,nT} = O_p(1) + O_p(\sqrt{n}) + O_p\left(\frac{n}{\sqrt{T}}\right) + O_p\left(\frac{\sqrt{T}}{n}\right)$.

Theorem 1 shows that, under the respective null hypotheses, both average-type statistics diverge, and therefore cannot be employed.

3 Tests based on the Hausman principle

Building on Bai (2009, Section 9), tests could be constructed indirectly using a pooled estimator of the β_is. In order to illustrate the idea, define the average slope $\beta = E(\beta_i)$. Estimation of β could be based on pooling the estimates of the individual β_is:

$$\hat{\beta}^{CCE/IE} = \frac{1}{n}\sum_{i=1}^{n}\tilde{\beta}_i^{CCE/IE}.$$

We use the notation $\hat{\beta}^{CCE}$ and $\hat{\beta}^{IE}$ according as the $\tilde{\beta}_i$s are computed using the individual CCE estimators (Pesaran, 2006) or the individual IE estimators (see Song, 2013) respectively. One can expect that under either null H_0^a and H_0^b, both the CCE and the IE estimators are consistent, since no assumption for the consistency of either estimator is violated. The Hausman principle can therefore be applied upon finding another estimator which is consistent, and more efficient, under the null - Bai (2009) points out that, in the context of slope homogeneity, estimators based on the "between" and "within" transformation should be more efficient under the null.

Testing for $H_0^a : \gamma_i = \gamma$

Under H_0^a, an alternative estimator for β is

$$\hat{\beta}^{bw} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{y}_{it}\right),$$

with $\dot{x}_{it} = x_{it} - n^{-1}\sum_{i=1}^{n}x_{it}$ and $\dot{y}_{it} = y_{it} - n^{-1}\sum_{i=1}^{n}y_{it}$; this is the Mean-Group version of the "between" estimator, as also suggested in Bai (2009). It can be expected that, under H_0^a, $\hat{\beta}^{bw}$ is consistent and should be more efficient than $\hat{\beta}^{CCE}$ and $\hat{\beta}^{IE}$. Hence, tests for H_0^a could be based on

$$S_{\gamma,nT}^{IE/CCE} = n\left(\hat{\beta}^{IE/CCE} - \hat{\beta}^{bw}\right)'\left[\text{Var}\left(\hat{\beta}^{IE/CCE} - \hat{\beta}^{bw}\right)\right]^{-1}\left(\hat{\beta}^{IE/CCE} - \hat{\beta}^{bw}\right).$$

Let the critical value $c_{\alpha,n}$ be defined such that $P\left(S_{\gamma,nT}^{IE/CCE} \leq c_{\alpha,n}\right) = 1-\alpha$ under H_0^a. It holds that:

Theorem 2 *Let Assumptions 1, 2, 3(i)-(ii) and 4 in CRT hold. As* $(n,T)\to\infty$ *with* $\frac{\sqrt{n}}{T}\to 0$, *under* H_0^a, $S_{\gamma,nT}^{IE} \xrightarrow{d} \chi_m^2$. *Assume further that* γ_i *is i.i.d. (and independent of all other quantities) with mean* γ *and* $E\|\gamma_i\|^{2+\delta} < \infty$. *Then, under the alternative* $H_1^a : \gamma_i \neq \gamma_j$ *for* $i \neq j$, *as* $(n,T)\to\infty$ *it holds that* $P\left(S_{\gamma,nT}^{IE} > c_{\alpha,n}\right) < 1$. *The same results holds for* $S_{\gamma,nT}^{CCE}$ *as* $\min\{n,T\}\to\infty$.

Theorem 2 is, in essence, a negative result. It is possible to construct a test statistic that does not diverge under the null, and which has a "standard" limiting distribution - this can be contrasted with Theorem 3 below. However, the test is inconsistent, i.e. the power does not tend to 1 as the sample size passes to infinity. Heuristically, this is due to the fact that, under the alternative, the estimation error of

$\hat{\beta}^{bw}$ (rescaled by $\sqrt{n}$) has the extra term

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}f_t'\left(\gamma_i-\gamma\right)\right];$$

under the quite standard (see e.g. Assumption 3 in Pesaran, 2006) random coefficients assumption for γ_i, such term has the same order of magnitude as the leading term (thus ruling out power versus local alternatives), and it does not converge to a constant; rather, it can be shown to converge to a normally distributed random variable. This has the effect of inflating the variance of $\sqrt{n}\left(\hat{\beta}^{IE}-\hat{\beta}^{bw}\right)$, but it does not introduce any non-centrality parameter that would diverge under alternatives, whence the result in the theorem.

