Dynamic spatial panel data models with common shocks

Jushan Bai* and and Kunpeng Li†

September 19, 2017

Abstract

Real data often have complicated correlation over cross section and time. Modeling, estimating and interpreting the correlations in data are particularly important in economic analysis. This paper integrates several correlation-modeling techniques and propose dynamic spatial panel data models with common shocks to accommodate possibly complicated correlation structure over cross section and time. A large number of incidental parameters exist within the model. The quasi maximum likelihood method (ML) is proposed to estimate the model. Heteroskedasticity is explicitly estimated. The asymptotic properties of the quasi maximum likelihood estimator (MLE) are investigated. Our analysis indicates that the MLE has a non-negligible bias. We propose a bias correction method for the MLE. The simulations further reveal the excellent finite sample properties of the quasi-MLE after bias correction.

Key Words: Panel data models, spatial interactions, common shocks, cross-sectional dependence, incidental parameters, maximum likelihood estimation

JEL Classification: C3; C13

*Department of Economics, Columbia University, NewYork

†School of International Economics and Management, Capital University of Economics and Business, Beijing, China,

1 Introduction

Real data often have complicated correlation over cross section and time. These correlations contain important information on the relationship among economic variables. Modeling, estimating and interpreting the correlations in data are particularly important in economic analysis. In econometric literature, the correlations over time are typically dealt with by the autoregressive models (e.g., Brockwell and Davis (1991), Fuller (1996), etc), among other models. The correlations over cross section are typically captured by spatial models or factor models (e.g., Anselin (1988), Bai and Li (2012), Fan et al. (2011), etc), among other models. In this paper, we integrate these correlation-modeling techniques and propose dynamic spatial panel data models with common shocks to accommodate possibly complicated correlation structure over cross section and time.

Spatial models are one of primary tools to study cross-sectional interactions among units. In these models, cross sectional dependence is captured by spatial weights matrices based either on physical distance, and relative position in a social network or on other types of economic distance[①]. Early development of spatial models has been summarized by a number of books, including Cliff and Ord (1973), Anselin (1988), and Cressie (1993). Generalized method of moments (GMM) estimation of spatial models are studied by Kelijian and Prucha (1998, 1999, 2010), and Kapoor et al. (2007), among others. The maximum likelihood method (ML) is considered by Ord (1975), Anselin (1988), Lee (2004a), Yu et al. (2008) and Lee and Yu (2010), and so on.

Cross-sectional dependence may also arises from the response of individuals to common shocks. This motivates common shocks models, which are widely used in applied studies, see, e.g., Ross (1976), Chamberlain and Rothschild (1983), Stock and Watson (1998), to name a few. For panel data models with multiple common shocks, Ahn et al. (2013) consider the fixed-T GMM estimation. Pesaran (2006) proposes the correlated random effects method by including additional regressors obtained from cross-sectionally averaging on dependent and the explanatory variables. The principal components method is studied by Bai (2009) and reinvestigated with perturbation theory by Moon and Weidner (2009). Bai and Li (2014b) consider the maximum likelihood method in the presence of heteroskedasticity.

A popular approach to dealing with temporal dependence is dynamic panel data models. In these models, the presence of individual time-invariant intercepts (fixed-effect) causes the so-called "incidental parameters problem" (Neyman and Scott (1948)), which is the primary concern in the related studies. A consequence of the incidental parameters problem is the inconsistency of the within group estimator under fixed-T (Nickell (1981)). Anderson and Hsiao (1981) suggests taking time difference to eliminate the fixed effects and use two-periods lagged dependent variable as instrument to estimate the model. Arellano and Bond (1991) extend the Anderson and Hsiao's idea with the GMM method. Under large-N and large-T setup, Hahn and Kuersteiner (2002) shows that the within-group estimator is still consistent but has a $O(\frac{1}{T})$ bias. After bias correc-

[①] For spatial interaction and economic distance, see, e.g., Case (1991), Case et al. (1993), Conley (1999), Conley and Dupor (2003), and Topa (2001).

tion, the corrected estimator achieves the efficiency bound under normality assumption of errors. Alvarez and Arellano (2003) investigate the asymptotic properties of the within group, GMM and limited information ML estimators under large-N and large-T.

In this paper, we consider jointly modeling spatial interactions, dynamic interactions and common shocks within the following model:

$$y_{it} = \alpha_i + \rho \sum_{j=1}^{N} w_{ij,N} y_{jt} + \delta y_{it-1} + x_{it}'\beta + \lambda_i' f_t + e_{it}. \tag{1.1}$$

where y_{it} is the dependent variable; $x_{it} = (x_{it1}, x_{it2}, \ldots, x_{itk})'$ is a k-dimensional vector of explanatory variables; f_t is an r-dimensional vector of unobservable common shocks; λ_i is the corresponding heterogenous response to the common shocks; $W_N = (w_{ij,N})_{N\times N}$ is a specified spatial weights matrix whose diagonal elements $w_{ii,N}$ are 0; and e_{it} are the idiosyncratic errors. In model (1.1), term $\lambda_i' f_t$ captures the common-shocks effects, $\rho \sum_{j=1}^{N} w_{ij,N} y_{jt}$ captures the spatial effects, and δy_{it-1} captures the dynamic effects. The joint modeling allows one to test which type of effects is present within data. We may test $\rho = 0$ while allowing common-shocks effects and dynamic effects; or similarly, we may determine if the number of factors is zero in a model with spatial effects and dynamic effects. It may be possible that all the three effects are present. The features of model (1.1) make it flexible enough to cover a wide range of applications. The applicability of the model is discussed in Section 2.

An additional feature of the model is the allowance of cross sectional heteroskedasticity. The importance of permitting heteroskedasticity is noted by Kelejian and Prucha (2010) and Lin and Lee (2010). The heteroskedastic variances can be empirically important, e.g., Glaeser et al. (1996) and Anselin (1988). In addition, if heteroskedasticity exists but homoskedasticity is imposed, then MLE can be inconsistent. Under large-N, the consistency analysis for MLE under heteroskedasticity is challenging even for spatial panel models without common shocks, owing to the simultaneous estimation of a large number of variance parameters along with (ρ, δ, β). The existing quasi maximum likelihood studies, such as Yu et al. (2008) and Lee and Yu (2010), typically assume homoskedasticity. These authors show that the limiting variance of MLE has a sandwich formula unless normality is assumed. Interestingly, we show that the limiting variance of the MLE is not of a sandwich form if heteroskedasticity is allowed.

Spatial correlations and common shocks are also considered by Pesaran and Tosetti (2011). Except that the dynamics is allowed in our model but not in theirs, another key difference is that they specify the spatial autocorrelation on the unobservable errors e_{it} while we specify the spatial autocorrelation on the observable dependent variable y_{it}. Both specifications are of practical relevance. Spatial specification on observable data makes explicit the empirical implication of the coefficient ρ. From a theoretical perspective, the spatial interaction on the dependent variable gives rise to the endogeneity problem, while the spatial interaction on the errors, in general, does not. As a result, under the Pesaran and Tosetti setup, existing estimation methods on the common shocks models such as Pesaran (2006) and Bai (2009) can be applied to estimate the model. As a comparison, these methods cannot be directly applied to model (1.1) due to the endogeneity from the spatial interactions.

In this study, we consider the pseudo-Gaussian maximum likelihood method (MLE), which simultaneously estimates all parameters of the model, including heteroskedasticity. We give a rigorous analysis of the MLE including the consistency, the rate of convergence and limiting distributions. Since the proposed model has several sources of incidental parameters (individual-dependent intercepts, interactive effects, heteroskedasticity), the incidental parameters problem exists and the MLE is shown to have a non-negligible bias. Following Hahn and Kuersteiner (2002), we conduct bias correction on the MLE to make it center around zero. The simulations show that the bias-corrected MLE has good finite sample performance.

The rest of the paper is organized as follows. Section 2 gives some potential show-case examples of the model. Section 3 lists the assumptions needed for the asymptotic analysis. Section 4 presents the objective function and the associated first order conditions. The asymptotic properties including the consistency, the convergence rates and the limiting distributions are derived in Section 5. Section 6 discusses the ML estimation on spatial models with heteroskedasticity. Section 7 reports simulation results. Section 8 discusses extensions of the model. The last section concludes. Technical proofs are given in a supplementary document. In subsequent exposition, the matrix norms are defined in the following way. For any $m \times n$ matrix A, $\|A\|$ denotes the Frobenius norm of A, i.e., $\|A\| = [\text{tr}(A'A)]^{1/2}$. In addition, $\|A\|_\infty$ is defined as $\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$ and $\|A\|_1$ is defined as $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$, where a_{ij} is the (i, j)th element of A. We use $\dot{a}_t$ to denote $\dot{a}_t = a_t - \frac{1}{T} \sum_{t=1}^{T} a_t$ for any column vector a_t. Throughout the paper, we assume the data of Y at time 0 are observed.

2 Some application examples

The proposed model can be applied in a variety of economic and social setups. In this section, we list two typical examples.

Finance. Recent studies pay much attention on financial network and financial contagion. Let y_{it} be the stock price (or profit) of firm i at period t. In financial market, one firm may hold shares of other firms and other firms may hold shares of this firm. This generates a financial network (Elliott et al. (2014)). Let $W_N = [w_{ij,N}]$ be some metric, which measures the cross-holding pattern among firms in market. Then $\rho \sum_{j=1}^{N} w_{ij,N} y_{jt}$ captures the cross-holding effects on firm i. In addition, as implied in asset pricing theory (see, e.g., Ross (1976), Conner and Korajczyk (1986, 1988), Geweke and Zhou (1996)), there are systematic shocks and risks affecting all the stocks, which we denote by f_t. The individual-dependent responses to these shocks are captured by λ_i. This leads to term $\lambda_i' f_t$. Furthermore, the adaptive expectation of firms gives rise to δy_{it-1}. Let x_{it} be a vector of explanatory variables, which are thought useful to explain the behaviors of stock prices. We allow that x_{it} has arbitrary correlations with systematic shocks f_t. Putting these ingredients together, we have the model specification like (1.1).

Macroeconomics. Standard economic theory asserts if other countries grow with high rates, the outside demand would drive up the growth rate of home country through trade. Recent studies shows that international trade exhibits some spatial pattern, not only due to the distance cost as illustrated by "gravity" theory, but also duo to regional trade agreement as well as ethnical, cultural and social network among the firms, see, e.g., Baltagi et al. (2008), Lawless (2009), Rauch and Trindade (2002), Defever et al. (2015), etc. Let y_{it} be growth rate of country i at period t, and $W_N = [w_{ij,N}]$ be some metric, which measures the closeness of countries based on the bilateral trade. Then term $\rho \sum_{j=1}^{N} w_{ij,N} y_{jt}$ captures the companion-driving effect in growth. Similarly as in the previous example, the growth rates of countries over the world are subject to global economic shocks, such as technological advances and financial crisis (Kose, Otrok and Whiteman 2003). We therefore introduce term $\lambda_i' f_t$ to adapt to this fact. Term δy_{it-1} is also necessary because of the inertia of growth. With these considerations, we have the specification of model (1.1).

Besides the above economic applications, the proposed model also has its applications in social science. In a pioneer study, Manski (1993) distinguishes three effects within social interactions, *endogenous effects, contextual effects* and *correlated effects*. In empirical studies, endogenous effects are estimated by the spatial term, controlling correlated effects through the usually additive fixed effects (Lin (2010)). In the proposed model, we can deal with correlated effects in a more general and plausible way by factor models. In addition, we allow the dynamics. In Appendix, we show that, with some slight modifications, our model specification can be motivated by the quadratic utility model of Calvó-Armengol et al. (2009).

Apart from the above specific applications, model (1.1) can also be used, as the first step, to determine which model should be used in analysis. For example, it is known that knowledge spills over after it is generated. The spill-over pattern may exhibit some ad hoc weak one, as specified by spatial models, or a general strong one, as specified by common shock models. There are some debates on this issue (Eberhardt et al. (2013)). Our model is helpful to solve this issue.

3 Assumptions

We first introduce a set of normalization conditions, which facilitate the analysis of the asymptotic properties. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)'$ and $Y_t = (y_{1t}, y_{2t}, \ldots, y_{Nt})'$. The symbols Y_{t-1}, X_t and e_t are defined similarly as Y_t. Then we can rewrite model (1.1) into matrix form

$$Y_t = \alpha + \rho W_N Y_t + \delta Y_{t-1} + X_t \beta + \Lambda f_t + e_t. \tag{3.1}$$

The above model can always be written as

$$Y_t = \underbrace{(\alpha + \Lambda \bar{f})}_{\alpha^\dagger} + \rho W_N Y_t + \delta Y_{t-1} + X_t \beta + \underbrace{\Lambda Q^{-1/2}}_{\Lambda^\dagger} \underbrace{Q^{1/2}(f_t - \bar{f})}_{f_t^\dagger} + e_t$$

where $Q = \frac{1}{N} \Lambda' \Sigma_{ee}^{-1} \Lambda$ and $\bar{f} = \frac{1}{T} \sum_{t=1}^{T} f_t$. Let $\alpha^\dagger$, $\Lambda^\dagger$ and $f_t^\dagger$ be defined as illustrated in the above equation. We see that $\sum_{t=1}^{T} f_t^\dagger = 0$ and $\frac{1}{N} \Lambda^{\dagger\prime} \Sigma_{ee}^{-1} \Lambda^\dagger = I_r$. So it is no loss of generality to assume

Normalization conditions: $\sum_{t=1}^{T} f_t = 0$; $\frac{1}{N}\Lambda'\Sigma_{ee}^{-1}\Lambda = I_r$.

We shall use $(\rho^*, \delta^*, \beta^*)$ to denote the true values for (ρ, δ, β), and we use (Λ^*, f_t^*) to denote the true values for (Λ, f_t). So the data generating process is

$$Y_t = \alpha^* + \rho^* W_N Y_t + \delta^* Y_{t-1} + X_t \beta^* + \Lambda^* f_t^* + e_t.$$

Let C be a generic constant large enough. We make following assumptions for the asymptotic analysis.

Assumption A: The x_{it} is either a fixed constant or a random variable. If x_{it} is fixed, we assume $\|x_{it}\| \leq C$; if x_{it} is random, we assume $E(\|x_{it}\|^4) \leq C$ for all i and t. If x_{it} is random, it is independent with the idiosyncratic error e_{js} for all i, j, t and s.

Assumption B: The λ_i^* and f_t^* can be either fixed constants and random variables. If λ_i^* is fixed, we assume that $\|\lambda_i^*\| \leq C$ for all i and $\frac{1}{N}\Lambda^{*\prime}\Sigma_{ee}^{*-1}\Lambda^* \to \Omega_\Lambda^*$ where $\Lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_N^*)'$, otherwise we assume that $E(\|\lambda_i^*\|^4) \leq C$ for all i and $\frac{1}{N}\Lambda^{*\prime}\Sigma_{ee}^{*-1}\Lambda^* \xrightarrow{p} \Omega_\Lambda^*$, where Σ_{ee}^* is defined in Assumption C and Ω_Λ^* is some matrix positive definite. If f_t^* is fixed, we assume that $\|f_t^*\| \leq C$ for all t and $\frac{1}{T}F^{*\prime}F^* \to \Omega_F^*$, otherwise we assume that $E\|f_t^*\|^4 \leq C$ for all t and $\frac{1}{T}F^{*\prime}F^* \xrightarrow{p} \Omega_F^*$, where Ω_F^* is some matrix positive definite.

Assumption C: The e_{it} is independent and identically distributed over t and independent over i with $E(e_{it}) = 0$, $C^{-1} \leq \sigma_i^{*2} \leq C$ and $E(e_{it}^8) \leq C$ for all i, where $\sigma_i^{*2} = E(e_{it}^2)$. Let $\Sigma_{ee}^* = \text{diag}(\sigma_1^{*2}, \sigma_2^{*2}, \ldots, \sigma_N^{*2})$ be the variance of $e_t = (e_{1t}, e_{2t}, \ldots, e_{Nt})'$. In addition, if $\{\lambda_i^*\}$ and $\{f_t^*\}$ are random, we assume that $\{e_{it}\}$ are independent with $\{\lambda_i^*\}$ and $\{f_t^*\}$.

Assumption D: The underlying value $\omega^* = (\rho^*, \delta^*, \beta^{*\prime})'$ is an interior point of parameters space $\Theta_\omega = (-1, 1) \times \mathbb{S}_\delta \times \mathbb{S}_\beta$, where $\mathbb{S}_\delta$ and $\mathbb{S}_\beta$ are the two compact subsets of $\mathbb{R}$ and $\mathbb{R}^k$.

Remark 3.1 Assumption A impose restrictions on the explanatory variables x_{it}. Although it requires that x_{it} be independent with e_{js}, it does allow x_{it} to have arbitrary correlations with λ_i or f_t or $\lambda_i' f_t$. This extends the traditional panel data analysis. Assumption B is about factors and factor loadings. This assumption is standard in pure factor analysis, see Bai (2003) and Bai and Li (2012). Assumption C assumes that the idiosyncratic error e_{it} is independent over the cross section and the time. In the present scenario, such an assumption is not restrictive as it looks to be since the weak correlations over the cross section and the time in data have been dealt with by the spatial term and the lag dependent term. However, if the cross sectional correlation of e_{it} is a major concern in empirical studies, our analysis can be extended to accommodate it, see the related discussion on SAR disturbances in Section 7. Assumptions D impose restrictions on the underlying coefficients. This assumption is standard.

Assumption E: The weights matrix W_N satisfies that $I_N - \rho^* W_N$ is invertible and

$$\limsup_{N\to\infty} \|W_N\|_\infty \leq C; \qquad \limsup_{N\to\infty} \|W_N\|_1 \leq C; \tag{3.2}$$

$$\limsup_{N\to\infty} \|(I_N - \rho^* W_N)^{-1}\|_\infty \leq C; \qquad \limsup_{N\to\infty} \|(I_N - \rho^* W_N)^{-1}\|_1 \leq C. \tag{3.3}$$

In addition, all the diagonal elements of W_N are zeros.

Assumption F: Let $G_N^* = (I_N - \rho^* W_N)^{-1}$. We assume

$$\limsup_{N\to\infty} \sum_{l=0}^{\infty} \left\|(\delta^* G_N^*)^l\right\|_\infty \le C; \qquad \limsup_{N\to\infty} \sum_{l=0}^{\infty} \left\|(\delta^* G_N^*)^l\right\|_1 \le C.$$

Remark 3.2 Assumptions E and F are imposed on the spatial weights matrix. Assumption E is standard in spatial econometrics, see Kelejian and Prucha (1998), Lee (2004a), Yu et al. (2008), Lee and Yu (2010), to name a few. Under this assumption, some key matrices, which play important roles in asymptotic analysis such as G_N^* in Assumption F and S_N^* in Assumption G, can be handled in a tractable way. Assumption F implicitly guarantees that y_{it} has a well-defined MA(∞) expression. Similar assumption also appears in Yu et al. (2008). A set of sufficient conditions for Assumptions E and F are $\limsup_{N\to\infty} \|W_N\|_\infty \le 1$, $\limsup_{N\to\infty} \|W_N\|_1 \le 1$ and $|\rho^*| + |\delta^*| < 1$ because

$$\limsup_{N\to\infty} \|G_N^*\|_\infty = \limsup_{N\to\infty} \|(I - \rho^* W_N)^{-1}\|_\infty \le \limsup_{N\to\infty} \sum_{j=0}^{\infty} (\|\rho^* W_N\|_\infty)^j \le \frac{1}{1-|\rho^*|} < \infty,$$

and the argument for $\limsup_{N\to\infty} \|G_N^*\|_1 \le \frac{1}{1-|\rho^*|} < \infty$ is the same. Similarly

$$\limsup_{N\to\infty} \sum_{l=0}^{\infty} \left\|(\delta^* G_N^*)^l\right\|_\infty \le \limsup_{N\to\infty} \sum_{l=0}^{\infty} (|\delta^*| \cdot \|G_N^*\|_\infty)^l \le \sum_{l=0}^{\infty} \left[\frac{|\delta^*|}{1-|\rho^*|}\right]^l = \frac{1-|\rho^*|}{1-|\delta^*|-|\rho^*|} < \infty,$$

and the argument for $\limsup_{N\to\infty} \sum_{l=0}^{\infty} (|\delta^*| \cdot \|G_N^*\|_1)^l \le \frac{1-|\rho^*|}{1-|\delta^*|-|\rho^*|} < \infty$ is the same.

To state Assumption G, we first introduce some notations for ease of exposition. Let $\ddot{Y} = (\ddot{y}_{it})_{N\times T}$ be the data matrix for $\ddot{y}_{it}$ with $\ddot{y}_{it} = \sum_{j=1}^N w_{ij,N} \dot{y}_{jt}$ and $\dot{y}_{jt} = y_{jt} - T^{-1}\sum_{s=1}^T y_{js}$, $\dot{Y}_{-1} = (\dot{y}_{it-1})_{N\times T}$ with $\dot{y}_{it-1} = y_{it-1} - T^{-1}\sum_{s=1}^T y_{is-1}$ and $\dot{X}_1, \dot{X}_2, \dots, \dot{X}_k$ be defined similarly as $\dot{Y}_{-1}$. Furthermore, let $(k+1)\times(k+1)$ matrix $\mathbb{D}_b$ be defined as

$$\mathbb{D}_b = \frac{1}{NT}\begin{bmatrix} \text{tr}(\dot{Y}_{-1}'\ddot{M}\dot{Y}_{-1}M_{F^*}) & \text{tr}(\dot{Y}_{-1}'\ddot{M}\dot{X}_1 M_{F^*}) & \cdots & \text{tr}(\dot{Y}_{-1}'\ddot{M}\dot{X}_k M_{F^*}) \\ \text{tr}(\dot{X}_1'\ddot{M}\dot{Y}_{-1}M_{F^*}) & \text{tr}(\dot{X}_1'\ddot{M}\dot{X}_1 M_{F^*}) & \cdots & \text{tr}(\dot{X}_1'\ddot{M}\dot{X}_k M_{F^*}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{tr}(\dot{X}_k'\ddot{M}\dot{Y}_{-1}M_{F^*}) & \text{tr}(\dot{X}_k'\ddot{M}\dot{X}_1 M_{F^*}) & \cdots & \text{tr}(\dot{X}_k'\ddot{M}\dot{X}_k M_{F^*}) \end{bmatrix}.$$

Assumption G: Let $S_N^* = W_N(I_N - \rho^* W_N)^{-1}$ and $S_{ij,N}^*$ be the (i,j)th element of S_N^*. Let $\Im$ be parameters space for Λ and Σ_{ee}, which satisfies the normalization conditions, i.e.,

$$\Im = \left\{(\Lambda, \Sigma_{ee}) \;\middle|\; C^{-1} \le \sigma_i^2 \le C, \forall\, i;\ \text{and}\ \frac{1}{N}\Lambda'\Sigma_{ee}^{-1}\Lambda = I_r\right\},$$

We assume one of the following conditions:

(i) $\delta^* \neq 0$ or $\beta^* \neq 0$. Let $\widetilde{Y} = S_N^*(\delta^*\dot{Y}_{-1} + \sum_{p=1}^k \beta_p^* \dot{X}_p)$ and

$$\zeta = \left[\frac{1}{NT}\text{tr}(\widetilde{Y}'\ddot{M}\dot{Y}_{-1}M_{F^*}), \frac{1}{NT}\text{tr}(\widetilde{Y}'\ddot{M}\dot{X}_1 M_{F^*}), \cdots, \frac{1}{NT}\text{tr}(\widetilde{Y}'\ddot{M}\dot{X}_k M_{F^*})\right],$$

where ζ is a $(k+1)$-dimensional row vector. The matrix $\mathbb{D}_a = \begin{bmatrix} \frac{1}{NT}\mathrm{tr}(\widetilde{Y}'\ddot{M}\widetilde{Y}M_{F^*}) & \zeta \\ \zeta' & \mathbb{D}_b \end{bmatrix}$ is positive definite on $\Im$, where $M_{F^*} = I_T - F^*(F^{*\prime}F^*)^{-1}F^{*\prime}$ and $\ddot{M} = \Sigma_{ee}^{-1} - N^{-1}\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}$.

(ii) For all $\rho \in \mathbb{S}_\rho$ and all N,

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\Big(S_{ij,N}^{*}\sigma_j^{*2} + S_{ji,N}^{*}\sigma_i^{*2} - (\rho-\rho^*)\sum_{p=1}^{N} S_{ip,N}^{*}S_{jp,N}^{*}\sigma_p^{*2}\Big)^2 > 0, \tag{3.4}$$

and $\mathbb{D}_b$ is positive definite on $\Im$, where $\ddot{M}$ and M_{F^*} are defined the same as in (i).

Remark 3.3 Assumption G imposes the conditions for the identification of ρ and δ, β. The identification for the coefficient of spatial term is a non-trivial problem in spatial econometrics. This problem is investigated in a thorough way in Lee (2004a). Assumption G(i) can be viewed as a version of Assumption 8 of Lee (2004a) in the common shocks setting. Since the identification of ρ in Assumption G(i) depends on the underlying value of δ and β, it is a local identification condition. In contrast, Assumption G(ii) is a global identification condition. Condition (3.4) corresponds to Assumption 9 in Lee (2004a) and the condition in Theorem 2 of Yu et al. (2008), but it is different from theirs because we allow heteroskedasticity. To see this, we show in Appendix A that condition (3.4) is related to the unique solution of $\mathcal{T}_{1N}(\rho, \sigma_1^2, \ldots, \sigma_N^2) = 0$ with

$$\mathcal{T}_{1N}(\rho, \sigma_1^2, \ldots, \sigma_N^2) = -\frac{1}{2N}\mathrm{tr}[\mathcal{R}\Sigma_{ee}^{*}\mathcal{R}'\Sigma_{ee}^{-1}] + \frac{1}{2N}\ln|\mathcal{R}\Sigma_{ee}^{*}\mathcal{R}'\Sigma_{ee}^{-1}| + \frac{1}{2},$$

where $\mathcal{R} = (I_N - \rho W_N)(I_N - \rho^* W_N)^{-1}$. When homoskedasticity is assumed, $\mathcal{T}_{1N}$ reduces to $T_{1,n}$ in Yu et al. (2008). After concentrating out the common variance σ^2, $T_{1,n}$ leads to Assumption 9 in Lee (2004a) and the assumption of Theorem 2 in Yu et al. (2008). Because of heteroskedasticity our identification condition takes a different form.

Assumption H: The parameters ω and σ_i^2 for $i = 1, 2, \ldots, N$ are estimated in compact sets.

Remark 3.4 Assumption H assumes that partial parameters are estimated in compact sets. This assumption guarantees that the maximizer of the objective function is well defined. In pure factor analysis, it is known that the global maximizer of the quasi likelihood function with allowance of cross sectional heteroskedasticity do not exist, but the local maximizers are well defined and are consistent estimators for the underlying parameters under large N and large T, see, e.g., Andreson (2003). The objective function in the present paper is an extended version of the one in pure factor models and inherits the same problem. We therefore impose Assumption H to confine our analysis on local maximizers.

4 Objective function and first order conditions

Let $Z_t(\alpha, \omega, \Lambda, F) = Y_t - \alpha - \rho W_N Y_t - \delta Y_{t-1} - X_t\beta - \Lambda f_t$ with $\omega = (\rho, \delta, \beta')'$. Conditional on Y_0 which we assume are observed, the quasi likelihood function, by assuming the normality of e_{it}, is

$$\mathcal{L}^*(\theta) = -\frac{1}{2NT}\sum_{t=1}^{T} Z_t(\alpha,\omega,\Lambda,F)'\Sigma_{ee}^{-1}Z_t(\alpha,\omega,\Lambda,F) - \frac{1}{2N}\ln|\Sigma_{ee}| + \frac{1}{N}\ln|I_N - \rho W_N|.$$

where $\theta = \big(\omega, \Lambda, \mathrm{diag}(\Sigma_{ee})\big)$.[②] Given Σ_{ee}, ω and Λ, it is seen that α and f_t maximize the above function at

$$\alpha = \bar{Y} - \rho W\bar{Y} - \delta\bar{Y}_{-1} - \bar{X}\beta - \Lambda\bar{f}$$

and

$$f_t = (\Lambda'\Sigma_{ee}^{-1}\Lambda)^{-1}\Lambda'\Sigma_{ee}^{-1}(\dot{Y}_t - \rho W\dot{Y}_t - \delta\dot{Y}_{t-1} - \dot{X}_t\beta).$$

Substituting the above two equation into the preceding likelihood function to concentrate out α and f_t, the objective function can therefore be simplified as

$$\begin{aligned}\mathcal{L}(\theta) = & -\frac{1}{2NT}\sum_{t=1}^{T}(\dot{Y}_t - \rho\ddot{Y}_t - \delta\dot{Y}_{t-1} - \dot{X}_t\beta)'\ddot{M}(\dot{Y}_t - \rho\ddot{Y}_t - \delta\dot{Y}_{t-1} - \dot{X}_t\beta) \\ & -\frac{1}{2N}\ln|\Sigma_{ee}| + \frac{1}{N}\ln|I_N - \rho W_N|.\end{aligned}$$

where $\ddot{M} = \Sigma_{ee}^{-1} - \Sigma_{ee}^{-1}\Lambda(\Lambda'\Sigma_{ee}^{-1}\Lambda)^{-1}\Lambda'\Sigma_{ee}^{-1} = \Sigma_{ee}^{-1} - \frac{1}{N}\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}$ and $\ddot{Y}_t = W_N\dot{Y}_t$. The maximizer, defined by

$$\hat{\theta} = \underset{\theta\in\Theta}{\mathrm{argmax}}\ \mathcal{L}(\theta),$$

is referred to as the quasi maximum likelihood estimator or MLE, where Θ is the parameters space specified by Assumptions G and H. More specifically, Θ is defined as

$$\Theta = \left\{\theta = (\omega, \Sigma_{ee}, \Lambda)\ \middle|\ \|\omega\| \le C;\ C^{-1} \le \sigma_i^2 \le C, \forall i;\ \frac{1}{N}\Lambda'\Sigma_{ee}^{-1}\Lambda = I_r\right\}.$$

The first order condition for Λ gives

$$\left[\frac{1}{NT}\sum_{t=1}^{T}(\dot{Y}_t - \hat{\rho}\ddot{Y}_t - \hat{\delta}\dot{Y}_{t-1} - \dot{X}_t\hat{\beta})(\dot{Y}_t - \hat{\rho}\ddot{Y}_t - \hat{\delta}\dot{Y}_{t-1} - \dot{X}_t\hat{\beta})'\right]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = \hat{\Lambda}\hat{V}. \tag{4.1}$$

where $\hat{V}$ is a diagonal matrix. The first order condition for σ_i^2 gives

$$\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\left[\dot{y}_{it} - \hat{\rho}\ddot{y}_{it} - \hat{\delta}\dot{y}_{it-1} - \dot{x}_{it}'\hat{\beta} - \hat{\lambda}_i'\hat{f}_t\right]^2$$

where $\ddot{y}_{it} = \sum_{j=1}^{N} w_{ij,N}\dot{y}_{jt}$ and

$$\hat{f}_t = (\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda})^{-1}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\dot{Y}_t - \hat{\rho}\ddot{Y}_t - \hat{\delta}\dot{Y}_{t-1} - \dot{X}_t\hat{\beta}) = \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\dot{Y}_t - \hat{\rho}\ddot{Y}_t - \hat{\delta}\dot{Y}_{t-1} - \dot{X}_t\hat{\beta}).$$

The first order condition for ρ is

$$\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\dot{Y}_t - \hat{\rho}\ddot{Y}_t - \hat{\delta}\dot{Y}_{t-1} - \dot{X}_t\hat{\beta}) - \frac{1}{N}\mathrm{tr}[W_N(I_N - \hat{\rho}W_N)^{-1}] = 0.$$

[②]Strictly speaking, θ should be written as θ_N since it also depends on N. But we drop this dependence from the symbol for notational simplicity. The symbols Θ and $\Im$ below are treated in a similar way.

The first order condition for δ is

$$\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\dot{Y}_t-\hat{\rho}\ddot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\dot{X}_t\hat{\beta})=0.$$

The first order condition for β is

$$\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\dot{Y}_t-\hat{\rho}\ddot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\dot{X}_t\hat{\beta})=0.$$

We emphasize that in computing the MLE, we do not need to solve the above first order conditions. They are just for theoretical analysis.

5 Asymptotic properties of the MLE

In this section, we first show that the MLE is consistent, we then derive the convergence rates, the asymptotic representation and the limiting distributions.

Proposition 5.1 *Under Assumptions A-H, when $N,T\to\infty$*[③]*, we have*

$$\hat{\omega}\xrightarrow{p}\omega^*;$$

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^{*2})^2\xrightarrow{p}0;$$

$$\frac{1}{N}\Lambda^{*\prime}\widehat{\ddot{M}}\Lambda^*\xrightarrow{p}0.$$

where $\omega^*=(\rho^*,\delta^*,\beta^{*\prime})'$ and $\widehat{\ddot{M}}=\hat{\Sigma}_{ee}^{-1}-N^{-1}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}$.

In the analysis of panel data models with common shocks but without spatial effects, a difficult problem is to establish consistency. The parameters of interest (δ,β) are simultaneously estimated with high dimensional nuisance parameters Λ and Σ_{ee}. The usual arguments need some modifications to accommodate this feature. The presence of spatial effects further compounds the difficult. Our proof of Proposition 5.1 consists of three steps. First we show there exists a function $\mathcal{L}_1(\theta)$ such that

$$\sup_{\theta\in\Theta}|\mathcal{L}(\theta)-\mathcal{L}_1(\theta)|\xrightarrow{p}0.$$

Then we show that the function $\mathcal{L}_1(\theta)$ possesses the property that there exists an $\epsilon>0$, which depends on the $\mathcal{N}^c(\omega^*)$, such that

$$\sup_{(\Lambda,\Sigma_{ee})\in\Im}\ \sup_{\omega\in\mathcal{N}^c(\omega^*)}\mathcal{L}_1(\theta)-\mathcal{L}_1(\theta^*)<-\epsilon,$$

[③] In this paper, when we say the limit we mean the joint limit, which is the limit by letting N and T pass to infinity simultaneously, without naming the order that which index diverges first and which one diverges next. The latter case is called the sequential limit in the literature. Readers are referred to Phillips and Moon (1999) for a formal and precise definition of the two types of limit. See also the definition O_p and o_p in Appendix A.

where $\mathcal{N}^c(\omega^*)$ is the complement of an open neighborhood of ω^*. Given the above two results, we have $\hat{\omega} \xrightarrow{p} \omega^*$. After obtaining the consistency of $\hat{\omega}$, in the third step we show the remaining two results in Proposition 5.1.

Notice that ω is low-dimensional but Σ_{ee} and Λ are high dimensional. So the usual consistency concept applies for ω. But for Σ_{ee} and Λ, their consistencies can only be defined under some chosen norm. The second result is equivalent to $\frac{1}{N}\|\hat{\Sigma}_{ee} - \Sigma_{ee}^*\|^2 \xrightarrow{p} 0$. So the chosen norm is dimension-adjusted frobenious norm. The norm used in the last result can be viewed as an extension of generalized square coefficient between two high-dimensional vectors. We choose this norm to take account of rotational indeterminacy on factor and factor loadings, see Bai and Li (2012) for discussions on rotational indeterminacy in factor analysis.

The consistency result allows us to further derive the rates of convergence.

Theorem 5.1 *Let $H = \frac{1}{NT}\hat{V}^{-1}(\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda^*)(F^{*\prime}F^*)$. Under Assumptions A-H, when $N, T \to \infty$, we have*

$$\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i^*\|^2 = O_p(\frac{1}{N^2}) + O_p(T^{-1});$$
$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^{*2})^2 = O_p(\frac{1}{N^2}) + O_p(T^{-1});$$
$$\hat{\omega} - \omega^* = O_p(N^{-1}) + O_p(T^{-1}).$$

where $\hat{V}$ is defined in (4.1).

It is well documented in econometric literature that the MLE for dynamic panel data models has a $O(\frac{1}{T})$ bias term, see, for example, Hahn and Kuersteiner (2002) and Alvarez and Arellano (2003). The case with inclusion of spatial term and lag spatial term has been investigated by Yu et al. (2008), which shows that the bias term is still $O(\frac{1}{T})$ but the expression is related with spatial weights matrix. This bias term is inherited by our MLE, as we can see that model (1.1) is an extension of classical spatial dynamic models. Apart from this $O(\frac{1}{T})$ bias term, our analysis indicates that there is another $O(\frac{1}{N})$ bias arising from common shocks part Λf_t. The presence of biases in the MLE is due to incidental parameters problem, see Neyman and Scott (1948) for a general discussion.

To state the asymptotic properties of the MLE, we define the following notations:

$$B_t = \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^* \dot{X}_{t-l}\beta^* + \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^* \Lambda^* \dot{f}_{t-l}^*, \qquad \dot{B}_t = B_t - \frac{1}{T}\sum_{s=1}^{T} B_s$$
$$\ddot{B}_t = W_N \dot{B}_t, \qquad Q_t = \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^* e_{t-l}, \qquad J_t = S_N^* \sum_{l=1}^{\infty}(\delta^* G_N^*)^l e_{t-l}.$$

Now we state the main theorem in this paper, which gives the asymptotic representation of $\hat{\omega} - \omega$.

Theorem 5.2 *Under Assumptions A-H, when $N, T \to \infty$ and $\sqrt{N}/T \to 0, \sqrt{T}/N \to 0$, we have*

$$\sqrt{NT}(\hat{\omega} - \omega^* + b) = \mathbb{D}^{-1}\xi + o_p(1),$$

where

$$\xi = \frac{1}{\sqrt{NT}} \begin{bmatrix} \sum_{t=1}^{T} \ddot{B}_t' \ddot{M}^* e_t - \sum_{t=1}^{T}\sum_{s=1}^{T} \ddot{B}_t' \ddot{M}^* e_s \pi_{st}^* + \sum_{t=1}^{T} J_t' \Sigma_{ee}^{*-1} e_t + \eta \\ \sum_{t=1}^{T} \dot{B}_{t-1}' \ddot{M}^* e_t - \sum_{t=1}^{T}\sum_{s=1}^{T} \dot{B}_{t-1}' \ddot{M}^* e_s \pi_{st}^* + \sum_{t=1}^{T} Q_{t-1}' \Sigma_{ee}^{*-1} e_t \\ \sum_{t=1}^{T} \dot{X}_t' \ddot{M}^* e_t - \sum_{t=1}^{T}\sum_{s=1}^{T} \dot{X}_t' \ddot{M}^* e_s \pi_{st}^* \end{bmatrix}$$

with $\pi_{st}^* = f_s^{*\prime}(F^{*\prime}F^*)^{-1}f_t^*$ *and* $\ddot{M}^* = \Sigma_{ee}^{*-1} - \frac{1}{N}\Sigma_{ee}^{*-1}\Lambda^*\Lambda^{*\prime}\Sigma_{ee}^{*-1}$. *The* $(k+2)\times(k+2)$ *matrix* $\mathbb{D}$ *is defined as*

$$\mathbb{D} = \frac{1}{NT} \begin{bmatrix} \mathrm{tr}(\ddot{Y}'\ddot{M}^*\ddot{Y}M_{F^*}) + \Phi & \mathrm{tr}(\ddot{Y}'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \mathrm{tr}(\ddot{Y}'\ddot{M}^*\dot{X}_1 M_{F^*}) & \cdots & \mathrm{tr}(\ddot{Y}'\ddot{M}^*\dot{X}_k M_{F^*}) \\ \mathrm{tr}(\dot{Y}_{-1}'\ddot{M}^*\ddot{Y}M_{F^*}) & \mathrm{tr}(\dot{Y}_{-1}'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \mathrm{tr}(\dot{Y}_{-1}'\ddot{M}^*\dot{X}_1 M_{F^*}) & \cdots & \mathrm{tr}(\dot{Y}_{-1}'\ddot{M}^*\dot{X}_k M_{F^*}) \\ \mathrm{tr}(\dot{X}_1'\ddot{M}^*\ddot{Y}M_{F^*}) & \mathrm{tr}(\dot{X}_1'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \mathrm{tr}(\dot{X}_1'\ddot{M}^*\dot{X}_1 M_{F^*}) & \cdots & \mathrm{tr}(\dot{X}_1'\ddot{M}^*\dot{X}_k M_{F^*}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{tr}(\dot{X}_k'\ddot{M}^*\ddot{Y}M_{F^*}) & \mathrm{tr}(\dot{X}_k'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \mathrm{tr}(\dot{X}_k'\ddot{M}^*\dot{X}_1 M_{F^*}) & \cdots & \mathrm{tr}(\dot{X}_k'\ddot{M}^*\dot{X}_k M_{F^*}) \end{bmatrix}$$

with $\Phi = T[\mathrm{tr}(S_N^{*2}) - 2\sum_{i=1}^{N} S_{ii,N}^{*2}]$. *The* $(k+2)$*-dimensional vector* b *is defined as*

$$b = \mathbb{D}^{-1} \begin{bmatrix} \frac{1}{N}\mathrm{tr}[\Lambda^{*\prime}S_N^{\circ\prime}\Sigma_{ee}^{*-1}\Lambda^*(\Lambda^{*\prime}\Sigma_{ee}^{*-1}\Lambda^*)^{-1}] + \frac{1}{NT}\mathrm{tr}[P_{\widetilde{F}}K] \\ \frac{1}{NT}\mathrm{tr}[P_{\widetilde{F}}L] \\ 0_{k\times 1} \end{bmatrix}$$

with $P_{\widetilde{F}} = \widetilde{F}(\widetilde{F}'\widetilde{F})^{-1}\widetilde{F}'$ *and* $\widetilde{F} = (F^*, \mathbf{1}_T)$. *Here* $\mathbf{1}_T$ *is a* T*-dimensional vector with all its elements being 1. In addition,*

$$K = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \mathrm{tr}[S_N^*(\delta^* G_N^*)] & 0 & 0 & \cdots & 0 \\ \mathrm{tr}[S_N^*(\delta^* G_N^*)^2] & \mathrm{tr}[S_N^*(\delta^* G_N^*)] & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathrm{tr}[S_N^*(\delta^* G_N^*)^{T-1}] & \mathrm{tr}[S_N^*(\delta^* G_N^*)^{T-2}] & \mathrm{tr}[S_N^*(\delta^* G_N^*)^{T-3}] & \cdots & 0 \end{bmatrix},$$

and

$$L = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \mathrm{tr}(G_N^*) & 0 & 0 & \cdots & 0 \\ \mathrm{tr}[G_N^*(\delta^* G_N^*)] & \mathrm{tr}(G_N^*) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathrm{tr}[G_N^*(\delta^* G_N^*)^{T-2}] & \mathrm{tr}[G_N^*(\delta^* G_N^*)^{T-3}] & \mathrm{tr}[G_N^*(\delta^* G_N^*)^{T-4}] & \cdots & 0 \end{bmatrix},$$

and

$$\eta = \sum_{t=1}^{T} e_t' S_N^{*\circ\prime}\Sigma_{ee}^{*-1} e_t = \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} \frac{1}{\sigma_i^{*2}} 1(i \neq j) e_{it} e_{jt} S_{ij,N}^*.$$

Here $S_N^{*\circ}$ *is an* $N \times N$ *matrix which is obtained by setting all the diagonal elements of* S_N^* *to zeros.*

Although ξ has a relatively complicated expression, it can be shown that $\mathbb{D}^{-1/2}\xi \xrightarrow{d} N(0, I_{k+2})$ by resorting to the martingale difference central limit theorem (see Corollary 3.1 in Hall and Heyde (1980)). Appendix E gives a detailed derivation. Given this result, we have the following corollary.

