Westerlund, Joakim; Norkute, Milda

Working Paper

A Factor Analytical Method to Interactive Effects Dynamic Panel Models with or without Unit Root

Working Paper, No. 2014:12

Provided in Cooperation with:
Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Westerlund, Joakim; Norkute, Milda (2014) : A Factor Analytical Method to Interactive Effects Dynamic Panel Models with or without Unit Root, Working Paper, No. 2014:12, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:
https://hdl.handle.net/10419/260109

Working Paper 2014:12

Department of Economics
School of Economics and Management

A Factor Analytical Method to Interactive Effects Dynamic Panel Models with or without Unit Root

Joakim Westerlund
Milda Norkute

April 2014

LUND UNIVERSITY

A Factor Analytical Method to Interactive Effects Dynamic Panel Models with or without Unit Root*

Joakim Westerlund
Deakin University
Australia

Milda Norkute
Lund University
Sweden

April 14, 2014

Abstract

In a recent study, Bai (Fixed-Effects Dynamic Panel Models, A Factor Analytical Method. *Econometrica* **81**, 285–314, 2013a) proposes a new factor analytic (FA) method to the estimation of dynamic panel data models, which has the unique and very useful property that it is completely bias-free. However, while certainly appealing, it is restricted to fixed effects models without a unit root. In many situations of practical relevance this is a rather restrictive consideration. The purpose of the current study is therefore to extend the FA approach to cover models with multiple interactive effects and a possible unit root.

JEL Classification: C12; C13; C33; C36.

Keywords: Interactive fixed effects; Dynamic panel data models; Unit root; Factor analytical method.

1 Introduction

Consider the panel data variable $y_{i,t}$, observable for $t = 1, ..., T$ time series and $i = 1, ..., N$ cross-sectional units. The data generating process (DGP) of this variable is assumed to be given by the following dynamic panel data model:

$$y_{i,t} = c_{i,t} + \rho y_{i,t-1} + \varepsilon_{i,t}, \tag{1}$$

*The authors would like to thank Jörg Breitung, David Edgerton and Vasilis Sarafidis for many valuable comments and suggestions.

where $\rho \in (-1, 1]$, $y_{1,0} = ... = y_{N,0} = 0$, $c_{i,t}$ is the common component of the data, and $\varepsilon_{i,t}$ is an error term. Two specifications of $c_{i,t}$ will be considered; (C1) $c_{i,t} = \lambda_i' F_t$, and (C2) $c_{i,t} = \lambda_i'(F_t - \rho F_{t-1})$ for $t = 2, ..., T$ and $c_{i,1} = \lambda_i' F_1$, where F_t is an $m \times 1$ vector of common factors and λ_i is a conformable vector of loading coefficients. Both specifications presume that $m \geq 1$; if $m = 0$, we define $c_{i,t} = 0$. The DGP that arises under C2 can be seen as emanating from $y_{i,t} = \lambda_i' F_t + s_{i,t}$, where $s_{i,t} = \rho s_{i,t-1} + \varepsilon_{i,t}$, which differs only slightly from the more common DGP under C1. Note in particular how C1 and C2 are indistinguishable for $|\rho| < 1$. Since the analysis of C1 is simplest we therefore assume throughout this paper that C1 holds whenever $|\rho| < 1$. The appropriate model to consider under $\rho = 1$ is less obvious and in the present paper we therefore consider both.

Using the terminology of Bai (2009), (1) constitutes a fixed "interactive" effects model under either C1 or C2, which is more general than many of the fixed effects models previously considered in the literature (see Bun and Sarafidis, 2013; Chudik and Pesaran, 2013, for recent surveys). Suppose, for example, that $F_t = 1$, which implies that under C1, $c_{i,t} = \lambda_i' F_t = \lambda_i$. This means that under the additional assumption of $|\rho| < 1$, (1) reduces to what can only be described as the "classical" dynamic panel data model with unit-specific fixed effects, which has attracted considerable attention in the literature. One reason for this is the existence of the so-called "incidental parameters bias", or "Nickell bias" (Nickell, 1981), which arises because of the increasing number of fixed effects. In the classical micro panel setting with T fixed and $N \to \infty$ this bias is a severe problem, as in this case least squares (LS) is inconsistent. This has led to the development of alternative estimators such as the generalized method of moments (GMM) (see, for example, Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998), which is now the most common approach in empirical work with dynamic panels.

This paper focuses on the case when $T \to \infty$, which lessens the problem of bias. However, while consistent, the asymptotic distribution of the LS estimator is still miscentered (see, for example, Hahn and Kuersteiner, 2002). In fact, all estimation approaches known to us are biased in one way or another (see Moon et al., 2013, for an overview of this literature). This includes GMM, which suffers from problems of weak instrumentation and instrument proliferation (see, for example, Roodman, 2009).

The presence of bias has recently motivated Bai (2013a, b) to propose a new factor analytical (FA) approach to the estimation of (1). The name stems from the fact that the estimator,

which is based on quasi-maximum likelihood (quasi-ML), coincides with the one used in factor analysis (see, for example, Anderson and Amemiya, 1988). A key feature of FA is that it does not require estimation of the individual effects themselves, but only estimation of their second moment, $S_\lambda = N^{-1}\sum_{i=1}^{N}\lambda_i\lambda_i'$.[1] Since under $F_t = 1$ this moment is just a scalar, the incidental parameter problem caused by the fixed effects is effectively removed, leading to an estimator that is completely bias-free. It is also instrumentation-free, which means that the usual difficulties associated with weak instruments and instrument proliferation do not arise in GMM. The work of Bai (2013a, b) is restricted to the classical setup with $F_t = 1$ and $|\rho| < 1$, but is otherwise very general with regard to the idiosyncratic error term, $\varepsilon_{i,t}$, which is allowed to be both cross-section and time series heteroskedastic. Time-specific effects[2], non-zero initial values and predetermined regressors can also be accommodated.

In this paper we extend the work of Bai (2013a, b) to the case when F_t is not necessarily just unity and $\rho \in (-1, 1]$. This is important for (at least) two reasons. One reason is that while fixed effects are commonly used, there are many situations in which they are unlikely to be sufficient. An example is the consumption model based on the life-cycle and rational-expectation hypotheses, which predicts that consumers' marginal utility of wealth should vary over time. Other examples include asset pricing models that assume time-varying risk premia, and models of economic growth in which the state of technology is assumed to follow a linear trend (see Bai, 2009, Section 3, for additional motivating examples). In such cases the fixed effects assumption is almost surely mistaken. The challenge from a theory perspective is that the property of unbiasedness in the fixed effects case does not necessarily carry over to the more general interactive effects model considered in this paper. Indeed, even the introduction of a linear trend, which is arguably the simplest departure from fixed effects, causes the LS bias to double in size regardless of of whether $|\rho| < 1$ or $\rho = 1$ (see Phillips and Sul, 2007).

The extension to the model with interactive effects is related to the recent working paper of Bai (2013c), which appeared after the first version of this paper was written. Bai (2013c) considers a dynamic interactive effects model that is similar to (1) under C1 with $|\rho| < 1$ imposed, which is estimated by a version of the estimator considered here. The focus of this paper, however, is on the relatively challenging unit root case, although we also consider

[1] In this sense, FA is very similar to the GMM approaches considered by Ahn et al. (2001, 2013) and Robertson et al. (2010) in the fixed-T case.

[2] Time-specific effects amounts to setting $F_t = (1, \eta_t)'$ and $\lambda_i = (\alpha_i, 1)'$, such that $c_{i,t} = \lambda_i' F_t = \alpha_i + \eta_t$.

the case when $|\rho| < 1$. Our extension holds considerable promise, from both applied and theoretical viewpoints. From an applied point of view, even variables that on theoretical grounds are expected to be stationary tend to be highly persistent, and the evidence that they do not contain a unit root is weak, at best, as is evident from the large and increasing literature on non-stationary panels (see Breitung and Pesaran, 2008; Baltagi, 2008, Chapter 12, for surveys of this literature). One would therefore not like to exclude the possibility of unit roots when working with real data. From a theoretical point of view, the main problem is the presence of bias, which is even more potent in the unit root case than under $|\rho| < 1$ (see, for example, Moon et al., 2013; Phillips and Sul, 2007). Indeed, as is now well understood, in the unit root case the presence of deterministic terms in the fitted model affects the asymptotic distribution of all estimators of ρ, and does so in both the time series and panel contexts. In panels, this implies that different deterministic specifications have their own bias expressions. For example, if the chosen specification involves structural break dummy variables, then the bias depends on the location of the break(s). This poses serious problems in implementation, as not only is there a need to bias-correct, but the appropriate correction factors also critically depend on the particular model being estimated. Moreover, the complexity of the calculations involved in obtaining these factors increases very quickly with both the number and non-linearity of the fitted deterministic terms. Even for simple LS the required moment calculations are in fact basically impossible, except in the simple case of (at most) a linear trend. Researchers therefore typically only provide correction factors for this case, thereby constraining the use of their estimators to panels that are characterized by similarly simplistic deterministic behavior.

Our findings show that the "FA estimator" has a normal limit for all values of ρ, including unity, and that it is unbiased.[3] The limiting distribution of the estimator considered here is thus continuous as ρ passes through unity, in contrast to what happens for most existing approaches. In fact, the only other estimator known to support asymptotically normal inference for all values of $\rho \in (-1, 1]$ in the current large-T context is the one of Han and Phillips (2010). This estimator is only $\sqrt{NT}$-consistent, however, in contrast to FA, which is (at least) $\sqrt{NT}$-consistent. The fact that FA is unbiased when $\rho = 1$ means that the standard requirement of (at most) a liner trend is not needed, and the otherwise common bias cor-

[3] Bai (2013c), and Bai and Li (2012) refer to the estimator as a "(quasi-)ML estimator". In this paper, however, we follow Bai (2013a, b) and refer to it as an "FA estimator".

rection factors can be completely avoided. In terms of model specification, this means that researchers can proceed just as in the classical regression context. Indeed, all one has to do is to augment the test regression with whatever deterministic specification is felt to be appropriate. The only requirement is that the chosen specification is general enough to include the true one. Interestingly, the usual empirical problem of deciding on which deterministic terms to include does not arise since the common factors, and hence also the deterministic part of the model, can be treated as unknown. Our approach is therefore not only general, but is in this sense also remarkably simple. However, this advantage is at the same time the main drawback of the approach. In the unit root case it is usually desirable to restrict the deterministic part of the model (see Schmidt and Phillips, 1992), but since in FA deterministic and stochastic factors are treated in the same manner, this is not possible without at the same time also restricting the other factors. We therefore consider both C1 and C2 when $\rho = 1$ in this paper.

2 Assumptions

It is useful to write (1) in vector notation. Let us therefore introduce $y_i = (y_{i,1}, ..., y_{i,T})'$, $F = (F_1', ..., F_T')'$ and $\varepsilon_i = (\varepsilon_{i,1}, ..., \varepsilon_{i,T})'$, where y_i and ε_i are $T \times 1$, while F is $T \times m$. Define

$$J = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}, \qquad L = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \rho & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \rho^{T-2} & \dots & \rho & 1 & 0 \end{bmatrix},$$

which are both $T \times T$. It is useful to think of J and L as "lag" and "accumulation" matrices, respectively. Let us further denote by $c_i = (c_{i,1}, ..., c_{i,T})'$ the $T \times 1$ vector of stacked observations on $c_{i,t}$, which under C1 and C2 is given by $c_i = F\lambda_i$ and $c_i = (I_T - \rho J)F\lambda_i$, respectively. In this notation,

$$y_i = c_i + \rho J y_i + \varepsilon_i, \tag{2}$$

which can be solved for y_i, giving

$$y_i = \Gamma c_i + \Gamma \varepsilon_i = \Gamma u_i, \tag{3}$$

where $\Gamma = (I_T - \rho J)^{-1} = I_T + \rho L$ and $u_i = c_i + \varepsilon_i$. Note that L, and hence also Γ, are functions of ρ. In order to emphasize this, we write $L = L(\rho)$ and $\Gamma = \Gamma(\rho)$ whenever appropriate.

The conditions that we are going to be working under are summarized in Assumptions EPS, F and LAM. Throughout, $C < \infty$, $\operatorname{tr} A$ and $||A|| = \sqrt{\operatorname{tr}(A'A)}$ will be used to denote a generic positive constant, and the trace and Frobenius (Euclidean) norm of the matrix A, respectively.

Assumption EPS. $\varepsilon_{i,t}$ is independent and identically distributed (iid) across both i and t with $E(\varepsilon_{i,t}) = E(\varepsilon_{i,t}^3) = 0$, $E(\varepsilon_{i,t}^2) = \sigma^2 > 0$, and $\sigma^{-4}E(\varepsilon_{i,t}^4) = \kappa \leq C$.

Assumption F. If $|\rho| < 1$, then $T^{-1}F'F \to \Sigma_F$, $T^{-1}F'L'F \to \Sigma_F^1$, $T^{-1}F'LL'F \to \Sigma_F^2$ and $T^{-1}F'L'LF \to \Sigma_F^3$ as $T \to \infty$ for some $m \times m$ positive definite matrices Σ_F, Σ_F^1, Σ_F^2 and Σ_F^3, whereas if $\rho = 1$, then $T^{-1}F'F \to \Sigma_F$, $T^{-2}F'L'F \to \Sigma_F^1$, $T^{-3}F'LL'F \to \Sigma_F^2$ and $T^{-3}F'L'LF \to \Sigma_F^3$ as $T \to \infty$. In both cases, $||F_t|| \leq C$ for all t.

Assumption LAM. $||\lambda_i|| \leq C$ for all i, and $S_\lambda \to \Sigma_\lambda$ as $N \to \infty$ for some $m \times m$ positive definite matrix Σ_λ.

Remark 1. Assumption F is significantly less restrictive than the fixed effects assumption of Bai (2013a, b). Although the way that Assumption F is stated supposes that F is fixed, this is not necessary. F can also be random. In this case, we assume F to be independent of $\varepsilon_{i,t}$ for all i and t, and also that Assumption F is satisfied in expectation, in the sense that the expected value of the various sample moments are assumed to behave as in Assumption F. Moreover, $E(||F_t||^4) \leq C$ instead of $||F_t|| \leq C$. This means that there are basically no restrictions on F at all. It can, for example, include both fixed and random elements, have nonzero mean and/or arbitrary dynamics. As we discuss in Section 3, the required moment conditions should be satisfied in most models of empirical relevance.

Remark 2. Since the focus in this paper is the treatment of F when $\rho \in (-1, 1]$, in interest of transparency of the results, some of the other assumptions are quite restrictive. Many of these can, however, be relaxed in the way suggested by Bai (2013a, c). For example, while Bai (2013b) requires that $\varepsilon_{i,t}$ is normal, in Bai (2013a, c) this assumption is relaxed to allow for iid but not necessarily normal innovations. In this paper we do not assume that $\varepsilon_{i,t}$ is normal, but do require that it has a symmetric distribution. The reason for this is that the information matrix is no longer diagonal when $E(\varepsilon_{i,t}^3) \neq 0$, thus adding to the complexity of an already quite complicated problem. Applications are not limited to models

with symmetric innovations, however, since all the relevant second-order derivatives are provided (see Appendix B). Similarly, although we follow Bai (2013b) here and assume that $\varepsilon_{i,t}$ is homoskedastic, the results can be extended along the lines of Bai (2013a, c) to allow for heteroskedasticity (over both time and cross-section). Nonzero initial values and regressors that are exogenous can also be permitted (see Bai, 2013a, b).

3 Asymptotic results

We begin by considering the scenario when $|\rho| < 1$ and F is known. We then show how the results are affected when $\rho = 1$ and/or F is unknown. Unless otherwise stated, we assume throughout that $m \geq 1$, and thus that there are at least some effects present. As mentioned in Section 1, the analytical results under C1 are substantially simpler than those that apply under C2. Since C1 and C2 are indistinguishable for all values of ρ but one, we will only be using C1 whenever $|\rho| < 1$.

3.1 $|\rho| < 1$ and F known

When F is known the vector of parameters is given by $\theta = [(\text{vech}\, S_\lambda)', \rho, \sigma^2]' = (\theta_1', \theta_2')'$, where $\theta_1 = \text{vech}\, S_\lambda$, $\theta_2 = (\rho, \sigma^2)'$, and vech is the half-vec operator that eliminates all supra-diagonal elements of A from vec A. The purpose of this paper is to make inference regarding this vector, and in so doing we follow the FA approach of Bai (2013a, b, c), which is based on the following "discrepancy" function (between $\Sigma(\theta)$ and S_y):

$$Q(\theta) = \log(|\Sigma(\theta)|) + \text{tr}\,(S_y \Sigma(\theta)^{-1}),$$

where $|A|$ is the determinant of A, $S_y = N^{-1}\sum_{i=1}^{N} y_i y_i'$, $\Sigma(\theta) = \sigma^2 \Gamma(\rho)\Lambda(S_\lambda, \sigma^2)\Gamma(\rho)'$ and $\Lambda(S_\lambda, \sigma^2) = I_T + \sigma^{-2} F S_\lambda F'$. To simplify notation we may at times write Q, Σ and Λ for $Q(\theta)$, $\Sigma(\theta)$ and $\Lambda(S_\lambda, \sigma^2)$, respectively. The objective function, denoted $\ell(\theta)$, is just $-N/2$ times $Q(\theta)$;

$$\ell(\theta) = -\frac{N}{2} Q(\theta) = -\frac{N}{2}[\log(|\Sigma(\theta)|) + \text{tr}\,(S_y \Sigma(\theta)^{-1})].$$

Remark 3. The objective function considered here is very similar to those considered by Bai and Li (2012) in the context of a pure common factor model, Ahn et al. (2001, 2013) in the context of a small-T static panel data regression model with (weakly) exogenous regressors,

and Robertson et al. (2010) in the context of a small-T dynamic panel data model. Note in particular how, as in these other papers, θ does not contain $\lambda_1, ..., \lambda_N$, but only S_λ. This means that the dimension of θ remains fixed as $N \to \infty$, which is also the reason for the unbiasedness of FA in the fixed effects case (see Bai, 2013a, for a detailed discussion).

Let us define $G(\rho) = \Gamma(\rho)^{-1} S_y \Gamma(\rho)^{-1\prime}$ and $\hat{\Lambda}(\theta_2) = I_T + \sigma^{-2} F \hat{S}_\lambda(\theta_2) F'$, where $\hat{S}_\lambda(\theta_2) = \sigma^2 F^-(\sigma^{-2} G(\rho) - I_T) F^{-\prime}$ and $A^- = (A'A)^{-1}A'$ for any matrix A. In Appendix A we show that concentration with respect to S_λ leads to the following concentrated objective function:

$$\ell_c(\theta_2) = -\frac{N}{2} Q_c(\theta_2), \tag{4}$$

with

$$Q_c(\theta) = T \log(\sigma^2) + \log(|\hat{\Lambda}(\theta_2)|) + \sigma^{-2} \text{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}]$$

being the correspondingly concentrated discrepancy function. The objective is to maximize $\ell_c(\theta_2)$ with respect to θ_2. Let us therefore denote the true values of ρ, σ^2 and κ by ρ_0, σ_0^2 and κ_0, respectively. Let $\hat{\theta}_2 = (\hat{\rho}, \hat{\sigma}^2)'$ be the FA estimator of $\theta_2^0 = (\rho_0, \sigma_0^2)'$ obtained by maximizing $\ell_c(\theta_2)$ over the parameter space $\Theta_2 = \{\theta_2 : \rho \in (-1, 1], \sigma^2 > 0\}$, that is, $\hat{\theta}_2 = \arg\max_{\theta_2 \in \Theta_2} \ell_c(\theta_2)$. It is assumed that θ_2^0 is an interior point in Θ_2.

Lemma 1. *Under C1, $|\rho_0| < 1$, and Assumptions EPS, F and LAM,*

$$(NT)^{-1}\ell_c(\theta_2) = -\frac{1}{2}\left(\log(\sigma^2) + \frac{\sigma_0^2}{\sigma^2}\right) - \frac{\sigma_0^2}{2\sigma^2}(\rho_0 - \rho)^2\omega_1^2 + O_p((NT)^{-1/2}) + O_p(T^{-1}\log(T)),$$

where $\omega_1^2 = T^{-1}\text{tr}\,(L_0 L_0' + \sigma_0^{-2} S_\lambda F' L_0' M_F L_0 F) \geq 0$, $L_0 = L(\rho_0)$, $M_F = I_T - P_F$ and $P_F = F(F'F)^{-1}F'$.

The second term in the expansion of $(NT)^{-1}\ell_c(\theta_2)$ is obviously maximized at $\rho = \rho_0$. The derivative of the first term with respect to σ^2 is given by $-(1 - \sigma_0^2/\sigma^2)/(2\sigma^2)$, which attains its maximum at $\sigma^2 = \sigma_0^2$. This implies the consistency of $\hat{\theta}_2$, that is, $||\hat{\theta}_2 - \theta_2^0|| = o_p(1)$. Interestingly, consistency does not require $N \to \infty$, but holds also when N is fixed, provided that $T \to \infty$. Bai (2013a, b, c) provides results that are similar to Lemma 1 (see in particular his Lemmas 2 and S.1). Unlike our Lemma 1, however, these results are based on letting $N, T \to \infty$. Moreover, is not apparent that the estimator is consistent also under a fixed N, since the accuracy of approximation is not given by Bai (2013a, b, c).

Theorem 1. *Under the conditions of Lemma 1, as $T \to \infty$ for any N, including $N \to \infty$ with $\sqrt{N}T^{-3/2} \to 0$,*

$$H_{1/2}(\hat{\theta}_2 - \theta_2^0) \sim N\left(0_{2\times 1}, \begin{bmatrix} \omega_1^{-2} & 0 \\ 0 & \sigma_0^4(\kappa_0 - 1) \end{bmatrix}\right),$$

where $\sim$ signifies asymptotic equivalence and $H_p = \text{diag}(\sqrt{N}T^p, \sqrt{NT})$.

According to Theorem 1 there is no asymptotic bias, despite the generality of the conditions placed on F; $(\hat{\theta}_2 - \theta_2^0)$ is centered at zero even when scaled by $\sqrt{NT}$. The condition that $\sqrt{N}T^{-3/2} \to 0$ is the same as in Bai (2013a, b). What is new, however, is the fact that asymptotic normality does not require $N \to \infty$, but holds even when N is fixed. The magnitude of N is not irrelevant, though, as $N \to \infty$ leads to an increase in the rate of consistency, from $\sqrt{T}$ to the $\sqrt{NT}$ rate given in Theorem 1.

The covariance matrix given in Theorem 1 is different from the one reported by Bai (2013b, Theorem S.2); what are here $\sigma_0^4(\kappa_0 - 1)$ and ω_1^{-2} are in Bai (2013b) $2\sigma_0^4$ and $(1 - \rho_0^2)$, respectively. The first difference is due to the fact that the results reported in Bai (2013b) assume that $\varepsilon_{i,t}$ is normally distributed. Under normality the two expressions coincide, since in this case $\kappa_0 = 3$, giving $\sigma_0^4(\kappa_0 - 1) = 2\sigma_0^4$. The second difference is due to the general formulation of F considered here, which includes the fixed effects consideration of Bai (2013a, b) as a special case. In order to appreciate this, note first that by Proof of Lemma 1 (see Appendix C), we have $T^{-1}\text{tr}\,(L_0 L_0') = 1/(1 - \rho_0^2) + O(T^{-1})$. Moreover, under fixed effects, $F = 1_T = (1, ..., 1)'$, a $T \times 1$ vector of ones, suggesting that

$$\begin{aligned} T^{-1}F'L_0'M_F L_0 F &= T^{-1}F'L_0'L_0F - T^{-1}F'L_0'F(F'F)^{-1}(F'L_0'F)' \\ &= T^{-1}1_T'L_0'L_0 1_T - (T^{-1}1_T'L_0'1_T)(T^{-1}1_T'L_0'1_T)', \end{aligned}$$

where, by Proof of Lemma D.1, $T^{-1}1_T'L_0'L_0 1_T = 1/(1 - \rho_0)^2 + O(T^{-1})$ and $T^{-1}1_T'L_0'1_T = 1/(1 - \rho_0) + O(T^{-1})$. Since the leading terms cancel out, $T^{-1}F'L_0'M_F L_0 F = O(T^{-1})$, which in turn implies

$$\omega_1^2 = T^{-1}\text{tr}\,(L_0 L_0' + \sigma_0^{-2} S_\lambda F' L_0' M_F L_0 F) = \frac{1}{(1 - \rho_0^2)} + O(T^{-1}).$$

The results reported in Theorem 1 are therefore identical to those reported in Theorem S.2 of Bai (2013b) under normality.

Although the above discussion refers to the case where $F = 1_T$, we expect $T^{-1}F'L_0'M_F L_0 F$ to be negligible in most other specifications of empirical relevance. In Appendix D we consider as examples the cases where F consists of an intercept with a possible break, and where

it consists of an intercept and (normalized) trend. In both cases, we show that the required sample moments satisfy

$$\begin{aligned} T^{-1}F'F &= \Sigma_F + O(T^{-1}), \\ T^{-1}F'L_0'F &= \frac{1}{(1-\rho_0)}\Sigma_F + O(T^{-1}), \\ T^{-1}F'L_0'L_0F &= \frac{1}{(1-\rho_0)^2}\Sigma_F + O(T^{-1}), \end{aligned}$$

with Σ_F depending on the particular specification of F being considered. Hence, in these cases the leading terms also cancel out, leading to $T^{-1}F'L_0'M_FL_0F = O(T^{-1})$.

Remark 4. The result given in Theorem 1 is similar to the one given in Theorem 2 of Bai (2013c). One difference is that his analysis is based on an approximation where the dependence on F, and hence also the presence of $T^{-1}F'L_0'M_FL_0F$ in ω_1^2, is treated as negligible (see Bai, 2013c, Theorem 1). Our Theorem 1 retains the dependence on F and is therefore more accurate in this regard. In particular, and despite our best efforts, we have not been able to prove that the dependence on F is in fact negligible in general. Moreover, as we explain in Section 3.2, when $\rho_0 = 1$ the dependence on F becomes more apparent and in fact drives some of the results. Another difference in comparison to Bai (2013c) is the method of proof. The proof given in Appendix C is based on formal derivation and evaluation of all the relevant derivatives, and is in fact interesting in itself.

3.2 $\rho_0 = 1$ and F known

In the common factor strand of the so-called "second-generation" panel unit root literature (see Breitung and Pesaran, 2008; Baltagi, 2008, Chapter 12) it is common to decompose F into two parts; (i) a deterministic part, and (ii) a random part that is mean zero. While the latter part is supposed to satisfy C1, the former is restricted as in C2 (see, for example, Moon and Perron, 2004; Pesaran, 2007; Peasaran et al., 2013; Phillips and Sul, 2003). In our case, both parts are given the same treatment, which is also the reason for considering both C1 and C2. Note in particular how under $\rho = 1$,

$$y_{i,t} = \sum_{n=1}^{t} c_{i,n} + \sum_{n=1}^{t} \varepsilon_{i,n}, \tag{5}$$

where the first term on the right equals $\lambda_i' \sum_{n=1}^{t} F_n$ in C1 and $\lambda_i' F_t$ in C2. For example, if $F_t = 1$, while in C2 $\lambda_1, ..., \lambda_N$ represent fixed effects, in C1 they represent unit-specific trend

slopes. Thus, while under C2 the interpretation of the loadings is the same for all values of ρ, including unity, this is not the case under C1. This is also the main reason why deterministic terms are typically supposed to satisfy C2 in the previous literature (see Schmidt and Phillips, 1992, for a discussion). On the other hand, if F_t is iid with zero mean and positive definite covariance matrix, then $\sum_{n=1}^{t} F_n$ represents a common stochastic trend, which is of the same order of magnitude as $\sum_{n=1}^{t} \varepsilon_{i,n}$. Under C2 the idiosyncratic part of the model will thus tend to dominate, but under C1 this is not the case. That which is desirable about C1 (C2) when F_t is stochastic (deterministic) is therefore undesirable when F_t is deterministic (stochastic). In this section we therefore consider both models.

We begin by considering the results under C1. The following lemma shows that $\hat{\theta}_2$ is consistent.

Lemma 2. *Under C1, $\rho_0 = 1$, and Assumptions EPS, F and LAM,*

$$N^{-1}T^{-3}\ell_c(\theta_2) = -\frac{1}{2T^2}\left(\log(\sigma^2) + \frac{\sigma_0^2}{\sigma^2}\right) - \frac{\sigma_0^2}{2\sigma^2}(\rho_0 - \rho)^2 T^{-2}\omega_1^2 + O_p(T^{-2}) + O_p((NT)^{-1/2}),$$

where $T^{-2}\omega_1^2 \geq 0$.

The result reported in Lemma 2 is similar to the one reported by Moon and Phillips (1999, equation (8)) for the Gaussian log-likelihood function in the fixed effects near-unit root case. A difference when compared to Lemma 1 is that in Lemma 2 the first term on the right-hand side of $N^{-1}T^{-3}\ell_c(\theta_2)$ is negligible. This does not mean that $\hat{\sigma}^2$ is inconsistent, but merely that it is consistent at a slower rate than $\hat{\rho}$. This is shown in Theorem 2, which provides the relevant asymptotic distribution.

Theorem 2. *Under the conditions of Lemma 2, as N, $T \to \infty$ with $\sqrt{N}T^{-3/2} \to 0$,*

$$H_{3/2}(\hat{\theta}_2 - \theta_2^0) \sim N\left(0_{2\times 1}, \begin{bmatrix} T^2\omega_1^{-2} & 0 \\ 0 & \sigma_0^4(\kappa_0 - 1) \end{bmatrix}\right).$$

Note how the rate of consistency of $(\hat{\rho} - \rho_0)$ is $\sqrt{N}T^{3/2}$, which is higher than the usual panel "superconsistency" rate of $\sqrt{N}T$. As mentioned in the above, the reason for this extraordinarily fast rate of consistency is that under C1 and Assumption F, while $\sum_{n=1}^{t} \varepsilon_{i,n} = O_p(\sqrt{T})$, we have $||\sum_{n=1}^{t} F_n|| \leq \sum_{n=1}^{t} ||F_n|| = O(T)$. The asymptotic distribution is therefore dominated by the common component. In order to appreciate the effect of this we look at $T^{-2}\omega_1^2 = T^{-3}\text{tr}\,(L_0 L_0' + \sigma_0^{-2} S_\lambda F' L_0' M_F L_0 F)$, the inverse of the asymptotic variance of

$\sqrt{N}T^{3/2}(\hat{\rho}-1)$. While the first term is due to $\sum_{n=1}^{t}\varepsilon_{i,n}$, the second term is due to $\sum_{n=1}^{t}F_n$. A direct calculation shows that

$$T^{-2}\text{tr}\,(L_0L_0') = \frac{1}{T^2}\sum_{t=1}^{T}(T-t) = \int_{v=0}^{1}(1-v)dv + o(1) = \frac{1}{2} + o(1),$$

implying that $T^{-2}\omega_1^2 = T^{-3}\text{tr}\,(\sigma_0^{-2}S_\lambda F'L_0'M_FL_0F) + o(1)$. Earlier we showed that the effect of $F'L_0'M_FL_0F$ was negligible in the special case of $|\rho_0| < 1$ and $F = 1_T$. When $\rho_0 = 1$ this is no longer the case. Indeed, it is not difficult to see that with $F = 1_T$ and $t = \lfloor vT \rfloor$ for $v \in [0,1]$,

$$\begin{aligned} T^{-3}F'L_0'M_FL_0F &= T^{-3}1_T'L_0'L_01_T - (T^{-2}1_T'L_0'1_T)(T^{-2}1_T'L_0'1_T)' \\ &= \frac{2}{T^3}\sum_{t=2}^{T}(T-t)t - \left(\frac{1}{T^2}\sum_{t=1}^{T-1}(T-t)\right)^2 + o(1) \\ &= 2\int_{v=0}^{1}(1-v)vdv - \left(\int_{v=0}^{1}(1-v)dv\right)^2 + o(1) = \frac{1}{12} + o(1), \end{aligned}$$

and therefore

$$T^{-2}\omega_1^2 = \frac{S_\lambda}{12\sigma_0^2} + o(1).$$

Hence, $\sqrt{N}T^{3/2}(\hat{\rho}-1) \sim N(0, 12\sigma_0^2/S_\lambda)$ under $F = 1_T$, a result that is again driven by the common component.

Theorem 1 and the discussion that follows it make use of Assumption F, which is very general. It is therefore interesting to consider a few special cases. Suppose for example that $m = 0$, such that the model can be fitted without factors. In this case, $T^{-2}\omega_1^2 = T^{-3}\text{tr}\,(L_0L_0') = O(T^{-1})$, suggesting that the rate of consistency is reduced from $\sqrt{N}T^{3/2}$ to $\sqrt{N}T$. In fact, since $T^{-2}\text{tr}\,(L_0L_0') \to 1/2$, it is not difficult to show that $\sqrt{N}T(\hat{\rho}-1) \to_d N(0,2)$ as $N, T \to \infty$ with $\sqrt{N}T^{-3/2} \to 0$, where $\to_d$ signifies convergence in distribution, which is in agreement with existing results for the ML and LS estimators of ρ_0 (see, for example, Levin and Lin, 1992, Theorem 3.2). If $m > 0$ but F_t is iid with zero mean and positive definite covariance matrix, such that $\sum_{n=1}^{t}\varepsilon_{i,n}$ and $\sum_{n=1}^{t}F_n$ are of the same order, then $\sqrt{N}T(\hat{\rho}-1) \sim N(0, T\omega_1^{-2})$. Moreover, since $T^{-2}\text{tr}\,(S_\lambda F'L_0'M_FL_0F) = T^{-2}\text{tr}\,(S_\lambda F'L_0'L_0F) + o_p(1) \leq C$ and $T^{-2}\text{tr}\,(L_0L_0') \to 1/2$, both the common and idiosyncratic components contribute to the asymptotic distribution.

Remark 5. Theorem 2 requires that $N, T \to \infty$ with $\sqrt{N}T^{-3/2} \to 0$, which is stronger than the corresponding condition in Theorem 1. The reason for this is the usual dependence on

Brownian motion as $T \to \infty$ when $\rho_0 = 1$, which is effectively smoothed out by passing $N \to \infty$, thereby enabling asymptotic normality.

The fact that $T^{-3}\text{tr}\left(S_\lambda F' L_0' M_F L_0 F\right)$ under C1 drives the results is important, not only for the rate of consistency, but also because of what it implies for λ_i. In the unit root literature it is quite common to assume that C1 holds, but to restrict the order of the deterministic trend polynomial to be the same under $\rho_0 = 1$ as when $|\rho_0| < 1$ (see, for example, Levin et al., 2002). For example, it is assumed that $\lambda_1 = ... = \lambda_N = 0$ when $F_t = 1$, for otherwise $y_{i,t}$ would contain a linear trend. Unfortunately, this is not possible in FA, at least not under C1, as the effect of S_λ on ω_1^2 is non-negligible and $S_\lambda \to \Sigma_\lambda > 0$ under Assumption LAM. Hence, if $F_t = \rho_0 = 1$, then $y_{i,t}$ must contain a linear trend, which is clearly very restrictive. With this is mind, we now continue to the results obtained under C2.

The required derivatives and the resulting asymptotic derivations become extremely tedious under C2 due to the way that the inverse of Γ enters into the expressions. Intuitively the extension of the above results for C1 to C2 follows from simply replacing F by $\Gamma^{-1}F$. Note in particular how the concentrated objective function has the same form as in (4) but with F replaced by $\Gamma^{-1}F$ in $\hat{\Lambda}(\theta_2)$ and $\hat{S}_\lambda(\theta_2)$; hence, $\hat{\Lambda}(\theta_2) = I_T + \sigma^{-2}\Gamma(\rho)^{-1}F\hat{S}_\lambda(\theta_2)F'\Gamma(\rho)^{-1\prime}$ and $\hat{S}_\lambda(\theta_2) = \sigma^2(\Gamma(\rho)^{-1}F)^-(\sigma^{-2}G(\rho) - I_T)(\Gamma(\rho)^{-1}F)^{-\prime}$.

Lemma 3. *Under C2, $\rho_0 = 1$, and Assumptions EPS, F and LAM,*

$$N^{-1}T^{-2}\ell_c(\theta_2) = -\frac{\sigma_0^2}{2\sigma^2}(\rho_0 - \rho)^2 T^{-2}\text{tr}\left(L_0 L_0'\right) + O_p(N^{-1/2}) + O_p(T^{-1}).$$

The rate of consistency of $\hat{\rho}$ under C2 is generally lower than under C1 as the normalization of $\ell_c(\theta_2)$ with respect to N and T indicates. Theorem 3 confirms this.

Theorem 3. *Under the conditions of Lemma 3, as N, $T \to \infty$ with $\sqrt{N}T^{-1} \to 0$,*

$$H_1(\hat{\theta}_2 - \theta_2^0) \sim N\left(0_{2\times 1}, \begin{bmatrix} T\omega_2^{-2} & 0 \\ 0 & \sigma_0^4(\kappa_0 - 1) \end{bmatrix}\right),$$

where $\omega_2^2 = T^{-1}\text{tr}\left(L_0 L_0' + \sigma_0^{-2} S_\lambda F' \Gamma^{-1\prime} L_0' M_{\Gamma^{-1}F} L_0 \Gamma^{-1} F\right)$.

It can be shown that $||T^{-1}F'\Gamma^{-1\prime}L_0' M_{\Gamma^{-1}F} L_0 \Gamma^{-1}F|| \leq C$, giving

$$T^{-1}\omega_2^2 = T^{-2}\text{tr}\left(L_0 L_0' + \sigma_0^{-2} S_\lambda F' \Gamma^{-1\prime} L_0' M_{\Gamma^{-1}F} L_0 \Gamma^{-1} F\right) = T^{-2}\text{tr}\left(L_0 L_0'\right) + o(1).$$

Hence, since $T^{-2}\text{tr}\left(L_0 L_0'\right) \to 1/2$, we can show that $\sqrt{N}T(\hat{\rho}-1) \to_d N(0,2)$ as $N, T \to \infty$ with $\sqrt{N}T^{-1} \to 0$, which is the same result as obtained under C1 with $m = 0$. Since F is completely unrestricted here, the specification in C2 therefore leads to simplified results when compared to C1. Note in particular how the asymptotic distribution does not depend on S_λ. This means that the requirement that $S_\lambda \to \Sigma_\lambda > 0$ is no longer necessary (a formal proof is available upon request). Under C2 some (or indeed all) of the loadings may be zero for all units, which was not possible under C1.

The fact that the limiting distribution is asymptotically invariant with respect to F is wort discussion. As explained in Section 1, most existing estimators of ρ are biased in ways that depend on the deterministic specification being fitted. Valid inference in these cases therefore requires bias-correction. Typically these correction factors are only available for the simple case of (at most) a linear trend, which obviously limits the applicability of these estimators. Here we are also assuming that F is known. In practice, however, there is uncertainty over F, and in such cases researchers have to adopt a liberal modeling strategy to ensure that the deterministic behaviors of all the units are captured. The conventional specification with (at most) a linear is clearly inadequate if one allows for the possibility that some of the units may be trending non-linearly. This will be the case when, for example, working with variables where trending behavior is evident, such as GDP, industrial production, money supply and consumer or commodity prices. The invariance property of FA is therefore not only very convenient from an applied point of view (as there is no bias to correct for), but also enables inference in cases previously not possible.

Remark 6. Hahn and Kuersteiner (2002) study the asymptotic distribution of the LS estimator of ρ_0 under C2 in the fixed effects unit root case. According to their Theorem 4, not only is LS biased, but is asymptotic variance ($51/5 \approx 10$) is also substantially higher then for FA (2).

Remark 7. The requirement that $\sqrt{N}T^{-1} \to 0$ is stronger than in Theorem 2. The reason for this is the relatively slow rate of consistency in this case.

In Section 3.1 ($|\rho_0| < 1$) we assumed that C1 held true. We then showed that the asymptotic distribution under $\rho_0 = 1$ can be written in exactly the same way but with a different rate of consistency. In this sense the results are continuous as ρ_0 passes through unity. The same is true under C2, that is, the asymptotic distribution of $\hat{\theta}_2$ when $|\rho_0| < 1$ is the same as

in Theorem 3 but with H_1 and $T\omega_2^{-2}$ replaced by $H_{1/2}$ and ω_2^{-2}, respectively (a formal proof is available upon request).

The above results imply that normal inference is possible for all $\rho_0 \in (-1, 1]$ under both C1 and C2. Consider C2. Denote by $\hat{\omega}_2^2$ an estimator of ω_2^2. This estimator can be based on either numerical or analytical evaluation of the Hessian at $\theta_2 = \hat{\theta}_2$ (the elements of which are given in Appendix B), but it can also be based on direct estimation of asymptotic formula for ω_2^2, that is,

$$\hat{\omega}_2^2 = T^{-1}\text{tr}\,(L_0 L_0' + \hat{\sigma}^{-2}\hat{S}_\lambda^* F'\hat{\Gamma}^{-1\prime} L_0' M_{\hat{\Gamma}^{-1}F} L_0 \hat{\Gamma}^{-1} F),$$

where $\hat{\Gamma} = \Gamma(\hat{\rho})$ and $\hat{S}_\lambda^* = \hat{S}_\lambda(\hat{\theta}_2)$. By using the results provided in Proof of Lemma 1, it is not difficult to show that $||\hat{S}_\lambda(\theta_2^0) - S_\lambda|| = o_p(1)$. But we also have $||\hat{\theta}_2 - \theta_2^0|| = o_p(1)$, suggesting that, by the continuous mapping theorem, $||\hat{S}_\lambda(\hat{\theta}_2) - \hat{S}_\lambda(\theta_2^0)|| = o_p(1)$. Hence, $||\hat{S}_\lambda(\hat{\theta}_2) - S_\lambda|| \leq ||\hat{S}_\lambda(\theta_2^0) - S_\lambda|| + ||\hat{S}_\lambda(\hat{\theta}_2) - \hat{S}_\lambda(\theta_2^0)|| = o_p(1)$, showing that $\hat{S}_\lambda^*$ is consistent for S_λ (see also Bai, 2013a, Theorem 1; Bai, 2013b, Corollary S.1). A similar argument can be used to show that $||\hat{\Gamma}(\hat{\rho})^{-1} - \Gamma^{-1}|| = o_p(1)$. Therefore, $\hat{\omega}_2^2$ is a consistent estimator of ω_2^2. Regardless of how $\hat{\omega}^2$ is constructed the FA-based t-statistic for testing $H_0 : \rho_0 = \rho^0$ is given by

$$t(\rho^0) = \hat{\omega}_2\sqrt{NT}(\hat{\rho} - \rho^0).$$

The asymptotic distribution (as N, $T \to \infty$) of this t-statistic under the null hypothesis is an immediate consequence of the above results and is given by

$$t(\rho^0) \to_d N(0, 1),$$

which holds for all values of $\rho^0 \in (-1, 1]$.[4] Note in particular how $t(\rho^0)$ can be used as a unit root test.