Testing for $H_0^b : f_t = f$

Under H_0^b, β can be estimated as

$$\hat{\beta}^{FE}=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}\bar{x}_{it}\bar{x}_{it}'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\bar{x}_{it}\bar{y}_{it}\right),$$

where $\bar{x}_{it}=x_{it}-T^{-1}\sum_{t=1}^{T}x_{it}$ and $\bar{y}_{it}=y_{it}-T^{-1}\sum_{t=1}^{T}y_{it}$; $\hat{\beta}^{FE}$ is the Mean-Group version of the "within" estimator. Based on this, testing for H_0^b could be done using either

$$S_{f,nT}^{IE/CCE}=nT\left(\hat{\beta}^{IE/CCE}-\hat{\beta}^{FE}\right)'\left[Var\left(\hat{\beta}^{IE/CCE}-\hat{\beta}^{FE}\right)\right]^{-1}\left(\hat{\beta}^{IE/CCE}-\hat{\beta}^{wn}\right).$$

It holds that

Theorem 3 *Let Assumptions 1-4 in CRT hold. As $(n,T)\to\infty$, under H_0^b*

$$S_{f,nT}^{IE} = O_p(1)+O_p\left(\sqrt{\frac{T}{n}}\right)+O_p\left(\sqrt{\frac{n}{T}}\right), \tag{9}$$

$$S_{f,nT}^{CCE} = O_p(1)+O_p\left(\sqrt{\frac{T}{n}}\right). \tag{10}$$

More specifically, as far as $S_{f,nT}^{IE}$ is concerned, equation (9) states that Hausman-type tests based on the IE estimator cannot be employed, as they always diverge under the null. The reason is that, in the expansion of $\tilde{\beta}_i^{IE}-\beta_i$, there are terms of order $O_p\left(n^{-1}\right)+O_p\left(T^{-1}\right)$, which do not get averaged out when calculating the cross-sectional averages. Thus, the impact of such terms on $\sqrt{nT}\left(\hat{\beta}^{IE}-\beta\right)$ is of order $O_p\left(\sqrt{\frac{n}{T}}\right)+O_p\left(\sqrt{\frac{T}{n}}\right)$, which diverges as $(n,T)\to\infty$. As far as $S_{f,nT}^{CCE}$ is concerned, equation (10) states that $S_{f,nT}^{CCE}$ could potentially be employed, at least under the restriction that $\frac{T}{n} \to 0$. As we point out in the proof in Appendix, the problem with this approach is that, in general, the distribution of the $O_p(1)$ term is degenerate, and it anyway depends on several nuisance parameters in the DGP of the x_{it}s, and on f_t and γ_i. In essence, equation (10) states that testing for no factor structure using $S_{f,nT}^{CCE}$ is fraught with difficulties and, in general, not feasible.

References

Bai, J., 2009, Panel data models with interactive fixed effects. Econometrica, vol. 77, 1229-1279.

Castagnetti, C., Rossi, E., Trapani, L., 2014, Inference on Factor Structures in Heterogeneous Panels. Journal of Econometrics (forthcoming).

Pesaran, M. H., 2006, Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica, vol. 74, 967-1012.

Pesaran, M.H., Yamagata, T., 2008, Testing slope homogeneity in large panels. Journal of Econometrics, vol. 142, 50-93.

Song, M., 2013. Asymptotic theory for dynamic heterogeneous panels with cross-sectional dependence and its applications. Mimeo, January 2013.

Appendix to: "Testing for no factor structures: on the use of average-type and Hausman-type statistics".

Throughout this Appendix we rely upon the following decomposition - see Proposition A.1 in Bai (2009):

$$
\begin{aligned}
\hat{F}-F &= \frac{1}{nT}\sum_{j=1}^{n}X_j\left(\tilde{\beta}_j-\beta_j\right)\left(\tilde{\beta}_j-\beta_j\right)'X_j'\hat{F} \\
&\quad -\frac{1}{nT}\sum_{j=1}^{n}X_j\left(\tilde{\beta}_j-\beta_j\right)\gamma_j'F'\hat{F}-\frac{1}{nT}\sum_{j=1}^{n}X_j\left(\tilde{\beta}_j-\beta_j\right)\epsilon_j'\hat{F} \\
&\quad -\frac{1}{nT}\sum_{j=1}^{n}F\gamma_j\left(\tilde{\beta}_j-\beta_j\right)'X_j'\hat{F}-\frac{1}{nT}\sum_{j=1}^{n}\epsilon_j\left(\tilde{\beta}_j-\beta_j\right)'X_j'\hat{F} \\
&\quad +\frac{1}{nT}\sum_{j=1}^{n}F\gamma_j\epsilon_j'\hat{F}+\frac{1}{nT}\sum_{j=1}^{n}\epsilon_j\gamma_j'F'\hat{F}+\frac{1}{nT}\sum_{j=1}^{n}\epsilon_j\epsilon_j'\hat{F}.
\end{aligned}
\tag{A.1}
$$

In (A.1), the main difference with Bai (2009) is the presence of the unit specific estimates, $\tilde{\beta}_j$. Consider also the following notation. We define $\Upsilon_i \equiv \left(X_i'\bar{M}_wX_i\right)^{-1}\left(X_i'\bar{M}_w\epsilon_i\right)$, so that we can write

$$
\begin{aligned}
\tilde{\beta}_i-\beta_i &= \left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_w\epsilon_i}{T}\right)+\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_wF}{T}\gamma_i\right) \\
&= \Upsilon_i+\bar{\Upsilon}_i,
\end{aligned}
\tag{A.2}
$$

for every i; by construction, $\bar{\Upsilon}_i = O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right)$. We extensively use the notation $\delta_{nT} = \min\left\{\sqrt{n},\sqrt{T}\right\}$ and $\phi_{nT}=\min\left\{n,\sqrt{T}\right\}$.