Corollary 5.1 *Under the assumptions in Theorem 5.2, when $N, T \to \infty$ and $N/T \to \kappa^2$, we have*

$$\sqrt{NT}(\hat{\omega} - \omega^*) \xrightarrow{d} N\Big(- b^{\diamond}, \Big[\operatorname*{plim}_{N,T\to\infty} \mathbb{D} \Big]^{-1} \Big),$$

where

$$b^{\diamond} = \operatorname*{plim}_{N,T\to\infty} \left\{ \mathbb{D}^{-1} \begin{bmatrix} \frac{1}{\kappa}\mathrm{tr}[\Lambda^{*\prime} S_N^{\circ\prime} \Sigma_{ee}^{*-1} \Lambda^* (\Lambda^{*\prime}\Sigma_{ee}^{*-1}\Lambda^*)^{-1} + \kappa \mathrm{tr}[\frac{1}{N} P_{\widetilde{F}} K] \\ \kappa \mathrm{tr}[\frac{1}{N} P_{\widetilde{F}} L] \\ 0_{k\times 1} \end{bmatrix} \right\}.$$

Theorem 5.2 include some important models as special cases. If there are no lag dependent term and spatial term in model (1.1), i.e.,

$$y_{it} = \alpha_i + x_{it}'\beta + \lambda_i' f_t + e_{it},$$

the present analysis indicates that under $\sqrt{N}/T \to 0$, $\sqrt{T}/N \to 0$ as well as other regularity conditions, the asymptotic representation of $\hat{\beta} - \beta$ is

$$\sqrt{NT}(\hat{\beta} - \beta^*) = \mathbb{D}_{\beta}^{-1} \frac{1}{\sqrt{NT}} \left(\sum_{t=1}^{T} \dot{X}_t' \ddot{M}^* e_t - \sum_{t=1}^{T}\sum_{s=1}^{T} \dot{X}_t' \ddot{M}^* e_s \pi_{st}^* \right) + o_p(1),$$

where

$$\mathbb{D}_{\beta} = \frac{1}{NT} \begin{bmatrix} \mathrm{tr}(\dot{X}_1' \ddot{M}^* \dot{X}_1 M_{F^*}) & \cdots & \mathrm{tr}(X_1' \ddot{M}^* X_k M_{F^*}) \\ \vdots & \ddots & \vdots \\ \mathrm{tr}(\dot{X}_k' \ddot{M}^* \dot{X}_1 M_{F^*}) & \cdots & \mathrm{tr}(X_k' \ddot{M}^* X_k M_{F^*}) \end{bmatrix}.$$

It is seen that the MLE is asymptotically free of bias. This extends the analysis of Bai (2009), which shows that the profile MLE has no bias in asymptotics if the error e_{it} is independent and identically distributed over the time and cross section dimensions. When lag dependent variable is included but the spatial term is absent, the MLE would have an identical limiting variance representation as the above, if we treat lag dependent variable as an additional exogenous regressor. But the MLE is no longer unbiased. The bias term is $\frac{1}{T}\mathrm{tr}(P_{\widetilde{F}} L^{\dagger}) \mathbb{D}_{\phi}^{-1} \iota_{k+1}$ if we label the lag dependent variable as the first regressor, where $\mathbb{D}_{\phi}^{-1}$ is the limiting variance of $\hat{\phi} = (\hat{\delta}, \hat{\beta}')'$; $L^{\dagger}$ is defined as

$$L^{\dagger} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \delta & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \delta^{T-2} & \delta^{T-3} & \delta^{T-4} & \cdots & 0 \end{bmatrix};$$

and ι_{k+1} is the first column of the $k+1$ dimensional identity matrix. Moon and Weidner (2013) consider a similar model by assuming cross sectional homoskedasticity. Our results are derived under cross sectional heteroskedasticity.

Remark 5.1 A specification of practical relevance, which is widely used in social interaction studies, is

$$Y_t = \alpha + \rho W_N Y_t + \delta Y_{t-1} + X_t \beta + W_N X_t \gamma + \Lambda f_t + e_t.$$

As pointed out by an array of studies (Lee (2007), Bramoullé et al. (2009), Lin (2010), ect.), ρ captures the endogenous effect and γ the contextual effect in terms of Manski (1993). Let $\widetilde{X}_t = (X_t, W_N X_t)$ and $\widetilde{\beta} = (\beta', \gamma')'$, we see that the above model is equivalent to

$$Y_t = \alpha + \rho W_N Y_t + \delta Y_{t-1} + \widetilde{X}_t \widetilde{\beta} + \Lambda f_t + e_t.$$

If $\widetilde{X}_t$ satisfies Assumption G, Theorem 5.2 applies.

Remark 5.2 Under Assumptions E and F, Y_t has a well-defined MA(∞) expression:

$$Y_t = \sum_{l=0}^{\infty} (\delta^* G_N^*)^l G_N^* \alpha^* + \sum_{l=0}^{\infty} (\delta^* G_N^*)^l G_N^* X_{t-l} \beta^* + \sum_{l=0}^{\infty} (\delta^* G_N^*)^l G_N^* \Lambda^* f_{t-l}^* + \sum_{l=0}^{\infty} (\delta^* G_N^*)^l G_N^* e_{t-l}.$$

Given the above results, we have

$$\frac{\partial Y_t}{\partial X'_{t-s,p}} = (\delta^* G_N^*)^s G_N^* \beta_p^*; \qquad \frac{\partial Y_t}{\partial e'_{t-s}} = (\delta^* G_N^*)^s G_N^*.$$

where $X_{t-s,p}$ denote the pth column of X_{t-s} ($p = 1, 2, \ldots, k$). The above result implies

$$\frac{\partial y_{it}}{\partial x_{j(t-s)p}} = [(\delta^* G_N^*)^s G_N^* \beta_p^*]_{ij}; \qquad \frac{\partial y_{it}}{\partial e_{j(t-s)}} = [(\delta^* G_N^*)^s G_N^*]_{ij}.$$

where we use $[M]_{ij}$ to denote the (i, j)th element of M. So the marginal effects of $x_{j(t-s)p}$ and $e_{j(t-s)}$ on y_{it} can be estimated according to the above formulas by plug-in method. The limiting distributions of the marginal effects can be easily calculated by the delta method via Theorem 5.2.

Remark 5.3 The limiting variance and the bias term can be estimated by plug-in method. More specifically, matrix $\mathbb{D}$ can be consistently estimated by

$$\widehat{\mathbb{D}} = \frac{1}{NT} \begin{bmatrix} \mathrm{tr}(\ddot{Y}' \widehat{\ddot{M}} \ddot{Y} M_{\hat{F}}) + \hat{\Phi} & \mathrm{tr}(\ddot{Y}' \widehat{\ddot{M}} \dot{Y}_{-1} M_{\hat{F}}) & \mathrm{tr}(\ddot{Y}' \widehat{\ddot{M}} \dot{X}_1 M_{\hat{F}}) & \cdots & \mathrm{tr}(\ddot{Y}' \widehat{\ddot{M}} \dot{X}_k M_{\hat{F}}) \\ \mathrm{tr}(\dot{Y}'_{-1} \widehat{\ddot{M}} \ddot{Y} M_{\hat{F}}) & \mathrm{tr}(\dot{Y}'_{-1} \widehat{\ddot{M}} \dot{Y}_{-1} M_{\hat{F}}) & \mathrm{tr}(\dot{Y}'_{-1} \widehat{\ddot{M}} \dot{X}_1 M_{\hat{F}}) & \cdots & \mathrm{tr}(\dot{Y}'_{-1} \widehat{\ddot{M}} \dot{X}_k M_{\hat{F}}) \\ \mathrm{tr}(\dot{X}'_1 \widehat{\ddot{M}} \ddot{Y} M_{\hat{F}}) & \mathrm{tr}(\dot{X}'_1 \widehat{\ddot{M}} \dot{Y}_{-1} M_{\hat{F}}) & \mathrm{tr}(\dot{X}'_1 \widehat{\ddot{M}} \dot{X}_1 M_{\hat{F}}) & \cdots & \mathrm{tr}(\dot{X}'_1 \widehat{\ddot{M}} \dot{X}_k M_{\hat{F}}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{tr}(\dot{X}'_k \widehat{\ddot{M}} \ddot{Y} M_{\hat{F}}) & \mathrm{tr}(\dot{X}'_k \widehat{\ddot{M}} \dot{Y}_{-1} M_{\hat{F}}) & \mathrm{tr}(\dot{X}'_k \widehat{\ddot{M}} \dot{X}_1 M_{\hat{F}}) & \cdots & \mathrm{tr}(\dot{X}'_k \widehat{\ddot{M}} \dot{X}_k M_{\hat{F}}) \end{bmatrix}$$

where

$$\hat{F} = \frac{1}{N} \Big(\dot{Y} - \hat{\delta} \dot{Y}_{-1} - \hat{\rho} \ddot{Y} - \sum_{p=1}^{k} \dot{X}_p \hat{\beta}_p \Big)' \hat{\Sigma}_{ee}^{-1} \hat{\Lambda}$$

and $\hat{\Phi} = T \cdot \mathrm{tr}[(\hat{S}_N^2) - 2 \sum_{i=1}^N \hat{S}_{ii,N}^2]$ with $\hat{S}_N = W_N \hat{G}_N$, $\hat{G}_N = (I_N - \hat{\rho} W_N)^{-1}$ and $\hat{S}_{ii,N}$ being the ith diagonal element of $\hat{S}_N$. In addition, the bias term b can be consistently estimated by

$$\hat{b} = \hat{\mathbb{D}}^{-1} \begin{bmatrix} \frac{1}{N} \mathrm{tr}[\hat{\Lambda}' \hat{S}_N^{\circ\prime} \hat{\Sigma}_{ee}^{-1} \hat{\Lambda} (\hat{\Lambda}' \hat{\Sigma}_{ee}^{-1} \hat{\Lambda})^{-1}] + \frac{1}{NT} \mathrm{tr}[P_{\widehat{\widehat{F}}} \hat{K}] \\ \frac{1}{NT} \mathrm{tr}[\frac{1}{NT} \mathrm{tr}[P_{\widehat{\widehat{F}}} \hat{L}] \\ 0_{k \times 1} \end{bmatrix}.$$

where $\hat{K}$ and $\hat{L}$ are defined similarly as K and L except that δ^*, G_N^* and S_N^* are replaced with $\hat{\delta}$, $\hat{G}_N$ and $\hat{S}_N$ respectively, and $\widehat{\widehat{F}} = [\hat{F}, \mathbf{1}_T]$.

6 Discussions on spatial models with heteroskedasticity

Allowance of heteroskedasticy in pure spatial models is of theoretical and practical relevance. As pointed out by Kelejian and Prucha (2010) and Lin and Lee (2010) among others, if heteroskedasticity exists but homoskedasticity is imposed, the MLE generally is inconsistent. In viewpoint of applied studies, assuming homoskedasticity seems too restrictive to be true. However, to the best of our knowledge, the MLE under heteroskedasticity has not been investigated so far in the literature. In this section, we give some discussions on this issue, which is of independent interest.

6.1 Dynamic spatial models

Consider the following dynamic spatial model,

$$y_{it} = \alpha_i + \rho \sum_{j=1}^{N} w_{ij,N} y_{jt} + \delta y_{it-1} + e_{it}. \tag{6.1}$$

The above model is special case of model (1.1). Under some regularity conditions stated in Section 2, the analysis of Theorem 5.2 indicates that the MLE for (6.1) has the following asymptotic representation:

$$\sqrt{NT}(\hat{\omega} - \omega + v_1) = \mathcal{D}_1^{-1} \frac{1}{\sqrt{NT}} \begin{bmatrix} \sum_{t=1}^{T} (\ddot{B}_t + J_t + S_N^{*\circ} e_t)' \Sigma_{ee}^{*-1} e_t \\ \sum_{t=1}^{T} (\dot{B}_{t-1} + J_{t-1})' \Sigma_{ee}^{*-1} e_t \\ \sum_{t=1}^{T} \dot{X}_t' \Sigma_{ee}^{*-1} e_t \end{bmatrix} + o_p(1), \tag{6.2}$$

where

$$\mathcal{D}_1 = \frac{1}{NT} \begin{bmatrix} \sum_{t=1}^{T} \ddot{Y}_t' \Sigma_{ee}^{*-1} \ddot{Y}_t + \Phi & \sum_{t=1}^{T} \ddot{Y}_t' \Sigma_{ee}^{*-1} \dot{Y}_{t-1} & \sum_{t=1}^{T} \ddot{Y}_t' \Sigma_{ee}^{*-1} \dot{X}_t \\ \sum_{t=1}^{T} \dot{Y}_{t-1}' \Sigma_{ee}^{*-1} \ddot{Y}_t & \sum_{t=1}^{T} \dot{Y}_{t-1}' \Sigma_{ee}^{*-1} \dot{Y}_{t-1} & \sum_{t=1}^{T} \dot{Y}_{t-1}' \Sigma_{ee}^{*-1} \dot{X}_t \\ \sum_{t=1}^{T} \dot{X}_t' \Sigma_{ee}^{*-1} \ddot{Y}_t & \sum_{t=1}^{T} \dot{X}_t' \Sigma_{ee}^{*-1} \dot{Y}_{t-1} & \sum_{t=1}^{T} \dot{X}_t' \Sigma_{ee}^{*-1} \dot{X}_t \end{bmatrix},$$

with Φ defined the same as in Theorem 5.2 and

$$v_1 = \mathcal{D}_1^{-1} \begin{bmatrix} \frac{1}{NT} \mathrm{tr}[\delta^* S_N^* G_N^* (I_N - \delta^* G_N^*)^{-1}] \\ \frac{1}{NT} \mathrm{tr}[G_N^* (I_N - \delta^* G_N^*)^{-1}] \\ 0_{k\times 1} \end{bmatrix}.$$

Given the above asymptotic representation, invoking the central limiting theorem for quadratic form (Kelejian and Prucha (2001), Giraitis and Taqqu (1998)), we have

$$\sqrt{NT}(\hat{\omega} - \omega^* + v_1) \xrightarrow{d} N\Big(0, \Big[\operatorname*{plim}_{N,T\to\infty} \mathcal{D}_1\Big]^{-1}\Big).$$

6.2 Spatial panel data models with SAR disturbances

Another interesting spatial model, which receive much attention in practice, is spatial panel data model with SAR disturbances, i.e.,

$$\begin{aligned} Y_t &= \alpha + \rho W_N Y_t + X_t \beta + u_t; \\ u_t &= \varrho M_N u_t + e_t. \end{aligned} \tag{6.3}$$

where M_N is another spatial weights matrix. Lee and Yu (2010) make a rigorous analysis for the ML estimation of (6.3) under the assumption that e_{it} is cross-sectionally homoskedastic. Using the method in this paper to deal with high dimensional variance parameters, [4] we can extend Lee and Yu's analysis to heteroskedasticity. For ease of exposition, we further introduce the following notations. Let

$$\mathcal{F} = M_N S_N^*(I_N - \varrho^* M_N)^{-1}, \ \mathcal{G} = (I_N - \varrho^* M_N) S_N^*(I_N - \varrho^* M_N)^{-1}, \ \mathcal{H} = M_N(I_N - \varrho^* M_N)^{-1};$$

$$\mathcal{P}_t = (I_N - \varrho^* M_N) W_N \dot{Y}_t, \quad \mathcal{Q}_t = M_N[(I_N - \rho^* W_N)\dot{Y}_t - \dot{X}_t \beta^*], \quad \mathcal{R}_t = (I_N - \varrho^* M_N)\dot{X}_t.$$

Define the $(k+2) \times (k+2)$ matrix $\mathcal{D}_2$ as

$$\mathcal{D}_2 = \frac{1}{NT} \begin{bmatrix} \sum_{t=1}^T \mathcal{P}_t' \Sigma_{ee}^{-1} \mathcal{P}_t + \varsigma_1 & \sum_{t=1}^T \mathcal{P}_t' \Sigma_{ee}^{-1} \mathcal{Q}_t + \varsigma_2 & \sum_{t=1}^T \mathcal{P}_t' \Sigma_{ee}^{-1} \mathcal{R}_t \\ \sum_{t=1}^T \mathcal{Q}_t' \Sigma_{ee}^{-1} \mathcal{P}_t + \varsigma_2 & \sum_{t=1}^T \mathcal{Q}_t' \Sigma_{ee}^{-1} \mathcal{Q}_t + \varsigma_3 & \sum_{t=1}^T \mathcal{Q}_t' \Sigma_{ee}^{-1} \mathcal{R}_t \\ \sum_{t=1}^T \mathcal{R}_t' \Sigma_{ee}^{-1} \mathcal{P}_t & \sum_{t=1}^T \mathcal{R}_t' \Sigma_{ee}^{-1} \mathcal{Q}_t & \sum_{t=1}^T \mathcal{R}_t' \Sigma_{ee}^{-1} \mathcal{R}_t \end{bmatrix}$$

with $\varsigma_1 = T[\text{tr}(\mathcal{G}^2) - 2\text{tr}(\mathcal{G} \circ \mathcal{G})]$, $\varsigma_2 = T[\text{tr}(\mathcal{F}) - 2\text{tr}(\mathcal{G} \circ \mathcal{H})]$ and $\varsigma_3 = T[\text{tr}(\mathcal{H}^2) - 2\text{tr}(\mathcal{H} \circ \mathcal{H})]$, where "$\circ$" denotes the Hadamard product.

Under some regularity conditions, we can show that the MLE for $\omega = (\rho, \varrho, \beta')'$ in (6.3) under cross sectional heteroskedasticity has the following asymptotic representation,

$$\sqrt{NT}(\hat{\omega} - \omega^*) = \mathcal{D}_2^{-1} \frac{1}{\sqrt{NT}} \begin{bmatrix} \sum_{t=1}^T [\beta^{*\prime} \dot{X}_t' S_N^{*\prime} (I_N - \varrho^* M_N)' + e_t' \mathcal{G}^{\circ\prime}] \Sigma_{ee}^{*-1} e_t \\ \sum_{t=1}^T e_t' \mathcal{H}^{\circ\prime} \Sigma_{ee}^{*-1} e_t \\ \sum_{t=1}^T \dot{X}_t' (I_N - \varrho^* M_N)' \Sigma_{ee}^{*-1} e_t \end{bmatrix} + o_p(1), \tag{6.4}$$

where $\mathcal{G}^\circ$ and $\mathcal{H}^\circ$ are defined similarly as $S_N^{*\circ}$. Given the above result, invoking the central limit theorem for quadratic form, we have

$$\sqrt{NT}(\hat{\omega} - \omega) \xrightarrow{d} N\Big(0, \Big[\operatorname*{plim}_{N,T\to\infty} \mathcal{D}_2\Big]^{-1}\Big).$$

6.3 Homoskedasticity versus heteroskedasticity

It is seen from the above that the limiting variance of the MLE is not a sandwich form. This result contrasts with the existing results in the literature such as Yu et al. (2008) and Lee and Yu (2010), in which the limiting variance of the MLE has a sandwich formula. The reason for the difference is the heteroskedasticity estimation. In the present paper we allow cross-sectional heteroskedasticity, while Yu et al. (2008) assume homoskedasticity. Under heteroskedasticity, the asymptotic expression does not involve e_{it}^2, as seen in (6.2) and (6.4). But under homoskedasticity, the situation is different. Still consider model (6.1). If homoskedasticity is assumed and is imposed in estimation (let $\sigma^{*2} = E(e_{it}^2)$), the

[4] The method to deal with high dimensional variance parameters σ_i^2 is as follows: First show $\frac{1}{N}\sum_{i=1}^N (\hat{\sigma}_i^2 - \sigma_i^2)^2 = o_p(1)$, see Proposition 5.1; then derive its convergence rate, see Propositions B.4 and B.6; then use this result to show that the magnitude of the difference between the term involving $\hat{\Sigma}_{ee}$ and the term involving Σ_{ee} is asymptotically negligible.

asymptotic expression for the MLE is

$$\sqrt{NT}(\tilde{\omega} - \omega^* + v_2) = \mathcal{D}_3^{-1} \frac{1}{\sqrt{NT}\sigma^{*2}} \begin{bmatrix} \sum_{t=1}^T e_t' S_N^{*\circ\prime} e_t + \sum_{t=1}^T J_t' e_t + \sum_{t=1}^T \ddot{B}_t^{\star\prime} e_t + \vartheta \\ \sum_{t=1}^T \dot{B}_{t-1}^{\star\prime} e_t + \sum_{t=1}^T Q_{t-1}' e_t \\ \sum_{t=1}^T \dot{X}_t' e_t \end{bmatrix} + o_p(1),$$

where

$$\vartheta = \sum_{i=1}^N \sum_{t=1}^T \left[S_{ii,N}^* - \frac{1}{N} \mathrm{tr}(S_N^*) \right] (e_{it}^2 - \sigma^{*2}), \quad B_t^\star = \sum_{l=0}^\infty (\delta^* G_N^*)^l X_{t-l} \beta^*, \quad \ddot{B}_t^\star = W_N \dot{B}_t^\star,$$

$$v_2 = \mathcal{D}_3^{-1} \left[\frac{1}{NT} \mathrm{tr}\Big(\delta^* S_N^* G_N^* (I_N - \delta^* G_N^*)^{-1} \Big), \frac{1}{NT} \mathrm{tr}\Big(G_N^* (I_N - \delta^* G_N^*)^{-1} \Big), 0_{1\times k} \right]',$$

and

$$\mathcal{D}_3 = \frac{1}{NT\sigma^{*2}} \begin{bmatrix} \sum_{t=1}^T \ddot{Y}_t' \ddot{Y}_t + T\sigma^{*2} \{ \mathrm{tr}(S_N^{*2}) - \frac{2}{N} [\mathrm{tr}(S_N^*)]^2 \} & \sum_{t=1}^T \ddot{Y}_t' \dot{Y}_{t-1} & \sum_{t=1}^T \ddot{Y}_t' \dot{X}_t \\ \sum_{t=1}^T \dot{Y}_{t-1}' \ddot{Y}_t & \sum_{t=1}^T \dot{Y}_{t-1}' \dot{Y}_{t-1} & \sum_{t=1}^T \dot{Y}_{t-1}' \dot{X}_t \\ \sum_{t=1}^T \dot{X}_t' \ddot{Y}_t & \sum_{t=1}^T \dot{X}_t' \dot{Y}_{t-1} & \sum_{t=1}^T \dot{X}_t' \dot{X}_t \end{bmatrix}$$

From the above, we can see that the asymptotic expression under the homoskedasticity involves e_{it}^2. So the limiting variance of $\tilde{\omega} - \omega^*$ will depend on the kurtosis of e_{it}. Because $\mathcal{D}_3$ does not depend on the kurtosis, the limiting variance of $\tilde{\omega} - \omega^*$ has a sandwich formula. In contrast, the MLE under heteroskedasticity has a limiting variance not of a sandwich form, regardless of normality. The same phenomenon also occurs for the spatial panel data models with SAR disturbances, see Lee and Yu (2010) for the asymptotic result of the MLE under homoskedasticity. This results is interesting. Thus estimating heteroskedasticity is desirable from two considerations: the limiting distribution is robust to the underlying distributions; it avoids potential inconsistency when homoskedasticity is incorrectly imposed.

7 Finite sample properties

In this section, we run Monte Carlo simulations to investigate the finite sample properties of the MLE. The data are generated according to

$$y_{it} = \alpha_i + \rho \sum_{j=1}^N w_{ij,N} y_{jt} + \delta y_{it-1} + x_{it1}\beta_1 + x_{it2}\beta_2 + \lambda_i' f_t + e_{it}$$

with $(\rho, \delta, \beta_1, \beta_2) = (0.5, 0.4, 1, 2)$. The number of factors is fixed to 2. The explanatory variable x_{itp} is generated according to

$$x_{itp} = \left[(\lambda_i + \gamma_{ip})' f_t + u_{itp} \right] 1 \left[(\lambda_i + \gamma_{ip})' f_t + u_{itp} \geq -3.5 \right]$$

for $p = 1, 2$. All the elements of $\alpha_i, \lambda_i, f_t, \gamma_{ip}$ and u_{itp} are all generated independently from $N(0,1)$. The way to generate the explanatory variables here is similar as in Moon and Weidner (2013). To generate the errors and heteroskedasticity, we consider the method similar as in Bai and Li (2014b). More specifically, we set $e_{it} = \sqrt{\psi_i}\varepsilon_{it}$ where ψ_i is defined as

$$\psi_i = 0.5 + \frac{1-\nu_i}{\nu_i}\lambda_i'\lambda_i,$$

where ν_i is drawn independently from $U[0.2, 0.8]$. The error ε_{it} is equal to $(\chi_2^2 - 2)/2$, where χ_2^2 denotes the chi-squared distribution with two degrees of freedom, which is normalized to zero mean and unit variance.

The generated data exhibit heteroskedasticity. The generated x_{it} does not have a factor structure and is correlated with the factors and factor loadings, and the two regressors x_{it1} and x_{it2} are also correlated; the errors are non-normal and skewed.

The spatial weights matrices generated in the simulation are similar to Kelejian and Prucha (1999) and Kapoor et al. (2007). More specifically, all the units are arranged in a circle and each unit is affected only by the q units immediately before it and immediately after it with equal weight. Following Kelejian and Prucha (1999), we normalize the spatial weights matrix by letting the sum of each row be equal to 1 (so the weight is $\frac{1}{2q}$) and call this specification of spatial weights matrix "q ahead and q behind."

Adapting a criterion in Bai and Li (2014b), the number of factors is determined by

$$\hat{r} = \underset{0\le m\le r_{max}}{\operatorname{argmin}} IC(m)$$

with

$$IC(m) = \frac{1}{2N}\sum_{i=1}^{N}\ln|(\hat{\sigma}_i^m)^2| - \frac{1}{N}\ln|I_N - \hat{\rho}^m W_N| + m\frac{N+T}{2NT}\ln[\min(N,T)].$$

and

$$(\hat{\sigma}_i^m)^2 = \frac{1}{T}\sum_{t=1}^{T}(\dot{y}_{it} - \hat{\rho}^m \ddot{y}_{it} - \hat{\delta}^m \dot{y}_{it-1} - \dot{x}_{it}'\hat{\beta} - \hat{\lambda}_i^{m\prime}\hat{f}_t^m)^2,$$

where the hat symbols with superscript "m" denotes the MLE when the number of factors is set to m. We set $r_{\max} = 4$.

The following four tables present the simulation results from 1000 repetitions under the combinations of $N = 100, 125, 150$ and $T = 75, 100, 125$. To measure the performance, we compute biases and root mean square errors (RMSE), which are defined as follows. We take ρ as the example to illustrate.

$$\text{Bias} = \frac{1}{\nu}\sum_{s=1}^{\nu}\hat{\rho}^{(s)} - \rho^*, \qquad \text{RMSE} = \sqrt{\frac{1}{\nu}\sum_{s=1}^{\nu}(\hat{\rho}^{(s)} - \rho^*)^2}.$$

where $\hat{\rho}^{(s)}$ is the estimator for ρ^* in the sth repetition and ν is the number of repetitions.

In all the simulations, the number of factors can be correctly estimated with probability almost one. The first two tables report the performance of the MLE before and after the bias correction under "1 ahead and 1 behind" spatial weights matrix. From Table 1, we see that the MLE are consistent. As the sample size becomes larger, the RMSEs of the MLE decrease stably. However, we also find that the ratio of the bias relative to the RMSE for the MLE of δ is considerably large, the ratio for ρ, albeit not as large as δ, is still pronounced, especially when N/T is large. This causes problems in statistical inference. We then investigates the performance of the bias-corrected MLE. From Table 2, we see that the bias-correct estimator performs well. The biases of the original estimators have been effectively reduced. To evaluate the estimator of the limiting variance, we calculate in simulations the t-statistics of the four regression coefficients and F-statistic for the null $\omega = 0$ based on the original estimators and the bias-corrected estimators. It is seen that the t-test would suffer a mild size distortion based on the original estimator and this issue has been alleviated after bias correction. Overall, the empirical sizes obtained from the 1000 repetitions are close to the nominal size. The next three tables report the performance of the MLE under '3 ahead and 3 behind" spatial weights matrix. The simulation results are similar as the case under "1 ahead and 1 behind" weights matrix. So we do not repeat the detailed analysis.

Table 1: The performance of the MLE before bias correction
with "1 ahead and 1 behind" spatial weights matrix

N	T	δ		ρ		β_1		β_2	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
100	75	-0.0014	0.0032	0.0007	0.0034	0.0003	0.0132	-0.0001	0.0133
125	75	-0.0014	0.0029	0.0006	0.0030	-0.0003	0.0118	-0.0002	0.0118
150	75	-0.0014	0.0027	0.0007	0.0028	0.0001	0.0106	-0.0006	0.0103
100	100	-0.0010	0.0026	0.0004	0.0029	0.0002	0.0111	0.0003	0.0107
125	100	-0.0009	0.0023	0.0004	0.0025	0.0007	0.0102	0.0001	0.0098
150	100	-0.0010	0.0022	0.0006	0.0024	-0.0001	0.0091	-0.0003	0.0092
100	125	-0.0008	0.0024	0.0004	0.0027	-0.0001	0.0099	0.0006	0.0099
125	125	-0.0007	0.0021	0.0004	0.0023	0.0000	0.0086	-0.0002	0.0090
150	125	-0.0008	0.0019	0.0004	0.0021	0.0000	0.0079	0.0002	0.0083

Table 2: The performance of the MLE after bias correction
with "1 ahead and 1 behind" spatial weights matrix

N	T	δ		ρ		β_1		β_2	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
100	75	-0.0002	0.0028	0.0002	0.0033	0.0007	0.0132	0.0006	0.0132
125	75	-0.0001	0.0025	0.0000	0.0029	0.0001	0.0118	0.0004	0.0118
150	75	-0.0001	0.0023	0.0001	0.0027	0.0005	0.0106	-0.0000	0.0102
100	100	-0.0000	0.0024	-0.0000	0.0028	0.0005	0.0111	0.0007	0.0108
125	100	0.0000	0.0021	-0.0001	0.0025	0.0010	0.0102	0.0006	0.0098
150	100	-0.0000	0.0020	0.0001	0.0023	0.0002	0.0091	0.0002	0.0091
100	125	-0.0000	0.0022	0.0000	0.0026	0.0002	0.0099	0.0010	0.0099
125	125	0.0001	0.0020	-0.0000	0.0023	0.0003	0.0086	0.0002	0.0089
150	125	-0.0000	0.0017	0.0001	0.0020	0.0003	0.0079	0.0005	0.0083

Table 3: The empirical sizes of t and F statistics
with "1 ahead and 1 behind" spatial weights matrix under 5% nominal size

N	T	δ	ρ	β_1	β_2	F	δ	ρ	β_1	β_2	F
		before bias correction					after bias correction				
100	75	9.5%	6.9%	7.4%	7.6%	10.7%	6.7%	5.7%	7.3%	7.5%	8.7%
125	75	10.5%	8.5%	8.6%	6.5%	11.2%	7.2%	7.7%	8.8%	6.3%	8.4%
150	75	11.3%	7.5%	7.2%	6.3%	11.6%	6.1%	6.2%	7.2%	6.3%	8.4%
100	100	7.8%	6.2%	6.7%	5.3%	7.2%	5.3%	6.1%	6.5%	5.4%	5.9%
125	100	8.5%	6.3%	6.8%	6.5%	9.9%	6.0%	5.9%	7.0%	6.0%	8.4%
150	100	9.2%	7.8%	7.0%	6.8%	9.6%	5.6%	7.0%	7.2%	6.4%	7.4%
100	125	8.3%	8.0%	6.1%	6.9%	8.2%	6.1%	7.0%	6.1%	7.2%	7.2%
125	125	8.3%	7.0%	6.1%	7.1%	7.3%	7.5%	5.7%	5.9%	6.8%	6.6%
150	125	8.3%	6.1%	5.6%	6.9%	9.2%	5.5%	5.4%	5.9%	6.5%	6.1%

Table 4: The performance of the MLE before bias correction
with "3 ahead and 3 behind" spatial weights matrix

N	T	ρ		δ		β_1		β_2	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
100	75	-0.0015	0.0034	0.0007	0.0039	0.0003	0.0124	0.0001	0.0124
125	75	-0.0016	0.0031	0.0008	0.0035	0.0002	0.0114	-0.0001	0.0113
150	75	-0.0015	0.0030	0.0009	0.0033	0.0001	0.0104	-0.0001	0.0103
100	100	-0.0013	0.0029	0.0007	0.0034	0.0004	0.0108	0.0001	0.0106
125	100	-0.0011	0.0026	0.0006	0.0031	0.0003	0.0098	-0.0008	0.0098
150	100	-0.0010	0.0023	0.0006	0.0026	0.0004	0.0087	-0.0003	0.0087
100	125	-0.0009	0.0026	0.0005	0.0030	0.0003	0.0099	0.0007	0.0097
125	125	-0.0008	0.0023	0.0005	0.0027	0.0003	0.0087	0.0004	0.0086
150	125	-0.0009	0.0021	0.0006	0.0024	-0.0001	0.0078	0.0003	0.0078

Table 5: The performance of the MLE after bias correction
with "3 ahead and 3 behind" spatial weights matrix

N	T	ρ		δ		β_1		β_2	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
100	75	-0.0001	0.0030	-0.0000	0.0038	0.0006	0.0124	0.0006	0.0124
125	75	-0.0002	0.0027	0.0001	0.0034	0.0005	0.0114	0.0004	0.0113
150	75	-0.0001	0.0026	0.0001	0.0031	0.0004	0.0104	0.0004	0.0103
100	100	-0.0002	0.0026	0.0001	0.0033	0.0007	0.0108	0.0005	0.0106
125	100	-0.0001	0.0024	0.0000	0.0030	0.0006	0.0098	-0.0004	0.0097
150	100	0.0001	0.0020	-0.0000	0.0025	0.0007	0.0087	0.0000	0.0087
100	125	-0.0001	0.0024	-0.0000	0.0029	0.0005	0.0099	0.0010	0.0097
125	125	0.0000	0.0022	-0.0000	0.0026	0.0006	0.0087	0.0007	0.0086
150	125	-0.0001	0.0019	0.0001	0.0023	0.0001	0.0078	0.0006	0.0078

Table 6: The empirical sizes of t and F statistics
with "3 ahead and 3 behind" spatial weights matrix under 5% nominal size

N	T	δ	ρ	β_1	β_2	F	δ	ρ	β_1	β_2	F
		before bias correction					after bias correction				
100	75	8.5%	7.3%	6.5%	5.5%	8.8%	5.4%	6.1%	6.5%	5.4%	6.6%
125	75	11.0%	7.1%	7.9%	7.5%	10.8%	6.1%	7.0%	7.7%	7.4%	8.5%
150	75	12.3%	7.6%	7.0%	7.0%	12.6%	7.1%	6.8%	6.8%	6.8%	9.4%
100	100	9.3%	7.8%	6.7%	5.9%	9.4%	6.0%	6.7%	6.8%	6.0%	7.5%
125	100	9.9%	8.3%	7.2%	7.1%	9.4%	6.2%	7.0%	7.1%	6.6%	6.8%
150	100	7.8%	6.0%	6.6%	5.4%	7.8%	5.1%	4.9%	6.4%	5.4%	5.9%
100	125	8.4%	6.5%	6.8%	5.8%	7.9%	4.8%	5.8%	6.8%	6.0%	6.8%
125	125	8.4%	6.4%	6.4%	6.0%	7.4%	6.5%	5.7%	6.5%	5.6%	6.5%
150	125	8.5%	6.5%	6.0%	6.6%	8.6%	5.3%	5.7%	6.0%	6.4%	6.0%

8 Some extensions

The analysis of the paper can be extended to more complex dynamics of the model. Consider the following model

$$Y_t = \alpha + \rho W_N Y_t + \delta Y_{t-1} + \varrho M_N Y_{t-1} + X_t \beta + \Lambda f_t + e_t. \tag{8.1}$$

where M_N is another spatial weights matrix, which is assumed to have similar properties as W_N (Assumption E). If M_N is identical to W_N and the common shocks part Λf_t is absent from (8.1), the model reduces to the one consider by Yu et al. (2008). To accommodate the new dynamics of the model, we make the following assumption to replace Assumption F:

Assumption F$'$. Let G_N^* be defined the same as in Assumption F. We assume

$$\limsup_{N\to\infty} \sum_{l=0}^{\infty} \left\| (\delta^* G_N^* + \varrho^* G_N^* M_N)^l \right\|_\infty \leq C; \qquad \limsup_{N\to\infty} \sum_{l=0}^{\infty} \left\| (\delta^* G_N^* + \varrho^* G_N^* M_N)^l \right\|_1 \leq C.$$

Using the methods stated in Section 4, we can derive the asymptotic representation of the MLE for model (8.1) in a similar way. In fact, the MLE has a similar limiting variance expression as in Theorem 5.1. But the bias expression is different, due to the different dynamics of the model. Let $\phi = (\rho, \delta, \varrho, \beta')'$ and $\hat{\phi}$ be the MLE. Define $\ddot{Y}_{t-1} = M_N \dot{Y}_{t-1}$ and $\ddot{Y}_{-1} = (\ddot{Y}_0, \ddot{Y}_1, \ldots, \ddot{Y}_{T-1})$. We state the result in the following theorem.

Theorem 8.1 *Under Assumptions A-E, F$'$, G-H, when $N, T \to \infty$, $\sqrt{N}/T \to 0$ and $\sqrt{T}/N \to 0$, we have*

$$\sqrt{NT}(\hat{\phi} - \phi^* + b_\phi) \xrightarrow{d} N\Big(0, \Big[\operatorname*{plim}_{N,T\to\infty} \mathbb{D}_\phi\Big]^{-1}\Big),$$

where

$$\mathbb{D}_\phi = \frac{1}{NT}$$

$$\times \begin{bmatrix} \text{tr}(\ddot{Y}'\ddot{M}^*\ddot{Y}M_{F^*})+\Phi & \text{tr}(\ddot{Y}'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \text{tr}(\ddot{Y}'\ddot{M}^*\ddot{Y}_{-1}M_{F^*}) & \cdots & \text{tr}(\ddot{Y}'\ddot{M}^*\dot{X}_kM_{F^*}) \\ \text{tr}(\dot{Y}_{-1}'\ddot{M}^*\ddot{Y}M_{F^*}) & \text{tr}(\dot{Y}_{-1}'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \text{tr}(\dot{Y}_{-1}'\ddot{M}^*\ddot{Y}_{-1}M_{F^*}) & \cdots & \text{tr}(\dot{Y}_{-1}'\ddot{M}^*\dot{X}_kM_{F^*}) \\ \text{tr}(\ddot{Y}_{-1}'\ddot{M}^*\ddot{Y}M_{F^*}) & \text{tr}(\ddot{Y}_{-1}'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \text{tr}(\ddot{Y}_{-1}'\ddot{M}^*\ddot{Y}_{-1}M_{F^*}) & \cdots & \text{tr}(\ddot{Y}_{-1}'\ddot{M}^*\dot{X}_kM_{F^*}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{tr}(\dot{X}_k'\ddot{M}^*\ddot{Y}M_{F^*}) & \text{tr}(\dot{X}_k'\ddot{M}^*\dot{Y}_{-1}M_{F^*}) & \text{tr}(\dot{X}_k'\ddot{M}^*\ddot{Y}_{-1}M_{F^*}) & \cdots & \text{tr}(\dot{X}_k'\ddot{M}^*\dot{X}_kM_{F^*}) \end{bmatrix}$$

with Φ *defined the same as in Theorem 5.2 and*

$$b_\phi = \mathbb{D}_\phi^{-1} \begin{bmatrix} \frac{1}{N}\text{tr}[\Lambda^{*\prime}S_N^{\circ\prime}\Sigma_{ee}^{*-1}\Lambda^*(\Lambda^{*\prime}\Sigma_{ee}^{*-1}\Lambda^*)^{-1}] + \frac{1}{NT}\text{tr}[P_{\widetilde{F}}K^\star] \\ \frac{1}{NT}\text{tr}[P_{\widetilde{F}}L^\star] \\ \frac{1}{NT}\text{tr}[P_{\widetilde{F}}J^\star] \\ 0_{k\times 1} \end{bmatrix}$$

with

$$K^\star = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \text{tr}[W_N\Gamma G_N^*] & 0 & 0 & \cdots & 0 \\ \text{tr}[W_N\Gamma^2 G_N^*] & \text{tr}[W_N\Gamma G_N^*] & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \text{tr}[W_N\Gamma^{T-1}G_N^*] & \text{tr}[W_N\Gamma^{T-2}G_N^*] & \text{tr}[W_N\Gamma^{T-3}G_N^*] & \cdots & 0 \end{bmatrix},$$

and

$$L^\star = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \text{tr}(G_N^*) & 0 & 0 & \cdots & 0 \\ \text{tr}[G_N^*\Gamma] & \text{tr}(G_N^*) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \text{tr}[G_N^*\Gamma^{T-2}] & \text{tr}[G_N^*\Gamma^{T-3}] & \text{tr}[G_N^*\Gamma^{T-4}] & \cdots & 0 \end{bmatrix},$$

and

$$J^\star = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \text{tr}(M_NG_N^*) & 0 & 0 & \cdots & 0 \\ \text{tr}[M_N\Gamma G_N^*] & \text{tr}(M_NG_N^*) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \text{tr}[M_N\Gamma^{T-2}G_N^*] & \text{tr}[M_N\Gamma^{T-3}G_N^*] & \text{tr}[M_N\Gamma^{T-4}G_N^*] & \cdots & 0 \end{bmatrix},$$

where $\Gamma = \delta^* G_N^* + \varrho^* G_N^* M_N^*$.