3.3 F unknown

The above presumes that F is known. This is not necessary. If F is unknown we define $\theta = [(\text{vech}\, S_\lambda)', \rho, \sigma^2, (\text{vec}\, F)']' = (\theta_1', \theta_2')'$, where $\theta_1 = \text{vech}\, S_\lambda$ is as before and $\theta_2 = [\rho, \sigma^2, (\text{vec}\, F)']'$. Let us denote the true value of F by $F^0 = (F_1^0, ..., F_T^0)'$, and the corresponding estimator by $\hat{F}$. The estimation of θ_2^0 can proceed exactly as before. The main difference

[4]Note that the asymptotic distribution of $t(\rho^0)$ holds even if $\rho_0 = 1$ so that the rate of consistency is faster than the $\sqrt{NT}$ rate used in normalization of $(\hat{\rho} - \rho^0)$. This is due to the "self-normalizing" property of $t(\rho^0)$. For example, if $\rho_0 = 1$ under C2, then $t(\rho^0) = \hat{\omega}\sqrt{NT}(\hat{\rho} - \rho^0) = \sqrt{T^{-1}\hat{\omega}^2}\sqrt{N}T(\hat{\rho} - \rho^0) \to_d N(0, 1)$.

is that since now both λ_i and F are unknown there is an identification issue, which can be resolved by imposing m^2 restrictions (see, for example, Bai and Li, 2012, Section 4, for a detailed discussion). In the Monte Carlo study of Section 4 this is accomplished by setting $F = (I_m, G')'$, where G is $(T-m) \times m$.

Proposition 1. *Under C1 or C2, $\rho_0 \in (-1, 1]$ and Assumptions EPS, F and LAM, uniformly in t,*

$$||\hat{F}_t - F_t^0|| = o_p(1).$$

In most applications the coefficient of interest is ρ_0, not F, and in such cases the main concern is how to control for F. For this reason we only provide a consistency result here, although the asymptotic distribution of $\sqrt{N}(\hat{F}_t - F_t^0)$ can be obtained as in Bai (2013c, Proposition 2). The fact that F can be treated as unknown means that applied researchers are spared from the problem of having to decide on which deterministic components to include. For example, if structural shifts are present, then there is no need for any a priori knowledge regarding their locations, which are obtained as part of the estimation process.

Remark 8. Proposition 1 supposes that the number of factors, m, is known. However, the asymptotic results also hold when m is replaced by a consistent estimator, $\hat{m}$ say. Write $\hat{\rho}(m)$ for $\hat{\rho}$. Consider for simplicity the case when $|\rho_0| < 1$. To see that $\hat{\rho}(\hat{m})$ has the same asymptotic distribution as $\hat{\rho} = \hat{\rho}(m)$, consider

$$\begin{aligned} P(\sqrt{NT}[\hat{\rho}(\hat{m}) - \rho_0] \leq \delta) &= P(\sqrt{NT}[\hat{\rho}(\hat{m}) - \rho_0] \leq \delta | \hat{m} = m) P(\hat{m} = m) \\ &+ P(\sqrt{NT}[\hat{\rho}(\hat{m}) - \rho_0] \leq \delta | \hat{m} \neq m) P(\hat{m} \neq m), \end{aligned}$$

where $\delta > 0$ is a small number. Because $P(\hat{m} = m) \to 1$ and $P(\hat{m} \neq m) \to 0$, the second term on the right-hand side converges to zero, and $P(\sqrt{NT}(\hat{\rho}(\hat{m}) - \rho_0) \leq \delta) = 1 + o(1)$. Moreover, conditional on $\hat{m} = m$, $\hat{\rho}(\hat{m}) = \hat{\rho}(m)$. Thus,

$$|P(\sqrt{NT}[\hat{\rho}(\hat{m}) - \rho_0] \leq \delta) - P(\sqrt{NT}[\hat{\rho}(m) - \rho_0] \leq \delta)| \to 0.$$

Bai and Ng (2002) consider the problem of consistent estimation of m in the context of a pure common factor model, and make several suggestions toward this end. It is conjectured that these estimators are consistent also in the present setup.

4 Monte Carlo results

A small-scale Monte Carlo simulation exercise was carried out to evaluate the small-sample performance of FA. The DGP is given by (1), where $\varepsilon_{i,t} \sim N(0,1)$, $\lambda_i \sim U(1,2)$ and $\rho_0 \in \{0, 0.5, 0.95, 1\}$. Three DGP's for F_t were considered:

F1. $F_t = 1$;

F2. $F_t = (1,0)'$ if $t < \lfloor T/2 \rfloor$ and $F_t = (1,1)'$ otherwise;

F3. $F_t \sim N(0,1)$.

While we assume that F_t is known in F1 and F2, we treat F_t as an unknown parameter to be estimated along with the other parameters of the model in F3. The estimation in F3 is carried out in two steps. According to (3), under C1, $y_i = \Gamma F \lambda_i + \Gamma \varepsilon_i$, which is merely a static common factor model for y_i. The first step of the estimation procedure therefore involves the use of the method of principal components to estimate $G = \Gamma F$. Since y_i need not be stationary, we follow Bai and Ng (2004), and apply the principal components method to $\Delta y_{i,t}$ rather than to $y_{i,t}$. This gives an estimator of (the space spanned by) G in first-differenced form, which is then accumulated to levels. In the second step, θ_2^0 is estimated conditional on the first-step estimator $\hat{G}$ of G. Under C2 and $\rho_0 = 1$, $G = F$.

In addition to FA, in F1 the fixed effects LS estimator, the bias-adjusted LS (BALS) estimator of Hahn and Kuersteiner (2002), and the Anderson and Hsiao (1981) instrumental variables (IV) estimator using both lagged levels (AHL) and differences (AHD) as instruments are simulated. A large number of results were produced, but in interest of space we focus on the bias and root mean squared error (RMSE) of $\hat{\rho}$, and the size of a nominal 5% level t-test. Some of the unreported results are described in the end of this section. The number of replications was set to 5,000. All computational work was done in GAUSS 11.[5]

Table 1 presents the results for F1 when $|\rho_0| < 1$. We see that the bias and RMSE of FA is very small and that this is true for all the sample sizes considered. In fact, performance is very good even for N and T as small as 10. The values of T and N are not irrelevant, however. In particular, we see how the bias and RMSE tend to zero when T and/or N increase, which agrees with the $\sqrt{NT}$-consistency of FA. As expected, this improvement in performance

[5]In implementing FA we used the BFGS algorithm for constrained optimization with non-negativity constraints on σ^2.

holds irrespectively of the relative expansion rate of T and N, and therefore even when N is held fixed and only T increases. In fact, FA is uniformly better than the competing estimators in terms of bias and RMSE. The performance of BALS is very similar, though, especially for the larger values of N and/or T, which is consistent with the fact that both estimators are asymptotically efficient (see Bai, 2013a, Section 4). We can also see that the size of the FA-based t-test is close to the nominal 5% level for all values of ρ_0 and sample sizes considered. The same cannot be said about the other estimators, however. Indeed, AHL is consistently undersized, and LS is consistently oversized. The results for AHD and BALS are generally better, although there is a tendency for the distortions to vary quite markedly with ρ_0; when $\rho_0 = 0$ the tests are oversized, whereas when $\rho_0 = 0.95$ they are undersized.

Since most of the estimators considered are designed specifically for the fixed effects case, we only consider FA in experiments F2 and F3. The results are reported in Table 2. The first thing to note is that the performance in F2 and F3 is almost as good as in F1. In fact, the results for F1 and F2 are almost identical. The results for F3 are slightly worse, which is as expected since in small samples the estimation of F will lead to increased variance. Performance is still very good, however, and gets better as N and/or T increases, which is presumably a reflection of the consistency of $\hat{F}_t$.

The results for the case when $\rho_0 = 1$ are summarized in Table 3. As expected in view of the relatively high rate of consistency in this case, the results are generally much better than when $|\rho_0| < 1$; the bias and RMSE values are very close to zero, and the size distortions are minimal. Comparing across the two specifications of the common component, we see that the results for C1 are generally much better than those for C2, which is again due to the difference in the rate of consistency. The only exception is F3 in which the results for C2 look best. The reason for this is that $F_t \sim N(0,1)$ here, which, as we explained in Section 3.2, implies that the rate of consistency under C1 is reduced from $\sqrt{N}T^{3/2}$ to $\sqrt{N}T$. Looking next across the three DGP's considered for F_t under C2 we see that the results are very similar, which we take as support for the theoretical prediction that FA should be asymptotically invariant with respect to F.

As mentioned in the beginning of this section, the complete set of Monte Carlo results is huge (they are available upon request). However, since most of the results are very similar to the ones reported in Tables 1–3, we do not include them here, but only briefly describe them. First, FA performs well even when $\varepsilon_{i,t}$ is drawn from a fat-tailed distribution. For

example, the results based on drawing $\varepsilon_{i,t}$ from a t-distribution with seven degrees of freedom are almost indistinguishable from those reported in Tables 1–3. Second, performance is not affected by the presence of heteroscedasticity provided FA is modified as outlined in Bai (2013a), and that T and N are sufficiently large, which is accordance with our expectations. Third, performance is also unaffected by the presence of time-specific fixed effects when appropriately accounted for as explained in Bai (2013a). Fourth, under C2 the presence of a unit root causes serious problems for the competing estimators, especially for AHL and AHD, where the bias and RMSE results are hundreds of times larger than those found under stationarity. LS and BALS perform better in terms of bias and RMSE, but their size distortions are still unacceptably large with sizes that are close to 100% in the majority of cases. The fact that the asymptotic distribution of the FA-based t-statistic is the same regardless of the value taken by ρ_0 is therefore a great advantage.

5 Conclusion

The FA approach of Bai (2013a, b) was extended to the case with interactive effects and a possible unit root. It was shown that the estimator is unbiased and asymptotically normal for all values of $\rho_0 \in (-1, 1]$. The unbiasedness property not only makes the estimator easy to compute, but also enables estimation and inference in situations previously not possible. In FA the deterministic terms are treated as additional common factors that may be estimated from the data. It was argued that while this makes for very simple implementation (in the sense that no modeling of the deterministic component is required), it is also a drawback in the unit root case when compared to other approaches that enable a separate treatment of deterministic and random factors.

References

Abadir, K. M., and J. R. Magnus (2005). *Matrix Algebra, Econometric Exercises 1*. Cambridge University Press, New York.

Ahn, S. C., Y. H. Lee and P. Schmidt (2001). GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects. *Journal of Econometrics* **101**, 219–255.

Ahn, S. C., Y. H. Lee and P. Schmidt (2013). Panel Data Models with Multiple Time-Varying Individual Effects. *Journal of Econometrics* **174**, 1–14.

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press, Cambridge.

Anderson, T. W., and Y. Amemiya (1988). The Asymptotic Normal Distribution of Estimators in Factor Analysis Under General Conditions. *Annals of Statistics* **16**, 759–771.

Anderson, T. W., and C. Hsiao (1981). Estimation of Dynamic Models with Error Components. *Journal of American Statistical Association* **76**, 598–606.

Arellano, M., and S. Bond (1991). Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* **58**, 277–297.

Arellano, M., and O. Bover (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics* **68**, 29–51.

Bai, J. (2009). Panel Data Models with Interactive Fixed Effects. *Econometrica* **77**, 1229–1279.

Bai, J. (2013a). Fixed-Effects Dynamic Panel Models, a Factor Analytical Method. *Econometrica* **81**, 285–314.

Bai, J. (2013b). Supplement to "Fixed-Effects Dynamic Panel Models, a Factor Analytical Method". (*Econometrica* **81**, 285–314). Available on the Econometric Society website.

Bai, J. (2013c). Likelihood Approach to Dynamic Panel Models with Interactive Effects. MPRA Paper 50267.

Bai, J., and K. P. Li (2012). Statistical Analysis of Factor Models of High Dimension. *Annals of Statistics* **40**, 436–465.

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

Bai, J., and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration. *Econometrica* **72**, 1127–1177.

Baltagi, B. (2008). *Econometric Analysis of Panel Data, Fourth Edition*. John Wiley and Sons, New York.

Blundell, R., and S. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* **87**, 115–143.

Breitung J., and M. H. Pesaran (2008). Unit Roots and Cointegration in Panels. In Matyas, L., and P. Sevestre (Eds.), *The Econometrics of Panel Data*, Klüwer Academic Publishers, 279322.

Bun, M., and V. Sarafidis (2013). Dynamic Panel Data Models. Forthcoming in Baltagi, B. (Ed.), *Oxford Handbook on Panel Data*, Oxford University Press.

Chudik, A., and M. H. Pesaran (2013). Large Panel Data Models with Cross-Sectional Dependence: A Survey. CESifo Working Paper No. 4371. Forthcoming in Baltagi, B. H. (Ed.). *The Oxford Handbook on Panel Data*, Oxford University Press.

Han, C., and P. C. B. Phillips (2010). GMM Estimation for Dynamic Panels with Fixed Effects and Strong Instruments at Unity. *Econometric Theory* **26**, 119–151.

Hahn, J., and G. Kuersteiner (2002). Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects when Both n and T are Large. *Econometrica* **70**, 1639–1657.

Levin, A., and C. Lin (1992). Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties. University of California working paper 9223.

Levin, A., C. Lin, and C.-J. Chu (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties. *Journal of Econometrics* **108**, 1–24.

Moon, H. R., and B. Perron (2004). Testing for Unit Root in Panels with Dynamic Factors. *Journal of Econometrics* **122**, 81–126.

Moon, H. R., and P. C. B. Phillips (1999). Maximum Likelihood Estimation in Panels with Incidental Trends. *Oxford Bulletin of Economics and Statistics* **61**, 748–771.

Moon, H. R., B. Perron and P. C. B. Phillips (2013). Incidental Parameters and Dynamic Panel Modeling. Forthcoming in Baltagi, B. (Ed.), *Oxford Handbook on Panel Data*, Oxford University Press.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica* **49**, 1417–1426.

Pesaran, M. H. (2007). A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. *Journal of Applied Econometrics* **22**, 265–312.

Pesaran, H. M., L. V. Smith, and T. Yamagata (2013). Panel Unit Root Tests in the Presence of a Multifactor Error Structure. *Journal of Econometrics* **175**, 94–115.

Phillips, P. C. B., and D. Sul (2003). Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence. *Econometrics Journal* **6**, 217–259.

Phillips, P. C. B., and D. Sul (2007). Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence. *Journal of Econometrics* **137**, 162–188.

Robertson, D., V. Sarafidis and J. Symons (2010). IV Estimation of Panels with Factor Residuals. MPRA Paper No. 26166.

Roodman, D. (2009). A Note on the Theme of too Many Instruments. *Oxford Bulletin of Economics and Statistics* **71**, 135–158.

Schmidt, P. and P. C. B. Phillips (1992). LM Tests for a Unit Root in the Presence of Deterministic Trends. *Oxford Bulletin of Economics and Statistics* **54**, 257–287.

Appendix A: Preliminaries

We start with some notation. It is convenient to write $L = L(\rho)$ as $L = (l_1, ..., l_T)$, where $l_t = l_t(\rho) = (0'_{t\times 1}, 1, \rho, ..., \rho^{T-1-t})'$ is $T \times 1$ and $0_{n\times k}$ is a $n \times k$ matrix of zeroes. In this notation, $\Gamma = I_T + \rho L$ and $\Gamma^{-1} = I_T - \rho J$. Let $\Gamma_0 = I_T + \rho_0 L_0$, where $L_0 = L(\rho_0) = (l_{1,0}, ..., l_{T,0})$. It follows that

$$\Gamma^{-1}\Gamma_0 = (I_T - \rho J)(I_T + \rho_0 L_0) = I_T - \rho J + \rho_0 L_0 - \rho\rho_0 J L_0 = I_T + (\rho_0 - \rho)L_0.$$

At times it will be useful to be able to rewrite L_0 as $L_0 = (l_{-T,0}, ..., l_{-1,0})'$, where $l_{-t,0} = (\rho^{T-1-t}, ..., \rho, 1, 0'_{t\times 1})'$ is the reverse version of $l_{t,0}$. Let $A^*_t = \sum_{s=1}^{t} \rho_0^{t-s} A_s$ for any $m \times 1$ vector A_s. In this notation, letting $A = (A_1, ..., A_T)'$,

$$L_0 A = \begin{bmatrix} l'_{-T,0}A \\ \vdots \\ l'_{-1,0}A \end{bmatrix} = \begin{bmatrix} 0'_{m\times 1} \\ A'_1 \\ \vdots \\ \sum_{s=1}^{T-1} \rho_0^{T-1-s} A'_s \end{bmatrix} = \begin{bmatrix} 0'_{m\times 1} \\ A^{*\prime}_1 \\ \vdots \\ A^{*\prime}_{T-1} \end{bmatrix}.$$

This result will be used frequently in the sequel.

The matrix treatment builds heavily on Abadir and Magnus (2005), especially the matrix calculus. It is convenient to define the matrix derivative operator D, which is such that if the matrix function $F(x)$ is $m \times p$ and x is $n \times q$, then $\mathrm{D}\,F(x) = \partial \mathrm{vec}\, F(x)/\partial(\mathrm{vec}\, x)'$ is $mp \times nq$. Hence, denoting by d the matrix differential, then $\mathrm{d}\,\mathrm{vec}\, F(x) = F(x)\mathrm{d}\,\mathrm{vec}\, x$, or $\mathrm{D}\,F(x) = \mathrm{d}\,\mathrm{vec}\, F(x)/\mathrm{d}\,\mathrm{vec}\, x$.

Throughout, A, B and C will be used to denote generic matrices. a, b and c denote generic scalars.

Proof of (4).

Consider $\log(|\Sigma|)$ in $Q(\theta)$. By using $|AB| = |A||B|$ and $|\Gamma| = 1$, we obtain $|\Sigma| = |\sigma^2\Lambda| = (\sigma^2)^T|\Lambda|$, and therefore

$$\log(|\Sigma|) = T\log(\sigma^2) + \log(|\Lambda|).$$

Making use of this, the definition of S_y, and then $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$, we obtain

$$Q(\theta) = T\log(\sigma^2) + \log(|\Lambda(S_\lambda, \sigma^2)|) + \sigma^{-2}\mathrm{tr}\,[G(\rho)\Lambda(S_\lambda, \sigma^2)^{-1}]$$

where $G(\rho) = \Gamma(\rho)^{-1}S_y\Gamma(\rho)^{-1\prime}$.

We now consider the first order condition with respect to S_λ. Since $\text{vec}(ABC) = (C' \otimes A)\text{vec}\, B$, we have $\text{D}\,(FS_\lambda F') = F \otimes F'$. By using this and $\text{D}\,\log|A| = [\text{vec}\,(A'^{-1})]'\text{D}\,A$, we obtain

$$\text{D}\,\log(|\Lambda|) = [\text{vec}\,(\Lambda'^{-1})]'(F \otimes F').$$

The derivative of $\text{tr}\,[G(\rho)\Lambda(S_\lambda, \sigma^2)^{-1}]$ is given by

$$\text{D}\,\text{tr}\,(G\Lambda^{-1}) = -[\text{vec}\,(\Lambda^{-1}G\Lambda^{-1})']'(F \otimes F'),$$

as follows from noting that $\text{D}\,\text{tr}\,|AB^{-1}| = -[\text{vec}\,(B^{-1}AB^{-1})']'\text{D}\,B$. Solving for S_λ from the resulting first order condition gives

$$\begin{aligned}
&\text{D}\,\log(|(I_T + \sigma^{-2}F\hat{S}_\lambda F')|) + \sigma^{-2}\text{D}\,\text{tr}\,(G(I_T + \sigma^{-2}F\hat{S}_\lambda F')^{-1}) \\
&= [\text{vec}\,((I_T + \sigma^{-2}F\hat{S}_\lambda F')'^{-1})]'(F \otimes F') \\
&- \sigma^{-2}[\text{vec}\,((I_T + \sigma^{-2}F\hat{S}_\lambda F')^{-1}G(I_T + \sigma^{-2}F\hat{S}_\lambda F')^{-1})']'(F \otimes F') = 0,
\end{aligned}$$

or $I_T + \sigma^{-2}F\hat{S}_\lambda F' = \sigma^{-2}G$, giving $\hat{S}_\lambda(\theta_2) = \sigma^2 F^-(\sigma^{-2}G(\rho) - I_T)F^{-\prime}$, where $A^- = (A'A)^{-1}A'$ for any matrix A. Let $\hat{\Lambda}(\theta_2) = I_T + \sigma^{-2}F\hat{S}_\lambda(\theta_2)F'$. The concentrated discrepancy function is

$$Q_c(\theta) = T\log(\sigma^2) + \log(|\hat{\Lambda}(\theta_2)|) + \sigma^{-2}\text{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}],$$

as required. ■

Appendix B: Derivatives

Derivatives under C1

The concentrated objective function is

$$\ell_c(\theta_2) = -\frac{NT}{2}\log(\sigma^2) - \frac{N}{2}\log(|\hat{\Lambda}(\theta_2)|) - \frac{N}{2\sigma^2}\text{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}].$$

We begin by taking partial derivative with respect to ρ;

$$\frac{\partial \ell_c(\theta_2)}{\partial \rho} = -\frac{N}{2}\text{D}\,\log(|\hat{\Lambda}(\theta_2)|) - \frac{N}{2\sigma^2}\text{D}\,\text{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}]. \tag{A1}$$

Consider $\text{D}\log(|\hat{\Lambda}(\theta_2)|)$. From

$$\text{d}\,\hat{S}_\lambda(\theta_2) = \sigma^2\text{d}\,[F^-(\sigma^{-2}G(\rho) - I_T)F^{-\prime}] = F^-\text{d}\,G(\rho)F^{-\prime},$$

we have

$$\text{d}\,\hat{\Lambda}(\theta_2) = \text{d}\,[I_T + \sigma^{-2}F\hat{S}_\lambda(\theta_2)F'] = \sigma^{-2}F[\text{d}\,\hat{S}_\lambda(\theta_2)]F' = \sigma^{-2}FF^-\text{d}\,G(\rho)F^{-\prime}F'.$$

Moreover, from $\mathrm{d}\,AB = (\mathrm{d}\,A)B + A(\mathrm{d}\,B)$ and $\mathrm{d}\,(A') = (\mathrm{d}\,A)'$,

$$\mathrm{d}\,G(\rho) = \mathrm{d}\,[\Gamma(\rho)^{-1}S_y\Gamma(\rho)^{-1\prime}] = [\mathrm{d}\,\Gamma(\rho)^{-1}]S_y\Gamma(\rho)^{-1\prime} + \Gamma(\rho)^{-1}S_y[\mathrm{d}\,\Gamma(\rho)^{-1}]',$$

and so, via $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}\,B$,

$$\mathrm{vec}\,\mathrm{d}\,G(\rho) = [\Gamma(\rho)^{-1} \otimes \mathrm{d}\,\Gamma(\rho)^{-1} + \mathrm{d}\,\Gamma(\rho)^{-1} \otimes \Gamma(\rho)^{-1}]\mathrm{vec}\,S_y.$$

Here,

$$\frac{\mathrm{d}\,\Gamma(\rho)^{-1}}{\mathrm{d}\rho} = \frac{\mathrm{d}}{\mathrm{d}\rho}\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & \dots & 0 \\ 0 & -\rho & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -\rho & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 0 \end{bmatrix} = -J,$$

from which it follows that

$$\mathrm{D}\,G(\rho) = -[\Gamma(\rho)^{-1} \otimes J + J \otimes \Gamma(\rho)^{-1}]\mathrm{vec}\,S_y = -C(\rho). \tag{A2}$$

with an obvious definition of $C(\rho)$. Hence, since $\mathrm{vec}\,\mathrm{d}\,\hat{\Lambda}(\theta_2) = \sigma^{-2}(FF^- \otimes FF^-)\mathrm{vec}\,\mathrm{d}\,G(\rho)$, we can show that

$$\mathrm{D}\,\hat{\Lambda}(\theta_2) = \sigma^{-2}(FF^- \otimes FF^-)\mathrm{D}\,G(\rho) = -\sigma^{-2}(FF^- \otimes FF^-)C(\rho). \tag{A3}$$

Application of $\mathrm{D}\,\log|A| = [\mathrm{vec}\,(A'^{-1})]'\mathrm{D}\,A$ now yields

$$\begin{aligned} \mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|) &= [\mathrm{vec}\,(\hat{\Lambda}(\theta_2)'^{-1})]'\mathrm{D}\,\hat{\Lambda}(\theta_2) = -\sigma^{-2}[\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1})]'(FF^- \otimes FF^-)C(\rho) \\ &= -\sigma^{-2}[\mathrm{vec}\,(F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-)]'C(\rho) \end{aligned} \tag{A4}$$

where the second equality holds because $\hat{\Lambda}(\theta_2)$ is symmetric, while the third is due to $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}\,B$, or $(\mathrm{vec}\,B)'(C \otimes A') = [\mathrm{vec}(ABC)]'$.

In order to obtain $\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}]$ we use the fact that $\mathrm{tr}\,(A'B) = (\mathrm{vec}\,A)'\mathrm{vec}\,B$, from which it follows that

$$\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] = (\mathrm{vec}\,I_T)'\mathrm{D}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}].$$

By using this, $\mathrm{d}\,AB = (\mathrm{d}\,A)B + A(\mathrm{d}\,B)$, $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}\,B$, and the symmetry of $\hat{\Lambda}(\theta_2)$,

$$\mathrm{D}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] = (\hat{\Lambda}(\theta_2)^{-1} \otimes I_T)\mathrm{D}\,G(\rho) + (I_T \otimes G(\rho))\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}],$$

and, by further use of $\mathrm{d}\, A^{-1} = -A^{-1}(\mathrm{d}\, A)A^{-1}$ and $(A \otimes B)(C \otimes D) = AC \otimes BD$, we obtain

$$\begin{aligned} \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] &= -[\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}]\mathrm{D}\,\hat{\Lambda}(\theta_2) \\ &= \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}](FF^- \otimes FF^-)C(\rho) \\ &= \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}FF^-]C(\rho). \end{aligned} \tag{A5}$$

This implies

$$\begin{aligned} &\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\ &= (\mathrm{vec}\, I_T)'\mathrm{D}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\ &= (\mathrm{vec}\, I_T)'[(\hat{\Lambda}(\theta_2)^{-1} \otimes I_T)\mathrm{D}\, G(\rho) + (I_T \otimes G(\rho))\mathrm{D}\,(\hat{\Lambda}(\theta_2)^{-1})] \\ &= (\mathrm{vec}\, I_T)'[\sigma^{-2}(I_T \otimes G(\rho))[\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}FF^-] - (\hat{\Lambda}(\theta_2)^{-1} \otimes I_T)]C(\rho) \\ &= (\mathrm{vec}\, I_T)'[\sigma^{-2}(\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^-) - (\hat{\Lambda}(\theta_2)^{-1} \otimes I_T)]C(\rho) \\ &= [\mathrm{vec}\,(\sigma^{-2}F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^- - \hat{\Lambda}(\theta_2)^{-1})]'C(\rho), \end{aligned} \tag{A6}$$

where the last equality follows from $(\mathrm{vec}\, B)'(C \otimes A') = [\mathrm{vec}(ABC)]'$, and the symmetry of G. Define $B(\theta_2) = F^{-\prime}F'(\hat{\Lambda}(\theta_2)^{-1} - \sigma^{-2}\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1})FF^- + \hat{\Lambda}(\theta_2)^{-1}$. Insertion of this and above expression for $\mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|)$ into (A1) now yields

$$\begin{aligned} \frac{\partial \ell_c(\theta_2)}{\partial \rho} &= -\frac{N}{2}\mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|) - \frac{N}{2\sigma^2}\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\ &= \frac{N}{2\sigma^2}[\mathrm{vec}\,(F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-)]'C(\rho) \\ &- \frac{N}{2\sigma^2}\mathrm{vec}\,[\sigma^{-2}F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^- - \hat{\Lambda}(\theta_2)^{-1}]'C(\rho) \\ &= \frac{N}{2\sigma^2}[\mathrm{vec}\, B(\theta_2)]'C(\rho), \end{aligned} \tag{A7}$$

as required.

$\partial \ell_c(\theta_2)/\partial \sigma^2$ can be obtained using exactly the same arguments as for $\partial \ell_c(\theta_2)/\partial \rho$. From $\hat{S}_\lambda(\theta_2) = \sigma^2 F^-(\sigma^{-2}G(\rho) - I_T)F^{-\prime}$, we obtain

$$\hat{\Lambda}(\theta_2) = I_T + \sigma^{-2}F\hat{S}_\lambda(\theta_2)F' = I_T + \sigma^{-2}FF^-G(\rho)F^{-\prime}F' - FF^-F^{-\prime}F', \tag{A8}$$

and therefore,

$$\mathrm{d}\,\hat{\Lambda}(\theta_2) = -\sigma^{-4}FF^-G(\rho)F^{-\prime}F',$$

from which it follows that

$$\mathrm{D}\,\hat{\Lambda}(\theta_2) = -\sigma^{-4}(FF^- \otimes FF^-)\mathrm{vec}\, G(\rho). \tag{A9}$$

Hence,

$$\begin{aligned} \mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|) &= -\sigma^{-4}[\mathrm{vec}\,(F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^{-})]'\mathrm{vec}\,G(\rho), \qquad \text{(A10)} \\ \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] &= \sigma^{-4}[\hat{\Lambda}(\theta_2)^{-1}FF^{-}\otimes\hat{\Lambda}(\theta_2)^{-1}FF^{-}]\mathrm{vec}\,G(\rho), \qquad \text{(A11)} \end{aligned}$$

which in turn implies

$$\begin{aligned} \mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] &= (\mathrm{vec}\,I_T)'(I_T\otimes G(\rho))\mathrm{D}\,(\hat{\Lambda}(\theta_2)^{-1}) \\ &= \sigma^{-4}(\mathrm{vec}\,I_T)'(I_T\otimes G(\rho))[\hat{\Lambda}(\theta_2)^{-1}FF^{-}\otimes\hat{\Lambda}(\theta_2)^{-1}FF^{-}]\mathrm{vec}\,G(\rho) \\ &= \sigma^{-4}(\mathrm{vec}\,I_T)'[\hat{\Lambda}(\theta_2)^{-1}FF^{-}\otimes G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^{-}]\mathrm{vec}\,G(\rho) \\ &= \sigma^{-4}\mathrm{vec}\,[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^{-}]'\mathrm{vec}\,G(\rho), \qquad \text{(A12)} \end{aligned}$$

Hence, since

$$\mathrm{D}\,\sigma^{-2}\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] = -\sigma^{-4}\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] + \sigma^{-2}\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}],$$

we can show that

$$\begin{aligned} &\frac{\partial\ell_c(\theta_2)}{\partial\sigma^2} \\ &= -\frac{NT}{2\sigma^2} - \frac{N}{2}\mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|) + \frac{N}{2\sigma^4}\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] - \frac{N}{2\sigma^2}\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\ &= -\frac{NT}{2\sigma^2} + \frac{N}{2}\sigma^{-4}[\mathrm{vec}\,(F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^{-})]'\mathrm{vec}\,G(\rho) + \frac{N}{2\sigma^4}[\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1})]'\mathrm{vec}\,G(\rho) \\ &- \frac{N}{2\sigma^2}\sigma^{-4}\mathrm{vec}\,[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^{-}]'\mathrm{vec}\,G(\rho) \\ &= -\frac{NT}{2\sigma^2} + \frac{N}{2\sigma^4}[\mathrm{vec}\,B(\theta_2)]'\mathrm{vec}\,G(\rho). \qquad \text{(A13)} \end{aligned}$$

Consider $\partial^2\ell_c(\theta_2)/(\partial\rho)^2$. The starting point is

$$\frac{2\sigma^2}{N}\frac{\partial\ell_c(\theta_2)}{\partial\rho} = [\mathrm{vec}\,B(\theta_2)]'C(\rho). \qquad \text{(A14)}$$

Since $\mathrm{vec}\,B(\theta_2)$ and $C(\rho)$ are vectors, we can apply $\mathrm{D}\,A'B = B'(\mathrm{D}\,A) + A'(\mathrm{D}\,B)$ to obtain

$$\frac{2\sigma^2}{N}\frac{\partial^2\ell_c(\theta_2)}{(\partial\rho)^2} = C(\rho)'\mathrm{D}\,\mathrm{vec}\,B(\theta_2) + [\mathrm{vec}\,B(\theta_2)]'\mathrm{D}\,C(\rho). \qquad \text{(A15)}$$

We start with $\mathrm{D}\,C(\rho)$, which is simplest. Indeed, from $\mathrm{d}\,(A\otimes B) = (\mathrm{d}\,A)\otimes B + A\otimes(\mathrm{d}\,B)$ and $\mathrm{d}\,\Gamma(\rho)^{-1}/\mathrm{d}\,\rho = -J$,

$$\mathrm{D}\,C(\rho) = \mathrm{D}\,[\Gamma(\rho)^{-1}\otimes J + J\otimes\Gamma(\rho)^{-1}]\mathrm{vec}\,S_y = -2(J\otimes J)\mathrm{vec}\,S_y. \qquad \text{(A16)}$$

Consider $\mathrm{D}\,\mathrm{vec}\,B(\theta_2)$. We have

$$\begin{aligned}
&\mathrm{vec}\,B(\theta_2)\\
&\quad = \mathrm{vec}\,[F^{-\prime}F'(\hat{\Lambda}(\theta_2)^{-1} - \sigma^{-2}\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1})FF^{-} + \hat{\Lambda}(\theta_2)^{-1}]\\
&\quad = (F^{-\prime}F' \otimes F^{-\prime}F')(\mathrm{vec}\,[\hat{\Lambda}(\theta_2)^{-1}] - \sigma^{-2}\mathrm{vec}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}]) + \mathrm{vec}\,[\hat{\Lambda}(\theta_2)^{-1}],
\end{aligned}$$

implying

$$\mathrm{D}\,\mathrm{vec}\,B(\theta_2) = (F^{-\prime}F' \otimes F^{-\prime}F')(\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] - \sigma^{-2}\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}]) + \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}],$$

where, by repeated use of $\mathrm{d}\,AB = (\mathrm{d}\,A)B + A(\mathrm{d}\,B)$,

$$\begin{aligned}
&\mathrm{d}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}]\\
&\quad = [\mathrm{d}\,\hat{\Lambda}(\theta_2)^{-1}]G(\rho)\hat{\Lambda}(\theta_2)^{-1} + \hat{\Lambda}(\theta_2)^{-1}\mathrm{d}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}]\\
&\quad = [\mathrm{d}\,\hat{\Lambda}(\theta_2)^{-1}]G(\rho)\hat{\Lambda}(\theta_2)^{-1} + \hat{\Lambda}(\theta_2)^{-1}([\mathrm{d}\,G(\rho)]\hat{\Lambda}(\theta_2)^{-1} + G(\rho)[\mathrm{d}\,\hat{\Lambda}(\theta_2)^{-1}]).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}]\\
&= [\hat{\Lambda}(\theta_2)^{-1}G(\rho) \otimes I_T + I_T \otimes \hat{\Lambda}(\theta_2)^{-1}G(\rho)][\mathrm{D}\,\hat{\Lambda}(\theta_2)^{-1}] + [\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}][\mathrm{D}\,G(\rho)]\\
&= \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^{-} \otimes \hat{\Lambda}(\theta_2)^{-1}FF^{-}]C(\rho)\\
&+ \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}FF^{-} \otimes \hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^{-}]C(\rho)\\
&- [\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}]C(\rho), \qquad \text{(A17)}
\end{aligned}$$

where we have used $\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] = \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}FF^{-} \otimes \hat{\Lambda}(\theta_2)^{-1}FF^{-}]C(\rho)$ and $\mathrm{D}\,G(\rho) = -C(\rho)$. It follows that

$$\begin{aligned}
\mathrm{D}\,\mathrm{vec}\,B(\theta_2) &= (F^{-\prime}F' \otimes F^{-\prime}F')\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}]\\
&- \sigma^{-2}(F^{-\prime}F' \otimes F^{-\prime}F')\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}] + \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}]\\
&= \sigma^{-2}[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^{-} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^{-}]C(\rho)\\
&- \sigma^{-4}[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^{-} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^{-}]C(\rho)\\
&- \sigma^{-4}[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^{-} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^{-}]C(\rho)\\
&+ \sigma^{-2}[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}]C(\rho)\\
&+ \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}FF^{-} \otimes \hat{\Lambda}(\theta_2)^{-1}FF^{-}]C(\rho), \qquad \text{(A18)}
\end{aligned}$$

which in turn implies

$$\begin{aligned}
&\frac{2\sigma^2}{N}\frac{\partial^2 \ell_c(\theta_2)}{(\partial\rho)^2} \\
&= \sigma^{-2}C(\rho)'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]C(\rho) \\
&- \sigma^{-4}C(\rho)'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]C(\rho) \\
&- \sigma^{-4}C(\rho)'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^-]C(\rho) \\
&+ \sigma^{-2}C(\rho)'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}]C(\rho) \\
&+ \sigma^{-2}C(\rho)'[\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}FF^-]C(\rho) \\
&- 2[\mathrm{vec}\, B(\theta_2)]'(J \otimes J)\mathrm{vec}\, S_y. \qquad \text{(A19)}
\end{aligned}$$

For $\partial^2 \ell_c(\theta_2)/(\partial\sigma^2)^2$,

$$\frac{\partial^2 \ell_c(\theta_2)}{(\partial\sigma^2)^2} = \frac{NT}{2\sigma^4} - \frac{N}{\sigma^6}[\mathrm{vec}\, B(\theta_2)]'\mathrm{vec}\, G(\rho) + \frac{N}{2\sigma^4}[\mathrm{vec}\, G(\rho)]'\mathrm{D}\,\mathrm{vec}\, B(\theta_2), \qquad \text{(A20)}$$

where

$$\begin{aligned}
&\mathrm{D}\,\mathrm{vec}\, B(\theta_2) \\
&= (F^{-\prime}F' \otimes F^{-\prime}F')(\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] - \sigma^{-2}\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}]) + \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] \\
&= \sigma^{-4}[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&- \sigma^{-6}[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&- \sigma^{-6}[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&+ \sigma^{-4}[\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho), \qquad \text{(A21)}
\end{aligned}$$

as follows from noting that

$$\begin{aligned}
&\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\
&= [\hat{\Lambda}(\theta_2)^{-1}G(\rho) \otimes I_T + I_T \otimes \hat{\Lambda}(\theta_2)^{-1}G(\rho)]\mathrm{D}\,\hat{\Lambda}(\theta_2)^{-1} \\
&= \sigma^{-4}[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&+ \sigma^{-4}[\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho).
\end{aligned}$$

Insertion into the expression for $\partial^2\ell_c(\theta_2)/(\partial\sigma^2)^2$ yields

$$\begin{aligned}
&\frac{\partial^2\ell_c(\theta_2)}{(\partial\sigma^2)^2} \\
&= \frac{NT}{2\sigma^4} - \frac{N}{\sigma^6}[\mathrm{vec}\, B(\theta_2)]'\mathrm{vec}\, G(\rho) \\
&+ \frac{N}{2\sigma^8}[\mathrm{vec}\, G(\rho)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&- \frac{N}{2\sigma^{10}}[\mathrm{vec}\, G(\rho)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&- \frac{N}{2\sigma^{10}}[\mathrm{vec}\, G(\rho)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&+ \frac{N}{2\sigma^8}[\mathrm{vec}\, G(\rho)]'[\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho). \qquad \text{(A22)}
\end{aligned}$$

It remains to consider $\partial^2\ell_c(\theta_2)/(\partial\rho\partial\sigma^2)$. Taking partial derivative of $\partial\ell_c(\theta_2)/\partial\rho$ with respect to σ^2,

$$\begin{aligned}
\frac{\partial^2\ell_c(\theta_2)}{\partial\rho\partial\sigma^2} &= -\frac{N}{2\sigma^4}[\mathrm{vec}\, B(\theta_2)]'C(\rho) + \frac{N}{2\sigma^2}C(\rho)'\mathrm{D}\,\mathrm{vec}\, B(\theta_2) \\
&= -\frac{N}{2\sigma^4}[\mathrm{vec}\, B(\theta_2)]'C(\rho) \\
&+ \frac{N}{2\sigma^6}C(\rho)'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&- \frac{N}{2\sigma^8}C(\rho)'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&- \frac{N}{2\sigma^8}C(\rho)'[F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho) \\
&+ \frac{N}{2\sigma^6}C(\rho)'[\hat{\Lambda}(\theta_2)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2)^{-1}FF^-]\mathrm{vec}\, G(\rho). \qquad \text{(A23)}
\end{aligned}$$

This establishes the last of the required derivatives under C1.