Proof of Theorem 1. We start with $\tilde{S}_{\gamma,nT}$. Under H_0^a we have

$$
\begin{aligned}
\sqrt{2r}\tilde{S}_{\gamma,nT} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[T\left(\hat{\gamma}_i-\gamma\right)'\hat{\Sigma}_{\gamma i}^{-1}\left(\hat{\gamma}_i-\gamma\right)-r\right]+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}T\left(\widehat{\overline{\gamma}}-\gamma\right)'\hat{\Sigma}_{\gamma i}^{-1}\left(\widehat{\overline{\gamma}}-\gamma\right) \\
&\quad -\frac{2}{\sqrt{n}}\sum_{i=1}^{n}T\left(\hat{\gamma}_i-\gamma\right)'\hat{\Sigma}_{\gamma i}^{-1}\left(\widehat{\overline{\gamma}}-\gamma\right) \\
&= I+II-III.
\end{aligned}
\tag{A.3}
$$

Consider I; by definition, under H_0^a

$$
\sqrt{T}\left(\hat{\gamma}_i-\gamma\right)=\left(\frac{\hat{F}'M_{Xi}\hat{F}}{T}\right)^{-1}\left[\frac{\hat{F}'M_{Xi}\epsilon_i}{\sqrt{T}}-\frac{\hat{F}'M_{Xi}\left(\hat{F}-F\right)\gamma}{\sqrt{T}}\right]; \tag{A.4}
$$

hence, we can write

$$
\begin{aligned}
I &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{\epsilon_i' M_{Xi}F}{\sqrt{T}}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\hat{\Sigma}_{\gamma i}^{-1}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\frac{F'M_{Xi}\epsilon_i}{\sqrt{T}} - r\right] \\
&\quad + \frac{1}{\sqrt{n}}T\sum_{i=1}^{n}\frac{\epsilon_i' M_{Xi}\left(\hat{F}-F\right)}{T}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\hat{\Sigma}_{\gamma i}^{-1}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\frac{\left(\hat{F}-F\right)' M_{Xi}\epsilon_i}{T} \\
&\quad + \frac{1}{\sqrt{n}}T\sum_{i=1}^{n}\frac{\gamma'\left(\hat{F}-F\right)' M_{Xi}\hat{F}}{T}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\hat{\Sigma}_{\gamma i}^{-1}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\frac{\hat{F}'M_{Xi}\left(\hat{F}-F\right)\gamma}{T} \\
&\quad + \frac{2}{\sqrt{n}}\sqrt{T}\sum_{i=1}^{n}\frac{\epsilon_i' M_{Xi}\left(\hat{F}-F\right)}{T}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\hat{\Sigma}_{\gamma i}^{-1}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\frac{F'M_{Xi}\epsilon_i}{\sqrt{T}} \\
&\quad + \frac{2}{\sqrt{n}}T\sum_{i=1}^{n}\frac{\epsilon_i' M_{Xi}\left(\hat{F}-F\right)}{T}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\hat{\Sigma}_{\gamma i}^{-1}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\frac{\hat{F}'M_{Xi}\left(\hat{F}-F\right)\gamma}{T} \\
&\quad + \frac{2}{\sqrt{n}}\sqrt{T}\sum_{i=1}^{n}\frac{\gamma'\left(\hat{F}-F\right)' M_{Xi}\hat{F}}{T}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\hat{\Sigma}_{\gamma i}^{-1}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\frac{F'M_{Xi}\epsilon_i}{\sqrt{T}} \\
&= I_a + I_b + I_c + I_d + I_e + I_f.
\end{aligned}
$$

By (7), I_a is $O_p(1)$. Turning to I_b, it is bounded by

$$
\sqrt{n}TE\left[\frac{\epsilon_i' M_{Xi}\left(\hat{F}-F\right)}{T}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\hat{\Sigma}_{\gamma i}^{-1}\left(\frac{F'M_{Xi}F}{T}\right)^{-1}\frac{\left(\hat{F}-F\right)' M_{Xi}\epsilon_i}{T}\right] \leq M\sqrt{n}TE\left\|\frac{\epsilon_i'\left(\hat{F}-F\right)}{T}\right\|^2,
$$

where we have used the consistency of $\hat{\Sigma}_{\gamma i}$ and Assumptions 3*(i)* and 4*(i)* in Castagnetti, Rossi and Trapani (2014) (CRT). Applying Lemma A.2*(i)* in CRT, we have $I_b = O_p\left(\sqrt{n}T\delta_{nT}^{-4}\right)$. By a similar logic, it can be shown that I_c is bounded by $\sqrt{n}TE\left\|T^{-1}\left(\hat{F}-F\right)F'\right\|^2$, which is $O_p\left(\sqrt{n}T\delta_{nT}^{-4}\right)$ by virtue of Lemma A.3*(ii)* in CRT. Turning to I_d, similar passages as above entails that it is bounded by