We use simulations to illustrate the performance of the MLE. The data are generated according to (8.1) with $(\rho, \delta, \varrho) = (0.2, 0.4, 0.3)$. The factors, factor loadings, errors and heteroskedasticity are generated in the same way as in Section 7. Other prespecified parameters such as the number of factors, the number of regressors and the true values of β are also the same. W_N is a "3 ahead and 3 behind" weights matrix and M_N is a "1 ahead and 1 behind" one. For simplicity, the number of factors is assumed to be known. Tables 7 and 8 reports the simulation results based on 1000 repetitions.

Tables 7 and 8 show that the maximum likelihood method continue to perform well. The RMSE decreases as the sample size becomes larger, implying that the MLE is consistent. The bias has been effectively reduced after the bias correction.

Table 7: The performance of the MLE before bias correction

N	T	ρ		δ		ϱ		β_1		β_2	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
100	50	0.0005	0.0068	-0.0025	0.0049	-0.0000	0.0043	0.0008	0.0155	0.0003	0.0157
100	100	0.0002	0.0044	-0.0012	0.0031	0.0001	0.0030	0.0005	0.0107	0.0001	0.0108
100	150	0.0004	0.0036	-0.0008	0.0024	-0.0000	0.0024	0.0004	0.0091	0.0002	0.0089
200	50	0.0009	0.0045	-0.0027	0.0040	0.0002	0.0029	0.0005	0.0110	-0.0007	0.0114
200	100	0.0004	0.0031	-0.0013	0.0024	0.0000	0.0020	-0.0001	0.0076	0.0002	0.0077
200	150	0.0004	0.0025	-0.0009	0.0018	0.0001	0.0017	0.0000	0.0060	-0.0004	0.0063
300	50	0.0010	0.0040	-0.0025	0.0034	-0.0001	0.0024	0.0001	0.0088	0.0001	0.0089
300	100	0.0004	0.0026	-0.0013	0.0021	0.0001	0.0017	-0.0000	0.0059	0.0003	0.0064
300	150	0.0004	0.0021	-0.0009	0.0016	-0.0000	0.0014	-0.0001	0.0049	0.0001	0.0051

Table 8: The performance of the MLE after bias correction

N	T	ρ		δ		ϱ		β_1		β_2	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
100	50	-0.0001	0.0067	-0.0001	0.0042	-0.0001	0.0043	0.0011	0.0155	0.0007	0.0157
100	100	-0.0002	0.0044	0.0000	0.0028	0.0001	0.0030	0.0006	0.0107	0.0003	0.0108
100	150	0.0001	0.0035	-0.0000	0.0023	-0.0001	0.0024	0.0005	0.0091	0.0003	0.0089
200	50	0.0003	0.0044	-0.0003	0.0029	0.0001	0.0029	0.0008	0.0110	-0.0003	0.0113
200	100	0.0000	0.0031	-0.0001	0.0020	-0.0001	0.0020	0.0000	0.0076	0.0003	0.0077
200	150	0.0001	0.0025	-0.0001	0.0016	0.0000	0.0017	0.0001	0.0060	-0.0003	0.0063
300	50	0.0004	0.0039	-0.0001	0.0024	-0.0002	0.0024	0.0004	0.0088	0.0006	0.0089
300	100	0.0001	0.0026	-0.0000	0.0017	0.0000	0.0017	0.0001	0.0059	0.0005	0.0065
300	150	0.0001	0.0020	-0.0000	0.0013	-0.0001	0.0014	-0.0000	0.0049	0.0002	0.0051

The present analysis can be also extended to allow SAR disturbance. Suppose $e_t = \varpi M_N e_t + \varepsilon_t$, where ε satisfies Assumption C. Under this specification, e_t has weak cross sectional correlation. To derive a tractable expression, pre-multiplying $I_N - \varpi M_N$ on both sides of (3.1), we have

$$Y_t = \alpha^\star + \rho W_N Y_t + \varpi M_N Y_t - \rho\varpi M_N W_N Y_t + \delta Y_{t-1} - \varpi\delta M_N Y_{t-1} + X_t\beta - M_N X_t \beta\varpi + \Lambda^\star f_t + \varepsilon_t$$

where $\alpha^\star = (I_N - \varpi M_N)\alpha$ and $\Lambda^\star = (I_N - \varpi M_N)\Lambda$. Now we see that the above model is similar as (3.1) except for high order spatial lags. The analysis of the MLE for the above model is similar as that of (3.1).

9 Conclusion

This paper considers spatial panel data models with common shocks, in which the spatial lag term is endogenous and the explanatory variables are correlated with the unobservable common factors and factor loadings. The proposed maximum likelihood estimator is capable of handling of both types of cross sectional dependence. The results make it possible to determine which type of cross-section dependence or both are present. Heteroskedasticity is explicitly allowed. It is found that when heteroskedasticity is estimated, the limiting variance of MLE is no longer of a sandwich form regardless of normality. We provide a rigorous analysis for the asymptotic theory of the MLE, demonstrating its desirable properties. The Monte Carlo simulations show that the MLE has good finite sample properties.

References

Ahn, S. G., Lee, Y. H. and Schmidt, P. (2013): Panel data models with multiple time-varying effects, *Journal of Econometrics*, 174, 1-14.

Alvarez, J., and Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica*, 71(4), 1121-1159.

Anderson, T.W. (2003). *Introduction to Multivariate Statistical Analysis.* Hoboken, New Jersey: John Wiley & Sons.

Anderson, T. W., and Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76(375), 598-606.

Anselin, L. (1988). Spatial econometrics: methods and models. The Netherlands: Kluwer Academic Publishers.

Arellano, M., and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277-297.

Bai, J. (2003) Inferential theory for factor models of large dimensions. *Econometrica*, 71(1), 135-171.

Bai, J. (2009) Panel data models with interactive fixed effects. *Econometrica*, 77(4), 1229–1279.

Bai, J. and Li, K. (2012) Statistical analysis of factor models of high dimension, *The Annals of Statistics*, 40(1), 436–465.

Bai, J. and Li, K. (2014a) Spatial panel data models with common shocks, *Manuscript*, Columbia University.

Bai, J. and Li, K. (2014b) Theory and methods of panel data models with interactive effects, *The Annals of Statistics*, 42(1), 142–170.

Baltagi, B., Song, S.H., and Koh, W. (2003) Testing panel data regression models with spatial error correlation, *Journal of Econometrics*, 117, 123-150.

Baltagi, B., Song, S.H., Jung, B.C., and Koh, W. (2007) Testing for serial correlation, spatial autocorrelation and random effects using panel data, *Journal of Econometrics*, 140, 5-51.

Baltagi, B., Egger, P., Pfaffermayr, M. (2008). Estimating regional trade agreement effects on FDI in an interdependent world. *Journal of Econometrics*, 145(1), 194-208.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of econometrics*, 150(1), 41-55.

Brockwell, P. J., and Davis, R. A. (1991). *Time series: Theory and methods*. Springer Series in Statistics, Berlin, New York: Springer, 2nd edition.

Calvó-Armengol, A., Patacchini, E., And Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies*, 76(4), 1239-1267.

Case, A. C. (1991) Spatial patterns in household demand. *Econometrica*, 59(4), 953–965.

Case, A. C., Rosen, H. S., and Hines, J. R. (1993) Budget spillovers and fiscal policy interdependence: Evidence from the states. *Journal of Public Economics*, 52(3), 285-307.

Chamberlain, G. and Rothschild, M. (1983): Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica*, **51:5**, 1281–1304.

Cliff, A. D., and Ord, J. K. (1973) *Spatial autocorrelation*, London: Pion Ltd.

Conley, T. G. (1999). GMM Estimation with Cross Sectional Dependence, *Journal of Econometrics*, 92, 1-45.

Conley, T.G., and Dupor, B. (2003) "A Spatial Analysis of Sectoral Complementarity," *Journal of Political Economy*, 111, 311-352.

Connor, G., and Korajczyk, R. A. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of financial economics*, 15(3), 373-394.

Connor, G., and Korajczyk, R. A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics*, 21(2), 255-289.

Cressie, N. (1993): Statistics for Spatial Data. New York: John Wiley & Sons.

Defever, F., Heid, B., and Larch, M. (2015). Spatial exporters. *Journal of International Economics*, 95(1), 145šC156.

Eberhardt, M., Helmers, C., and Strauss, H. (2013). Do spillovers matter when estimating private returns to R&D?. *Review of Economics and Statistics*, 95(2), 436-448.

Elliott, M., Golub, B., and Jackson, M. O. (2014). Financial Networks and Contagion.*American Economic Review*, 104(10), 3115-53.

Fan, J., Liao, Y., and Mincheva, M. (2011). High dimensional covariance matrix estimation in approximate factor models. *Annals of statistics*, 39(6), 3320.

Fuller, W. A. (1996). *Introduction to Statistical Time Series* (Vol. 230). John Wiley & Sons.

Geweke, J., and Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies*, 9(2), 557-587.

Giraitis, L., and Taqqu, M. S. (1998). Central limit theorems for quadratic forms with time-domain conditions. *Annals of Probability*, 377-398.

Glaeser, E. L., Sacerdote, B., and Scheinkman, J. A. (1996) Crime and social interactions. *The Quarterly Journal of Economics*, 111(2), 507-548.

Hahn, J., and Kuersteiner, G. (2002). Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large. *Econometrica*, 70(4), 1639-1657.

Hall P, and Heyde C C. *Martingale limit theory and applications*. Academic, New York, 1980.

Jennrich, R. I. (1969) Asymptotic properties of non-linear least squares estimators. *The Annals of Mathematical Statistics*, 40(2), 633-643.

Kapoor, M., Kelejian, H. H., and Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1), 97-130.

Kelejian, H. H., and Prucha, I. R. (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1), 99-121.

Kelejian, H. H., and Prucha, I. R. (1999) A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*, 40(2), 509-533.

Kelejian, H. H., and Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1), 53-67.

Kelejian, H. H., and Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics*, 104(2), 219-257.

Lawless, M. (2009). Firm export dynamics and the geography of trade. *Journal of International Economics*, 77(2), 245-254.

Lee, L. (2004a) Asymptotic distribution of quasi-maximum likelihood estiators for spatial augoregressive models, *Econometrica*, 72(6), 1899–1925.

Lee, L. (2004b) A supplement to "Asymptotic distribution of quasi-maximum likelihood estiators for spatial augoregressive models", *Manuscipt*.

Lee, L. F. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140(2), 333-374.

Lee, L. and Yu, J. (2010) Estimation of spatial autoregressive panel data models with fixed effects, *Journal of Econometrics*, 154(2), 165-185.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4), 825-860.

Lin, X., and Lee, L. F. (2010) GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *Journal of Econometrics*, 157(1), 34-52.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3), 531-542.

Moon, H. and Weidner, M. (2009) Likelihood expansion for panel regression models with factors, *manuscipt, University of Southern California*.

Newey, W. and McFadden, D. (1994): Large Sample Estimation and Hypothesis Testing, in Engle, R.F. and D. McFadden (eds.) *Handbook of Econometrics,* North Holland.

Neyman, J., and Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrica*. 1-32.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417-1426.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70(349), 120-126.

Pesaran, M. H. (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure, *Econometrica*, 74(4), 967–1012.

Pesaran, M. H., and Tosetti, E. (2011). Large panels with spatial correlations and common factors, *Journal of Econometrics*, 161, 182–202.

Phillips, P. C., and Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, 67(5), 1057-1111.

Rauch, J. E., and Trindade, V. (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics*, 84(1), 116-130.

Ross, S. A. (1976): The arbitrage theory of capital asset pricing, *Journal of Economic Theory*,**13(3)**, 341–360.

Stock, J. H., and Watson, M. W. (1998). Diffusion indexes (No. w6702). *National Bureau of Economic Research.*

Topa, G. (2001). Social Interactions, Local Spillovers and Unemployment, *Review of Economic Studies*, 68, 261-295.

Yu, J., de Jong, R., and Lee, L. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 146(1), 118-134.

Appendix A: The proof of consistency

This section provide a detailed proof on the consistency. Throughout the proof, we say $M_1 \geq M_2$ for two matrices M_1 and M_2, if $M_1 - M_2$ is semi-positively definite. In addition, Let X_{NT} be a generic random variable which depends on N and T, we say $X_{NT} = O_p(a_{NT})$ or $X_{NT} = o_p(a_{NT})$, where a_{NT} may be $T^{-1}, N^{-1}, N^{-1/2}T^{-1/2}$ or other magnitudes appearing in the paper, if and only if for every $\epsilon > 0$, there exists a constant M_ϵ such that

$$P\Big(|a_{NT}^{-1}X_{NT}| \geq M_\epsilon\Big) \leq \epsilon$$

for all N and T; or

$$\lim_{N,T\to\infty} P\Big(|a_{NT}^{-1}X_{NT}| \geq \epsilon\Big) = 0.$$

where $N, T \to \infty$ denotes that N and T pass to infinity simultaneously (joint limit), see the detailed discussion on joint limit in Phillips and Moon (1999). More specifically, let $\{N_m\}$ denote any increasing sequence of N and $\{T_m\}$ denote any increasing sequence of T. Let $\{b_m\}$ be a sequence whose mth element is $b_m = (N_m, T_m)$. Then the preceding limit is equivalent to

$$\lim_{m\to\infty} P\Big(|a_{N_mT_m}^{-1}X_{N_mT_m}| \geq \epsilon\Big) = 0$$

for all sequences $\{b_m\}$.

The uniform version of $O_p(a_{NT})$ and $o_p(a_{NT})$ are defined similarly. We say $X_{NT}(\theta_N)$ is $O_p(a_{NT})$ or $o_p(a_{NT})$ uniformly on Θ_N, if and only if for every $\epsilon > 0$, there exists a constant M_ϵ such that

$$P\Big(\sup_{\theta_N\in\Theta_N} |a_{NT}^{-1}X_{NT}(\theta_N)| \geq M_\epsilon\Big) \leq \epsilon$$

for all N and T; or

$$\lim_{N,T\to\infty} P\Big(\sup_{\theta_N\in\Theta_N} |a_{NT}^{-1}X_{NT}(\theta_N)| \geq \epsilon\Big) = 0.$$

For notational simplicity, in the presentation of the following uniform results, we drop the superscript "N" from the symbols θ_N and Θ_N for notational simplicity.

In addition, we define the following notations for ease of exposition in Appendix A:

$$\begin{aligned} &\dot{X}_{t0} = \dot{Y}_{t-1}, && \beta_0 = \delta, && \beta_0^* = \delta^*; \\ &\dot{X}_{tk+1} = S_N^*(\dot{Y}_{t-1}\delta^* + \dot{X}_t\beta^*), && \beta_{k+1} = \rho, && \beta_{k+1}^* = \rho^*, \end{aligned} \tag{A.1}$$

The following lemmas are useful for the proof of consistency.

Lemma A.1 *Under Assumptions A-H, we have*

$$(a)\quad \sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{p=0}^{k+1}(\beta_p - \beta_p^*)\sum_{t=1}^{T}\dot{X}_{tp}'\ddot{M}[I_N - (\rho-\rho^*)S_N^*]\dot{e}_t\Big| = o_p(1),$$

$$(b)\quad \sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}[I_N - (\rho-\rho^*)S_N^*]'\ddot{M}[I_N - (\rho-\rho^*)S_N^*]\dot{e}_t\Big| = o_p(1),$$

$$(c)\quad \sup_{\theta\in\Theta}\Big|\frac{1}{N^2T}\sum_{t=1}^{T}\dot{e}_t'[I_N - (\rho-\rho^*)S_N^*]'\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}[I_N - (\rho-\rho^*)S_N^*]\dot{e}_t\Big| = o_p(1),$$

$$(d)\quad \sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{t=1}^{T}\mathrm{tr}\Big[[I_N-(\rho-\rho^*)S_N^*]'\Sigma_{ee}^{-1}[I_N-(\rho-\rho^*)S_N^*](\dot{e}_t\dot{e}_t'-\Sigma_{ee}^*)\Big]\Big|=o_p(1).$$

where $\beta_0,\beta_{k+1},\dot{X}_{t0}$ and $\dot{X}_{tk+1}$ are defined in (A.1) and the parameters space Θ is defined as

$$\Theta=\Big\{\theta=(\omega,\Sigma_{ee},\Lambda)\ \Big|\ \|\omega\|\le C;\ C^{-1}\le\sigma_i^2\le C,\forall i;\ \frac{1}{N}\Lambda'\Sigma_{ee}^{-1}\Lambda=I_r\Big\}.$$

PROOF OF LEMMA A.1. Consider (a). The left hand side is bounded by

$$\sum_{p=0}^{k+1}\Big[\sup_{\theta\in\Theta}|\beta_p-\beta_p^*|\cdot\Big|\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_{tp}'\ddot{M}[I_N-(\rho-\rho^*)S_N^*]\dot{e}_t\Big|\Big]$$

Since β is in a compact set by Assumption H, it suffices to show

$$\sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_{tp}'\ddot{M}[I_N-(\rho-\rho^*)S_N^*]\dot{e}_t\Big|=o_p(1)\tag{A.2}$$

for all $p=0,1,\dots,k+1$. The left hand side of (A.2) is bounded by

$$\sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{t=1}^{T}X_{tp}'\ddot{M}[I_N-(\rho-\rho^*)S_N^*]e_t\Big|+\sup_{\theta\in\Theta}\Big|\frac{1}{N}\bar{X}_p'\ddot{M}[I_N-(\rho-\rho^*)S_N^*]\bar{e}\Big|=I_1+I_2,\ \text{say,}$$

where $\bar{X}_p=(\bar{x}_{1p},\bar{x}_{2p},\dots,\bar{x}_{Np})'$ with $\bar{x}_{ip}=T^{-1}\sum_{t=1}^{T}x_{itp}$ and $\bar{e}=(\bar{e}_1,\bar{e}_2,\dots,\bar{e}_N)$ with $\bar{e}_i=T^{-1}\sum_{t=1}^{T}e_{it}$. Consider term I_1. By $\ddot{M}=\Sigma_{ee}^{-1}-N^{-1}\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}$, we can rewrite it as

$$\begin{aligned}&\sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{t=1}^{T}X_{tp}'\Sigma_{ee}^{-1}e_t-(\rho-\rho^*)\frac{1}{NT}\sum_{t=1}^{T}X_{tp}'\Sigma_{ee}^{-1}S_N^*e_t\\&-\frac{1}{N^2T}\sum_{t=1}^{T}X_{tp}'\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}e_t+(\rho-\rho^*)\frac{1}{N^2T}\sum_{t=1}^{T}X_{tp}'\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}S_N^*e_t\Big| \qquad (A.3)\\&\le\sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^2}x_{itp}e_{it}\Big|+\sup_{\theta\in\Theta}\Big|\mathrm{tr}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\sigma_i^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{t=1}^{T}x_{itp}e_{jt}\Big]\Big|\\&+\sup_{\theta\in\Theta}|\rho-\rho^*|\cdot\Big|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^2}x_{itp}\ddot{e}_{it}\Big|+\sup_{\theta\in\Theta}|\rho-\rho^*|\cdot\Big|\mathrm{tr}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\sigma_i^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{t=1}^{T}x_{itp}\ddot{e}_{jt}\Big]\Big|,\end{aligned}$$

where $\ddot{e}_{it}=\sum_{o=1}^{N}S_{io,N}e_{ot}$. We use $I_3,I_4,\dots,I_6$ to denote the four expressions on right hand side. Term I_3, by the Cauchy-Schwarz inequality, is bounded by

$$\Big\{\sup_{\theta\in\Theta}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\Big]^{1/2}\Big\}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}x_{itp}e_{it}\Big|^2\Big]^{1/2}\le C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}x_{itp}e_{it}\Big|^2\Big]^{1/2}=O_p(T^{-1/2}).$$

where the first inequality is due to Assumption H and the second result is due to $E(N^{-1}\sum_{i=1}^{N}|T^{-1}\sum_{t=1}^{T}x_{itp}e_{it}|^2)=O(T^{-1})$ for $p=0,1,\dots,k+1$ under Assumptions A and C. Thus $I_3=O_p(T^{-1/2})$. Consider I_4. Ignore $\sup_{\theta\in\Theta}$, the expression of I_4 in the trace operator is bounded in norm by

$$\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\|\lambda_i\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}x_{itp}e_{jt}\Big|^2\Big]^{1/2}.$$

However,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\|\lambda_i\|^2=\operatorname{tr}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i\lambda_i'\Big]=\operatorname{tr}\Big[\frac{1}{N}\Lambda'\Sigma_{ee}^{-1}\Lambda\Big]=\operatorname{tr}[I_r]=r. \tag{A.4}$$

Given this result, by the boundedness of σ_i^2 by Assumption H,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\|\lambda_i\|^2\le C\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\|\lambda_i\|^2=Cr.$$

So $I_4=O_p(T^{-1/2})$ by $E(\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|T^{-1}\sum_{t=1}^{T}x_{itp}e_{jt}|^2)=O(T^{-1})$ under Assumptions A and C. The remaining two terms of the right hand side of (A.3) can be proved to be $O_p(T^{-1/2})$ similarly as the first two terms by noticing $|\rho-\rho^*|$ is bounded by Assumption H and, for $p=0,1,\dots,k+1$,

$$E\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}x_{itp}\ddot{e}_{it}\Big|^2\Big]=O(T^{-1}),\quad E\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}x_{itp}\ddot{e}_{jt}\Big|^2\Big]=O(T^{-1}), \tag{A.5}$$

where the two results in (A.5) are shown in Lemma A.2(e) and (f) of Bai and Li (2014a). Given the above results, we have $I_1=O_p(T^{-1/2})$. Term I_2 can be proved to be $O_p(T^{-1/2})$ similarly as I_1. So we have (a).

Consider (b). By the normalization condition $\sum_{t=1}^{T}f_t^*=0$, $\dot{e}_t$ in the expression can be replaced with e_t. So the expression on the left hand side is

$$\begin{aligned}&\sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}\ddot{M}e_t-(\rho-\rho^*)\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}S_N^{*\prime}\ddot{M}e_t\\&-(\rho-\rho^*)\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}\ddot{M}S_N^{*}e_t+(\rho-\rho^*)^2\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}S_N^{*\prime}\ddot{M}S_N^{*}e_t\Big|\\&\le\sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}\ddot{M}e_t\Big|+\sup_{\theta\in\Theta}|\rho-\rho^*|\cdot\Big|\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}S_N^{*\prime}\ddot{M}e_t\Big|\\&+\sup_{\theta\in\Theta}|\rho-\rho^*|\cdot\Big|\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}\ddot{M}S_N^{*}e_t\Big|+\sup_{\theta\in\Theta}|\rho-\rho^*|^2\cdot\Big|\frac{1}{NT}\sum_{t=1}^{T}f_t^{*\prime}\Lambda^{*\prime}S_N^{*\prime}\ddot{M}S_N^{*}e_t\Big|\end{aligned}$$

We use $I_7,I_8,\dots,I_{10}$ to denote the four expressions on right hand side. By the definition of $\ddot{M}$, term I_7 is bounded by

$$\sup_{\theta\in\Theta}\Big|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^2}f_t^{*\prime}\lambda_i^*e_{it}\Big|+\sup_{\theta\in\Theta}\Big|\operatorname{tr}\Big[\Big(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i^*\lambda_i'\Big)\Big(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^2}\lambda_ie_{it}f_t^{*\prime}\Big)\Big]\Big|.$$

By the Cauchy-Schwarz inequality, the first term is bounded by

$$\sup_{\theta\in\Theta}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\|\lambda_i^*\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_t^*e_{it}\Big\|^2\Big]^{1/2}=O_p(T^{-1/2})$$

by Assumptions B and H. The expression of the second term in the trace operator is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i^*\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\|\lambda_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_t^*e_{it}\Big\|^2\Big]^{1/2}=O_p(T^{-1/2}).$$

So we have $I_7 = O_p(T^{-1/2})$. Treating $S_N^*\Lambda^*$ as a new Λ^* and $S_N^* e_t$ as a new e_t and noticing that $|\rho - \rho^*|$ is bounded by Assumption H as well as $E(N^{-1}\sum_{i=1}^N \|T^{-1}\sum_{t=1}^T f_t^* \ddot{e}_{it}\|^2) = O(T^{-1})$, terms I_8, I_9 and I_{10} can be proved to be $O_p(T^{-1/2})$ similarly as I_7. Then we have (b).

Consider (c). The left hand side of (c) is bounded by

$$\sup_{\theta\in\Theta}\Big|\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^T \Lambda'\Sigma_{ee}^{-1}\dot{e}_t\dot{e}_t'\Sigma_{ee}^{-1}\Lambda\Big]\Big| + 2\sup_{\theta\in\Theta}|\rho-\rho^*|\cdot\Big|\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^T \Lambda'\Sigma_{ee}^{-1}\dot{e}_t\dot{e}_t' S_N^{*\prime}\Sigma_{ee}^{-1}\Lambda\Big]\Big|$$

$$+\sup_{\theta\in\Theta}|\rho-\rho^*|^2\cdot\Big|\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^T \Lambda'\Sigma_{ee}^{-1}S_N^*\dot{e}_t\dot{e}_t' S_N^{*\prime}\Sigma_{ee}^{-1}\Lambda\Big]\Big|$$

We use I_{11}, I_{12} and I_{13} to denote the above three expression. First consider I_{13}. Since $|\rho - \rho^*|$ is bounded by Assumption H, it suffices to prove

$$\sup_{\theta\in\Theta}\Big|\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^T \Lambda'\Sigma_{ee}^{-1}S_N^*\dot{e}_t\dot{e}_t' S_N^{*\prime}\Sigma_{ee}^{-1}\Lambda\Big]\Big| = o_p(1).$$

The left hand side of the above equation is bounded by

$$\sup_{\theta\in\Theta}\Big|\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^T \Lambda'\Sigma_{ee}^{-1}(\ddot{e}_t\ddot{e}_t' - S_N^*\Sigma_{ee}^* S_N^{*\prime})\Sigma_{ee}^{-1}\Lambda\Big]\Big| \tag{A.6}$$

$$+\sup_{\theta\in\Theta}\Big|\mathrm{tr}\Big[\frac{1}{N^2}\Lambda'\Sigma_{ee}^{-1}S_N^*\Sigma_{ee}^* S_N^{*\prime}\Sigma_{ee}^{-1}\Lambda\Big]\Big| + \sup_{\theta\in\Theta}\Big|\mathrm{tr}\Big[\frac{1}{N^2}\Lambda'\Sigma_{ee}^{-1}S_N^*\bar{e}\bar{e}' S_N^{*\prime}\Sigma_{ee}^{-1}\Lambda\Big]\Big|.$$

where $\ddot{e}_t = S_N^*\dot{e}_t$. The expression of the first term in the trace operator can be written as

$$\frac{1}{N^2T}\sum_{i=1}^N\sum_{j=1}^N \frac{1}{\sigma_i^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{t=1}^T[\ddot{e}_{it}\ddot{e}_{jt} - E(\ddot{e}_{it}\ddot{e}_{jt})]$$

which is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^N\frac{1}{\sigma_i^2}\|\lambda_i\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Big|\frac{1}{T}\sum_{t=1}^T[\ddot{e}_{it}\ddot{e}_{jt} - E(\ddot{e}_{it}\ddot{e}_{jt})]\Big|^2\Big]^{1/2}.$$

The first factor is r by (A.4) and the second factor is $O_p(T^{-1/2})$ by Lemma A.2(f) of Bai and Li (2014a). So we have that the first term of (A.6) is $O_p(T^{-1/2})$. Consider the second term. By the boundedness of σ_i^2 by Assumption H, there exists a constant C such that $\Sigma_{ee}^{-1}S_N^{*\prime}\Sigma_{ee}^* S_N^* \le C\cdot I_N$. Given this result, the second term is bounded by $\frac{1}{N}Cr$, which is $O(N^{-1})$. Consider the third term. By $\ddot{M} = \Sigma_{ee}^{-1} - N^{-1}\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}$ is semi-positive definite, we have

$$0 \le \mathrm{tr}\Big[\frac{1}{N^2}\Lambda'\Sigma_{ee}^{-1}S_N^*\bar{e}\bar{e}' S_N^{*\prime}\Sigma_{ee}^{-1}\Lambda\Big] \le \frac{1}{N^2}\bar{e}' S_N^{*\prime}\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}S_N^*\bar{e}$$
$$\le \frac{1}{N}\bar{e}' S_N^{*\prime}\Sigma_{ee}^{-1}S_N^*\bar{e} \le C\frac{1}{N}\bar{e}' S_N^{*\prime}S_N^*\bar{e}$$

Notice the last expression is independent with parameters and is $O_p(T^{-1})$. Given the above results, we have $I_{13} = O_p(T^{-1/2}) + O_p(N^{-1})$. Terms I_{11} and I_{12} can be shown to be $O_p(T^{-1/2}) + O_p(N^{-1})$ similarly as I_{13}. So we have (c).

Consider (d). The left hand side of (d) can be written as

$$\sup_{\theta\in\Theta}\Big|\text{tr}\Big[\Sigma_{ee}^{-1}\frac{1}{NT}\sum_{t=1}^{T}(\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)\Big] - 2(\rho-\rho^*)\text{tr}\Big[\Sigma_{ee}^{-1}\frac{1}{NT}\sum_{t=1}^{T}S_N^*(\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)\Big]$$
$$+(\rho-\rho^*)^2\text{tr}\Big[\Sigma_{ee}^{-1}\frac{1}{NT}\sum_{t=1}^{T}S_N^*(\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)S_N^{*\prime}\Big]\Big|,$$

which is bounded by

$$\sup_{\theta\in\Theta}\Big|\text{tr}\Big[\Sigma_{ee}^{-1}\frac{1}{NT}\sum_{t=1}^{T}(\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)\Big]\Big| + 2\sup_{\theta\in\Theta}|\rho-\rho^*|\cdot\Big|\text{tr}\Big[\Sigma_{ee}^{-1}\frac{1}{NT}\sum_{t=1}^{T}S_N^*(\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)\Big]\Big|$$
$$+\sup_{\theta\in\Theta}|\rho-\rho^*|^2\cdot\Big|\text{tr}\Big[\Sigma_{ee}^{-1}\frac{1}{NT}\sum_{t=1}^{T}S_N^*(\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)S_N^{*\prime}\Big]\Big| = I_{14} + 2I_{15} + I_{16}, \quad \text{say.}$$

Consider I_{16}. Since $|\rho - \rho^*|$ is bounded by Assumption H, it suffices to prove

$$\sup_{\theta\in\Theta}\Big|\text{tr}\Big[\Sigma_{ee}^{-1}\frac{1}{NT}\sum_{t=1}^{T}S_N^*(\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)S_N^{*\prime}\Big]\Big| = o_p(1).$$

The left hand side of the above equation is further bounded by

$$\sup_{\theta\in\Theta}\Big|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\frac{1}{T}\sum_{t=1}^{T}[\ddot{e}_{it}^2 - E(\ddot{e}_{it}^2)]\Big| + \sup_{\theta\in\Theta}\Big|\text{tr}\Big[\frac{1}{N}\sum_{i=1}^{N}\Sigma_{ee}^{-1}S_N^*\bar{e}\bar{e}'S_N^{*\prime}\Big]\Big|. \tag{A.7}$$

By the Cauchy-Schwarz inequality, the first term of (A.7) is bounded by

$$\Big\{\sup_{\theta\in\Theta}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\Big]^{1/2}\Big\}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}[\ddot{e}_{it}^2 - E(\ddot{e}_{it}^2)]\Big|^2\Big]^{1/2},$$

which is $O_p(T^{-1/2})$ by Lemma A.2(c) of Bai and Li (2014a). By the boundedness of σ_i^2, there exists a constant C such that $\Sigma_{ee}^{-1} \le CI_N$. The second term of (A.7) is therefore bounded by

$$C\frac{1}{N}\bar{e}'S_N^{*\prime}S_N^*\bar{e} = O_p(T^{-1}).$$

So we have $I_{16} = O_p(T^{-1/2})$. Terms I_{14} and I_{15} can be proved to be $O_p(T^{-1/2})$ similarly as I_{16}. So we have (d). This completes the proof of Lemma A.1. □

Lemma A.2 *Under Assumptions A-H, we have*

(*a*) $\frac{1}{NT}\text{tr}[\Xi_1'\widehat{\widehat{M}}\Xi_1] = O_p(\|\hat{\omega} - \omega^*\|^2)$,

(*b*) $\frac{1}{NT}\text{tr}[\Xi_1'\widehat{\widehat{M}}\Xi_2] = O_p(\|\hat{\omega} - \omega^*\|^2)$,

(*c*) $\frac{1}{NT}\text{tr}[\Xi_2'\widehat{\widehat{M}}\Xi_2] = O_p(\|\hat{\omega} - \omega^*\|^2)$, (*d*) $\frac{1}{NT}\text{tr}[(F^{*\prime}F^*)^{1/2}\Lambda^{*\prime}\widehat{\ddot{M}}\Xi_1] = O_p(\|\hat{\omega}-\omega^*\|)$,

(*e*) $\frac{1}{NT}\text{tr}[(F^{*\prime}F^*)^{1/2}\Lambda^{*\prime}\widehat{\ddot{M}}\Xi_2] = O_p(\|\hat{\omega}-\omega^*\|)$.

where Ξ_1 and Ξ_2 are defined in (A.21) below.

Proof of Lemma A.2. Consider (a). Notice that $\widehat{\ddot{M}} \le \hat{\Sigma}_{ee}^{-1} \le C\cdot I_N$ for some constant C, the left hand side of (a) is bounded by

$$C(\hat{\rho}-\rho^*)^2\text{tr}\Big[\frac{1}{NT}(F^{*\prime}F^*)^{1/2}\Lambda^{*\prime}S_N^{*\prime}S_N^*\Lambda^*(F^{*\prime}F^*)^{1/2}\Big] = O_p(\|\hat{\omega}-\omega^*\|^2)$$

by Assumption B.

Consider (b). The left hand side is equal to

$$\sum_{p=0}^{k+1}(\hat{\rho}-\rho^*)(\hat{\beta}_p-\beta_p^*)\text{tr}\Big[\frac{1}{NT}\Lambda^{*\prime}S_N^{*\prime}\widehat{\ddot{M}}\dot{X}_pF^*\Big]$$

which can be further written as

$$\begin{aligned}&\sum_{p=0}^{k+1}(\hat{\rho}-\rho^*)(\hat{\beta}_p-\beta_p^*)\text{tr}\Big[\frac{1}{NT}\Lambda^{*\prime}S_N^{*\prime}\hat{\Sigma}_{ee}^{-1}\dot{X}_pF^*\Big] \\ &+\sum_{p=0}^{k+1}(\hat{\rho}-\rho^*)(\hat{\beta}_p-\beta_p^*)\text{tr}\Big[\frac{1}{N^2T}\Lambda^{*\prime}S_N^{*\prime}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_pF^*\Big]\end{aligned} \tag{A.8}$$

Notice that

$$\Big\|\frac{1}{NT}\Lambda^{*\prime}S_N^{*\prime}\hat{\Sigma}_{ee}^{-1}\dot{X}_pF^*\Big\| \le \Big[\frac{1}{N}\sum_{i=1}^N\frac{1}{\hat{\sigma}_i^4}\Big|\sum_{j=1}^N S_{ij,N}\lambda_j^*\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^N\Big\|\frac{1}{T}\sum_{t=1}^T\dot{x}_{itp}f_t^*\Big\|^2\Big]^{1/2} = O_p(1).$$

So the first term of (A.8) is $O_p(\|\hat{\omega}-\omega^*\|^2)$. Also notice that

$$\begin{aligned}&\Big\|\frac{1}{N^2T}\Lambda^{*\prime}S_N^{*\prime}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_pF^*\Big\| \\ &\le C\Big[\frac{1}{N}\sum_{i=1}^N\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^N\Big\|\sum_{j=1}^N S_{ij,N}^*\lambda_j^*\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^N\Big\|\frac{1}{T}\sum_{t=1}^T\dot{x}_{itp}f_t^*\Big\|^2\Big]^{1/2} = O_p(1).\end{aligned}$$

So the second term of (A.8) is $O_p(\|\hat{\omega}-\omega^*\|^2)$. Given the above results, we have (b).