Derivatives under C2

The concentrated objective function has the same form as before, except that F should be replaced by $\Gamma^{-1}F$. Hence, if we let $\hat{\Lambda}(\theta_2) = I_T + \sigma^{-2}\Gamma(\rho)^{-1}F\hat{S}_\lambda(\theta_2)F'\Gamma(\rho)^{-1\prime}$ and $\hat{S}_\lambda(\theta_2) = \sigma^2(\Gamma(\rho)^{-1}F)^-(\sigma^{-2}G(\rho) - I_T)(\Gamma(\rho)^{-1}F)^{-\prime}$, then

$$\ell_c(\theta_2) = -\frac{NT}{2}\log(\sigma^2) - \frac{N}{2}\log(|\hat{\Lambda}(\theta_2)|) - \frac{N}{2\sigma^2}\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}],$$

and therefore

$$\frac{\partial\ell_c(\theta_2)}{\partial\rho} = -\frac{N}{2}\mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|) - \frac{N}{2\sigma^2}\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}].$$

Consider $\mathrm{D}\log(|\hat{\Lambda}(\theta_2)|)$. Repeated use of $\mathrm{d}\,AB = (\mathrm{d}\,A)B + A(\mathrm{d}\,B)$ yields

$$\begin{aligned}
\mathrm{d}\,\hat{S}_\lambda(\theta_2) &= \sigma^2 \mathrm{d}\,[(\Gamma(\rho)^{-1}F)^-(\sigma^{-2}G(\rho) - I_T)(\Gamma(\rho)^{-1}F)^{-\prime}] \\
&= \sigma^2[\mathrm{d}\,(\Gamma(\rho)^{-1}F)^-](\sigma^{-2}G(\rho) - I_T)(\Gamma(\rho)^{-1}F)^{-\prime} \\
&+ \sigma^2(\Gamma(\rho)^{-1}F)^-[\mathrm{d}\,(\sigma^{-2}G(\rho) - I_T)](\Gamma(\rho)^{-1}F)^{-\prime} \\
&+ \sigma^2(\Gamma(\rho)^{-1}F)^-(\sigma^{-2}G(\rho) - I_T)[\mathrm{d}\,(\Gamma(\rho)^{-1}F)^{-\prime}]. \qquad \text{(A24)}
\end{aligned}$$

Letting $H(\rho) = F'\Gamma(\rho)^{-1\prime}\Gamma(\rho)^{-1}F$, we have

$$\mathrm{d}\,(\Gamma(\rho)^{-1}F)^- = \mathrm{d}\,H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime} = [\mathrm{d}\,H(\rho)^{-1}]F'\Gamma(\rho)^{-1\prime} + H(\rho)^{-1}F'\mathrm{d}\,\Gamma(\rho)^{-1\prime}.$$

From $\mathrm{d}\,A^{-1} = -A^{-1}(\mathrm{d}\,A)A^{-1}$, $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}\,B$ and the symmetry of $H(\rho)$,

$$\mathrm{D}\,H(\rho)^{-1} = -[H(\rho)^{-1} \otimes H(\rho)^{-1}]\mathrm{D}\,H(\rho),$$

where

$$\mathrm{d}\,H(\rho) = F'[(\mathrm{d}\,\Gamma(\rho)^{-1\prime})\Gamma(\rho)^{-1} + \Gamma(\rho)^{-1\prime}(\mathrm{d}\,\Gamma(\rho)^{-1})]F,$$

and therefore, by further use of $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}\,B$, $\mathrm{d}\,(A') = (\mathrm{d}\,A)'$ and $A' \otimes B' = (A \otimes B)'$,

$$\begin{aligned}
\mathrm{D}\,H(\rho) &= (F'\Gamma(\rho)^{-1\prime} \otimes F')\mathrm{D}\,\Gamma(\rho)^{-1\prime} + (F' \otimes F'\Gamma(\rho)^{-1\prime})\mathrm{D}\,\Gamma(\rho)^{-1} \\
&= -(\Gamma(\rho)^{-1}F \otimes F)'\mathrm{vec}\,J' - (F \otimes \Gamma(\rho)^{-1}F)'\mathrm{vec}\,J. \qquad \text{(A25)}
\end{aligned}$$

Hence, since $(A \otimes B)(C \otimes D) = AC \otimes BD$ and $\mathrm{d}\,\Gamma(\rho)^{-1}/\mathrm{d}\,\rho = -J$,

$$\begin{aligned}
&\mathrm{D}\,H(\rho)^{-1} \\
&= [H(\rho)^{-1} \otimes H(\rho)^{-1}][(F'\Gamma(\rho)^{-1\prime} \otimes F')\mathrm{vec}\,J' + (F' \otimes F'\Gamma(\rho)^{-1\prime})\mathrm{vec}\,J] \\
&= (H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime} \otimes H(\rho)^{-1}F')\mathrm{vec}\,J' + (H(\rho)^{-1}F' \otimes H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime})\mathrm{vec}\,J,
\end{aligned}$$

which in turn implies, via $A \otimes B + A \otimes C = A \otimes (B + C)$,

$$\begin{aligned}
\mathrm{D}\,(\Gamma(\rho)^{-1}F)^- &= (\Gamma(\rho)^{-1}F \otimes I_m)\mathrm{D}\,H(\rho)^{-1} - (I_T \otimes H(\rho)^{-1}F')\mathrm{vec}\,J' \\
&= (\Gamma(\rho)^{-1}FH(\rho)^{-1}F'\Gamma(\rho)^{-1\prime} \otimes H(\rho)^{-1}F')\mathrm{vec}\,J' \\
&+ (\Gamma(\rho)^{-1}FH(\rho)^{-1}F' \otimes H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime})\mathrm{vec}\,J - (I_T \otimes H(\rho)^{-1}F')\mathrm{vec}\,J' \\
&= [(\Gamma(\rho)^{-1}FH(\rho)^{-1}F'\Gamma(\rho)^{-1\prime} - I_T) \otimes H(\rho)^{-1}F']\mathrm{vec}\,J' \\
&+ (\Gamma(\rho)^{-1}FH(\rho)^{-1}F' \otimes H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime})\mathrm{vec}\,J. \qquad \text{(A26)}
\end{aligned}$$

From this we can further deduce that

$$\begin{aligned}\mathrm{D}\,(\Gamma(\rho)^{-1}F)^{-\prime} &= (H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime}\otimes\Gamma(\rho)^{-1}FH(\rho)^{-1}F')\mathrm{vec}\,J' \\ &+ [H(\rho)^{-1}F'\otimes(\Gamma(\rho)^{-1}FH(\rho)^{-1}F'\Gamma(\rho)^{-1\prime}-I_T)]\mathrm{vec}\,J. \qquad \text{(A27)}\end{aligned}$$

Also, from the results for C1, $\mathrm{D}\,G(\rho) = -C(\rho) = C_0(\rho) = -[\Gamma(\rho)^{-1}\otimes J + J\otimes\Gamma(\rho)^{-1}]\mathrm{vec}\,S_y$. Insertion of this, $\mathrm{D}\,(\Gamma(\rho)^{-1}F)^{-}$ and $\mathrm{D}\,(\Gamma(\rho)^{-1}F)^{-\prime}$ into $\mathrm{D}\,\hat{S}_\lambda(\theta_2)$ now yields

$$\begin{aligned}&\mathrm{D}\,\hat{S}_\lambda(\theta_2) \\ &= \sigma^2[(\Gamma(\rho)^{-1}F)^{-}(\sigma^{-2}G(\rho)-I_T)\otimes I_m]\mathrm{D}\,(\Gamma(\rho)^{-1}F)^{-} \\ &+ [(\Gamma(\rho)^{-1}F)^{-}\otimes(\Gamma(\rho)^{-1}F)^{-}]\mathrm{D}\,G(\rho) \\ &+ \sigma^2[I_m\otimes(\Gamma(\rho)^{-1}F)^{-}(\sigma^{-2}G(\rho)-I_T)]\mathrm{D}\,(\Gamma(\rho)^{-1}F)^{-\prime} \\ &= \sigma^2[(\Gamma(\rho)^{-1}F)^{-}(\sigma^{-2}G(\rho)-I_T)(\Gamma(\rho)^{-1}FH(\rho)^{-1}F'\Gamma(\rho)^{-1\prime}-I_T)\otimes H(\rho)^{-1}F']\mathrm{vec}\,J' \\ &+ \sigma^2[(\Gamma(\rho)^{-1}F)^{-}(\sigma^{-2}G(\rho)-I_T)\Gamma(\rho)^{-1}FH(\rho)^{-1}F'\otimes H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime}]\mathrm{vec}\,J \\ &+ [(\Gamma(\rho)^{-1}F)^{-}\otimes(\Gamma(\rho)^{-1}F)^{-}]C_0(\rho) \\ &+ \sigma^2[H(\rho)^{-1}F'\Gamma(\rho)^{-1\prime}\otimes(\Gamma(\rho)^{-1}F)^{-}(\sigma^{-2}G(\rho)-I_T)\Gamma(\rho)^{-1}FH(\rho)^{-1}F']\mathrm{vec}\,J' \\ &+ \sigma^2[H(\rho)^{-1}F'\otimes(\Gamma(\rho)^{-1}F)^{-}(\sigma^{-2}G(\rho)-I_T)(\Gamma(\rho)^{-1}FH(\rho)^{-1}F'\Gamma(\rho)^{-1\prime}-I_T)]\mathrm{vec}\,J.\end{aligned}$$

Suppressing for simplicity any dependence on θ_2, defining $V = V(\rho) = \Gamma^{-1}FH^{-1}$, and using $A\otimes B + A\otimes C = A\otimes(B+C)$, $a(A\otimes B) = aA\otimes B = A\otimes aB$, $A'\otimes B' = (A\otimes B)'$, $V'\Gamma^{-1}S_y = V'G\Gamma'$ and $H^{-1}F' = V'\Gamma'$,

$$\begin{aligned}&\mathrm{D}\,\hat{S}_\lambda \\ &= \sigma^2[V'(\sigma^{-2}G-I_T)(VF'\Gamma^{-1\prime}-I_T)\otimes H^{-1}F']\mathrm{vec}\,J' \\ &+ \sigma^2[H^{-1}F'\otimes V'(\sigma^{-2}G-I_T)(VF'\Gamma^{-1\prime}-I_T)]\mathrm{vec}\,J \\ &+ \sigma^2[V'(\sigma^{-2}G-I_T)VF'\otimes V']\mathrm{vec}\,J + \sigma^2[V'\otimes V'(\sigma^{-2}G-I_T)VF']\mathrm{vec}\,J' + (V'\otimes V')C_0 \\ &= (V\otimes V)'[\Gamma'\otimes(G-\sigma^2I_T)(VF'\Gamma^{-1\prime}-I_T)]\mathrm{vec}\,J \\ &+ (V\otimes V)'[(G-\sigma^2I_T)(VF'\Gamma^{-1\prime}-I_T)\otimes\Gamma']\mathrm{vec}\,J' \\ &+ (V\otimes V)'[(G-\sigma^2I_T)VF'\otimes I_T]\mathrm{vec}\,J + (V\otimes V)'[I_T\otimes(G-\sigma^2I_T)VF']\mathrm{vec}\,J' \\ &+ (V\otimes V)'C_0,\end{aligned}$$

where

$$
\begin{aligned}
&[\Gamma' \otimes (G - \sigma^2 I_T)(VF'\Gamma^{-1\prime} - I_T)]\mathrm{vec}\, J \\
&= \mathrm{vec}\,[(G - \sigma^2 I_T)(VF'\Gamma^{-1\prime} - I_T)J\Gamma] \\
&= \mathrm{vec}\,[\Gamma^{-1} S_y \Gamma^{-1\prime}(VF'\Gamma^{-1\prime} - I_T)J\Gamma] - \sigma^2 \mathrm{vec}\,[(VF'\Gamma^{-1\prime} - I_T)J\Gamma] \\
&= [\Gamma' J'(VF'\Gamma^{-1\prime} - I_T)\Gamma^{-1} \otimes \Gamma^{-1}]\mathrm{vec}\, S_y - \sigma^2 \mathrm{vec}\,[(VF'\Gamma^{-1\prime} - I_T)J\Gamma].
\end{aligned}
$$

and

$$
\begin{aligned}
&[(G - \sigma^2 I_T)VF' \otimes I_T]\mathrm{vec}\, J \\
&= \mathrm{vec}\,[JFV'(G - \sigma^2 I_T)] = \mathrm{vec}\,(JFV'G) - \sigma^2 \mathrm{vec}\,(JFV') \\
&= \mathrm{vec}\,(JFV'\Gamma^{-1} S_y \Gamma^{-1\prime}) - \sigma^2 \mathrm{vec}\,(JFV') = (\Gamma^{-1} \otimes JFV'\Gamma^{-1})\mathrm{vec}\, S_y - \sigma^2 \mathrm{vec}\,(JFV').
\end{aligned}
$$

Similar calculations reveal that,

$$
\begin{aligned}
&[(G - \sigma^2 I_T)(VF'\Gamma^{-1\prime} - I_T) \otimes \Gamma']\mathrm{vec}\, J' \\
&= [\Gamma^{-1} \otimes \Gamma' J'(VF'\Gamma^{-1\prime} - I_T)\Gamma^{-1}]\mathrm{vec}\, S_y - \sigma^2 \mathrm{vec}\,[\Gamma' J'(\Gamma^{-1} FV' - I_T)],
\end{aligned}
$$

and

$$
[I_T \otimes (G - \sigma^2 I_T)VF']\mathrm{vec}\, J' = (JFV'\Gamma^{-1} \otimes \Gamma^{-1})\mathrm{vec}\, S_y - \sigma^2 \mathrm{vec}\,(VF'J').
$$

Hence, letting

$$
\begin{aligned}
C_1 &= [(\Gamma' J'(VF'\Gamma^{-1\prime} - I_T) + JFV')\Gamma^{-1} \otimes \Gamma^{-1}]\mathrm{vec}\, S_y - \sigma^2 \mathrm{vec}\,[(VF'\Gamma^{-1\prime} - I_T)J\Gamma + VF'J'], \\
C_2 &= [\Gamma^{-1} \otimes (\Gamma' J'(VF'\Gamma^{-1\prime} - I_T) + JFV')\Gamma^{-1}]\mathrm{vec}\, S_y - \sigma^2 \mathrm{vec}\,[\Gamma' J'(\Gamma^{-1} FV' - I_T) + JFV'],
\end{aligned}
$$

which are both functions of θ_2, we obtain

$$
\mathrm{D}\,\hat{S}_\lambda = (V \otimes V)'(C_0 + C_1 + C_2). \tag{A28}
$$

Let us now consider $\hat{\Lambda}(\theta_2) = I_T + \sigma^{-2}\Gamma(\rho)^{-1} F \hat{S}_\lambda(\theta_2) F'\Gamma(\rho)^{-1\prime}$;

$$
\begin{aligned}
\mathrm{d}\,\hat{\Lambda}(\theta_2) &= \sigma^{-2}\mathrm{d}\,[\Gamma(\rho)^{-1} F \hat{S}_\lambda(\theta_2) F'\Gamma(\rho)^{-1\prime}] \\
&= \sigma^{-2}[\mathrm{d}\,\Gamma(\rho)^{-1}] F \hat{S}_\lambda(\theta_2) F'\Gamma(\rho)^{-1\prime} + \sigma^{-2}\Gamma(\rho)^{-1} F [\mathrm{d}\,\hat{S}_\lambda(\theta_2)] F'\Gamma(\rho)^{-1\prime} \\
&+ \sigma^{-2}\Gamma(\rho)^{-1} F \hat{S}_\lambda(\theta_2) F' [\mathrm{d}\,\Gamma(\rho)^{-1\prime}],
\end{aligned}
$$

from which it follows that

$$\begin{aligned}
\mathrm{D}\,\hat{\Lambda}(\theta_2) &= \sigma^{-2}[\Gamma(\rho)^{-1}F\hat{S}_\lambda(\theta_2)F'\otimes I_T]\mathrm{D}\,\Gamma(\rho)^{-1} + \sigma^{-2}[\Gamma(\rho)^{-1}F\otimes\Gamma(\rho)^{-1}F]\mathrm{D}\,\hat{S}_\lambda(\theta_2)\\
&+ \sigma^{-2}[I_T\otimes\Gamma(\rho)^{-1}F\hat{S}_\lambda(\theta_2)F']\mathrm{D}\,\Gamma(\rho)^{-1\prime}\\
&= -\sigma^{-2}[\Gamma(\rho)^{-1}F\hat{S}_\lambda(\theta_2)F'\otimes I_T]\mathrm{vec}\,J + \sigma^{-2}[\Gamma(\rho)^{-1}F\otimes\Gamma(\rho)^{-1}F]\mathrm{D}\,\hat{S}_\lambda(\theta_2)\\
&- \sigma^{-2}[I_T\otimes\Gamma(\rho)^{-1}F\hat{S}_\lambda(\theta_2)F']\mathrm{vec}\,J'. \qquad \text{(A29)}
\end{aligned}$$

Use of $\hat{S}_\lambda = (\Gamma^{-1}F)^-(G-\sigma^2 I_T)(\Gamma^{-1}F)^{-\prime} = V'(G-\sigma^2 I_T)V$ and noting that $\Gamma^{-1}FV'$ is symmetric,

$$\begin{aligned}
\mathrm{D}\,\hat{\Lambda} &= -\sigma^{-2}[\Gamma^{-1}FV'(G-\sigma^2 I_T)VF'\otimes I_T]\mathrm{vec}\,J - \sigma^{-2}[I_T\otimes\Gamma^{-1}FV'(G-\sigma^2 I_T)VF']\mathrm{vec}\,J'\\
&+ \sigma^{-2}(\Gamma^{-1}F\otimes\Gamma^{-1}F)(V\otimes V)'(C_0+C_1+C_2).
\end{aligned}$$

From $V'V = H^{-1}$,

$$\begin{aligned}
&[\Gamma^{-1}FV'(G-\sigma^2 I_T)VF'\otimes I_T]\mathrm{vec}\,J\\
&= \mathrm{vec}\,[JFV'(G-\sigma^2 I_T)VF'\Gamma^{-1\prime}] = \mathrm{vec}\,(JFV'GVF'\Gamma^{-1\prime}) - \sigma^2\mathrm{vec}\,(JFV'VF'\Gamma^{-1\prime})\\
&= \mathrm{vec}\,(JFV'\Gamma^{-1}S_y\Gamma^{-1\prime}VF'\Gamma^{-1\prime}) - \sigma^2\mathrm{vec}\,(JFV')\\
&= (\Gamma^{-1}\otimes J)(FV'\Gamma^{-1}\otimes FV'\Gamma^{-1})\mathrm{vec}\,S_y - \sigma^2\mathrm{vec}\,(JFV'),
\end{aligned}$$

and, by following the same steps,

$$\begin{aligned}
&[I_T\otimes\Gamma^{-1}FV'(G-\sigma^2 I_T)VF']\mathrm{vec}\,J'\\
&= (J\otimes\Gamma^{-1})(FV'\Gamma^{-1}\otimes FV'\Gamma^{-1})\mathrm{vec}\,S_y - \sigma^2\mathrm{vec}\,(VF'J').
\end{aligned}$$

Hence, defining

$$C_3(\theta_2) = C_3 = (\Gamma^{-1}\otimes J + J\otimes\Gamma^{-1})(FV'\Gamma^{-1}\otimes FV'\Gamma^{-1})\mathrm{vec}\,S_y - \sigma^2\mathrm{vec}\,(JFV' + VF'J'),$$

we have

$$\mathrm{D}\,\hat{\Lambda} = -\sigma^{-2}[-(\Gamma^{-1}FV'\otimes\Gamma^{-1}FV')(C_0+C_1+C_2)+C_3] = -\sigma^{-2}\overline{C}, \qquad \text{(A30)}$$

with an implicit definition of $\overline{C} = \overline{C}(\theta_2)$. Application of $\mathrm{D}\,\log|A| = [\mathrm{vec}\,(A'^{-1})]'\mathrm{D}\,A$ now yields

$$\mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|) = [\mathrm{vec}\,(\hat{\Lambda}(\theta_2)'^{-1})]'\mathrm{D}\,\hat{\Lambda}(\theta_2) = -\sigma^{-2}[\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1})]'\overline{C}(\theta_2). \qquad \text{(A31)}$$

The $\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}]$ term has the same form as before. We can therefore make use of the results reported in Proof of Lemma A.1 to arrive at

$$\begin{aligned}\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] &= (\mathrm{vec}\, I_T)'\mathrm{D}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\ &= (\mathrm{vec}\, I_T)'[(\hat{\Lambda}(\theta_2)^{-1}\otimes I_T)\mathrm{D}\, G(\rho) + (I_T\otimes G(\rho))\mathrm{D}\,(\hat{\Lambda}(\theta_2)^{-1})],\end{aligned}$$

where, via $\mathrm{d}\, A^{-1} = -A^{-1}(\mathrm{d}\, A)A^{-1}$,

$$\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] = -[\hat{\Lambda}(\theta_2)^{-1}\otimes\hat{\Lambda}(\theta_2)^{-1}]\mathrm{D}\,\hat{\Lambda}(\theta_2) = \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}\otimes\hat{\Lambda}(\theta_2)^{-1}]\overline{C}(\theta_2). \tag{A32}$$

Hence, since $(\mathrm{vec}\, B)'(C\otimes A') = [\mathrm{vec}(ABC)]'$,

$$\begin{aligned}&\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\ &= (\mathrm{vec}\, I_T)'[(\hat{\Lambda}(\theta_2)^{-1}\otimes I_T)\mathrm{D}\, G(\rho) + (I_T\otimes G(\rho))\mathrm{D}\,(\hat{\Lambda}(\theta_2)^{-1})] \\ &= (\mathrm{vec}\, I_T)'(\hat{\Lambda}(\theta_2)^{-1}\otimes I_T)C_0(\rho) + \sigma^{-2}(\mathrm{vec}\, I_T)'(I_T\otimes G(\rho))[\hat{\Lambda}(\theta_2)^{-1}\otimes\hat{\Lambda}(\theta_2)^{-1}]\overline{C}(\theta_2) \\ &= [\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1})]'C_0(\rho) + \sigma^{-2}[\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1})]'\overline{C}(\theta_2)\end{aligned} \tag{A33}$$

The above results lead to the following expression for $\partial\ell_c(\theta_2)/\partial\rho$:

$$\begin{aligned}\frac{\partial\ell_c(\theta_2)}{\partial\rho} &= -\frac{N}{2}\mathrm{D}\,\log(|\hat{\Lambda}(\theta_2)|) - \frac{N}{2\sigma^2}\mathrm{D}\,\mathrm{tr}\,[G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\ &= \frac{N}{2\sigma^2}[\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1})]'\overline{C}(\theta_2) - \frac{N}{2\sigma^2}[\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1})]'C_0(\rho) \\ &- \frac{N}{2\sigma^2}\sigma^{-2}[\mathrm{vec}\,(\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1})]'\overline{C}(\theta_2) \\ &= -\frac{N}{2\sigma^2}[(\mathrm{vec}\, B_1(\theta_2))'C_0(\rho) - (\mathrm{vec}\, B_2(\theta_2))'\overline{C}(\theta_2)],\end{aligned} \tag{A34}$$

where

$$\begin{aligned}B_1(\theta_2) &= \hat{\Lambda}(\theta_2)^{-1}, \\ B_2(\theta_2) &= \sigma^{-2}\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1} - \hat{\Lambda}(\theta_2)^{-1}.\end{aligned}$$

Consider $\partial^2\ell_c(\theta_2)/(\partial\rho)^2$. As before, the starting point is

$$\frac{2\sigma^2}{N}\frac{\partial\ell_c(\theta_2)}{\partial\rho} = -[(\mathrm{vec}\, B_1(\theta_2))'C_0(\rho) - (\mathrm{vec}\, B_2(\theta_2))'\overline{C}(\theta_2)].$$

From $\mathrm{D}\, A'B = B'(\mathrm{D}\, A) + A'(\mathrm{D}\, B)$,

$$\begin{aligned}\frac{2\sigma^2}{N}\frac{\partial^2\ell_c(\theta_2)}{(\partial\rho)^2} &= -C_0(\rho)'\mathrm{D}\, B_1(\theta_2) - [\mathrm{vec}\, B_1(\theta_2)]'\mathrm{D}\, C_0(\rho) - \overline{C}(\theta_2)'\mathrm{D}\, B_2(\theta_2) \\ &- [\mathrm{vec}\, B_2(\theta_2)]'\mathrm{D}\,\overline{C}(\theta_2).\end{aligned} \tag{A35}$$

We have already shown that $\mathrm{D}\,B_1(\theta_2) = \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] = \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}]\overline{C}(\theta_2)$ and $\mathrm{D}\,C_0(\rho) = -\mathrm{D}\,C(\rho) = 2(J \otimes J)\mathrm{vec}\,S_y$. Let us consider $\mathrm{D}\,B_2(\theta_2)$;

$$\mathrm{D}\,B_2(\theta_2) = \sigma^{-2}\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}] - \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}].$$

Since the second term on the right is already known, we only need to consider the first term, which has the same form as under C1. As in that case,

$$\begin{aligned}
&\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}] \\
&= [\hat{\Lambda}(\theta_2)^{-1}G(\rho) \otimes I_T + I_T \otimes \hat{\Lambda}(\theta_2)^{-1}G(\rho)][\mathrm{D}\,\hat{\Lambda}(\theta_2)^{-1}] + [\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}][\mathrm{D}\,G(\rho)] \\
&= \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1} + \hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}]\overline{C}(\theta_2) \\
&+ [\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}]C_0(\rho), \qquad \text{(A36)}
\end{aligned}$$

showing that, suppressing again the dependence on θ_2,

$$\begin{aligned}
&\mathrm{D}\,B_2 \\
&= \sigma^{-2}\mathrm{D}\,[\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1}] - \mathrm{D}\,(\hat{\Lambda}^{-1}) \\
&= \sigma^{-4}(\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1} + \hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1}G\hat{\Lambda}^{-1})\overline{C} + \sigma^{-2}(\hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1})C_0 \\
&- \sigma^{-2}(\hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1})\overline{C} \\
&= \sigma^{-4}(\hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1})(G\hat{\Lambda}^{-1} \otimes I_T + I_T \otimes G\hat{\Lambda}^{-1} - \sigma^2 I_{T^2})\overline{C} + \sigma^{-2}(\hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1})C_0. \qquad \text{(A37)}
\end{aligned}$$

The only term missing now in $\partial^2\ell_c(\theta_2)/(\partial\rho)^2$ is $\mathrm{D}\,\overline{C}(\theta_2)$. When evaluating this term it is convenient to write C_0, C_1, C_2 and C_3 in vectorized matrix format;

$$C_k = \mathrm{vec}\,(c_k),$$

where

$$\begin{aligned}
c_0 &= -(J\Gamma G + G\Gamma' J'), \\
c_1 &= G((\Gamma^{-1}FV' - I_T)J\Gamma + VF'J') - \sigma^2(VF'\Gamma^{-1\prime} - I_T)J\Gamma - \sigma^2 VF'J', \\
c_2 &= (\Gamma' J'(VF'\Gamma^{-1\prime} - I_T) + JFV')G - \sigma^2\Gamma' J'(\Gamma^{-1}FV' - I_T) - \sigma^2 JFV', \\
c_3 &= JFV'GVF'\Gamma^{-1\prime} + \Gamma^{-1}FV'GVF'J' - \sigma^2 JFV' - \sigma^2 VF'J'.
\end{aligned}$$

In this notation

$$\begin{aligned}
\overline{C} &= -(\Gamma^{-1}FV' \otimes \Gamma^{-1}FV')(C_0 + C_1 + C_2) + C_3 \\
&= -\mathrm{vec}\,[\Gamma^{-1}FV'(c_0 + c_1 + c_2)VF'\Gamma^{-1\prime}] + C_3, \qquad \text{(A38)}
\end{aligned}$$

leading to the following expression for $\mathrm{D}\,\overline{C}(\theta_2)$:

$$\mathrm{D}\,\overline{C} = -\mathrm{D}\,[\Gamma^{-1}FV'(c_0+c_1+c_2)VF'\Gamma^{-1\prime}] + \mathrm{D}\,C_3. \tag{A39}$$

Here

$$\begin{aligned}
&\mathrm{d}\,[\Gamma^{-1}FV'(c_0+c_1+c_2)VF'\Gamma^{-1\prime}] \\
&\quad = \quad (\mathrm{d}\,\Gamma^{-1})FV'(c_0+c_1+c_2)VF'\Gamma^{-1\prime} + \Gamma^{-1}F(\mathrm{d}\,V')(c_0+c_1+c_2)VF'\Gamma^{-1\prime} \\
&\quad + \quad \Gamma^{-1}FV'[\mathrm{d}\,(c_0+c_1+c_2)]VF'\Gamma^{-1\prime} + \Gamma^{-1}FV'(c_0+c_1+c_2)(\mathrm{d}\,V)F'\Gamma^{-1\prime} \\
&\quad + \quad \Gamma^{-1}FV'(c_0+c_1+c_2)VF'(\mathrm{d}\,\Gamma^{-1\prime}),
\end{aligned}$$

and therefore

$$\begin{aligned}
&\mathrm{D}\,[\Gamma^{-1}FV'(c_0+c_1+c_2)VF'\Gamma^{-1\prime}] \\
&\quad = \quad [\Gamma^{-1}FV'(c_0+c_1+c_2)'VF'\otimes I_T]\mathrm{D}\,\Gamma^{-1} + [\Gamma^{-1}FV'(c_0+c_1+c_2)'\otimes\Gamma^{-1}F]\mathrm{D}\,V' \\
&\quad + \quad [\Gamma^{-1}FV'\otimes\Gamma^{-1}FV']\mathrm{D}\,(c_0+c_1+c_2) + [\Gamma^{-1}F\otimes\Gamma^{-1}FV'(c_0+c_1+c_2)]\mathrm{D}\,V \\
&\quad + \quad [I_T\otimes\Gamma^{-1}FV'(c_0+c_1+c_2)VF']\mathrm{D}\,\Gamma^{-1\prime},
\end{aligned} \tag{A40}$$

where all the required derivatives are known, except for $\mathrm{D}\,(c_0+c_1+c_2)$. Let us therefore consider $\mathrm{D}\,c_0$. Since $\mathrm{d}\,\mathrm{vec}\,c_0 = \mathrm{d}\,C_0$, we have that $\mathrm{D}\,c_0 = \mathrm{D}\,C_0 = 2(J\otimes J)\mathrm{vec}\,S_y = 2\mathrm{vec}\,(JS_yJ') = 2\mathrm{vec}\,(J\Gamma G\Gamma'J')$. For $\mathrm{D}\,c_1 = \mathrm{D}\,C_1$, we use

$$\begin{aligned}
\mathrm{d}\,c_1 \quad &= \quad (\mathrm{d}\,G)((\Gamma^{-1}FV'-I_T)J\Gamma+VF'J') + G[\mathrm{d}\,((\Gamma^{-1}FV'-I_T)J\Gamma+VF'J')] \\
&- \quad \sigma^2[\mathrm{d}\,(VF'\Gamma^{-1\prime}-I_T)]J\Gamma - \sigma^2(VF'\Gamma^{-1\prime}-I_T)J(\mathrm{d}\,\Gamma) - \sigma^2(\mathrm{d}\,V)F'J',
\end{aligned}$$

from which it follows that

$$\begin{aligned}
\mathrm{D}\,c_1 \quad &= \quad [(\Gamma'J'(VF'\Gamma^{-1\prime}-I_T)+JFV')\otimes I_T]\mathrm{D}\,G + (I_T\otimes G)\mathrm{D}\,((\Gamma^{-1}FV'-I_T)J\Gamma+VF'J') \\
&- \quad \sigma^2(\Gamma'J'\otimes I_T)\mathrm{D}\,(VF'\Gamma^{-1\prime}-I_T) - \sigma^2[I_T\otimes(VF'\Gamma^{-1\prime}-I_T)J]\mathrm{D}\,\Gamma \\
&- \quad \sigma^2(JF\otimes I_T)\mathrm{D}\,V.
\end{aligned} \tag{A41}$$

Here

$$\mathrm{d}\,((\Gamma^{-1}FV'-I_T)J\Gamma+VF'J') = [\mathrm{d}\,(\Gamma^{-1}FV'-I_T)]J\Gamma + (\Gamma^{-1}FV'-I_T)J(\mathrm{d}\,\Gamma) + (\mathrm{d}\,V)F'J'.$$

giving

$$\begin{aligned}
\mathrm{D}\,((\Gamma^{-1}FV'-I_T)J\Gamma+VF'J') \quad &= \quad (\Gamma'J'\otimes I_T)\mathrm{D}\,(\Gamma^{-1}FV'-I_T) \\
&+ \quad [I_T\otimes(\Gamma^{-1}FV'-I_T)J]\mathrm{D}\,\Gamma + (JF\otimes I_T)\mathrm{D}\,V.
\end{aligned}$$

Let $P_{\Gamma^{-1}F} = \Gamma^{-1}FH^{-1}F'\Gamma^{-1\prime}$ and $M_{\Gamma^{-1}F} = I_T - P_{\Gamma^{-1}F}$. In this notation,

$$\begin{aligned}
\mathrm{D}\,V &= \mathrm{D}\,(\Gamma^{-1}F'H^{-1}) = (H^{-1}F \otimes I_T)\mathrm{D}\,\Gamma^{-1} + (I_T \otimes \Gamma^{-1}F)\mathrm{D}\,H^{-1} \\
&= -(H^{-1}F' \otimes I_T)\mathrm{vec}\,J \\
&+ (I_T \otimes \Gamma^{-1}F)[(H^{-1}F'\Gamma^{-1\prime} \otimes H^{-1}F')\mathrm{vec}\,J' + (H^{-1}F' \otimes H^{-1}F'\Gamma^{-1\prime})\mathrm{vec}\,J] \\
&= (H^{-1}F'\Gamma^{-1\prime} \otimes \Gamma^{-1}FH^{-1}F')\mathrm{vec}\,J' + [H^{-1}F' \otimes (\Gamma^{-1}FH^{-1}F'\Gamma^{-1\prime} - I_T)]\mathrm{vec}\,J \\
&= (H^{-1}F'\Gamma^{-1\prime} \otimes \Gamma^{-1}FH^{-1}F')\mathrm{vec}\,J' - (H^{-1}F' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\,J. \qquad \text{(A42)}
\end{aligned}$$

and therefore

$$\mathrm{D}\,V' = (\Gamma^{-1}FH^{-1}F' \otimes H^{-1}F'\Gamma^{-1\prime})\mathrm{vec}\,J - (M_{\Gamma^{-1}F} \otimes H^{-1}F')\mathrm{vec}\,J', \qquad \text{(A43)}$$

from which we obtain

$$\begin{aligned}
&\mathrm{D}\,(\Gamma^{-1}FV' - I_T) \\
&= (VF' \otimes I_T)\mathrm{D}\,\Gamma^{-1} + (I_T \otimes \Gamma^{-1}F)\mathrm{D}\,V' \\
&= (\Gamma^{-1}FH^{-1}F' \otimes P_{\Gamma^{-1}F} - VF' \otimes I_T)\mathrm{vec}\,J - (M_{\Gamma^{-1}F} \otimes \Gamma^{-1}FH^{-1}F')\mathrm{vec}\,J' \\
&= -(VF' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\,J - (M_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J',
\end{aligned}$$

and

$$\mathrm{D}\,(VF'\Gamma^{-1\prime} - I_T) = -(M_{\Gamma^{-1}F} \otimes VF')\mathrm{vec}\,J' - (P_{\Gamma^{-1}F}\Gamma' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\,J.$$

But we also have

$$\frac{\mathrm{d}\,\Gamma(\rho)}{\mathrm{d}\rho} = \frac{\mathrm{d}}{\mathrm{d}\rho}\begin{bmatrix} 1 & 0 & \dots & 0 \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \rho^{T-1} & \dots & \rho & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 2\rho & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ (T-1)\rho^{T-2} & \dots & 2\rho & 1 & 0 \end{bmatrix} = \Gamma(\rho)J\Gamma(\rho),$$

such that

$$\mathrm{D}\,\Gamma(\rho) = \mathrm{vec}\,[\Gamma(\rho)J\Gamma(\rho)] = [\Gamma(\rho)' \otimes \Gamma(\rho)]\mathrm{vec}\,J. \qquad \text{(A44)}$$

By adding these results,

$$
\begin{aligned}
&\mathrm{D}\left((\Gamma^{-1}FV' - I_T)J\Gamma + VF'J'\right) \\
&\quad = \; -(\Gamma'J'VF' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\, J - (\Gamma'J'M_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\, J' \\
&\quad + \; [I_T \otimes (\Gamma^{-1}FV' - I_T)J](\Gamma' \otimes \Gamma)\mathrm{vec}\, J + (JFH^{-1}F'\Gamma^{-1\prime} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\, J' \\
&\quad - \; (JFH^{-1}F' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\, J \\
&\quad = \; -[(\Gamma'J' + J\Gamma)P_{\Gamma^{-1}F}\Gamma' \otimes M_{\Gamma^{-1}F}]\mathrm{vec}\, J + [(J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F}) \otimes P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\, J' \\
&\quad - \; (I_T \otimes M_{\Gamma^{-1}F}J)(\Gamma' \otimes \Gamma)\mathrm{vec}\, J,
\end{aligned}
$$

which can in turn be inserted into the expression for $\mathrm{D}\, c_1$ giving

$$
\begin{aligned}
&\mathrm{D}\, c_1 \\
&\quad = \; -[(\Gamma'J'M_{\Gamma^{-1}F} - JFV') \otimes I_T]C_0 \\
&\quad - \; [(\Gamma'J' + J\Gamma)P_{\Gamma^{-1}F}\Gamma' \otimes GM_{\Gamma^{-1}F}]\mathrm{vec}\, J + \sigma^2[(\Gamma'J' + J\Gamma)P_{\Gamma^{-1}F}\Gamma' \otimes M_{\Gamma^{-1}F}]\mathrm{vec}\, J \\
&\quad + \; [(J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F}) \otimes GP_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\, J' - \sigma^2[(J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F}) \otimes P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\, J' \\
&\quad - \; [I_T \otimes GM_{\Gamma^{-1}F}J](\Gamma' \otimes \Gamma)\mathrm{vec}\, J + \sigma^2[I_T \otimes M_{\Gamma^{-1}F}J](\Gamma' \otimes \Gamma)\mathrm{vec}\, J \\
&\quad = \; [(J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F}) \otimes I_T]C_0 - [(\Gamma'J' + J\Gamma)P_{\Gamma^{-1}F}\Gamma' \otimes (G - \sigma^2 I_T)M_{\Gamma^{-1}F}]\mathrm{vec}\, J \\
&\quad + \; [(J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F}) \otimes (G - \sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\, J' \\
&\quad - \; [\Gamma' \otimes (G - \sigma^2 I_T)M_{\Gamma^{-1}F}J\Gamma]\mathrm{vec}\, J, \qquad \text{(A45)}
\end{aligned}
$$

and because of symmetry we also have

$$
\begin{aligned}
\mathrm{D}\, c_2 \;&=\; [I_T \otimes (J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F})]C_0 - [(G - \sigma^2 I_T)M_{\Gamma^{-1}F} \otimes (\Gamma'J' + J\Gamma)P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\, J' \\
&+\; [(G - \sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma' \otimes (J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F})]\mathrm{vec}\, J \\
&-\; [(G - \sigma^2 I_T)M_{\Gamma^{-1}F}J\Gamma \otimes \Gamma']\mathrm{vec}\, J'. \qquad \text{(A46)}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\mathrm{D}\,(c_0+c_1+c_2)\\
&\quad= \mathrm{D}\,C_0+[(J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})\otimes I_T + I_T\otimes (J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})]C_0\\
&\quad- [\Gamma'\otimes (G-\sigma^2 I_T)M_{\Gamma^{-1}F}J\Gamma]\mathrm{vec}\,J - [(G-\sigma^2 I_T)M_{\Gamma^{-1}F}J\Gamma\otimes\Gamma']\mathrm{vec}\,J'\\
&\quad- [(\Gamma' J'+J\Gamma)P_{\Gamma^{-1}F}\Gamma'\otimes (G-\sigma^2 I_T)M_{\Gamma^{-1}F}]\mathrm{vec}\,J\\
&\quad- [(G-\sigma^2 I_T)M_{\Gamma^{-1}F}\otimes (\Gamma' J'+J\Gamma)P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\,J'\\
&\quad+ [(G-\sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma'\otimes (J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})]\mathrm{vec}\,J\\
&\quad+ [(J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})\otimes (G-\sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\,J'.
\end{aligned}
$$

Insertion into (A40) now gives

$$
\begin{aligned}
&\mathrm{D}\,[\Gamma^{-1}FV'(c_0+c_1+c_2)VF'\Gamma^{-1\prime}]\\
&\quad= [\Gamma^{-1}FV'(c_0+c_1+c_2)'VF'\otimes I_T]\mathrm{D}\,\Gamma^{-1} + [\Gamma^{-1}FV'(c_0+c_1+c_2)'\otimes\Gamma^{-1}F]\mathrm{D}\,V'\\
&\quad+ [\Gamma^{-1}FV'\otimes\Gamma^{-1}FV']\mathrm{D}\,(c_0+c_1+c_2) + [\Gamma^{-1}F\otimes\Gamma^{-1}FV'(c_0+c_1+c_2)]\mathrm{D}\,V\\
&\quad+ [I_T\otimes\Gamma^{-1}FV'(c_0+c_1+c_2)VF']\mathrm{D}\,\Gamma^{-1\prime}\\
&\quad= -[P_{\Gamma^{-1}F}(c_0+c_1+c_2)'P_{\Gamma^{-1}F}\Gamma'\otimes M_{\Gamma^{-1}F}]\mathrm{vec}\,J\\
&\quad- [P_{\Gamma^{-1}F}(c_0+c_1+c_2)'M_{\Gamma^{-1}F}\otimes P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\,J'\\
&\quad+ [P_{\Gamma^{-1}F}\otimes P_{\Gamma^{-1}F}]\mathrm{D}\,C_0\\
&\quad+ [P_{\Gamma^{-1}F}(J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})\otimes P_{\Gamma^{-1}F} + P_{\Gamma^{-1}F}\otimes P_{\Gamma^{-1}F}(J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})]C_0\\
&\quad- [P_{\Gamma^{-1}F}\Gamma'\otimes P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F}J\Gamma]\mathrm{vec}\,J\\
&\quad- [P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F}J\Gamma\otimes P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\,J'\\
&\quad- [P_{\Gamma^{-1}F}(\Gamma' J'+J\Gamma)P_{\Gamma^{-1}F}\Gamma'\otimes P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F}]\mathrm{vec}\,J\\
&\quad- [P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F}\otimes P_{\Gamma^{-1}F}(\Gamma' J'+J\Gamma)P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\,J'\\
&\quad+ [P_{\Gamma^{-1}F}(G-\sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma'\otimes P_{\Gamma^{-1}F}(J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})]\mathrm{vec}\,J\\
&\quad+ [P_{\Gamma^{-1}F}(J\Gamma P_{\Gamma^{-1}F}-\Gamma' J' M_{\Gamma^{-1}F})\otimes P_{\Gamma^{-1}F}(G-\sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\,J'\\
&\quad- [M_{\Gamma^{-1}F}\otimes P_{\Gamma^{-1}F}(c_0+c_1+c_2)P_{\Gamma^{-1}F}\Gamma']\mathrm{vec}\,J'\\
&\quad- [P_{\Gamma^{-1}F}\Gamma'\otimes P_{\Gamma^{-1}F}(c_0+c_1+c_2)M_{\Gamma^{-1}F}]\mathrm{vec}\,J. \qquad \text{(A47)}
\end{aligned}
$$

It remains to consider $\mathrm{D}\,C_3$, which we expand as follows:

$$
\mathrm{D}\,C_3 = \mathrm{D}\,(JFV'GVF'\Gamma^{-1\prime}) + \mathrm{D}\,(\Gamma^{-1}FV'GVF'J') - \sigma^2\mathrm{D}\,(JFV') - \sigma^2\mathrm{D}\,(VF'J'), \qquad \text{(A48)}
$$

where

$$\begin{aligned}\mathrm{d}\,(JFV'GVF'\Gamma^{-1\prime}) &= JF(\mathrm{d}\,V')GVF'\Gamma^{-1\prime} + JFV'(\mathrm{d}\,G)VF'\Gamma^{-1\prime} + JFV'G(\mathrm{d}\,V)F'\Gamma^{-1\prime} \\ &+ JFV'GVF'(\mathrm{d}\,\Gamma^{-1\prime})\end{aligned}$$

giving

$$\begin{aligned}&\mathrm{D}\,(JFV'GVF'\Gamma^{-1\prime}) \\ &= (\Gamma^{-1}FV'G \otimes JF)\mathrm{D}\,V' + (\Gamma^{-1}FV' \otimes JFV')\mathrm{D}\,G + (\Gamma^{-1}F \otimes JFV'G)\mathrm{D}\,V \\ &+ (I_T \otimes JFV'GVF')\mathrm{D}\,\Gamma^{-1\prime} \\ &= (P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F}\Gamma' \otimes J\Gamma P_{\Gamma^{-1}F})\mathrm{vec}\,J - (P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F} \otimes J\Gamma P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' \\ &+ (P_{\Gamma^{-1}F} \otimes J\Gamma P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' - (P_{\Gamma^{-1}F}\Gamma' \otimes J\Gamma P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F})\mathrm{vec}\,J \\ &- (I_T \otimes J\Gamma P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' + (P_{\Gamma^{-1}F} \otimes J\Gamma P_{\Gamma^{-1}F})C_0 \\ &= (P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F}\Gamma' \otimes J\Gamma P_{\Gamma^{-1}F})\mathrm{vec}\,J - (P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F} \otimes J\Gamma P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' \\ &- (M_{\Gamma^{-1}F} \otimes J\Gamma P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' - (P_{\Gamma^{-1}F}\Gamma' \otimes J\Gamma P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F})\mathrm{vec}\,J \\ &+ (P_{\Gamma^{-1}F} \otimes J\Gamma P_{\Gamma^{-1}F})C_0.\end{aligned}$$

The matrix product in second term of $\mathrm{D}\,C_3$ is just the transpose of the product in the first term. Hence,

$$\begin{aligned}&\mathrm{D}\,(\Gamma^{-1}FV'GVF'J') \\ &= (J\Gamma P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' - (J\Gamma P_{\Gamma^{-1}F}\Gamma' \otimes P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F})\mathrm{vec}\,J \\ &- (J\Gamma P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F}\Gamma' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\,J - (J\Gamma P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' \\ &+ (J\Gamma P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F})C_0.\end{aligned}$$

The third and fourth terms are given by

$$\begin{aligned}\mathrm{D}\,(JFV') &= (I_T \otimes JF)\mathrm{D}\,V' \\ &= (\Gamma^{-1}FH^{-1}F' \otimes JFH^{-1}F'\Gamma^{-1\prime})\mathrm{vec}\,J - (M_{\Gamma^{-1}F} \otimes JFH^{-1}F')\mathrm{vec}\,J' \\ &= (P_{\Gamma^{-1}F}\Gamma' \otimes J\Gamma P_{\Gamma^{-1}F})\mathrm{vec}\,J - (M_{\Gamma^{-1}F} \otimes J\Gamma P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J',\end{aligned}$$

and

$$\mathrm{D}\,(VF'J') = (J\Gamma P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' - (J\Gamma P_{\Gamma^{-1}F}\Gamma' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\,J,$$

respectively. Insertion into (A48) gives, after simplification,

$$\begin{aligned}
\mathrm{D}\,C_3 &= (I_T \otimes J\Gamma P_{\Gamma^{-1}F} - J\Gamma \otimes M_{\Gamma^{-1}F})(P_{\Gamma^{-1}F}(G-\sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma' \otimes I_T)\mathrm{vec}\,J \\
&+ (J\Gamma P_{\Gamma^{-1}F} \otimes I_T - M_{\Gamma^{-1}F} \otimes J\Gamma)(I_T \otimes P_{\Gamma^{-1}F}(G-\sigma^2 I_T)P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' \\
&- (I_T \otimes J\Gamma + J\Gamma \otimes I_T)(P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' \\
&- (I_T \otimes J\Gamma + J\Gamma \otimes I_T)(P_{\Gamma^{-1}F}\Gamma' \otimes P_{\Gamma^{-1}F}GM_{\Gamma^{-1}F})\mathrm{vec}\,J \\
&+ (I_T \otimes J\Gamma + J\Gamma \otimes I_T)(P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F})C_0. \qquad \text{(A49)}
\end{aligned}$$

The corresponding expression for $\mathrm{D}\,\overline{C}$ in (A39) can be obtained by using this and the result for $\mathrm{D}\,[\Gamma^{-1}FV'(c_0+c_1+c_2)VF'\Gamma^{-1\prime}]$. The expression for $\mathrm{D}\,\overline{C}$, together with those for C_0, $\mathrm{D}\,B_1$, B_1, $\mathrm{D}\,C_0$, $\overline{C}$, $\mathrm{D}\,B_2$ and B_2, can in turn be inserted into (A35) to obtain the required expression for $\partial^2\ell_c(\theta_2)/(\partial\rho)^2$.