$$
\begin{aligned}
\sqrt{n}TE\left[\frac{\epsilon_i' M_{Xi}\left(\hat{F}-F\right)}{T}\frac{F'M_{Xi}\epsilon_i}{\sqrt{T}}\right] &\leq \sqrt{n}T\left[E\left(\left\|\frac{\epsilon_i'\left(\hat{F}-F\right)}{T}\right\|^2\right)\right]^{1/2}\left[E\left(\left\|\frac{F'\epsilon_i}{\sqrt{T}}\right\|^2\right)\right]^{1/2} \\
&= \sqrt{n}TO_p\left(\delta_{nT}^{-2}\right)O_p(1).
\end{aligned}
$$

Similarly, I_e is bounded by

$$\begin{aligned}
& \sqrt{n}TE\left[\frac{\epsilon_i' M_{Xi}\left(\hat{F}-F\right)}{T}\frac{\hat{F}' M_{Xi}\left(\hat{F}-F\right)\gamma}{T}\right] \\
\leq \quad & \sqrt{n}T\left[E\left(\left\|\frac{\epsilon_i'\left(\hat{F}-F\right)}{T}\right\|^2\right)\right]^{1/2}\left[E\left(\left\|\frac{\hat{F}'\left(\hat{F}-F\right)}{T}\right\|^2\right)\right]^{1/2} = \sqrt{n}TO_p\left(\delta_{nT}^{-2}\right)O_p\left(\delta_{nT}^{-2}\right);
\end{aligned}$$

using a similar logic, it can be shown that $I_f = O_p\left(\sqrt{nT}\delta_{nT}^{-2}\right)$. Putting all together, $I = O_p(1)$ $+O_p\left(\sqrt{nT}\delta_{nT}^{-2}\right)$ $+O_p\left(\sqrt{n}T\delta_{nT}^{-4}\right)$. Finally, consider II and III in (A.3). As far as II is concerned, note that $II = \sqrt{n}T\left(\widehat{\bar{\gamma}}-\bar{\gamma}\right)'\Sigma_{\gamma i}^{-1}\left(\widehat{\bar{\gamma}}-\bar{\gamma}\right)$ $+o_p(1)$ by consistency of $\hat{\Sigma}_{\gamma i}$. Thus, Lemma A.4 in CRT entails that $II = O_p\left(\sqrt{n}T\delta_{nT}^{-4}\right)$. Turning to III, this is bounded by $\sqrt{n}T\max_i\Sigma_{\gamma i}^{-1}\left\|\widehat{\bar{\gamma}}-\bar{\gamma}\right\|\left\|\frac{1}{n}\sum_{i=1}^n\left(\hat{\gamma}_i-\bar{\gamma}\right)\right\|$, which has the same order of magnitude as II. Putting all together, it holds that $\sqrt{2r}\tilde{S}_{\gamma,nT} = O_p(1)$ $+O_p\left(\sqrt{nT}\delta_{nT}^{-2}\right)$ $+O_p\left(\sqrt{n}T\delta_{nT}^{-4}\right) = O_p(1)$ $+O_p\left(\sqrt{\frac{n}{T}}\right)$ $+O_p\left(\sqrt{\frac{T}{n}}\right)$.

We now turn to studying $\tilde{S}_{f,nT}$. Under H_0^b, we have

$$\begin{aligned}
\sqrt{2r}\tilde{S}_{f,nT} &= \frac{1}{\sqrt{T}}\sum_{t=1}^T\left[n\left(\hat{f}_t-f\right)'\hat{\Sigma}_{ft}^{-1}\left(\hat{f}_t-f\right)-r\right] \\
&\quad +\frac{1}{\sqrt{T}}\sum_{t=1}^T n\left(\widehat{\bar{f}}-f\right)'\hat{\Sigma}_{ft}^{-1}\left(\widehat{\bar{f}}-f\right)-\frac{2}{\sqrt{T}}\sum_{t=1}^T n\left(\hat{f}_t-f\right)'\hat{\Sigma}_{ft}^{-1}\left(\widehat{\bar{f}}-f\right) \\
&= I+II-III.
\end{aligned} \tag{A.5}$$

Now define $b_t = \hat{f}_t - f$; using (A.1) we can write

$$b_t = \left(\frac{\hat{F}'F}{T}\right)\frac{1}{n}\sum_{i=1}^n\gamma_i\epsilon_{it}+b_{2t} = b_{1t}+b_{2t}, \tag{A.6}$$

where b_{2t} contains terms $I-VI$ and $VIII$ in (A.1).