The proof of result (c) is similar as that of result (a). The proofs of results (d) and (e) are similar as that of result (b). The details are therefore omitted. This completes the proof of Lemma A.2. □

Proof of Proposition 5.1: Consider the following function

$$\begin{aligned}\mathcal{L}(\theta) = &-\frac{1}{2NT}\sum_{t=1}^T Z_t(\delta,\rho,\beta,\Lambda,F)'\Sigma_{ee}^{-1}Z_t(\delta,\rho,\beta,\Lambda,F) - \frac{1}{2N}\ln|\Sigma_{ee}| \\ &+\frac{1}{N}\ln|I_N-\rho W_N| + \frac{1}{2N}\ln|\Sigma_{ee}^*| - \frac{1}{N}\ln|I_N-\rho^*W| + \frac{1}{2}.\end{aligned} \tag{A.9}$$

where

$$Z_t(\delta, \rho, \beta, \Lambda, F) = \dot{Y}_t - \delta \dot{Y}_{t-1} - \rho W \dot{Y}_t - \dot{X}_t \beta - \Lambda f_t.$$

The above function is a centered objective function and will be used in the subsequent analysis. Given Λ, δ, ρ and β, it is seen that the factors F maximize (A.9) at

$$f_t(\delta, \rho, \beta, \Lambda) = (\Lambda' \Sigma_{ee}^{-1} \Lambda)^{-1} \Lambda' \Sigma_{ee}^{-1} (\dot{Y}_t - \delta \dot{Y}_{t-1} - \rho W_N \dot{Y}_t - \dot{X}_t \beta). \tag{A.10}$$

Substituting (A.10) into (A.9) to concentrate out F, the objective function now is

$$\begin{aligned} \mathcal{L}(\theta) = &-\frac{1}{2NT} \sum_{t=1}^{T} (\dot{Y}_t - \delta \dot{Y}_{t-1} - \rho W \dot{Y}_t - \dot{X}_t \beta - \Lambda f_t)' \ddot{M} (\dot{Y}_t - \delta Y_{t-1} - \rho W \dot{Y}_t - \dot{X}_t \beta - \Lambda f_t) \\ &- \frac{1}{2N} \ln |\Sigma_{ee}| + \frac{1}{N} \ln |I_N - \rho W| + \frac{1}{2N} \ln |\Sigma_{ee}^*| - \frac{1}{N} \ln |I_N - \rho^* W| + \frac{1}{2}, \end{aligned} \tag{A.11}$$

where $\ddot{M} = \Sigma_{ee}^{-1} - \Sigma_{ee}^{-1} \Lambda (\Lambda' \Sigma_{ee}^{-1} \Lambda)^{-1} \Lambda' \Sigma_{ee}^{-1} = \Sigma_{ee}^{-1} - \frac{1}{N} \Sigma_{ee}^{-1} \Lambda \Lambda' \Sigma_{ee}^{-1}$, where the second equation is due to the normalization condition $\frac{1}{N} \Lambda' \Sigma_{ee}^{-1} \Lambda = I_r$. By $\dot{Y}_t = \delta^* \dot{Y}_{t-1} + \rho^* W_N \dot{Y}_t + \dot{X}_t \beta^* + \Lambda^* f_t^* + \dot{e}_t$, we have

$$\dot{Y}_t = (I_N - \rho^* W_N)^{-1} (\delta^* \dot{Y}_{t-1} + \dot{X}_t \beta^* + \Lambda^* f_t^* + \dot{e}_t),$$

which implies

$$\begin{aligned} \dot{Y}_t - \delta \dot{Y}_{t-1} - \rho W_N \dot{Y}_t - \dot{X}_t \beta &= (I_N - \rho W_N) \dot{Y}_t - \delta \dot{Y}_{t-1} - \dot{X}_t \beta \\ = (I_N - \rho W_N)(I_N - \rho^* W_N)^{-1} &(\delta^* \dot{Y}_{t-1} + \dot{X}_t \beta^* + \Lambda^* f_t^* + \dot{e}_t) - \delta \dot{Y}_{t-1} - \dot{X}_t \beta \end{aligned}$$

Notice that $(I_N - \rho W_N)(I_N - \rho^* W_N)^{-1} = I_N - (\rho - \rho^*) S_N^*$ with $S_N^* = W_N (I_N - \rho^* W_N)^{-1}$. Then we can rewrite the preceding equation in terms of the notations in (A.1) as

$$\begin{aligned} &\dot{Y}_t - \delta \dot{Y}_{t-1} - \rho W_N \dot{Y}_t - \dot{X}_t \beta \\ &\quad = -\sum_{p=0}^{k+1} \dot{X}_{tp} (\beta_p - \beta_p^*) + [I_N - (\rho - \rho^*) S_N^*] \Lambda^* f_t^* + [I_N - (\rho - \rho^*) S_N^*] \dot{e}_t. \end{aligned} \tag{A.12}$$

Using (A.12), the objective function (A.11) can be further written as

$$\mathcal{L}(\theta) = \mathcal{L}_1(\theta) + \mathcal{L}_2(\theta),$$

where

$$\begin{aligned} \mathcal{L}_1(\theta) = &-\frac{1}{2} \sum_{p=0}^{k+1} \sum_{q=0}^{k+1} (\beta_p - \beta_p^*)(\beta_q - \beta_q^*) \frac{1}{NT} \operatorname{tr}(\dot{X}_p' \ddot{M} \dot{X}_q) \\ &- \frac{1}{2NT} \sum_{t=1}^{T} f_t^{*\prime} \Lambda^{*\prime} [I_N - (\rho - \rho^*) S_N^*]' \ddot{M} [I_N - (\rho - \rho^*) S_N^*] \Lambda^* f_t^* \\ &+ \sum_{p=0}^{k+1} (\beta_p - \beta_p^*) \frac{1}{NT} \sum_{t=1}^{T} \dot{X}_{tp}' \ddot{M} [I_N - (\rho - \rho^*) S_N^*] \Lambda^* f_t^* - \frac{1}{2N} \operatorname{tr}(R) + \frac{1}{2N} \ln |R| + \frac{1}{2} \end{aligned}$$

and

$$\mathcal{L}_2(\theta) = \frac{1}{NT}\sum_{p=0}^{k+1}(\beta_p - \beta_p^*)\sum_{t=1}^{T}\dot{X}_{tp}'\ddot{M}[I_N - (\rho-\rho^*)S_N^*]\dot{e}_t$$
$$+ \frac{1}{NT}\sum_{t=1}^{T} f_t^{*\prime}\Lambda^{*\prime}[I_N - (\rho-\rho^*)S_N^*]'\ddot{M}[I_N - (\rho-\rho^*)S_N^*]\dot{e}_t$$
$$+ \frac{1}{2N^2T}\sum_{t=1}^{T}\dot{e}_t'[I_N - (\rho-\rho^*)S_N^*]'\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}[I_N - (\rho-\rho^*)S_N^*]\dot{e}_t$$
$$- \frac{1}{2NT}\sum_{t=1}^{T}\operatorname{tr}\Big[[I_N - (\rho-\rho^*)S_N^*]'\Sigma_{ee}^{-1}[I_N - (\rho-\rho^*)S_N^*](\dot{e}_t\dot{e}_t' - \Sigma_{ee}^*)\Big]$$

with

$$R = (I_N - \rho W_N)(I_N - \rho^* W_N)^{-1}\Sigma_{ee}^*(I_N - \rho^* W_N)^{-1\prime}(I_N - \rho W_N)'\Sigma_{ee}^{-1}.$$

Since $\hat{\theta}$ maximizes the objective function, we have $\mathcal{L}_1(\hat{\theta}) + \mathcal{L}_2(\hat{\theta}) \geq \mathcal{L}_1(\theta^*) + \mathcal{L}_2(\theta^*)$. It is easy to see that $\mathcal{L}_1(\theta^*) = 0$. Given this result, we have $\mathcal{L}_1(\hat{\theta}) \geq \mathcal{L}_2(\theta^*) - \mathcal{L}_2(\hat{\theta}) \geq -2\sup_{\theta\in\Theta}|\mathcal{L}_2(\theta)|$. However, Lemma A.1 shows that $|\mathcal{L}_2(\theta)| = o_p(1)$ uniformly on Θ. Thus, $\mathcal{L}_1(\hat{\theta}) \geq -|o_p(1)|$. Now consider $\mathcal{L}_1(\hat{\theta})$, which can be alternatively written as

$$\mathcal{L}_1(\hat{\theta}) = -\frac{1}{2}\sum_{p=0}^{k+1}\sum_{q=0}^{k+1}(\hat{\beta}_p - \beta_p^*)(\hat{\beta}_q - \beta_q^*)\frac{1}{NT}\operatorname{tr}(\dot{X}_p'\widehat{\ddot{M}}\dot{X}_q M_{F^*})$$
$$- \frac{1}{2NT}\operatorname{tr}(\eta'\widehat{\ddot{M}}\eta) - \Big\{\frac{1}{2N}\operatorname{tr}(\hat{R}) - \frac{1}{2N}\ln|\hat{R}| - \frac{1}{2}\Big\}$$

with

$$\eta = \sum_{p=0}^{k+1}(\hat{\beta}_p - \beta_p^*)\dot{X}_p F^*(F^{*\prime}F^*)^{-1/2} - [I_N - (\hat{\rho}-\rho^*)S_N^*]\Lambda^*(F^{*\prime}F^*)^{1/2}$$

and

$$\hat{R} = (I_N - \hat{\rho} W_N)(I_N - \rho^* W_N)^{-1}\Sigma_{ee}^*(I_N - \rho^* W_N)^{-1\prime}(I_N - \hat{\rho} W_N)'\hat{\Sigma}_{ee}^{-1}.$$

It is seen that the three expressions of $\mathcal{L}_1(\theta)$ are all non-positive. The first two expressions are apparent to be non-positive. Now consider the third expression, let τ_i be the ith-largest eigenvalue of the matrix

$$A = \hat{\Sigma}_{ee}^{-1/2}(I_N - \hat{\rho} W_N)(I_N - \rho^* W_N)^{-1}\Sigma_{ee}^*(I_N - \rho^* W_N)^{-1\prime}(I_N - \hat{\rho} W_N)'\hat{\Sigma}_{ee}^{-1/2}. \quad \text{(A.13)}$$

Since A is symmetric, all the eigenvalues are real. Now the third expression of $\mathcal{L}_1(\hat{\theta})$ is equivalent to

$$-\Big[\frac{1}{2N}\sum_{i=1}^{N}\tau_i - \frac{1}{2N}\sum_{i=1}^{N}\ln\tau_i - \frac{1}{2}\Big] = -\frac{1}{2N}\sum_{i=1}^{N}(\tau_i - \ln\tau_i - 1) \leq 0. \quad \text{(A.14)}$$

by the fact that $f(x) = x - \ln x - 1$ achieves its minimum value 0 at $x = 1$. Given $\mathcal{L}_1(\hat{\theta}) \leq 0$ for all $\hat{\theta}$ and $\mathcal{L}_1(\hat{\theta}) \geq -|o_p(1)|$, we have

$$\sum_{p=0}^{k+1}\sum_{q=0}^{k+1}(\hat{\beta}_p - \beta_p^*)(\hat{\beta}_q - \beta_q^*)\frac{1}{NT}\operatorname{tr}(\dot{X}_p'\widehat{\ddot{M}}\dot{X}_q M_{F^*}) = o_p(1); \quad \text{(A.15)}$$

$$\frac{1}{NT}\text{tr}(\eta'\widehat{\ddot{M}}\eta) = o_p(1); \tag{A.16}$$

$$\frac{1}{2N}\text{tr}(\hat{R}) - \frac{1}{2N}\ln|\hat{R}| - \frac{1}{2} = o_p(1). \tag{A.17}$$

We first prove consistency of $\hat{\omega}$ under the local identification conditions, i.e., under Assumption G(i). Notice that the left hand side of (A.15) is equivalent to

$$(\hat{\omega}-\omega^*)'\widehat{\mathbb{D}}_a(\hat{\omega}-\omega^*) = o_p(1).$$

where $\widehat{\mathbb{D}}_a$ is the matrix $\mathbb{D}_a$ when $\Lambda = \hat{\Lambda}$ and $\Sigma_{ee} = \hat{\Sigma}_{ee}$. By Assumption G(i), $\widehat{\mathbb{D}}_a$ is positive definite, we have $\hat{\omega} \xrightarrow{p} \omega^*$.

When Assumption G(i) fails, we show that the consistency of $\hat{\omega}$ can still be obtained by Assumption G(ii). We first prove $\hat{\rho} \xrightarrow{p} \rho^*$ under the global identification condition (3.4). By (A.14), equation (A.17) is equal to

$$o_p(1) = \frac{1}{2N}\sum_{i=1}^{N}(\tau_i - \ln\tau_i - 1). \tag{A.18}$$

Consider matrix A in (A.13). By the boundedness of $\hat{\rho}, \hat{\sigma}_i^2$, it is easy to see $\tau_i \in [C^{-1}, C]$ for all i for some large constant C. In addition, there exists a constant b (for example $b = \frac{1}{4C^2}$), such that $x - \ln x - 1 \geq b(x-1)^2$ for all $x \in [C^{-1}, C]$. Given this result, we have

$$\frac{1}{2N}\text{tr}(\hat{R}) - \frac{1}{2N}\ln|\hat{R}| - \frac{1}{2} = \frac{1}{2N}\sum_{i=1}^{N}(\tau_i - \ln\tau_i - 1) \geq b\frac{1}{2N}\sum_{i=1}^{N}(\tau_i-1)^2 = b\frac{1}{2N}\big\|A - I_N\big\|^2,$$

implying

$$\frac{1}{N}\big\|A - I_N\big\|^2 = o_p(1).$$

Let $\hat{\Psi}_N = (I_N - \hat{\rho}W_N)(I_N - \rho^* W_N)^{-1} = I_N - (\hat{\rho}-\rho^*)S_N^*$. Now $A = \hat{\Sigma}_{ee}^{-1/2}\Psi_N\Sigma_{ee}^*\Psi_N'\hat{\Sigma}_{ee}^{-1/2}$. The above result is equivalent to

$$\frac{1}{N}\text{tr}\Big[(\hat{\Sigma}_{ee}^{-1/2}\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N'\hat{\Sigma}_{ee}^{-1/2} - I_N)'(\hat{\Sigma}_{ee}^{-1/2}\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N'\hat{\Sigma}_{ee}^{-1/2} - I_N)\Big] = o_p(1),$$

which can be written as

$$\frac{1}{N}\text{tr}\Big[\hat{\Sigma}_{ee}^{-1/2}(\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee})\hat{\Sigma}_{ee}^{-1}(\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee})'\hat{\Sigma}_{ee}^{-1/2}\Big] = o_p(1).$$

However, by the boundedness of $\hat{\sigma}_i^2$, there exists some constant c such that $\hat{\Sigma}_{ee}^{-1/2} \geq cI_N$. Thus

$$o_p(1) = \frac{1}{N}\text{tr}\Big[\hat{\Sigma}_{ee}^{-1/2}(\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee})\hat{\Sigma}_{ee}^{-1}(\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee})'\hat{\Sigma}_{ee}^{-1/2}\Big]$$
$$\geq c^4\frac{1}{N}\text{tr}\Big[(\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee})(\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee})'\Big] = c^4\frac{1}{N}\big\|\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee}\big\|^2 > 0.$$

So we have

$$\frac{1}{N}\big\|\hat{\Psi}_N\Sigma_{ee}^*\hat{\Psi}_N' - \hat{\Sigma}_{ee}\big\|^2 = o_p(1).$$

By $\hat{\Psi}_N = I_N - (\hat{\rho} - \rho^*)S_N^*$, the above result is equivalent to

$$\frac{1}{N}\sum_{i=1}^{N}\Big(\sigma_i^{*2} - \hat{\sigma}_i^2 - 2(\hat{\rho} - \rho^*)S_{ii,N}^*\sigma_i^{*2} + (\hat{\rho} - \rho^*)^2\sum_{j=1}^{N}S_{ij,N}^{*2}\sigma_j^{*2}\Big)^2$$

$$+(\hat{\rho} - \rho^*)^2\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\Big(S_{ij,N}^*\sigma_j^{*2} + S_{ji,N}^*\sigma_i^{*2} - (\hat{\rho} - \rho^*)\sum_{p=1}^{N}S_{ip,N}^*S_{jp,N}^*\sigma_p^{*2}\Big)^2 = o_p(1).$$

The two expressions on the left hand side are both nonnegative, so we have

$$\frac{1}{N}\sum_{i=1}^{N}\Big(\sigma_i^{*2} - \hat{\sigma}_i^2 - 2(\hat{\rho} - \rho^*)S_{ii,N}^*\sigma_i^{*2} + (\hat{\rho} - \rho^*)^2\sum_{j=1}^{N}S_{ij,N}^{*2}\sigma_j^{*2}\Big)^2 = o_p(1), \tag{A.19}$$

$$(\hat{\rho} - \rho^*)^2\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\Big(S_{ij,N}^*\sigma_j^{*2} + S_{ji,N}^*\sigma_i^{*2} - (\hat{\rho} - \rho^*)\sum_{p=1}^{N}S_{ip,N}^*S_{jp,N}^*\sigma_p^{*2}\Big)^2 = o_p(1). \tag{A.20}$$

Result (A.20) implies $\hat{\rho} \xrightarrow{p} \rho^*$ in view of (3.4). Given the consistency of $\hat{\rho}$, equation (A.15) now can be simplified as

$$\sum_{p=0}^{k}\sum_{q=0}^{k}(\hat{\beta}_p - \beta_p^*)(\hat{\beta}_q - \beta_q^*)\frac{1}{NT}\text{tr}(\dot{X}_p'\widehat{\widehat{M}}\dot{X}_qM_{F^*}) = o_p(1).$$

Let $\hat{\omega}^\dagger = (\hat{\delta}, \hat{\beta}')'$ and $\omega^{\dagger *} = (\delta^*, \beta^{*\prime})'$. By the definition of $\mathbb{D}_b$, the preceding equation is equivalent to

$$(\hat{\omega}^\dagger - \omega^{\dagger *})'\hat{\mathbb{D}}_b(\hat{\omega}^\dagger - \omega^{\dagger *}) = o_p(1).$$

where $\hat{\mathbb{D}}_b$ is the matrix of $\mathbb{D}_b$ when $\Lambda = \hat{\Lambda}$ and $\Sigma_{ee} = \hat{\Sigma}_{ee}$. By Assumption G(ii), we have $\hat{\omega}^\dagger \xrightarrow{p} \omega^{\dagger *}$. Given the consistency of $\hat{\rho}$ and $\hat{\omega}^\dagger$, we have proved $\hat{\omega} \xrightarrow{p} \omega^*$ under Assumption G(ii).

Given the consistency of $\hat{\rho}$, equation (A.17) now can be simplified as

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sigma_i^{*2}}{\hat{\sigma}_i^2} - \ln\frac{\sigma_i^{*2}}{\hat{\sigma}_i^2} - 1\right) = o_p(1).$$

Since $\hat{\sigma}_i^2$ and σ_i^{*2} are both bounded by Assumption H, by the similar arguments following (A.18), we have that there exists a constant b such that

$$o_p(1) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sigma_i^{*2}}{\hat{\sigma}_i^2} - \ln\frac{\sigma_i^{*2}}{\hat{\sigma}_i^2} - 1\right) = b\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sigma_i^{*2}}{\hat{\sigma}_i^2} - 1\right)^2 \geq bC^{-2}\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^{*2})^2,$$

which gives

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^{*2})^2 = o_p(1).$$

We further consider (A.16). By the definition of η, we have

$$\begin{aligned}\eta &= -\Lambda^*(F^{*\prime}F^*)^{1/2} + (\hat{\rho} - \rho^*)S_N^*\Lambda^*(F^{*\prime}F^*)^{1/2} + \sum_{p=0}^{k+1}(\hat{\beta}_p - \beta_p^*)\dot{X}_pF^*(F^{*\prime}F^*)^{-1/2} \\ &= -\Lambda^*(F^{*\prime}F^*)^{1/2} + \Xi_1 + \Xi_2, \qquad \text{say.}\end{aligned} \tag{A.21}$$

Given the consistency of $\hat{\omega}$, together with Lemma A.2, we can simplify (A.16) as

$$\operatorname{tr}\Big\{\Big(\frac{1}{T}F^{*\prime}F^{*}\Big)^{-1}\Big[\frac{1}{N}\Lambda^{*\prime}\widehat{\ddot{M}}\Lambda^{*}\Big]\Big(\frac{1}{T}F^{*\prime}F^{*}\Big)^{-1}\Big\}=o_p(1).$$

Since the matrix in the trace operator is positive definite, we have

$$\Big(\frac{1}{T}F^{*\prime}F^{*}\Big)^{-1}\Big[\frac{1}{N}\Lambda^{*\prime}\widehat{\ddot{M}}\Lambda^{*}\Big]\Big(\frac{1}{T}F^{*\prime}F^{*}\Big)^{-1}=o_p(1),$$

implying $\frac{1}{N}\Lambda^{*\prime}\widehat{\ddot{M}}\Lambda^{*}=o_p(1)$ by Assumption B on F^{*}. This completes the proof of Proposition 5.1.

Appendix B: Detailed proofs for the convergence rates

From Appendices B to F, we drop the superscript "*" from the parameters of the underlying true values for notational simplicity. The following lemmas are useful for the subsequent analysis.

Lemma B.1 *Let d_{NT} be defined in (B.2) below. Then we have*

$$\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}d_{NT}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=o_p(1).$$

Proof of Lemma B.1. By definition, d_{NT} composes of 26 terms. We only choose the first, tenth and thirteenth terms to prove. The proofs of the remaining terms are similar and simpler.

Consider the first term, which is

$$(\hat{\delta}-\delta)^2\frac{1}{N^2T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\sum_{t=1}^{T}\dot{Y}_{t-1}\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=(\hat{\delta}-\delta)^2\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\hat{\lambda}_i\hat{\lambda}_j'\sum_{t=1}^{T}\dot{y}_{it-1}\dot{y}_{jt-1},$$

which, by the boundedness of $\hat{\sigma}_i^2$ by Assumption H, is bounded in norm by

$$(\hat{\delta}-\delta)^2C\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}\dot{y}_{jt-1}\Big|^2\Big]^{1/2}.$$

By Assumptions A-F, it is seen that

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}\dot{y}_{jt-1}\Big|^2=O_p(1).$$

In addition, we also have

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2=\operatorname{tr}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\hat{\lambda}_i\hat{\lambda}_i'\Big]=\operatorname{tr}\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]=\operatorname{tr}[I_r]=r.$$

Given the above three results, together with $\hat{\delta}-\delta=o_p(1)$ by Proposition 5.1, we have that the first term is $o_p(1)$.

Consider the tenth term, which is

$$\frac{1}{N^2T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\sum_{t=1}^{T}f_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=\Big(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\hat{\lambda}_i\lambda_i'\Big)\Big(\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_j^2}f_te_{jt}\hat{\lambda}_j'\Big).$$

The above expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_te_{jt}\Big\|^2\Big]^{1/2}=O_p(T^{-1/2})$$

by Assumptions B and C.

Consider the thirteenth term, which is

$$\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Sigma_{ee}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\le C\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=C\frac{1}{N}I_r=o_p(1),$$

where we have used the fact $\hat{\Sigma}_{ee}^{-1}\Sigma_{ee}\le CI_N$ by the boundedness of $\hat{\sigma}_i^2$ and σ_i^2 by Assumptions C and H. This completes the proof. $\Box$

Proposition B.1 *Under Assumptions A-H, we have*

$$\hat{V}\xrightarrow{p}\Sigma_F,\qquad \hat{D}\triangleq\hat{V}^{-1}\xrightarrow{p}\Sigma_F^{-1},\qquad H\equiv\hat{V}^{-1}\Big(\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big)\Big(\frac{1}{T}F'F\Big)=O_p(1),$$

where Σ_F is the diagonal matrix whose diagonal elements are the eigenvalues of $\lim\limits_{T\to\infty}\frac{1}{T}F'F$ arranged in a descending order.

Proof of Proposition B.1. The first order condition for Λ gives

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}(\dot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\hat{\rho}\ddot{Y}_t-\dot{X}_t\hat{\beta})(\dot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\hat{\rho}\ddot{Y}_t-\dot{X}_t\hat{\beta})'\Big]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=\hat{\Lambda}\hat{V}.$$

By $\dot{Y}_t=\delta\dot{Y}_{t-1}+\rho\ddot{Y}_t+\dot{X}_t\beta+\Lambda f_t+\dot{e}_t$, we have

$$\dot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\hat{\rho}\ddot{Y}_t-\dot{X}_t\hat{\beta}=-(\hat{\delta}-\delta)\dot{Y}_{t-1}-(\hat{\rho}-\rho)\ddot{Y}_t-\dot{X}_t(\hat{\beta}-\beta)+\Lambda f_t+\dot{e}_t \tag{B.1}$$

where $\ddot{Y}_t=W_N\dot{Y}_t$. Using the above expression, we can rewrite the preceding first order condition as

$$\Big\{(\hat{\delta}-\delta)^2\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}\dot{Y}_{t-1}'+(\hat{\rho}-\rho)^2\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t\ddot{Y}_t'+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{X}_t' \tag{B.2}$$

$$+(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}\ddot{Y}_t'+(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t\dot{Y}_{t-1}'+(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}(\hat{\beta}-\beta)'\dot{X}_t'$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)\dot{Y}_{t-1}'(\hat{\delta}-\delta)+(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t(\hat{\beta}-\beta)'\dot{X}_t'+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)\ddot{Y}_t'(\hat{\rho}-\rho)$$

$$+\frac{1}{N}\Lambda\Big[\frac{1}{T}F'F\Big]\Lambda'+\Lambda\frac{1}{NT}\sum_{t=1}^{T}f_te_t'+\frac{1}{NT}\sum_{t=1}^{T}e_tf_t'\Lambda'+\frac{1}{NT}\sum_{t=1}^{T}[e_te_t'-\Sigma_{ee}]+\frac{1}{N}\Sigma_{ee}-\frac{1}{N}\bar{e}\bar{e}'$$

$$-(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}f_t'\Lambda'-(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t f_t'\Lambda'-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)f_t'\Lambda'$$

$$-\frac{1}{NT}\Lambda\sum_{t=1}^{T}f_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'-\frac{1}{NT}\Lambda\sum_{t=1}^{T}f_t(\hat{\rho}-\rho)\ddot{Y}_t'-\frac{1}{NT}\Lambda\sum_{t=1}^{T}f_t(\hat{\beta}-\beta)'\dot{X}_t'$$

$$-(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}e_t'-(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t e_t'-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)e_t'$$

$$-\frac{1}{NT}\sum_{t=1}^{T}e_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'-\frac{1}{NT}\sum_{t=1}^{T}e_t(\hat{\rho}-\rho)\ddot{Y}_t'-\frac{1}{NT}\sum_{t=1}^{T}e_t(\hat{\beta}-\beta)'\dot{X}_t'\Big\}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=\hat{\Lambda}\hat{V}.$$

Let d_{NT} denote the expression in the bracket excluding the term $\frac{1}{NT}\Lambda F'F\Lambda'$. Pre-multiplying $\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}$ on both sides of the preceding equation, together with the normalization condition $\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=I_r$, it follows that

$$\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Big[\frac{1}{NT}\Lambda F'F\Lambda'+d_{NT}\Big]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=\hat{V}. \tag{B.3}$$

Given Lemma B.1, we have

$$\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\Big[\frac{1}{T}F'F\Big]\Big[\frac{1}{N}\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]-\hat{V}=o_p(1). \tag{B.4}$$

Since $\hat{V}$ is a diagonal matrix, the above equation implies that the eigenvalues of the matrix

$$\Big[\frac{1}{T}F'F\Big]\Big[\frac{1}{N}\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]$$

are equal to the counterparts of $\hat{V}$ plus a $o_p(1)$ term by the fact that M_1M_2 and M_2M_1 have the same eigenvalues for any square matrix M_1 and M_2. The last result of Proposition 5.1 is equivalent to

$$\Big[\frac{1}{N}\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]-\frac{1}{N}\Lambda'\hat{\Sigma}_{ee}^{-1}\Lambda=o_p(1).$$

However,

$$\frac{1}{N}\Lambda'\hat{\Sigma}_{ee}^{-1}\Lambda=\frac{1}{N}\Lambda'\Sigma_{ee}^{-1}\Lambda-\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i\lambda_i'=I_r-\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i\lambda_i'.$$

Since

$$\Big\|\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i\lambda_i'\Big\|\le C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^4\Big]^{1/2}=o_p(1)$$

by the second result of Proposition 5.1 and Assumption B, we have

$$\Big[\frac{1}{N}\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\xrightarrow{p}I_r.$$

The above result, together with (B.4), gives the first two results of the proposition. The third result is directly from the definition of H and the fact

$$\Big\|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big\|\le C\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}=C\sqrt{r}\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}. \tag{B.5}$$

This completes the proof of Proposition B.1. □

Lemma B.2 *Under Assumptions A-H, we have*

$$\frac{1}{N}\sum_{i=1}^{N}\|\mathbb{T}_{i1}\|^2 = O_p(\|\hat{\omega}-\omega\|^4), \quad \frac{1}{N}\sum_{i=1}^{N}\|\mathbb{T}_{i2}\|^2 = O_p(\|\hat{\omega}-\omega\|^2), \quad \frac{1}{N}\sum_{i=1}^{N}\|\mathbb{T}_{i,3}\|^2 = O_p(\|\hat{\omega}-\omega\|^2),$$

where $\mathbb{T}_{i1}, \mathbb{T}_{i2}$ and $\mathbb{T}_{i3}$ are defined in (B.6) below.

Proof of Lemma B.2. Consider the first result. By definition, $\mathbb{T}_{i1}$ consists of nigh terms. We use $I_{i1}, \dots, I_{i9}$ to denote them. By the Cauchy-Schwarz inequality, we have

$$\frac{1}{N}\sum_{i=1}^{N}\|\mathbb{T}_{i1}\|^2 \le 9\Big(\frac{1}{N}\sum_{i=1}^{N}\|I_{i1}\|^2 + \cdots + \|I_{i9}\|^2\Big).$$

The proofs of these nigh terms are similar. We only choose the first one to prove. Notice

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\|I_{i1}\|^2 &\le (\hat{\delta}-\delta)^4\|\hat{D}\|^2\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\dot{y}_{it-1}\Big\|^2 \\ &\le C(\hat{\delta}-\delta)^4\|\hat{D}\|^2\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}\dot{y}_{jt-1}\Big|^2\Big],\end{aligned}$$

which is $O_p(\|\hat{\omega}-\omega\|^4)$. Then we have the first result.

Consider the second result. By definition, $\mathbb{T}_{i2}$ consists of eight terms. We use $II_{i1}, \dots, II_{i8}$ to denote them. By the Cauchy-Schwarz inequality, we have

$$\frac{1}{N}\sum_{i=1}^{N}\|\mathbb{T}_{i2}\|^2 \le 8\Big(\frac{1}{N}\sum_{i=1}^{N}\|II_{i1}\|^2 + \cdots + \|II_{i8}\|^2\Big).$$

The proofs of these eight terms are similar. We only choose the first one to prove. Notice

$$\frac{1}{N}\sum_{i=1}^{N}\|II_{i1}\| \le (\delta-\delta)^2\|\hat{D}\|^2\cdot\Big\|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big\|^2\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_t\dot{y}_{it-1}\Big\|^2\Big] = O_p(\|\hat{\omega}-\omega\|^2)$$

by (B.5) and Proposition B.1. So we obtain the second result.

The proof of the third result is similar as that of the second one. The details are therefore omitted. □

Proposition B.2 *Under Assumptions A-H,*

$$(a)\ \ \frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2 = O_p(\frac{1}{N^2}) + O_p(T^{-1}) + O_p(\|\hat{\omega}-\omega\|^2),$$

$$(b)\ \ \frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^4 = O_p(N^{-4}) + O_p(T^{-2}) + O_p(\|\hat{\omega}-\omega\|^4).$$

Proof of Proposition B.2. Consider (a). Equation (B.2) can be rewritten as

$$\hat{\lambda}_i - H\lambda_i = \hat{D}\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\Big[\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\Big] + \hat{D}\Big[\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_tf_t'\Big]\lambda_i \qquad \text{(B.6)}$$

$$+\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}[e_t e_{it}-E(e_t e_{it})]+\frac{1}{N}\hat{D}\hat{\lambda}_i\frac{\sigma_i^2}{\hat{\sigma}_i^2}-\frac{1}{N}\hat{D}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}\bar{e}_i+\mathbb{T}_{i1}+\mathbb{T}_{i2}+\mathbb{T}_{i3}$$

where

$$H=\hat{V}^{-1}\Big(\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big)\Big(\frac{1}{T}F'F\Big)=\hat{D}\Big(\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big)\Big(\frac{1}{T}F'F\Big),$$

and

$$\mathbb{T}_{i1}=(\hat{\delta}-\delta)^2\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\dot{y}_{it-1}+(\hat{\rho}-\rho)^2\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\ddot{y}_{it} \tag{B.7}$$

$$+\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{x}_{it}+(\hat{\delta}-\delta)(\hat{\rho}-\rho)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\dot{y}_{it-1}$$

$$(\hat{\delta}-\delta)(\hat{\rho}-\rho)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\ddot{y}_{it}+\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\dot{x}_{it}$$

$$+\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\dot{x}_{it}+\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)\dot{y}_{it-1}(\hat{\delta}-\delta)$$

$$+\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)\ddot{y}_{it}(\hat{\rho}-\rho).$$

and

$$\mathbb{T}_{i2}=-(\hat{\delta}-\delta)\hat{D}\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\frac{1}{T}\sum_{t=1}^{T}f_t\dot{y}_{it-1}-(\hat{\rho}-\rho)\hat{D}\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\frac{1}{T}\sum_{t=1}^{T}f_t\ddot{y}_{it} \tag{B.8}$$

$$-\hat{D}\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{it}'(\hat{\beta}-\beta)-(\hat{\delta}-\delta)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}f_t'\lambda_i$$

$$-(\hat{\rho}-\rho)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t f_t'\lambda_i-\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)f_t'\lambda_i$$

and

$$\mathbb{T}_{i3}=-(\hat{\delta}-\delta)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t\dot{y}_{it-1}-(\hat{\rho}-\rho)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t\ddot{y}_{it} \tag{B.9}$$

$$-\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t\dot{x}_{it}'(\hat{\beta}-\beta)-(\hat{\delta}-\delta)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}e_{it}$$

$$-(\hat{\rho}-\rho)\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t e_{it}-\hat{D}\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)e_{it}$$

There are eight terms on the right hand side of (B.6), we use $I_{i1},\dots,I_{i8}$ to denote them. By the Cauchy-Schwarz inequality,

$$\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^2\le 8\frac{1}{N}\sum_{i=1}^{N}(\|I_{i1}\|^2+\|I_{i2}\|^2+\cdots+\|I_{i8}\|^2).$$

By Lemma B.2, we have

$$\frac{1}{N}\sum_{i=1}^{N}(\|I_{i6}\|^2+\|I_{i7}\|^2+\|I_{i8}\|^2)=O_p(\|\hat{\omega}-\omega\|^2).$$

So it suffices to examine the first five terms. Consider the first term, which is bounded in norm by

$$\|\hat{D}\|^2 \cdot \|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\|^2 \frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t e_{it}\Big\|^2 = O_p(T^{-1})$$

by Proposition B.1 and (B.5). The second term can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\Big\|\hat{D}\Big[\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_j^2}\hat{\lambda}_j f_t' e_{jt}\Big]\lambda_i\Big\|^2,$$

which is bounded in norm by

$$C\|\hat{D}\|^2\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t e_{it}\|^2\Big] = O_p(T^{-1})$$

by Proposition B.1 and Assumption B. Similarly the third term is bounded in norm by

$$C\|\hat{D}\|^2\Big[\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}\|\hat{\lambda}_j\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}[e_{it}e_{jt} - E(e_{it}e_{jt})]\Big|^2\Big] = O_p(T^{-1}).$$

The fourth term is bounded in norm by

$$\|\hat{D}\|^2\frac{1}{N^3}\sum_{i=1}^{N}\frac{\sigma_i^4}{\hat{\sigma}_i^4}\|\hat{\lambda}_i\|^2 \le C\|\hat{D}\|^2\frac{1}{N^3}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2 = C\frac{1}{N^2}r\|\hat{D}\|^2 = O_p(\frac{1}{N^2}).$$

where we use the fact that there exists a constant C large enough such that $\hat{\sigma}_i^{-4}\sigma_i^4 \le C\hat{\sigma}_i^{-2}$ (i.e. $\hat{\sigma}_i^{-2}\sigma_i^4 \le C$). The last term is bounded in norm by

$$\frac{1}{T^2}\|\hat{D}\|^2\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}(\sqrt{T}\bar{e}_i)^2\Big]^2 = O_p(T^{-2}).$$

Summarizing all the results, we have (a).

Consider (b). By the inequality

$$\|I_{i1} + I_{i2} + \cdots + I_{i8}\|^4 \le 8^4\|I_{i1}\|^4 + \|I_{i2}\|^4 + \cdots + \|I_{i8}\|^4,$$

we have

$$\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^4 \le 8^4\frac{1}{N}\sum_{i=1}^{N}(\|I_{i1}\|^4 + \|I_{i2}\|^4 + \cdots + \|I_{i8}\|^4).$$

Now the proof proceeds similarly as (a). The first term is bounded in norm by

$$\|\hat{D}\|^4 \cdot \|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\|^4 \frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t e_{it}\Big\|^4 = O_p(T^{-2}).$$

The second term is bounded in norm by

$$C\|\hat{D}\|^4\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^4\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]^2\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t e_{it}\|^2\Big]^2 = O_p(T^{-2}).$$

The third term is bounded in norm by

$$C\|\hat{D}\|^4\Big[\frac{1}{N}\sum_{i=1}^N\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]^2\Big[\frac{1}{N}\sum_{i=1}^N\Big(\frac{1}{N}\sum_{j=1}^N\Big|\frac{1}{T}\sum_{t=1}^T[e_{jt}e_{it}-E(e_{jt}e_{it})]\Big|^2\Big)^2\Big]=O_p(T^{-2}).$$

The fourth term is

$$\frac{1}{N}\sum_{i=1}^N\Big\|\frac{1}{N}\hat{D}\hat{\lambda}_i\frac{\sigma_i^2}{\hat{\sigma}_i^2}\Big\|^4\le 2^4\frac{1}{N}\sum_{i=1}^N\Big\|\frac{1}{N}\hat{D}(\hat{\lambda}_i-H\lambda_i)\frac{\sigma_i^2}{\hat{\sigma}_i^2}\Big\|^4+2^4\frac{1}{N}\sum_{i=1}^N\Big\|\frac{1}{N}\hat{D}H\lambda_i\frac{\sigma_i^2}{\hat{\sigma}_i^2}\Big\|^4.$$

By the boundedness of $\hat{\sigma}_i^2,\sigma_i^2$, the first term is $C\|\hat{D}\|^4\frac{1}{N^5}\sum_{i=1}^N\|\hat{\lambda}_i-H\lambda_i\|^4$, which is of smaller order term than $\frac{1}{N}\sum_{i=1}^N\|\hat{\lambda}_i-H\lambda_i\|^4$ and hence negligible. The second term is $O_p(N^{-4})$ by the boundedness of $\hat{\sigma}_i^2,\sigma_i^2$ and λ_i. So the fourth term is $O_p(N^{-4})$. The fifth term is bounded in norm by

$$C\|\hat{D}\|^4\Big[\frac{1}{N}\sum_{i=1}^N\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]^2\Big[\frac{1}{N}\sum_{i=1}^N(\sqrt{T}\bar{e}_i)^2\Big]^2\frac{1}{N}\sum_{i=1}^N(\sqrt{T}\bar{e}_i)^4=O_p(T^{-4})$$

The last three terms can be proved to be $O_p(\|\hat{\omega}-\omega\|^4)$ by the similar method as in proving Lemma B.2. This proves (b). □

Lemma B.3 *Let $\mathbb{V}$ be defined in (B.10) below. Under Assumptions A-H,*

$$\mathbb{V}=H^{-1\prime}\cdot O_p(\|\hat{\omega}-\omega\|).$$

The proof of Lemma B.3 is similar as that of Lemma B.2. The details are therefore omitted.

Proposition B.3 *Under Assumptions A-H,*

$$\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda-H^{-1\prime}=O_p(N^{-1})+O_p(T^{-1/2})+O_p(\|\hat{\omega}-\omega\|).$$

Proof of Proposition B.3. Consider equation (B.2). Pre-multiplying $\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}$ and post-multiplying $\hat{D}=\hat{V}^{-1}$, we have

$$\begin{aligned}I_r-\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]H'&=\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_tf_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\frac{1}{NT}\sum_{t=1}^Tf_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}\\&+(\hat{\delta}-\delta)^2\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+(\hat{\rho}-\rho)^2\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}\\&+\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}\\&+(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}\\&+\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+\frac{1}{N^2T}\sum_{t=1}^T\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}\end{aligned}$$

$$+\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-(\hat{\delta}-\delta)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}f_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_tf_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)f_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-\left[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\right]\frac{1}{NT}\sum_{t=1}^{T}f_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\left[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\right]\frac{1}{NT}\sum_{t=1}^{T}f_t(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-\left[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\right]\frac{1}{NT}\sum_{t=1}^{T}f_t(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-(\hat{\delta}-\delta)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}e_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)e_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$+\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\dot{e}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$

Post-multiplying $H^{-1\prime}$ on both sides,

$$\left[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\right]=H^{-1\prime}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_tf_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime} \tag{B.10}$$
$$-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\frac{1}{NT}\sum_{t=1}^{T}f_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\dot{e}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+\mathbb{V}$$

where

$$\mathbb{V}=-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$-(\hat{\delta}-\delta)^2\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}-(\hat{\rho}-\rho)^2\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$+\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$+(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+(\hat{\delta}-\delta)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}f_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$+(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_tf_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)f_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$+\left[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\right]\frac{1}{NT}\sum_{t=1}^{T}f_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+\left[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\right]\frac{1}{NT}\sum_{t=1}^{T}f_t(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$+\left[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\right]\frac{1}{NT}\sum_{t=1}^{T}f_t(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+(\hat{\delta}-\delta)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}e_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$

$$-\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)e_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$-(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$
$$-(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}+\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}$$

Consider the second term on the right hand side of (B.10). Ignore $H^{-1\prime}$, this term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\Big\|^2\Big]^{1/2}\|\hat{D}\|=O_p(T^{-1/2})$$

by Proposition B.1. So the second term is $H^{-1}\cdot O_p(T^{-1/2})$. The third term can be proved to be $H^{-1}\cdot O_p(T^{-1/2})$ similarly. Consider the fourth term. Ignore $H^{-1\prime}$, this term can be written as

$$\frac{1}{N^2T}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}[e_te_t'-\Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Sigma_{ee}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}\bar{e}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}.$$

The first expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}[e_{it}e_{jt}-E(e_{it}e_{jt})]\Big|^2\Big]^{1/2}\|\hat{D}\|=O_p(T^{-1/2}).$$

By the boundedness of $\hat{\sigma}_i^2$ and σ_i^2, we have $\hat{\Sigma}_{ee}^{-1}\Sigma_{ee}\leq CI_N$ for some constant C. Then the second expression is bounded by

$$C\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=C\frac{1}{N}I_r=O(N^{-1}).$$

The last expression is easy to see $O_p(T^{-1})$. Given the above results, we have that the fourth term on the right hand side of (B.10) is $H^{-1\prime}\cdot[O_p(N^{-1})+O_p(T^{-1/2})]$.

Now consider the expression on the right hand side of (B.10) again. The 2nd-4th terms are $H^{-1\prime}\cdot[O_p(N^{-1})+O_p(T^{-1/2})]$ and the last term is $H^{-1\prime}\cdot O_p(\|\hat{\omega}-\omega\|)$ by Lemma B.3. So $H^{-1\prime}$ dominates the remaining four terms. Given $\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda$ is $O_p(1)$ by (B.5), we have $H^{-1\prime}=O_p(1)$. Given $H^{-1\prime}=O_p(1)$, the second and third terms are now $O_p(T^{-1/2})$; the fourth term is $O_p(N^{-1})+O_p(T^{-1/2})$ and the last term is $O_p(\|\hat{\omega}-\omega\|)$. This proves the proposition. □

Lemma B.4 *Under Assumptions A-H, we have*

$$\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t-H^{-1\prime}f_t\|^2=O_p\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]+O_p(N^{-1})+O_p(T^{-1})+O_p(\|\hat{\omega}-\omega\|^2).$$

The above result, together with Proposition 5.1, implies

$$\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t-H^{-1\prime}f_t\|^2=o_p(1),\qquad \frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\hat{f}_t'=H^{-1\prime}\Big(\frac{1}{T}\sum_{t=1}^{T}f_tf_t'\Big)H^{-1}+o_p(1).$$

Proof of Lemma B.4. By definition, we have

$$\hat{f}_t = (\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda})^{-1}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\dot{Y}_t - \hat{\delta}\dot{Y}_{t-1} - \hat{\rho}\ddot{Y}_t - \dot{X}_t\hat{\beta}) = \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\dot{Y}_t - \hat{\delta}\dot{Y}_{t-1} - \hat{\rho}\ddot{Y}_t - \dot{X}_t\hat{\beta}),$$

where the second equality is due to the normalization condition $N^{-1}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = I_r$. By (B.1), the preceding equation can be written as

$$\hat{f}_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda f_t = \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t - (\hat{\delta}-\delta)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1} - (\hat{\rho}-\rho)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta).$$

The above equation can be alternatively rewritten as

$$\begin{aligned}\hat{f}_t - H^{-1\prime}f_t = & -\Big[H^{-1\prime} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]f_t + \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t - (\hat{\delta}-\delta)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1} \\ & - (\hat{\rho}-\rho)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta).\end{aligned} \tag{B.11}$$

We use $II_{t1}, II_{t2}, \dots, II_{t5}$ to denote the five terms on the right hand side. By the Cauchy-Schwarz inequality, we have

$$\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t - H^{-1\prime}f_t\|^2 = 5\frac{1}{T}\sum_{t=1}^{T}(\|II_{t1}\|^2 + \|II_{t2}\|^2 + \cdots + \|II_{t5}\|^2).$$

The first term is bounded in norm by $\frac{1}{T}\sum_{t=1}^{T}\|f_t\|^2 \cdot \|H^{-1\prime} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\|^2$, which is $O_p(\frac{1}{N^2}) + O_p(T^{-1}) + O_p(\|\hat{\omega}-\omega\|^2)$ by Proposition B.3. The third term is bounded in norm by

$$C(\hat{\delta}-\delta)^2\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\dot{y}_{it-1}^2\Big] = O_p(\|\hat{\omega}-\omega\|^2).$$

The fourth and fifth terms can be proved to be $O_p(\|\hat{\omega}-\omega\|^2)$ similarly. Now consider the second term. Since

$$\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}(\hat{\lambda}_i - H\lambda_i)e_{it} - H\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i e_{it} + H\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i e_{it} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}$$

by the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big\|^2 = & \, 4\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}(\hat{\lambda}_i - H\lambda_i)e_{it}\Big\|^2 + 4\frac{1}{T}\sum_{t=1}^{T}\Big\|H\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i e_{it}\Big\|^2 \\ & + 4\frac{1}{T}\sum_{t=1}^{T}\Big\|H\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i e_{it}\Big\|^2 + 4\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}\Big\|^2.\end{aligned}$$

We use a, b, c, d to denote the four terms on the right hand side. Term a is bounded by

$$4C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2\Big]\Big[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}e_{it}^2\Big] = O_p(\frac{1}{N^2}) + O_p(T^{-1}) + O_p(\|\hat{\omega}-\omega\|)$$

by Proposition B.2. Term b, by the boundedness of $\hat{\sigma}_i^2, \sigma_i^2$ and λ_i, is bounded by

$$4C\|H\|^2\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]\Big[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}e_{it}^2\Big] = O_p\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big].$$

Term c is $O_p(N^{-1})$ and term d is $O_p(T^{-1})$. So we have

$$\frac{1}{T}\sum_{t=1}^{T}\|II_{t2}\|^2 = O_p\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]+O_p(N^{-1})+O_p(T^{-1})+O_p(\|\hat{\omega}-\omega\|).$$

Summarizing all the results, we have prove the first part of the lemma. The second part is the direct result of the first part. This completes the proof. □

Lemma B.5 *Let $\mathbb{U}_{i1}$ and $\mathbb{U}_{i2}$ be defined in (B.12) below. Under Assumptions A-H,*

$$\frac{1}{N}\sum_{i=1}^{N}\|\mathbb{U}_{i1}\|^2 = O_p(\|\hat{\omega}-\omega\|^4), \qquad \frac{1}{N}\sum_{i=1}^{N}\|\mathbb{U}_{i2}\|^2 = O_p(\|\hat{\omega}-\omega\|^2).$$

The proof of Lemma B.5 is similar as that of Lemma B.2. The details are omitted.