$\partial^2\ell_c(\theta_2)/(\partial\rho\partial\sigma^2)$ remains. Note that

$$\frac{2\sigma^2}{N}\frac{\partial^2\ell_c(\theta_2)}{\partial\rho\partial\sigma^2} = -C_0(\rho)'\mathrm{D}\,B_1(\theta_2) - \overline{C}(\theta_2)'\mathrm{D}\,B_2(\theta_2) - [\mathrm{vec}\,B_2(\theta_2)]'\mathrm{D}\,\overline{C}(\theta_2). \qquad \text{(A50)}$$

As under C1,

$$\mathrm{D}\,B_1(\theta_2) = \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] = \sigma^{-4}[\hat{\Lambda}(\theta_2)^{-1}P_{\Gamma^{-1}F} \otimes \hat{\Lambda}(\theta_2)^{-1}P_{\Gamma^{-1}F}]\mathrm{vec}\,G(\rho). \qquad \text{(A51)}$$

Also,

$$\begin{aligned}
\mathrm{d}\,B_2(\theta_2) &= \mathrm{d}\,[\sigma^{-2}\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}] - \mathrm{d}\,[\hat{\Lambda}(\theta_2)^{-1}] \\
&= (\mathrm{d}\,\sigma^{-2})\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1} + \sigma^{-2}\mathrm{d}\,[\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1}] - \mathrm{d}\,[\hat{\Lambda}(\theta_2)^{-1}] \\
&= \sigma^{-2}\hat{\Lambda}(\theta_2)^{-1}G(\rho)\hat{\Lambda}(\theta_2)^{-1} + \sigma^{-2}[\mathrm{d}\,\hat{\Lambda}(\theta_2)^{-1}]G(\rho)\hat{\Lambda}(\theta_2)^{-1} \\
&+ \sigma^{-2}\hat{\Lambda}(\theta_2)^{-1}G(\rho)\mathrm{d}\,[\hat{\Lambda}(\theta_2)^{-1}] - \mathrm{d}\,[\hat{\Lambda}(\theta_2)^{-1}].
\end{aligned}$$

from which we obtain

$$\begin{aligned}
\mathrm{D}\,B_2(\theta_2) &= -\sigma^{-4}[\hat{\Lambda}(\theta_2)^{-1} \otimes \hat{\Lambda}(\theta_2)^{-1}]\mathrm{vec}\,G(\rho) + \sigma^{-2}[\hat{\Lambda}(\theta_2)^{-1}G(\rho) \otimes I_T]\mathrm{D}\,\hat{\Lambda}(\theta_2)^{-1} \\
&+ \sigma^{-2}[I_T \otimes \hat{\Lambda}(\theta_2)^{-1}G(\rho)]\mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}] - \mathrm{D}\,[\hat{\Lambda}(\theta_2)^{-1}],
\end{aligned}$$

or, suppressing the dependence on θ_2,

$$\begin{aligned}
\mathrm{D}\,B_2 &= -\sigma^{-4}(\hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1})\mathrm{vec}\,G + [\sigma^{-2}(\hat{\Lambda}^{-1}G \otimes I_T) + \sigma^{-2}(I_T \otimes \hat{\Lambda}^{-1}G) - I_{T^2}]\mathrm{D}\,(\hat{\Lambda}^{-1}) \\
&= -\sigma^{-4}(\hat{\Lambda}^{-1} \otimes \hat{\Lambda}^{-1})\mathrm{vec}\,G + [\sigma^{-2}(\hat{\Lambda}^{-1}G \otimes I_T) + \sigma^{-2}(I_T \otimes \hat{\Lambda}^{-1}G) - I_{T^2}] \\
&\times (\hat{\Lambda}^{-1}P_{\Gamma^{-1}F} \otimes \hat{\Lambda}^{-1}P_{\Gamma^{-1}F})\mathrm{vec}\,G. \qquad \text{(A52)}
\end{aligned}$$

For $\overline{C}$, since C_1, C_2 and C_3 are already in vector format, we have

$$\mathrm{D}\,C_1 = -\sigma^{-2}\mathrm{vec}\,[(VF'\Gamma^{-1\prime} - I_T)J\Gamma + VF'J'] = -\sigma^{-2}\mathrm{vec}\,(P_{\Gamma^{-1}F}\Gamma'J' - M_{\Gamma^{-1}F}J\Gamma), \quad \text{(A53)}$$

$$\mathrm{D}\,C_2 = -\sigma^{-2}\mathrm{vec}\,[\Gamma'J'(\Gamma^{-1}FV' - I_T) + JFV'] = -\sigma^{-2}\mathrm{vec}\,(J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F}), \quad \text{(A54)}$$

$$\mathrm{D}\,C_3 = -\sigma^{-2}\mathrm{vec}\,(JFV' + VF'J') = -\sigma^{-2}\mathrm{vec}\,(J\Gamma P_{\Gamma^{-1}F} + P_{\Gamma^{-1}F}\Gamma'J'). \quad \text{(A55)}$$

With $\mathrm{D}\,C_0 = 0_{T^2\times 1}$, this yields

$$\begin{aligned}
\mathrm{D}\,\overline{C} &= -(\Gamma^{-1}FV' \otimes \Gamma^{-1}FV')\mathrm{D}\,(C_1 + C_2) + \mathrm{D}\,C_3 \\
&= \sigma^{-2}(P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F})\mathrm{vec}\,(P_{\Gamma^{-1}F}\Gamma'J' - M_{\Gamma^{-1}F}J\Gamma + J\Gamma P_{\Gamma^{-1}F} - \Gamma'J'M_{\Gamma^{-1}F}) \\
&- \sigma^{-2}\mathrm{vec}\,(J\Gamma P_{\Gamma^{-1}F} + P_{\Gamma^{-1}F}\Gamma'J') \\
&= \sigma^{-2}(P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F})(I_T \otimes P_{\Gamma^{-1}F}\Gamma' - M_{\Gamma^{-1}F} \otimes \Gamma')\mathrm{vec}\,J' \\
&+ \sigma^{-2}(P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F})(P_{\Gamma^{-1}F}\Gamma' \otimes I_T - \Gamma' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\,J \\
&- \sigma^{-2}(P_{\Gamma^{-1}F}\Gamma' \otimes I_T)\mathrm{vec}\,J - \sigma^{-2}(I_T \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' \\
&= \sigma^{-2}(P_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' + \sigma^{-2}(P_{\Gamma^{-1}F}\Gamma' \otimes P_{\Gamma^{-1}F})\mathrm{vec}\,J \\
&- \sigma^{-2}(P_{\Gamma^{-1}F}\Gamma' \otimes I_T)\mathrm{vec}\,J - \sigma^{-2}(I_T \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' \\
&= -\sigma^{-2}(M_{\Gamma^{-1}F} \otimes P_{\Gamma^{-1}F}\Gamma')\mathrm{vec}\,J' - \sigma^{-2}(P_{\Gamma^{-1}F}\Gamma' \otimes M_{\Gamma^{-1}F})\mathrm{vec}\,J. \quad \text{(A56)}
\end{aligned}$$

The required expression for $\partial^2\ell_c(\theta_2)/(\partial\rho\partial\sigma^2)$ is implied by this.

Appendix C: Proofs of main results

Proof of Lemma 1.

Let $K = K(\theta_2) = (\sigma^2\hat{S}_\lambda(\theta_2)^{-1} + F'F)^{-1}$. Application of $(A + CBC')^{-1} = A^{-1} - A^{-1}C(B^{-1} + C'A^{-1}C)^{-1}C'A^{-1}$ to $\hat{\Lambda}^{-1}$ yields

$$\hat{\Lambda}^{-1} = I_T - F(\sigma^2\hat{S}_\lambda^{-1} + F'F)^{-1}F' = I_T - FKF'.$$

Since $\mathrm{tr}\,(A + B) = \mathrm{tr}\,A + \mathrm{tr}\,B$ and $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$, we can show that

$$Q_c = T\log(\sigma^2) + \log(|\hat{\Lambda}|) + \sigma^{-2}\mathrm{tr}\,G - \sigma^{-2}\mathrm{tr}\,(GFKF'), \quad \text{(A57)}$$

where $G = G(\rho) = \Gamma(\rho)^{-1}S_y\Gamma(\rho)^{-1\prime}$.

In order to establish the required result we need to evaluate each of the right-hand side terms of (A150). We begin with $\sigma^{-2}\mathrm{tr}\,(GFKF')$. By the definition of F^-,

$$\begin{aligned}
\hat{S}_\lambda &= \sigma^2F^-(\sigma^{-2}G - I_T)F^{-\prime} = \sigma^2(F'F)^{-1}F'(\sigma^{-2}G - I_T)F(F'F)^{-1} \\
&= (F'F)^{-1}F'GF(F'F)^{-1} - \sigma^2(F'F)^{-1}. \quad \text{(A58)}
\end{aligned}$$

By using this and $(A+CBC')^{-1} = A^{-1} - A^{-1}C(B^{-1}+C'A^{-1}C)^{-1}C'A^{-1}$ we obtain

$$\begin{aligned} K &= (\sigma^2\hat{S}_\lambda^{-1} + F'F)^{-1} = (F'F)^{-1} - (F'F)^{-1}(\sigma^{-2}\hat{S}_\lambda + (F'F)^{-1})^{-1}(F'F)^{-1} \\ &= (F'F)^{-1} - \sigma^2(F'GF)^{-1}, \end{aligned} \tag{A59}$$

suggesting that

$$\begin{aligned} \text{tr}\,(GFKF') &= \text{tr}\,(F'GFK) = \text{tr}\,[F'GF((F'F)^{-1} - \sigma^2(F'GF)^{-1})] \\ &= \text{tr}\,[F'GF(F'F)^{-1}] - \sigma^2\text{tr}\,I_m = \text{tr}\,[F'GF(F'F)^{-1}] - \sigma^2 m. \end{aligned} \tag{A60}$$

Consider $F'GF$. In particular, let us consider S_y. Clearly, this quantity only depends on the true values of ρ and σ^2, ρ_0 and σ_0^2. Hence, writing Γ_0 for $\Gamma(\rho_0)$, we have

$$\begin{aligned} S_y &= \Gamma_0 S_u \Gamma_0' = \Gamma_0 \frac{1}{N}\sum_{i=1}^N (F\lambda_i + \varepsilon_i)(F\lambda_i + \varepsilon_i)'\Gamma_0' \\ &= \Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0' + \Gamma_0 \frac{1}{N}\sum_{i=1}^N F\lambda_i\varepsilon_i'\Gamma_0' + \Gamma_0\frac{1}{N}\sum_{i=1}^N \varepsilon_i\lambda_i'F'\Gamma_0' \\ &+ \Gamma_0 \frac{1}{N}\sum_{i=1}^N (\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)\Gamma_0', \end{aligned} \tag{A61}$$

where the third equality follows from adding and subtracting $\sigma_0^2\Gamma_0\Gamma_0'$. It follows that

$$\begin{aligned} T^{-2}F'GF &= T^{-2}F'\Gamma^{-1}S_y\Gamma^{-1\prime}F = T^{-1}F'\Gamma^{-1}\Gamma_0 S_u\Gamma_0'\Gamma^{-1\prime}F \\ &= T^{-2}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}F + T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^N F\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}F \\ &+ T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^N \varepsilon_i\lambda_i'F'\Gamma_0'\Gamma^{-1\prime}F \\ &+ T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^N (\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}F, \end{aligned} \tag{A62}$$

We now evaluate each term on the right-hand side. The first term can be expanded in the following fashion:

$$\begin{aligned} &T^{-2}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}F \\ &= \sigma_0^2 T^{-2}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F + T^{-2}F'\Gamma^{-1}\Gamma_0 FS_\lambda F'\Gamma_0'\Gamma^{-1\prime}F. \end{aligned} \tag{A63}$$

Consider $T^{-2}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F$. From $\Gamma^{-1}\Gamma_0 = I_T + (\rho_0 - \rho)L_0$, $||AB|| \leq ||A||||B||$, $||A+B|| \leq ||A|| + ||B||$, and the assumed properties of the moments in F and L_0, we obtain

$$\begin{aligned} &||T^{-1}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F|| \\ &= ||T^{-1}F'[I_T + (\rho_0-\rho)L_0][I_T + (\rho_0-\rho)L_0]'F|| \\ &\leq ||T^{-1}F'F|| + 2|\rho_0 - \rho|||T^{-1}F'L_0'F|| + (\rho_0-\rho)^2||T^{-1}F'L_0L_0'F|| \leq C, \end{aligned}$$

implying

$$||T^{-2}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F|| = O(T^{-1}).$$

As for the second term on the right-hand side of (A63), by substitution of $\Gamma^{-1}\Gamma_0 = I_T + (\rho_0 - \rho)L_0$,

$$\begin{aligned} F'\Gamma^{-1}\Gamma_0 F S_\lambda F'\Gamma_0'\Gamma^{-1\prime}F &= F'[I_T + (\rho_0-\rho)L_0]FS_\lambda F'[I_T + (\rho_0-\rho)L_0]'F \\ &= F'FS_\lambda F'F + (\rho_0-\rho)F'FS_\lambda F'L_0'F + (\rho_0-\rho)F'L_0FS_\lambda F'F \\ &+ (\rho_0-\rho)^2F'L_0FS_\lambda F'L_0'F, \end{aligned} \tag{A64}$$

which can be substituted back into (A63), giving

$$\begin{aligned} & T^{-2}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}F \\ &= \sigma_0^2 T^{-2}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F + T^{-2}F'\Gamma^{-1}\Gamma_0 FS_\lambda F'\Gamma_0'\Gamma^{-1\prime}F \\ &= (T^{-1}F'F)S_\lambda(T^{-1}F'F) + (\rho_0-\rho)(T^{-1}F'F)S_\lambda T^{-1}F'L_0'F \\ &+ (\rho_0-\rho)T^{-1}F'L_0FS_\lambda(T^{-1}F'F) + (\rho_0-\rho)^2T^{-1}F'L_0FS_\lambda T^{-1}F'L_0'F + O(T^{-1}). \end{aligned} \tag{A65}$$

The effect of the second term on the right of (A62) can be deduced from

$$\begin{aligned} & \left|\left|T^{-1}F'\Gamma^{-1}\Gamma_0 F\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}F\right|\right| \\ &= \left|\left|T^{-1}F'[I_T + (\rho_0-\rho)L_0]F\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'[I_T + (\rho_0-\rho)L_0]'F\right|\right| \\ &\le ||T^{-1}F'F||\left|\left|\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'F\right|\right| + |\rho_0-\rho|||T^{-1}F'F||\left|\left|\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'L_0'F\right|\right| \\ &+ |\rho_0-\rho|||T^{-1}F'L_0F||\left|\left|\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'F\right|\right| + (\rho_0-\rho)^2||T^{-1}F'L_0F||\left|\left|\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'L_0'F\right|\right|. \end{aligned}$$

By using $E(\varepsilon_i\varepsilon_i') = \sigma_0^2 I_T$ and the fact that $\varepsilon_i'FF'\varepsilon_j$ is just a scalar,

$$\begin{aligned} & E\left(\left|\left|\frac{1}{\sqrt{NT}}\sum_{i=1}^N \lambda_i\varepsilon_i'F\right|\right|^2\right) \\ &= \frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N E[\mathrm{tr}(\lambda_i\varepsilon_i'FF'\varepsilon_j\lambda_j')] = \frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N E(\varepsilon_i'FF'\varepsilon_j)\mathrm{tr}(\lambda_i\lambda_j') \\ &= \frac{1}{NT}\sum_{i=1}^N\sum_{j=1}^N \mathrm{tr}[E(\varepsilon_j\varepsilon_i')FF']\mathrm{tr}(\lambda_i\lambda_j') = \frac{1}{NT}\sum_{i=1}^N \mathrm{tr}[E(\varepsilon_i\varepsilon_i')FF']\mathrm{tr}(\lambda_i\lambda_i') \\ &= \sigma_0^2\mathrm{tr}(T^{-1}F'F)\mathrm{tr}(S_\lambda) \le C, \end{aligned}$$

and by repeated use of the same argument,

$$E\left(\left|\left|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i\varepsilon_i'L_0'F\right|\right|^2\right) = \sigma_0^2\text{tr}(T^{-1}F'L_0L_0'F)\text{tr}(S_\lambda) \leq C,$$

suggesting that $||(NT)^{-1/2}\sum_{i=1}^{N}\lambda_i\varepsilon_i'F||$ and $||(NT)^{-1/2}\sum_{i=1}^{N}\lambda_i\varepsilon_i'L_0'F||$ are $O_p(1)$. The order of the second term on the right-hand side of (A62) is therefore given by

$$\left|\left|T^{-1}F'\Gamma^{-1}\Gamma_0F\frac{1}{NT}\sum_{i=1}^{N}\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}F\right|\right| = O_p((NT)^{-1/2}). \tag{A66}$$

The effect of the third term is of the same order.

It remains to consider the fourth term, which can be expanded in the following fashion:

$$\begin{aligned}
&T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)\Gamma_0'\Gamma^{-1\prime}F \\
&= T^{-1}F'[I_T+(\rho_0-\rho)L_0]\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)[I_T+(\rho_0-\rho)L_0]'F \\
&= T^{-1}F'\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)F+(\rho_0-\rho)T^{-1}F'\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)L_0'F \\
&+ (\rho_0-\rho)T^{-1}F'L_0\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)F \\
&+ (\rho_0-\rho)^2T^{-1}F'L_0\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)L_0'F.
\end{aligned} \tag{A67}$$

Consider the first term on the right-hand side. We have

$$\begin{aligned}
&E\left(\left|\left|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}F'(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)F\right|\right|^2\right) \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\text{tr}(E[F'(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)FF'(\varepsilon_j\varepsilon_j'-\sigma_0^2I_T)F]) \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}\text{tr}(E[F'(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)FF'(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)F]) \\
&+ \frac{1}{NT^2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\text{tr}(F'E[(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)]FF'E[(\varepsilon_j\varepsilon_j'-\sigma_0^2I_T)]F) \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}\text{tr}[F'E(\varepsilon_i\varepsilon_i'FF'\varepsilon_i\varepsilon_i')F-\sigma_0^2F'FF'E(\varepsilon_i\varepsilon_i')F-2\sigma_0^2F'E(\varepsilon_i\varepsilon_i')FF'F+\sigma_0^4F'FF'F] \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}\text{tr}[F'E(\varepsilon_i\varepsilon_i'FF'\varepsilon_i\varepsilon_i')F-\sigma_0^4F'FF'F]
\end{aligned} \tag{A68}$$

Here,

$$
\begin{aligned}
&\mathrm{tr}[F'E(\varepsilon_i\varepsilon_i'FF'\varepsilon_i\varepsilon_i')F] \\
&= E(\mathrm{tr}[F'\varepsilon_i\varepsilon_i'FF'\varepsilon_i\varepsilon_i'F]) = E[(\varepsilon_i'FF'\varepsilon_i)^2] = E\left[\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{i,t}\varepsilon_{i,s}F_t'F_s\right)^2\right] \\
&= \sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{m=1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{m=1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,t}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_tF_m'F_n + \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&+ \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \kappa_0\sum_{t=1}^{T}F_t'F_tF_t'F_t + 4\sigma_0^4\sum_{t=1}^{T}\sum_{s=1}^{t-1}F_t'F_sF_s'F_t + 2\sigma_0^4\sum_{s=n+1}^{T}\sum_{n=1}^{s-1}F_n'F_sF_s'F_n,
\end{aligned}
$$

as follows from nothing that

$$
\begin{aligned}
&\sum_{t=1}^{T}\sum_{m=1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,t}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_tF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{m=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,t}\varepsilon_{i,m}\varepsilon_{i,m})F_t'F_tF_m'F_m + \sum_{t=1}^{T}\sum_{m=1}^{T}\sum_{n=1}^{m-1}E(\varepsilon_{i,t}\varepsilon_{i,t}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_tF_m'F_n \\
&+ \sum_{t=1}^{T}\sum_{m=1}^{T}\sum_{n=m+1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,t}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_tF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{m=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,t}\varepsilon_{i,m}\varepsilon_{i,m})F_t'F_tF_m'F_m \\
&= \sum_{t=1}^{T}E(\varepsilon_{i,t}^4)F_t'F_tF_t'F_t + 2\sum_{t=1}^{T}\sum_{m=1}^{t-1}E(\varepsilon_{i,t}^2)E(\varepsilon_{i,m}^2)F_t'F_tF_m'F_m \\
&= \kappa_0\sum_{t=1}^{T}F_t'F_tF_t'F_t + 2\sigma_0^4\sum_{t=1}^{T}\sum_{m=1}^{t-1}F_t'F_tF_m'F_m,
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=1}^{T}\sum_{n=1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{n=1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,n})F_t'F_sF_s'F_n + \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=1}^{s-1}\sum_{n=1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&+ \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=s+1}^{T}\sum_{n=1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{n=1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,n})F_t'F_sF_s'F_n + \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=s+1}^{T}\sum_{n=1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{s=1}^{t-1} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,t})F_t'F_sF_s'F_t \\
&+ \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=s+1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,s})F_t'F_sF_m'F_s + \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=s+1}^{T}\sum_{n=1}^{s-1} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&+ \sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{m=s+1}^{T}\sum_{n=s+1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= 2\sum_{t=1}^{T}\sum_{s=1}^{t-1} E(\varepsilon_{i,t}^2)E(\varepsilon_{i,s}^2)F_t'F_sF_s'F_t = 2\sigma_0^4\sum_{t=1}^{T}\sum_{s=1}^{t-1} F_t'F_sF_s'F_t,
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,n})F_t'F_sF_s'F_n+\sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=1}^{s-1}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&+ \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=s+1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,n})F_t'F_sF_s'F_n+\sum_{s=t+1}^{T}\sum_{m=1}^{s-1}E(\varepsilon_{i,m}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,s})F_m'F_sF_m'F_s \\
&+ \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=s+1}^{T}\sum_{n=1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= \sum_{t=1}^{T}\sum_{s=t+1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,s})F_t'F_sF_s'F_s+\sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{n=1}^{s-1}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,n})F_t'F_sF_s'F_n \\
&+ \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{n=s+1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,s}\varepsilon_{i,n})F_t'F_sF_s'F_n+\sum_{s=t+1}^{T}\sum_{m=1}^{s-1}E(\varepsilon_{i,m}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,s})F_m'F_sF_m'F_s \\
&+ \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=s+1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,m})F_t'F_sF_m'F_m+\sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=s+1}^{T}\sum_{n=1}^{m-1}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&+ \sum_{t=1}^{T}\sum_{s=t+1}^{T}\sum_{m=s+1}^{T}\sum_{n=m+1}^{T}E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,m}\varepsilon_{i,n})F_t'F_sF_m'F_n \\
&= 2\sum_{s=n+1}^{T}\sum_{n=1}^{s-1}E(\varepsilon_{i,n}^2)E(\varepsilon_{i,s}^2)F_n'F_sF_s'F_n = 2\sigma_0^4\sum_{s=n+1}^{T}\sum_{n=1}^{s-1}F_n'F_sF_s'F_n.
\end{aligned}
$$

Hence, since

$$
\mathrm{tr}(F'FF'F)=\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{tr}(F_tF_t'F_sF_s')=\sum_{t=1}^{T}\sum_{s=1}^{T}F_s'F_tF_t'F_s=\sum_{t=1}^{T}F_t'F_tF_t'F_t+2\sum_{t=1}^{T}\sum_{s=1}^{t-1}F_s'F_tF_t'F_s,
$$

we can show that

$$
\begin{aligned}
&E\left(\left|\left|\frac{1}{\sqrt{N}T}\sum_{i=1}^{N}F'(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)F\right|\right|^2\right) \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}\mathrm{tr}[F'E(\varepsilon_i\varepsilon_i'FF'\varepsilon_i\varepsilon_i')F-\sigma_0^4F'FF'F] \\
&= (\kappa_0-\sigma_0^4)\sum_{t=1}^{T}F_t'F_tF_t'F_t+2\sigma_0^4\left(\sum_{t=1}^{T}\sum_{s=1}^{t-1}F_t'F_sF_s'F_t+\sum_{s=n+1}^{T}\sum_{n=1}^{s-1}F_n'F_sF_s'F_n\right)\leq C,
\end{aligned}
$$

implying

$$
\left|\left|\frac{1}{NT^2}\sum_{i=1}^{N}F'(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)F\right|\right|=O_p(N^{-1/2}T^{-1}). \tag{A69}
$$

Multiplication by L_0 does not affect this result. The other terms in (A67) are therefore of the same order. Therefore,

$$\left|\left|T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}F\right|\right| = O_p(N^{-1/2}T^{-1}). \tag{A70}$$

Hence, by adding the results, and using $O_p(N^{-1/2}T^{-1}) < O_p((NT)^{-1/2})$,

$$\begin{aligned}
&T^{-2}F'GF \\
&= T^{-2}F'\Gamma^{-1}\Gamma_0 FS_\lambda F'\Gamma_0'\Gamma^{-1\prime}F + O(T^{-1}) + O_p((NT)^{-1/2}) \\
&= (T^{-1}F'F)S_\lambda(T^{-1}F'F) + (\rho_0-\rho)(T^{-1}F'F)S_\lambda T^{-1}F'L_0'F + (\rho_0-\rho)T^{-1}F'L_0FS_\lambda(T^{-1}F'F) \\
&+ (\rho_0-\rho)^2T^{-1}F'L_0FS_\lambda T^{-1}F'L_0'F + O(T^{-1}) + O_p((NT)^{-1/2}),
\end{aligned} \tag{A71}$$

which in turn implies

$$\begin{aligned}
&T^{-1}\text{tr}\,(GFKF') \\
&= \text{tr}\,[T^{-2}F'GF(T^{-1}F'F)^{-1}] - \sigma^2T^{-1}m \\
&= \text{tr}\,[T^{-2}F'\Gamma^{-1}\Gamma_0FS_\lambda F'\Gamma_0'\Gamma^{-1\prime}F(T^{-1}F'F)^{-1}] + O(T^{-1}) + O_p((NT)^{-1/2}) \\
&= \text{tr}\,[(T^{-1}F'F)S_\lambda(T^{-1}F'F)(T^{-1}F'F)^{-1}] + (\rho_0-\rho)\text{tr}\,[(T^{-1}F'F)S_\lambda T^{-1}F'L_0'F(T^{-1}F'F)^{-1}] \\
&+ (\rho_0-\rho)\text{tr}\,[T^{-1}F'L_0FS_\lambda(T^{-1}F'F)(T^{-1}F'F)^{-1}] \\
&+ (\rho_0-\rho)^2\text{tr}\,[T^{-1}F'L_0FS_\lambda T^{-1}F'L_0'F(T^{-1}F'F)^{-1}] + O(T^{-1}) + O_p((NT)^{-1/2}) \\
&= \text{tr}\,[(T^{-1}F'F)S_\lambda] + 2(\rho_0-\rho)\text{tr}\,(T^{-1}F'L_0FS_\lambda) \\
&+ (\rho_0-\rho)^2\text{tr}\,[T^{-1}F'L_0FS_\lambda T^{-1}F'L_0'F(T^{-1}F'F)^{-1}] + O(T^{-1}) + O_p((NT)^{-1/2}).
\end{aligned} \tag{A72}$$

Next, consider $\text{tr}\,G$, the third term in Q_c. By using the above results regarding the order of the cross-sectional sums in $\varepsilon_i\lambda_i'$ and $(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)$, we can show that

$$\begin{aligned}
&T^{-1}\text{tr}\,G \\
&= T^{-1}\text{tr}\,(\Gamma^{-1}S_y\Gamma^{-1\prime}) \\
&= T^{-1}\text{tr}\,[\Gamma^{-1}\Gamma_0(\sigma_0^2I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}] + T^{-1}\text{tr}\left(\Gamma^{-1}\Gamma_0\frac{1}{N}\sum_{i=1}^{N}F\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}\right) \\
&+ T^{-1}\text{tr}\left(\Gamma^{-1}\Gamma_0\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i\lambda_i'F'\Gamma_0'\Gamma^{-1\prime}\right) \\
&+ T^{-1}\text{tr}\left(\Gamma^{-1}\Gamma_0\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2I_T)\Gamma_0'\Gamma^{-1\prime}\right) \\
&= T^{-1}\text{tr}\,[\Gamma^{-1}\Gamma_0(\sigma_0^2I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}] + O_p((NT)^{-1/2}).
\end{aligned} \tag{A73}$$

For the remaining term, via $T^{-1}\text{tr}\, I_T = 1$, $\text{tr}\, L_0 = \text{tr}\, L_0' = 0$,

$$\begin{aligned}
&T^{-1}\text{tr}\,[\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}] \\
&= T^{-1}\text{tr}\,([I_T + (\rho_0-\rho)L_0](\sigma_0^2 I_T + FS_\lambda F')[I_T + (\rho_0-\rho)L_0]') \\
&= T^{-1}\text{tr}\,(\sigma_0^2 I_T + FS_\lambda F') + (\rho_0-\rho)T^{-1}\text{tr}\,[(\sigma_0^2 I_T + FS_\lambda F')L_0'] \\
&+ (\rho_0-\rho)T^{-1}\text{tr}\,[L_0(\sigma_0^2 I_T + FS_\lambda F')] + (\rho_0-\rho)^2 T^{-1}\text{tr}\,[L_0(\sigma_0^2 I_T + FS_\lambda F')L_0'] \\
&= \sigma_0^2[1 + (\rho_0-\rho)^2 T^{-1}\text{tr}\,(L_0 L_0')] + T^{-1}\text{tr}\,(FS_\lambda F') + (\rho_0-\rho)T^{-1}\text{tr}\,(FS_\lambda F' L_0') \\
&+ (\rho_0-\rho)T^{-1}\text{tr}\,(L_0 FS_\lambda F') + (\rho_0-\rho)^2 T^{-1}\text{tr}\,(L_0 FS_\lambda F' L_0') \\
&= \sigma_0^2[1 + (\rho_0-\rho)^2 \text{tr}\,(T^{-1}L_0 L_0')] + \text{tr}\,(T^{-1}F'FS_\lambda) + 2(\rho_0-\rho)\text{tr}\,(T^{-1}F'L_0 FS_\lambda) \\
&+ (\rho_0-\rho)^2 \text{tr}\,(T^{-1}F'L_0' L_0 FS_\lambda),
\end{aligned}$$

giving

$$\begin{aligned}
T^{-1}\text{tr}\, G &= \sigma_0^2[1 + (\rho_0-\rho)^2 \text{tr}\,(T^{-1}L_0 L_0')] + \text{tr}\,(T^{-1}F'FS_\lambda) + 2(\rho_0-\rho)\text{tr}\,(T^{-1}F'L_0 FS_\lambda) \\
&+ (\rho_0-\rho)^2 \text{tr}\,(T^{-1}F'L_0' L_0 FS_\lambda) + O((NT)^{-1/2}). \qquad \text{(A74)}
\end{aligned}$$

The order of the second term in Q_c is given by

$$T^{-1}\log(|\hat{\Lambda}|) = T^{-1}\log(|I_T + \sigma^{-2}F\hat{S}_\lambda F'|) = O_p(T^{-1}\log(T)), \qquad \text{(A75)}$$

as is clear from noting that

$$\begin{aligned}
T^{-1}||F\hat{S}_\lambda F'|| &= T^{-1}||F[\sigma^2 F^-(\sigma^{-2}G - I_T)F^{-\prime}]F|| \\
&= \sigma^2 T^{-1}||F(F'F)^{-1}F'(\sigma^{-2}G - I_T)F(F'F)^{-1}F'|| \\
&\leq T^{-1}||F(F'F)^{-1}F'GF(F'F)^{-1}F'|| + \sigma^2 T^{-1}||F(F'F)^{-1}F'|| \\
&= ||T^{-2}F'GF(T^{-1}F'F)^{-1}|| + \sigma^2 T^{-1}||I_m|| = O_p(1).
\end{aligned}$$

Hence, by putting everything together, with $O(T^{-1}) < O_p(T^{-1}\log(T))$,

$$
\begin{aligned}
&T^{-1}Q_c \\
&= \log(\sigma^2) + T^{-1}\log(|\hat{\Lambda}|) + \sigma^{-2}T^{-1}\text{tr}\, G - \sigma^{-2}T^{-1}\text{tr}\,(GFKF') \\
&= \log(\sigma^2) + \sigma^{-2}T^{-1}\text{tr}\, G - \sigma^{-2}T^{-1}\text{tr}\,(GFKF') + O_p(T^{-1}\log(T)) \\
&= \log(\sigma^2) + \sigma^{-2}\sigma_0^2[1 + (\rho_0-\rho)^2\text{tr}\,(T^{-1}L_0L_0')] + \sigma^{-2}\text{tr}\,(T^{-1}F'FS_\lambda) \\
&+ 2\sigma^{-2}(\rho_0-\rho)\text{tr}\,(T^{-1}F'L_0FS_\lambda) + \sigma^{-2}(\rho_0-\rho)^2\text{tr}\,(T^{-1}F'L_0'L_0FS_\lambda) \\
&- \sigma^{-2}\text{tr}\,[(T^{-1}F'F)S_\lambda] - 2\sigma^{-2}(\rho_0-\rho)\text{tr}\,(T^{-1}F'L_0FS_\lambda) \\
&- \sigma^{-2}(\rho_0-\rho)^2\text{tr}\,[T^{-1}F'L_0FS_\lambda T^{-1}F'L_0'F(T^{-1}F'F)^{-1}] + O_p((NT)^{-1/2}) + O_p(T^{-1}\log(T)) \\
&= \log(\sigma^2) + \sigma^{-2}\sigma_0^2 + \sigma^{-2}(\rho_0-\rho)^2[\sigma_0^2\text{tr}\,(T^{-1}L_0L_0') + \text{tr}\,(T^{-1}F'L_0'L_0FS_\lambda) \\
&- \text{tr}\,(T^{-1}F'L_0FS_\lambda T^{-1}F'L_0'F(T^{-1}F'F)^{-1})] + O_p((NT)^{-1/2}) + O_p(T^{-1}\log(T)) \\
&= \log(\sigma^2) + \sigma^{-2}\sigma_0^2 + \sigma^{-2}\sigma_0^2(\rho_0-\rho)^2\omega_1^2 + O_p((NT)^{-1/2}) + O_p(T^{-1}\log(T)), \qquad \text{(A76)}
\end{aligned}
$$

where $\omega_1^2 = T^{-1}\text{tr}\,(L_0L_0' + \sigma_0^{-2}S_\lambda F'L_0'M_FL_0F)$, $M_F = I_T - P_F$ and $P_F = F(F'F)^{-1}F'$. If A and B are positive semidefinite, then $0 \leq \text{tr}\,(AB) \leq (\text{tr}\,A)(\text{tr}\,B)$. Since M_F is idempotent, $T^{-1}F'L_0'M_FL_0F$ positive semidefinite. Hence, because S_λ is positive definite too, we have that $\text{tr}\,(S_\lambda T^{-1}F'L_0'M_FL_0F) \geq 0$. By using this and

$$
\begin{aligned}
T^{-1}\text{tr}\,(L_0L_0') &= \frac{1}{T}\sum_{t=1}^{T}\text{tr}\,(l_{t,0}l_{t,0}') = \frac{1}{T}\sum_{t=1}^{T}\sum_{n=0}^{T-t}\rho_0^{2n} = \frac{1}{T(1-\rho_0^2)}\sum_{t=1}^{T}(1-\rho_0^{2(T+1-t)}) \\
&= \frac{1}{(1-\rho_0^2)} + O(T^{-1}), \qquad \text{(A77)}
\end{aligned}
$$

from which it follows that $\text{tr}\,(T^{-1}L_0L_0') > 0$, we obtain $\omega_1^2 \geq 0$. Hence,

$$
\begin{aligned}
(NT)^{-1}\ell_c &= -\frac{1}{2T}Q_c \\
&= -\frac{1}{2}\left(\log(\sigma^2) + \frac{\sigma_0^2}{\sigma^2}\right) - \frac{\sigma_0^2}{2\sigma^2}(\rho_0-\rho)^2\omega_1^2 + O_p((NT)^{-1/2}) \\
&+ O_p(T^{-1}\log(T)), \qquad \text{(A78)}
\end{aligned}
$$

as required for the proof. ■

Lemma C.1. *Under C1, $|\rho_0| < 1$, and Assumptions EPS, F and LAM, as $T \to \infty$ for any N, including $N \to \infty$, provided that $\sqrt{N}T^{-3/2} \to 0$,*

$$
H_{1/2}^{-1}\frac{\partial \ell_c(\theta_2^0)}{\partial \theta_2} \sim N\left(0_{2\times 1}, \begin{bmatrix} \omega_1^2 & 0 \\ 0 & \frac{(\kappa_0-1)}{4\sigma_0^4} \end{bmatrix}\right),
$$

where $H_p = \text{diag}(\sqrt{N}T^p, \sqrt{NT})$, $\omega_1^2 = T^{-1}\text{tr}\,(L_0L_0' + \sigma_0^{-2}S_\lambda F'L_0'M_FL_0F)$, $M_F = I_T - P_F$ *and* $P_F = F(F'F)^{-1}F'$.

Proof of Lemma C.1.

We have

$$\frac{2\sigma_0^2}{N}\frac{\partial \ell_c(\theta_2^0)}{\partial \rho} = [\text{vec}\, B(\theta_2^0)]'C(\rho_0), \tag{A79}$$

where $G(\rho_0) = \Gamma_0^{-1}S_y\Gamma_0^{-1\prime}$ and $\hat{\Lambda}(\theta_2^0) = I_T + \sigma_0^{-2}F\hat{S}_\lambda(\theta_2^0)F'$. Consider $C(\rho_0)$. From $S_y = \Gamma_0 S_u \Gamma_0'$, we have $\text{vec}\, S_y = (\Gamma_0 \otimes \Gamma_0)\text{vec}\, S_u$. Note that $J\Gamma_0 = L_0$. Hence, since $(A \otimes B)(C \otimes D) = AC \otimes BD$ and $(C' \otimes A)\text{vec}\, B = \text{vec}(ABC)$,

$$\begin{aligned} C(\rho_0) &= (\Gamma_0^{-1} \otimes J + J \otimes \Gamma_0^{-1})\text{vec}\, S_y = (\Gamma_0^{-1} \otimes J + J \otimes \Gamma_0^{-1})(\Gamma_0 \otimes \Gamma_0)\text{vec}\, S_u \\ &= (I_T \otimes L_0 + L_0 \otimes I_T)\text{vec}\, S_u = \text{vec}\,(L_0S_u + S_uL_0'). \end{aligned} \tag{A80}$$

By using this and $\text{tr}\,(A'B) = (\text{vec}\, A)'\text{vec}\, B$, we obtain

$$\frac{2\sigma_0^2}{\sqrt{NT}}\frac{\partial \ell_c(\theta_2^0)}{\partial \rho} = \sqrt{N}T^{-1/2}\text{tr}\,[B(\theta_2^0)'(L_0S_u + S_uL_0')]. \tag{A81}$$

Consider $B(\theta_2^0)$. Let $K_0 = K(\theta_2^0) = (\sigma_0^2\hat{S}_\lambda(\theta_2^0)^{-1} + F'F)^{-1} = (F'F)^{-1} - \sigma_0^2T^{-2}(T^{-2}F'G(\rho_0)F)^{-1}$, such that $\hat{\Lambda}(\theta_2^0)^{-1} = I_T - FK_0F'$ (see Proof of Lemma 1). This implies

$$\begin{aligned} F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^- &= F(F'F)^{-1}F'(I_T - FK_0F')F(F'F)^{-1}F' = F((F'F)^{-1} - K_0)F' \\ &= \sigma_0^2T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F', \end{aligned} \tag{A82}$$

and

$$\begin{aligned} &F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^- \\ &= F(F'F)^{-1}F'(I_T - FK_0F')G(\rho_0)(I_T - FK_0F')F(F'F)^{-1}F' \\ &= F(F'F)^{-1}F'(G(\rho_0) - G(\rho_0)FK_0F' - FK_0F'G(\rho_0) + FK_0F'G(\rho_0)FK_0F')F(F'F)^{-1}F' \\ &= F((F'F)^{-1} - K_0)F'G(\rho_0)F((F'F)^{-1} - K_0)F' \\ &= \sigma_0^4T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F', \end{aligned} \tag{A83}$$

suggesting that $B(\theta_2^0)$ simplifies to

$$\begin{aligned} B(\theta_2^0) &= F^{-\prime}F'(\hat{\Lambda}(\theta_2^0)^{-1} - \sigma_0^{-2}\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1})FF^- + \hat{\Lambda}(\theta_2^0)^{-1} \\ &= \sigma_0^2T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F' - \sigma_0^{-2}\sigma_0^4T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F' + \hat{\Lambda}(\theta_2^0)^{-1} \\ &= \hat{\Lambda}(\theta_2^0)^{-1}. \end{aligned} \tag{A84}$$

Therefore,

$$\begin{aligned}\frac{2\sigma_0^2}{\sqrt{NT}}\frac{\partial \ell_c(\theta_2^0)}{\partial \rho} &= \sqrt{N}T^{-1/2}\text{tr}\left[B(\theta_2^0)'(L_0S_u + S_uL_0')\right] \\ &= \sqrt{N}T^{-1/2}\text{tr}\left[\hat{\Lambda}(\theta_2^0)^{-1}(L_0S_u + S_uL_0')\right] = Q. \end{aligned} \tag{A85}$$

Note that $\text{tr}\,(S_uL_0') = \text{tr}\,(L_0S_u)$, as is clear from using $\text{tr}\,A = \text{tr}\,A'$ and the symmetry S_u. But K_0 is symmetric too, and therefore $\text{tr}\,(K_0F'S_uL_0'F) = \text{tr}\,(F'L_0S_uFK_0) = \text{tr}\,(K_0F'L_0S_uF)$, which in turn implies that $\text{tr}\left[FK_0F'(L_0S_u + S_uL_0')\right] = \text{tr}\left[K_0F'(L_0S_u + S_uL_0')F\right] = 2\text{tr}\,(K_0F'L_0S_uF)$. By using this and $\hat{\Lambda}(\theta_2^0)^{-1} = I_T - FK_0F'$, we obtain

$$\begin{aligned} Q &= \sqrt{N}T^{-1/2}\text{tr}\left[\hat{\Lambda}(\theta_2^0)^{-1}(L_0S_u + S_uL_0')\right] \\ &= \sqrt{N}T^{-1/2}\text{tr}\left[(I_T - FK_0F')(L_0S_u + S_uL_0')\right] \\ &= \sqrt{N}T^{-1/2}\text{tr}\left[L_0S_u + S_uL_0' - FK_0F'(L_0S_u + S_uL_0')\right] \\ &= 2\sqrt{N}T^{-1/2}[\text{tr}\,(L_0S_u) - \text{tr}\,(K_0F'L_0S_uF)] = 2(Q_1 - Q_2), \end{aligned} \tag{A86}$$

with implicit definitions of Q_1 and Q_2.