Consider I; using (A.6) we may write

$$I = \frac{1}{\sqrt{T}}\sum_{t=1}^T\left(nb_{1t}'\hat{\Sigma}_{ft}^{-1}b_{1t}-r\right)+\frac{1}{\sqrt{T}}\sum_{t=1}^T nb_{2t}'\hat{\Sigma}_{ft}^{-1}b_{2t}+\frac{2}{\sqrt{T}}\sum_{t=1}^T nb_{1t}'\hat{\Sigma}_{ft}^{-1}b_{2t} = I_a+I_b+I_c.$$

By (8), $T^{-1/2}\sum_{t=1}^T\left(nb_{1t}'\hat{\Sigma}_{ft}^{-1}b_{1t}-r\right) = O_p(1)$. As far as I_b is concerned, by virtue of the consistency of Σ_{ft}, it is bounded by $n\sqrt{T}E\left\|b_{2t}\right\|^2 = n\sqrt{T}\min\left\{T^{-1},n^{-2}\right\} = O_p\left(\frac{n}{\sqrt{T}}\right)+O_p\left(\frac{\sqrt{T}}{n}\right)$, which follows from the proof of Theorem 2 in CRT. Finally, turning to I_c and setting $\hat{\Sigma}_{Ft}^{-1} = I_r$ for simplicity, we may write

$$
\begin{aligned}
\frac{1}{2} I_c &= n \left(\frac{F'\hat{F}}{T} \right) \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i' \left(\frac{\hat{F}' X_j}{T} \right) \left(\tilde{\beta}_j - \beta_j \right) \left(\tilde{\beta}_j - \beta_j \right)' \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_{jt} \epsilon_{it} \right) \\
&\quad - n \left(\frac{F'\hat{F}}{T} \right) \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i' \left(\frac{\hat{F}' F}{T} \right) \gamma_j \left(\tilde{\beta}_j - \beta_j \right)' \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_{jt} \epsilon_{it} \right) \\
&\quad - n \left(\frac{F'\hat{F}}{T} \right) \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i' \frac{1}{T} \left(\hat{F}' \epsilon_j \right) \left(\tilde{\beta}_j - \beta_j \right)' \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_{jt} \epsilon_{it} \right) \\
&\quad - n \left(\frac{F'\hat{F}}{T} \right) \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i' \left(\frac{\hat{F}' X_j}{T} \right) \left(\tilde{\beta}_j - \beta_j \right) \gamma_j' \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} f_t \epsilon_{it} \right) \\
&\quad - n \left(\frac{F'\hat{F}}{T} \right) \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i' \left(\frac{\hat{F}' X_j}{T} \right) \left(\tilde{\beta}_j - \beta_j \right) \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \epsilon_{jt} \epsilon_{it} \right) \\
&\quad + n \left(\frac{F'\hat{F}}{T} \right) \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i' \frac{\hat{F}' \epsilon_j}{T} \gamma_j' \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} f_t \epsilon_{it} \right) \\
&\quad + n \left(\frac{F'\hat{F}}{T} \right) \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i' \frac{\hat{F}' \epsilon_j}{T} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \epsilon_{jt} \epsilon_{it} \right) \\
&= I_{c,1} - I_{c,2} - I_{c,3} - I_{c,4} - I_{c,5} + I_{c,6} + I_{c,7}.
\end{aligned}
$$

Studying the order of magnitude of each term is based on similar passages to the ones in the proof of Theorem 2 in CRT. The only differences are: the summation across t; the normalization by $T^{-1/2}$; and the multiplication by n.

We have $I_{c,1} = O_p\left(\sqrt{n} T^{-1}\right)$; $I_{c,2} = O_p\left(\delta_{nT}^{-1}\right)$; $I_{c,3} = O_p\left(\sqrt{n} T^{-1}\right)$, $I_{c,4} = O_p\left(\delta_{nT}^{-1}\right)$ and $I_{c,6} = O_p\left(T^{-1/2}\right)$. As far as $I_{c,5}$ and $I_{c,7}$ are concerned, the contribution of $T^{-1/2} \sum_{t=1}^{T} \epsilon_{jt} \epsilon_{it}$ is at most $O_p\left(\sqrt{T}\right)$; thus, $I_{c,5} = O_p\left(\sqrt{n}\right)$ and $I_{c,7} = O_p\left(\sqrt{T} \delta_{nT}^{-2}\right)$. We now turn to analyzing II and III in (A.5). By Lemma A.5 in CRT, $II = n\sqrt{T} O_p\left(\delta_{nT}^{-4}\right)$. As far as III is concerned, using the consistency of $\hat{\Sigma}_{ft}$ and the invertibility of Σ_{ft}, it is bounded by $n\sqrt{T} \max_t \left\| \Sigma_{ft}^{-1} \right\| \left\| \hat{\bar{f}} - f \right\| \left\| \frac{1}{T} \sum_{t=1}^{T} \left(\hat{f}_t - f \right) \right\| = n\sqrt{T} O_p\left(\delta_{nT}^{-4}\right)$, again by Lemma A.5 in CRT. Putting all together, the result follows. QED

Proof of Theorem 2. We report the proof for $S_{\gamma,nT}^{CCE}$ only - the proof for $S_{\gamma,nT}^{IE}$ is almost identical; the only difference is the need for the restriction $\frac{\sqrt{n}}{T} \to 0$, which can be shown based on the passages in the proof of Theorem 3.