Proposition B.4 *Under Assumptions A-H,*

$$(a)\ \ \frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t - H^{-1\prime}f_t\|^2 = O_p(N^{-1})+O_p(T^{-1})+O_p(\|\hat{\omega}-\omega\|^2);$$

$$(b)\ \ \frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2 = O_p(\frac{1}{N^2})+O_p(T^{-1})+O_p(\|\hat{\omega}-\omega\|^2).$$

Proof of Proposition B.4. The first order condition for σ_i^2 gives

$$\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\Big[-(\hat{\delta}-\delta)\dot{y}_{it-1}-(\hat{\rho}-\rho)\ddot{y}_{it}-\dot{x}_{it}'(\hat{\beta}-\beta)-(\hat{\lambda}_i-H\lambda_i)'\hat{f}_t-\lambda_i'H'(\hat{f}_t-H^{-1\prime}f_t)+\dot{e}_{it}\Big]^2.$$

The above equation can be written as

$$\begin{aligned}\hat{\sigma}_i^2-\sigma_i^2 &= \frac{1}{T}\sum_{t=1}^{T}(e_{it}^2-\sigma_i^2)-2(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\dot{e}_{it}-2\lambda_i'H'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)\dot{e}_{it}\\ &+2(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t(\hat{f}_t-H^{-1\prime}f_t)'H\lambda_i+(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\hat{f}_t'(\hat{\lambda}_i-H\lambda_i)\\ &+\lambda_i'H'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)(\hat{f}_t-H^{-1\prime}f_t)'H\lambda_i-\bar{e}_i^2+\mathbb{U}_{i1}+\mathbb{U}_{i2} \qquad \text{(B.12)}\end{aligned}$$

where

$$\mathbb{U}_{i1} = (\hat{\delta}-\delta)^2\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}^2+(\hat{\rho}-\rho)^2\frac{1}{T}\sum_{t=1}^{T}\ddot{y}_{it}^2+(\hat{\beta}-\beta)\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{x}_{it}'(\hat{\beta}-\beta)$$

$$+2(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}\ddot{y}_{it}+2(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\dot{y}_{it-1}+2(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\ddot{y}_{it}.$$

and

$$\mathbb{U}_{i2} = 2(\hat{\delta}-\delta)\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}\hat{f}_t'(\hat{\lambda}_i-H\lambda_i)+2(\hat{\rho}-\rho)\frac{1}{T}\sum_{t=1}^{T}\ddot{y}_{it}\hat{f}_t'(\hat{\lambda}_i-H\lambda_i)$$

$$+2(\hat{\beta}-\beta)'\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}\hat{f}_t'(\hat{\lambda}_i-H\lambda_i)+2(\hat{\delta}-\delta)\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}(\hat{f}_t-H^{-1\prime}f_t)'H\lambda_i$$

$$+2(\hat{\rho}-\rho)\frac{1}{T}\sum_{t=1}^{T}\ddot{y}_{it}(\hat{f}_t-H^{-1\prime}f_t)'H\lambda_i+2(\hat{\beta}-\beta)'\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}(\hat{f}_t-H^{-1\prime}f_t)'H\lambda_i$$

$$-2(\hat{\delta}-\delta)\frac{1}{T}\sum_{t=1}^{T}\dot{y}_{it-1}e_{it}-2(\hat{\rho}-\rho)\frac{1}{T}\sum_{t=1}^{T}\ddot{y}_{it}e_{it}-2(\hat{\beta}-\beta)'\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}e_{it}.$$

There are nine terms on the right hand side of (B.12). We use $III_{i1}, III_{i2}, \dots, III_{i9}$ to denote them. By the Cauchy-Schwarz inequality, we have

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\leq 9\frac{1}{N}\sum_{i=1}^{N}(\|III_{i1}\|^2+\|III_{i2}\|^2+\cdots+\|III_{i9}\|^2).$$

The term $\frac{1}{N}\sum_{i=1}^{N}(\|III_{i8}\|^2+\|III_{i9}\|^2)$ is $O_p(\|\hat{\omega}-\omega\|^2)$, which is implied by Lemma B.5. It suffices to consider the first seven terms. The first term $\frac{1}{T}\sum_{t=1}^{T}\|III_{i1}\|^2$ is apparent to be $O_p(T^{-1})$. Consider the second term. Since

$$(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\dot{e}_{it}=(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)\dot{e}_{it}+(\hat{\lambda}_i-H\lambda_i)'H^{-1\prime}\frac{1}{N}\sum_{i=1}^{N}f_t\dot{e}_{it},$$

we have

$$\frac{1}{N}\sum_{i=1}^{N}\|III_{i2}\|^2\leq 8\frac{1}{N}\sum_{i=1}^{N}\Big\|(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)\dot{e}_{it}\Big\|^2$$
$$+8\frac{1}{N}\sum_{i=1}^{N}\Big\|(\hat{\lambda}_i-H\lambda_i)'H^{-1\prime}\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\Big\|^2.$$

The first term on right hand side is bounded in norm by

$$8\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t-H^{-1\prime}f_t\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^4\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big(\frac{1}{T}\sum_{t=1}^{T}\dot{e}_{it}^2\Big)^2\Big]^{1/2},$$

which is $o_p(\frac{1}{N^2})+o_p(T^{-1})+o_p(\|\hat{\omega}-\omega\|^2)$ by Proposition B.2. The second term is bounded in norm by

$$8\|H^{-1}\|^2\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^4\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\Big\|^4\Big]^{1/2},$$

which is $O_p(T^{-2})+O_p(\frac{1}{N^2}T^{-1})+o_p(\|\hat{\omega}-\omega\|^2)$ by Proposition B.2. So we have

$$\frac{1}{N}\sum_{i=1}^{N}\|III_{i2}\|^2=o_p(\frac{1}{N^2})+o_p(T^{-1})+o_p(\|\hat{\omega}-\omega\|^2).$$

The third term is analyzed later and we consider the fourth term. By the Cauchy-Schwarz inequality,

$$\frac{1}{N}\sum_{i=1}^{N}\|III_{i4}\|^2\leq 4C\|H\|^2\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^2\Big]\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t\|^2\Big]\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t-H^{-1\prime}f_t\|^2\Big]$$

which is also $o_p(\frac{1}{N^2}) + o_p(T^{-1}) + o_p(\|\hat{\omega} - \omega\|^2)$ by Proposition B.2 and Lemma B.4. The fifth term is bounded in norm by

$$\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^4\Big]\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t\|^2\Big]^2 = O_p(N^{-4}) + O_p(T^{-2}) + O_p(\|\hat{\omega} - \omega\|^4)$$

by Proposition B.2 and Lemma B.4. The sixth term is bounded in norm by

$$\|H\|^4\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^4\Big]\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t - H^{-1\prime}f_t\|^2\Big]^2 = O_p\Big(\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t - H^{-1\prime}f_t\|^2\Big]^2\Big). \tag{B.13}$$

The seventh term is apparent to be $O_p(T^{-2})$. Now we consider the third term. Substituting (B.11) into the third term, we have

$$\begin{aligned}\lambda_i'H'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t - H^{-1\prime}f_t)\dot{e}_{it} &= -\lambda_i'H'\Big[H^{-1\prime} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\\ &+\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\dot{e}_{it} - (\hat{\delta}-\delta)\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\dot{e}_{it}\\ &-(\hat{\rho}-\rho)\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\dot{e}_{it} - \lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)\dot{e}_{it}.\end{aligned}$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\|III_{i3}\|^2 &\le 5\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H'\Big[H^{-1\prime} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\Big\|^2 \\ &+5\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\dot{e}_{it}\Big\|^2 + 5\frac{1}{N}\sum_{i=1}^{N}\Big\|(\hat{\delta}-\delta)\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\dot{e}_{it}\Big\|^2\\ &+5\frac{1}{N}\sum_{i=1}^{N}\Big\|(\hat{\rho}-\rho)\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\dot{e}_{it}\Big\|^2 + 5\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)\dot{e}_{it}\Big\|^2.\end{aligned} \tag{B.14}$$

By Proposition B.3, the first term is $O_p(\frac{1}{N^2}T^{-1}) + O_p(T^{-2}) + o_p(\|\hat{\omega} - \omega\|^2)$. Using the similar arguments in Lemma B.4, the last three terms can be proved to be $O_p(\|\hat{\omega} - \omega\|^2)$. Consider the second term. Notice that

$$\begin{aligned}\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\dot{e}_{it} &= -\lambda_i'H'H\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_j^2\sigma_j^2}\lambda_j[e_{jt}e_{it} - E(e_{jt}e_{it})]\\ &+\lambda_i'H'H\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_j^2}\lambda_j[e_{jt}e_{it} - E(e_{jt}e_{it})] + \lambda_i'H'\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_j^2}(\hat{\lambda}_j - H\lambda_j)[e_{jt}e_{it} - E(e_{jt}e_{it})]\\ &+\frac{1}{N}\frac{\sigma_i^2}{\hat{\sigma}_i^2}\lambda_i'H'(\hat{\lambda}_i - H\lambda_i) + \frac{1}{N}\frac{\sigma_i^2}{\hat{\sigma}_i^2}\lambda_i'H'H\lambda_i - \lambda_i'H\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}\hat{\lambda}_j\bar{e}_j\bar{e}_i.\end{aligned}$$

Again using the Cauchy-Schwarz inequality, we have

$$\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\dot{e}_{it}\Big\|^2 \le 6\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}\hat{\lambda}_j\bar{e}_j\bar{e}_i\Big\|^2$$

$$+6\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H'H\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_j^2}\lambda_j[e_{jt}e_{it}-E(e_{jt}e_{it})]\Big\|^2+6\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{N}\frac{\sigma_i^2}{\hat{\sigma}_i^2}\lambda_i'H'(\hat{\lambda}_i-H\lambda_i)\Big\|^2$$

$$+6\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H'\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_j^2}(\hat{\lambda}_j-H\lambda_j)[e_{jt}e_{it}-E(e_{jt}e_{it})]\Big\|^2+6\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{N}\frac{\sigma_i^2}{\hat{\sigma}_i^2}\lambda_i'H'H\lambda_i\Big\|^2$$

$$+6\frac{1}{N}\sum_{i=1}^{N}\Big\|\lambda_i'H'H\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_j^2-\sigma_j^2}{\hat{\sigma}_j^2\sigma_j^2}\lambda_j[e_{jt}e_{it}-E(e_{jt}e_{it})]\Big\|^2.$$

The first term is bounded by

$$6C\frac{1}{T^2}\|H\|^2\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\cdot|\sqrt{T}\bar{e}_i|^2\Big]\Big[\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}\|\hat{\lambda}_j\|^2\Big]\Big[\frac{1}{N}\sum_{j=1}^{N}|\sqrt{T}\bar{e}_j|^2\Big]=O_p(T^{-2}).$$

The second term is $O_p(\frac{1}{NT})$. The third term is apparent to be $o_p(\frac{1}{N^2})$. The fourth term is bounded in norm by

$$C\|H\|^2\Big[\frac{1}{N}\sum_{j=1}^{N}\|\hat{\lambda}_j-H\lambda_j\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}[e_{jt}e_{it}-E(e_{jt}e_{it})]\Big|^2\Big],$$

which is $O_p(\frac{1}{N^2}T^{-1})+O_p(T^{-2})+o_p(\|\hat{\omega}-\omega\|^2)$. The fifth term is apparent to be $O_p(\frac{1}{N^2})$ by the boundedness of $\hat{\sigma}_i^2$ and σ_i^2. The last term is bounded by

$$C\|H\|^4\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}[e_{jt}e_{it}-E(e_{jt}e_{it})]\Big|^2\Big], \tag{B.15}$$

which is $O_p(T^{-1})\cdot O_p\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]$. Summarizing all the results, we have that the second term on the right hand side of (B.14) is $O_p(T^{-1})\cdot O_p\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]+O_p(\frac{1}{N^2})+O_p(T^{-2})$. This result, together with the results on the remaining four terms on the right hand side of (B.14), gives

$$\frac{1}{N}\sum_{i=1}^{N}\|III_{i3}\|^2=O_p(\frac{1}{N^2})+O_p(T^{-2})+O_p(T^{-1})\cdot O_p\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]+O_p(\|\hat{\omega}-\omega\|^2).$$

Summarizing the results on $\frac{1}{N}\sum_{i=1}^{N}\|III_{i1}\|^2,\ldots,\frac{1}{N}\sum_{i=1}^{N}\|III_{i9}\|^2$, we have

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2=O_p(\frac{1}{N^2})+O_p(T^{-1})+O_p\Big(\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t-H^{-1\prime}f_t\|^2\Big]^2\Big)+O_p(\|\hat{\omega}-\omega\|^2),$$

where we neglect the term $O_p(T^{-1})\cdot O_p\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]$ since it is of smaller order term than $\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2$. Substituting the result of Lemma B.4 into the above equation, we have

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2=O_p(\frac{1}{N^2})+O_p(T^{-1})+O_p(\|\hat{\omega}-\omega\|^2).$$

The above result, together with Lemma B.4, gives

$$\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t - H^{-1\prime}f_t\|^2 = O_p(N^{-1}) + O_p(T^{-1}) + O_p(\|\hat{\omega}-\omega\|^2).$$

This completes the proof of Proposition B.4. □

Lemma B.6 *Under Assumptions A-H,*

$$(a)\quad \frac{1}{NT}\sum_{t=1}^{T} f_t e_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = O_p(\frac{1}{\sqrt{NT}}) + O_p(T^{-1}) + o_p(\|\hat{\omega}-\omega\|);$$

$$(b)\quad \frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\frac{1}{T}\sum_{t=1}^{T}[e_te_t' - \Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

Proof of Lemma B.6. Using the results in Propositions B.2 and (B.4), the proof of the first result is similar as that of Lemma C.1(e) in the supplement of Bai and Li (2012).

Consider (b). Notice that

$$\begin{aligned}
\frac{1}{N^2T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\sum_{s=1}^{T}(e_se_s' - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\hat{\lambda}_i\hat{\lambda}_j'\frac{1}{T}\sum_{s=1}^{T}[e_{is}e_{js} - E(e_{is}e_{jt})] \\
= H\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\sigma_i^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s}H' &+ \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\hat{\lambda}_i(\hat{\lambda}_j - H\lambda_j)'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s} \\
+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}(\hat{\lambda}_i - H\lambda_i)\lambda_j'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s}H' &- H\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s}H' \\
-H\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s}H', & \qquad \text{(B.16)}
\end{aligned}$$

where $u_{ij,s} = e_{is}e_{js} - E(e_{is}e_{jt})$. The first term on the right hand side is $O_p(\frac{1}{N\sqrt{T}})$. The second term can be written as

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}(\hat{\lambda}_i - H\lambda_i)(\hat{\lambda}_j - H\lambda_j)'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s} + H\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\sigma_i^2\hat{\sigma}_j^2}\lambda_i(\hat{\lambda}_j - H\lambda_j)'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s}$$
$$-H\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2\hat{\sigma}_j^2}\lambda_i(\hat{\lambda}_j - H\lambda_j)'\frac{1}{T}\sum_{s=1}^{T}u_{ij,s}.$$

The first term of the above expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{s=1}^{T}u_{ij,s}\Big\|^2\Big]^{1/2} = O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|)$$

by Proposition B.2. The second term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{j=1}^{N}\|\hat{\lambda}_j - H\lambda_j\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}\Big\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}\lambda_iu_{ij,s}\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-3/2}T^{-1/2}) + O_p(\frac{1}{\sqrt{NT}}) + o_p(\|\hat{\omega} - \omega\|)$. The third term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^N(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^N\|\hat{\lambda}_i - H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Big|\frac{1}{T}\sum_{s=1}^T u_{ij,s}\Big\|^2\Big]^{1/2}$$

which is $O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Propositions B.2 and B.4. So the second term on the right hand side of (B.16) is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. Consider the third term, which can be written as

$$\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\frac{1}{\hat{\sigma}_i^2\sigma_j^2}(\hat{\lambda}_i - H\lambda_i)\lambda_j'\frac{1}{T}\sum_{s=1}^T u_{ij,s}H' - \frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^2}(\hat{\lambda}_i - H\lambda_i)\lambda_j'\frac{1}{T}\sum_{s=1}^T u_{ij,s}H'$$

The first term of the above expression is bounded in norm by

$$C\|H\|\Big[\frac{1}{N}\sum_{i=1}^N\|\hat{\lambda}_i - H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^N\Big\|\frac{1}{NT}\sum_{j=1}^N\sum_{s=1}^T\frac{1}{\sigma_j^2}\lambda_j'u_{ij,s}\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-3/2}T^{-1/2}) + O_p(\frac{1}{\sqrt{NT}}) + o_p(\|\hat{\omega} - \omega\|)$. The second term of the above expression is bounded in norm by

$$C\|H\|\Big[\frac{1}{N}\sum_{i=1}^N\|\hat{\lambda}_i - H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^N(\hat{\sigma}_j^2 - \sigma_j^2)^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Big|\frac{1}{T}\sum_{s=1}^T u_{ij,s}\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Propositions B.2 and B.4. So the third term on the right hand side of (B.16) is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. Consider the fourth term, which can be written as

$$\begin{aligned}&-H\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\frac{(\hat{\sigma}_i^2 - \sigma_i^2)(\hat{\sigma}_j^2 - \sigma_j^2)}{\hat{\sigma}_i^2\sigma_i^2\hat{\sigma}_j^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{s=1}^T u_{ij,s}H' \\ &\quad + H\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\sigma_i^2\hat{\sigma}_j^2\sigma_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{s=1}^T u_{ij,s}H'. \end{aligned} \tag{B.17}$$

The first term is bounded in norm by

$$C\|H\|^2\Big[\frac{1}{N}\sum_{i=1}^N(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\Big|\frac{1}{T}\sum_{s=1}^T u_{ij,s}\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. The second term is bounded in norm by

$$C\|H\|^2\Big[\frac{1}{N}\sum_{j=1}^N(\hat{\sigma}_j^2 - \sigma_j^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^N\Big\|\frac{1}{NT}\sum_{i=1}^N\sum_{s=1}^T\frac{1}{\sigma_i^2}\lambda_i u_{ij,s}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{\sqrt{NT}}) + o_p(\|\hat{\omega} - \omega\|)$. So the fourth term on the right hand side of (B.16) is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. The fifth term can be proved to be $O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{\sqrt{NT}}) + o_p(\|\hat{\omega} - \omega\|)$ similarly as (B.17). Summarizing all the results, we have

$$\frac{1}{N^2T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\sum_{s=1}^{T}(e_se_s' - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

This completes the proof of Lemma B.6. □

Using the results in Lemma B.6, we can strengthen Proposition B.3. The strengthened result is given in the following proposition.

Proposition B.5 *Under Assumptions A-H,*

(*a*) $\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda - H^{-1\prime} = O_p(N^{-1}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|)$;

(*b*) $HH' - I_r = O_p(N^{-1}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|)$;

(*c*) $H'H - I_r = O_p(N^{-1}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|)$.

Proof of Proposition B.5. The proof of (a) is similar as that of Proposition B.3, except that when dealing with $\frac{1}{NT}\sum_{t=1}^{T} f_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}$ and $\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\frac{1}{T}\sum_{t=1}^{T}[e_te_t' - \Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}$, we use the more sharper convergence rates in Lemma B.6.

Consider (b). Notice that

$$\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = \frac{1}{N}\Lambda'\Sigma_{ee}^{-1}\Lambda = I_r,$$

which is equivalent to

$$-\frac{1}{N}(\hat{\Lambda} - \Lambda H')'\hat{\Sigma}_{ee}^{-1}(\hat{\Lambda} - \Lambda H') + \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\hat{\Lambda} - \Lambda H')$$
$$+\frac{1}{N}(\hat{\Lambda} - \Lambda H')'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} + H\Big[\frac{1}{N}\Lambda'(\hat{\Sigma}_{ee}^{-1} - \Sigma_{ee}^{-1})\Lambda\Big]H' + HH' = I_r. \qquad \text{(B.18)}$$

The first term on the left hand side is bounded in norm by

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i - H\lambda_i\|^2 \le C\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2 = O_p(\frac{1}{N^2}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|^2)$$

by Proposition B.2, where the first inequality is due to the boundedness of $\hat{\sigma}_i^2$. The second and third term are of the same magnitude. So it suffices to investigate one of them. Consider the third term. Substituting (B.6) into it, we can rewrite the third term as

$$\hat{D}\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big]\Big[\frac{1}{NT}\sum_{t=1}^{T} f_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big] + \hat{D}\Big[\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_tf_t'\Big]\Big[\frac{1}{N}\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]$$
$$+\hat{D}\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\frac{1}{T}\sum_{t=1}^{T}[e_te_t' - \Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} + \hat{D}\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\hat{\lambda}_i\hat{\lambda}_i' - \frac{1}{N^2}\hat{D}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}\bar{e}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}$$
$$+\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}(\mathbb{T}_{i1} + \mathbb{T}_{i2} + \mathbb{T}_{i3})\hat{\lambda}_i',$$

where $\mathbb{T}_{i1}, \mathbb{T}_{i2}$ and $\mathbb{T}_{i3}$ are given in (B.7)-(B.9). The first and second terms of the preceding expression are both $O_p(\frac{1}{\sqrt{NT}}) + O_p(T^{-1}) + o_p(\|\hat{\omega} - \omega\|)$ by (B.5) and Lemma B.6(a). The third term is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Lemma B.6(b). The fourth term is bounded in norm by

$$\|\hat{D}\| \cdot \frac{1}{N^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \|\hat{\lambda}_i\|^2 \leq C\|\hat{D}\| \cdot \frac{1}{N^2} \sum_{i=1}^{N} \frac{1}{\hat{\sigma}_i^2} \|\hat{\lambda}_i\|^2 = \frac{1}{N} \|\hat{D}\| r = O_p(N^{-1}),$$

where the first inequality is due to $\sigma_i^{-2} \leq C\hat{\sigma}_i^{-2}$ by the boundedness of $\hat{\sigma}_i^2$ and σ_i^2. The fifth term is apparent to be $O_p(T^{-1})$. The last term is $O_p(\|\hat{\omega} - \omega\|)$ which is implicitly given in Lemma B.2. Given these results, we have that the third term of (B.18) is $O_p(N^{-1}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|)$. Consider the fourth term, which can be written as

$$-\frac{1}{N} \sum_{i=1}^{N} \frac{\hat{\sigma}_i^2 - \sigma_i^2}{\sigma_i^4} \lambda_i \lambda_i' + \frac{1}{N} \sum_{i=1}^{N} \frac{(\hat{\sigma}_i^2 - \sigma_i^2)^2}{\hat{\sigma}_i^2 \sigma_i^4} \lambda_i \lambda_i'.$$

The second term of the above expression is bounded in norm by $C\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2 = O_p(\frac{1}{N^2}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|^2)$ by Proposition B.4(b). The first term is $O_p(\frac{1}{\sqrt{NT}}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|)$, which can be proved similarly as Lemma S.12 of Bai and Li (2015) (see also the proof of Lemma E.1 below). Given all the results, we have proved (b). Given (b), pre-multiplying H^{-1} and post-multiplying H and noticing that H^{-1} and H are both $O_p(1)$, we have (c). This completes the proof. □

The following proposition, which can be viewed as the strengthened version of Propositions B.2 and B.4, are useful for the subsequent analysis.

Proposition B.6 *Under Assumptions A-H, we have*

$$(a)\quad \frac{1}{N} \sum_{i=1}^{N} \Big\|\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1} \sum_{t=1}^{T} f_t e_{it}\Big\|^2 = O_p(\frac{1}{N^2}) + O_p(T^{-2}) + O_p(\|\hat{\omega} - \omega\|^2),$$

$$(b)\quad \frac{1}{N} \sum_{i=1}^{N} \Big[\hat{\sigma}_i^2 - \sigma_i^2 - \frac{1}{T} \sum_{t=1}^{T} (e_{it}^2 - \sigma_i^2)\Big]^2 = O_p(\frac{1}{N^2}) + O_p(T^{-2}) + O_p(\|\hat{\omega} - \omega\|^2),$$

$$(c)\quad \frac{1}{T} \sum_{t=1}^{T} \Big\|\hat{f}_t - H^{-1\prime} f_t - \frac{1}{N} \hat{\Lambda}' \hat{\Sigma}_{ee}^{-1} \dot{e}_t\Big\|^2 = O_p(\frac{1}{N^2}) + O_p(T^{-2}) + O_p(\|\hat{\omega} - \omega\|^2).$$

Proof of Proposition B.6. *By the definition of H, we can rewrite (B.6) as*

$$\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1} \sum_{t=1}^{T} f_t e_{it} = \hat{D}\Big[\frac{1}{NT} \sum_{t=1}^{T} \hat{\Lambda}' \hat{\Sigma}_{ee}^{-1} e_t f_t'\Big] \lambda_i - \frac{1}{N} \hat{D} \hat{\Lambda}' \hat{\Sigma}_{ee}^{-1} \bar{e} \bar{e}_i$$

$$+ \hat{D} \frac{1}{NT} \sum_{t=1}^{T} \hat{\Lambda}' \hat{\Sigma}_{ee}^{-1} [e_t e_{it} - E(e_t e_{it})] + \frac{1}{N} \hat{D} \hat{\lambda}_i \frac{\hat{\sigma}_i^2}{\sigma_i^2} + \mathbb{T}_{i1} + \mathbb{T}_{i2} + \mathbb{T}_{i3}.$$

where $\mathbb{T}_{i1}, \mathbb{T}_{i2}$ and $\mathbb{T}_{i3}$ are defined in (B.7)-(B.9). Using the symbols in Proposition B.2, we use $I_{i1}, I_{i2}, \ldots, I_{i7}$ to denote the seven terms on right hand side. By the Cauchy-schwarz inequality, we have

$$\frac{1}{N} \sum_{i=1}^{N} \Big\|\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1} \sum_{t=1}^{T} f_t e_{it}\Big\|^2 \leq 7\frac{1}{N} \sum_{i=1}^{N} (\|I_{i1}\|^2 + \|I_{i2}\|^2 + \cdots + \|I_{i7}\|^2).$$

Consider the first term, which is bounded by

$$\|\hat{D}\|^2 \cdot \Big\|\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t f_t'\Big\|^2 \cdot \Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big] = O_p(\frac{1}{NT}) + O_p(T^{-2}) + o_p(\|\hat{\omega}-\omega\|^2)$$

by Lemma B.6(a). The second term is $O_p(T^{-2})$, which is shown in the proof of Proposition B.2. Consider the third term. Notice that

$$\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}[e_t e_{it} - E(e_t e_{it})] = \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_j^2}(\hat{\lambda}_j - H\lambda_j)[e_{jt}e_{it} - E(e_{jt}e_{it})]$$
$$-H\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_j^2-\sigma_j^2}{\hat{\sigma}_j^2\sigma_j^2}\lambda_j[e_{jt}e_{it} - E(e_{jt}e_{it})] + \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_j^2}\lambda_j[e_{jt}e_{it} - E(e_{jt}e_{it})].$$

So we have

$$\frac{1}{N}\sum_{i=1}^{N}\|I_{i3}\|^2 \le 3\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_j^2}(\hat{\lambda}_j - H\lambda_j)[e_{jt}e_{it} - E(e_{jt}e_{it})]\Big\|^2$$
$$+3\frac{1}{N}\sum_{i=1}^{N}\Big\|H\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_j^2-\sigma_j^2}{\hat{\sigma}_j^2\sigma_j^2}\lambda_j[e_{jt}e_{it} - E(e_{jt}e_{it})]\Big\|^2$$
$$+3\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_j^2}\lambda_j[e_{jt}e_{it} - E(e_{jt}e_{it})]\Big\|^2.$$

The first term is bounded by

$$C\Big[\frac{1}{N}\sum_{j=1}^{N}\|\hat{\lambda}_j - H\lambda_j\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}[e_{it}e_{jt} - E(e_{it}e_{jt})]\Big|^2\Big],$$

which is $O_p(\frac{1}{N^2}T^{-1}) + O_p(T^{-2}) + o_p(\|\hat{\omega}-\omega\|^2)$ by Proposition B.2. The second term is bounded by

$$C\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}[e_{it}e_{jt} - E(e_{it}e_{jt})]\Big|^2\Big],$$

which is $O_p(\frac{1}{N^2}T^{-1}) + O_p(T^{-2}) + o_p(\|\hat{\omega}-\omega\|^2)$ by Proposition B.4. The third term is $O_p(\frac{1}{NT})$. Given these results, we have

$$\frac{1}{N}\sum_{i=1}^{N}\|I_{i3}\|^2 = O_p(\frac{1}{NT}) + O_p(T^{-2}) + o_p(\|\hat{\omega}-\omega\|^2).$$

The fourth term is $O_p(\frac{1}{N^2})$ and the remaining three terms are $O_p(\|\hat{\omega}-\omega\|^2)$, which are shown in the proof of Proposition B.2. Summarizing all the results, we have (a).

The proof of result (b) is almost the same as that of Proposition B.4. The only difference is that when dealing with $\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t - H^{-1\prime}f_t\|^2$ and $\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2$, we use the convergence rates given in Proposition B.4.

Consider (c). By (B.11), we have

$$\hat{f}_t - H^{-1\prime} f_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t = -\Big[H^{-1\prime} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big] f_t - (\hat{\delta}-\delta)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}$$
$$-(\hat{\rho}-\rho)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta).$$

By the Cauchy-Schwarz inequality, we have

$$\frac{1}{T}\sum_{t=1}^{T}\Big\|\hat{f}_t - H^{-1\prime} f_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big\|^2$$
$$\leq 4\frac{1}{T}\sum_{t=1}^{T}\Big\|\Big[H^{-1\prime} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big] f_t\Big\|^2 + 4\frac{1}{T}\sum_{t=1}^{T}\Big\|(\hat{\delta}-\delta)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{Y}_{t-1}\Big\|^2$$
$$+4\frac{1}{T}\sum_{t=1}^{T}\Big\|(\hat{\rho}-\rho)\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\ddot{Y}_t\Big\|^2 + 4\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{X}_t(\hat{\beta}-\beta)\Big\|^2.$$

The first term is bounded by

$$4\Big(\frac{1}{T}\sum_{t=1}^{T}\|f_t\|^2\Big)\Big\|H^{-1\prime} - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big\|^2 = O_p(\frac{1}{N^2}) + O_p(T^{-2}) + O_p(\|\hat{\omega}-\omega\|^2)$$

by Proposition B.5. The second term is bounded by

$$C(\hat{\delta}-\delta)^2\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\dot{y}_{it-1}^2\Big] = O_p(\|\hat{\omega}-\omega\|^2).$$

The third and fourth terms are $O_p(\|\hat{\omega}-\omega\|^2)$, which can be proved similarly as the second term. Given the above results, we have (c).

This completes the proof of Proposition B.6. □

Appendix C: Analyzing the first order condition for β

In this section, we give a detailed analysis on the first order condition for β. We first derive some results, which will be used in the subsequent analysis. By (3.1), we have

$$Y_t = G_N^*\alpha^* + \delta^* G_N^* Y_{t-1} + G_N^* X_t\beta^* + G_N^*\Lambda^* f_t^* + G_N^* e_t.$$

with $G_N = (I_N - \rho^* W_N)^{-1}$, which implies

$$Y_t = \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^*\alpha^* + \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^* X_{t-l}\beta^* + \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^*\Lambda^* f_{t-l}^* + \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^* e_{t-l}.$$

Given the above result, we can rewrite $\dot{Y}_t = Y_t - T^{-1}\sum_{s=1}^{T} Y_s$ as

$$\dot{Y}_t = \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^*\dot{X}_{t-l}\beta^* + \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^*\Lambda^*\dot{f}_{t-l}^* + \sum_{l=0}^{\infty}(\delta^* G_N^*)^l G_N^*\dot{e}_{t-l}.$$

We can write $\ddot{Y}_t = W_N \dot{Y}_t$ as

$$\ddot{Y}_t = S_N^* \sum_{l=0}^{\infty} (\delta^* G_N^*)^l \dot{X}_{t-l} \beta^* + S_N^* \sum_{l=0}^{\infty} (\delta^* G_N^*)^l \dot{f}_{t-l}^* + S_N^* \sum_{l=1}^{\infty} (\delta^* G_N^*)^l \dot{e}_{t-l} + S_N^* \dot{e}_t.$$

Define the following notations (see also the main text):

$$B_t = \sum_{l=0}^{\infty} (\delta^* G_N^*)^l G_N^* \dot{X}_{t-l} \beta^* + \sum_{l=0}^{\infty} (\delta^* G_N^*)^l G_N^* \Lambda^* \dot{f}_{t-l}^*, \qquad \dot{B}_t = B_t - \frac{1}{T} \sum_{s=1}^{T} B_s$$

$$\ddot{B}_t = W_N \dot{B}_t, \qquad Q_t = \sum_{l=0}^{\infty} (\delta^* G_N^*)^l G_N^* e_{t-l}, \qquad J_t = S_N^* \sum_{l=1}^{\infty} (\delta^* G_N^*)^l e_{t-l}.$$

Given the above notation, we have

$$\dot{Y}_{t-1} = \dot{B}_{t-1} + \dot{Q}_{t-1}; \qquad \ddot{Y}_t = \ddot{B}_t + \dot{J}_t + S_N \dot{e}_t.$$

The following lemmas are useful for the subsequent analysis.

Lemma C.1 *Let $\mathbb{S}_{\beta 1}$ and $\mathbb{S}_{\beta 2}$ be defined in (C.12). Under Assumptions A-H,*

$$\mathbb{S}_{\beta 1} = O_p(\|\hat{\omega} - \omega\|^2), \qquad \mathbb{S}_{\beta 2} = o_p(\|\hat{\omega} - \omega\|).$$

The proof of Lemma C.1 is similar as that of Lemma B.2. See also the proof of Lemma C.1 of Bai and Li (2014a) for more details.

Lemma C.2 *Under Assumptions A-H,*

(a) $\displaystyle \frac{1}{NT} \sum_{t=1}^{T} \dot{X}_t' \widehat{\ddot{M}} \dot{Y}_{t-1} = \frac{1}{NT} \sum_{t=1}^{T} \dot{X}_t' \ddot{M} \dot{Y}_{t-1} + o_p(1);$

(b) $\displaystyle \frac{1}{NT} \sum_{t=1}^{T} \dot{X}_t' \widehat{\ddot{M}} \ddot{Y}_t = \frac{1}{NT} \sum_{t=1}^{T} \dot{X}_t' \ddot{M} \ddot{Y}_t + o_p(1);$

(c) $\displaystyle \frac{1}{NT} \sum_{t=1}^{T} \dot{X}_t' \widehat{\ddot{M}} \dot{X}_t = \frac{1}{NT} \sum_{t=1}^{T} \dot{X}_t' \ddot{M} \dot{X}_t + o_p(1);$

(d) $\displaystyle \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{X}_t' \widehat{\ddot{M}} \dot{Y}_{s-1} \pi_{st} = \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{X}_t' \ddot{M} \dot{Y}_{s-1} \pi_{st} + o_p(1);$

(e) $\displaystyle \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{X}_t' \widehat{\ddot{M}} \ddot{Y}_s \pi_{st} = \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{X}_t' \ddot{M} \ddot{Y}_s \pi_{st} + o_p(1);$

(f) $\displaystyle \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{X}_t' \widehat{\ddot{M}} \dot{X}_s \pi_{st} = \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{X}_t' \ddot{M} \dot{X}_s \pi_{st} + o_p(1).$

where $\pi_{st} = f_s' (F'F)^{-1} f_t$.

The proof of Lemma C.2 is similar as (actually easier than) that of Lemma C.3. The details are therefore omitted.

Lemma C.3 *Under Assumptions A-H,*

$$(a)\ \ \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\Lambda\frac{1}{NT}\sum_{s=1}^{T}f_se_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$= O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);$$

$$(b)\ \ \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{e}_t = \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\ddot{M}e_t - \Delta$$
$$+ O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);$$

$$(c)\ \ \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\ddot{M}e_s\pi_{st} - \Delta$$
$$+ O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);$$

$$(d)\ \ \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{e}_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$= O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

where $\Delta = \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\Sigma_{ee}^{-1}\Lambda(F'F)^{-1}f_t$.

Proof of Lemma C.3. Consider (a). The left hand side can be written as

$$\mathrm{tr}\Big[\Big(\frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_t'\widehat{\ddot{M}}\Lambda\Big)\Big(\frac{1}{NT}\sum_{s=1}^{T}f_se_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big)\hat{D}H^{-1\prime}\Big].$$

Consider the term $\frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_t'\widehat{\ddot{M}}\Lambda$, which is equal to

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\Lambda H' - \frac{1}{N^2T}\sum_{t=1}^{T}f_t\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\Lambda H'\Big]H^{-1\prime}$$
$$= -\frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_t'\hat{\Sigma}_{ee}^{-1}(\hat{\Lambda}-\Lambda H')H^{-1\prime} + \frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big[\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\hat{\Lambda}-\Lambda H')\Big]H^{-1\prime}.$$

The first term is bounded in norm by

$$C\|H^{-1\prime}\|\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{it}\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-1})+O_p(T^{-1/2})+O_p(\|\hat{\omega}-\omega\|)$ by Proposition B.2. The second term is bounded in norm by

$$C\|H^{-1\prime}\|\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{it}\Big\|^2\Big]^{1/2},$$

which is also $O_p(N^{-1}) + O_p(T^{-1/2}) + O_p(\|\hat{\omega} - \omega\|)$ by Proposition B.2. Given the above result, we have

$$\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_t' \widehat{\ddot{M}} \Lambda = O_p(N^{-1}) + O_p(T^{-1/2}) + O_p(\|\hat{\omega} - \omega\|). \tag{C.1}$$

Given (C.1), together with Lemma B.6(a), we obtain (a).