Consider Q_1, which, via $\text{tr}\,(A+B) = \text{tr}\,A + \text{tr}\,B$ and $\text{tr}\,(AB) = \text{tr}\,(BA)$, can be expanded in the following fashion:

$$\begin{aligned} Q_1 &= \sqrt{N}T^{-1/2}\text{tr}\,(L_0S_u) \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}[\text{tr}\,(L_0F\lambda_i\lambda_i'F') + \text{tr}\,(L_0F\lambda_i\varepsilon_i') + \text{tr}\,(L_0\varepsilon_i\lambda_i'F') + \text{tr}\,(L_0\varepsilon_i\varepsilon_i')] \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}[\text{tr}\,(\lambda_i'F'L_0F\lambda_i) + \text{tr}\,(\varepsilon_i'L_0F\lambda_i) + \text{tr}\,(\lambda_i'F'L_0\varepsilon_i) + \text{tr}\,(\varepsilon_i'L_0\varepsilon_i)] \\ &= Q_{11} + Q_{11} + Q_{12}, \end{aligned} \tag{A87}$$

where

$$\begin{aligned} Q_{11} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i'F'L_0F\lambda_i, \\ Q_{12} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}(\varepsilon_i'L_0F\lambda_i + \lambda_i'F'L_0\varepsilon_i), \\ Q_{13} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\varepsilon_i'L_0\varepsilon_i. \end{aligned}$$

From $\lambda_i'F'L_0F\lambda_i = \text{tr}\,(\lambda_i'F'L_0F\lambda_i) = \text{tr}\,(\lambda_i\lambda_i'F'L_0F)$, Q_{11} can be written as

$$\begin{aligned} Q_{11} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i'F'L_0F\lambda_i = \frac{\sqrt{N}}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\text{tr}\,(\lambda_i\lambda_i'F'L_0F) \\ &= \sqrt{N}T^{-1/2}\text{tr}\,(S_\lambda F'L_0F). \end{aligned} \tag{A88}$$

Consider Q_{12}, where $T^{-1/2}\varepsilon_i' L_0 F$ and $T^{-1/2}F' L_0 \varepsilon_i$ are clearly mean zero, and, by a central limit theorem (CLT), also normal. As for the variance of these normals, by using $E(\varepsilon_i \varepsilon_i') = \sigma_0^2 I_T$, we have

$$\begin{aligned} & E[(\varepsilon_i' L_0 F \lambda_i + \lambda_i' F' L_0 \varepsilon_i)^2] \\ &= E(\lambda_i' F' L_0' \varepsilon_i \varepsilon_i' L_0 F \lambda_i) + 2E(\lambda_i' F' L_0 \varepsilon_i \varepsilon_i' L_0 F \lambda_i) + E(\lambda_i' F' L_0 \varepsilon_i \varepsilon_i' L_0' F \lambda_i) \\ &= \sigma_0^2 \lambda_i' F' (L_0' L_0 + 2 L_0 L_0 + L_0 L_0') F \lambda_i, \end{aligned}$$

suggesting that

$$\begin{aligned} E(Q_{12}^2) &= E\left[\left(\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} (\varepsilon_i' L_0 F \lambda_i + \lambda_i' F' L_0 \varepsilon_i)\right)^2\right] \\ &= \frac{1}{NT} \sum_{i=1}^{N} \sum_{j=1}^{N} E[(\varepsilon_i' L_0 F \lambda_i + \lambda_i' F' L_0 \varepsilon_i)(\varepsilon_j' L_0 F \lambda_j + \lambda_j' F' L_0 \varepsilon_j)] \\ &= \frac{1}{NT} \sum_{i=1}^{N} E[(\varepsilon_i' L_0 F \lambda_i + \lambda_i' F' L_0 \varepsilon_i)^2] \\ &= \frac{1}{NT} \sum_{i=1}^{N} \lambda_i' F' E(L_0' \varepsilon_i \varepsilon_i' L_0 + 2 L_0 \varepsilon_i \varepsilon_i' L_0 + L_0 \varepsilon_i \varepsilon_i' L_0') F \lambda_i \\ &= \sigma_0^2 \frac{1}{NT} \sum_{i=1}^{N} \lambda_i' F' (L_0' L_0 + 2 L_0 L_0 + L_0 L_0') F \lambda_i \\ &= \sigma_0^2 \mathrm{tr}\left[S_\lambda T^{-1} F' (L_0' L_0 + 2 L_0 L_0 + L_0 L_0') F\right] = \Sigma_{12} \end{aligned}$$

Hence, using $\sim$ to signify asymptotic equivalence,

$$Q_{12} \sim N(0, \Sigma_{12}), \tag{A89}$$

which holds for any N, including $N \to \infty$, provided that $T \to \infty$.

When evaluating Q_{13} it is useful to write $\varepsilon_i' L_0 \varepsilon_i = \sum_{t=2}^{T} \varepsilon_{i,t} \varepsilon_{i,t-1}^*$, where $\varepsilon_{i,t}^*$ is as in Appendix A. Clearly,

$$\frac{1}{T} \sum_{t=2}^{T} E[(\varepsilon_{i,t-1}^*)^2] = T^{-1} E[\mathrm{tr}\,(\varepsilon_i' L_0' L_0 \varepsilon_i)] = T^{-1} \mathrm{tr}\,[E(\varepsilon_i \varepsilon_i') L_0' L_0] = \sigma_0^2 T^{-1} \mathrm{tr}\,(L_0 L_0'),$$

from which we obtain

$$\begin{aligned} T^{-1} E[(\varepsilon_i' L_0 \varepsilon_i)^2] &= \frac{1}{T} \sum_{t=2}^{T} \sum_{s=2}^{T} E(\varepsilon_{i,t} \varepsilon_{i,s} \varepsilon_{i,t-1}^* \varepsilon_{i,s-1}^*) \\ &= \frac{1}{T} \sum_{t=2}^{T} E(\varepsilon_{i,t}^2) E[(\varepsilon_{i,t-1}^*)^2] + \frac{2}{T} \sum_{t=3}^{T} \sum_{s=2}^{t-1} E(\varepsilon_{i,t}) E(\varepsilon_{i,s} \varepsilon_{i,t-1}^* \varepsilon_{i,s-1}^*) \\ &= \sigma_0^2 \frac{1}{T} \sum_{t=2}^{T} E[(\varepsilon_{i,t-1}^*)^2] = \sigma_0^4 T^{-1} \mathrm{tr}\,(L_0 L_0') = \Sigma_{13}. \end{aligned}$$

Hence, by a CLT for martingale difference sequences,

$$T^{-1/2}\varepsilon_i' L_0 \varepsilon_i \sim N(0, \Sigma_{13}) \tag{A90}$$

as $T \to \infty$. This means that

$$Q_{13} = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \varepsilon_i' L_0 \varepsilon_i \sim N(0, \Sigma_{13}) \tag{A91}$$

as $T \to \infty$ for any N, including $N \to \infty$.

Let us now consider the covariance between Q_{12} and Q_{13}. Note first that if $k \geq t$,

$$\begin{aligned}
&E(\varepsilon_{i,t}\varepsilon^*_{i,t-1}\varepsilon^*_{i,k-1}) \\
&= \sum_{s=1}^{t-1}\sum_{n=1}^{k-1} \rho_0^{t+k-(2+n+s)} E(\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{i,n}) \\
&= \sum_{s=1}^{t-1} \rho_0^{t+k-(2+t+s)} E(\varepsilon_{i,t}^2)E(\varepsilon_{i,s}) + \sum_{s=1}^{t-1}\sum_{n=t+1}^{k-1} \rho_0^{t+k-(2+n+s)} E(\varepsilon_{i,t})E(\varepsilon_{i,s})E(\varepsilon_{i,n}) \\
&+ \sum_{s=1}^{t-1}\sum_{n=1}^{t-1} \rho_0^{t+k-(2+n+s)} E(\varepsilon_{i,t})E(\varepsilon_{i,s}\varepsilon_{i,n}) = 0,
\end{aligned}$$

whereas if $k < t$, then

$$E(\varepsilon_{i,t}\varepsilon^*_{i,t-1}\varepsilon^*_{i,k-1}) = \sum_{s=1}^{t-1}\sum_{n=1}^{k-1} \rho_0^{t+k-(2+n+s)} E(\varepsilon_{i,t})E(\varepsilon_{i,s}\varepsilon_{i,n}) = 0,$$

suggesting

$$\begin{aligned}
E(F'L_0\varepsilon_i\varepsilon_i'L_0\varepsilon_i) &= \sum_{k=2}^{T}\sum_{t=2}^{T} F_k E(\varepsilon_{i,t}\varepsilon^*_{i,t-1}\varepsilon^*_{i,k-1}) \\
&= \sum_{t=3}^{T}\sum_{k=2}^{t-1} F_k E(\varepsilon_{i,t})E(\varepsilon^*_{i,t-1}\varepsilon^*_{i,k-1}) + \sum_{t=2}^{T}\sum_{k=t}^{T} F_k E(\varepsilon_{i,t}\varepsilon^*_{i,t-1}\varepsilon^*_{i,k-1}) = 0.
\end{aligned}$$

Moreover,

$$\begin{aligned}
E(\varepsilon_i'L_0\varepsilon_i\varepsilon_i'L_0F) &= \sum_{k=2}^{T}\sum_{t=2}^{T} E(\varepsilon_{i,t}\varepsilon^*_{i,t-1}\varepsilon_{i,k})F^{*\prime}_{k-1} \\
&= \sum_{t=2}^{T} E(\varepsilon_{i,t}^2)E(\varepsilon^*_{i,t-1})F^{*\prime}_{t-1} + \sum_{t=3}^{T}\sum_{k=2}^{t-1} E(\varepsilon_{i,t})E(\varepsilon^*_{i,t-1}\varepsilon_{i,k})F^{*\prime}_{k-1} \\
&+ \sum_{t=2}^{T}\sum_{k=t+1}^{T} E(\varepsilon_{i,t}\varepsilon^*_{i,t-1})E(\varepsilon_{i,k})F^{*\prime}_{k-1} = 0.
\end{aligned}$$

Therefore, $E[\varepsilon_i'L_0\varepsilon_i(\varepsilon_i'L_0F\lambda_i + \lambda_i'F'L_0\varepsilon_i)] = 0$, suggesting that

$$E(Q_{12}Q_{13}) = \frac{1}{NT}\sum_{i=1}^{N} E[\varepsilon_i'L_0\varepsilon_i(\varepsilon_i'L_0F\lambda_i + \lambda_i'F'L_0\varepsilon_i)] = 0. \tag{A92}$$

Hence, putting everything together,

$$Q_1 - Q_{11} = Q_{12} + Q_{13} \sim N(0, \Sigma_{12} + \Sigma_{13}). \tag{A93}$$

Next, consider Q_2;

$$\begin{aligned} Q_2 &= \sqrt{N}T^{-1/2}\text{tr}\,(K_0F'L_0S_uF) \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}[\text{tr}\,(K_0F'L_0F\lambda_i\lambda_i'F'F) + \text{tr}\,(K_0F'L_0F\lambda_i\varepsilon_i'F) + \text{tr}\,(K_0F'L_0\varepsilon_i\lambda_i'F'F) \\ &+ \text{tr}\,(K_0F'L_0\varepsilon_i\varepsilon_i'F)] \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}[\text{tr}\,(\lambda_i'F'FK_0F'L_0F\lambda_i) + \text{tr}\,(\varepsilon_i'FK_0F'L_0F\lambda_i) + \text{tr}\,(\lambda_i'F'FK_0F'L_0\varepsilon_i) \\ &+ \text{tr}\,(K_0F'L_0\varepsilon_i\varepsilon_i'F)] \\ &= Q_{21} + Q_{22} + Q_{23}, \end{aligned} \tag{A94}$$

with

$$\begin{aligned} Q_{21} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i'F'FK_0F'L_0F\lambda_i, \\ Q_{22} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}(\varepsilon_i'FK_0F'L_0F\lambda_i + \lambda_i'F'FK_0F'L_0\varepsilon_i), \\ Q_{23} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\text{tr}\,(K_0F'L_0\varepsilon_i\varepsilon_i'F). \end{aligned}$$

From Proof of Lemma 1, $T^{-2}F'G(\rho_0)F = (T^{-1}F'F)S_\lambda(T^{-1}F'F) + O(T^{-1}) + O_p((NT)^{-1/2})$, and therefore

$$\begin{aligned} TK_0 &= (T^{-1}F'F)^{-1} - \sigma_0^2T^{-1}(T^{-2}F'G(\rho_0)F)^{-1} \\ &= (T^{-1}F'F)^{-1} - \sigma_0^2T^{-1}(T^{-1}F'F)^{-1}S_\lambda^{-1}(T^{-1}F'F)^{-1} + O(T^{-2}) \\ &+ O_p(N^{-1/2}T^{-3/2}). \end{aligned} \tag{A95}$$

Here, $||(T^{-1}F'F)^{-1}S_\lambda^{-1}(T^{-1}F'F)^{-1}|| = O_p(1)$, which means that (A95) may be written as

$$TK_0 = (T^{-1}F'F)^{-1} + O(T^{-1}). \tag{A96}$$

Substitution of (A95) into the expression for Q_{21} yields

$$\begin{aligned} Q_{21} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i' T^{-1}F'F(TK_0)F'L_0F\lambda_i \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i' F'L_0F\lambda_i \\ &- \sigma_0^2\frac{\sqrt{N}}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\lambda_i' S_\lambda^{-1}(T^{-1}F'F)^{-1}T^{-1}F'L_0F\lambda_i + O(\sqrt{N}T^{-5/2}) + O_p(T^{-2}) \\ &= Q_{11} - \sqrt{N}T^{-1/2}\sigma_0^2\mathrm{tr}\,(P_FL_0) + O(\sqrt{N}T^{-5/2}) + O_p(T^{-2}), \end{aligned} \tag{A97}$$

where the last equality holds, because

$$\begin{aligned} \sigma_0^2\frac{1}{N}\sum_{i=1}^{N}\lambda_i' S_\lambda^{-1}(T^{-1}F'F)^{-1}T^{-1}F'L_0F\lambda_i &= \sigma_0^2\frac{1}{N}\sum_{i=1}^{N}\mathrm{tr}\,[\lambda_i' S_\lambda^{-1}(T^{-1}F'F)^{-1}T^{-1}F'L_0F\lambda_i] \\ &= \sigma_0^2\mathrm{tr}\,[(T^{-1}F'F)^{-1}T^{-1}F'L_0F] \\ &= \sigma_0^2\mathrm{tr}\,(P_FL_0), \end{aligned}$$

with $P_F = F(F'F)^{-1}F'$.

Consider Q_{22}. This term is mean zero and, by a CLT, also asymptotically normal. As for the variance, via $TK_0 = (T^{-1}F'F)^{-1} + O(T^{-1})$,

$$\begin{aligned} &T^{-1}E[(\varepsilon_i'FK_0F'L_0F\lambda_i + \lambda_i'F'FK_0F'L_0\varepsilon_i)^2] \\ &= T^{-1}\lambda_iF'L_0'FK_0F'E(\varepsilon_i\varepsilon_i')FK_0F'L_0F\lambda_i + 2T^{-1}\lambda_i'F'FK_0F'L_0E(\varepsilon_i\varepsilon_i')FK_0F'L_0F\lambda_i \\ &+ T^{-1}\lambda_i'F'FK_0F'L_0E(\varepsilon_i\varepsilon_i')L_0'FK_0F'F\lambda_i \\ &= \sigma_0^2\lambda_iT^{-1}F'L_0'F(TK_0)T^{-1}F'F(TK_0)T^{-1}F'L_0F\lambda_i \\ &+ 2\sigma_0^2\lambda_i'T^{-1}F'F(TK_0)T^{-1}F'L_0F(TK_0)T^{-1}F'L_0F\lambda_i \\ &+ \sigma_0^2\lambda_i'T^{-1}F'F(TK_0)T^{-1}F'L_0L_0'F(TK_0)T^{-1}F'F\lambda_i \\ &= \sigma_0^2\lambda_i'T^{-1}F'L_0'F(T^{-1}F'F)^{-1}T^{-1}F'L_0F\lambda_i + 2\sigma_0^2\lambda_i'T^{-1}F'L_0F(T^{-1}F'F)^{-1}T^{-1}F'L_0F\lambda_i \\ &+ \sigma_0^2\lambda_i'T^{-1}F'L_0L_0'F\lambda_i + O(T^{-1}). \end{aligned}$$

Hence, letting

$$\begin{aligned} \Sigma_{22} &= \sigma_0^2\frac{1}{N}\sum_{i=1}^{N}E[\lambda_i'T^{-1}F'[L_0'F(F'F)^{-1}F'L_0 + 2L_0F(F'F)^{-1}F'L_0 + L_0L_0']F\lambda_i] \\ &= \sigma_0^2\frac{1}{N}\sum_{i=1}^{N}\mathrm{tr}\,(\lambda_i'T^{-1}F'[L_0'F(F'F)^{-1}F'L_0 + 2L_0F(F'F)^{-1}F'L_0 + L_0L_0']F\lambda_i) \\ &= \sigma_0^2\mathrm{tr}\,[S_\lambda T^{-1}F'(L_0'P_FL_0 + 2L_0P_FL_0 + L_0L_0')F], \end{aligned}$$

we can show that

$$Q_{22} = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} (\varepsilon_i' F K_0 F' L_0 F \lambda_i + \lambda_i' F' F K_0 F' L_0 \varepsilon_i) \sim N(0, \Sigma_{22}), \tag{A98}$$

which again only requires $T \to \infty$; N may be fixed but can also tend to infinity.

For Q_{23},

$$E(K_0 F' L_0 \varepsilon_i \varepsilon_i' F) = \sigma_0^2 (T K_0) T^{-1} F' L_0 F = \sigma_0^2 (T^{-1} F' F)^{-1} T^{-1} F' L_0 F + O(T^{-1}),$$

giving

$$\begin{aligned}
Q_{23} &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \text{tr}\,(K_0 F' L_0 \varepsilon_i \varepsilon_i' F) \\
&= \sigma_0^2 \sqrt{N} T^{-1/2} \text{tr}\,(K_0 F' L_0 F) + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \text{tr}\,[K_0 (F' L_0 \varepsilon_i \varepsilon_i' F - \sigma_0^2 F' L_0 F)] \\
&= \sqrt{N} T^{-1/2} \sigma_0^2 \text{tr}\,[(T^{-1} F' F)^{-1} T^{-1} F' L_0 F] \\
&+ T^{-1/2} \text{tr} \left(T K_0 \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (T^{-1} F' L_0 \varepsilon_i \varepsilon_i' F - \sigma_0^2 T^{-1} F' L_0 F) \right) + O_p(\sqrt{N} T^{-3/2}) \\
&= \sqrt{N} T^{-1/2} \sigma_0^2 \text{tr}\,(P_F L_0) + O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}).
\end{aligned} \tag{A99}$$

The results for Q_{21} and Q_{23} implies

$$\begin{aligned}
&Q_{21} + Q_{23} \\
&= Q_{11} - \sqrt{N} T^{-1/2} \sigma_0^2 \text{tr}\,(P_F L_0) + \sqrt{N} T^{-1/2} \sigma_0^2 \text{tr}\,(P_F L_0) + O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}) \\
&= Q_{11} + O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}).
\end{aligned} \tag{A100}$$

Hence, if we assume that N is fixed or $N \to \infty$ with $\sqrt{N} T^{-3/2} = o(1)$, provided that $T \to \infty$, then

$$\begin{aligned}
Q_2 &= Q_{21} + Q_{22} + Q_{23} = Q_{11} + Q_{22} + O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}) \\
&\sim Q_{11} + N(0, \Sigma_{22}).
\end{aligned} \tag{A101}$$

Q_1 and Q_2 are not uncorrelated. Note in particular how

$$E[(Q_1 - Q_{11})(Q_2 - Q_{11})] = E[(Q_{12} + Q_{13}) Q_{22}] + o(1) = E(Q_{12} Q_{22}) + o(1),$$

where the first equality requires $\sqrt{N} T^{-3/2} = o(1)$ for the remainder to be negligible. The last equality is due to the fact that $E(Q_{13} Q_{22}) = 0$. In order to see that this is so, write

$$\begin{aligned}
E(Q_{13} Q_{22}) &= \frac{1}{NT} \sum_{i=1}^{N} E[\varepsilon_i' L_0 \varepsilon_i (\varepsilon_i' F K_0 F' L_0 F \lambda_i + \lambda_i' F' F K_0 F' L_0 \varepsilon_i)] \\
&= \frac{1}{NT} \sum_{i=1}^{N} [E(\varepsilon_i' L_0 \varepsilon_i \varepsilon_i' F K_0 F' L_0 F \lambda_i) + E(\lambda_i' F' F K_0 F' L_0 \varepsilon_i \varepsilon_i' L_0 \varepsilon_i)],
\end{aligned}$$

where we know from before that $E(F'L_0\varepsilon_i\varepsilon_i'L_0\varepsilon_i) = 0$, suggesting that the second term is zero. We can similarly show that

$$\begin{aligned} E(\varepsilon_i'L_0\varepsilon_i\varepsilon_i'F) &= \sum_{k=2}^{T}\sum_{t=2}^{T} E(\varepsilon_{i,t}\varepsilon_{i,t-1}^{*}\varepsilon_{i,k})F_k' \\ &= \sum_{t=2}^{T} E(\varepsilon_{i,t}^2)E(\varepsilon_{i,t-1}^{*})F_t' + \sum_{t=3}^{T}\sum_{k=2}^{t-1} E(\varepsilon_{i,t})E(\varepsilon_{i,t-1}^{*}\varepsilon_{i,k})F_k' \\ &+ \sum_{t=2}^{T}\sum_{k=t+1}^{T} E(\varepsilon_{i,t}\varepsilon_{i,t-1}^{*})E(\varepsilon_{i,k})F_k' = 0, \end{aligned}$$

showing that the first term is zero too. In sum, therefore,

$$E(Q_{13}Q_{22}) = 0. \tag{A102}$$

For $E(Q_{12}Q_{22})$,

$$\begin{aligned} &E(Q_{12}Q_{22}) \\ &= \frac{1}{NT}\sum_{i=1}^{N} E[(\varepsilon_i'L_0F\lambda_i + \lambda_i'F'L_0\varepsilon_i)(\varepsilon_i'FK_0F'L_0F\lambda_i + \lambda_i'F'FK_0F'L_0\varepsilon_i)] \\ &= \frac{1}{NT}\sum_{i=1}^{N} E(\lambda_i'F'L_0'\varepsilon_i\varepsilon_i'FK_0F'L_0F\lambda_i + \lambda_i'F'L_0\varepsilon_i\varepsilon_i'FK_0F'L_0F\lambda_i + \lambda_i'F'L_0'\varepsilon_i\varepsilon_i'L_0'FK_0F'F\lambda_i \\ &+ \lambda_i'F'L_0\varepsilon_i\varepsilon_i'L_0'FK_0F'F\lambda_i), \end{aligned}$$

where

$$\begin{aligned} &\frac{1}{NT}\sum_{i=1}^{N} E(\lambda_i'F'L_0'\varepsilon_i\varepsilon_i'FK_0F'L_0F\lambda_i) \\ &= \frac{1}{NT}\sum_{i=1}^{N} \lambda_i'F'L_0'E(\varepsilon_i\varepsilon_i')FK_0F'L_0F\lambda_i \\ &= \sigma_0^2\frac{1}{N}\sum_{i=1}^{N} \lambda_i'T^{-1}F'L_0'F(TK_0)T^{-1}F'L_0F\lambda_i \\ &= \sigma_0^2\frac{1}{N}\sum_{i=1}^{N} \mathrm{tr}\left[\lambda_i'T^{-1}F'L_0'F(T^{-1}F'F)^{-1}T^{-1}F'L_0F\lambda_i\right] + O(T^{-1}) \\ &= \sigma_0^2\mathrm{tr}\left(S_\lambda T^{-1}F'L_0'P_FL_0F\right) + O(T^{-1}), \end{aligned}$$

and, by the same steps,

$$\begin{aligned} \frac{1}{NT}\sum_{i=1}^{N} E(\lambda_i'F'L_0\varepsilon_i\varepsilon_i'FK_0F'L_0F\lambda_i) &= \sigma_0^2\mathrm{tr}\left(S_\lambda T^{-1}F'L_0P_FL_0F\right) + O(T^{-1}), \\ \frac{1}{NT}\sum_{i=1}^{N} E(\lambda_i'F'L_0'\varepsilon_i\varepsilon_i'L_0'FK_0F'F\lambda_i) &= \sigma_0^2\mathrm{tr}\left(S_\lambda T^{-1}F'L_0'L_0'P_FF\right) + O(T^{-1}), \\ \frac{1}{NT}\sum_{i=1}^{N} E(\lambda_i'F'L_0\varepsilon_i\varepsilon_i'L_0'FK_0F'F\lambda_i) &= \sigma_0^2\mathrm{tr}\left(S_\lambda T^{-1}F'L_0L_0'P_FF\right) + O(T^{-1}). \end{aligned}$$

It follows that

$$E(Q_{12}Q_{22}) = \sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'(L_0' P_F L_0 + L_0 P_F L_0 + L_0' L_0' P_F + L_0 L_0' P_F) F\right] + O(T^{-1}), \quad \text{(A103)}$$

and so we obtain

$$\begin{aligned} & E[(Q_1 - Q_{11})(Q_2 - Q_{11})] \\ &= E(Q_{12}Q_{22}) + o(1) \\ &= \sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'(L_0' P_F L_0 + L_0 P_F L_0 + L_0' L_0' P_F + L_0 L_0' P_F) F\right] + o(1), \end{aligned} \quad \text{(A104)}$$

Hence, by combining the results,

$$\begin{aligned} & E[(Q_1 - Q_2)^2] \\ &= E[((Q_1 - Q_{11}) - (Q_2 - Q_{11}))^2] \\ &= E[(Q_1 - Q_{11})^2] + E[(Q_2 - Q_{11})^2] - 2E[(Q_1 - Q_{11})(Q_2 - Q_{11})] \\ &= \Sigma_{12} + \Sigma_{13} + \Sigma_{22} - 2E(Q_{12}Q_{22}) + o(1) \\ &= \sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'(L_0' L_0 + 2L_0 L_0 + L_0 L_0') F\right] + \sigma_0^4 T^{-1} \text{tr}\left(L_0 L_0'\right) \\ &+ \sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'(L_0' P_F L_0 + 2L_0 P_F L_0 + L_0 L_0') F\right] \\ &- 2\sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'(L_0' P_F L_0 + L_0 P_F L_0 + L_0' L_0' P_F + L_0 L_0' P_F) F\right] + o(1) \\ &= \sigma_0^4 T^{-1} \text{tr}\left(L_0 L_0'\right) \\ &+ \sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'(L_0' L_0 + 2L_0 L_0 + 2L_0 L_0' - L_0' P_F L_0 - 2L_0' L_0' P_F - 2L_0 L_0' P_F) F\right] + o(1) \\ &= \sigma_0^4 T^{-1} \text{tr}\left(L_0 L_0'\right) + \sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'(L_0' M_F L_0 + 2M_F L_0 L_0 + 2L_0 L_0' M_F) F\right] + o(1) \\ &= T^{-1} \text{tr}\left(\sigma_0^4 L_0 L_0' + \sigma_0^2 S_\lambda F' L_0' M_F L_0 F\right) + o(1) = \sigma_0^4 \omega_1^2 + o(1), \end{aligned} \quad \text{(A105)}$$

where $\omega_1^2 = T^{-1} \text{tr}\left(L_0 L_0' + \sigma_0^{-2} S_\lambda F' L_0' M_F L_0 F\right)$. The sixth equality is a direct consequence of $\text{tr}\left(S_\lambda T^{-1} F' L_0' L_0' P_F F\right) = \text{tr}\left(S_\lambda T^{-1} F' P_F L_0 L_0 F\right)$, while the seventh is due to $F' M_F = 0_{m \times T}$ and $M_F F = 0_{T \times m}$.

Thus, putting everything together,

$$\begin{aligned} Q_1 - Q_2 &= Q_{11} + Q_{12} + Q_{13} - (Q_{11} + Q_{22}) + O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}) \\ &= Q_{12} + Q_{13} - Q_{22} + O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}) \\ &\sim N(0, \sigma_0^4 \omega_1^2), \end{aligned} \quad \text{(A106)}$$

which holds for $T \to \infty$ and any N, including $N \to \infty$, provided that $\sqrt{N} T^{-3/2} = o(1)$. The implication of this result is that

$$\frac{1}{\sqrt{NT}} \frac{\partial \ell_c(\theta_2^0)}{\partial \rho} = \frac{1}{2\sigma_0^2} Q = \sigma_0^{-2}(Q_1 - Q_2) \sim N(0, \omega_1^2). \quad \text{(A107)}$$

Next, consider $\partial \ell_c(\theta_2^0)/\partial \sigma^2$, which we write as

$$\frac{\partial \ell_c(\theta_2^0)}{\partial \sigma^2} = -\frac{NT}{2\sigma_0^2} + \frac{N}{2\sigma_0^4}[\text{vec}\, B(\theta_2^0)]'\text{vec}\, G(\rho_0) = -\frac{NT}{2\sigma_0^2} + \frac{N}{2\sigma_0^4}\text{tr}\,[B(\theta_2^0)G(\rho_0)], \tag{A108}$$

where we know from before that $B(\theta_2^0) = \hat{\Lambda}(\theta_2^0)^{-1} = I_T - FK_0F' = I_T - F(F'F)^{-1}F' + \sigma_0^2 T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'$, suggesting

$$\begin{aligned}
&\text{tr}\,[\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)] \\
&= \text{tr}\, G(\rho_0) - \text{tr}\,[F(F'F)^{-1}F'G(\rho_0)] + \sigma_0^2\text{tr}\,[T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'G(\rho_0)] \\
&= \text{tr}\, G(\rho_0) - \text{tr}\,[(F'F)^{-1}F'G(\rho_0)F] + \sigma_0^2\text{tr}\, I_2,
\end{aligned}$$

which can be substituted back into $\partial \ell_c(\theta_2^0)/\partial \sigma^2$;

$$\begin{aligned}
\frac{\partial \ell_c(\theta_2^0)}{\partial \sigma^2} &= -\frac{NT}{2\sigma_0^2} + \frac{N}{2\sigma_0^4}\text{tr}\,[B(\theta_2^0)G(\rho_0)] \\
&= -\frac{N(T-2)}{2\sigma_0^2} + \frac{N}{2\sigma_0^4}(\text{tr}\, G(\rho_0) - \text{tr}\,[(F'F)^{-1}F'G(\rho_0)F]) \\
&= \frac{N}{2\sigma_0^4}(\text{tr}\,[G(\rho_0) - \sigma_0^2 I_T] - \text{tr}\,[(F'F)^{-1}F'G(\rho_0)F - \sigma_0^2 I_2]).
\end{aligned} \tag{A109}$$

Clearly, $G(\rho_0) = \Gamma_0^{-1}S_y\Gamma_0^{-1\prime} = \Gamma_0^{-1}(\Gamma_0 S_u \Gamma_0')\Gamma_0^{-1\prime} = S_u$, from which it follows that

$$\begin{aligned}
\text{tr}\,[(F'F)^{-1}F'G(\rho_0)F - \sigma_0^2 I_2] &= \text{tr}\,[(F'F)^{-1}F'S_uF - \sigma_0^2 I_2] \\
&= \frac{1}{N}\sum_{i=1}^{N}(\lambda_i'F'F\lambda_i + 2\varepsilon_i'F\lambda_i + \text{tr}\,[(F'F)^{-1}F'\varepsilon_i\varepsilon_i'F - \sigma_0^2 I_2]), \\
\text{tr}\,[G(\rho_0) - \sigma_0^2 I_T] &= \text{tr}\,(S_u - \sigma_0^2 I_T) \\
&= \frac{1}{N}\sum_{i=1}^{N}[\lambda_i'F'F\lambda_i + 2\varepsilon_i'F\lambda_i + \text{tr}\,(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)].
\end{aligned}$$

By using $E[(\varepsilon_{i,t}^2 - \sigma_0^2)^2] = E(\varepsilon_{i,t}^4 - 2\sigma_0^2\varepsilon_{i,t}^2 + \sigma_0^4) = E(\varepsilon_{i,t}^4) - \sigma_0^4$, it is possible to show that, by a CLT, as $N \to \infty$ or $T \to \infty$, or both,

$$\begin{aligned}
\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\text{tr}\,(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T) &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}(\varepsilon_i'\varepsilon_i - \sigma_0^2 T) = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{i,t}^2 - \sigma_0^2) \\
&\to_d N(0, E(\varepsilon_{i,t}^4) - \sigma_0^4).
\end{aligned}$$

But we also have

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\text{tr}\,[(F'F)^{-1}F'\varepsilon_i\varepsilon_i'F - \sigma_0^2 I_2] = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\text{tr}\,[(T^{-1}F'F)^{-1}T^{-1}F'\varepsilon_i\varepsilon_i'F - \sigma_0^2 I_2] = O_p(1),$$

and therefore, with $\kappa_0 = \sigma_0^{-4} E(\varepsilon_{i,t}^4)$,

$$\begin{aligned}
\frac{1}{\sqrt{NT}} \frac{\partial \ell_c(\theta_2^0)}{\partial \sigma^2} &= \frac{1}{2\sigma_0^4} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} [\text{tr}\,(\varepsilon_i \varepsilon_i' - \sigma_0^2 I_T) - \text{tr}\,((F'F)^{-1} F' \varepsilon_i \varepsilon_i' F - \sigma_0^2 I_2)] \\
&= \frac{1}{2\sigma_0^4} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} (\varepsilon_{i,t}^2 - \sigma_0^2) + O_p(T^{-1/2}) \\
&\to_d N\left(0, \frac{(\kappa_0 - 1)}{4\sigma_0^4}\right), \qquad \text{(A110)}
\end{aligned}$$

which requires $T \to \infty$, but not necessarily $N \to \infty$.

In what remains we show that

$$\begin{aligned}
&\frac{1}{NT} E\left(\frac{\partial \ell_c(\theta_2^0)}{\partial \sigma^2} \frac{\partial \ell_c(\theta_2^0)}{\partial \rho}\right) \\
&= \frac{1}{2\sigma_0^6} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} E[(\varepsilon_{i,t}^2 - \sigma_0^2)(Q_{12} + Q_{13} - Q_{22})] + O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}) \\
&= O_p(T^{-1/2}) + O_p(\sqrt{N} T^{-3/2}). \qquad \text{(A111)}
\end{aligned}$$

The proof begins by the following observation:

$$\begin{aligned}
\sum_{t=1}^{T} \sum_{s=2}^{T} E[(\varepsilon_{i,t}^2 - \sigma_0^2) \varepsilon_{i,s} \varepsilon_{i,s-1}^*] &= \sum_{t=2}^{T} E[(\varepsilon_{i,t}^2 - \sigma_0^2) \varepsilon_{i,t}] E(\varepsilon_{i,t-1}^*) + \sum_{t=3}^{T} \sum_{s=2}^{t-1} E(\varepsilon_{i,t}^2 - \sigma_0^2) E(\varepsilon_{i,s} \varepsilon_{i,s-1}^*) \\
&+ \sum_{t=1}^{T} \sum_{s=t+1}^{T} E(\varepsilon_{i,s}) E[(\varepsilon_{i,t}^2 - \sigma_0^2) \varepsilon_{i,s-1}^*] = 0,
\end{aligned}$$

suggesting that $(NT)^{-1/2} \partial \ell_c(\theta_2^0)/\partial \sigma^2$ is uncorrelated with Q_{13}. But we also have

$$\begin{aligned}
\sum_{t=1}^{T} \sum_{s=2}^{T} E[(\varepsilon_{i,t}^2 - \sigma_0^2) \varepsilon_{i,s-1}^* F_s] &= \sum_{t=2}^{T} E(\varepsilon_{i,t}^2 - \sigma_0^2) E(\varepsilon_{i,t-1}^*) F_t + \sum_{t=3}^{T} \sum_{s=2}^{t-1} E(\varepsilon_{i,t}^2 - \sigma_0^2) E(\varepsilon_{i,s-1}^*) F_s \\
&+ \sum_{t=1}^{T} \sum_{s=t+1}^{T} F_s E[(\varepsilon_{i,t}^2 - \sigma_0^2) \varepsilon_{i,s-1}^*] = \sum_{t=1}^{T} \sum_{s=t+1}^{T} F_s E(\varepsilon_{i,t}^2 \varepsilon_{i,s-1}^*), \\
\sum_{t=1}^{T} \sum_{s=2}^{T} E[(\varepsilon_{i,t}^2 - \sigma_0^2) \varepsilon_{i,s} F_{s-1}^*] &= \sum_{t=2}^{T} E[(\varepsilon_{i,t}^2 - \sigma_0^2) \varepsilon_{i,t}] F_{t-1}^* + \sum_{t=3}^{T} \sum_{s=2}^{t-1} E(\varepsilon_{i,t}^2 - \sigma_0^2) E(\varepsilon_{i,s}) F_{i,s-1}^* \\
&+ \sum_{t=1}^{T} \sum_{s=t+1}^{T} E(\varepsilon_{i,s}) E(\varepsilon_{i,t}^2 - \sigma_0^2) F_{i,s-1}^* = \sum_{t=2}^{T} E(\varepsilon_{i,t}^3) F_{t-1}^*,
\end{aligned}$$

which are both zero if $E(\varepsilon_{i,t}^3) = 0$. Hence, under this condition ,$(NT)^{-1/2} \partial \ell_c(\theta_2^0)/\partial \sigma^2$ is not only uncorrelated with Q_{13}, but also with Q_{12} and Q_{22}. It follows that $(NT)^{-1/2} \partial \ell_c(\theta_2^0)/\partial \sigma^2$ and $(NT)^{-1/2} \partial \ell_c(\theta_2^0)/\partial \rho$ are asymptotically uncorrelated, and hence independent by normality. Therefore,

$$\frac{1}{\sqrt{NT}} \begin{bmatrix} \frac{\partial \ell_c(\theta_2^0)}{\partial \rho} \\ \frac{\partial \ell_c(\theta_2^0)}{\partial \sigma^2} \end{bmatrix} \sim N\left(0_{2\times 1}, \begin{bmatrix} \omega_1^2 & 0 \\ 0 & \frac{(\kappa_0 - 1)}{4\sigma_0^4} \end{bmatrix}\right) \qquad \text{(A112)}$$

as $T \to \infty$ and any N, including $N \to \infty$, provided that $\sqrt{N}T^{-3/2} = o(1)$. ■

Lemma C.2. *Under C1, $|\rho_0| < 1$, and Assumptions EPS, F and LAM,*

$$-H_{1/2}^{-1}\frac{\partial^2 \ell_c(\theta_2^0)}{\partial\theta_2(\partial\theta_2)'}H_{1/2}^{-1} = \begin{bmatrix} \omega_1^2 & 0 \\ 0 & \frac{1}{2\sigma_0^4} \end{bmatrix} + O_p(T^{-1}) + O_p((NT)^{-1/2}).$$

Proof of Lemma C.2.

Write

$$\frac{2\sigma_0^2}{NT}\frac{\partial^2 \ell_c(\theta_2^0)}{(\partial\rho)^2} = U_1 + ... + U_6, \tag{A113}$$

where

$$\begin{aligned}
U_1 &= \sigma_0^{-2}T^{-1}C(\rho_0)'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-]C(\rho_0), \\
U_2 &= -\sigma_0^{-4}T^{-1}C(\rho_0)'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-]C(\rho_0), \\
U_3 &= -\sigma_0^{-4}T^{-1}C(\rho_0)'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^-]C(\rho_0), \\
U_4 &= \sigma_0^{-2}T^{-1}C(\rho_0)'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}]C(\rho_0), \\
U_5 &= \sigma_0^{-2}T^{-1}C(\rho_0)'[\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2^0)^{-1}FF^-]C(\rho_0), \\
U_6 &= -2T^{-1}[\mathrm{vec}\, B(\theta_2^0)]'(J \otimes J)\mathrm{vec}\, S_y.
\end{aligned}$$

From Proof of Lemma C.1, $C(\rho_0) = \mathrm{vec}\,(L_0S_u + S_uL_0')$. By using this, $(C' \otimes A)\mathrm{vec}\, B = \mathrm{vec}(ABC)$ and $\mathrm{tr}\,(A'B) = (\mathrm{vec}\, A)'\mathrm{vec}\, B$, we obtain

$$\begin{aligned}
U_1 &= \sigma_0^{-2}T^{-1}C(\rho_0)'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-]C(\rho_0) \\
&= \sigma_0^{-2}T^{-1}C(\rho_0)'\mathrm{vec}\,[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-(L_0S_u + S_uL_0')F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-] \\
&= \sigma_0^{-2}T^{-1}\mathrm{tr}\,[(L_0S_u + S_uL_0')'F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-(L_0S_u + S_uL_0')F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-].
\end{aligned}$$

From Proof of Lemma C.1, we further learn that $F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^- = \sigma_0^2T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'$. Also,

$$\begin{aligned}
T^{-2}F'L_0S_uF &= \frac{1}{NT^2}\sum_{i=1}^N(F'L_0F\lambda_i\lambda_i'F'F + F'L_0F\lambda_i\varepsilon_i'F + F'L_0\varepsilon_i\lambda_i'F'F + F'L_0\varepsilon_i\varepsilon_i'F) \\
&= \frac{1}{NT^2}\sum_{i=1}^N F'L_0F\lambda_i\lambda_i'F'F + (NT)^{-1/2}(T^{-1}F'L_0F)\frac{1}{\sqrt{NT}}\sum_{i=1}^N\lambda_i\varepsilon_i'F \\
&+ (NT)^{-1/2}\frac{1}{\sqrt{NT}}\sum_{i=1}^N F'L_0\varepsilon_i\lambda_i'(T^{-1}F'F) + \frac{1}{T}\frac{1}{NT}\sum_{i=1}^N F'L_0\varepsilon_i\varepsilon_i'F \\
&= T^{-2}F'L_0FS_\lambda F'F + O_p((NT)^{-1/2}) + O_p(T^{-1}).
\end{aligned} \tag{A114}$$

The limit of $T^{-2}F'S_uL_0'F$ is just the transpose of this, suggesting that

$$\begin{aligned} & T^{-2}F'(L_0S_u + S_uL_0')'F \\ &= T^{-2}F'S_uL_0'F + T^{-2}F'L_0S_uF \\ &= T^{-2}(F'L_0FS_\lambda F'F + F'FS_\lambda F'L_0'F) + O_p((NT)^{-1/2}) + O_p(T^{-1}). \end{aligned} \tag{A115}$$

By using this and the fact that $T^{-1}F'L_0F$ and $T^{-1}F'F$ are bounded by assumption, we can show that $||T^{-2}F'(L_0S_u + S_uL_0')'F|| = O_p(1)$. But $||T^{-2}F'G(\rho_0)F||$ is of the same order, and therefore

$$\begin{aligned} U_1 &= T^{-1}\sigma_0^2 \text{tr}\,[T^{-2}F'(L_0S_u + S_uL_0')'F(T^{-2}F'G(\rho_0)F)^{-1}T^{-2}F'(L_0S_u + S_uL_0')F \\ &\times (T^{-2}F'G(\rho_0)F)^{-1}] = O_p(T^{-1}). \end{aligned} \tag{A116}$$

Consider U_2. By using the results reported in Proof of Lemma C.1 for $F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-$ and $F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^-$, we obtain

$$\begin{aligned} & U_2 \\ &= -\sigma_0^{-4}T^{-1}C(\rho_0)'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-]C(\rho_0), \\ &= -\sigma_0^{-4}T^{-1}C(\rho_0)'\text{vec}\,[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-(L_0S_u + S_uL_0')F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^-] \\ &= -\sigma_0^{-4}T^{-1}\text{tr}\,[(L_0S_u + S_uL_0')'F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-(L_0S_u + S_uL_0') \\ &\times F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^-] \\ &= -\sigma_0^2T^{-1}\text{tr}\,[T^{-2}F'(L_0S_u + S_uL_0')'F(T^{-2}F'G(\rho_0)F)^{-1}T^{-2}F'(L_0S_u + S_uL_0')F \\ &\times (T^{-2}F'G(\rho_0)F)^{-1}] = -U_1, \end{aligned} \tag{A117}$$

and it is not difficult to see that, $U_3 = U_2 = -U_1$. Moreover,

$$\begin{aligned} F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1} &= F(F'F)^{-1}F'(I_T - FK_0F') = F((F'F)^{-1} - K_0)F' \\ &= \sigma_0^2T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F' = \hat{\Lambda}(\theta_2^0)^{-1}FF^-, \end{aligned}$$

suggesting

$$\begin{aligned} U_4 &= \sigma_0^{-2}T^{-1}C(\rho_0)'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}]C(\rho_0) \\ &= \sigma_0^{-2}T^{-1}C(\rho_0)'\text{vec}\,[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}(L_0S_u + S_uL_0')\hat{\Lambda}(\theta_2^0)^{-1}FF^-] \\ &= \sigma_0^2T^{-1}\text{tr}\,[T^{-2}F'(L_0S_u + S_uL_0')'F(T^{-2}F'G(\rho_0)F)^{-1}T^{-2}F'(L_0S_u + S_uL_0')F \\ &\times (T^{-2}F'G(\rho_0)F)^{-1}] = U_1, \end{aligned} \tag{A118}$$

which is true also for U_5, that is, $U_5 = U_1$.