Consider the building block of the test statistic, viz. $\hat{\beta}^{bw}-\beta$:

$$\begin{aligned}\hat{\beta}^{bw}-\beta &= \frac{1}{n}\sum_{i=1}^{n}(\beta_i-\beta)+\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}x_{it}'(\beta_i-\beta)\right]\\ &-\frac{1}{n}\sum_{i=1}^{n}\frac{1}{n}\sum_{j=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{jt}\dot{x}_{jt}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{jt}x_{it}'(\beta_i-\beta)\right]\\ &+\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{\epsilon}_{it}\right].\end{aligned}$$

Note also that

$$\begin{aligned}\hat{\beta}^{CCE}-\beta &= \frac{1}{n}\sum_{i=1}^{n}(\beta_i-\beta)+\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_w\epsilon_i}{T}\right)\\ &+\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_wF}{T}\gamma_i\right),\end{aligned}$$

so that

$$\begin{aligned}\hat{\beta}^{bw}-\hat{\beta}^{CCE} &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}x_{it}'(\beta_i-\beta)\right]\\ &-\frac{1}{n}\sum_{i=1}^{n}\frac{1}{n}\sum_{j=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{jt}\dot{x}_{jt}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{jt}x_{it}'(\beta_i-\beta)\right]\\ &+\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{\epsilon}_{it}\right]-\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_w\epsilon_i}{T}\right)\\ &-\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_wF}{T}\gamma_i\right)\\ &= I+II+III-IV-V. \qquad \text{(A.7)}\end{aligned}$$

Terms $IV+V$ have magnitude $O_p\left(\frac{1}{\sqrt{nT}}\right)+O_p\left(\frac{1}{n}\right)$, as discussed above. Also, in a similar way it can be shown that $III=O_p\left(\frac{1}{\sqrt{nT}}\right)$. Finally, we have

$$I+II=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\tilde{D}_{it}x_{it}'(\beta_i-\beta),$$

where $\tilde{D}_{it}=\left[T^{-1}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\dot{x}_{it}-n^{-1}\sum_{j=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{jt}\dot{x}_{jt}'\right]^{-1}\dot{x}_{jt}$. By Assumption 3 in CRT, the sequence $T^{-1}\sum_{t=1}^{T}\tilde{D}_{it}x_{it}'(\beta_i-\beta)$ is uncorrelated across i, so that the magnitude of $I+II$ is proportional to the square root of $n^{-2}\sum_{i=1}^{n}E\left\|\tilde{D}_{it}x_{it}'(\beta_i-\beta)\right\|^2\leq n^{-2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}E\left\|\tilde{D}_{it}x_{it}'\right\|^2 E\|\beta_i-\beta\|^2$. Using Assumptions 3 and 2*(i)* in CRT, this is of order $O\left(n^{-1}\right)$,so that $I+II=O_p\left(n^{-1/2}\right)$. The limiting distribution follows from standard arguments, upon noting that the sequence $T^{-1}\sum_{t=1}^{T}\tilde{D}_{it}x_{it}'(\beta_i-\beta)$ is conditionally independent across i by Assumption 3 in CRT, and has finite moment of order $2+\delta$ for $\delta>0$. Putting all together, the null distribution follows.

As far as power is concerned, the CCE estimator is consistent under alternatives; as far as the between estimator is concerned, $\hat{\beta}^{bw}-\beta$ has the same expansion as above with the extra term

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}f_t'\left(\gamma_i-\frac{1}{n}\sum_{i=1}^{n}\gamma_i\right)\right]\\
=\;&\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}f_t'(\gamma_i-\gamma)\right]-\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}f_t'\left(\frac{1}{n}\sum_{i=1}^{n}\gamma_i-\gamma\right)\right]\\
=\;&I+II,
\end{aligned}$$

where term II is clearly dominated. As far as I is concerned, when premultiplied by $\sqrt{n}$, we have $I=n^{-1/2}\sum_{i=1}^{n}C_{iT}$. By assumption, C_{iT} has mean zero, it can be shown to have finite moment of order $2+\delta$ for $\delta>0$, and it is conditionally independent across i. It is also conditionally independent of $II+III$ in (A.7). This entails that, under alternatives, $\sqrt{n}\left(\hat{\beta}^{bw}-\hat{\beta}^{CCE}\right)$ converges to a normally distributed random variable with mean zero, and a higher variance than under the null. Standard passages ensure the validity of the theorem. QED

Proof of Theorem 3. Consider first equation (10); we start with $\sqrt{nT}\left(\hat{\beta}^{CCE}-\hat{\beta}^{FE}\right)$ under H_0^a. Recall that under the null H_0^b, $f_t=f=ci_T$, where c is a constant. Therefore, $M_F=I_T-T^{-1}i_Ti_T'$. This entails that

$$\hat{\beta}^{FE}-\beta=\frac{1}{n}\sum_{i=1}^{n}(\beta_i-\beta)+\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'M_FX_i}{T}\right)^{-1}\left(\frac{X_i'M_F\epsilon_i}{T}\right). \tag{A.8}$$