Consider (b). The left hand side is equivalent to

$$\frac{1}{NT}\sum_{t=1}^{T} \dot{X}_t' \hat{\Sigma}_{ee}^{-1} e_t - \frac{1}{N^2 T}\sum_{t=1}^{T} \dot{X}_t' \hat{\Sigma}_{ee}^{-1} \hat{\Lambda}\hat{\Lambda}' \hat{\Sigma}_{ee}^{-1} e_t = I_1 - I_2, \quad \text{say}$$

Consider I_1, which can be written as

$$I_1 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \frac{1}{\sigma_i^2} \dot{x}_{it} e_{it} - \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2 \sigma_i^2} \dot{x}_{it} e_{it} = I_3 - I_4, \quad \text{say.}$$

Term I_4 can be written as

$$I_4 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \frac{1}{\hat{\sigma}_i^2 \sigma_i^2}\Big[\hat{\sigma}_i^2 - \sigma_i^2 - \frac{1}{T}\sum_{s=1}^{T}(e_{is}^2 - \sigma_i^2)\Big]\dot{x}_{it} e_{it} + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T} \frac{1}{\hat{\sigma}_i^2 \sigma_i^2}\sum_{s=1}^{T}(e_{is}^2 - \sigma_i^2)\dot{x}_{it} e_{it}$$

By the boundedness of $\hat{\sigma}_i^2$ and σ_i^2 and the Cauchy-Schwarz inequality, the first term is bounded by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\hat{\sigma}_i^2 - \sigma_i^2 - \frac{1}{T}\sum_{s=1}^{T}(e_{is}^2 - \sigma_i^2)\Big|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it} e_{it}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.6(b). The second term can be further written as

$$\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^4}\dot{x}_{it} e_{it}(e_{is}^2 - \sigma_i^2) - \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2 \sigma_i^4}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{x}_{it} e_{it}(e_{is}^2 - \sigma_i^2). \tag{C.2}$$

The second term of (C.2) is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\Big(\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it} e_{it}\Big)\Big(\frac{1}{T}\sum_{s=1}^{T} e_{is}^2 - \sigma_i^2\Big)\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. Consider the first term of (C.2), which can be written as

$$\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^4}\dot{x}_{it}\Big[e_{it}(e_{is}^2 - \sigma_i^2) - E[e_{it}(e_{is}^2 - \sigma_i^2)]\Big] + \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^4}\dot{x}_{it}E(e_{it}^3).$$

The first term of the above expression is $O_p(\frac{1}{\sqrt{NT}})$ and the second term is 0 due to $\sum_{t=1}^{T}\dot{x}_{it} = 0$. So the first term of (C.2) is $O_p(\frac{1}{\sqrt{NT}})$. Given these results, we have

$$I_4 = O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Now consider I_2, which can be written as

$$I_2 = \frac{1}{N^2T}\sum_{t=1}^{T}\dot{X}_t'\hat{\Sigma}_{ee}^{-1}(\hat{\Lambda}-\Lambda H')\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t + \frac{1}{N^2T}\sum_{t=1}^{T}\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\Lambda H'(\hat{\Lambda}-\Lambda H')'\hat{\Sigma}_{ee}^{-1}e_t$$

$$+\frac{1}{N^2T}\sum_{t=1}^{T}\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\Lambda(H'H-I_r)\Lambda'\hat{\Sigma}_{ee}^{-1}e_t + \frac{1}{N^2T}\sum_{t=1}^{T}\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\Lambda\Lambda'(\hat{\Sigma}_{ee}^{-1}-\Sigma_{ee}^{-1})e_t$$

$$+\frac{1}{N^2T}\sum_{t=1}^{T}\dot{X}_t'(\hat{\Sigma}_{ee}^{-1}-\Sigma_{ee}^{-1})\Lambda\Lambda'\Sigma_{ee}^{-1}e_t + \frac{1}{N^2T}\sum_{t=1}^{T}\dot{X}_t'\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}e_t = I_5+\cdots+I_{10}, \quad \text{say}$$

First consider I_7, which is k-dimensional vector. Its pth element ($p=1,2\ldots,k$) is equal to

$$\operatorname{tr}\Big[(H'H-I_r)\Big(\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\lambda_i\lambda_j'\frac{1}{T}\sum_{t=1}^{T}e_{it}\dot{x}_{jtp}\Big)\Big],$$

where $\dot{x}_{jtp}$ is the pth element of $\dot{x}_{jt}$. The expression in the second bracket is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}e_{it}\dot{x}_{jtp}\Big|^2\Big]^{1/2} = O_p(T^{-1/2}).$$

This result, together with Proposition B.5(c), gives

$$I_7 = O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

Consider I_6. It is the pth element is equal to

$$\frac{1}{N^2T}\operatorname{tr}\Big[H'(\hat{\Lambda}-\Lambda H')'\hat{\Sigma}_{ee}^{-1}\sum_{t=1}^{T}e_t\dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}\Lambda\Big].$$

Ignore the trace operator, the expression can be written as

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\lambda_i'H'(\hat{\lambda}_j-H\lambda_j)\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{itp}e_{jt}$$

$$=\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\lambda_i'H'\Big[\hat{\lambda}_j-H\lambda_j-H(F'F)^{-1}\sum_{s=1}^{T}f_se_{js}\Big]\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{itp}e_{jt}$$

$$+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\lambda_i'H'H(F'F)^{-1}\sum_{s=1}^{T}f_se_{js}\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{itp}e_{jt} = I_{11}+I_{12}, \quad \text{say.}$$

Term I_{11} is bounded in norm by

$$C\Big\|H\Big\|\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{itp}e_{jt}\Big|^2\Big]^{1/2}$$

$$\times\Big[\frac{1}{N}\sum_{j=1}^{N}\Big\|\hat{\lambda}_j-H\lambda_j-H(F'F)^{-1}\sum_{s=1}^{T}f_se_{js}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.6(a). Term I_{12} can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\lambda_i' H' H\Big[\frac{1}{T}F'F\Big]^{-1}\Big[\frac{1}{NT^2}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_j^2}f_s\dot{x}_{itp}[e_{jt}e_{js} - E(e_{jt}e_{js})]\Big]$$

$$-\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^2}\lambda_i' H' H\Big[\frac{1}{T}F'F\Big]^{-1}\Big[\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}f_s\dot{x}_{itp}[e_{jt}e_{js} - E(e_{jt}e_{js})]\Big]$$

$$+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\lambda_i' H' H(F'F)^{-1}\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{itp}.$$

The first expression is bounded in norm by

$$C\Big\|H'H[\frac{1}{T}F'F]^{-1}\Big\|\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT^2}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_j^2}f_s\dot{x}_{itp}[e_{jt}e_{js} - E(e_{jt}e_{js})]\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{\sqrt{NT}})$. The second expression is bounded in norm by

$$C\Big\|H'H[\frac{1}{T}F'F]^{-1}\Big\|\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2 - \sigma_j^2)^2\Big]^{1/2}$$

$$\times\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big\|\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}f_s\dot{x}_{itp}[e_{jt}e_{js} - E(e_{jt}e_{js})]\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.2. The last expression can be written as

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\lambda_i'(H'H - I_r)(F'F)^{-1}\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{itp} - \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\lambda_i'(F'F)^{-1}\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{itp}$$

$$-\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i'(F'F)^{-1}\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{itp} + \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i'(F'F)^{-1}\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{itp}.$$

The first term of the above expression is $O_p(\frac{1}{NT}) + O_p(T^{-2}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.5. The second term is bounded in norm by

$$\Big\|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\lambda_i'(F'F)^{-1}\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{itp}\Big\|\cdot\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2 - \sigma_j^2)^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.4. The third term is bounded in norm by

$$C\frac{1}{T}\Big\|(\frac{1}{T}F'F)^{-1}\Big\|\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big(\frac{1}{T}\sum_{t=1}^{T}\|f_t\|\cdot\|\dot{x}_{itp}\|\Big)^2\Big]^{1/2},$$

which is also $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.4. The last term is

$$\frac{1}{NT}\sum_{t=1}^{T} \dot{X}_t' \Sigma_{ee}^{-1} \Lambda (F'F)^{-1} f_t \triangleq \Delta.$$

Summarizing all the results, we have

$$I_6 = \Delta + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Consider I_5.

$$\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} \frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}(\hat{\lambda}_i - H\lambda_i)'(\hat{\lambda}_j - H\lambda_j)\dot{x}_{it}e_{jt}$$

$$-\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} \frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^2}(\hat{\lambda}_i - H\lambda_i)'H\lambda_j\dot{x}_{it}e_{jt}$$

$$+\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} \frac{1}{\hat{\sigma}_i^2\sigma_j^2}\dot{x}_{it}e_{jt}\lambda_j'H'(\hat{\lambda}_i - H\lambda_i).$$

The first expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}\|\hat{\lambda}_j - H\lambda_j\|^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}e_{jt}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. The second expression is bounded in norm by

$$C\|H\| \cdot \Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2 - \sigma_j^2)^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}e_{jt}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Propositions B.2 and B.4. The third expression is bounded in norm by

$$C\Big\|H\Big\| \cdot \Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_j^2}\dot{x}_{it}\lambda_j'e_{jt}\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-3/2}T^{-1/2}) + O_p(\frac{1}{\sqrt{NT}}) + o_p(\|\hat{\omega} - \omega\|)$. So we have

$$I_5 = O_p(\frac{1}{\sqrt{NT}}) + O_p(N^{-3/2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Consider I_8, which can be written as

$$-\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^2}\Big[\hat{\sigma}_j^2 - \sigma_j^2 - \frac{1}{T}\sum_{s=1}^{T}(e_{js}^2 - \sigma_j^2)\Big]\lambda_i'\lambda_j\dot{x}_{it}e_{jt}$$

$$+\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^4}\Big[\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{x}_{it}e_{jt}(e_{js}^2 - \sigma_j^2)\lambda_j'\Big]\lambda_i$$

$$-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\Big[\frac{1}{NT^2}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_j^4}\dot{x}_{it}e_{jt}(e_{js}^2-\sigma_j^2)\lambda_j'\Big]\lambda_i.$$

The first expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}\Big\|\hat{\sigma}_j^2-\sigma_j^2-\frac{1}{T}\sum_{s=1}^{T}(e_{js}^2-\sigma_j^2)\Big\|^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}\dot{x}_{it}e_{jt}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. The second expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big\|\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{x}_{it}e_{jt}(e_{js}^2-\sigma_j^2)\lambda_j'\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. The third expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT^2}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_j^4}\dot{x}_{it}e_{jt}(e_{js}^2-\sigma_j^2)\lambda_j'\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{\sqrt{NT}})$. So we have

$$I_8=O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

Consider I_9, which is equivalent to

$$-\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i\Big[\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_j^2}\dot{x}_{it}e_{jt}\lambda_j'\Big].$$

The above expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_j^2}\dot{x}_{it}e_{jt}\lambda_j'\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{\sqrt{NT}})+O_p(N^{-3/2}T^{-1/2})+o_p(\|\hat{\omega}-\omega\|)$. Summarizing all the results, we have (b).

Consider (c). Treating $\sum_{s=1}^{T}e_s\pi_{st}=\sum e_sf_s'(F'F)^{-1}f_t$ as a new e_t, the proof of (c) is similar as that of (b). The details are therefore omitted.

Consider (d). The left hand side of (d) is a k-dimensional vector. It suffices to consider its pth element.

$$\mathrm{tr}\Big[\frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_{tp}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_se_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big]-\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T}f_t\dot{X}_{tp}'\widehat{\ddot{M}}\bar{e}\bar{e}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}H^{-1\prime}\Big].$$

where $\dot{X}_{tp}=(\dot{x}_{1tp},\dot{x}_{2tp},\dots,x_{Ntp})'$ is an N-dimensional vector. We use III_1 and III_2 to denote the above two terms. We first consider the second term. Let $\widetilde{e}_i=T^{-1/2}\sum_{t=1}^{T}e_{it}$ and $\widetilde{e}=(\widetilde{e}_1,\widetilde{e}_2,\dots,\widetilde{e}_N)'$. The expression in the trace operator can be written as

$$III_2=\frac{1}{N^2T^2}\sum_{t=1}^{T}f_t\dot{X}_{tp}'\widehat{\ddot{M}}[\widetilde{e}\widetilde{e}'-\Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}H^{-1\prime}+\frac{1}{N^2T^2}\sum_{t=1}^{T}f_t\dot{X}_{tp}'\widehat{\ddot{M}}\hat{\Lambda}H^{-1\prime}$$

$$-\frac{1}{N^2T^2}\sum_{t=1}^{T} f_t\dot{X}_{tp}'\widehat{\ddot{M}}(\hat{\Sigma}_{ee}-\Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}H^{-1\prime} = III_3 + III_4 - III_5, \qquad \text{say}$$

Term $III_4 = 0$ by $\widehat{\ddot{M}}\hat{\Lambda} = 0$. Consider III_3, which is equivalent to

$$\begin{aligned} &\frac{1}{N^2T^2}\sum_{t=1}^{T} f_t\dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}[\widetilde{e}\widetilde{e}'-\Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}H^{-1\prime} \\ &-\frac{1}{N^3T^2}\sum_{t=1}^{T} f_t\dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}[\widetilde{e}\widetilde{e}'-\Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}H^{-1\prime}. \end{aligned} \tag{C.3}$$

The first term (ignore $H^{-1\prime}$) is

$$\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\Big(\frac{1}{T}\sum_{t=1}^{T} f_t\dot{x}_{itp}\Big)\hat{\lambda}_j[\widetilde{e}_i\widetilde{e}_j - E(\widetilde{e}_i\widetilde{e}_j)]$$

which can be written as

$$\begin{aligned} &\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\Big(\frac{1}{T}\sum_{t=1}^{T} f_t\dot{x}_{itp}\Big)(\hat{\lambda}_j - H\lambda_j)'[\widetilde{e}_i\widetilde{e}_j - E(\widetilde{e}_i\widetilde{e}_j)] \\ &-\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2-\sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^2}\Big(\frac{1}{T}\sum_{t=1}^{T} f_t\dot{x}_{itp}\Big)\lambda_j'[\widetilde{e}_i\widetilde{e}_j - E(\widetilde{e}_i\widetilde{e}_j)]H' \\ &+\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\sigma_j^2}\Big(\frac{1}{T}\sum_{t=1}^{T} f_t\dot{x}_{itp}\Big)\lambda_j'[\widetilde{e}_i\widetilde{e}_j - E(\widetilde{e}_i\widetilde{e}_j]H \end{aligned}$$

The first term of the above expression is bounded in norm by

$$C\frac{1}{T}\Big[\frac{1}{N}\sum_{j=1}^{N}\|\hat{\lambda}_j - H\lambda_j\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t\dot{x}_{itp}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\widetilde{e}_i\widetilde{e}_j - E(\widetilde{e}_i\widetilde{e}_j)\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.2(a). The second term is bounded in norm by

$$C\|H\|\frac{1}{T}\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t\dot{x}_{itp}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\widetilde{e}_i\widetilde{e}_j - E(\widetilde{e}_i\widetilde{e}_j)\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.4(b). The third term is bounded in norm by

$$C\|H\|\frac{1}{T}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t\dot{x}_{itp}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\sigma_j^2}\lambda_j[\widetilde{e}_i\widetilde{e}_j - E(\widetilde{e}_i\widetilde{e}_j)]\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{\sqrt{N}T})$. Given the above results, we have that the first term of (C.3) is $O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|)$. For the second term of (C.3), first notice that $\frac{1}{NT}\sum_{t=1}^{T} f_t\dot{X}_{tp}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = O_p(1)$. Therefore we only need to consider $\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}[\widetilde{e}\widetilde{e}'-\Sigma_{ee}]\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}$, which, using the arguments in the proof of the first term, is also $O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. Given these results, we have

$$III_3 = O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Consider III_5, which is equal to

$$\begin{aligned} &\frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^4} f_t \dot{x}_{itp} \hat{\lambda}_i' H^{-1\prime} \\ -&\frac{1}{N^3T^2}\sum_{t=1}^{T} f_t \dot{X}_{tp} \hat{\Sigma}_{ee}^{-1} \hat{\Lambda}\hat{\Lambda}' \hat{\Sigma}_{ee}^{-1} (\hat{\Sigma}_{ee} - \Sigma_{ee}) \hat{\Sigma}_{ee}^{-1} \hat{\Lambda} H^{-1\prime}. \end{aligned} \tag{C.4}$$

The first term of (C.4) (ignore $H^{-1\prime}$) can be written as

$$\frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^4} f_t \dot{x}_{itp} (\hat{\lambda}_i - H\lambda_i)' + \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^4} f_t \dot{x}_{itp} \lambda_i' H'$$

By the boundedness of $\hat{\sigma}_i^2$ and σ_i^2, term $|\hat{\sigma}_i^{-4}(\hat{\sigma}_i^2 - \sigma_i^2)|$ is uniformly bounded by some constant C. Given this result, the first term is bounded in norm by

$$C\frac{1}{NT}\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i - H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t \dot{x}_{itp}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2}T^{-1}) + O_p(N^{-1}\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. By the boundedness of λ_i and $\hat{\sigma}_i^2$, the second term is bounded in norm by

$$C\frac{1}{NT}\|H\|\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t \dot{x}_{itp}\Big\|^2\Big]^{1/2},$$

which is also $O_p(\frac{1}{N^2}T^{-1}) + O_p(N^{-1}\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. Given these two results, we have the first term of (C.4) is $O_p(\frac{1}{N^2}T^{-1}) + O_p(N^{-1}\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. Consider the second term. Notice that by the boundedness of $\hat{\sigma}_i^2$ and σ_i^2, there exists a constant C large enough such that $C \cdot I_N - \hat{\Sigma}_{ee}^{-1}(\hat{\Sigma}_{ee} - \Sigma_{ee})$ is positive definite. The second term of (C.4) is bounded in norm by

$$\begin{aligned} &\frac{1}{NT}\Big\|\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_{tp}' \hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big\| \cdot \frac{1}{N}\sum_{i=1}^{N}\Big|\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^4}\Big|\|\hat{\lambda}_i\|^2 \\ \le C&\frac{1}{NT}\Big\|\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_{tp}' \hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big\| \cdot \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2 = Cr\frac{1}{NT}\Big\|\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_{tp}' \hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big\| \end{aligned}$$

which is $O_p(\frac{1}{NT})$. Summarizing the above results, we have

$$III_5 = O_p(\frac{1}{NT}) + o_p(\|\hat{\omega} - \omega\|).$$

The results on III_3 and III_5 implies that

$$III_2 = O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Now consider III_1, which can be written as

$$\operatorname{tr}\Big[\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}(e_s e_s' - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big] + \operatorname{tr}\Big[\frac{1}{N^2 T}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\widehat{\ddot{M}}\Sigma_{ee}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big]. \tag{C.5}$$

Consider the first term of (C.5). The expression in the trace operator (ignore $\hat{D}H^{-1\prime}$) is equal to

$$\frac{1}{N^2T^2}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}\sum_{s=1}^{T}(e_s e_s' - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} - \Big[\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]\Big[\frac{1}{N^2T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\sum_{s=1}^{T}(e_s e_s' - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]. \tag{C.6}$$

By Lemma B.6(b), together with $\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} = O_p(1)$, the second term of (C.6) is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. As for the first term. Since $\dot{X}_{tp}$ is exogenous, replacing Λ' by $\frac{1}{T}\sum_{t=1}^{T} f_t \dot{X}_{tp}'$, the proof of the first term being $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ is almost the same as that of the second one. So we have

$$\operatorname{tr}\Big[\frac{1}{NT}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}(e_s e_s' - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big] = O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|). \tag{C.7}$$

Consider the second term of (C.5), which can be written as

$$\operatorname{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\widehat{\ddot{M}}\hat{\Lambda}\hat{D}H^{-1\prime}\Big] - \operatorname{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\widehat{\ddot{M}}(\hat{\Sigma}_{ee} - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big]$$

The first term is 0 by $\widehat{\ddot{M}}\hat{\Lambda} = 0$. For the second term, the expression in the trace operator (ignore $\hat{D}H^{-1\prime}$) is equivalent to

$$\frac{1}{N^2T}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}(\hat{\Sigma}_{ee} - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda} - \frac{1}{N^3T}\sum_{t=1}^{T} f_t \dot{X}_{tp}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\hat{\Sigma}_{ee} - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}. \tag{C.8}$$

The first term of (C.8) is equal to

$$\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^4} f_t \dot{x}_{itp}\hat{\lambda}_i' = \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^4} f_t \dot{x}_{itp}\lambda_i' H' + \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^4} f_t \dot{x}_{itp}(\hat{\lambda}_i - H\lambda_i)'.$$

The first term on right hand side is bounded in norm by

$$C\frac{1}{N}\|H\|\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T} f_t \dot{x}_{itp}\Big\|^2\Big]^{1/2} = O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$$

by Proposition B.4(b). The second term is bounded in norm by

$$C\frac{1}{N}\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_t\dot{x}_{itp}\Big\|^2\Big]^{1/2}=O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$$

by Proposition B.2. So the first term of (C.8) is $O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. Further consider the second term of (C.8), which is equal to

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_{tp}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]\Big[\frac{1}{N^2}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^4}\hat{\lambda}_i\hat{\lambda}_i'\Big].$$

Notice that

$$\frac{1}{N^2}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^4}\hat{\lambda}_i\hat{\lambda}_i'=\frac{1}{N^2}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^4}\hat{\lambda}_i(\hat{\lambda}_i-H\lambda_i)'+\frac{1}{N^2}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^4}\hat{\lambda}_i\lambda_i'H'.$$

By the boundedness of $\hat{\sigma}_i^2$ and σ_i^2, we have $|\hat{\sigma}_i^{-4}(\hat{\sigma}_i^2-\sigma_i^2)|^2\le C\hat{\sigma}_i^2$ for some large constant C. Given this result, the first term on the right hand side is bounded in norm by

$$C\frac{1}{N}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^2\Big]^{1/2}=O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$$

by Proposition B.2. The second term is bounded in norm by

$$C\frac{1}{N}\|H\|\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]^{1/2}=O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$$

by Proposition B.4. Given the above results, together with $\frac{1}{NT}\sum_{t=1}^{T}f_t\dot{X}_{tp}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}=O_p(1)$, we have that the second term of (C.8) is $O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. Summarizing the above results, we have

$$\operatorname{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T}f_t\dot{X}_{tp}'\widehat{\ddot{M}}\Sigma_{ee}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big]=O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).\qquad\text{(C.9)}$$

Given (C.7) and (C.9), together with (C.5), we have

$$III_1=O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

The above result, together with the result on III_2, gives (d). □

Analyzing the first order condition for β. The first order condition for β is

$$\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\dot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\hat{\rho}\ddot{Y}_t-\dot{X}_t\hat{\beta})=0.$$

By $\dot{Y}_t=\delta\dot{Y}_{t-1}+\rho\ddot{Y}_t+\dot{X}_t\beta+\Lambda f_t+\dot{e}_t$, we can rewrite the above equation as

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{Y}_{t-1}\Big](\hat{\delta}-\delta)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\ddot{Y}_t\Big](\hat{\rho}-\rho)$$

$$+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{X}_t\Big](\hat{\beta}-\beta)=\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\Lambda f_t+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{e}_t.$$

By $\widehat{\ddot{M}}\hat{\Lambda}=0$, the above equation can be further written as

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{Y}_{t-1}\Big](\hat{\delta}-\delta)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\ddot{Y}_t\Big](\hat{\rho}-\rho) \tag{C.10}$$
$$+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{X}_t\Big](\hat{\beta}-\beta)=-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\hat{\Lambda}-\Lambda H')H^{-1\prime}f_t+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{e}_t.$$

However, the first order condition for Λ gives (which can also be derived from (B.2))

$$\hat{\Lambda}-\Lambda H'=\frac{1}{NT}\sum_{t=1}^{T}e_tf_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+\Lambda\frac{1}{NT}\sum_{t=1}^{T}f_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+\frac{1}{NT}\sum_{t=1}^{T}\dot{e}_t\dot{e}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$+(\hat{\delta}-\delta)^2\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+(\hat{\rho}-\rho)^2\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D} \tag{C.11}$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$+(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t(\hat{\beta}-\beta)(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}f_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_tf_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{NT}\sum_{t=1}^{T}X_t(\hat{\beta}-\beta)f_t'\Lambda'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-\frac{1}{NT}\Lambda\sum_{t=1}^{T}f_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{NT}\Lambda\sum_{t=1}^{T}f_t(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-\frac{1}{NT}\Lambda\sum_{t=1}^{T}f_t(\hat{\beta}-\beta)'\dot{X}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}e_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_te_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{NT}\sum_{t=1}^{T}X_t(\hat{\beta}-\beta)e_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-\frac{1}{NT}\sum_{t=1}^{T}e_t(\hat{\delta}-\delta)\dot{Y}_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}-\frac{1}{NT}\sum_{t=1}^{T}e_t(\hat{\rho}-\rho)\ddot{Y}_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$
$$-\frac{1}{NT}\sum_{t=1}^{T}e_t(\hat{\beta}-\beta)'X_t'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$

Substituting (C.11) into (C.10), we have

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{Y}_{t-1}\Big](\hat{\delta}-\delta)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\ddot{Y}_t\Big](\hat{\rho}-\rho)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{X}_t\Big](\hat{\beta}-\beta)$$

$$
\begin{aligned}
&= \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}e_t - \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\Lambda\frac{1}{NT}\sum_{s=1}^{T}f_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t \qquad \text{(C.12)}\\
&-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st} - \frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{e}_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&+(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{Y}_{s-1}\pi_{st} + (\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\ddot{Y}_s\pi_{st}\\
&+\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\dot{X}_s(\hat{\beta}-\beta)\pi_{st} + \mathbb{S}_{\beta 1} + \mathbb{S}_{\beta 2}.
\end{aligned}
$$

where $\pi_{st} = f_s'(F'F)^{-1}f_t$, and

$$
\begin{aligned}
\mathbb{S}_{\beta 1} = &-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)^2\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\hat{\rho}-\rho)^2\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{S}_{\beta 2} = &\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t
\end{aligned}
$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}X_s(\hat{\beta}-\beta)e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\beta}-\beta)'X_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

Using the results in Lemmas C.1, C.2 and C.3, we have

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\ddot{M}\dot{Y}_{t-1}-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\ddot{M}\dot{Y}_{s-1}\pi_{st}\Big](\hat{\delta}-\delta)$$
$$+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\ddot{M}\ddot{Y}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\ddot{M}\ddot{Y}_s\pi_{st}\Big](\hat{\rho}-\rho)$$
$$+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\ddot{M}\dot{X}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\ddot{M}\dot{X}_s\pi_{st}\Big](\hat{\beta}-\beta)$$
$$=\frac{1}{NT}\sum_{t=1}^{T}\dot{X}_t'\ddot{M}e_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\ddot{M}e_s\pi_{st} \tag{C.13}$$
$$+O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

This completes the whole derivation. □

Appendix D: Analyzing the first order condition for δ

In this section, we give a detailed analysis on the first order condition for δ. The following lemmas are useful for the subsequent analysis.

Lemma D.1 *Let $\mathbb{S}_{\delta 1}$ and $\mathbb{S}_{\delta 2}$ be defined in (D.6) below. Under Assumptions A-H,*

$$\mathbb{S}_{\delta 1}=O_p(\|\hat{\omega}-\omega\|^2); \qquad \mathbb{S}_{\delta 2}=o_p(\|\hat{\omega}-\omega\|).$$

The proof of Lemma D.1 is similar as that of Lemma B.2. The details are omitted.

Lemma D.2 *Under Assumptions A-H,*

$$(a)\ \ \frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{Y}_{t-1}=\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\ddot{M}\dot{Y}_{t-1}+o_p(1);$$
$$(b)\ \ \frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\ddot{Y}_t=\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\ddot{M}\ddot{Y}_t+o_p(1);$$

(c) $\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{Y}_{s-1}\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\ddot{M}\dot{Y}_{s-1}\pi_{st} + o_p(1);$

(d) $\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\ddot{Y}_{s}\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\ddot{M}\ddot{Y}_{s}\pi_{st} + o_p(1).$

The proof of this lemma is actually easier than that of Lemma D.3 below. The details are therefore omitted.

Lemma D.3 *Under Assumptions A-H,*

$$(a)\ \frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\Lambda\frac{1}{NT}\sum_{s=1}^{T}f_s e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$= O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);$$

$$(b)\ \frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{e}_t = \frac{1}{NT}\sum_{t=1}^{T}\dot{B}_{t-1}'\ddot{M}e_t + \frac{1}{NT}\sum_{t=1}^{T}Q_{t-1}'\Sigma_{ee}^{-1}e_t - \Delta^{\star} - \frac{1}{NT}(\mathbf{1}_T'\mathbf{1}_T)^{-1}\mathbf{1}_T'L\mathbf{1}_T$$
$$+ O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);$$

$$(c)\ \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{e}_s\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{B}_{t-1}'\ddot{M}e_s\pi_{st} - \Delta^{\star} + \frac{1}{NT}\mathrm{tr}\Big[(F'F)^{-1}F'LF\Big]$$
$$+ O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);$$

$$(d)\ \frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{e}_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$= O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

with

$$\Delta^{\star} = \frac{1}{NT}\sum_{t=1}^{T}\dot{B}_{t-1}'\Sigma_{ee}^{-1}\Lambda(F'F)^{-1}f_t$$

and L is defined in Theorem 5.2.

Proof of Lemma D.3. The proof of result (a) is similar as that of Lemma C.3(a). The details are omitted.

Consider (b). The left hand side is equal to $\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}e_t$, which, by $\dot{Y}_{t-1} = \dot{B}_{t-1} + \dot{Q}_{t-1}$, can be further written as

$$\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}e_t = \frac{1}{NT}\sum_{t=1}^{T}\dot{B}_{t-1}'\widehat{\ddot{M}}e_t + \frac{1}{NT}\sum_{t=1}^{T}\dot{Q}_{t-1}'\widehat{\ddot{M}}e_t = I_1 + I_2, \quad \text{say.}$$

Notice that $\dot{B}_{t-1}$ is exogenous, the derivation for the first term is therefore almost the same as that of Lemma C.3(b). So we have

$$I_1 = \frac{1}{NT}\sum_{t=1}^{T}\dot{B}_{t-1}'\ddot{M}e_t - \Delta^{\star} + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

Consider the second term, which can be written as

$$\frac{1}{NT}\sum_{t=1}^{T}Q_{t-1}'\hat{\Sigma}_{ee}^{-1}e_t-\frac{1}{N}\bar{Q}_{-1}'\hat{\Sigma}_{ee}^{-1}\bar{e}-\frac{1}{N^2T}\sum_{t=1}^{T}Q_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t+\frac{1}{N^2}\bar{Q}_{-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}.$$

Ignore the signs of the above four terms, we use I_3, I_4, I_5 and I_6 to denote them. Consider I_3, which can be written as

$$\frac{1}{NT}\sum_{t=1}^{T}Q_{t-1}'\Sigma_{ee}^{-1}e_t-\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\frac{1}{T}\sum_{t=1}^{T}Q_{it-1}e_{it}.$$

The second term of the above expression can be written as

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2\sigma_i^2}\Big[\hat{\sigma}_i^2-\sigma_i^2-\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]\frac{1}{T}\sum_{t=1}^{T}Q_{it-1}e_{it}\\&\quad+\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]\frac{1}{T}\sum_{t=1}^{T}Q_{it-1}e_{it}\\&\quad-\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]\frac{1}{T}\sum_{t=1}^{T}Q_{it-1}e_{it}.\end{aligned}\tag{D.1}$$

The first term of (D.1) is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\hat{\sigma}_i^2-\sigma_i^2-\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T}\sum_{t=1}^{T}Q_{it-1}e_{it}\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.6. The second term is $O_p(\frac{1}{\sqrt{N}T})$. The third term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\frac{1}{T}\sum_{t=1}^{T}Q_{it-1}e_{it}\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.4. Summarizing all the results, we have

$$I_3=\frac{1}{NT}\sum_{t=1}^{T}Q_{t-1}'\Sigma_{ee}^{-1}e_t+O_p(\frac{1}{\sqrt{N}T})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

Consider I_4, which is equivalent to $\frac{1}{NT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}Q_{t-1}'\hat{\Sigma}_{ee}^{-1}e_s$. This term can be further written as

$$\begin{aligned}&\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}[Q_{it-1}e_{is}-E(Q_{it-1}e_{is})]+\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}E(Q_{it-1}e_{is})\\&-\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}[Q_{it-1}e_{is}-E(Q_{it-1}e_{is})]-\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}E(Q_{it-1}e_{is}).\end{aligned}\tag{D.2}$$

The first term of (D.2) is $O_p(\frac{1}{\sqrt{NT}})$. The second term of (D.2) is

$$\frac{1}{NT^2}\sum_{t=1}^{T}\Big[\sum_{s=0}^{t-2}\text{tr}[G_N(\delta G_N)^s]\Big] = \frac{1}{NT}(\mathbf{1}_T'\mathbf{1}_T)^{-1}\mathbf{1}_T'L\mathbf{1}_T$$

with L defined in Theorem 5.2. The third term of (D.2) is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}[Q_{it-1}e_{is}-E(Q_{it-1}e_{is})]\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. The last term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}E(Q_{it-1}e_{is})\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. Given the above result, we have

$$I_4 = \frac{1}{NT}(\mathbf{1}_T'\mathbf{1}_T)^{-1}\mathbf{1}_T'L\mathbf{1}_T + O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

Consider I_5, which can be written as $\text{tr}[\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\frac{1}{T}\sum_{t=1}^{T}e_tQ_{t-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}]$. Using the argument in proving Lemma B.6(b), we can show that the above term is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. Consider I_6. We first show that

$$\begin{aligned}
&\text{(i)}\ \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}(\hat{\lambda}_i-H\lambda_i)\bar{e}_i = O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|);\\
&\text{(ii)}\ \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i\bar{e}_i = O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|); \qquad \text{(D.3)}\\
&\text{(iii)}\ \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i\bar{e}_i = O_p(\frac{1}{\sqrt{NT}}).
\end{aligned}$$

First consider result (i) of (D.3), which can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\Big[\hat{\lambda}_i-H\lambda_i-H(F'F)^{-1}\sum_{t=1}^{T}f_te_{it}\Big]\bar{e}_i + H\Big(\frac{1}{T}F'F\Big)^{-1}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^2}f_te_{it}\bar{e}_i$$
$$-H\Big(\frac{1}{T}F'F\Big)^{-1}\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\Big(\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\Big)\bar{e}_i$$

The first term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\hat{\lambda}_i-H\lambda_i-H(F'F)^{-1}\sum_{t=1}^{T}f_te_{it}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\bar{e}_i^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.6. By $\sum_{t=1}^{T}f_t=0$, the second term is equivalent to

$$H\Big(\frac{1}{T}F'F\Big)^{-1}\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}f_t[e_{it}e_{is}-E(e_{it}e_{is})] = O_p(\frac{1}{\sqrt{NT}}).$$

The third term is bounded in norm by

$$C\|H\| \cdot \Big\| \Big(\frac{1}{T}F'F\Big)^{-1} \Big\| \cdot \Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\Big(\frac{1}{T}\sum_{t=1}^{T}f_t e_{it}\Big)\bar{e}_i\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.2. Given the above result, we obtain result (i) of (D.3). Consider result (ii) of (D.3). The left hand side can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2\sigma_i^2}\Big[\hat{\sigma}_i^2 - \sigma_i^2 - \frac{1}{T}\sum_{s=1}^{T}(e_{is}^2 - \sigma_i^2)\Big]\lambda_i\bar{e}_i + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^4}(e_{it}^2 - \sigma_i^2)\lambda_i\bar{e}_i$$
$$-\frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2 - \sigma_i^2)\lambda_i\bar{e}_i.$$

The first term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\hat{\sigma}_i^2 - \sigma_i^2 - \frac{1}{T}\sum_{s=1}^{T}(e_{is}^2 - \sigma_i^2)\Big|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\|\lambda_i\bar{e}_i\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.6. The second term can be written as

$$\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^4}\lambda_i(e_{it}^2 - \sigma_i^2)e_{is} = \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\Big[\frac{1}{\sigma_i^4}\lambda_i(e_{it}^2 - \sigma_i^2)e_{is} - \frac{1}{\sigma_i^4}\lambda_i E[(e_{it}^2 - \sigma_i^2)e_{is}]\Big]$$
$$+ \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^4}\lambda_i E(e_{it}^3).$$

which is $O_p(T^{-1})$. The third term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2 - \sigma_i^2)\lambda_i\bar{e}_i\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.4. Summarizing the above results, we have (ii). Result (iii) is apparent. Notice that

$$\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}\hat{\lambda}_j\bar{e}_j = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}(\hat{\lambda}_j - H\lambda_j)\bar{e}_j - H\frac{1}{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_j^2\sigma_j^2}\lambda_j\bar{e}_j + H\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\sigma_j^2}\lambda_j\bar{e}_j.$$

Using the results in (D.3), we have

$$\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}\hat{\lambda}_j\bar{e}_j = O_p(\frac{1}{\sqrt{NT}}) + O_p(T^{-1}) + o_p(\|\hat{\omega} - \omega\|). \tag{D.4}$$

Similarly, we can show that

$$\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}\hat{\lambda}_j\bar{Q}_{j,-1} = O_p(\frac{1}{\sqrt{NT}}) + O_p(T^{-1}) + o_p(\|\hat{\omega} - \omega\|).$$

The above two results implies $I_6 = O_p(\frac{1}{NT}) + O_p(T^{-3}) + o_p(\|\hat{\omega} - \omega\|)$. Summarizing all the results on $I_3, \ldots, I_6$, we have

$$I_2 = \frac{1}{NT}\sum_{t=1}^{T} Q'_{t-1}\Sigma_{ee}^{-1}e_t - \xi_1 + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Given the results on I_1 and I_2, we have (b).

Consider (c). By $\sum_{t=1}^{T} f_t = 0$, the left hand side of (c) is equivalent to

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T} Y'_{t-1}\widehat{\ddot{M}} e_s \pi_{st}.$$

Treating $\sum_{s=1}^{T} e_s \pi_{st}$ as v_t, the analysis of (c) is very similar as that of result (b). The detailed proof is therefore omitted.

Consider (d). By $\dot{Y}_{t-1} = \dot{B}_{t-1} + \dot{Q}_{t-1}$, the left hand side is equal to

$$\frac{1}{NT}\sum_{t=1}^{T} \dot{B}'_{t-1}\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T} \dot{e}_s\dot{e}'_s\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t + \frac{1}{NT}\sum_{t=1}^{T} \dot{Q}'_{t-1}\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T} \dot{e}_s\dot{e}'_s\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t.$$

We use II_1 and II_2 to denote the above two expressions. Since $\dot{B}_{t-1}$ is exogenous, the derivation on II_1 is almost the same as that of Lemma C.3(d). So we have

$$II_1 = O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{N}T}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\|\hat{\omega} - \omega\|).$$

Consider II_2, which can be written as

$$\operatorname{tr}\Big[\frac{1}{NT}\sum_{t=1}^{T} f_t\dot{Q}'_{t-1}\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}(e_s e'_s - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big]$$
$$-\operatorname{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T} f_t\dot{Q}'_{t-1}\widehat{\ddot{M}}(\hat{\Sigma}_{ee} - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big] - \operatorname{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T} f_t\dot{Q}'_{t-1}\widehat{\ddot{M}}\bar{e}\bar{e}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big].$$

We use II_3, II_4 and II_5 to denote the above three expressions. First consider II_3, which is equivalent to

$$\operatorname{tr}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\Big(\frac{1}{T}\sum_{t=1}^{T} f_t Q_{it-1}\Big)\hat{\lambda}'_j\frac{1}{T}\sum_{s=1}^{T}[e_{is}e_{js} - E(e_{is}e_{js})]\hat{D}H^{-1\prime}\Big] \tag{D.5}$$
$$-\operatorname{tr}\Big[\Big(\frac{1}{NT}\sum_{t=1}^{T} f_t Q_{t-1}\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big)\Big(\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\frac{1}{T}\sum_{s=1}^{T}(e_s e'_s - \Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big)\hat{D}H^{-1\prime}\Big]\Big].$$

The first expression of the first term of (D.5) in the trace operator (ignore $\hat{D}H^{-1\prime}$) can be further written as

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\Big(\frac{1}{T}\sum_{t=1}^{T} f_t Q_{it-1}\Big)(\hat{\lambda}_j - H\lambda_j)'\frac{1}{T}\sum_{s=1}^{T}[e_{is}e_{js} - E(e_{is}e_{js})]$$
$$-\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2 - \sigma_j^2}{\hat{\sigma}_i^2\hat{\sigma}_j^2\sigma_j^2}\Big(\frac{1}{T}\sum_{t=1}^{T} f_t Q_{it-1}\Big)\lambda'_j\frac{1}{T}\sum_{s=1}^{T}[e_{is}e_{js} - E(e_{is}e_{js})]H'$$

$$+\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\Big(\frac{1}{T}\sum_{t=1}^{T}f_tQ_{it-1}\Big)\Big[\frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T}\frac{1}{\sigma_j^2}\lambda_j'[e_{is}e_{js}-E(e_{is}e_{js})]\Big]H'$$

The first term of the above expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_tQ_{it-1}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}\|\hat{\lambda}_j-H\lambda_j\|^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{s=1}^{T}[e_{is}e_{js}-E(e_{is}e_{js})]\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.2. The second term is bounded in norm by

$$C\|H\|\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_tQ_{it-1}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]^{1/2}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big|\frac{1}{T}\sum_{s=1}^{T}[e_{is}e_{js}-E(e_{is}e_{js})]\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.4. The third term is bounded in norm by

$$C\|H\|\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_tQ_{it-1}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T}\frac{1}{\sigma_j^2}\lambda_j'[e_{is}e_{js}-E(e_{is}e_{js})]\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{\sqrt{NT}})$. Given the above result, we have that the first term of (D.5) is $O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. By Lemma B.6, the second term is $O_p(N^{-1/2}\frac{1}{T\sqrt{T}})+O_p(\frac{1}{NT})+O_p(T^{-2})+o_p(\|\hat{\omega}-\omega\|)$. Given these two results, we have

$$II_3=O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

The derivations on II_4 and II_5 are similar as those of

$$\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T}f_t\dot{X}_{tp}'\widehat{\ddot{M}}(\hat{\Sigma}_{ee}-\Sigma_{ee})\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big]$$

and

$$\mathrm{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T}f_t\dot{X}_{tp}'\widehat{\ddot{M}}\bar{e}\bar{e}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}\Big],$$

which are given in Lemma C.3(d), we therefore omit the details. Given the results on II_1 and II_2, we obtain (d). □

Analyzing the first order condition for δ. The first order condition for δ is

$$\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\dot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\hat{\rho}\ddot{Y}_t-\dot{X}_t\hat{\beta})=0.$$

By $\dot{Y}_t=\delta\dot{Y}_{t-1}+\rho\ddot{Y}_t+\dot{X}_t\beta+\Lambda f_t+\dot{e}_t$, we have

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{Y}_{t-1}\Big](\hat{\delta}-\delta)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\ddot{Y}_t\Big](\hat{\rho}-\rho)$$

$$+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{X}_t\Big](\hat{\beta}-\beta)=\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\Lambda f_t+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{e}_t.$$

Using the arguments in deriving (C.12), we have

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{Y}_{t-1}\Big](\hat{\delta}-\delta)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\ddot{Y}_t\Big](\hat{\rho}-\rho)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{X}_t\Big](\hat{\beta}-\beta)$$

$$=\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{e}_t-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\Lambda\frac{1}{NT}\sum_{s=1}^{T}f_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t \tag{D.6}$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{e}_s\pi_{st}-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{e}_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{Y}_{s-1}\pi_{st}+(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\ddot{Y}_s\pi_{st}$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\dot{X}_s(\hat{\beta}-\beta)\pi_{st}+\mathbb{S}_{\delta 1}+\mathbb{S}_{\delta 2}.$$

where $\pi_{st}=f_s'(F'F)^{-1}f_t$ and

$$\mathbb{S}_{\delta 1}=-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\hat{\delta}-\delta)^2\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\hat{\rho}-\rho)^2\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

and

$$\mathbb{S}_{\delta 2}=\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\hat{\delta}-\delta)\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}X_s(\hat{\beta}-\beta)e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\beta}-\beta)'X_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

Using the results in Lemmas D.1, D.2 and D.3, together with the fact

$$\frac{1}{NT}(\mathbf{1}_T'\mathbf{1}_T)^{-1}\mathbf{1}_T'L\mathbf{1}_T+\frac{1}{NT}\mathrm{tr}\Big[(F'F)^{-1}F'LF\Big]=\frac{1}{NT}\mathrm{tr}[(\widetilde{F}'\widetilde{F})^{-1}\widetilde{F}'L\widetilde{F}],$$

where $\widetilde{F}=(F,\mathbf{1}_T)$, equation (D.6) can be simplified as

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\ddot{M}\dot{Y}_{t-1}-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\ddot{M}\dot{Y}_{s-1}\pi_{st}\Big](\hat{\delta}-\delta)$$
$$+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\ddot{M}\ddot{Y}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\ddot{M}\ddot{Y}_s\pi_{st}\Big](\hat{\rho}-\rho)$$
$$+\Big[\frac{1}{NT}\sum_{t=1}^{T}\dot{Y}_{t-1}'\ddot{M}\dot{X}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{Y}_{t-1}'\ddot{M}\dot{X}_s\pi_{st}\Big](\hat{\beta}-\beta)$$
$$=\frac{1}{NT}\sum_{t=1}^{T}\dot{B}_{t-1}'\ddot{M}e_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{B}_{t-1}'\ddot{M}\dot{e}_s\pi_{st}+\frac{1}{NT}\sum_{t=1}^{T}Q_{t-1}'\Sigma_{ee}^{-1}e_t \tag{D.7}$$
$$-\frac{1}{NT}\mathrm{tr}[(\widetilde{F}'\widetilde{F})^{-1}\widetilde{F}'L\widetilde{F}]+O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

This completes the whole analysis. □

Appendix E: Analyzing the first order condition for ρ

In this section, we give a detailed analysis on the first order condition for ρ. The following lemmas are useful for the subsequent analysis.