It remains to consider U_6. From $J\Gamma_0 = L_0$ and $(C' \otimes A)\text{vec}\, B = \text{vec}(ABC)$,

$$(J \otimes J)\text{vec}\, S_y = (J \otimes J)(\Gamma_0 \otimes \Gamma_0)\text{vec}\, S_u = (L_0 \otimes L_0)\text{vec}\, S_u = \text{vec}\,(L_0 S_u L_0').$$

Also, $\hat{\Lambda}(\theta_2^0)^{-1} = I_T - F(F'F)^{-1}F' + \sigma_0^2 T^{-2} F(T^{-2}F'G(\rho_0)F)^{-1}F'$, which we can use to show that

$$\begin{aligned}
U_6 &= -2T^{-1}[\text{vec}\, B(\theta_2^0)]'(J \otimes J)\text{vec}\, S_y \\
&= -2T^{-1}\text{tr}\,[(F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-} - \sigma_0^{-2}F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^{-} + \hat{\Lambda}(\theta_2)^{-1}) \\
&\times\ L_0 S_u L_0'] \\
&= -2T^{-1}\text{tr}\,[(I_T - F(F'F)^{-1}F' + \sigma_0^2 T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F')L_0 S_u L_0'] \\
&= -2\text{tr}\,(T^{-1}L_0 S_u L_0') + 2\text{tr}\,[((T^{-1}F'F)^{-1} - \sigma_0^2 T^{-1}(T^{-2}F'G(\rho_0)F)^{-1})T^{-2}F'L_0 S_u L_0'F] \\
&= -2\text{tr}\,(T^{-1}L_0 S_u L_0') + 2\text{tr}\,[(T^{-1}F'F)^{-1}T^{-2}F'L_0 S_u L_0'F] + O_p(T^{-1}).
\end{aligned}$$

Here

$$\begin{aligned}
&T^{-2}F'L_0 S_u L_0'F \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}(F'L_0 F\lambda_i\lambda_i'F'L_0'F + F'L_0 F\lambda_i\varepsilon_i'L_0'F + F'L_0\varepsilon_i\lambda_i'F'L_0'F + F'L_0\varepsilon_i\varepsilon_i'L_0'F) \\
&= \frac{1}{NT^2}\sum_{i=1}^{N}F'L_0 F\lambda_i\lambda_i'F'L_0'F + (NT)^{-1/2}(T^{-1}F'L_0F)\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i\varepsilon_i'L_0'F \\
&+ \frac{1}{\sqrt{NT}}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}F'L_0\varepsilon_i\lambda_i'(T^{-1}F'L_0'F) + \frac{1}{T}\frac{1}{NT}\sum_{i=1}^{N}F'L_0\varepsilon_i\varepsilon_i'L_0'F \\
&= T^{-2}F'L_0 F S_\lambda F'L_0'F + O_p((NT)^{-1/2}) + O_p(T^{-1}),
\end{aligned} \tag{A119}$$

and therefore

$$\begin{aligned}
&\text{tr}\,[(T^{-1}F'F)^{-1}T^{-2}F'L_0 S_u L_0'F] \\
&= \text{tr}\,[(T^{-1}F'F)^{-1}T^{-2}F'L_0 F S_\lambda F'L_0'F] + O_p((NT)^{-1/2}) + O_p(T^{-1}) \\
&= \text{tr}\,(S_\lambda F'L_0'P_F L_0 F) + O_p((NT)^{-1/2}) + O_p(T^{-1}).
\end{aligned}$$

For $\text{tr}\,(T^{-1}L_0 S_u L_0')$, we use that $T^{-1/2}\varepsilon_i'L_0'L_0F\lambda_i = T^{-1/2}\sum_{t=2}^{T}\varepsilon_{i,t-1}^* F_{t-1}^{*\prime}\lambda_i$, which is asymptotically normal as $T \to \infty$, suggesting that $(NT)^{-1/2}\sum_{i=1}^{N}\varepsilon_i'L_0'L_0F\lambda_i = O_p(1)$ for any N, including $N \to \infty$, provided that $T \to \infty$. But the same is true for $(NT)^{-1/2}\sum_{i=1}^{N}\text{tr}\,[L_0(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)L_0']$, because $T^{-1/2}\text{tr}\,[L_0(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)L_0'] = T^{-1/2}\sum_{t=2}^{T}[(\varepsilon_{i,t-1}^*)^2 - \sigma_0^2(T-1)^{-1}\text{tr}\,(L_0L_0')]$

is asymptotically normal too. It follows that

$$\begin{aligned}
&\text{tr}\,(T^{-1}L_0S_uL_0') \\
&= \frac{1}{NT}\sum_{i=1}^{N}[\text{tr}\,(L_0F\lambda_i\lambda_i'F'L_0') + 2\text{tr}\,(L_0F\lambda_i\varepsilon_i'L_0') + \text{tr}\,(L_0\varepsilon_i\varepsilon_i'L_0')] \\
&= \frac{1}{N}\sum_{i=1}^{N}\text{tr}\,(\lambda_i\lambda_i'T^{-1}F'L_0'L_0F) + \frac{1}{\sqrt{NT}}\frac{2}{\sqrt{NT}}\sum_{i=1}^{N}\varepsilon_i'L_0'L_0F\lambda_i \\
&+ \sigma_0^2\text{tr}\,(T^{-1}L_0'L_0) + \frac{1}{\sqrt{NT}}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\text{tr}\,[(\varepsilon_i\varepsilon_i' - \sigma_0^2I_T)L_0'L_0] \\
&= \text{tr}\,(S_\lambda T^{-1}F'L_0'L_0F) + \sigma_0^2\text{tr}\,(T^{-1}L_0'L_0) + O_p((NT)^{-1/2}),
\end{aligned} \tag{A120}$$

from which the following result for U_6 is obtained:

$$\begin{aligned}
U_6 &= -2\text{tr}\,(T^{-1}L_0S_uL_0') + 2\text{tr}\,[(T^{-1}F'F)^{-1}T^{-2}F'L_0S_uL_0'F] + O_p(T^{-1}) \\
&= -2T^{-1}\text{tr}\,(\sigma_0^2L_0'L_0 + S_\lambda F'L_0'M_FL_0F) + O_p(T^{-1}) + O_p((NT)^{-1/2}).
\end{aligned} \tag{A121}$$

Hence,

$$\begin{aligned}
-\frac{1}{NT}\frac{\partial^2\ell_c(\theta_2^0)}{(\partial\rho)^2} &= -\frac{1}{2\sigma_0^2}(U_1 + ... + U_6) = -\frac{1}{2\sigma_0^2}(U_1 + U_6) = -\frac{1}{2\sigma_0^2}U_6 + O_p(T^{-1}) \\
&= T^{-1}\text{tr}\,(L_0'L_0 + \sigma_0^{-2}S_\lambda F'L_0'M_FL_0F) + O_p(T^{-1}) + O_p((NT)^{-1/2}) \\
&= \omega_1^2 + O_p(T^{-1}) + O_p((NT)^{-1/2}).
\end{aligned} \tag{A122}$$

Next, consider $\partial^2\ell_c(\theta_2^0)/(\partial\sigma^2)^2$, which we write as

$$\frac{2\sigma_0^4}{NT}\frac{\partial^2\ell_c(\theta_2)}{(\partial\sigma^2)^2} = P_1 + ... + P_5, \tag{A123}$$

with

$$\begin{aligned}
P_1 &= 1 - 2\sigma_0^{-2}T^{-1}[\text{vec}\,B(\theta_2^0)]'\text{vec}\,G(\rho_0), \\
P_2 &= \sigma_0^{-4}T^{-1}[\text{vec}\,G(\rho_0)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-]\text{vec}\,G(\rho_0), \\
P_3 &= -\sigma_0^{-6}T^{-1}[\text{vec}\,G(\rho_0)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^-] \\
&\times \text{vec}\,G(\rho_0), \\
P_4 &= -\sigma_0^{-6}T^{-1}[\text{vec}\,G(\rho_0)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^-] \\
&\times \text{vec}\,G(\rho_0), \\
P_5 &= \sigma_0^{-4}T^{-1}[\text{vec}\,G(\rho_0)]'[\hat{\Lambda}(\theta_2^0)^{-1}FF^- \otimes \hat{\Lambda}(\theta_2^0)^{-1}FF^-]\text{vec}\,G(\rho_0).
\end{aligned}$$

From Proof of Lemma 1,

$$\begin{aligned} T^{-1}\text{tr}\left[\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\right] &= \text{tr}\left[T^{-1}G(\rho_0)\right] - \text{tr}\left[(T^{-1}F'F)^{-1}T^{-2}F'G(\rho_0)F\right] + \sigma_0^2 T^{-1}\text{tr}\, I_2 \\ &= \sigma_0^2 + \text{tr}\left(S_\lambda T^{-1}F'F\right) - \text{tr}\left(S_\lambda T^{-1}F'F\right) + \sigma_0^2 T^{-1}\text{tr}\, I_2 + O_p(T^{-1}) \\ &= \sigma_0^2 + O_p(T^{-1}), \end{aligned}$$

leading to the following result for P_1:

$$\begin{aligned} P_1 &= 1 - 2\sigma_0^{-2}T^{-1}[\text{vec}\, B(\theta_2^0)]'\text{vec}\, G(\rho_0) = 1 - 2\sigma_0^{-2}T^{-1}\text{tr}\left[B(\theta_2^0)G(\rho_0)\right], \\ &= 1 - 2\sigma_0^{-2}T^{-1}\text{tr}\left[\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\right] = -1 + O_p(T^{-1}). \end{aligned} \tag{A124}$$

For P_2, via $F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-} = \sigma_0^2 T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'$,

$$\begin{aligned} P_2 &= \sigma_0^{-4}T^{-1}[\text{vec}\, G(\rho_0)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}]\text{vec}\, G(\rho_0) \\ &= \sigma_0^{-4}T^{-1}[\text{vec}\, G(\rho_0)]'\text{vec}\left[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}G(\rho_0)F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}\right] \\ &= \sigma_0^{-4}T^{-1}\text{tr}\left[G(\rho_0)F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}G(\rho_0)F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}\right] \\ &= T^{-1}\text{tr}\left[G(\rho_0)T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'G(\rho_0)T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'\right] \\ &= T^{-1}\text{tr}\, I_2, \end{aligned} \tag{A125}$$

whereas for P_3,

$$\begin{aligned} P_3 &= -\sigma_0^{-6}T^{-1}[\text{vec}\, G(\rho_0)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^{-} \otimes F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}] \\ &\times \text{vec}\, G(\rho_0) \\ &= -\sigma_0^{-6}T^{-1}[\text{vec}\, G(\rho_0)]'[F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}G(\rho_0)F^{-\prime}F'\hat{\Lambda}(\theta_2^0)^{-1}G(\rho_0)\hat{\Lambda}(\theta_2^0)^{-1}FF^{-}] \\ &= -T^{-1}\text{tr}\left[G(\rho_0)T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'G(\rho_0)T^{-2}F(T^{-2}F'G(\rho_0)F)^{-1}F'\right] \\ &= -T^{-1}\text{tr}\, I_2 = -P_2. \end{aligned} \tag{A126}$$

We can similarly show that $P_4 = -P_2$ and $P_5 = P_2$. Therefore,

$$-\frac{1}{NT}\frac{\partial^2 \ell_c(\theta_2^0)}{(\partial\sigma^2)^2} = -\frac{1}{2\sigma_0^4}(P_1 + ... + P_5) = -\frac{1}{2\sigma_0^4}P_1 + O_p(T^{-1}) = \frac{1}{2\sigma_0^4} + O_p(T^{-1}). \tag{A127}$$

$\partial^2 \ell_c(\theta_2^0)/(\partial\rho\partial\sigma^2)$ satisfies

$$\frac{2\sigma_0^4}{NT}\frac{\partial^2 \ell_c(\theta_2^0)}{\partial\rho\partial\sigma^2} = R_1 + ... + R_5,$$

where

$$
\begin{aligned}
R_1 &= -T^{-1}[\text{vec}\, B(\theta_2^0)]' C(\rho_0), \\
R_2 &= \sigma_0^{-2} T^{-1} C(\rho_0)' [F^{-\prime} F' \hat{\Lambda}(\theta_2^0)^{-1} F F^- \otimes F^{-\prime} F' \hat{\Lambda}(\theta_2^0)^{-1} F F^-] \text{vec}\, G(\rho_0), \\
R_3 &= -\sigma_0^{-4} T^{-1} C(\rho_0)' [F^{-\prime} F' \hat{\Lambda}(\theta_2^0)^{-1} G(\rho_0) \hat{\Lambda}(\theta_2^0)^{-1} F F^- \otimes F^{-\prime} F' \hat{\Lambda}(\theta_2^0)^{-1} F F^-] \text{vec}\, G(\rho_0), \\
R_4 &= -\sigma_0^{-4} T^{-1} C(\rho_0)' [F^{-\prime} F' \hat{\Lambda}(\theta_2^0)^{-1} F F^- \otimes F^{-\prime} F' \hat{\Lambda}(\theta_2^0)^{-1} G(\rho_0) \hat{\Lambda}(\theta_2^0)^{-1} F F^-] \text{vec}\, G(\rho_0), \\
R_5 &= \sigma_0^{-2} T^{-1} C(\rho_0)' [\hat{\Lambda}(\theta_2^0)^{-1} F F^- \otimes \hat{\Lambda}(\theta_2^0)^{-1} F F^-] \text{vec}\, G(\rho_0).
\end{aligned}
$$

As in $\partial^2 \ell_c(\theta_2^0)/(\partial \sigma^2)^2$, the first term in the expansion is also the dominant term. We therefore focus on this. From $C(\rho_0) = \text{vec}\,(L_0 S_u + S_u L_0')$,

$$
\begin{aligned}
R_1 &= -T^{-1}[\text{vec}\, B(\theta_2^0)]' C(\rho_0) = -T^{-1} \text{tr}\,[\hat{\Lambda}(\theta_2^0)^{-1}(L_0 S_u + S_u L_0')] \\
&= -T^{-1} \text{tr}\,[(I_T - F(F'F)^{-1}F' + \sigma_0^2 T^{-2} F (T^{-2} F' G(\rho_0) F)^{-1} F')(L_0 S_u + S_u L_0')] \\
&= -2\text{tr}\,(T^{-1} L_0 S_u) + 2\text{tr}\,[(T^{-1} F'F)^{-1} T^{-2} F' L_0 S_u F] \\
&- 2\sigma_0^2 T^{-1} \text{tr}\,[(T^{-2} F' G(\rho_0) F)^{-1} T^{-2} F' L_0 S_u F] \\
&= -2\text{tr}\,(T^{-1} L_0 S_u) + 2\text{tr}\,[(T^{-1} F'F)^{-1} T^{-2} F' L_0 S_u F] + O_p(T^{-1}).
\end{aligned}
$$

Here

$$
\begin{aligned}
\text{tr}\,(T^{-1} L_0 S_u) &= \frac{1}{NT} \sum_{i=1}^{N} [\text{tr}\,(\lambda_i \lambda_i' F' L_0 F) + \text{tr}\,(L_0 F \lambda_i \varepsilon_i') + \text{tr}\,(L_0 \varepsilon_i \lambda_i' F') + \text{tr}\,(L_0 \varepsilon_i \varepsilon_i')] \\
&= \text{tr}\,(S_\lambda T^{-1} F' L_0 F) + O_p((NT)^{-1/2}), \qquad \text{(A128)}
\end{aligned}
$$

suggesting that, with $T^{-2} F' L_0 S_u F = T^{-2} F' L_0 F S_\lambda F' F + O_p((NT)^{-1/2}) + O_p(T^{-1})$,

$$
\begin{aligned}
R_1 &= -2\text{tr}\,(T^{-1} L_0 S_u) + 2\text{tr}\,[(T^{-1} F'F)^{-1} T^{-2} F' L_0 S_u F] + O_p(T^{-1}) \\
&= -2\text{tr}\,(S_\lambda T^{-1} F' L_0 F) + 2\text{tr}\,(S_\lambda T^{-1} F' L_0 F) + O_p(T^{-1}) + O_p((NT)^{-1/2}) \\
&= O_p(T^{-1}) + O_p((NT)^{-1/2}). \qquad \text{(A129)}
\end{aligned}
$$

The other terms are all of smaller order in magnitude than this. Therefore,

$$
\frac{2\sigma_0^4}{NT} \frac{\partial^2 \ell_c(\theta_2^0)}{\partial \rho \partial \sigma^2} = O_p(T^{-1}) + O_p((NT)^{-1/2}), \qquad \text{(A130)}
$$

as was to be shown. This completes the proof of the lemma. ■

Proof of Theorem 1.

In view of Lemmas C.1 and C.2, the proof of Theorem 1 follows by standard arguments (see, for example, Amemiya, 1985, Chapter 4). It can be shown that the third-order derivative of $\ell_c(\theta_2^0)$ is bounded. By using this and Taylor expansion of $\ell_c(\theta_2)$ about $\theta_2 = \theta_2^0$;

$$\ell_c(\theta_2) = \ell_c(\theta_2^0) + \frac{\partial \ell_c(\theta_2^0)}{\partial \theta_2'}(\theta_2 - \theta_2^0) + \frac{1}{2}(\theta_2 - \theta_2^0)' \frac{\partial^2 \ell_c(\theta_2^0)}{\partial \theta_2 \partial \theta_2'}(\theta_2 - \theta_2^0) + O_p((NT)^{-1/2}), \quad \text{(A131)}$$

suggesting that

$$\begin{aligned} NT[\ell_c(\theta_2)(\rho) - \ell_c(\theta_2^0)] &= \sqrt{NT}\frac{\partial \ell_c(\theta_2^0)}{\partial \theta_2'}\sqrt{NT}(\theta_2 - \theta_2^0) \\ &+ \frac{1}{2}\sqrt{NT}(\theta_2 - \theta_2^0)' \frac{\partial^2 \ell_c(\theta_2^0)}{\partial \theta_2 \partial \theta_2'}\sqrt{NT}(\theta_2 - \theta_2^0) + O_p((NT)^{-1/2}). \end{aligned}$$

$\hat{\theta}_2$ is the minimizer of $NT[\ell_c(\theta_2) - \ell_c(\theta_2^0)]$. Thus, treating this as a function of $\sqrt{NT}(\theta_2 - \theta_2^0)$, we obtain the following first order condition:

$$\begin{aligned} &\sqrt{NT}\frac{\partial \ell_c(\theta_2^0)}{\partial \theta_2'}\sqrt{NT}(\theta_2 - \theta_2^0) \\ &+ \frac{1}{2}\sqrt{NT}(\theta_2 - \theta_2^0)' \frac{\partial^2 \ell_c(\theta_2^0)}{\partial \theta_2 \partial \theta_2'}\sqrt{NT}(\theta_2 - \theta_2^0) + O_p((NT)^{-1/2}) = 0, \end{aligned}$$

implying

$$\sqrt{NT}(\hat{\theta}_2 - \theta_2^0) = -\left(\frac{1}{NT}\frac{\partial^2 \ell_c(\theta_2^0)}{\partial \theta_2 \partial \theta_2'}\right)^{-1} \frac{1}{\sqrt{NT}}\frac{\partial \ell_c(\theta_2^0)}{\partial \theta_2'} + O_p((NT)^{-1/2}). \quad \text{(A132)}$$

The required result is now a direct consequence of Lemmas C.1 and C.2. ■

Proof of Lemma 2.

The proof of Lemma 2 follows from simple manipulations of Proof of Lemma 1. In particular, since $\rho_0 = 1$ affects the order of all quantities involving L_0, all the results involving such terms will have to be reevaluated.

We start by considering $T^{-4}F'GF$ (earlier $T^{-2}F'GF$), which has the same expansion as in (A62), repeated here for convenience;

$$\begin{aligned} T^{-4}F'GF &= T^{-4}F'\Gamma^{-1}S_y\Gamma^{-1\prime}F = T^{-4}F'\Gamma^{-1}\Gamma_0 S_u \Gamma_0'\Gamma^{-1\prime}F \\ &= T^{-4}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}F + T^{-3}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^N F\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}F \\ &+ T^{-3}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^N \varepsilon_i\lambda_i'F'\Gamma_0'\Gamma^{-1\prime}F \\ &+ T^{-3}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^N(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}F. \end{aligned}$$

By using the known orders of the sample moments in F and L_0,

$$\begin{aligned}
&T^{-4}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}F \\
&= \sigma_0^2 T^{-4}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F + T^{-4}F'\Gamma^{-1}\Gamma_0 FS_\lambda F'\Gamma_0'\Gamma^{-1\prime}F \\
&= T^{-2}(T^{-1}F'F)S_\lambda(T^{-1}F'F) + T^{-1}(\rho_0-\rho)T^{-1}F'FS_\lambda T^{-2}F'L_0'F \\
&+ T^{-1}(\rho_0-\rho)T^{-2}F'L_0FS_\lambda T^{-1}F'F + (\rho_0-\rho)^2T^{-2}F'L_0FS_\lambda T^{-2}F'L_0'F + O(T^{-2}).
\end{aligned}$$

Also,

$$\begin{aligned}
&T^{-3}F'\Gamma^{-1}\Gamma_0 F\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}F \\
&= T^{-3}F'[I_T + (\rho_0-\rho)L_0]F\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'[I_T + (\rho_0-\rho)L_0]'F \\
&= N^{-1/2}T^{-5/2}(T^{-1}F'F)\frac{1}{\sqrt{NT}}\sum_{i=1}^N \lambda_i\varepsilon_i'F \\
&+ N^{-1/2}T^{-3/2}(\rho_0-\rho)(T^{-1}F'F)\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^N \lambda_i\varepsilon_i'L_0'F \\
&+ N^{-1/2}T^{-3/2}(\rho_0-\rho)(T^{-2}F'L_0F)\frac{1}{\sqrt{NT}}\sum_{i=1}^N \lambda_i\varepsilon_i'F \\
&+ (NT)^{-1/2}(\rho_0-\rho)^2(T^{-2}F'L_0F)\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^N \lambda_i\varepsilon_i'L_0'F,
\end{aligned}$$

whose order is determined by the last term on the right-hand side. By using the fact that $\varepsilon_i'L_0'FF'L_0\varepsilon_j$ is a scalar we can show that the normalized sum in this term is $O_p(1)$;

$$\begin{aligned}
&E\left(\left|\left|\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^N \lambda_i\varepsilon_i'L_0'F\right|\right|^2\right) \\
&= \frac{1}{NT^3}\sum_{i=1}^N\sum_{j=1}^N E[\mathrm{tr}(\lambda_i\varepsilon_i'L_0'FF'L_0\varepsilon_j\lambda_j)] = \frac{1}{NT^3}\sum_{i=1}^N\sum_{j=1}^N E(\varepsilon_i'L_0'FF'L_0\varepsilon_j)\lambda_i'\lambda_j \\
&= \frac{1}{NT^3}\sum_{i=1}^N\sum_{j=1}^N \mathrm{tr}[F'L_0E(\varepsilon_j\varepsilon_i')L_0'F]\lambda_i'\lambda_j = \sigma_0^2\frac{1}{NT^3}\sum_{i=1}^N \mathrm{tr}(F'L_0L_0'F)\lambda_i'\lambda_i \\
&= \sigma_0^2\mathrm{tr}(T^{-3}F'L_0L_0'F)\mathrm{tr}(S_\lambda) \leq C.
\end{aligned}$$

Hence,

$$\left|\left|T^{-3}F'\Gamma^{-1}\Gamma_0 F\frac{1}{NT}\sum_{i=1}^N \lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}F\right|\right| = O_p((NT)^{-1/2}), \tag{A133}$$

Moreover,

$$\begin{aligned}
&\left|\left|T^{-4}F'\Gamma^{-1}\Gamma_0\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}F\right|\right| \\
&= \left|\left|T^{-4}F'[I_T+(\rho_0-\rho)L_0]\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)[I_T+(\rho_0-\rho)L_0]'F\right|\right| \\
&\leq N^{-1/2}T^{-3}\left|\left|F'\frac{1}{\sqrt{N}T}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)F\right|\right| \\
&+ 2N^{-1/2}T^{-5/2}|\rho_0-\rho|\left|\left|F'\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)L_0'F\right|\right| \\
&+ N^{-1/2}T^{-1}(\rho_0-\rho)^2\left|\left|F'L_0\frac{1}{\sqrt{N}T^3}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)L_0'F\right|\right| \\
&= O_p(N^{-1/2}T^{-1}),
\end{aligned}$$

which holds because each of the three terms are $O_p(1)$, as follows from using the same steps as in Proof of Lemma 1.

Insertion of the above results into the expression for $T^{-4}F'GF$ yields

$$\begin{aligned}
T^{-4}F'GF &= T^{-2}(T^{-1}F'F)S_\lambda(T^{-1}F'F)+T^{-1}(\rho_0-\rho)T^{-1}F'FS_\lambda T^{-2}F'L_0'F \\
&+ T^{-1}(\rho_0-\rho)T^{-2}F'L_0FS_\lambda T^{-1}F'F+(\rho_0-\rho)^2T^{-2}F'L_0FS_\lambda T^{-2}F'L_0'F \\
&+ O(T^{-2})+O((NT)^{-1/2}),
\end{aligned} \tag{A134}$$

which in turn implies

$$\begin{aligned}
&T^{-3}\text{tr}\,(GFKF') \\
&= \text{tr}\,[T^{-4}F'GF(T^{-1}F'F)^{-1}]+O(T^{-3}) \\
&= T^{-2}\text{tr}\,(T^{-1}F'FS_\lambda T^{-1}F'F)+T^{-1}(\rho_0-\rho)\text{tr}\,(T^{-1}F'FS_\lambda T^{-2}F'L_0'F) \\
&+ T^{-1}(\rho_0-\rho)\text{tr}\,(T^{-2}F'L_0FS_\lambda T^{-1}F'F)+(\rho_0-\rho)^2\text{tr}\,(T^{-2}F'L_0FS_\lambda T^{-2}F'L_0'F) \\
&+ O(T^{-2})+O((NT)^{-1/2}).
\end{aligned} \tag{A135}$$

For $\text{tr}\,G$, by the same expansion as in (A73), and then the above results to evaluate the

order of the three last terms,

$$\begin{aligned} T^{-3}\text{tr}\, G &= T^{-3}\text{tr}\,(\Gamma^{-1}S_y\Gamma^{-1\prime}) \\ &= T^{-3}\text{tr}\,[\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}] + O((NT)^{-1/2}) \\ &= T^{-2}\sigma_0^2[1+(\rho_0-\rho)^2\text{tr}\,(T^{-1}L_0L_0')] + T^{-2}\text{tr}\,(T^{-1}F'FS_\lambda) \\ &+ 2T^{-1}(\rho_0-\rho)\text{tr}\,(T^{-2}F'L_0FS_\lambda) + (\rho_0-\rho)^2\text{tr}\,(T^{-3}F'L_0'L_0FS_\lambda) \\ &+ O((NT)^{-1/2}). \end{aligned} \tag{A136}$$

Note that the three first terms in this expression are actually $O(T^{-1})$. The order of $T^{-1}\log(|\hat{\Lambda}|)$ is the same as when $|\rho_0| < 1$ (see (A75)). By using this, the above results regarding $T^{-3}\text{tr}\, G$ and $T^{-3}\text{tr}\,(GFKF')$, and the same algebra as in (A76),

$$\begin{aligned} T^{-3}Q_c &= T^{-2}\log(\sigma^2) + T^{-3}\log(|\hat{\Lambda}|) + \sigma^{-2}T^{-3}\text{tr}\, G - \sigma^{-2}T^{-3}\text{tr}\,(GFKF') \\ &= \sigma^{-2}T^{-3}\text{tr}\, G - \sigma^{-2}T^{-3}\text{tr}\,(GFKF') + O_p(T^{-2}) \\ &= T^{-2}[\log(\sigma^2) + \sigma^{-2}\sigma_0^2 + \sigma^{-2}\sigma_0^2(\rho_0-\rho)^2\omega_1^2] \\ &+ O_p((NT)^{-1/2}) + O_p(T^{-2}), \end{aligned} \tag{A137}$$

which in turn implies

$$\begin{aligned} N^{-1}T^{-3}\ell_c &= -\frac{1}{2T^3}Q_c \\ &= -\frac{1}{2T^2}\left(\log(\sigma^2) + \frac{\sigma_0^2}{\sigma^2}\right) - \frac{\sigma_0^2}{2\sigma^2}(\rho_0-\rho)^2T^{-2}\omega_1^2 \\ &+ O_p(T^{-2}) + O_p((NT)^{-1/2}), \end{aligned} \tag{A138}$$

where $T^{-2}\omega_1^2 \geq 0$. ■

Lemma C.3. *Under C1, $\rho_0 = 1$, and Assumptions EPS, F and LAM, as $T \to \infty$ for any N, including $N \to \infty$, provided that $\sqrt{N}T^{-3/2} \to 0$,*

$$H_{3/2}^{-1}\frac{\partial \ell_c(\theta_2^0)}{\partial \theta_2} \sim N\left(0_{2\times 1}, \begin{bmatrix} T^{-2}\omega_1^2 & 0 \\ 0 & \frac{(\kappa_0-1)}{4\sigma_0^4} \end{bmatrix}\right).$$

Proof of Lemma C.3.

This proof is almost identical to that of Lemma C.1, and hence only essential details are given. All variable definitions are the same as in Proof of Lemma C.1. Because of the change of normalization of $\partial\ell_c(\theta_2^0)/\partial\rho$ from $(NT)^{-1/2}$ to $N^{-1/2}T^{-3/2}$ the relevant quantity is no

longer Q but $T^{-1}Q$. Consider $T^{-1}Q_1$. As in Proof of Lemma C.1, the mean of this quantity is zero. For the variance, from Proof of Lemma C.1,

$$E[(T^{-1}Q_{12}^2)] = \sigma_0^2 \text{tr}\left[S_\lambda T^{-3} F'(L_0'L_0 + 2L_0L_0 + L_0L_0')F\right] = T^{-2}\Sigma_{12},$$

and therefore

$$T^{-1}Q_{12} \sim N(0, T^{-2}\Sigma_{12}), \tag{A139}$$

which holds for $T \to \infty$ and any N, including $N \to \infty$, as in Lemma C.1. $T^{-1/2}Q_{13}$ is mean zero too, and with variance

$$T^{-2}E[(\varepsilon_i' L_0 \varepsilon_i)^2] = \sigma_0^4 T^{-2} \text{tr}\,(L_0 L_0') = T^{-1}\Sigma_{13} \leq C,$$

suggesting that $T^{-1/2}Q_{13} = O_p(1)$. Hence,

$$T^{-1}(Q_1 - Q_{11}) = T^{-1}(Q_{12} + Q_{13}) = T^{-1}Q_{12} + O_p(T^{-1/2}) \sim N(0, T^{-2}\Sigma_{12}). \tag{A140}$$

$T^{-1}Q_2$ requires more work. TK_0 is the same as before; $TK_0 = (T^{-1}F'F)^{-1} + O(T^{-1})$. Therefore, all remainder terms that are driven by this result have the same order as in Proof of Lemma C.1. This implies

$$T^{-1}Q_{21} = T^{-1}Q_{11} - \sqrt{N}T^{-3/2}\sigma_0^2 \text{tr}\,(P_F L_0) + O(\sqrt{N}T^{-5/2}) + O_p(T^{-2}). \tag{A141}$$

Consider $T^{-1}Q_{22}$, whose variance is given by

$$\begin{aligned}
&T^{-3}E[(\varepsilon_i' F K_0 F' L_0 F \lambda_i + \lambda_i' F' F K_0 F' L_0 \varepsilon_i)^2] \\
&= \sigma_0^2 \lambda_i T^{-2} F' L_0' F (TK_0) T^{-1} F' F (TK_0) T^{-2} F' L_0 F \lambda_i \\
&+ 2\sigma_0^2 \lambda_i' T^{-1} F' F (TK_0) T^{-2} F' L_0 F (TK_0) T^{-2} F' L_0 F \lambda_i \\
&+ \sigma_0^2 \lambda_i' T^{-1} F' F (TK_0) T^{-3} F' L_0 L_0' F (TK_0) T^{-1} F' F \lambda_i \\
&= \sigma_0^2 \lambda_i' T^{-2} F' L_0' F (T^{-1}F'F)^{-1} T^{-2} F' L_0 F \lambda_i + 2\sigma_0^2 \lambda_i' T^{-2} F' L_0 F (T^{-1}F'F)^{-1} T^{-2} F' L_0 F \lambda_i \\
&+ \sigma_0^2 \lambda_i' T^{-3} F' L_0 L_0' F \lambda_i + O(T^{-1}).
\end{aligned}$$

Hence, noting that $T^{-2}\Sigma_{22} = \sigma_0^2 \text{tr}\left[S_\lambda T^{-3} F'(L_0' P_F L_0 + 2L_0 P_F L_0 + L_0 L_0')F\right]$, we have

$$T^{-1}Q_{22} = \frac{1}{\sqrt{N}T^{3/2}} \sum_{i=1}^{N} (\varepsilon_i' F K_0 F' L_0 F \lambda_i + \lambda_i' F' F K_0 F' L_0 \varepsilon_i) \sim N(0, T^{-2}\Sigma_{22}) \tag{A142}$$

as N, $T \to \infty$. The reason for the large-N requirement here is that $L_0\varepsilon_i$ is a random walk, suggesting a nonstandard limiting distribution for $T^{-1}Q_{22}$ as $T \to \infty$ with N fixed.

$T^{-1}Q_{23}$ can be expanded in the same was as in Proof of Lemma C.1;

$$\begin{aligned} T^{-1}Q_{23} &= \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\text{tr}\left(K_0F'L_0\varepsilon_i\varepsilon_i'F\right) \\ &= \sigma_0^2\sqrt{N}T^{-3/2}\text{tr}\left(K_0F'L_0F\right)+\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\text{tr}\left[K_0(F'L_0\varepsilon_i\varepsilon_i'F-\sigma_0^2F'L_0F)\right] \\ &= \sqrt{N}T^{-3/2}\sigma_0^2\text{tr}\left[(T^{-1}F'F)^{-1}T^{-1}F'L_0F\right] \\ &+ T^{-1/2}\text{tr}\left(TK_0\frac{1}{\sqrt{N}T^2}\sum_{i=1}^{N}(F'L_0\varepsilon_i\varepsilon_i'F-\sigma_0^2F'L_0F)\right)+O_p(\sqrt{N}T^{-3/2}) \\ &= \sqrt{N}T^{-3/2}\sigma_0^2\text{tr}\left(P_FL_0\right)+O_p(T^{-1/2})+O_p(\sqrt{N}T^{-3/2}). \end{aligned} \tag{A143}$$

Hence,

$$\begin{aligned} &T^{-1}(Q_{21}+Q_{23}) \\ &= T^{-1}Q_{11}-\sqrt{N}T^{-3/2}\sigma_0^2\text{tr}\left(P_FL_0\right)+\sqrt{N}T^{-3/2}\sigma_0^2\text{tr}\left(P_FL_0\right)+O_p(T^{-1/2})+O_p(\sqrt{N}T^{-3/2}) \\ &= T^{-1}Q_{11}+O_p(T^{-1/2})+O_p(\sqrt{N}T^{-3/2}), \end{aligned} \tag{A144}$$

which in turn implies, provided that $\sqrt{N}T^{-3/2}\to 0$ as $N,\ T\to\infty$,

$$\begin{aligned} T^{-1}Q_2 &= T^{-1}(Q_{21}+Q_{22}+Q_{23}) = T^{-1}(Q_{11}+Q_{22})+O_p(T^{-1/2})+O_p(\sqrt{N}T^{-3/2}) \\ &\sim T^{-1}Q_{11}+N(0,T^{-2}\Sigma_{22}). \end{aligned} \tag{A145}$$

Except for the rescaling by T the correlation between $T^{-1}Q_1$ and $T^{-1}Q_2$ is the same as in Proof of Lemma C.1, that is,

$$\begin{aligned} &T^{-2}E[(Q_1-Q_{11})(Q_2-Q_{11})] \\ &= \sigma_0^2\text{tr}\left[S_\lambda T^{-3}F'(L_0'P_FL_0+L_0P_FL_0+L_0'L_0'P_F+L_0L_0'P_F)F\right]+o(1), \end{aligned} \tag{A146}$$

The above results, together with the same algebra used in Proof of Lemma C.1,

$$\begin{aligned} &T^{-2}E[(Q_1-Q_2)^2] \\ &= T^{-2}E[((Q_1-Q_{11})-(Q_2-Q_{11}))^2] \\ &= T^{-2}E[(Q_1-Q_{11})^2]+T^{-2}E[(Q_2-Q_{11})^2]-2T^{-2}E[(Q_1-Q_{11})(Q_2-Q_{11})] \\ &= T^{-2}\Sigma_{12}+T^{-2}\Sigma_{13}+T^{-2}\Sigma_{22}-2T^{-2}E[(Q_1-Q_{11})(Q_2-Q_{11})]+o(1) \\ &= \sigma_0^4T^{-2}\omega_1^2+o(1), \end{aligned} \tag{A147}$$

which is just T^{-2} times the result obtained in Lemma C.1. Note also that the effect of $T^{-2}\Sigma_{13}$ in $T^{-2}\omega_1^2$ is $O(T^{-1})$. We consequently obtain

$$\begin{aligned} T^{-1}(Q_1 - Q_2) &= T^{-1}[Q_{11} + Q_{12} + Q_{13} - (Q_{11} + Q_{22})] + O_p(T^{-1/2}) + O_p(\sqrt{N}T^{-3/2}) \\ &= T^{-1}(Q_{12} + Q_{13} - Q_{22}) + O_p(T^{-1/2}) + O_p(\sqrt{N}T^{-3/2}) \\ &\sim N(0, \sigma_0^4 T^{-2}\omega_1^2) \end{aligned} \tag{A148}$$

as N, $T \to \infty$ with $\sqrt{N}T^{-3/2} \to 0$. Insertion into $N^{-1/2}T^{-3/2}\partial\ell_c(\theta_2^0)/\partial\rho$ gives

$$\frac{1}{\sqrt{N}T^{3/2}}\frac{\partial\ell_c(\theta_2^0)}{\partial\rho} = \sigma_0^{-2}T^{-1}(Q_1 - Q_2) \sim N(0, T^{-2}\omega_1^2). \tag{A149}$$

This establishes the desired result for $\partial\ell_c(\theta_2^0)/\partial\rho$.

The result for $\partial\ell_c(\theta_2^0)/\partial\rho$ is a consequence of the fact that $\partial\ell_c(\theta_2^0)/\partial\rho$ does not depend on L_0. Moreover, by using exactly the same calculations as in Proof of Lemma C.1 we can show that the expected value of the normalized cross-derivative is zero. ■

Lemma C.4. *Under C1, $\rho_0 = 1$, and Assumptions EPS, F and LAM, as $T \to \infty$ for any N, including $N \to \infty$,*

$$-H_{3/2}^{-1}\frac{\partial^2\ell_c(\theta_2^0)}{\partial\theta_2(\partial\theta_2)'}H_{3/2}^{-1} = \begin{bmatrix} T^{-2}\omega_1^2 & 0 \\ 0 & \frac{1}{2\sigma_0^2} \end{bmatrix} + O_p(T^{-1}) + O_p((NT)^{-1/2}).$$

Proof of Lemma C.4.

This proof is analogous to that of Lemma C.2. The only change is the rescaling by T, the effect of which can be traced out following the steps as in Proof of Lemma C.3. ■

Proof of Theorem 2.

In view of Lemmas C.3 and C.4 the desired result follows by the same line of argumentation used in Proof of Theorem 1. ■

Proof of Lemma 3.

From the first-order condition with respect to S_λ we obtain the following slightly modified expression for $\hat{\Lambda}(\theta_2)$: $\hat{\Lambda}(\theta_2) = I_T + \sigma^{-2}\Gamma(\rho)^{-1}F\hat{S}_\lambda(\theta_2)F'\Gamma(\rho)^{-1\prime}$. Letting $K = K(\theta_2) = (\sigma^2\hat{S}_\lambda(\theta_2)^{-1} + F'\Gamma(\rho)^{-1\prime}\Gamma(\rho)^{-1}F)^{-1}$, this implies

$$\hat{\Lambda}^{-1} = I_T - \Gamma^{-1}F(\sigma^2\hat{S}_\lambda^{-1} + F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime} = I_T - \Gamma^{-1}FKF'\Gamma^{-1\prime},$$

and therefore

$$Q_c = T\log(\sigma^2) + \log(|\hat{\Lambda}|) + \sigma^{-2}\text{tr}\, G - \sigma^{-2}\text{tr}\,(G\Gamma^{-1}FKF'\Gamma^{-1\prime}),$$

where $G = G(\rho) = \Gamma(\rho)^{-1}S_y\Gamma(\rho)^{-1\prime}$ is as before.