By using (A.2) and equation (56) in Pesaran (2006, p. 982):

$$\begin{aligned}
\sqrt{nT}\left(\hat{\beta}^{CCE}-\hat{\beta}^{FE}\right) \;=\;& \sqrt{nT}\frac{1}{n}\sum_{i=1}^{n}(\beta_i-\beta)+\sqrt{nT}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_w\epsilon_i}{T}\right)\\
&+\sqrt{nT}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_wF}{T}\gamma_i\right)\\
&-\sqrt{nT}\frac{1}{n}\sum_{i=1}^{n}(\beta_i-\beta)-\sqrt{nT}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'M_FX_i}{T}\right)^{-1}\left(\frac{X_i'M_F\epsilon_i}{T}\right).
\end{aligned}$$

Under the rank condition in Assumption 4*(ii)* in CRT, we have that

$$\begin{aligned}
\frac{X_i'\bar{M}_w\epsilon_i}{T} &= \frac{X_i'M_F\epsilon_i}{T}+O_p\left(\frac{1}{n}\right), \\
\frac{X_i'\bar{M}_wF}{T} &= O_p\left(\frac{1}{n}\right)+O_p\left(\frac{1}{\sqrt{nT}}\right);
\end{aligned}$$

note that the former equation does not need the rank condition in Assumption 4*(ii)*, whereas the latter does. Therefore

$$\begin{aligned}
\sqrt{nT}\left(\hat{\beta}^{CCE}-\hat{\beta}^{FE}\right) &= O_p\left(\sqrt{\frac{T}{n}}\right)+\sqrt{nT}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'\bar{M}_wX_i}{T}\right)^{-1}\left(\frac{X_i'\bar{M}_wF}{T}\gamma_i\right) \\
&= O_p\left(\sqrt{\frac{T}{n}}\right)+O_p\left(\sqrt{\frac{T}{n}}\right)+O_p\left(1\right),
\end{aligned}$$

which proves part 1 of the Theorem. The asymptotics of the terms $O_p\left(\sqrt{\frac{T}{n}}\right)$ and $O_p\left(1\right)$ depends on the DGP of x_{it} and y_{it} through $\bar{M}_w$ and $T^{-1}\left(X_i'\bar{M}_wF\right)$.

Consider now equation (9). Note that

$$\hat{\beta}^{IE}-\beta=\left(\hat{\beta}^{IE}-\frac{1}{n}\sum_{i=1}^{n}\beta_i\right)+\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i-\beta\right),$$

so that, using (A.8):

$$\begin{aligned}
\sqrt{nT}\left(\hat{\beta}^{IE}-\hat{\beta}^{FE}\right) &= \sqrt{nT}\left(\hat{\beta}^{IE}-\frac{1}{n}\sum_{i=1}^{n}\beta_i\right)-\sqrt{nT}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i'M_FX_i}{T}\right)^{-1}\left(\frac{X_i'M_F\epsilon_i}{T}\right) \\
&= \sqrt{nT}\left(\hat{\beta}^{IE}-\frac{1}{n}\sum_{i=1}^{n}\beta_i\right)+O_p\left(1\right),
\end{aligned}$$

where the $O_p\left(1\right)$ term holds by Assumptions 1 and 2 in CRT. Let $\Gamma=[\gamma_1|...|\gamma_n]$; using equation (42) in

Song (2013), we have

$$\begin{aligned}
& \hat{\beta}^{IE} - \frac{1}{n}\sum_{i=1}^{n}\beta_i \\
= & \frac{1}{n}\sum_{i=1}^{n}\left(\tilde{\beta}_i^{IE} - \beta_i\right) \\
= & \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_i}{T}\right)^{-1}\left(\frac{X_i' M_{\hat{F}} \epsilon_i}{T}\right) \\
& -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_i}{T}\right)^{-1}\frac{1}{n}\sum_{j=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_j}{T}\right)\left[\gamma_j'\left(\frac{\Gamma'\Gamma}{n}\right)^{-1}\gamma_i\right]\left(\tilde{\beta}_j - \beta_j\right) \\
& -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_i}{T}\right)^{-1}\left[\frac{1}{n}\sum_{j=1}^{n}\frac{X_i' M_{\hat{F}} X_j}{T}\left(\tilde{\beta}_j^{IE} - \beta_j\right)\left(\tilde{\beta}_j^{IE} - \beta_j\right)'\frac{X_j'\hat{F}}{T}\right. \\
& +\frac{1}{n}\sum_{j=1}^{n}\frac{X_i' M_{\hat{F}} X_j}{T}\left(\tilde{\beta}_j^{IE} - \beta_j\right)\frac{\epsilon_j'\hat{F}}{T} + \frac{1}{n}\sum_{j=1}^{n}\frac{X_i' M_{\hat{F}} F}{T}\gamma_j\left(\tilde{\beta}_j^{IE} - \beta_j\right)'\frac{X_j'\hat{F}}{T} + \\
& \frac{1}{n}\sum_{j=1}^{n}\frac{X_i' M_{\hat{F}} \epsilon_j}{T}\left(\tilde{\beta}_j^{IE} - \beta_j\right)'\frac{X_j'\hat{F}}{T} + \frac{1}{n}\sum_{j=1}^{n}\frac{X_i' M_{\hat{F}} F}{T}\gamma_j\frac{\epsilon_j'\hat{F}}{T} + \frac{1}{n}\sum_{j=1}^{n}\frac{X_i' M_{\hat{F}} \epsilon_j}{T}\gamma_j'\frac{F'\hat{F}}{T} \\
& \left.+\frac{1}{n}\sum_{j=1}^{n}\frac{X_i' M_{\hat{F}} \epsilon_j}{T}\frac{\epsilon_j'\hat{F}}{T}\right]\left(\frac{F'\hat{F}}{T}\right)^{-1}\left(\frac{\Gamma'\Gamma}{n}\right)\gamma_i \\
= & \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_i}{T}\right)^{-1}\left(\frac{X_i' M_{\hat{F}} \epsilon_i}{T}\right) + J_1 + J_2 + J_3 + J_4 + J_5 + J_6 + J_7 + J_8;
\end{aligned}$$