Lemma E.1 *Under Assumptions A-H,*

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}(\hat{\sigma}_i^2-\sigma_i^2)S_{ii,N}=\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\frac{1}{T}\sum_{t=1}^{T}(e_{it}^2-\sigma_i^2)-\operatorname{tr}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\lambda_i\lambda_i'\Big]$$

$$-(r+1)\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N}-2\Big(\frac{1}{N}\sum_{i=1}^{N}S_{ii,N}^2\Big)(\hat{\rho}-\rho)+O_p(N^{-3/2})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

Proof of Lemma E.1. By (B.12), the left hand side can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}(\hat{\sigma}_i^2-\sigma_i^2)S_{ii,N}=\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}\frac{1}{T}\sum_{t=1}^{T}(e_{it}^2-\sigma_i^2)-2\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\dot{e}_{it}$$

$$-2\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}\lambda_i'H'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)\dot{e}_{it}+2\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t(\hat{f}_t-H^{-1\prime}f_t)'H\lambda_i$$

$$+\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\hat{f}_t'(\hat{\lambda}_i-H\lambda_i)-\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}\bar{e}_i^2$$

$$+\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}\lambda_i'H'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)(\hat{f}_t-H^{-1\prime})'H\lambda_i+\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}\mathbb{U}_{i1}+\frac{1}{N}\sum_{i=1}^{N}\frac{S_{ii,N}}{\sigma_i^2}\mathbb{U}_{i2}$$

$$=I\!I_1+I\!I_2+\cdots+I\!I_9,\qquad \text{say.}$$

Consider $I\!I_2$, which can be written as

$$\begin{aligned}&-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}(\hat{\lambda}_i-H\lambda_i)'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)\dot{e}_{it}\\&-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}(\hat{\lambda}_i-H\lambda_i)'H^{-1\prime}\frac{1}{T}\sum_{t=1}^{T}f_te_{it}\end{aligned}\tag{E.1}$$

The first term of (E.1) can be written as

$$-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\Big[\hat{\lambda}_i-H\lambda_i-H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big]'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)\dot{e}_{it}$$
$$-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\sum_{s=1}^{T}e_{is}f_s'(F'F)^{-1}H'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-H^{-1\prime}f_t)\dot{e}_{it}.$$

The first term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\hat{\lambda}_i-H\lambda_i-H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big\|^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t-H^{-1\prime}f_t\|^2\Big]^{1/2}\Big[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\dot{e}_{it}^2\Big]^{1/2},$$

which is $O_p(N^{-3/2})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. Consider the second term, which can be written as

$$-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\sum_{s=1}^{T}e_{is}f_s'(F'F)^{-1}H'\frac{1}{T}\sum_{t=1}^{T}\Big(\hat{f}_t-H^{-1\prime}f_t-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)\dot{e}_{it}$$

$$-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\sum_{s=1}^{T}e_{is}f_s'(F'F)^{-1}H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_te_{it} \tag{E.2}$$
$$+\frac{2}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\sum_{s=1}^{T}e_{is}f_s'(F'F)^{-1}H'\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}\bar{e}_i.$$

The first term of (E.2) is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\sum_{s=1}^{T}e_{is}f_s'(F'F)^{-1}\Big\|^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\Big\|\hat{f}_t-H^{-1\prime}f_t-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big\|^2\Big]^{1/2}\Big[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\dot{e}_{it}^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.6. The second term of (E.2) can be written as

$$-2\text{tr}\Big[\Big(\frac{1}{T}F'F\Big)^{-1}H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_t\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}S_{ii,N}[e_{it}e_{is}-E(e_{it}e_{is})]f_s'\Big]$$
$$-2\text{tr}\Big[\frac{1}{T}\Big(\frac{1}{T}F'F\Big)^{-1}H'\Big(\frac{1}{N}\sum_{i=1}^{N}S_{ii,N}\Big)\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}e_tf_t'\Big]. \tag{E.3}$$

The expression of the first term in the trace operator (ignore $(\frac{1}{T}F'F)^{-1}H'$) can be written as

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_j^2}(\hat{\lambda}_j-H\lambda_j)e_{jt}\Big(\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{S_{ii,N}}{\sigma_i^2}[e_{it}e_{is}-E(e_{it}e_{is})]f_s'\Big)$$
$$-H\frac{1}{NT}\sum_{t=1}^{T}\sum_{j=1}^{N}\frac{\hat{\sigma}_j^2-\sigma_j^2}{\hat{\sigma}_j^2\sigma_j^2}\lambda_je_{jt}\Big(\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{S_{ii,N}}{\sigma_i^2}[e_{it}e_{is}-E(e_{it}e_{is})]f_s'\Big)$$
$$+H\frac{1}{NT}\sum_{t=1}^{T}\sum_{j=1}^{N}\frac{1}{\sigma_j^2}\lambda_je_{jt}\Big(\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{S_{ii,N}}{\sigma_i^2}[e_{it}e_{is}-E(e_{it}e_{is})]f_s'\Big).$$

The first term of the above expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{j=1}^{N}\|\hat{\lambda}_j-H\lambda_j\|^2\Big]^{1/2}\Big[\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}e_{jt}^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{S_{ii,N}}{\sigma_i^2}[e_{it}e_{is}-E(e_{it}e_{is})]f_s'\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-3/2}T^{-1/2})+O_p(\frac{1}{\sqrt{NT}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.2. The second term is bounded in norm by

$$C\|H\|\Big[\frac{1}{N}\sum_{j=1}^{N}(\hat{\sigma}_j^2-\sigma_j^2)^2\Big]^{1/2}\Big[\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}e_{jt}^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{S_{ii,N}}{\sigma_i^2}[e_{it}e_{is}-E(e_{it}e_{is})]f_s'\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-3/2}T^{-1/2})+O_p(\frac{1}{\sqrt{NT}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.4. The third term is bounded in norm by

$$C\Big[\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\sigma_j^2}\lambda_je_{jt}\Big\|^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\frac{S_{ii,N}}{\sigma_i^2}[e_{it}e_{is}-E(e_{it}e_{is})]f_s'\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}})$. Given the above result, we have that the first term of (E.3) is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + o_p(\|\hat{\omega} - \omega\|)$. The second term of (E.3) is $O_p(N^{-1/2}\frac{1}{T\sqrt{T}}) + O_p(T^{-2}) + o_p(\|\hat{\omega} - \omega\|)$ by Lemma B.6. So the second term of (E.2) is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(T^{-2}) + o_p(\|\hat{\omega} - \omega\|)$. The last term of (E.2) is $O_p(\frac{1}{\sqrt{NT}}) + O_p(T^{-5/2}) + o_p(\|\hat{\omega} - \omega\|)$ by (D.4). Summarizing all the results, we have that the first term of (E.1) is $O_p(N^{-3/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$.

Consider the second term of (E.1), which can be written as

$$-2\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\Big[\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1}\sum_{s=1}^{T}f_s e_{is}\Big]'H^{-1\prime}\frac{1}{T}\sum_{t=1}^{T}f_t e_{it}$$

$$-2\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}S_{ii,N}e_{is}e_{it}\pi_{st}.$$

The first term of the above expression is bounded in norm by

$$C\|H\|\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1}\sum_{s=1}^{T}f_s e_{is}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T}\sum_{t=1}^{T}f_t e_{it}\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.6. The second term can be written as

$$-2\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}S_{ii,N}[e_{is}e_{it} - E(e_{is}e_{it})]\pi_{st} - 2\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}S_{ii,N}\pi_{tt},$$

which is $O_p(\frac{1}{\sqrt{NT}}) - 2r\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N}$. Given the above result, we have that the first term of (E.1) is $-2r\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N} + O_p(N^{-3/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$. So we have

$$II_2 = -2r\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N} + O_p(N^{-3/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Consider II_3, which is equivalent to

$$-2\text{tr}\Big[H'\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t - H^{-1\prime}f_t)\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\dot{e}_{it}\lambda_i'\Big].$$

The expression in the trace operator can be further written as

$$H'\frac{1}{T}\sum_{t=1}^{T}\Big(\hat{f}_t - H^{-1\prime}f_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\dot{e}_{it}\lambda_i'$$

$$+H'\frac{1}{NT}\sum_{t=1}^{T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\dot{e}_{it}\lambda_i'.$$

The first term is bounded in norm by

$$C\|H\|\Big[\frac{1}{T}\sum_{t=1}^{T}\Big\|\hat{f}_t - H^{-1\prime}f_t - \frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big\|^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\dot{e}_{it}\lambda_i'\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-3/2}) + O_p(\frac{1}{\sqrt{NT}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.6. The second term can be shown to be

$$\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\lambda_i\lambda_i' + O_p(\frac{1}{N\sqrt{T}}\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$$

similarly as term I_6 in Lemma C.3(b). So we have

$$II_3 = -2\text{tr}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\lambda_i\lambda_i'\Big] + O_p(N^{-3/2}T^{-1/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Consider II_4, which can be written as

$$\text{tr}\Big[H\Big(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\lambda_i(\hat{\lambda}_i - H\lambda_i)'\Big)\Big(\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t(\hat{f}_t - H^{-1\prime}f_t)'\Big)\Big].$$

Using the arguments in proving Proposition B.5(b), the expression in the former bracket is $O_p(N^{-1}) + O_p(T^{-1}) + O_p(\|\hat{\omega} - \omega\|)$. The expression in the latter bracket is bounded in norm by

$$\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t\|^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t - H^{-1\prime}f_t\|^2\Big]^{1/2} = O_p(N^{-1/2}) + O_p(T^{-1/2}) + O_p(\|\hat{\omega} - \omega\|)$$

by Proposition B.4. Given the above result, we have

$$II_4 = O_p(N^{-3/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Consider II_5, which can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\Big[\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big]'\Big[\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\hat{f}_t'\Big]\Big[\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big]$$
$$+2\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\Big[H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big]'\Big[\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\hat{f}_t'\Big]\Big[\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big]$$
$$+\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\Big[H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big]'\Big[\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t\hat{f}_t'\Big]\Big[H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big].$$

The first term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big\|^2\Big]\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t\|^2\Big],$$

which is $O_p(\frac{1}{N^2}) + O_p(T^{-2}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.6. The second term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\hat{\lambda}_i - H\lambda_i - H(F'F)^{-1}\sum_{s=1}^{T}f_se_{is}\Big\|^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^{T}\|\hat{f}_t\|^2\Big],$$

which is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.6. The third term is apparent to be $r\frac{1}{NT}\sum_{i=1}^N S_{ii,N}+O_p(\frac{1}{\sqrt{NT}})$. Summarizing all the results we have

$$II_5 = r\frac{1}{NT}\sum_{i=1}^N S_{ii,N}+O_p(\frac{1}{N^2})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

Apparently,

$$II_6 = -\frac{1}{NT}\sum_{i=1}^N S_{ii,N}+O_p(\frac{1}{\sqrt{NT}}).$$

Consider II_7, which can be written as

$$\operatorname{tr}\Big[\Big(\frac{1}{N}\sum_{i=1}^N\frac{1}{\sigma_i^2}S_{ii,N}\lambda_i\lambda_i'\Big)\Big[H'\frac{1}{T}\sum_{t=1}^T(\hat{f}_t-H^{-1\prime}f_t)(\hat{f}_t-H^{-1\prime}f_t)'H\Big]\Big]$$

So it suffices to consider $H'\frac{1}{T}\sum_{t=1}^T(\hat{f}_t-H^{-1\prime}f_t)(\hat{f}_t-H^{-1\prime}f_t)'H$. This term can be written as

$$H'\frac{1}{T}\sum_{t=1}^T\Big(\hat{f}_t-H^{-1\prime}f_t-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)\Big(\hat{f}_t-H^{-1\prime}f_t-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)'H$$

$$+2H'\frac{1}{T}\sum_{t=1}^T\Big(\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)\Big(\hat{f}_t-H^{-1\prime}f_t-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)'H$$

$$+H'\frac{1}{T}\sum_{t=1}^T\Big(\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)\Big(\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big)'H.$$

The first term is bounded in norm by

$$C\|H\|^2\Big[\frac{1}{T}\sum_{t=1}^T\Big\|\hat{f}_t-H^{-1\prime}f_t-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big\|^2\Big],$$

which is $O_p(\frac{1}{N^2})+O_p(T^{-2})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.6. The second term is bounded in norm by

$$2\|H\|^2\Big[\frac{1}{T}\sum_{t=1}^T\Big\|\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big\|^2\Big]^{1/2}\Big[\frac{1}{T}\sum_{t=1}^T\Big\|\hat{f}_t-H^{-1\prime}f_t-\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\dot{e}_t\Big\|^2\Big]^{1/2},$$

which is $O_p(N^{-3/2})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.6. The third term can be written as

$$H'\frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\hat{\lambda}_i\hat{\lambda}_j'\frac{1}{T}\sum_{t=1}^T[e_{it}e_{jt}-E(e_{it}e_{jt})]H-H'\frac{1}{N^2}\sum_{i=1}^N\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^4}\hat{\lambda}_i\hat{\lambda}_i'H$$

$$+\frac{1}{N}H'H-H'\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}\bar{e}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}H$$

The first term is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Lemma B.6. The second term is $O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. The third term is $\frac{1}{N}I_r+O_p(\frac{1}{N^2})+O_p(\frac{1}{NT})+o_p(\|\hat{\omega}-$

$\omega\|)$ by Proposition B.5. The last term is $O_p(\frac{1}{NT}) + O_p(T^{-3}) + o_p(\|\hat{\omega} - \omega\|)$ by (D.4). Given the above results, we have

$$II_7 = \text{tr}\Big(\frac{1}{N}\sum_{i=1}^N \frac{1}{\sigma_i^2} S_{ii,N}\lambda_i\lambda_i'\Big) + O_p(N^{-3/2}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

For II_8, it is easy to see that $II_8 = O_p(\|\hat{\omega} - \omega\|^2) = o_p(\|\hat{\omega} - \omega\|)$. Consider II_9, which is equal to $-(\frac{2}{N}\sum_{i=1}^N S_{ii,N}^2)(\hat{\rho} - \rho) + o_p(\|\hat{\omega} - \omega\|)$. The derivation is similar as that of Lemma B.2. The details are therefore omitted. Summarizing all the results, we have Lemma E.1. □

Lemma E.2 *Under Assumptions A-H,*

$$\frac{1}{NT}\sum_{t=1}^T \dot{e}_t' S_N' \widehat{\ddot{M}} \dot{e}_t - \frac{1}{N}\text{tr}(S_N) = \frac{1}{NT}\sum_{t=1}^T e_t' S_N^{\circ\prime} \Sigma_{ee}^{-1} e_t + r\frac{1}{NT}\sum_{i=1}^N S_{ii,N} - \text{tr}\Big[\frac{1}{N^2}\Lambda' S_N^{\circ\prime}\Sigma_{ee}^{-1}\Lambda\Big]$$
$$+2\Big(\frac{1}{N}\sum_{i=1}^N S_{ii,N}^2\Big)(\hat{\rho} - \rho) + O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

Proof of Lemma E.2. The left hand side is equal to

$$\Big[\frac{1}{NT}\sum_{t=1}^T e_t' S_N' \hat{\Sigma}_{ee}^{-1} e_t - \frac{1}{N}\text{tr}(S_N)\Big] - \frac{1}{N^2 T}\sum_{t=1}^T e_t' S_N' \hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1} e_t - \frac{1}{N}\bar{e}' S_N' \widehat{\ddot{M}} \bar{e} = I_1 - I_2 - I_3, \text{ say.}$$

Consider I_1, which is equivalent to

$$I_1 = \frac{1}{NT}\sum_{t=1}^T e_t'(\hat{\Sigma}_{ee}^{-1} - \Sigma_{ee}^{-1}) S_N e_t + \frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \frac{1}{\sigma_i^2} S_{ij,N}[e_{it}e_{jt} - E(e_{it}e_{jt})] = I_4 + I_5, \quad \text{say.}$$

Consider I_4, which is equivalent to

$$\begin{aligned} I_4 &= -\frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2} S_{ij,N} e_{it} e_{jt} \\ &= -\frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2} S_{ij,N}[e_{it}e_{jt} - E(e_{it}e_{jt})] - \frac{1}{N}\sum_{i=1}^N \frac{\hat{\sigma}_i^2 - \sigma_i^2}{\hat{\sigma}_i^2} S_{ii,N} = -I_6 - I_7 \end{aligned}$$

Further consider I_6, which can be written as

$$\begin{aligned} &\frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \frac{1}{\hat{\sigma}_i^2\sigma_i^2}\Big[\hat{\sigma}_i^2 - \sigma_i^2 - \frac{1}{T}\sum_{s=1}^T (e_{is}^2 - \sigma_i^2)\Big] S_{ij,N}[e_{it}e_{jt} - E(e_{it}e_{jt})] \\ &\quad + \frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T \frac{1}{\hat{\sigma}_i^2\sigma_i^2}\Big[\frac{1}{T}\sum_{s=1}^T (e_{is}^2 - \sigma_i^2)\Big] S_{ij,N}[e_{it}e_{jt} - E(e_{it}e_{jt})]. \end{aligned} \tag{E.4}$$

The first term of (E.4) is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^N \Big|\hat{\sigma}_i^2 - \sigma_i^2 - \frac{1}{T}\sum_{s=1}^T (e_{is}^2 - \sigma_i^2)\Big|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^N \Big|\frac{1}{T}\sum_{t=1}^T [e_{it}\ddot{e}_{it} - E(e_{it}\ddot{e}_{it})]\Big|^2\Big]^{1/2}$$

with $\ddot{e}_{it} = \sum_{j=1}^{N} S_{ij,N} e_{jt}$, which is $O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.6. The second term of (E.4) can be written as

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{1}{\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]S_{ij,N}[e_{it}e_{jt}-E(e_{it}e_{jt})]$$

$$-\frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]S_{ij,N}[e_{it}e_{jt}-E(e_{it}e_{jt})].$$

The first term of the preceding expression is

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]^2 S_{ii,N} + O_p(\frac{1}{\sqrt{NT}}).$$

The second term is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\Big(\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big)\Big(\frac{1}{T}\sum_{t=1}^{T}[e_{it}\ddot{e}_{it}-E(e_{it}\ddot{e}_{it})]\Big)\Big|^2\Big]^{1/2}$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|)$ by Proposition B.4. So we have

$$\begin{aligned} I_6 = {} & \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]^2 S_{ii,N} + O_p(\frac{1}{\sqrt{N}T}) \\ & + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|). \end{aligned}$$

Further consider I_7, which is equal to

$$I_7 = -\frac{1}{N}\sum_{i=1}^{N}\frac{(\hat{\sigma}_i^2-\sigma_i^2)^2}{\hat{\sigma}_i^2\sigma_i^2}S_{ii,N} + \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\sigma_i^2}S_{ii,N} = -I_8 + I_9.$$

Consider the term I_8, which is equivalent to

$$\begin{aligned} I_8 = {} & \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2\sigma_i^2}\Big[\hat{\sigma}_i^2-\sigma_i^2-\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]^2 S_{ii,N} + \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2\sigma_i^2}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]^2 S_{ii,N} \\ & + \frac{2}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2\sigma_i^2}\Big[\hat{\sigma}_i^2-\sigma_i^2-\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)S_{ii,N}. \end{aligned} \tag{E.5}$$

By the boundedness of $\hat{\sigma}_i^2, \sigma_i^2$ and $S_{ii,N}$, the first term on the right hand side is bounded in norm by

$$C\frac{1}{N}\sum_{i=1}^{N}\Big|\hat{\sigma}_i^2-\sigma_i^2-\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big|^2 = O_p(\frac{1}{N^2}) + O_p(T^{-2}) + O_p(\|\hat{\omega}-\omega\|^2)$$

by Proposition B.6. The second term of (E.5) can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]^2 S_{ii,N} - \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]^2 S_{ii,N}. \tag{E.6}$$

The second term of the above expression is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}S_{ii,N}^2\Big|\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big|^4\Big]^{1/2},$$

which is $O_p(\frac{1}{NT})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.4. The third term of (E.5) is bounded in norm by

$$C\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\hat{\sigma}_i^2-\sigma_i^2-\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}S_{ii,N}^2\Big|\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.6. Summarizing all the result, we have

$$I_8=\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^4}\Big[\frac{1}{T}\sum_{s=1}^{T}(e_{is}^2-\sigma_i^2)\Big]^2S_{ii,N}+O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

The term I_9 is given in Lemma E.1. By the definitions of I_1 and $I_4,\dots,I_9$, we have $I_1=I_5-I_6+I_8-I_9$. Given the results on I_5, I_6, I_8 and I_9, we have

$$I_1=\frac{1}{NT}\sum_{t=1}^{T}e_t'S_N^{\circ\prime}\Sigma_{ee}^{-1}e_t+\operatorname{tr}\Big[\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}S_{ii,N}\lambda_i\lambda_i'\Big]+(r+1)\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N}+2\Big(\frac{1}{N}\sum_{i=1}^{N}S_{ii,N}^2\Big)(\hat{\rho}-\rho)$$
$$+O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|). \tag{E.7}$$

Now consider the term I_2, which can be written as

$$\operatorname{tr}\Big[\frac{1}{N^2T}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(e_te_t'-\Sigma_{ee})S_N'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]-\operatorname{tr}\Big[\frac{1}{N^2}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}(\hat{\Sigma}_{ee}-\Sigma_{ee})S_N'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big]+\operatorname{tr}\Big[\frac{1}{N^2}\hat{\Lambda}'S_N'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\Big].$$

Ignore the signs of the above three terms, we use I_{10},I_{11} and I_{12} to denote them. Term I_{10} can be written as

$$\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\hat{\sigma}_i^2\hat{\sigma}_j^2}\hat{\lambda}_i\hat{\lambda}_j'\frac{1}{T}\sum_{t=1}^{T}[e_{it}\ddot{e}_{it}-E(e_{it}\ddot{e}_{it})],$$

which can be proved to be $O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ similarly as Lemma B.6(b). To analyze I_{11} and I_{12}, we first note that

$$\frac{1}{N}\sum_{i=1}^{N}\Big\|\sum_{j=1}^{N}S_{ij,N}(\hat{\lambda}_j-H\lambda_j)\Big\|^2=O_p(\frac{1}{N^2})+O_p(T^{-1})+O_p(\|\hat{\omega}-\omega\|^2).$$

By $\sum_{j=1}^{N}|S_{ij,N}|<\infty$ for all i, the proof of the above result is almost the same as that of Proposition B.2(a) if we treating $\sum_{j=1}^{N}S_{ij,N}e_{jt}$ as a new e_{it}. Now first consider I_{12}, which can be written as

$$I_{12}=\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\hat{\lambda}_i\Big[\sum_{j=1}^{N}S_{ij,N}(\hat{\lambda}_j-H\lambda_j)\Big]'+\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}(\hat{\lambda}_i-H\lambda_i)\Big[\sum_{j=1}^{N}S_{ij,N}\lambda_j'\Big]H'$$

$$-H\frac{1}{N^2}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}\lambda_i\Big[\sum_{j=1}^{N}S_{ij,N}\lambda_j\Big]'H'+\frac{1}{N}\text{tr}\Big[H'H\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i\Big(\sum_{j=1}^{N}S_{ij,N}\lambda_j\Big)\Big].$$

The first term is bounded in norm by

$$C\frac{1}{N}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\hat{\sigma}_i^2}\|\hat{\lambda}_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\sum_{j=1}^{N}S_{ij,N}(\hat{\lambda}_j-H\lambda_j)\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. The second term is bounded in norm by

$$C\frac{1}{N}\|H\|\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}\|\hat{\lambda}_i-H\lambda_i\|^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\sum_{j=1}^{N}S_{ij,N}\lambda_j\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.2. The third term is bounded in norm by

$$C\frac{1}{N}\|H\|^2\cdot\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\sum_{j=1}^{N}S_{ij,N}\lambda_j\Big\|^2\Big]^{1/2},$$

which is also $O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.4. The last term is $\frac{1}{N}\text{tr}[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i(\sum_{j=1}^{N}S_{ij,N}\lambda_j)]+O_p(\frac{1}{N^2})+O_p(\frac{1}{NT})$ by Proposition B.5. So we have

$$I_{12}=\frac{1}{N}\text{tr}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i\Big(\sum_{j=1}^{N}S_{ij,N}\lambda_j\Big)\Big]+O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|).$$

Term I_{11} can be proved to be $O_p(\frac{1}{N^2})+O_p(\frac{1}{N\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ similarly as I_{12}. Given the above results, we have

$$I_2=\frac{1}{N}\text{tr}\Big[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i\Big(\sum_{j=1}^{N}S_{ij,N}\lambda_j\Big)\Big]+O_p(\frac{1}{N^2})$$
$$+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{N\sqrt{T}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|). \tag{E.8}$$

Consider I_3, which can be written as

$$\frac{1}{N}\bar{e}'S_N'\hat{\Sigma}_{ee}^{-1}\bar{e}-\frac{1}{N^2}\bar{e}'S_N'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}.$$

Notice $\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}\bar{e}=O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$ by (D.4). If we treating $S_N\bar{e}$ as a new $\bar{e}$, we can show in almost the same way that $\frac{1}{N}\hat{\Lambda}'\hat{\Sigma}_{ee}^{-1}S_N\bar{e}=O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T\sqrt{T}})+o_p(\|\hat{\omega}-\omega\|)$. So the second term is $O_p(\frac{1}{NT})+O_p(\frac{1}{T^3})+o_p(\|\hat{\omega}-\omega\|)$. The first term is $\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N}+O_p(\frac{1}{\sqrt{NT}})+o_p(\|\hat{\omega}-\omega\|)$, which can be shown similarly as term I_4. So we have

$$I_3=\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N}+O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{T^3})+o_p(\|\hat{\omega}-\omega\|). \tag{E.9}$$

Given the results (E.7), (E.8) and (E.9) and noticing that the left hand side of the lemma is equal to $I_1-I_2-I_3$, we obtain Lemma E.2. □

Lemma E.3 *Under Assumptions A-H,*

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{e}_t'S_N'\hat{\Sigma}_{ee}^{-1}\dot{e}_s\pi_{st} = r\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N} + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

where $\pi_{st} = f_s'(F'F)^{-1}f_t$.

Proof of Lemma E.3. The left hand side is equivalent to

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{e}_t'S_N'\hat{\Sigma}_{ee}^{-1}\dot{e}_s\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}e_t'S_N'\hat{\Sigma}_{ee}^{-1}e_s\pi_{st}, \tag{E.10}$$

where we use the fact that $\sum_{t=1}^{T}f_t = 0$. The right hand side of (E.10) can be written as

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\hat{\sigma}_i^2}[\ddot{e}_{it}e_{is} - E(\ddot{e}_{it}e_{it})]\pi_{st} - r\frac{1}{NT}\sum_{i=1}^{N}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2}S_{ii,N} + r\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N}, \tag{E.11}$$

where $\ddot{e}_{it} = \sum_{j=1}^{N}S_{ij,N}e_{jt}$. The first term of the preceding expression is equal to

$$\text{tr}\Big[\Big(\frac{1}{T}F'F\Big)^{-1}\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\hat{\sigma}_i^2}f_tf_s'[\ddot{e}_{it}e_{is} - E(\ddot{e}_{it}e_{it})]\Big],$$

which can be written as

$$\text{tr}\Big[\Big(\frac{1}{T}F'F\Big)^{-1}\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{\sigma_i^2}f_tf_s'[\ddot{e}_{it}e_{is} - E(\ddot{e}_{it}e_{it})]\Big]$$
$$-\text{tr}\Big[\Big(\frac{1}{T}F'F\Big)^{-1}\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{\hat{\sigma}_i^2-\sigma_i^2}{\hat{\sigma}_i^2\sigma_i^2}f_tf_s'[\ddot{e}_{it}e_{is} - E(\ddot{e}_{it}e_{it})]\Big].$$

The first term is $O_p(\frac{1}{\sqrt{NT}})$. The second term is bounded in norm by

$$C\Big\|\frac{1}{T}F'F\Big\|\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i^2)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}f_tf_s'[\ddot{e}_{it}e_{is} - E(\ddot{e}_{it}e_{it})]\Big\|^2\Big]^{1/2},$$

which is $O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|)$ by Proposition B.4. So the first term of (E.11) is $O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|)$. The second term of (E.11) is bounded in norm by

$$C\frac{1}{T}\Big[\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2-\sigma_i)^2\Big]^{1/2} = O_p(\frac{1}{NT}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|)$$

by Proposition B.4. Given the above result, we have Lemma E.3. □

Lemma E.4 *Let $\mathbb{S}_{\rho 1}$ and $\mathbb{S}_{\rho 2}$ be defined in (E.12). Under Assumptions A-H,*

$$\mathbb{S}_{\rho 1} = O_p(\|\hat{\omega}-\omega\|^2), \qquad \mathbb{S}_{\rho 2} = o_p(\|\hat{\omega}-\omega\|).$$

The proof of Lemma E.4 is similar as that of Lemma B.2. See also the proof of Lemma C.1 in Bai and Li (2014a) for more details.

Lemma E.5 *Under Assumptions A-H,*

$$(a)\ \ \frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\ddot{Y}_t = \frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\ddot{M}\ddot{Y}_t + o_p(1);$$

$$(b)\ \ \frac{1}{N}\mathrm{tr}[S_N^2(\widetilde{\rho})] = \frac{1}{N}\mathrm{tr}(S_N^2) + o_p(1);$$

$$(c)\ \ \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\ddot{Y}_s\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\ddot{M}\ddot{Y}_s\pi_{st} + o_p(1).$$

where $S_N(\widetilde{\rho}) = W_N(I_N - \widetilde{\rho}W_N)^{-1}$ *and* $\widetilde{\rho}$ *is some point between* $\hat{\rho}$ *and* ρ.

The proof of Lemma E.5 is similar and actually easier than that of Lemma E.6 below. The details are omitted.

Lemma E.6 *Under Assumptions A-H,*

$$\begin{aligned}(a)\ \ &\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\Lambda\frac{1}{NT}\sum_{s=1}^{T}f_se_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\ &= O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);\end{aligned}$$

$$\begin{aligned}(b)\ \ &\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{e}_t - \frac{1}{N}\mathrm{tr}(S_N)\\ &= \frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\ddot{M}e_t + \frac{1}{NT}\sum_{t=1}^{T}J_t'\Sigma_{ee}^{-1}e_t + \frac{1}{NT}\sum_{t=1}^{T}e_t'S_N^{\circ\prime}\Sigma_{ee}^{-1}e_t\\ &\quad - \Delta^{\diamond} - \frac{1}{NT}(\mathbf{1}_T'\mathbf{1}_T)^{-1}\mathbf{1}_T'K\mathbf{1}_T - \frac{1}{N^2}\mathrm{tr}[\Lambda'S_N^{\circ\prime}\Sigma_{ee}^{-1}\Lambda] + \frac{2}{N}\sum_{i=1}^{N}S_{ii,N}^2(\hat{\rho}-\rho)\\ &\quad + O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);\end{aligned}$$

$$\begin{aligned}(c)\ \ &\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st}\\ &= \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\ddot{M}e_s\pi_{st} - \Delta^{\diamond} + \frac{1}{NT}\mathrm{tr}[(F'F)^{-1}F'KF]\\ &\quad + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|);\end{aligned}$$

$$\begin{aligned}(d)\ \ &\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{e}_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\ &= O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).\end{aligned}$$

where

$$\Delta^{\diamond} = \frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\Sigma_{ee}^{-1}\Lambda(F'F)^{-1}f_t - r\frac{1}{NT}\sum_{i=1}^{N}S_{ii,N}$$

PROOF OF LEMMA E.6. The proof of result (a) is similar as that of Lemma C.3(a). The proof of result (d) is similar as that of Lemma D.3(d). The details are therefore omitted.

Consider (b). By $\ddot{Y}_t = \ddot{B}_t + \dot{J}_t + S_N\dot{e}_t$, the left hand side of (b) is equal to

$$\frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\widehat{\ddot{M}}\dot{e}_t + \frac{1}{NT}\sum_{t=1}^{T}\dot{J}_t'\widehat{\ddot{M}}\dot{e}_t + \left[\frac{1}{NT}\sum_{t=1}^{T}\dot{e}_t'S_N'\widehat{\ddot{M}}\dot{e}_t - \frac{1}{N}\mathrm{tr}(S_N)\right] = II_1 + II_2 + II_3 \quad \text{say.}$$

Notice that $\ddot{B}_t$ is exogenous, so the first term can be proved to be

$$II_1 = \frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\ddot{M}e_t - \frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\Sigma_{ee}^{-1}\Lambda(F'F)^{-1}f_t + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

similarly as Lemma C.3(b). The second term can be proved to be

$$II_2 = \frac{1}{NT}\sum_{t=1}^{T}J_t\Sigma_{ee}^{-1}e_t - \frac{1}{NT}(\mathbf{1}_T'\mathbf{1}_T)^{-1}\mathbf{1}_T'K\mathbf{1}_T + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|),$$

similarly as term I_2 in Lemma D.3(b). The third term is given in Lemma E.2. Summarizing all the results, we have (b).

Consider (c). By $\ddot{Y}_t = \ddot{B}_t + \dot{J}_t + S_N\dot{e}_t$, the left hand side of (c) is equal to

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st} + \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{J}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st} + \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{e}_t'S_N'\widehat{\ddot{M}}\dot{e}_s\pi_{st}.$$

The first term is

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\ddot{M}\dot{e}_s\pi_{st} - \frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\Sigma_{ee}^{-1}\Lambda(F'F)^{-1}f_t + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|),$$

which can be proved similarly as Lemma C.3(c). The second term can be show as

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{J}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}J_t'\ddot{M}e_s\pi_{st} + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

The first expression is $\frac{1}{NT}\mathrm{tr}[(F'F)^{-1}F'KF] + O_p(\frac{1}{\sqrt{NT}})$. Given this, we have

$$\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{J}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st} = \frac{1}{NT}\mathrm{tr}[(F'F)^{-1}F'KF] + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega}-\omega\|).$$

The third term is given in Lemma E.3. Given the above results, we have (c). □

Analyzing the first order condition for ρ. The first order condition for ρ is

$$\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\dot{Y}_t-\hat{\delta}\dot{Y}_{t-1}-\hat{\rho}\ddot{Y}_t-\dot{X}_t\hat{\beta})-\frac{1}{N}\text{tr}[(I_N-\hat{\rho}W_N)^{-1}W_N]=0.$$

By $\dot{Y}_t=\delta\dot{Y}_{t-1}+\rho\ddot{Y}_t+\dot{X}_t\beta+\dot{X}_t\beta+\Lambda f_t+\dot{e}_t$, we can rewrite the above equation as

$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{Y}_{t-1}\Big](\hat{\delta}-\delta)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\ddot{Y}_t+\frac{1}{N}\text{tr}[S_N^2(\widetilde{\rho})]\Big](\hat{\rho}-\rho)$$
$$\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{X}_t\Big](\hat{\beta}-\beta)=\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\Lambda f_t+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{e}_t-\frac{1}{N}\text{tr}(S_N);$$

where $\widetilde{\rho}$ is some point between $\hat{\rho}$ and ρ. Using the similar method in deriving (C.12), the above equation can be further written as

$$\begin{aligned}
&\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t\widehat{\ddot{M}}\dot{Y}_{t-1}\Big](\hat{\delta}-\delta)+\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\ddot{Y}_t+\frac{1}{N}\text{tr}[S_N^2(\widetilde{\rho})]\Big](\hat{\rho}-\rho)\\
&+\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{X}_t\Big](\hat{\beta}-\beta)=\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{e}_t-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\Lambda\frac{1}{NT}\sum_{s=1}^{T}f_se_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{e}_s\pi_{st}-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{e}_s\dot{e}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t \qquad \text{(E.12)}\\
&+(\hat{\delta}-\delta)\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{Y}_{s-1}\pi_{st}+(\hat{\rho}-\rho)\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\ddot{Y}_s\pi_{st}\\
&+\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\dot{X}_s(\hat{\beta}-\beta)\pi_{st}-\frac{1}{N}\text{tr}(S_N)+\mathcal{S}_{\rho1}+\mathcal{S}_{\rho2}.
\end{aligned}$$

where $\pi_{st}=f_s'(F'F)^{-1}f_t$, $\bar{Y}_{-1}=T^{-1}\sum_{t=1}^{T}Y_{t-1}$, $\bar{e}=T^{-1}\sum_{t=1}^{T}e_t$ and

$$\begin{aligned}
\mathcal{S}_{\rho1}=&-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)^2\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\hat{\rho}-\rho)^2\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t\\
&-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t
\end{aligned}$$

$$-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{X}_s(\hat{\beta}-\beta)(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}(\hat{\delta}-\delta)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s(\hat{\rho}-\rho)(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

and

$$\mathcal{S}_{\rho 2}=\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\Lambda\sum_{s=1}^{T}f_s(\hat{\beta}-\beta)'\dot{X}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\hat{\delta}-\delta)\frac{1}{NT}\sum_{s=1}^{T}\dot{Y}_{s-1}e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}(\hat{\rho}-\rho)\frac{1}{NT}\sum_{s=1}^{T}\ddot{Y}_s e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}X_s(\hat{\beta}-\beta)e_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\delta}-\delta)\dot{Y}_{s-1}'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\rho}-\rho)\ddot{Y}_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\widehat{\ddot{M}}\frac{1}{NT}\sum_{s=1}^{T}e_s(\hat{\beta}-\beta)'X_s'\hat{\Sigma}_{ee}^{-1}\hat{\Lambda}\hat{D}H^{-1\prime}f_t$$

Using the results in Lemmas E.4, E.5 and E.6, the above equation can be simplified as

$$\begin{aligned}
&\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\ddot{M}\dot{Y}_{t-1}-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\ddot{M}\dot{Y}_{s-1}\pi_{st}\Big](\hat{\delta}-\delta)\\
&\quad+\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\ddot{M}\ddot{Y}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\ddot{M}\ddot{Y}_s\pi_{st}+\frac{1}{N}\mathrm{tr}(S_N^2)-\frac{2}{N}\sum_{i=1}^{N}S_{ii,N}^2\Big](\hat{\rho}-\rho)\\
&\quad+\Big[\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\ddot{M}\dot{X}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{Y}_t'\ddot{M}\dot{X}_s\pi_{st}\Big](\hat{\beta}-\beta) \qquad \text{(E.13)}\\
&=\frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\ddot{M}e_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\ddot{M}\dot{e}_s\pi_{st}+\frac{1}{NT}\sum_{t=1}^{T}J_t'\Sigma_{ee}^{-1}e_t\\
&\quad+\frac{1}{NT}\sum_{t=1}^{T}e_t'S_N^{\circ}\Sigma_{ee}^{-1}e_t-\frac{1}{N^2}\mathrm{tr}[\Lambda'S_N^{\circ\prime}\Sigma_{ee}^{-1}\Lambda]-\frac{1}{NT}\mathrm{tr}[P_{\widetilde{F}}K]
\end{aligned}$$

$$+ O_p(\frac{1}{N^2}) + O_p(\frac{1}{N\sqrt{T}}) + O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{T\sqrt{T}}) + o_p(\|\hat{\omega} - \omega\|).$$

This completes the analysis. □

Appendix F: Proof of Theorem 5.2 and Corollary 5.1

Proof of Theorem 5.2. By (D.7), (E.13) and (C.13) and noticing that $O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{N\sqrt{T}})$ is dominated by $O_p(N^{-3/2}) + O_p(\frac{1}{T\sqrt{T}})$, as well as the terms of the order $o_p(\|\hat{\omega} - \omega\|)$ are negligible since they are dominated by the terms on the left hand sides of the three equations, we have

$$\begin{aligned}
&\hat{\omega} - \omega + b \\
&= \mathbb{D}^{-1} \frac{1}{NT} \begin{bmatrix} \sum_{t=1}^{T} \ddot{B}_t' \ddot{M} e_t - \sum_{t=1}^{T} \sum_{s=1}^{T} \ddot{B}_t' \ddot{M} e_s \pi_{st} + \sum_{t=1}^{T} J_t' \Sigma_{ee}^{-1} e_t + \eta \\ \sum_{t=1}^{T} \dot{B}_{t-1}' \ddot{M} e_t - \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{B}_{t-1}' \ddot{M} e_s \pi_{st} + \sum_{t=1}^{T} Q_{t-1}' \Sigma_{ee}^{-1} e_t \\ \sum_{t=1}^{T} \dot{X}_t' \ddot{M} e_t - \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{X}_t' \ddot{M} e_s \pi_{st} \end{bmatrix} \\
&+ O_p(N^{-3/2}) + O_p(\frac{1}{T\sqrt{T}}).
\end{aligned}$$

Theorem 5.2 is a direct result of the above expression. This completes the proof of Theorem 5.2. □

Now we show the corollary 5.1. Notice that if we can show that $\mathbb{D}^{-1/2}\xi \xrightarrow{d} N(0, I)$ conditional on the realizations of λ_i, f_t and x_{it} for every i and t, it would follow that $\mathbb{D}^{-1/2}\xi \xrightarrow{d} N(0, I)$ unconditionally. In this sense, it is no loss of generality to assume that λ_i, f_t and x_{it} are nonrandom. The following lemmas are useful for our analysis.