Consider $\sigma^{-2}\text{tr}\,(G\Gamma^{-1}FKF'\Gamma^{-1\prime})$. As in Proof of Lemma 1,

$$\begin{aligned}\hat{S}_\lambda &= \sigma^2(\Gamma^{-1}F)^-(\sigma^{-2}G - I_T)(\Gamma^{-1}F)^{-\prime} \\ &= (F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma^2(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}, \qquad \text{(A150)}\end{aligned}$$

which in turn implies

$$K = (\sigma^2\hat{S}_\lambda^{-1} + F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} = (F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}, \qquad \text{(A151)}$$

suggesting that

$$\begin{aligned}\text{tr}\,(G\Gamma^{-1}FKF'\Gamma^{-1\prime}) &= \text{tr}\,(F'\Gamma^{-1\prime}G\Gamma^{-1}FK) \\ &= \text{tr}\,[F'\Gamma^{-1\prime}G\Gamma^{-1}F((F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1})] \\ &= \text{tr}\,[F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}] - \sigma^2\text{tr}\, I_m \\ &= \text{tr}\,[F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}] - \sigma^2 m. \qquad \text{(A152)}\end{aligned}$$

Consider $F'\Gamma^{-1\prime}G\Gamma^{-1}F = F'\Gamma^{-1\prime}\Gamma^{-1}S_y\Gamma^{-1\prime}\Gamma^{-1}F$, where S_y can be expanded in the usual fashion as

$$\begin{aligned}S_y &= \Gamma_0 S_u\Gamma_0' = \Gamma_0\frac{1}{N}\sum_{i=1}^N(\Gamma_0^{-1}F\lambda_i + \varepsilon_i)(\Gamma_0^{-1}F\lambda_i + \varepsilon_i)'\Gamma_0' \\ &= \sigma_0^2\Gamma_0\Gamma_0' + FS_\lambda F' + \frac{1}{N}\sum_{i=1}^N F\lambda_i\varepsilon_i'\Gamma_0' + \Gamma_0\frac{1}{N}\sum_{i=1}^N\varepsilon_i\lambda_i'F' \\ &+ \Gamma_0\frac{1}{N}\sum_{i=1}^N(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)\Gamma_0'. \qquad \text{(A153)}\end{aligned}$$

Hence,

$$\begin{aligned}&T^{-3}F'\Gamma^{-1\prime}G\Gamma^{-1}F \\ &= T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}S_y\Gamma^{-1\prime}\Gamma^{-1}F \\ &= \sigma_0^2T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}\Gamma^{-1}F + T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}FS_\lambda F'\Gamma^{-1\prime}\Gamma^{-1}F \\ &+ T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}\frac{1}{N}\sum_{i=1}^N F\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}\Gamma^{-1}F + T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}\Gamma_0\frac{1}{N}\sum_{i=1}^N\varepsilon_i\lambda_i'F'\Gamma^{-1\prime}\Gamma^{-1}F \\ &+ T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}\Gamma_0\frac{1}{N}\sum_{i=1}^N(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}\Gamma^{-1}F, \qquad \text{(A154)}\end{aligned}$$

where

$$
\begin{aligned}
&T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}\Gamma^{-1}F \\
&= T^{-3}F'\Gamma^{-1\prime}[I_T+(\rho_0-\rho)L_0][I_T+(\rho_0-\rho)L_0]'\Gamma^{-1}F \\
&= T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}F+(\rho_0-\rho)T^{-3}F'\Gamma^{-1\prime}(L_0+L_0')\Gamma^{-1}F+(\rho_0-\rho)^2T^{-3}F'\Gamma^{-1\prime}L_0L_0'\Gamma^{-1}F.
\end{aligned}
$$

Consider the T-rowed matrix $A=(A_1,...,A_T)'$. If $\rho=\rho_0=1$, $\Gamma^{-1}A=(A_1,\Delta A_2...,\Delta A_T)'$, whereas if $\rho=\rho_0=0$, then $\Gamma^{-1}A=A$. This suggests that the (norm of) above sample moments are minimized for $\rho=\rho_0=1$ and maximized for $\rho=0$, in which case the orders are the same as in Proof of Lemma 2. Hence, $||T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F||$, $||T^{-2}F'\Gamma^{-1\prime}(L_0+L_0')\Gamma^{-1}F||$ and $||T^{-3}F'\Gamma^{-1\prime}L_0L_0'\Gamma^{-1}F||$ are all $O(1)$, suggesting that

$$
T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}\Gamma^{-1}F=(\rho_0-\rho)^2T^{-3}F'\Gamma^{-1\prime}L_0L_0'\Gamma^{-1}F+O(T^{-1}), \tag{A155}
$$

and, by the same argument,

$$
||T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}FS_\lambda F'\Gamma^{-1\prime}\Gamma^{-1}F||\leq T^{-1}||T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F||^2||S_\lambda||=O(T^{-1}). \tag{A156}
$$

Setting again $\rho=0$, we can also show that

$$
\begin{aligned}
&\left|\left|T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}F\frac{1}{N}\sum_{i=1}^N\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}\Gamma^{-1}F\right|\right| \\
&\leq ||T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F||\left|\left|\frac{1}{NT^2}\sum_{i=1}^N\lambda_i\varepsilon_i'[I_T+(\rho_0-\rho)L_0]'\Gamma^{-1}F\right|\right| \\
&\leq N^{-1/2}T^{-3/2}||T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F||\left|\left|\frac{1}{\sqrt{NT}}\sum_{i=1}^N\lambda_i\varepsilon_i'\Gamma^{-1}F\right|\right| \\
&+ (NT)^{-1/2}|\rho_0-\rho|||T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F||\left|\left|\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^N\lambda_i\varepsilon_i'L_0'\Gamma^{-1}F\right|\right| \\
&= O_p((NT)^{-1/2}),
\end{aligned} \tag{A157}
$$

and

$$\begin{aligned}
&\left|\left|T^{-3}F'\Gamma^{-1\prime}\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}\Gamma^{-1}F\right|\right| \\
&= \left|\left|T^{-3}F'\Gamma^{-1\prime}[I_T+(\rho_0-\rho)L_0]\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)[I_T+(\rho_0-\rho)L_0]'\Gamma^{-1}F\right|\right| \\
&\leq N^{-1/2}T^{-2}\left|\left|F'\Gamma^{-1\prime}\frac{1}{\sqrt{N}T}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)\Gamma^{-1}F\right|\right| \\
&+ 2N^{-1/2}T^{-3/2}|\rho_0-\rho|\left|\left|F'\Gamma^{-1\prime}\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)L_0'\Gamma^{-1}F\right|\right| \\
&+ N^{-1/2}|\rho_0-\rho|^2\left|\left|F'\Gamma^{-1\prime}L_0\frac{1}{\sqrt{N}T^3}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)L_0'\Gamma^{-1}F\right|\right| = O_p(N^{-1/2}). \quad \text{(A158)}
\end{aligned}$$

It follows that

$$T^{-3}F'\Gamma^{-1\prime}G\Gamma^{-1}F = \sigma_0^2(\rho_0-\rho)^2T^{-3}F'\Gamma^{-1\prime}L_0L_0'\Gamma^{-1}F + O(T^{-1}) + O_p(N^{-1/2}), \quad \text{(A159)}$$

which can be substituted back into $\text{tr}\,(G\Gamma^{-1}FKF'\Gamma^{-1\prime})$, giving

$$\begin{aligned}
&T^{-2}\text{tr}\,(G\Gamma^{-1}FKF'\Gamma^{-1\prime}) \\
&= \text{tr}\,[T^{-3}F'\Gamma^{-1\prime}G\Gamma^{-1}F(T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}] - \sigma^2T^{-2}m \\
&= \sigma_0^2(\rho_0-\rho)^2\text{tr}\,[T^{-3}F'\Gamma^{-1\prime}L_0L_0'\Gamma^{-1}F(T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}] + O(T^{-1}) + O_p(N^{-1/2}). \quad \text{(A160)}
\end{aligned}$$

Next, consider $\text{tr}\,G$. Since $T^{-2}\text{tr}\,(\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}) = (\rho_0-\rho)^2T^{-2}\text{tr}\,(L_0L_0') + O(T^{-1})$,

$$\begin{aligned}
T^{-2}\text{tr}\,G &= T^{-2}\text{tr}\,(\Gamma^{-1}S_y\Gamma^{-1\prime}) \\
&= T^{-2}\sigma_0^2\text{tr}\,(\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}) + T^{-1}\text{tr}\,(T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}FS_\lambda) \\
&+ 2(NT)^{-1/2}\text{tr}\left(\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\Gamma^{-1}F\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}\right) \\
&+ N^{-1/2}\text{tr}\left(\frac{1}{\sqrt{N}T^2}\sum_{i=1}^{N}\Gamma^{-1}\Gamma_0(\varepsilon_i\varepsilon_i'-\sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}\right) \\
&= \sigma_0^2(\rho_0-\rho)^2T^{-2}\text{tr}\,(L_0L_0') + O_p(T^{-1}) \\
&+ O_p(N^{-1/2}). \quad \text{(A161)}
\end{aligned}$$

Hence, by adding the results,

$$\begin{aligned}
&T^{-2}Q_c \\
&= T^{-1}\log(\sigma^2) + T^{-2}\log(|\hat{\Lambda}|) + \sigma^{-2}T^{-2}\text{tr}\, G - \sigma^{-2}T^{-2}\text{tr}\,(G\Gamma^{-1}FKF'\Gamma^{-1\prime}) \\
&= \sigma^{-2}T^{-2}\text{tr}\, G - \sigma^{-2}T^{-2}\text{tr}\,(F'\Gamma^{-1\prime}G\Gamma^{-1}FK) + O_p(T^{-1}) \\
&= \sigma^{-2}\sigma_0^2(\rho_0-\rho)^2T^{-2}\text{tr}\,(L_0L_0') \\
&- \sigma^{-2}\sigma_0^2(\rho_0-\rho)^2\text{tr}\,[T^{-3}F'\Gamma^{-1\prime}L_0L_0'\Gamma^{-1}F(T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}] + O_p(T^{-1}) + O_p(N^{-1/2}) \\
&= \sigma^{-2}\sigma_0^2(\rho_0-\rho)^2T^{-2}\text{tr}\,(L_0'M_{\Gamma^{-1}F}L_0) + O_p(T^{-1}) + O_p(N^{-1/2}),
\end{aligned}$$

with an obvious definition of $M_{\Gamma^{-1}F}$. In particular, note how for any $T \times m$ matrix $A = (A_1, ..., A_T)'$,

$$L_0'\Gamma^{-1}A = L_0'\begin{bmatrix} A_1' \\ (\Delta A_2)' \\ \vdots \\ (\Delta A_T)' \end{bmatrix} = \begin{bmatrix} A_T' - A_1' \\ \vdots \\ A_T' - A_{T-1}' \\ 0_{m\times 1}' \end{bmatrix} = 1_T A_T' - A.$$

Hence, since $||T^{-1}(1_TF_T' - F)'(1_TF_T' - F)|| \leq 2||T^{-1}F_T1_T'1_TF_T'|| + 2||T^{-1}F'F|| = O(1)$, we can show that

$$\begin{aligned}
&T^{-2}\text{tr}\,(L_0'M_{\Gamma^{-1}F}L_0) \\
&= T^{-2}\text{tr}\,(L_0L_0') - T^{-2}\text{tr}\,[L_0'\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0] \\
&= T^{-2}\text{tr}\,(L_0L_0') - T^{-2}\text{tr}\,[T^{-1}(1_TF_T' - F)'(1_TF_T' - F)(T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}] \\
&= T^{-2}\text{tr}\,(L_0L_0') + O(T^{-2})
\end{aligned}$$

It follows that

$$T^{-2}Q_c = \sigma^{-2}\sigma_0^2(\rho_0-\rho)^2T^{-2}\text{tr}\,(L_0L_0') + O_p(T^{-1}) + O_p(N^{-1/2}),$$

which in turn implies

$$N^{-1}T^{-2}\ell_c = -\frac{1}{2T^2}Q_c = -\frac{\sigma_0^2}{2\sigma^2}(\rho_0-\rho)^2T^{-2}\text{tr}\,(L_0L_0') + O_p(N^{-1/2}) + O_p(T^{-1}), \quad \text{(A162)}$$

as was to be shown. ■

Lemma C.5. *Under C2, $\rho_0 = 1$, and Assumptions EPS, F and LAM, as $N, T \to \infty$ with $\sqrt{N}T^{-1} \to 0$,*

$$H_1^{-1}\frac{\partial \ell_c(\theta_2^0)}{\partial \theta_2} \sim N\left(0_{2\times 1}, \begin{bmatrix} T^{-1}\omega_2^2 & 0 \\ 0 & \frac{(\kappa_0-1)}{4\sigma_0^4} \end{bmatrix}\right).$$

Proof of Lemma C.5.

In this proof we set $\theta_2 = \theta_2^0$. Therefore, to simplify notation, functions such as $\Gamma(\rho_0)$ and $\overline{C}(\theta_2^0)$ will be written Γ and $\overline{C}$, respectively. Analogous to Proof of Lemma C.1 and using the results provided in Appendix B, we may write

$$\frac{2\sigma_0^2}{\sqrt{NT}}\frac{\partial \ell_c}{\partial \rho} = \sqrt{N}T^{-1}\text{tr}\,(B_2'\bar{c} - B_1'c_0), \tag{A163}$$

where

$$\begin{aligned}
B_1 &= \hat{\Lambda}^{-1}, \\
B_2 &= \sigma_0^{-2}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1} - \hat{\Lambda}^{-1}, \\
\bar{c} &= -P_{\Gamma^{-1}F}(c_0 + c_1 + c_2)P_{\Gamma^{-1}F} + c_3, \\
c_0 &= -(L_0S_u + S_uL_0'), \\
c_1 &= (S_u - \sigma_0^2 I_T)P_{\Gamma^{-1}F}L_0' - (S_u - \sigma_0^2 I_T)M_{\Gamma^{-1}F}L_0, \\
c_2 &= L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T) - L_0'M_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T), \\
c_3 &= L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)P_{\Gamma^{-1}F} + P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)P_{\Gamma^{-1}F}L_0'.
\end{aligned}$$

where $M_A = I_T - A(A'A)^{-1}A' = I_T - P_A$ for any T-rowed matrix A, $G = \Gamma^{-1}S_y\Gamma^{-1\prime}$, $\hat{\Lambda} = I_T + \sigma_0^{-2}\Gamma^{-1}F\hat{S}_\lambda F'\Gamma^{-1\prime}$ and $\hat{S}_\lambda = \sigma_0^2(\Gamma^{-1}F)^-(\sigma_0^{-2}G - I_T)(\Gamma^{-1}F)^{-\prime}$. Here $c_1, ..., c_3$ are as in Appendix B with $\theta_2 = \theta_2^0$ imposed, which implies $J\Gamma = L_0$ and $G = \Gamma^{-1}S_y\Gamma^{-1\prime} = S_u$.

Consider $-\sqrt{N}T^{-1}\text{tr}\,(B_1'c_0)$, the second term on the right-hand side of (A163). Since $B_1 = \hat{\Lambda}^{-1} = I_T - \Gamma^{-1}FKF'\Gamma^{-1\prime}$ (see Proof of Lemma 1) and $c_0 = -(L_0S_u + S_uL_0')$ are both symmetric,

$$\begin{aligned}
-\sqrt{N}T^{-1}\text{tr}\,(B_1'c_0) &= \sqrt{N}T^{-1}\text{tr}\,[\hat{\Lambda}^{-1}(L_0S_u + S_uL_0')] = 2\sqrt{N}T^{-1}\text{tr}\,(\hat{\Lambda}^{-1}L_0S_u) \\
&= 2\sqrt{N}T^{-1}\text{tr}\,[(I_T - \Gamma^{-1}FKF'\Gamma^{-1\prime})L_0S_u] \\
&= 2\sqrt{N}T^{-1}\text{tr}\,(L_0S_u - KF'\Gamma^{-1\prime}L_0S_u\Gamma^{-1}F) \\
&= 2T^{-1/2}(Q_1 - Q_2),
\end{aligned} \tag{A164}$$

where Q_1 is the same as in Proof of Lemma C.1. The only difference is that F in Q_{11}, Q_{12} and Q_{13} is premultiplied by Γ^{-1}. The asymptotic distribution of $T^{-1/2}Q_1$ under $\rho_0 = 1$ can be obtained by using exactly the same steps as in Proof of Lemma 2, and can be shown to be

$$T^{-1/2}(Q_1 - Q_{11}) \sim N(0, T^{-1}(\Sigma_{12} + \Sigma_{13})), \tag{A165}$$

where

$$\begin{aligned}
Q_{11} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i'F'\Gamma^{-1\prime}L_0\Gamma^{-1}F\lambda_i, \\
\Sigma_{13} &= \sigma_0^4 T^{-1}\text{tr}\,(L_0L_0'), \\
\Sigma_{12} &= \sigma_0^2\text{tr}\,[S_\lambda T^{-1}F'\Gamma^{-1\prime}(L_0'L_0+2L_0L_0+L_0L_0')\Gamma^{-1}F].
\end{aligned}$$

This result, which requires N, $T \to \infty$, is similar to the one given in Proof of Lemma C.4, except that now the effect of Q_{13} is no longer negligible.

The analysis of Q_2 differs from before. Note in particular how

$$\begin{aligned}
T^{-1/2}Q_2 &= \sqrt{N}T^{-1}\text{tr}\,(KF'\Gamma^{-1\prime}L_0S_u\Gamma^{-1}F) \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}[\text{tr}\,(\lambda_i'F'\Gamma^{-1\prime}\Gamma^{-1}FKF'\Gamma^{-1\prime}L_0\Gamma^{-1}F\lambda_i)+\text{tr}\,(\varepsilon_i'\Gamma^{-1}FKF'\Gamma^{-1\prime}L_0\Gamma^{-1}F\lambda_i) \\
&+ \text{tr}\,(\lambda_i'F'\Gamma^{-1\prime}\Gamma^{-1}FKF'\Gamma^{-1\prime}L_0\varepsilon_i)+\text{tr}\,(KF'\Gamma^{-1\prime}L_0\varepsilon_i\varepsilon_i'\Gamma^{-1}F)] \\
&= T^{-1/2}(Q_{21}+Q_{22}+Q_{23}), \qquad \text{(A166)}
\end{aligned}$$

with

$$\begin{aligned}
Q_{21} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\lambda_i'F'\Gamma^{-1\prime}\Gamma^{-1}FKF'\Gamma^{-1\prime}L_0\Gamma^{-1}F\lambda_i, \\
Q_{22} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}(\varepsilon_i'\Gamma^{-1}FKF'\Gamma^{-1\prime}L_0\Gamma^{-1}F\lambda_i+\lambda_i'F'\Gamma^{-1\prime}\Gamma^{-1}FKF'\Gamma^{-1\prime}L_0\varepsilon_i), \\
Q_{23} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\text{tr}\,(KF'\Gamma^{-1\prime}L_0\varepsilon_i\varepsilon_i'\Gamma^{-1}F),
\end{aligned}$$

which are the same as in Proof of Lemma C.1, except for the premultiplication of F by Γ^{-1}. From Proof of Lemma 3, under $\rho = \rho_0$, $||T^{-3}F'\Gamma^{-1\prime}G\Gamma^{-1}F|| = O(T^{-1}) + O_p(N^{-1/2})$, and therefore, by Taylor expanding the inverse,

$$\begin{aligned}
TK &= (T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma_0^2T^{-2}(T^{-3}F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1} \\
&= (T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} + O(T^{-3}) + O_p(N^{-1/2}T^{-2}). \qquad \text{(A167)}
\end{aligned}$$

Substitution into the expression for Q_{21} yields

$$\begin{aligned}
T^{-1/2}Q_{21} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i'T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F(TK)T^{-1}F'\Gamma^{-1\prime}L_0\Gamma^{-1}F\lambda_i \\
&= Q_{11} + O(\sqrt{N}T^{-3}) + O_p(T^{-2}). \qquad \text{(A168)}
\end{aligned}$$

The last equality makes use of the fact that $L_0\Gamma^{-1} = J$, and therefore $T^{-1}F'\Gamma^{-1\prime}L_0\Gamma^{-1}F = T^{-1}F'\Gamma^{-1\prime}JF = T^{-1}\sum_{t=2}^{T}\Delta F_tF_{t-1}' = O(1)$.

The analysis of Q_{22} is similar to that in Proof of Lemma C.1, as is the end result;

$$T^{-1/2}Q_{22} \sim N(0, T^{-1}\Sigma_{22}),$$

where

$$\Sigma_{22} = \sigma_0^2 \text{tr}\left[S_\lambda T^{-1} F'\Gamma^{-1\prime}(L_0' P_{\Gamma^{-1}F} L_0 + 2L_0 P_{\Gamma^{-1}F} L_0 + L_0 L_0')\Gamma^{-1}F\right],$$

which requires $T \to \infty$ with N fixed or $N \to \infty$.

For Q_{23},

$$\begin{aligned}
T^{-1/2}Q_{23} &= \frac{1}{\sqrt{N}T}\sum_{i=1}^{N} \text{tr}\left(KF'\Gamma^{-1\prime}L_0\varepsilon_i\varepsilon_i'\Gamma^{-1}F\right) \\
&= \sigma_0^2\sqrt{N}T^{-1/2}\text{tr}\left(KF'\Gamma^{-1\prime}L_0\Gamma^{-1}F\right) \\
&+ \frac{1}{\sqrt{N}T}\sum_{i=1}^{N} \text{tr}\left[K(F'\Gamma^{-1\prime}L_0\varepsilon_i\varepsilon_i'\Gamma^{-1}F - \sigma_0^2 F'\Gamma^{-1\prime}L_0\Gamma^{-1}F)\right] \\
&= \sqrt{N}T^{-1}\sigma_0^2\text{tr}\left[(T^{-1}F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}T^{-1}F'\Gamma^{-1\prime}L_0\Gamma^{-1}F\right] \\
&+ T^{-1/2}\text{tr}\left(TK\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}(F'\Gamma^{-1\prime}L_0\varepsilon_i\varepsilon_i'\Gamma^{-1}F - \sigma_0^2 F'\Gamma^{-1\prime}L_0\Gamma^{-1}F)\right) \\
&+ O_p(\sqrt{N}T^{-4}) + O_p(T^{-3}) = O_p(\sqrt{N}T^{-1}) + O_p(T^{-1/2}),
\end{aligned} \tag{A169}$$

which is $o_p(1)$, provided that $\sqrt{N}T^{-1} = o(1)$. Hence, if we assume that $T \to \infty$ with N fixed or $N \to \infty$ such that $\sqrt{N}T^{-1} = o(1)$, then

$$\begin{aligned}
Q_2 &= Q_{21} + Q_{22} + Q_{23} = Q_{11} + Q_{22} + O_p(T^{-1/2}) + O_p(\sqrt{N}T^{-1}) \\
&\sim Q_{11} + N(0, \Sigma_{22}).
\end{aligned} \tag{A170}$$

The correlation between $T^{-1/2}Q_1$ and $T^{-1/2}Q_2$ is the same as in Proof of Lemma C.1, with F replaced by $\Gamma^{-1}F$;

$$\begin{aligned}
&T^{-1}E[(Q_1 - Q_{11})(Q_2 - Q_{11})] \\
&= \sigma_0^2\text{tr}\left[S_\lambda T^{-2}F'\Gamma^{-1\prime}(L_0' P_{\Gamma^{-1}F}L_0 + L_0 P_{\Gamma^{-1}F}L_0 + L_0'L_0' P_{\Gamma^{-1}F} + L_0 L_0' P_{\Gamma^{-1}F})\Gamma^{-1}F\right] \\
&+ o(1),
\end{aligned} \tag{A171}$$

which we can use to show that

$$\begin{aligned}
T^{-1}E[(Q_1 - Q_2)^2] &= T^{-2}\text{tr}\left(\sigma_0^4 L_0 L_0' + \sigma_0^2 S_\lambda F'\Gamma^{-1\prime}L_0' M_{\Gamma^{-1}F}L_0\Gamma^{-1}F\right) + o(1) \\
&= \sigma_0^4 T^{-1}\omega_2^2 + o(1).
\end{aligned} \tag{A172}$$

where $\omega_2^2 = T^{-1}\text{tr}\,(L_0 L_0' + \sigma_0^{-2} S_\lambda F'\Gamma^{-1\prime} L_0' M_{\Gamma^{-1}F} L_0 \Gamma^{-1} F)$. Hence,

$$-\frac{1}{2}\sqrt{N}T^{-1}\text{tr}\,(B_1' c_0) = T^{-1/2}(Q_1 - Q_2) \sim N(0, \sigma_0^4 T^{-1}\omega_2^2), \tag{A173}$$

as $N,\ T \to \infty$ with $\sqrt{N}T^{-1} = o(1)$.

Let us now consider $\sqrt{N}T^{-1}\text{tr}\,(B_2'\bar{c})$, the first term in (A163). From $\hat{\Lambda}^{-1} = I_T - \Gamma^{-1}FKF'\Gamma^{-1\prime}$, $K = (F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma_0^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}$, and the idempotency of $P_{\Gamma^{-1}F}$,

$$\begin{aligned}
&P_{\Gamma^{-1}F}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1}P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}(I_T - \Gamma^{-1}FKF'\Gamma^{-1\prime})G(I_T - \Gamma^{-1}FKF'\Gamma^{-1\prime})P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F} - \sigma_0^{-2}P_{\Gamma^{-1}F}G\Gamma^{-1}FKF'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&- P_{\Gamma^{-1}F}\Gamma^{-1}FKF'\Gamma^{-1\prime}GP_{\Gamma^{-1}F} + \sigma_0^{-2}P_{\Gamma^{-1}F}\Gamma^{-1}FKF'\Gamma^{-1\prime}G\Gamma^{-1}FKF'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F} - P_{\Gamma^{-1}F}G\Gamma^{-1}FKF'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&- P_{\Gamma^{-1}F}\Gamma^{-1}FKF'\Gamma^{-1\prime}GP_{\Gamma^{-1}F} + P_{\Gamma^{-1}F}\Gamma^{-1}FKF'\Gamma^{-1\prime}G\Gamma^{-1}FKF'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F} - P_{\Gamma^{-1}F}G\Gamma^{-1}F[(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma_0^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}]F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&- P_{\Gamma^{-1}F}\Gamma^{-1}F[(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma_0^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}]F'\Gamma^{-1\prime}GP_{\Gamma^{-1}F} \\
&+ P_{\Gamma^{-1}F}\Gamma^{-1}F[(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma_0^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}]F'\Gamma^{-1\prime}G \\
&\times \Gamma^{-1}F[(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma_0^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}]F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}GP_{\Gamma^{-1}F} - P_{\Gamma^{-1}F}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&+ \sigma_0^2 P_{\Gamma^{-1}F}G\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&- P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}GP_{\Gamma^{-1}F} \\
&+ \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}GP_{\Gamma^{-1}F} \\
&+ P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&- \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&- \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&+ \sigma_0^4 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&= \sigma_0^4 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F},
\end{aligned}$$

and

$$
\begin{aligned}
& P_{\Gamma^{-1}F}\hat{\Lambda}^{-1}P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}(I_T - \Gamma^{-1}FKF'\Gamma^{-1\prime})P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}P_{\Gamma^{-1}F} - P_{\Gamma^{-1}F}\Gamma^{-1}F[(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - \sigma_0^2(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}]F'\Gamma^{-1\prime}P_{\Gamma^{-1}F} \\
&= \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}P_{\Gamma^{-1}F}.
\end{aligned}
$$

We similarly have

$$
\begin{aligned}
& P_{\Gamma^{-1}F}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1}L_0P_{\Gamma^{-1}F} \\
&= P_{\Gamma^{-1}F}GL_0P_{\Gamma^{-1}F} - P_{\Gamma^{-1}F}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F} \\
&+ \sigma_0^2 P_{\Gamma^{-1}F}G\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F} \\
&- P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}GL_0P_{\Gamma^{-1}F} \\
&+ \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}GL_0P_{\Gamma^{-1}F} \\
&+ P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F} \\
&- \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F} \\
&- \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F} \\
&+ \sigma_0^4 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F} \\
&= \sigma^2 P_{\Gamma^{-1}F}[\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G - I_T]L_0P_{\Gamma^{-1}F} \\
&+ \sigma_0^4 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F},
\end{aligned}
$$

and

$$
P_{\Gamma^{-1}F}\hat{\Lambda}^{-1}L_0P_{\Gamma^{-1}F} = \sigma_0^2 P_{\Gamma^{-1}F}\Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}L_0P_{\Gamma^{-1}F}.
$$

It follows that

$$
\begin{aligned}
&\text{tr}\,(B_2'\bar{c}) \\
&= \text{tr}\,[(\sigma_0^{-2}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1} - \hat{\Lambda}^{-1})c_3] \\
&- \text{tr}\,[P_{\Gamma^{-1}F}(\sigma_0^{-2}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1} - \hat{\Lambda}^{-1})P_{\Gamma^{-1}F}(c_0 + c_1 + c_2)] \\
&= \text{tr}\,[(\sigma_0^{-2}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1} - \hat{\Lambda}^{-1})c_3] \\
&= \text{tr}\,[(\sigma_0^{-2}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1} - \hat{\Lambda}^{-1})(L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)P_{\Gamma^{-1}F} + P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)P_{\Gamma^{-1}F}L_0')] \\
&= 2\text{tr}\,[P_{\Gamma^{-1}F}(\sigma_0^{-2}\hat{\Lambda}^{-1}G\hat{\Lambda}^{-1} - \hat{\Lambda}^{-1})L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)] \\
&= -2\text{tr}\,[P_{\Gamma^{-1}F}(I_T - \Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G)L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)] \\
&= -2\text{tr}\,[P_{\Gamma^{-1}F}M_{\Gamma^{-1}F}L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)] \\
&- 2\text{tr}\,[P_{\Gamma^{-1}F}(P_{\Gamma^{-1}F} - \Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G)L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)] \\
&= -2\text{tr}\,[P_{\Gamma^{-1}F}(P_{\Gamma^{-1}F} - \Gamma^{-1}F(F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G)L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)].
\end{aligned}
$$

By adding and subtracting terms, and noting that $G = S_u$,

$$
\begin{aligned}
&(F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime} - (F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G \\
&= [(\sigma_0^2F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1} - (F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}]\sigma_0^2F'\Gamma^{-1\prime} - (F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}(G - \sigma_0^2 I_T) \\
&= (F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}(G - \sigma_0^2 I_T)\Gamma^{-1}F(\sigma_0^2F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}\sigma_0^2F'\Gamma^{-1\prime} \\
&- (F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}(G - \sigma_0^2 I_T) \\
&= -(F'\Gamma^{-1\prime}S_u\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}(S_u - \sigma_0^2 I_T)M_{\Gamma^{-1}F},
\end{aligned}
$$

which, together with $I_T + L_0 = \Gamma$ and $M_{\Gamma^{-1}F}P_{\Gamma^{-1}F} = 0_{T\times T}$, gives

$$
\begin{aligned}
&\sqrt{N}T^{-1}\text{tr}\,(B_2'\bar{c}) \\
&= -2\sqrt{N}T^{-1}\text{tr}\,[P_{\Gamma^{-1}F}\Gamma^{-1}F((F'\Gamma^{-1\prime}\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime} - (F'\Gamma^{-1\prime}G\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}G) \\
&\times \; L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)] \\
&= 2\sqrt{N}T^{-1}\text{tr}\,[\Gamma^{-1}F(F'\Gamma^{-1\prime}S_u\Gamma^{-1}F)^{-1}F'\Gamma^{-1\prime}(S_u - \sigma_0^2 I_T)M_{\Gamma^{-1}F}L_0P_{\Gamma^{-1}F}(S_u - \sigma_0^2 I_T)] \\
&= 2N^{-1/2}T^{-1}\text{tr}\,[(T^{-1}F'\Gamma^{-1\prime}S_u\Gamma^{-1}F)^{-1} \\
&\times \; T^{-1}F'\Gamma^{-1\prime}\sqrt{N}(S_u - \sigma_0^2 I_T)M_{\Gamma^{-1}F}\Gamma P_{\Gamma^{-1}F}\sqrt{N}(S_u - \sigma_0^2 I_T)\Gamma^{-1}F] \\
&= O_p(N^{-1/2}T^{-1}).
\end{aligned}
\tag{A174}
$$

By adding the results

$$\begin{aligned}\frac{1}{\sqrt{N}T}\frac{\partial \ell_c}{\partial \rho} &= \frac{1}{2\sigma_0^2}\sqrt{N}T^{-1}\text{tr}\left(B_2'\bar{c} - B_1'c_0\right) = -\frac{1}{2\sigma_0^2}\sqrt{N}T^{-1}\text{tr}\left(B_1'c_0\right) + O_p(N^{-1/2}T^{-1}) \\ &= -\frac{1}{\sigma_0^2}T^{-1/2}(Q_1 - Q_2) \sim N(0, T^{-1}\omega_2^2) \end{aligned} \tag{A175}$$

as N, $T \to \infty$ with $\sqrt{N}T^{-1} = o(1)$.

Since $\partial \ell_c / \partial \sigma^2$ is unaffected by the rescaling of F by $\Gamma^{-1}F$, the asymptotic distribution of $(NT)^{-1/2}\partial \ell_c / \partial \sigma^2$ is the unaffected too, as is its covariance with $N^{-1/2}T^{-1}\partial \ell_c / \partial \rho$. This completes the proof. ■

Lemma C.6. *Under C2, $\rho_0 = 1$, and Assumptions EPS, F and LAM, as N, $T \to \infty$,*

$$-H_1^{-1}\frac{\partial^2 \ell_c(\theta_2^0)}{\partial \theta_2 (\partial \theta_2)'}H_1^{-1} = \begin{bmatrix} T^{-1}\omega_2^2 & 0 \\ 0 & \frac{1}{2\sigma_0^4} \end{bmatrix} + o_p(1).$$

Proof of Lemma C.6.

The proof of Lemma C.6 is tedious, yet straightforward, following the same steps as in Proof of Lemma C.2. It is therefore omitted. ■

Proof of Proposition 1.

In this proof we only consider the case when $|\rho_0| < 1$ (under C1); the results for the case when $\rho_0 = 1$ are analogous (after suitable rescaling by T; see Proof of Lemma 2). The proof proceeds as follows. We begin by deriving the appropriate limit of $(NT)^{-1}\ell_c$. We then show that the leading term of this limit is minimized for $F = F^0$, where F^0 is the true value of F.

The first part of the proof is very similar to that of Lemma 1; hence, only essential details are given. Consider $F'GF = F'\Gamma^{-1}S_y\Gamma^{-1\prime}F$. Using F^0 to denote the true value of F, S_y may be expanded as

$$\begin{aligned} S_y &= \Gamma_0 \frac{1}{N}\sum_{i=1}^{N}(F^0\lambda_i + \varepsilon_i)(F^0\lambda_i + \varepsilon_i)'\Gamma_0' \\ &= \Gamma_0(\sigma_0^2 I_T + F^0 S_\lambda F^{0\prime})\Gamma_0' + \Gamma_0 \frac{1}{N}\sum_{i=1}^{N} F^0\lambda_i\varepsilon_i'\Gamma_0' + \Gamma_0\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i\lambda_i'F^{0\prime}\Gamma_0' \\ &+ \Gamma_0 \frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)\Gamma_0', \end{aligned} \tag{A176}$$

giving

$$\begin{aligned} T^{-2}F'GF &= T^{-1}F'\Gamma^{-1}(T^{-1}S_y)\Gamma^{-1\prime}F \\ &= T^{-2}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + F^0 S_\lambda F^{0\prime})\Gamma_0'\Gamma^{-1\prime}F + T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^{N}F^0\lambda_i\varepsilon_i'\Gamma_0'\Gamma^{-1\prime}F \\ &+ T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^{N}\varepsilon_i\lambda_i'F^{0\prime}\Gamma_0'\Gamma^{-1\prime}F \\ &+ T^{-1}F'\Gamma^{-1}\Gamma_0\frac{1}{NT}\sum_{i=1}^{N}(\varepsilon_i\varepsilon_i' - \sigma_0^2 I_T)\Gamma_0'\Gamma^{-1\prime}F. \end{aligned} \tag{A177}$$

The first term on the right is

$$\begin{aligned} &T^{-2}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + F^0 S_\lambda F^{0\prime})\Gamma_0'\Gamma^{-1\prime}F \\ &= \sigma_0^2 T^{-2}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F + T^{-2}F'\Gamma^{-1}\Gamma_0F^0S_\lambda F^{0\prime}\Gamma_0'\Gamma^{-1\prime}F, \end{aligned}$$

where we know from Proof of Lemma 1 that $T^{-2}F'\Gamma^{-1}\Gamma_0\Gamma_0'\Gamma^{-1\prime}F = O(T^{-1})$. But we also have

$$\begin{aligned} F'\Gamma^{-1}\Gamma_0F^0S_\lambda F^{0\prime}\Gamma_0'\Gamma^{-1\prime}F &= F'[I_T + (\rho_0-\rho)L_0]F^0S_\lambda F^{0\prime}[I_T + (\rho_0-\rho)L_0]'F \\ &= F'F^0S_\lambda F^{0\prime}F + (\rho_0-\rho)F'F^0S_\lambda F^{0\prime}L_0'F + (\rho_0-\rho)F'L_0F^0S_\lambda F^{0\prime}F \\ &+ (\rho_0-\rho)^2F'L_0F^0S_\lambda F^{0\prime}L_0'F, \end{aligned}$$

and therefore

$$\begin{aligned} &T^{-2}F'\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + F^0S_\lambda F^{0\prime})\Gamma_0'\Gamma^{-1\prime}F \\ &= (T^{-1}F'F^0)S_\lambda(T^{-1}F^{0\prime}F) + (\rho_0-\rho)(T^{-1}F'F^0)S_\lambda T^{-1}F^{0\prime}L_0'F \\ &+ (\rho_0-\rho)T^{-1}F'L_0F^0S_\lambda(T^{-1}F^{0\prime}F) + (\rho_0-\rho)^2T^{-1}F'L_0F^0S_\lambda T^{-1}F^{0\prime}L_0'F + O(T^{-1}). \end{aligned} \tag{A178}$$

By using the same arguments as in Proof of Lemma 1, while the second and third terms on the right of (A177) are $O_p((NT)^{-1/2})$, the fourth term is $O_p(N^{-1/2}T^{-1})$. It follows that

$$\begin{aligned} T^{-2}F'GF &= (T^{-1}F'F^0)S_\lambda(T^{-1}F^{0\prime}F) + (\rho_0-\rho)(T^{-1}F'F^0)S_\lambda T^{-1}F^{0\prime}L_0'F \\ &+ (\rho_0-\rho)T^{-1}F'L_0F^0S_\lambda(T^{-1}F^{0\prime}F) \\ &+ (\rho_0-\rho)^2T^{-1}F'L_0F^0S_\lambda T^{-1}F^{0\prime}L_0'F + O(T^{-1}) + O_p((NT)^{-1/2}), \end{aligned} \tag{A179}$$

which in turn implies

$$
\begin{aligned}
&T^{-1}\text{tr}\,(GFKF') \\
&= \text{tr}\,[T^{-2}F'GF(T^{-1}F'F)^{-1}] - \sigma^2 T^{-1}m \\
&= \text{tr}\,[T^{-2}F'\Gamma^{-1}\Gamma_0 F^0 S_\lambda F^{0\prime}\Gamma_0'\Gamma^{-1\prime}F(T^{-1}F'F)^{-1}] + O(T^{-1}) + O_p((NT)^{-1/2}) \\
&= \text{tr}\,[(T^{-1}F'F^0)S_\lambda(T^{-1}F^{0\prime}F)(T^{-1}F'F)^{-1}] + (\rho_0-\rho)\text{tr}\,[(T^{-1}F'F^0)S_\lambda T^{-1}F^{0\prime}L_0'F(T^{-1}F'F)^{-1}] \\
&+ (\rho_0-\rho)\text{tr}\,[T^{-1}F'L_0F^0S_\lambda(T^{-1}F^{0\prime}F)(T^{-1}F'F)^{-1}] \\
&+ (\rho_0-\rho)^2\text{tr}\,[T^{-1}F'L_0FS_\lambda T^{-1}F'L_0'F(T^{-1}F'F)^{-1}] + O(T^{-1}) + O_p((NT)^{-1/2}) \\
&= \text{tr}\,(S_\lambda T^{-1}F^{0\prime}P_F F^0) + 2(\rho_0-\rho)\text{tr}\,(S_\lambda T^{-1}F^{0\prime}P_F L_0 F^0) \\
&+ (\rho_0-\rho)^2\text{tr}\,(S_\lambda T^{-1}F'L_0'P_F L_0 F) + O(T^{-1}) + O_p((NT)^{-1/2}). \qquad \text{(A180)}
\end{aligned}
$$

For $\text{tr}\,G$, by the arguments of the proof of Lemma 1,

$$
\begin{aligned}
T^{-1}\text{tr}\,G &= T^{-1}\text{tr}\,(\Gamma^{-1}S_y\Gamma^{-1\prime}) = T^{-1}\text{tr}\,[\Gamma^{-1}\Gamma_0(\sigma_0^2 I_T + FS_\lambda F')\Gamma_0'\Gamma^{-1\prime}] + o_p(1) \\
&= \sigma_0^2[1 + (\rho_0-\rho)^2\text{tr}\,(T^{-1}L_0L_0')] + \text{tr}\,(S_\lambda T^{-1}F^{0\prime}F^0) + 2(\rho_0-\rho)\text{tr}\,(S_\lambda T^{-1}F^{0\prime}L_0F^0) \\
&+ (\rho_0-\rho)^2\text{tr}\,(S_\lambda T^{-1}F'L_0'L_0F^0) + O(T^{-1}). \qquad \text{(A181)}
\end{aligned}
$$

By using this and the fact that $T^{-1}\log(|\hat{\Lambda}|) = O_p(T^{-1}\log(T))$ (see Proof of Lemma 1), we obtain

$$
\begin{aligned}
&T^{-1}Q_c \\
&= \log(\sigma^2) + T^{-1}\log(|\hat{\Lambda}|) + \sigma^{-2}T^{-1}\text{tr}\,G - \sigma^{-2}T^{-1}\text{tr}\,(GFKF') \\
&= \log(\sigma^2) + \sigma^{-2}T^{-1}\text{tr}\,G - \sigma^{-2}T^{-1}\text{tr}\,(GFKF') + O_p(T^{-1}\log(T)) \\
&= \log(\sigma^2) + \sigma^{-2}\sigma_0^2[1 + (\rho_0-\rho)^2\text{tr}\,(T^{-1}L_0L_0')] + \sigma^{-2}\text{tr}\,(S_\lambda T^{-1}F^{0\prime}F^0) \\
&+ 2\sigma^{-2}(\rho_0-\rho)\text{tr}\,(S_\lambda T^{-1}F^{0\prime}L_0F^0) + \sigma^{-2}(\rho_0-\rho)^2\text{tr}\,(S_\lambda T^{-1}F'L_0'L_0F^0) \\
&- \sigma^{-2}\text{tr}\,(S_\lambda T^{-1}F^{0\prime}P_F F^0) - 2\sigma^{-2}(\rho_0-\rho)\text{tr}\,(S_\lambda T^{-1}F^{0\prime}P_F L_0 F^0) \\
&- \sigma^{-2}(\rho_0-\rho)^2\text{tr}\,(S_\lambda T^{-1}F'L_0'P_F L_0 F) + O_p((NT)^{-1/2}) + O_p(T^{-1}\log(T)) \\
&= \log(\sigma^2) + \sigma^{-2}\sigma_0^2[1 + (\rho_0-\rho)^2\text{tr}\,(T^{-1}L_0L_0')] + \sigma^{-2}\text{tr}\,(S_\lambda T^{-1}F^{0\prime}M_F F^0) \\
&+ 2\sigma^{-2}(\rho_0-\rho)\text{tr}\,(S_\lambda T^{-1}F^{0\prime}M_F L_0 F^0) + \sigma^{-2}(\rho_0-\rho)^2\text{tr}\,(S_\lambda T^{-1}F'L_0'M_F L_0 F^0) \\
&+ O_p((NT)^{-1/2}) + O_p(T^{-1}\log(T)). \qquad \text{(A182)}
\end{aligned}
$$

As in Lemma 1 we can show that all four terms involving the trace are nonnegative. Hence, since $\text{tr}\,(S_\lambda T^{-1}F^{0\prime}M_F F^0) = \text{tr}\,(S_\lambda T^{-1}F^{0\prime}M_F L_0 F^0) = 0$ for $F = F^0$, Q_c (ℓ_c) is minimized (maximized) for $F = F^0$. This completes the proof of the proposition. ■

Appendix D: Additional results

In Lemmas D.1 and D.2 we report the limits of $T^{-1}F'F$, $T^{-1}F'L_0'F$ and $T^{-1}F'L_0'L_0F$ in two cases with a specific F and $|\rho| < 1$ (under C1). These limits are useful for evaluating ω_1^2. In the first case, F is made up of an intercept subject to a single structural break at time $T_B = \lfloor \tau T \rfloor$, where $\tau \in (0,1)$ is the break fraction and $\lfloor x \rfloor$ is the integer part of x. That is, $F = (1_T, D_{T_B})$, where $1_T = (1, ..., 1)'$ is a $T \times 1$ vector of ones and $D_{T_B} = (1'_{T_B}, 0'_{(T-T_B)\times 1})'$. In the second case, F is made up of an intercept and (normalized) time trend; $F = (1_T, T^{-1}\tau_T)$, where $\tau_T = (1, ..., T)'$.