quantities like $\frac{F'\hat{F}}{T}$, $\frac{\Gamma'\Gamma}{n}$ and γ_i will be omitted henceforth, to simplify the notation. Similar passages as above yield $\sqrt{nT}\ \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_i}{T}\right)^{-1}\left(\frac{X_i' M_{\hat{F}} \epsilon_i}{T}\right) = O_p(1)$. As far as the other terms are concerned, note first that, by Proposition 1 in Song (2013), $\tilde{\beta}_i^{IE} - \beta_i = O_p\left(\phi_{nT}^{-1}\right)$. Since the order of magnitude of an average is bounded by the order of the summands, the same passages as in Song (2013) would entail $J_2 = O_p\left(\phi_{nT}^{-2}\right)$; J_3, J_5 and J_6 are all bounded by $O_p\left(\phi_{nT}^{-2}\right) + O_p\left(\phi_{nT}^{-1}\delta_{nT}^{-1}\right)$; $J_4 = O_p\left(\phi_{nT}^{-2}\right) + O_p\left(\phi_{nT}^{-1}\delta_{nT}^{-2}\right)$; $J_7 = O_p\left(T^{-1/2}\phi_{nT}^{-1}\right) + O_p\left(T^{-1/2}\phi_{nT}^{-1}\right)$. Putting all together, this entails that $\sqrt{nT}$ $(J_2 + J_3 + J_4 + J_5 + J_6 + J_7) = O_p\left(\sqrt{\frac{n}{T}}\right) + O_p\left(\sqrt{\frac{T}{n^3}}\right) + O_p(1)$. This bound is not the sharpest possible, but it suffices for our purposes. Finally, consider J_1 and J_8. As far as the former is concerned, we have

$$J_1 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_i}{T}\right)^{-1}\frac{1}{n}\sum_{j=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_j}{T}\right)\left[\gamma_j'\left(\frac{\Gamma'\Gamma}{n}\right)^{-1}\gamma_i\right]\left(\tilde{\beta}_j - \beta_j\right),$$

which, by virtue of Assumptions 2*(i)* and 4*(i)* in CRT, has the same order as derived in Song (2013);

thus, $J_1 = O_p\left(n^{-1/2}T^{-1/2}\right)$ and $\sqrt{nT}J_2 = O_p(1)$. Turning to J_8

$$
\begin{aligned}
-J_8 &= \frac{1}{nT}\sum_{i=1}^{n}\left(\frac{X_i' M_{\hat{F}} X_i}{T}\right)^{-1} X_i' M_{\hat{F}}\left(\frac{1}{nT}\sum_{k=1}^{n}\epsilon_k\epsilon_k'\hat{F}\right) \\
&\leq M\frac{1}{nT^2}\sum_{k=1}^{n}\left\|X_i' M_{\hat{F}}\left(\epsilon_k\epsilon_k'\hat{F}\right)\right\| = O_p\left(\frac{1}{\delta_{nT}^2}\right) + O_p\left(\frac{1}{\phi_{nT}\sqrt{T}}\right),
\end{aligned}
$$

where the last passage comes from the proof of Proposition 1 in Song (2013). Therefore, $\sqrt{nT}J_8 = O_p\left(\sqrt{\frac{n}{T}}\right) + O_p\left(\sqrt{\frac{T}{n}}\right)$. Putting all together, part 2 of the Theorem follows. The behaviour of the test statistics under the null H_0^a follows immediately from the passages above, since the term $\sqrt{nT}\left(\hat{\beta}^{IE} - \frac{1}{n}\sum_{i=1}^{n}\beta_i\right)$ is still $O_p\left(\sqrt{\frac{n}{T}}\right) + O_p\left(\sqrt{\frac{T}{n}}\right)$. QED

References

Bai, J., 2009, Panel data models with interactive fixed effects. Econometrica, vol. 77, 1229-1279.

Castagnetti, C., Rossi, E., Trapani, L., 2014, Inference on Factor Structures in Heterogeneous Panels. Journal of Econometrics (forthcoming).

Pesaran, M. H., 2006, Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica, vol. 74, 967-1012.

Song, M., 2013. Asymptotic theory for dynamic heterogeneous panels with cross-sectional dependence and its applications. Mimeo, January 2013.