Lemma F.1 *Under Assumptions A-G,*

$$(a) \quad \frac{1}{NT} \sum_{t=1}^{T} \dot{J}_t' \ddot{M} \dot{J}_t = \frac{1}{NT} \sum_{t=1}^{T} J_t' \Sigma_{ee}^{-1} J_t + o_p(1),$$

$$(b) \quad \frac{1}{NT} \sum_{t=1}^{T} \dot{e}_t' S_N' \ddot{M} (\ddot{B}_t + \dot{J}_t) = o_p(1),$$

$$(c) \quad \frac{1}{NT} \sum_{t=1}^{T} \dot{e}_t' S_N' \ddot{M} S_N \dot{e}_t = \frac{1}{N} \mathrm{tr}(\Sigma_{ee} S_N' \Sigma_{ee}^{-1} S_N) + o_p(1),$$

$$(d) \quad \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{J}_t' \ddot{M} \dot{J}_s \pi_{st} = o_p(1),$$

$$(e) \quad \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{e}_t' S_N' \ddot{M} (\ddot{B}_s + \dot{J}_s) \pi_{st} = o_p(1),$$

$$(f) \quad \frac{1}{NT} \sum_{t=1}^{T} \sum_{s=1}^{T} \dot{e}_t S_N' \ddot{M} S_N \dot{e}_s \pi_{st} = o_p(1).$$

Proof of Lemma F.1. Consider (a). The left hand side is equal to

$$\frac{1}{NT} \sum_{t=1}^{T} J_t' \Sigma_{ee}^{-1} J_t + \frac{1}{N^2 T} \sum_{t=1}^{T} J_t' \Sigma_{ee}^{-1} \Lambda \Lambda' \Sigma_{ee}^{-1} J_t - \frac{1}{N} \bar{J}' \ddot{M} \bar{J}. \tag{F.1}$$

For the second term, it's expectation is equal to

$$\text{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T}\Lambda'\Sigma_{ee}^{-1}E(J_tJ_t')\Sigma_{ee}^{-1}\Lambda\Big] = \text{tr}\Big[\frac{1}{N^2}\Lambda'\Sigma_{ee}^{-1}\Big(S_N\sum_{l=1}^{\infty}(\delta G_N)^l\Sigma_{ee}(\delta G_N')^l S_N'\Big)\Sigma_{ee}^{-1}\Lambda\Big].$$

by the definition of J_t. Since

$$\begin{aligned}\|E(J_tJ_t')\|_1 &= \Big\|S_N\sum_{l=1}^{\infty}(\delta G_N)^l\Sigma_{ee}(\delta G_N')^l S_N'\Big\|_1 \\ &\leq \|S_N\|_1\cdot\|S_N\|_\infty\cdot\|\Sigma_{ee}\|_1\Big[\sum_{l=1}^{\infty}(|\delta|\cdot\|G_N\|_\infty)^l\Big]\Big[\sum_{l=1}^{\infty}(|\delta|\cdot\|G_N\|_1)^l\Big],\end{aligned} \tag{F.2}$$

which is bounded by some constant C by Assumption F, there exists a $C > 0$, such that

$$\Sigma_{ee}^{-1/2}S_N\sum_{l=1}^{\infty}(\delta G_N)^l\Sigma_{ee}(\delta G_N')^l S_N'\Sigma_{ee}^{-1/2} \leq CI_N.$$

So we have

$$\text{tr}\Big[\frac{1}{N^2T}\sum_{t=1}^{T}\Lambda'\Sigma_{ee}^{-1}E(J_tJ_t')\Sigma_{ee}^{-1}\Lambda\Big] = O(\frac{1}{N}).$$

By the Markov's inequality, we have that the second term of (F.1) is $O_p(N^{-1})$. Consider the third term. Notice that

$$0 \leq \frac{1}{N}\bar{J}'\ddot{M}\bar{J} \leq \frac{1}{N}\bar{J}'\Sigma_{ee}^{-1}\bar{J} \leq C\frac{1}{N}\bar{J}'\bar{J} = O_p(\frac{1}{T}).$$

where the last result is due to $E(N^{-1}\bar{J}'\bar{J}) = O(T^{-1})$, which can be proved similarly as the second term with (F.2). Given these two results, we have (a).

Consider (b). The left hand side of (b) is equal to $\frac{1}{NT}\sum_{t=1}^{T}e_t'S_N'\ddot{M}(\ddot{B}_t + \dot{J}_t)$, which can be written as

$$\frac{1}{NT}\sum_{t=1}^{T}e_t'S_N'\ddot{M}\ddot{B}_t + \frac{1}{NT}\sum_{t=1}^{T}e_t'S_N'\ddot{M}J_t - \frac{1}{N}\sum_{t=1}^{T}\bar{e}'S_N'\ddot{M}\bar{J}$$

Notice that

$$E\Big[\frac{1}{NT}\sum_{t=1}^{T}e_t'S_N'\ddot{M}\ddot{B}_t\Big]^2 = \frac{1}{N^2T^2}\sum_{t=1}^{T}\ddot{B}_t'\ddot{M}S_N\Sigma_{ee}S_N'\ddot{M}\ddot{B}_t \leq C\frac{1}{N^2T^2}\sum_{t=1}^{T}\ddot{B}_t'\ddot{B}_t = O(\frac{1}{NT}).$$

where the inequality is due to the boundedness of $\|\ddot{M}S_N\Sigma_{ee}S_N'\ddot{M}\|_1$ and $\|\ddot{M}S_N\Sigma_{ee}S_N'\ddot{M}\|_\infty$. So the first term is $O_p(\frac{1}{\sqrt{NT}})$. Similarly,

$$E\Big[\frac{1}{NT}\sum_{t=1}^{T}e_t'S_N'\ddot{M}J_t\Big]^2 = \frac{1}{N^2T}\text{tr}\Big[\Sigma_{ee}^{1/2}S_N'\ddot{M}E(J_tJ_t')\ddot{M}S_N\Sigma_{ee}^{1/2}\Big] = O(\frac{1}{NT})$$

where the last equality is due to (F.2). Then the second term is $O_p(\frac{1}{\sqrt{NT}})$. The last term is $O_p(T^{-1})$ which can be easily proved. Then we have (b).

Consider (c). The left hand side is equal to

$$\frac{1}{NT}\sum_{t=1}^{T}e_t'S_N'\Sigma_{ee}^{-1}S_Ne_t - \frac{1}{N^2T}\sum_{t=1}^{T}e_t'S_N'\Sigma_{ee}^{-1}\Lambda\Lambda'\Sigma_{ee}^{-1}S_Ne_t - \frac{1}{N}\bar{e}'S_N'\ddot{M}S_N\bar{e}.$$

For the first term, we see that

$$E\Big[\frac{1}{NT}\sum_{t=1}^{T} e_t' S_N' \Sigma_{ee}^{-1} S_N e_t\Big] = \frac{1}{N}\mathrm{tr}(\Sigma_{ee} S_N' \Sigma_{ee}^{-1} S_N).$$

In addition, we also have that

$$\mathrm{var}\Big[\frac{1}{NT}\sum_{t=1}^{T} e_t' S_N' \Sigma_{ee}^{-1} S_N e_t\Big] = \frac{2}{N^2 T}\mathrm{tr}[(\Sigma_{ee} S_N' \Sigma_{ee}^{-1} S_N)^2]$$
$$+\frac{1}{N^2 T}\frac{\kappa_4 - 3\sigma^4}{\sigma^4}\mathrm{tr}[(\Sigma_{ee} S_N' \Sigma_{ee}^{-1} S_N)\circ(\Sigma_{ee} S_N' \Sigma_{ee}^{-1} S_N)] = O(\frac{1}{NT}).$$

So we have

$$\frac{1}{NT}\sum_{t=1}^{T} e_t' S_N' \Sigma_{ee}^{-1} S_N e_t = \frac{1}{N}\mathrm{tr}(\Sigma_{ee} S_N' \Sigma_{ee}^{-1} S_N) + O_p(\frac{1}{\sqrt{NT}}).$$

Consider the second term, which is bounded in norm by

$$\frac{1}{T}\sum_{t=1}^{T}\Big\|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\lambda_i \ddot{e}_{it}\|^2 = O_p(\frac{1}{N}),$$

where $\ddot{e}_{it} = \sum_{o=1}^{N} S_{io,N} e_{ot}$. Consider the third term. Similarly, there is a constant C such that $S_N' \ddot{M} S_N \leq C \cdot I_N$. Given this, we have $\frac{1}{N}\bar{e}' S_N' \ddot{M} S_N \bar{e} \leq C\frac{1}{N}\bar{e}'\bar{e} = O_p(T^{-1})$. Given these three results, we have (c).

Consider (d). The left hand side can be written as

$$\mathrm{tr}\Big[\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}(F'F)^{-1/2} f_t e_t' \ddot{M} e_s f_s' (F'F)^{-1/2}\Big]$$
$$\leq C\mathrm{tr}\Big[\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}(F'F)^{-1/2} f_t e_t' e_s f_s' (F'F)^{-1/2}\Big] = O_p(T^{-1}).$$

Then (d) follows.

Result (e) can be proved similarly as result (b) and result (f) can be proved similarly as result (d). The details are therefore omitted. This completes the proof of Lemma F.1. □

Lemma F.2 *Let $\mathcal{A}_{t-1}$ be defined in (F.4) below and $\mathcal{A}_{it-1}$ its i-th element. Under Assumptions A-F, we have $E(|\mathcal{A}_{it-1}|^{2+c}) \leq C$ for some $c > 0$ for all i, t.*

Proof of Lemma F.2. Let $\mathbf{v}_i$ be the ith column of the N-dimensional identity matrix. By definition, we have

$$\mathcal{A}_{it-1} = \mathbf{i}_1 \ddot{B}_t' \ddot{M}\mathbf{v}_i - \mathbf{i}_1 \sum_{s=1}^{T}\pi_{st}\ddot{B}_s' \ddot{M}\mathbf{v}_i + \mathbf{i}_1 J_t' \Sigma_{ee}^{-1}\mathbf{v}_i + \mathbf{i}_2 \dot{B}_{t-1}' \ddot{M}\mathbf{v}_i$$
$$- \mathbf{i}_2 \sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}' \ddot{M}\mathbf{v}_i + \mathbf{i}_2 Q_{t-1}' \Sigma_{ee}^{-1}\mathbf{v}_i + \mathbf{i}_3' \dot{X}_t' \ddot{M}\mathbf{v}_i - \mathbf{i}_3' \sum_{s=1}^{T}\pi_{st}\dot{X}_s' \ddot{M}\mathbf{v}_i$$

Then it follows that

$$|\mathcal{A}_{it-1}|^{2+c} \leq 8^{2+c}\Bigg[|\mathbf{i}_1\ddot{B}_t'\ddot{M}\mathbf{v}_i|^{2+c} + |\mathbf{i}_1\sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M}\mathbf{v}_i|^{2+c} + |\mathbf{i}_1 J_t'\Sigma_{ee}^{-1}\mathbf{v}_i|^{2+c} + |\mathbf{i}_2\dot{B}_{t-1}'\ddot{M}\mathbf{v}_i|^{2+c}$$
$$+|\mathbf{i}_2\sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M}\mathbf{v}_i|^{2+c} + |\mathbf{i}_2 Q_{t-1}'\Sigma_{ee}^{-1}\mathbf{v}_i|^{2+c} + |\mathbf{i}_3'\dot{X}_t'\ddot{M}\mathbf{v}_i|^{2+c} + |\mathbf{i}_3'\sum_{s=1}^{T}\pi_{st}\dot{X}_s'\ddot{M}\mathbf{v}_i|^{2+c}\Bigg]. \tag{F.3}$$

So it suffices to show that the expectations of the above eight terms on right hand side are bounded. Since the proofs for different $\mathbf{i}_j$ ($j = 1, 2, 3$) are similar, we only choose the three terms involving $\mathbf{i}_1$ to prove. Consider the first term. If x_{it} and f_t are bounded fixed values, the proof is easy. We only consider the random case. Notice that

$$\begin{aligned} E(|\ddot{B}_t'\ddot{M}\mathbf{v}_i|^{2+c}) &= E[(\mathbf{v}_i'\ddot{M}\ddot{B}_t\ddot{B}_t'\ddot{M}\mathbf{v}_i)^{1+\frac{c}{2}}] \leq E[(\mathbf{v}_i'\Sigma_{ee}^{-1}\ddot{B}_t\ddot{B}_t'\Sigma_{ee}^{-1}\mathbf{v}_i)^{1+\frac{c}{2}}] \\ &\leq CE[(\mathbf{v}_i'\ddot{B}_t\ddot{B}_t'\mathbf{v}_i)^{1+\frac{c}{2}}] \leq CE[|\mathbf{v}_i'\ddot{B}_t|^{2+c}] = CE[|\ddot{B}_{it}|^{2+c}]. \end{aligned}$$

Let $\widetilde{G}_l = (\delta G_N)^l G_N$ and $\widetilde{G}_{ij,l}$ be its (i, j)th element. By Assumption F, it is easy to verify that $\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\widetilde{G}_{ij,l}| \leq C$ for all i and $\sum_{l=0}^{\infty}\sum_{i=1}^{N}|\widetilde{G}_{ij,l}| \leq C$ for all j. Now, by definition,

$$\ddot{B}_{it} = \sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}\beta_p\widetilde{G}_{ij,l}\dot{x}_{jtp} + \sum_{l=0}^{\infty}\sum_{j=1}^{N}\widetilde{G}_{ij,l}\lambda_j'\dot{f}_{t-l}.$$

So we have

$$E[|\ddot{B}_{it}|^{2+c}] \leq 2^{2+c}E\Big[\Big|\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}\beta_p\widetilde{G}_{ij,l}\dot{x}_{jtp}\Big|^{2+c}\Big] + 2^{2+c}E\Big[\Big|\sum_{l=0}^{\infty}\sum_{j=1}^{N}\widetilde{G}_{ij,l}\lambda_j'\dot{f}_{t-l}\Big|^{2+c}\Big]$$

It suffices to show that the two terms on right hand side are bounded. The proofs of the two terms are similar. So we only choose the first term to prove. Notice that

$$\Big|\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}\beta_p\widetilde{G}_{ij,l}\dot{x}_{jtp}\Big| \leq \sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\beta_p\widetilde{G}_{ij,l}|\cdot|\dot{x}_{jtp}|$$

Let $\breve{G}_i = \sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\beta_p\widetilde{G}_{ij,l}|$. Then we have

$$\Bigg[\frac{1}{\breve{G}_i}\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\beta_p\widetilde{G}_{ij,l}|\cdot|\dot{x}_{jtp}|\Bigg]^{2+c} \leq \frac{1}{\breve{G}_i}\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\beta_p\widetilde{G}_{ij,l}|\cdot|\dot{x}_{jtp}|^{2+c}$$

by the Jensen's inequality. Then it follows that

$$\Bigg[\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\beta_p\widetilde{G}_{ij,l}|\cdot|\dot{x}_{jtp}|\Bigg]^{2+c} \leq \breve{G}_i^{1+c}\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\beta_p\widetilde{G}_{ij,l}|\cdot|\dot{x}_{jtp}|^{2+c}$$

Thus,

$$E\Big[\Big|\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}\beta_p\widetilde{G}_{ij,l}\dot{x}_{jtp}\Big|^{2+c} \leq \breve{G}_i^{1+c}\sum_{p=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\beta_p\widetilde{G}_{ij,l}|E(|\dot{x}_{jtp}|^{2+c}) \leq C\breve{G}_i^{2+c}$$

which is bounded since $\sum_{l=0}^{\infty}\sum_{j=1}^{N}|\widetilde{G}_{ij,l}| \le C$. This completes the proof for the first term on the right hand side of (F.3). By treating $\sum_{s=1}^{T}\pi_{st}\ddot{B}_s$ as a new $\ddot{B}_t$, the proof of the second term is similar as the first term. As regard the third term, by the similar arguments in the proof of the first term, we can show that

$$E(|J_t'\Sigma_{ee}^{-1}\mathbf{v}_i|^{2+c}) \le CE(|\mathbf{v}_i'J_t|^{2+c}).$$

The remaining proof is therefore similar as that of the first term by treating e_{t-l} as a new $\dot{X}_{t-l}\beta$. So we have proved that the terms involving $\mathbf{i}_1$ in (F.3) are bounded. The same arguments also apply to the terms involving $\mathbf{i}_2$ and $\mathbf{i}_3$. This completes the proof. □

Proof of Corollary 5.1. As defined in Theorem 5.2,

$$\xi = \frac{1}{\sqrt{NT}}\begin{bmatrix} \sum_{t=1}^{T}\ddot{B}_t'\ddot{M}e_t - \sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\ddot{M}e_s\pi_{st} + \sum_{t=1}^{T}J_t'\Sigma_{ee}^{-1}e_t + \eta \\ \sum_{t=1}^{T}\dot{B}_{t-1}'\ddot{M}e_t - \sum_{t=1}^{T}\sum_{s=1}^{T}\dot{B}_{t-1}'\ddot{M}e_s\pi_{st} + \sum_{t=1}^{T}Q_{t-1}'\Sigma_{ee}^{-1}e_t \\ \sum_{t=1}^{T}\dot{X}_t'\ddot{M}e_t - \sum_{t=1}^{T}\sum_{s=1}^{T}\dot{X}_t'\ddot{M}e_s\pi_{st}\end{bmatrix}.$$

We use the Cramér-Wold device to show that ξ converges in distribution to a multivariate normal distribution. For any nonrandom $(k+2)$-dimensional vector $\ell = (\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3')'$, where $\mathbf{i}_3$ is k-dimension, Consider $\ell'\xi$, which is equal to

$$\begin{aligned}\ell'\xi = \mathbf{i}_1\frac{1}{\sqrt{NT}}&\left\{\sum_{t=1}^{T}\ddot{B}_t'\ddot{M}e_t - \sum_{t=1}^{T}\sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M}e_t + \sum_{t=1}^{T}J_t'\Sigma_{ee}^{-1}e_t + \sum_{t=1}^{T}e_t'S_N^{\circ\prime}\Sigma_{ee}^{-1}e_t\right\} \\ &+ \mathbf{i}_2\frac{1}{\sqrt{NT}}\left\{\sum_{t=1}^{T}\dot{B}_{t-1}'\ddot{M}e_t - \sum_{t=1}^{T}\sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M}e_t + \sum_{t=1}^{T}Q_{t-1}'\Sigma_{ee}^{-1}e_t\right\} \\ &+ \mathbf{i}_3'\frac{1}{\sqrt{NT}}\left\{\sum_{t=1}^{T}\dot{X}_t'\ddot{M}e_t - \sum_{t=1}^{T}\sum_{s=1}^{T}\pi_{st}\dot{X}_s'\ddot{M}e_t\right\}\end{aligned}$$

Let $\mathcal{A}_{t-1}$ be defined as

$$\begin{aligned}\mathcal{A}_{t-1}' &= \mathbf{i}_1\Big[\ddot{B}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M} + J_t'\Sigma_{ee}^{-1}\Big] + \mathbf{i}_2\Big[\dot{B}_{t-1}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M} + Q_{t-1}'\Sigma_{ee}^{-1}\Big] \\ &\quad + \mathbf{i}_3'\Big[\dot{X}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{X}_s'\ddot{M}\Big] \\ &= \ell'\begin{bmatrix}\ddot{B}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M} + J_t'\Sigma_{ee}^{-1} \\ \dot{B}_{t-1}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M} + Q_{t-1}'\Sigma_{ee}^{-1} \\ \dot{X}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{X}_s'\ddot{M}\end{bmatrix} \qquad \text{(F.4)}\end{aligned}$$

and $\mathcal{E} = \mathbf{i}_1(S_N^{\circ\prime}\Sigma_{ee}^{-1} + \Sigma_{ee}^{-1}S_N^{\circ})/2$. By definition, it is easy to see $\|\mathcal{E}\|_\infty < \infty$. Then we can rewrite the above result as

$$\begin{aligned}\ell'\xi &= \frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\mathcal{A}_{t-1}'e_t + \frac{1}{\sqrt{NT}}\sum_{t=1}^{T}e_t'\mathcal{E}e_t \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathcal{A}_{it-1}e_{it} + 2\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}e_{it}\Big(\sum_{j=1}^{i-1}\mathcal{E}_{ij}e_{jt}\Big).\end{aligned}$$

Now we apply the martingale central limit theorem theorem to show that $\ell'\xi$ converges to a normal distribution. Let

$$z_{ti} = \frac{1}{\sqrt{N\bar{T}}}\Big[\mathcal{A}_{i,t-1}e_{it} + 2\Big(\sum_{j=1}^{i-1}\mathcal{E}_{ij}e_{jt}\Big)e_{it}\Big],$$

$$\mathcal{V}_{NT} = \Big[\frac{1}{N\bar{T}}\sum_{t=1}^{T}\mathcal{A}'_{t-1}\Sigma_{ee}\mathcal{A}_{t-1} + 2\frac{1}{N}\text{tr}(\Sigma_{ee}\mathcal{E}'\Sigma_{ee}\mathcal{E})\Big].$$

Let $\mathscr{F}_{ti}$ be the σ-field generated by

$$\mathscr{F}_{ti} = \sigma\Big\{e_{11},\ldots,e_{N1},e_{12},\ldots,e_{N2},\ldots,e_{1t-1},\ldots,e_{Nt-1},e_{1t},e_{2t},\ldots,e_{it}\Big\}.$$

Then $\mathscr{F}_{10},\mathscr{F}_{11},\ldots,\mathscr{F}_{1N},\mathscr{F}_{21},\ldots,\mathscr{F}_{2N},\ldots,\mathscr{F}_{T1},\ldots,\mathscr{F}_{TN}$ form a sequence of increasing σ-fields with $\mathscr{F}_{10} = \varnothing$. Given the above definition, it is easy to verify that $\mathbb{E}(z_{ti}|\mathscr{F}_{t,i-1}) = 0$. So $\{z_{ti},\mathscr{F}_{ti-1}\}$ forms a martingale difference array. According to Corollary 3.1 in Hall and Heyde (1980), we have $\sum_{i=1}^{N}\sum_{t=1}^{T} z_{ti}/\sqrt{\mathcal{V}_{NT}} \xrightarrow{d} N(0,1)$ if we can show that any $\epsilon > 0$,

$$\sum_{t=1}^{T}\sum_{i=1}^{N} E[z_{ti}^2(|z_{ti}| > \epsilon)|\mathscr{F}_{t,i-1}] \xrightarrow{p} 0 \tag{F.5}$$

and

$$\sum_{t=1}^{T}\sum_{i=1}^{N} E(z_{ti}^2|\mathscr{F}_{t,i-1}) - \mathcal{V}_{NT} \xrightarrow{p} 0 \tag{F.6}$$

Let $\breve{z}_{ti} = \mathcal{A}_{i,t-1}e_{it} + 2(\sum_{j=1}^{i-1}\mathcal{E}_{ij}e_{jt})e_{it}$, i.e., $z_{ti} = \breve{z}_{ti}/\sqrt{NT}$. A sufficient condition for (F.5) is $\mathbb{E}(\breve{z}_{ti}^{2+\delta}) \le C$ for some constant C for all i and t. To see this, notice that

$$\begin{aligned}E[z_{ti}^2(|z_{ti}| > \epsilon)] &= \int_{|z_{ti}|>\epsilon} z_{ti}^2 d\mathbb{P} \le \frac{1}{\epsilon^\delta}\int_{|z_{ti}|>\epsilon}|z_{ti}|^{2+\delta}d\mathbb{P}\\ &\le \frac{1}{\epsilon^\delta}\int |z_{ti}|^{2+\delta}d\mathbb{P} = \frac{1}{\epsilon^\delta}E(|z_{ti}|^{2+\delta}).\end{aligned}$$

Given this result,

$$\begin{aligned}E\left[\sum_{t=1}^{T}\sum_{i=1}^{N}E[z_{ti}^2(|z_{ti}| > \epsilon)|\mathscr{F}_{t,i-1}]\right] &= \sum_{t=1}^{T}\sum_{i=1}^{N}E[z_{ti}^2(|z_{ti}| > \epsilon)] \le \frac{1}{\epsilon^\delta}\sum_{t=1}^{T}\sum_{i=1}^{N}E(|z_{ti}|^{2+\delta})\\ &= \frac{1}{\epsilon^\delta}\frac{1}{(N\bar{T})^{1+\delta/2}}\sum_{t=1}^{T}\sum_{i=1}^{N}E(|\breve{z}_{ti}|^{2+\delta}) = O((NT)^{-\delta/2}).\end{aligned}$$

Thus, (F.5) follows by the markov inequality. Now consider $\breve{z}_{ti} = \mathcal{A}_{i,t-1}e_{it} + 2(\sum_{j=1}^{i-1}\mathcal{E}_{ij}e_{jt})e_{it}$. Let $u = \frac{2+\delta}{1+\delta}$ and $v = 2+\delta$, it is seen that $u^{-1} + v^{-1} = 1$. Notice that $|\breve{z}_{ti}|$ is bounded by

$$|\breve{z}_{ti}| \le |\mathcal{A}_{i,t-1}|\cdot|e_{it}| + \sum_{j=1}^{i-1}(|\mathcal{E}_{ij}|^{\frac{1}{u}}|e_{jt}|)\cdot(2|\mathcal{E}_{ij}|^{\frac{1}{v}}|e_{it}|).$$

By the Hölder inequality

$$\sum_i a_i b_i \le \Big(\sum_i |a_i|^u\Big)^{\frac{1}{u}}\Big(\sum_i |b_i|^v\Big)^{\frac{1}{v}},$$

we can further bound the preceding expression by

$$|\breve{z}_{ti}| \leq \left[|\mathcal{A}_{i,t-1}|^u + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}||e_{jt}|^u\right]^{\frac{1}{u}}\left[|e_{it}|^v + \sum_{j=1}^{i-1}2^v|\mathcal{E}_{ij}||e_{it}|^v\right]^{\frac{1}{v}}$$

or equivalently

$$|\breve{z}_{ti}|^v \leq \left[|\mathcal{A}_{i,t-1}|^u + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}||e_{jt}|^u\right]^{\frac{v}{u}}\left[|e_{it}|^v + \sum_{j=1}^{i-1}2^v|\mathcal{E}_{ij}||e_{it}|^v\right].$$

By $E(|\breve{z}_{it}|^v) = E\big[E(|\breve{z}_{it}|^v|\mathscr{F}_{t,i-1})\big]$, we have

$$E(|\breve{z}_{it}|^v) \leq E\left[|\mathcal{A}_{i,t-1}|^u + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}||e_{jt}|^u\right]^{\frac{v}{u}}\left[E(|e_{it}|^v) + \sum_{j=1}^{i-1}2^v|\mathcal{E}_{ij}|\cdot E(|e_{it}|^v)\right] \tag{F.7}$$

Since $E(|e_{it}|^8) < \infty$ by Assumption A, together with the boundedness of $\|\mathcal{E}\|_\infty$, we have

$$E(|e_{it}|^v) + \sum_{j=1}^{i-1}2^v|\mathcal{E}_{ij}|\cdot E(|e_{it}|^v) \leq C$$

for some constant C. Proceed to consider the first factor on the right hand side of (F.7). Since $f(x) = x^{v/u}$ is a convex function for $v \geq u$, it follows that

$$\Big[\sum_i \omega_i|x_i|^u\Big]^{\frac{1}{u}} \leq \Big[\sum_i \omega_i|x_i|^v\Big]^{\frac{1}{v}}$$

by the Jensen inequality, where $\omega_i \geq 0$ and $\sum_i \omega_i = 1$. Now let $\varsigma_i = 1 + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}|$, then

$$\left[\frac{1}{\varsigma_i}\Big(1\cdot|\mathcal{A}_{i,t-1}|^u + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}|\cdot|e_{jt}|^u\Big)\right]^{\frac{1}{u}} \leq \left[\frac{1}{\varsigma_i}\Big(1\cdot|\mathcal{A}_{i,t-1}|^v + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}|\cdot|e_{jt}|^v\Big)\right]^{\frac{1}{v}}$$

or equivalently

$$\left[|\mathcal{A}_{i,t-1}|^u + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}||e_{jt}|^u\right]^{\frac{v}{u}} \leq \varsigma_i^{\frac{v-u}{u}}\left[|\mathcal{A}_{i,t-1}|^v + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}|\cdot|e_{jt}|^v\right]$$

Since $\varsigma_i = 1 + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}|$ is bounded by $1 + \|\mathcal{E}\|_\infty$, which is further bounded by some constant C. Thus,

$$E\left[|\mathcal{A}_{i,t-1}|^u + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}||e_{jt}|^u\right]^{\frac{v}{u}} \leq C\left[E(|\mathcal{A}_{i,t-1}|^v) + \sum_{j=1}^{i-1}|\mathcal{E}_{ij}|\cdot\mathbb{E}(|e_{jt}|^v)\right]$$

By $E(|e_{it}|^8) < \infty$, there exists a constant C such that $\mathbb{E}(|e_{jt}|^v) < C$. Then the last two terms of the preceding display is bounded. This result, together with $E(|\mathcal{A}_{i,t-1}|^v) < \infty$

which is given in Lemma F.2, gives that the first factor on the right hand side of (F.7) is bounded for all i and t. Then by (F.7), we have $\mathbb{E}(|\breve{z}_{ti}|^v) < C$ for some $v > 2$ for all i and t. Then (F.5) follows.

Consider (F.6). By the definition of z_{ti} and $\mathscr{F}_{ti-1}$, it is seen that

$$\sum_{t=1}^{T}\sum_{i=1}^{N} E(z_{ti}^2|\mathscr{F}_{t,i-1}) = \frac{1}{NT}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\mathcal{A}_{t-1} + 4\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sigma_i^2\Big(\sum_{j=1}^{i-1}\mathcal{E}_{ij}e_{jt}\Big)^2 + 4\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sigma_i^2\mathcal{A}_{it-1}\Big(\sum_{j=1}^{i-1}\mathcal{E}_{ij}e_{jt}\Big)$$

Let $\widetilde{\mathcal{E}}$ be the matrix obtained by setting the elements above the diagonal to zeros. By definition, we have $\mathcal{E} = \widetilde{\mathcal{E}} + \widetilde{\mathcal{E}}'$. Then we can rewrite the above expression as

$$\sum_{t=1}^{T}\sum_{i=1}^{N} E(z_{ti}^2|\mathscr{F}_{t,i-1}) = \frac{1}{NT}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\mathcal{A}_{t-1} + 4\frac{1}{NT}\sum_{t=1}^{T}e_t'\widetilde{\mathcal{E}}'\Sigma_{ee}\widetilde{\mathcal{E}}e_t + 4\frac{1}{NT}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\widetilde{\mathcal{E}}e_t \tag{F.8}$$

For the second term, we have

$$4E\Big[\frac{1}{NT}\sum_{t=1}^{T}e_t'\widetilde{\mathcal{E}}'\Sigma_{ee}\widetilde{\mathcal{E}}e_t\Big] = 4\frac{1}{N}\mathrm{tr}(\Sigma_{ee}\widetilde{\mathcal{E}}'\Sigma_{ee}\widetilde{\mathcal{E}}) = 2\frac{1}{N}\mathrm{tr}(\Sigma_{ee}\mathcal{E}'\Sigma_{ee}\mathcal{E})$$

where we use the fact that $\mathrm{tr}(\Sigma_{ee}\widetilde{\mathcal{E}}\Sigma_{ee}\widetilde{\mathcal{E}}) = 0$, which is due to $\widetilde{\mathcal{E}}$ being a low-triangular matrix with diagonal elements zeros. By the well-known result that

$$\mathrm{var}(\boldsymbol{v}_t'A\boldsymbol{v}_t) = \mathrm{tr}(A^2) + \mathrm{tr}(A'A) + \kappa\mathrm{tr}(A\circ A)$$

where "$\circ$" denotes the Hadamard product and $\boldsymbol{v}_t$ are iid over t with zero mean and identity variance matrix, and the elements of v_t are also iid with the fourth moment $3+\kappa$, we have

$$\mathrm{var}\Big[\frac{1}{NT}\sum_{t=1}^{T}e_t'\widetilde{\mathcal{E}}'\Sigma_{ee}\widetilde{\mathcal{E}}e_t\Big] = 2\frac{1}{N^2T}\mathrm{tr}\Big[(\Sigma_{ee}\mathcal{E}'\Sigma_{ee}\mathcal{E})^2\Big]$$

which is $O_p(\frac{1}{N\sqrt{T}})$. So we have

$$4\frac{1}{NT}\sum_{t=1}^{T}e_t'\widetilde{\mathcal{E}}'\Sigma_{ee}\widetilde{\mathcal{E}}e_t = 2\frac{1}{N}\mathrm{tr}(\Sigma_{ee}\mathcal{E}'\Sigma_{ee}\mathcal{E}) + O_p(\frac{1}{\sqrt{NT}}). \tag{F.9}$$

For the third term of (F.8), we see

$$E\Big[\frac{1}{NT}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\widetilde{\mathcal{E}}e_t\Big]^2 = \frac{1}{N^2T^2}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\widetilde{\mathcal{E}}\Sigma_{ee}\widetilde{\mathcal{E}}'\Sigma_{ee}\mathcal{A}_{t-1} \le \frac{C}{N^2T^2}\sum_{t=1}^{T}E(\mathcal{A}_{t-1}'\Sigma_{ee}\mathcal{A}_{t-1})$$

where we use the fact that $\Sigma_{ee}^{1/2}\widetilde{\mathcal{E}}\Sigma_{ee}\widetilde{\mathcal{E}}'\Sigma_{ee}^{1/2} \le C\cdot I_N$ for some constant C. As will be shown below, $\frac{1}{NT}\sum_{t=1}^{T}E(\mathcal{A}_{t-1}'\Sigma_{ee}\mathcal{A}_{t-1}) = O(1)$. So we have

$$\frac{1}{NT}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\widetilde{\mathcal{E}}e_t = O_p(\frac{1}{\sqrt{NT}}) \tag{F.10}$$

Given (F.8), (F.9) and (F.10), we have

$$\sum_{t=1}^{T}\sum_{i=1}^{N} E(z_{ti}^2|\mathscr{F}_{t,i-1}) = \frac{1}{NT}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\mathcal{A}_{t-1} + 2\frac{1}{N}\mathrm{tr}(\Sigma_{ee}\mathcal{E}'\Sigma_{ee}\mathcal{E}) + o_p(1) = \mathcal{V}_{NT} + o_p(1).$$

Therefore, (F.6) is proved. Given (F.5) and (F.6), we have

$$\frac{\ell'\xi}{\sqrt{\mathcal{V}_{NT}}} \xrightarrow{d} N(0,1).$$

Given the above result, if we can show that $\mathcal{V}_{NT} - \ell'\mathbb{D}\ell = o_p(1)$, then by the Slutsky's lemma and the Cramér-Wold theorem, we have

$$\mathbb{D}^{-1/2}\xi \xrightarrow{d} N(0, I_{k+2}). \tag{F.11}$$

Now consider the expression of $\mathcal{V}_{NT}$. By $\mathcal{E} = \mathbf{i}_1(S_N^{\circ\prime}\Sigma_{ee}^{-1} + \Sigma_{ee}^{-1}S_N^{\circ})/2$, the expression of $\mathcal{V}_{NT}$ can be alternatively written as

$$\mathcal{V}_{NT} = \frac{1}{NT}\sum_{t=1}^{T}\mathcal{A}_{t-1}'\Sigma_{ee}\mathcal{A}_{t-1} + \mathbf{i}_1^2\left\{\frac{1}{N}\mathrm{tr}(\Sigma_{ee}S_N^{\circ\prime}\Sigma_{ee}^{-1}S_N^{\circ}) + \frac{1}{N}\mathrm{tr}(S_N^{\circ 2})\right\}$$

By (F.4), we can further rewrite $\mathcal{V}_{NT}$ as

$$\mathcal{V}_{NT} = \ell'\left\{\frac{1}{NT}\sum_{t=1}^{T}\begin{bmatrix}\ddot{B}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M} + J_t'\Sigma_{ee}^{-1}\\ \dot{B}_{t-1}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M} + Q_{t-1}'\Sigma_{ee}^{-1}\\ \dot{X}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{X}_s'\ddot{M}\end{bmatrix}\Sigma_{ee}\right.$$
$$\left.\times\begin{bmatrix}\ddot{B}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M} + J_t'\Sigma_{ee}^{-1}\\ \dot{B}_{t-1}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M} + Q_{t-1}'\Sigma_{ee}^{-1}\\ \dot{X}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{X}_s'\ddot{M}\end{bmatrix}'\right\}\ell + \mathbf{i}_1^2\left\{\frac{1}{N}\mathrm{tr}(\Sigma_{ee}S_N^{\circ\prime}\Sigma_{ee}^{-1}S_N^{\circ}) + \frac{1}{N}\mathrm{tr}(S_N^{\circ 2})\right\}.$$

So, to complete the proof, we need to show

$$\frac{1}{NT}\mathrm{tr}[\ddot{Y}'\ddot{M}\ddot{Y}M_F] + \frac{1}{N}\left[\mathrm{tr}(S_N^2) - 2\sum_{i=1}^{N}S_{ii,N}^2\right] = \frac{1}{N}\mathrm{tr}(\Sigma_{ee}S_N^{\circ\prime}\Sigma_{ee}^{-1}S_N^{\circ}) + \frac{1}{N}\mathrm{tr}(S_N^{\circ 2}) \tag{F.12}$$
$$+ \frac{1}{NT}\sum_{t=1}^{T}\left[\ddot{B}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M} + J_t'\Sigma_{ee}^{-1}\right]\Sigma_{ee}\left[\ddot{B}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M} + J_t'\Sigma_{ee}^{-1}\right]' + o_p(1),$$

$$\frac{1}{NT}\mathrm{tr}[\dot{Y}_{-1}'\ddot{M}\dot{Y}_{-1}M_F] = \left\{\frac{1}{NT}\sum_{t=1}^{T}\left[\dot{B}_{t-1}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M} + Q_{t-1}'\Sigma_{ee}^{-1}\right]\right.$$
$$\left.\times\Sigma_{ee}\left[\dot{B}_{t-1}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M} + Q_{t-1}'\Sigma_{ee}^{-1}\right]'\right\} + o_p(1),$$

$$\frac{1}{NT}\mathrm{tr}[\dot{X}_p'\ddot{M}\dot{X}_q'M_F] = \frac{1}{NT}\sum_{t=1}^{T}\left[\dot{X}_{tp}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{X}_{sp}'\ddot{M}\right]\Sigma_{ee}\left[\dot{X}_{tq}'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\dot{X}_{sq}'\ddot{M}\right]' + o_p(1),$$

$$\frac{1}{NT}\mathrm{tr}[\ddot{Y}'\ddot{M}\dot{Y}_{-1}M_F] = \left\{\frac{1}{NT}\sum_{t=1}^{T}\left[\ddot{B}_t'\ddot{M} - \sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M} + J_t'\Sigma_{ee}^{-1}\right]\right.$$

$$\times\Sigma_{ee}\left[\dot{B}_{t-1}'\ddot{M}-\sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M}+Q_{t-1}'\Sigma_{ee}^{-1}\right]'\Bigg\}+o_p(1),$$

$$\frac{1}{NT}\mathrm{tr}[\ddot{Y}'\ddot{M}\dot{X}_p'M_F]=\frac{1}{NT}\sum_{t=1}^{T}\left[\ddot{B}_t'\ddot{M}-\sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M}+J_t'\Sigma_{ee}^{-1}\right]\Sigma_{ee}\left[\dot{X}_{tp}'\ddot{M}-\sum_{s=1}^{T}\pi_{st}\dot{X}_{sp}'\ddot{M}\right]'+o_p(1),$$

$$\frac{1}{NT}\mathrm{tr}[\dot{Y}_{-1}'\ddot{M}\dot{X}_p'M_F]=\Bigg\{\frac{1}{NT}\sum_{t=1}^{T}\left[\dot{B}_{t-1}'\ddot{M}-\sum_{s=1}^{T}\pi_{st}\dot{B}_{s-1}'\ddot{M}+Q_{t-1}'\Sigma_{ee}^{-1}\right]$$
$$\times\Sigma_{ee}\left[\dot{X}_{tp}'\ddot{M}-\sum_{s=1}^{T}\pi_{st}\dot{X}_{sp}'\ddot{M}\right]'\Bigg\}+o_p(1),$$

where the six expressions on left hand side comes from $\mathbb{D}$ and the six ones on right hand side are the counterparts from $\mathcal{V}_{NT}$. Notice that once we have proven the above six results, we have implicitly shown that $\frac{1}{NT}\sum_{t=1}^{T}E(\mathcal{A}_{t-1}'\Sigma_{ee}\mathcal{A}_{t-1})=O(1)$, as required in the derivation of (F.6). The proofs of the above six results are similar, so we only choose the first one to prove. The right hand side of (F.12) is

$$\mathrm{rhs}=\frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\ddot{M}\ddot{B}_t+2\frac{1}{NT}\sum_{t=1}^{T}J_t'\ddot{M}\ddot{B}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M}\ddot{B}_t+\frac{1}{NT}\sum_{t=1}^{T}J_t'\Sigma_{ee}^{-1}J_t$$
$$-2\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\pi_{st}\ddot{B}_s'\ddot{M}J_t+\frac{1}{N}\mathrm{tr}(\Sigma_{ee}S_N^{\circ\prime}\Sigma_{ee}^{-1}S_N^{\circ})+\frac{1}{N}\mathrm{tr}(S_N^{\circ 2}).$$

By $\ddot{Y}_t=\ddot{B}_t+\dot{J}_t+S_N\dot{e}_t$, the left hand side of (F.12) is equal to

$$\mathrm{lhs}=\frac{1}{NT}\mathrm{tr}[\ddot{Y}'\ddot{M}\ddot{Y}M_F]+\frac{1}{N}\left[\mathrm{tr}(S_N^2)-2\sum_{i=1}^{N}S_{ii,N}^2\right]$$
$$=\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\ddot{M}\ddot{Y}_t-\frac{1}{NT}\sum_{t=1}^{T}\ddot{Y}_t'\ddot{M}\sum_{s=1}^{T}\ddot{Y}_s\pi_{st}+\frac{1}{N}\left[\mathrm{tr}(S_N^2)-2\sum_{i=1}^{N}S_{ii,N}^2\right]$$
$$=\frac{1}{NT}\sum_{t=1}^{T}\ddot{B}_t'\ddot{M}\ddot{B}_t+2\frac{1}{NT}\sum_{t=1}^{T}\dot{J}_t'\ddot{M}\ddot{B}_t+\frac{1}{NT}\sum_{t=1}^{T}\dot{J}_t'\ddot{M}\dot{J}_t+2\frac{1}{NT}\sum_{t=1}^{T}\dot{e}_t'S_N'\ddot{M}(\ddot{B}_t+\dot{J}_t)$$
$$+\frac{1}{NT}\sum_{t=1}^{T}\dot{e}_t'S_N'\ddot{M}S_N\dot{e}_t-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\ddot{M}\ddot{B}_s\pi_{st}-2\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\ddot{B}_t'\ddot{M}\dot{J}_s\pi_{st}$$
$$-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{J}_t'\ddot{M}\dot{J}_s\pi_{st}-2\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{e}_t'S_N'\ddot{M}(\ddot{B}_s+\dot{J}_s)\pi_{st}$$
$$-\frac{1}{NT}\sum_{t=1}^{T}\sum_{s=1}^{T}\dot{e}_t'S_N'\ddot{M}S_N\dot{e}_s\pi_{st}+\frac{1}{N}\mathrm{tr}(\Sigma_{ee}S_N^{\circ\prime}\Sigma_{ee}^{-1}S_N^{\circ})+\frac{1}{N}\mathrm{tr}(S_N^{\circ 2}).$$

Using the results in Lemma F.1, we see that $\mathrm{lhs}=\mathrm{rhs}+o_p(1)$. Thus, (F.11) follows. Given this result, together with Theorem 5.2, we have Corollary 5.1. This completes the proof.