Lemma D.1. *Suppose that $F = (1_T, D_{T_B})$. Under $|\rho_0| < 1$ and Assumption F,*

$$\begin{aligned} T^{-1}F'F &= \Sigma_F + O(T^{-1}), \\ T^{-1}F'L_0'F &= \frac{1}{(1-\rho_0)}\Sigma_F + O(T^{-1}), \\ T^{-1}F'L_0'L_0F &= \frac{1}{(1-\rho_0)^2}\Sigma_F + O(T^{-1}), \end{aligned}$$

where

$$\Sigma_F = \begin{bmatrix} 1 & \tau \\ \tau & \tau \end{bmatrix}.$$

Proof of Lemma D.1.

The first result is easy;

$$T^{-1}F'F = \frac{1}{T}\begin{bmatrix} 1'_T 1_T & 1'_T D_{T_B} \\ D'_{T_B} 1_T & D'_{T_B} D_{T_B} \end{bmatrix} = \begin{bmatrix} 1 & \tau \\ \tau & \tau \end{bmatrix} + O(T^{-1}) = \Sigma_F + O(T^{-1}). \tag{A183}$$

Consider $T^{-1}F'L_0'F$. According to the ratio test, if $\lim_{t\to\infty} |a_{t+1}/a_t| < 1$, then $\sum_{t=0}^{T} a_t$ is convergent. Hence, since $|\rho_0^{t+1}/\rho_0^t| = |\rho_0| < 1$, $\sum_{t=0}^{T}\rho_0^t$ converges, as does $\sum_{t=0}^{T}\rho_0^{2t}$. Specifically, $\sum_{t=0}^{T}\rho_0^t = (1-\rho_0^{T+1})/(1-\rho_0) \to 1/(1-\rho_0)$ and $\sum_{t=0}^{T}\rho_0^{2t} = (1-\rho_0^{2(T+1)})/(1-\rho_0^2) \to 1/(1-\rho_0^2)$ as $T \to \infty$. Also, for any sequence $\{a_t\}_{t=1}^{T}$, $\sum_{t=1}^{T}\sum_{s=t+1}^{T+1} a_{s-(t+1)} = \sum_{t=0}^{T-1}(T-t)a_t$. It follows that

$$\begin{aligned} T^{-1}1'_T L_0' 1_T &= \frac{1}{T}\sum_{t=1}^{T} l'_{t,0}1_T = \frac{1}{T}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\rho_0^{s-(t+1)} = \frac{1}{T}\sum_{t=0}^{T-2}(T-1-t)\rho_0^t \\ &= \frac{(T-1)}{T}\sum_{t=0}^{T-1}\rho_0^t - \frac{1}{T}\sum_{t=0}^{T-1}t\rho_0^t = \frac{(T-1)}{T}\sum_{t=0}^{T-1}\rho_0^t + O(T^{-1}) \\ &= \frac{1}{(1-\rho_0)} + O(T^{-1}), \end{aligned}$$

and, similarly, since $l'_{t,0}D_{T_B} = 0$ for all $t = T_B + 1, ..., T$,

$$\begin{aligned} & T^{-1}1'_T L'_0 D_{T_B} \\ &= \frac{1}{T}\sum_{t=1}^{T} l'_{t,0}D_{T_B} = \frac{1}{T}\sum_{t=1}^{T_B} l'_{t,0}D_{T_B} = \frac{1}{T}\sum_{t=1}^{T_B-1}\sum_{s=t+1}^{T_B}\rho_0^{s-(t+1)} = \frac{1}{T}\sum_{t=0}^{T_B-2}(T_B-1-t)\rho_0^t \\ &= \frac{(T_B-1)}{T}\sum_{t=0}^{T_B-1}\rho_0^t + O(T^{-1}) = \frac{\tau}{(1-\rho_0)} + O(T^{-1}). \end{aligned}$$

Consider $\sum_{t=0}^{T} t\rho_0^t$. Since $|(t+1)\rho_0^{t+1}/(t\rho_0^t)| = [(t+1)/t]|\rho_0| \to |\rho_0| < 1$, $\sum_{t=1}^{T} t\rho_0^t$ converges. Moreover, $\rho_0^{T-T_B} = \rho_0^{\lfloor(1-\tau)T\rfloor} \to 0$ and $\sum_{t=0}^{T-T_B}\rho_0^t = (1-\rho_0^{T-T_B+1})/(1-\rho_0) = (1-\rho_0^{\lfloor(1-\tau)T\rfloor+1})/(1-\rho_0) \to 1/(1-\rho_0)$. These results imply

$$\begin{aligned} T^{-1}D'_{T_B}L'_0 1_T &= \frac{1}{T}\sum_{t=1}^{T_B} l'_{t,0}1_T = \frac{1}{T}\sum_{t=1}^{T_B-1}\sum_{s=t+1}^{T}\rho_0^{s-(t+1)} \\ &= \frac{(T_B-1)}{T}\sum_{t=0}^{T-T_B}\rho_0^t + \frac{1}{T}\sum_{t=0}^{T_B-3}(T_B-2-t)\rho_0^{T-T_B+1+t} \\ &= \frac{(T_B-1)}{T}\sum_{t=0}^{T-T_B}\rho_0^t + \rho_0^{T-T_B+1}\frac{(T_B-2)}{T}\sum_{t=0}^{T_B-3}\rho_0^t - \rho_0^{T-T_B+1}\frac{1}{T}\sum_{t=0}^{T_B-3}t\rho_0^t \\ &= \frac{(T_B-1)}{T}\sum_{t=0}^{T-T_B}\rho_0^t + O(T^{-1}) = \frac{\tau}{(1-\rho_0)} + O(T^{-1}), \\ T^{-1}D'_{T_B}L'_0 D_{T_B} &= \frac{1}{T}\sum_{t=1}^{T_B} l'_{t,0}D_{T_B} = \frac{1}{T}\sum_{t=1}^{T_B-1}\sum_{s=t+1}^{T_B}\rho_0^{s-(t+1)} = \frac{\tau}{(1-\rho_0)} + O(T^{-1}). \end{aligned}$$

Therefore,

$$\begin{aligned} T^{-1}F'L'_0F &= T^{-1}\begin{bmatrix} 1'_T L'_0 1_T & 1'_T L'_0 D_{T_B} \\ D'_{T_B}L'_0 1_T & D'_{T_B}L'_0 D_{T_B} \end{bmatrix} \\ &= \frac{1}{(1-\rho_0)}\begin{bmatrix} 1 & \tau \\ \tau & \tau \end{bmatrix} + O(T^{-1}) = \frac{1}{(1-\rho_0)}\Sigma_F + O(T^{-1}). \end{aligned} \tag{A184}$$

It remains to consider

$$T^{-1}F'L'_0L_0F = \frac{1}{T}\begin{bmatrix} 1'_T L'_0 L_0 1_T & 1'_T L'_0 L_0 D_{T_B} \\ D'_{T_B}L'_0 L_0 1_T & D'_{T_B}L'_0 L_0 D_{T_B} \end{bmatrix}.$$

Note that $l'_{t,0}l_{s,0} = \sum_{n=t+1}^{T} \rho_0^{2n-(t+s+2)} = \sum_{n=0}^{T-1-t} \rho_0^{2n+t-s}$ for $t \geq s$, suggesting

$$\begin{aligned}
T^{-1}1'_T L'_0 L_0 1_T &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} l'_{t,0}l_{s,0} = \frac{1}{T}\sum_{t=1}^{T} l'_{t,0}l_{t,0} + \frac{2}{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1} l'_{t,0}l_{s,0} \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{n=0}^{T-1-t}\rho_0^{2n} + \frac{2}{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\rho_0^{t-s}\sum_{n=0}^{T-1-t}\rho_0^{2n} \\
&= \frac{1}{T(1-\rho_0^2)}\sum_{t=1}^{T}(1-\rho_0^{2(T-t)}) + \frac{2}{T(1-\rho_0^2)}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\rho_0^{t-s}(1-\rho_0^{2(T-t)}) \\
&= \frac{1}{(1-\rho_0^2)} - \frac{1}{T(1-\rho_0^2)}\sum_{t=1}^{T}\rho_0^{2(T-t)} + \frac{2}{T(1-\rho_0^2)}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\rho_0^{t-s} \\
&- \frac{2}{T(1-\rho_0^2)}\sum_{t=2}^{T}\rho_0^{2(T-t)}\sum_{s=1}^{t-1}\rho_0^{t-s} \\
&= \frac{1}{(1-\rho_0^2)} - \frac{1}{T(1-\rho_0^2)}\sum_{t=0}^{T-1}\rho_0^{2t} + \frac{2\rho_0}{T(1-\rho_0^2)}\sum_{t=0}^{T-2}(T-1-t)\rho_0^t \\
&- \frac{2}{T(1-\rho_0^2)}\sum_{t=2}^{T}\rho_0^{2(T-t)}\sum_{s=1}^{t-1}\rho_0^{s} \\
&= \frac{1}{(1-\rho_0^2)} + \frac{2\rho_0}{T(1-\rho_0^2)}\sum_{t=0}^{T-2}(T-1-t)\rho_0^t + O(T^{-1}) \\
&= \frac{1}{(1-\rho_0^2)} + \frac{2\rho_0}{(1-\rho_0^2)(1-\rho_0)} + O(T^{-1}) \\
&= \frac{(1+\rho_0)}{(1-\rho_0^2)(1-\rho_0)} + O(T^{-1}) = \frac{1}{(1-\rho_0)^2} + O(T^{-1}).
\end{aligned}$$

Moreover, since again $\rho_0^{T_B} \to 0$, $T_B/T \to \tau$, and $\sum_{t=1}^{T_B}\rho_0^t$ and $\sum_{t=1}^{T_B} t\rho_0^t$ are convergent,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=T_B+1}^{T}\sum_{s=1}^{T_B} l'_{t,0}l_{s,0} \\
&= \frac{1}{T}\sum_{t=T_B+1}^{T}\sum_{s=1}^{T_B}\rho_0^{t-s}\sum_{n=0}^{T-1-t}\rho_0^{2n} = \frac{1}{T(1-\rho_0^2)}\sum_{t=T_B+1}^{T}\sum_{s=1}^{T_B}\rho_0^{t-s}(1-\rho_0^{2(T-t)}) \\
&= \frac{1}{T(1-\rho_0^2)}\sum_{t=T_B+1}^{T}\sum_{s=1}^{T_B}\rho_0^{t-s} + O(T^{-1}) \\
&= \frac{1}{T(1-\rho_0^2)}T_B\rho_0^{T_B} + \frac{1}{T(1-\rho_0^2)}\sum_{s=1}^{T_B-1} t\rho_0^t + \frac{1}{T(1-\rho_0^2)}\sum_{t=1}^{T_B-1}(T_B-t)\rho_0^{T_B+t} + O(T^{-1}) \\
&= \frac{1}{T(1-\rho_0^2)}\rho_0^{T_B}\sum_{t=1}^{T_B-1}(T_B-t)\rho_0^t + O(T^{-1}) = \frac{1}{T(1-\rho_0^2)}\rho_0^{T_B}T_B\sum_{t=1}^{T_B-1}\rho_0^t + O(T^{-1}) \\
&= O(T^{-1}),
\end{aligned}$$

and therefore

$$\begin{aligned} T^{-1}1_T'L_0'L_0D_{T_B} &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T_B} l_{t,0}'l_{s,0} = \frac{1}{T}\sum_{t=T_B+1}^{T}\sum_{s=1}^{T_B} l_{t,0}'l_{s,0} + \frac{1}{T}\sum_{t=1}^{T_B}\sum_{s=1}^{T_B} l_{t,0}'l_{s,0} \\ &= \frac{1}{T}\sum_{t=T_B+1}^{T}\sum_{s=1}^{T_B}\rho_0^{t-s}\sum_{n=0}^{T-1-t}\rho_0^{2n} + \frac{1}{T}\sum_{t=1}^{T_B}\sum_{s=1}^{T_B} l_{t,0}'l_{s,0} \\ &= \frac{\tau}{(1-\rho_0^2)} + \frac{2\tau\rho_0}{(1-\rho_0^2)(1-\rho_0)} + O(T^{-1}) = \frac{\tau}{(1-\rho_0)^2} + O(T^{-1}), \\ T^{-1}D_{T_B}'L_0'L_0D_{T_B} &= \frac{1}{T}\sum_{t=1}^{T_B}\sum_{s=1}^{T_B} l_{t,0}'l_{s,0} = \frac{\tau}{(1-\rho_0)^2} + O(T^{-1}). \end{aligned}$$

It follows that

$$T^{-1}F'L_0'L_0F = \frac{1}{(1-\rho_0)^2}\Sigma_F + O(T^{-1}), \tag{A185}$$

and so we are done. ■

Lemma D.2. *Suppose that $F = (1_T, T^{-1}\tau_T)$. Under $|\rho_0| < 1$ and Assumption F the results of Lemma D.1 hold, but with Σ_F given by*

$$\Sigma_F = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}.$$

Proof of Lemma D.2.

For $T^{-1}F'F$,

$$T^{-1}F'F = T^{-1}\begin{bmatrix} 1_T'1_T & T^{-1}1_T'\tau_T \\ T^{-1}\tau_T'1_T & T^{-2}\tau_T'\tau_T \end{bmatrix} = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} 1 & T^{-1}t \\ T^{-1}t & T^{-2}t^2 \end{bmatrix} = \Sigma_F + O(T^{-1}), \tag{A186}$$

where

$$\Sigma_F = \int_0^1 \begin{bmatrix} 1 & r \\ r & r^2 \end{bmatrix} dr = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}, \tag{A187}$$

with $r \in [0,1]$ being the limit of $T^{-1}t$. The order of the error of approximation follows from $\sup_{1\le t\le T}\sup_{(t-1)/T\le r\le t/T}|(T^{-1}t)^k - r^k| = O(T^{-1})$ for all $k < \infty$.

Consider

$$T^{-1}F'L_0'F = T^{-1}\begin{bmatrix} 1_T'L_0'1_T & T^{-1}1_T'L_0'\tau_T \\ T^{-1}\tau_T'L_0'1_T & T^{-2}\tau_T'L_0'\tau_T \end{bmatrix},$$

where $T^{-1}1_T'L_0'1_T$ is know from Lemma D.2. Let us therefore consider $T^{-2}1_T'L_0'\tau_T$. Since $|(t+1)\rho_0^{t+1}/(t\rho_0^t)| = [(t+1)/t]|\rho_0| \to |\rho_0| < 1$, $\sum_{t=1}^{T} t\rho_0^t$ converges, as do $\sum_{t=1}^{T}\rho_0^t$, $\sum_{t=1}^{T} t^2\rho_0^t$,

$\sum_{t=1}^{T}\rho_0^{2t}$ and $\sum_{t=1}^{T}t\rho_0^{2t}$. By using these results and $\sum_{t=1}^{T}t = T(T+1)/2$, we obtain

$$\begin{aligned}
&T^{-2}1_T'L_0'\tau_T \\
&= \frac{1}{T^2}\sum_{t=1}^{T}l_{t,0}'\tau_T = \frac{1}{T^2}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}s\rho_0^{s-(t+1)} = \frac{1}{T^2}\sum_{t=0}^{T-2}\sum_{s=t+2}^{T}s\rho_0^{t} = \frac{1}{T^2}\sum_{t=0}^{T-2}\sum_{s=1}^{T-1-t}(s+t+1)\rho_0^{t} \\
&= \frac{1}{T^2}\sum_{t=0}^{T-2}\sum_{s=1}^{T-1-t}s\rho_0^{t} + \frac{1}{T^2}\sum_{t=0}^{T-2}(t+1)(T-1-t)\rho_0^{t} \\
&= \frac{1}{T^2}\sum_{t=0}^{T-2}\frac{(T-1-t)(T-t)}{2}\rho_0^{t} + \frac{1}{T^2}\sum_{t=0}^{T-2}(t+1)(T-1-t)\rho_0^{t} \\
&= \frac{(T-1)}{2T}\sum_{t=0}^{T-2}\rho_0^{t} - \frac{(2T-1)}{2T^2}\sum_{t=0}^{T-2}t\rho_0^{t} + \frac{1}{2T^2}\sum_{t=0}^{T-2}t^2\rho_0^{t} + \frac{1}{T^2}\sum_{t=0}^{T-2}(t+1)(T-1-t)\rho_0^{t} \\
&= \frac{(T-1)}{2T}\sum_{t=0}^{T-2}\rho_0^{t} + O(T^{-1}) = \frac{1}{2(1-\rho_0)} + O(T^{-1}),
\end{aligned}$$

and by further use of $\sum_{t=1}^{T}t^2 = T(T+1)(2T+1)/6$,

$$\begin{aligned}
T^{-2}\tau_T'L_0'1_T &= \frac{1}{T^2}\sum_{t=1}^{T}tl_{t,0}'1_T = \frac{1}{T^2}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}t\rho_0^{s-(t+1)} = \frac{1}{T^2}\sum_{t=0}^{T-2}\sum_{s=1}^{T-1-t}s\rho_0^{t} \\
&= \frac{(T-1)}{2T}\sum_{t=0}^{T-2}\rho_0^{t} + O(T^{-1}) = \frac{1}{2(1-\rho_0)} + O(T^{-1}), \\
T^{-3}\tau_T'L_0'\tau_T &= \frac{1}{T^3}\sum_{t=1}^{T}tl_{t,0}'\tau_T = \frac{1}{T^3}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}ts\rho_0^{s-(t+1)} = \frac{1}{T^3}\sum_{t=0}^{T-2}\sum_{s=t+2}^{T}(s-1-t)(s-t)\rho_0^{t} \\
&= \frac{1}{T^3}\sum_{t=0}^{T-2}\sum_{s=1}^{T-1-t}s(s+1)\rho_0^{t} = \frac{1}{T^3}\sum_{t=0}^{T-2}\sum_{s=1}^{T-1-t}s^2\rho_0^{t} + \frac{1}{T^3}\sum_{t=0}^{T-2}(T-1-t)\rho_0^{t} \\
&= \frac{1}{T^3}\sum_{t=0}^{T-2}\sum_{s=1}^{T-1-t}s^2\rho_0^{t} + O(T^{-2}) \\
&= \frac{1}{T^3}\sum_{t=0}^{T-2}\frac{(T-1-t)(T-t)[2(T-1-t)+1]}{6}\rho_0^{t} + O(T^{-2}) \\
&= \frac{1}{3}\sum_{t=0}^{T-2}\rho_0^{t} + O(T^{-1}) = \frac{1}{3(1-\rho_0)} + O(T^{-1}).
\end{aligned}$$

Therefore,

$$\begin{aligned}
T^{-1}F'L_0'F &= \begin{bmatrix} T^{-1}1_T'L_0'1_T & T^{-2}1_T'L_0'\tau_T \\ T^{-2}\tau_T'L_0'1_T & T^{-3}\tau_T'L_0'\tau_T \end{bmatrix} \\
&= \frac{1}{(1-\rho_0)}\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix} + O(T^{-1}) = \frac{1}{(1-\rho_0)}\Sigma_F + O(T^{-1}). \qquad \text{(A188)}
\end{aligned}$$

Finally, consider $T^{-1}F'L_0'L_0F$. Since $\sum_{t=1}^{T}t^2\rho_0^{2(T-t)}$ and $\sum_{t=1}^{T}t\rho_0^{2(T-t)}$ are convergent by

the ratio test,

$$\begin{aligned}
&\frac{1}{T^3}\sum_{t=2}^{T} t\rho_0^{2(T-t)}\sum_{s=1}^{t-1} s\rho_0^{t-s} \\
&= \frac{1}{T^3}\sum_{t=2}^{T} t\rho_0^{2(T-t)}\sum_{s=1}^{t-1}(t-s)\rho_0^{s} \\
&= \frac{1}{T^3}\sum_{t=2}^{T} t^2\rho_0^{2(T-t)}\sum_{s=1}^{t-1}\rho_0^{s} - \frac{1}{T^3}\sum_{t=2}^{T} t\rho_0^{2(T-t)}\sum_{s=1}^{t-1} s\rho_0^{s} \\
&= \frac{1}{T^3(1-\rho_0)}\sum_{t=2}^{T} t^2\rho_0^{2(T-t)}(1-\rho_0^{t}) - \frac{1}{T^3}\sum_{t=2}^{T} t\rho_0^{2(T-t)}\sum_{s=1}^{t-1} s\rho_0^{s} = O(T^{-2}).
\end{aligned}$$

This implies

$$\begin{aligned}
&T^{-3}\tau_T' L_0' L_0 \tau_T \\
&= \frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{T} tsl_{t,0}' l_{s,0} = \frac{1}{T^3}\sum_{t=1}^{T} t^2 l_{t,0}' l_{t,0} + \frac{2}{T^3}\sum_{t=2}^{T}\sum_{s=1}^{t-1} tsl_{t,0}' l_{s,0} \\
&= \frac{1}{T^3}\sum_{t=1}^{T} t^2\sum_{n=0}^{T-1-t}\rho_0^{2n} + \frac{2}{T^3}\sum_{t=2}^{T}\sum_{s=1}^{t-1} ts\rho_0^{t-s}\sum_{n=0}^{T-1-t}\rho_0^{2n} \\
&= \frac{1}{T^3(1-\rho_0^2)}\sum_{t=1}^{T} t^2(1-\rho_0^{2(T-t)}) + \frac{2}{T^3(1-\rho_0^2)}\sum_{t=2}^{T}\sum_{s=1}^{t-1} ts\rho_0^{t-s}(1-\rho_0^{2(T-t)}) \\
&= \frac{1}{T^3(1-\rho_0^2)}\frac{T(T+1)(2T+1)}{6} - \frac{1}{T^3(1-\rho_0^2)}\sum_{t=1}^{T} t^2\rho_0^{2(T-t)} + \frac{2}{T^3(1-\rho_0^2)}\sum_{t=2}^{T} t\sum_{s=1}^{t-1} s\rho_0^{t-s} \\
&- \frac{2}{T^3(1-\rho_0^2)}\sum_{t=2}^{T} t\rho_0^{2(T-t)}\sum_{s=1}^{t-1} s\rho_0^{t-s} \\
&= \frac{1}{3(1-\rho_0^2)} + \frac{2\rho_0}{3(1-\rho_0^2)(1-\rho_0)} + O(T^{-1}) \\
&= \frac{1+\rho_0}{3(1-\rho_0^2)(1-\rho_0)} + O(T^{-1}) = \frac{1}{3(1-\rho_0)^2} + O(T^{-1}),
\end{aligned}$$

where the last equality holds because

$$\begin{aligned}
\frac{1}{T^3}\sum_{t=2}^{T} t\sum_{s=1}^{t-1} s\rho_0^{t-s} &= \frac{1}{T^3}\sum_{t=2}^{T} t\sum_{s=1}^{t-1}(t-s)\rho_0^{s} = \frac{1}{T^3}\sum_{t=2}^{T} t^2\sum_{s=1}^{t-1}\rho_0^{s} - \frac{1}{T^3}\sum_{t=2}^{T} t\sum_{s=1}^{t-1} s\rho_0^{s} \\
&= \frac{\rho_0}{T^3}\sum_{t=2}^{T} t^2\sum_{s=0}^{t-2}\rho_0^{s} + O(T^{-1}) = \frac{\rho_0}{T^3(1-\rho_0)}\sum_{t=2}^{T} t^2(1-\rho_0^{t-1}) + O(T^{-1}) \\
&= \frac{\rho_0}{3(1-\rho_0)} + O(T^{-1}).
\end{aligned}$$

The same steps can be used to show that

$$T^{-2}\tau_T' L_0' L_0 1_T = T^{-2} 1_T' L_0' L_0 \tau_T = \frac{1}{2(1-\rho_0)^2} + O(T^{-1}).$$

giving

$$T^{-1}F'L_0'L_0F = \frac{1}{(1-\rho_0)^2}\Sigma_F + O(T^{-1}). \tag{A189}$$

as required. ■

Table 1: Bias, RMSE and 5% size results for F1 when $|\rho_0| < 1$.

		AHL			AHD			FA			LS			BCLS		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
								$\rho_0 = 0$								
10	10	0.0146	0.2180	0.2	0.0642	0.3366	8.7	0.0002	0.0892	5.3	−0.1100	0.1505	24.6	−0.0210	0.1149	8.0
10	50	0.0026	0.0677	0.0	0.0089	0.1141	9.1	0.0008	0.0437	4.7	−0.0194	0.0493	11.1	0.0002	0.0462	5.3
10	100	0.0007	0.0464	0.0	0.0025	0.0795	10.4	0.0005	0.0309	4.4	−0.0095	0.0328	8.2	0.0004	0.0317	4.8
10	200	0.0007	0.0322	0.0	0.0023	0.0562	9.8	0.0005	0.0221	4.3	−0.0045	0.0227	6.8	0.0004	0.0224	4.4
50	10	0.0029	0.0870	0.1	0.0121	0.1313	8.3	0.0001	0.0420	5.2	−0.1111	0.1203	73.8	−0.0222	0.0553	12.3
50	50	0.0008	0.0301	0.0	0.0027	0.0510	9.4	0.0000	0.0195	5.0	−0.0206	0.0286	25.7	−0.0010	0.0204	5.9
50	100	0.0002	0.0208	0.0	0.0010	0.0359	9.9	−0.0002	0.0141	5.2	−0.0103	0.0176	17.3	−0.0004	0.0144	5.3
50	200	0.0001	0.0142	0.0	0.0005	0.0247	8.9	0.0001	0.0099	5.3	−0.0050	0.0112	12.5	0.0000	0.0101	4.9
100	10	0.0009	0.0631	0.2	0.0034	0.0932	8.4	0.0001	0.0299	4.5	−0.1116	0.1164	94.3	−0.0228	0.0428	16.4
100	50	−0.0001	0.0216	0.0	−0.0003	0.0358	10.0	−0.0001	0.0141	5.1	−0.0205	0.0250	41.7	−0.0009	0.0147	5.9
100	100	0.0001	0.0148	0.0	0.0000	0.0249	9.6	0.0000	0.0101	5.2	−0.0101	0.0143	26.2	−0.0002	0.0103	5.6
100	200	0.0000	0.0103	0.0	−0.0001	0.0175	9.5	0.0001	0.0071	5.8	−0.0050	0.0087	17.1	0.0000	0.0072	5.1
200	10	0.0012	0.0445	0.1	0.0033	0.0645	7.4	−0.0003	0.0215	5.2	−0.1111	0.1135	99.8	−0.0222	0.0339	23.2
200	50	0.0003	0.0152	0.0	0.0007	0.0252	8.7	0.0001	0.0100	5.4	−0.0204	0.0228	63.6	−0.0008	0.0105	6.1
200	100	0.0002	0.0104	0.0	0.0002	0.0178	8.9	0.0001	0.0071	5.7	−0.0100	0.0123	40.7	−0.0001	0.0073	5.2
200	200	0.0002	0.0072	0.0	0.0002	0.0124	9.0	0.0001	0.0050	5.3	−0.0049	0.0071	25.5	0.0000	0.0051	4.9
								$\rho_0 = 0.5$								
10	10	0.0438	26.1063	0.0	−1.3524	82.6024	7.2	−0.0042	0.0621	6.1	−0.1255	0.1539	41.4	0.0119	0.0987	5.2
10	50	0.0029	0.1271	0.0	0.0415	0.2986	9.0	−0.0009	0.0343	5.1	−0.0293	0.0485	18.1	0.0001	0.0394	5.4
10	100	0.0009	0.0693	0.0	0.0143	0.1888	9.3	−0.0003	0.0257	4.8	−0.0148	0.0311	13.2	0.0001	0.0277	5.2
10	200	0.0007	0.0433	0.0	0.0088	0.1288	8.6	0.0000	0.0187	5.0	−0.0074	0.0207	9.5	0.0001	0.0194	5.1
50	10	−0.1001	35.6017	0.0	0.1275	0.7946	8.2	−0.0008	0.0299	5.2	−0.1286	0.1349	94.4	0.0085	0.0456	4.5
50	50	0.0006	0.0511	0.0	0.0091	0.1215	8.2	−0.0004	0.0158	5.2	−0.0294	0.0341	51.3	0.0000	0.0176	5.6
50	100	−0.0001	0.0298	0.0	0.0032	0.0821	8.6	−0.0003	0.0117	5.0	−0.0151	0.0194	33.9	−0.0002	0.0124	5.2
50	200	0.0000	0.0189	0.0	0.0016	0.0558	7.4	0.0000	0.0085	4.9	−0.0074	0.0115	21.9	0.0000	0.0088	5.3
100	10	−0.3243	21.2584	0.0	0.0475	0.3848	8.2	−0.0003	0.0214	5.1	−0.1285	0.1318	99.8	0.0086	0.0329	3.9
100	50	0.0003	0.0367	0.0	0.0013	0.0851	8.4	−0.0003	0.0113	5.2	−0.0291	0.0316	77.2	0.0003	0.0126	4.7
100	100	0.0002	0.0212	0.0	0.0012	0.0577	8.1	−0.0001	0.0083	5.1	−0.0148	0.0172	52.5	0.0000	0.0089	5.1
100	200	0.0002	0.0136	0.0	0.0003	0.0399	7.6	0.0000	0.0060	5.1	−0.0074	0.0097	33.3	0.0000	0.0063	5.2
200	10	−0.1975	22.3755	0.0	0.0245	0.2235	6.7	−0.0003	0.0154	5.2	−0.1289	0.1305	100.0	0.0082	0.0237	3.1
200	50	0.0003	0.0256	0.0	0.0026	0.0595	7.5	0.0000	0.0080	5.2	−0.0290	0.0303	95.7	0.0004	0.0090	5.3
200	100	0.0002	0.0149	0.0	0.0011	0.0407	7.3	0.0000	0.0059	4.7	−0.0147	0.0159	77.6	0.0002	0.0063	5.2
200	200	0.0002	0.0096	0.0	0.0008	0.0281	7.3	0.0000	0.0042	4.5	−0.0074	0.0086	52.3	0.0001	0.0044	4.9
								$\rho_0 = 0.95$								
10	10	0.0014	0.0417	0.0	0.0014	0.0623	0.3	−0.0001	0.0184	5.5	−0.0159	0.0265	19.3	0.1775	0.1791	0.0
10	50	0.0002	0.0159	0.0	0.0020	0.0684	1.3	−0.0001	0.0036	5.0	−0.0019	0.0043	12.4	0.0371	0.0373	0.0
10	100	0.0000	0.0150	0.0	0.0058	0.1007	3.1	−0.0001	0.0027	5.1	−0.0013	0.0031	11.5	0.0182	0.0184	0.0
10	200	0.0000	0.0160	0.0	0.0101	0.1537	5.4	−0.0001	0.0023	5.4	−0.0010	0.0026	11.8	0.0087	0.0090	0.0
50	10	0.0005	0.0205	0.0	0.0008	0.0327	0.2	0.0001	0.0090	5.0	−0.0195	0.0221	58.3	0.1736	0.1739	0.0
50	50	−0.0001	0.0078	0.0	0.0004	0.0384	1.8	0.0000	0.0018	4.6	−0.0023	0.0030	32.8	0.0367	0.0367	0.0
50	100	0.0000	0.0076	0.0	0.0020	0.0570	3.4	0.0000	0.0013	4.7	−0.0015	0.0021	30.3	0.0179	0.0180	0.0
50	200	0.0001	0.0084	0.0	0.0035	0.0855	4.9	0.0000	0.0011	4.6	−0.0012	0.0017	27.0	0.0085	0.0086	0.0
100	10	−0.0002	0.0146	0.0	0.0003	0.0232	0.3	−0.0001	0.0065	5.3	−0.0200	0.0213	85.1	0.1730	0.1732	0.0
100	50	−0.0001	0.0057	0.0	0.0010	0.0277	1.8	−0.0001	0.0013	5.0	−0.0023	0.0027	53.7	0.0366	0.0366	0.0
100	100	0.0000	0.0054	0.0	0.0010	0.0407	3.3	0.0000	0.0009	4.5	−0.0015	0.0018	46.6	0.0180	0.0180	0.0
100	200	−0.0001	0.0059	0.0	0.0024	0.0623	4.7	0.0000	0.0008	4.8	−0.0012	0.0015	43.0	0.0085	0.0086	0.0
200	10	−0.0002	0.0105	0.0	0.0000	0.0169	0.4	0.0000	0.0046	4.8	−0.0203	0.0210	98.8	0.1726	0.1727	0.0
200	50	0.0000	0.0040	0.0	0.0000	0.0203	1.8	0.0000	0.0009	4.9	−0.0024	0.0026	78.1	0.0366	0.0366	0.0
200	100	0.0000	0.0039	0.0	0.0003	0.0298	3.1	0.0000	0.0007	5.1	−0.0016	0.0017	71.2	0.0179	0.0179	0.0
200	200	−0.0001	0.0043	0.0	0.0003	0.0455	4.8	0.0000	0.0006	4.4	−0.0013	0.0014	68.5	0.0085	0.0085	0.0

Notes: "AHL" ("AHD"), "FA", "LS" and "BALS" refer to the Anderson and Hsiao (1981) IV estimator using lagged levels (differences) as instruments, the factor analytic estimator, the fixed effects LS estimator, and to the bias-adjusted LS estimator of Hahn and Kuersteiner (2002), respectively. "RMSE" refers to the root mean squared errors. The size results are for a nominal 5% level test. In experiment F1, $F_t = 1$.

Table 2: Bias, RMSE and 5% size results for F2 and F3 when $|\rho_0| < 1$.

		$\rho_0 = 0$			$\rho_0 = 0.5$			$\rho_0 = 0.95$		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
					F2					
10	10	0.0006	0.0859	5.5	−0.0010	0.0635	6.1	−0.0002	0.0245	6.0
10	50	0.0007	0.0431	4.6	−0.0008	0.0326	5.4	0.0000	0.0042	4.5
10	100	0.0004	0.0306	4.6	−0.0003	0.0248	5.0	0.0000	0.0027	4.9
10	200	0.0004	0.0218	4.8	−0.0001	0.0181	4.7	−0.0001	0.0020	5.3
50	10	0.0008	0.0396	4.9	0.0002	0.0293	5.3	0.0001	0.0117	5.2
50	50	0.0000	0.0192	5.0	−0.0002	0.0146	5.3	0.0000	0.0020	4.8
50	100	−0.0002	0.0139	5.5	−0.0003	0.0112	4.9	0.0000	0.0013	5.5
50	200	0.0000	0.0099	7.6	−0.0001	0.0083	5.0	0.0000	0.0009	4.7
100	10	0.0003	0.0287	5.4	0.0002	0.0212	5.6	0.0000	0.0084	5.1
100	50	0.0000	0.0136	4.8	−0.0001	0.0105	4.5	0.0000	0.0014	4.9
100	100	0.0000	0.0099	6.5	−0.0001	0.0079	5.3	0.0000	0.0009	4.9
100	200	0.0001	0.0071	7.1	0.0000	0.0059	5.2	0.0000	0.0007	5.0
200	10	−0.0001	0.0201	5.0	−0.0003	0.0147	5.2	−0.0002	0.0058	5.1
200	50	0.0000	0.0098	5.2	−0.0001	0.0075	4.9	0.0000	0.0010	4.7
200	100	0.0001	0.0071	5.9	0.0000	0.0057	5.0	0.0000	0.0006	4.6
200	200	0.0001	0.0050	6.9	0.0000	0.0042	4.8	0.0000	0.0005	5.1
					F3					
10	10	0.0018	0.1201	6.2	−0.0138	0.1191	6.5	−0.0463	0.1316	8.0
10	50	0.0010	0.0480	5.6	−0.0015	0.0429	5.9	−0.0063	0.0237	7.2
10	100	0.0005	0.0338	5.5	−0.0006	0.0299	5.8	−0.0031	0.0145	7.1
10	200	0.0006	0.0234	5.6	−0.0002	0.0207	5.6	−0.0015	0.0089	7.2
50	10	−0.0001	0.0517	5.5	−0.0039	0.0513	5.5	−0.0108	0.0460	3.2
50	50	−0.0004	0.0205	5.3	−0.0008	0.0184	5.5	−0.0010	0.0091	5.3
50	100	−0.0002	0.0145	5.7	−0.0003	0.0128	5.2	−0.0006	0.0059	4.7
50	200	0.0001	0.0103	7.4	0.0000	0.0090	5.2	−0.0003	0.0037	5.3
100	10	−0.0006	0.0367	5.0	−0.0024	0.0364	5.3	−0.0063	0.0307	3.5
100	50	0.0000	0.0145	4.8	−0.0002	0.0129	4.7	−0.0005	0.0063	4.8
100	100	−0.0001	0.0102	5.9	−0.0002	0.0089	5.0	−0.0004	0.0041	5.4
100	200	0.0001	0.0074	7.9	0.0000	0.0064	5.2	−0.0001	0.0026	5.3
200	10	0.0003	0.0256	4.6	−0.0012	0.0255	5.2	−0.0035	0.0207	4.1
200	50	0.0001	0.0105	5.7	−0.0001	0.0092	4.9	−0.0002	0.0045	4.5
200	100	0.0001	0.0074	6.6	0.0000	0.0064	5.1	−0.0001	0.0029	4.4
200	200	0.0001	0.0051	6.7	0.0000	0.0045	4.6	−0.0001	0.0018	4.7

Notes: In F2, $F_t = (1,0)'$ if $t < \lfloor T/2 \rfloor$ and $F_t = (1,1)'$ otherwise, whereas in F3, $F_t \sim N(0,1)$. See Table 1 for an explanation of the rest.

Table 3: Bias, RMSE and 5% size results for F1–F3 when $\rho_0 = 1$.

		F1			F2			F3		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
					C1					
10	10	0.0001	0.0152	5.3	−0.0001	0.0206	5.7	−0.0543	0.1430	9.0
10	50	0.0000	0.0013	4.9	0.0000	0.0018	4.5	−0.0085	0.0230	7.6
10	100	0.0000	0.0005	5.1	0.0000	0.0006	4.9	−0.0053	0.0133	9.3
10	200	0.0000	0.0002	4.3	0.0000	0.0002	5.3	−0.0026	0.0068	9.5
50	10	0.0001	0.0074	5.1	0.0001	0.0099	5.3	−0.0123	0.0483	2.7
50	50	0.0000	0.0007	4.4	0.0000	0.0008	4.8	−0.0012	0.0065	4.0
50	100	0.0000	0.0002	5.3	0.0000	0.0003	5.1	−0.0009	0.0038	3.6
50	200	0.0000	0.0001	4.3	0.0000	0.0001	4.9	−0.0004	0.0020	3.9
100	10	0.0000	0.0054	5.6	0.0000	0.0071	5.1	−0.0066	0.0299	2.5
100	50	0.0000	0.0005	4.9	0.0000	0.0006	4.5	−0.0006	0.0043	4.0
100	100	0.0000	0.0002	4.3	0.0000	0.0002	4.7	−0.0004	0.0025	3.8
100	200	0.0000	0.0001	4.7	0.0000	0.0001	4.5	−0.0002	0.0013	3.5
200	10	0.0000	0.0038	5.1	−0.0001	0.0049	4.9	−0.0036	0.0195	3.5
200	50	0.0000	0.0003	4.7	0.0000	0.0004	4.6	−0.0002	0.0030	4.2
200	100	0.0000	0.0001	4.7	0.0000	0.0002	4.7	−0.0002	0.0016	3.9
200	200	0.0000	0.0000	4.8	0.0000	0.0001	5.2	−0.0001	0.0009	3.9
					C2					
10	10	−0.0186	0.0696	3.2	−0.0294	0.1002	4.5	−0.0137	0.0588	7.6
10	50	−0.0030	0.0112	6.4	−0.0032	0.0118	6.02	−0.0031	0.0117	8.14
10	100	−0.0015	0.0054	6.5	−0.0015	0.0055	6.28	−0.0016	0.0057	8.16
10	200	−0.0007	0.0027	6.36	−0.0007	0.0027	6.36	−0.0007	0.0028	7.74
50	10	−0.0037	0.0241	4.88	−0.0043	0.0278	3.48	−0.0025	0.0218	5.44
50	50	−0.0005	0.0042	5.2	−0.0005	0.0043	5.14	−0.0004	0.0041	5.42
50	100	−0.0002	0.0021	5.38	−0.0003	0.0021	5.34	−0.0003	0.0021	5.74
50	200	−0.0001	0.0010	5.48	−0.0001	0.0010	5.56	−0.0001	0.0010	5.78
100	10	−0.0019	0.0168	5.46	−0.0025	0.0191	4.4	−0.0014	0.0152	5.38
100	50	−0.0003	0.0029	5.58	−0.0003	0.0030	5.32	−0.0003	0.0029	5.58
100	100	−0.0001	0.0015	5.64	−0.0001	0.0015	5.38	−0.0001	0.0014	5.58
100	200	−0.0001	0.0007	5.18	−0.0001	0.0007	5.02	−0.0001	0.0007	5.28
200	10	−0.0006	0.0117	4.8	−0.0009	0.0131	4.08	−0.0004	0.0108	5.5
200	50	−0.0001	0.0021	5.06	−0.0001	0.0021	4.78	−0.0001	0.0020	5.42
200	100	0.0000	0.0010	4.7	0.0000	0.0010	4.86	0.0000	0.0010	4.88
200	200	0.0000	0.0005	5.1	0.0000	0.0005	5.1	0.0000	0.0005	5.28

Notes: In C1, the common component is generated as $c_{i,t} = \lambda_i' F_t$, whereas in C2, $c_{i,t} = \lambda_i'(F_t - \rho F_{t-1})$ for $t = 2, ..., T$ and $c_{i,1} = \lambda_i' F_1$. See Tables 1 and 2 for an explanation of the rest.