



INTERNATIONAL ASSOCIATION FOR RESEARCH AND TEACHING  
Economics, Finance, Operations Research, Econometrics and Statistics

++ research ++ teaching ++

## ECORE DISCUSSION PAPER

2010/49

### **Rank-based Optimal Tests for Random Effects in Panel Data**

Nezar BENNALA  
Marc HALLIN  
Davy PAINDAVEINE

# RANK-BASED OPTIMAL TESTS FOR RANDOM EFFECTS IN PANEL DATA

Nezar BENNALA<sup>a</sup>, Marc HALLIN<sup>a,b,c,\*</sup>  
and  
Davy PAINDAVEINE<sup>a,d,\*</sup>

<sup>a</sup>Institut de Recherche en Statistique, E.C.A.R.E.S., Université Libre de Bruxelles,

<sup>b</sup>Académie Royale de Belgique,

<sup>c</sup>CentER, Tilburg University, and

<sup>d</sup>Département de Mathématique, Université Libre de Bruxelles.

## Abstract

We consider the problem of detecting unobserved heterogeneity, that is, the problem of testing the absence of random individual effects in a  $n \times T$  panel. We establish a local asymptotic normality property (with respect to intercept, regression coefficient, scale parameter  $\sigma^2$ , and the parameter of interest  $\sigma_u^2$ , for fixed density  $f_1$ ), when  $n$  tend to infinity and  $T$  is fixed. This result allows for developing asymptotically optimal parametric procedures for  $\sigma_u^2$  under specified densities  $f_1$ . The pseudo-Gaussian tests (optimal under Gaussian densities but valid under non-Gaussian ones) are investigated. Rank-based versions of the optimal parametric procedures are also provided. These tests are locally asymptotically optimal at correctly specified innovation densities. The limiting distribution of our test statistics is obtained both under the null and under sequences of contiguous alternatives. A local asymptotic linearity property is established in order to control for the effect of substituting estimated values for nuisance parameters. The asymptotic relative efficiencies of the proposed procedures with respect to the corresponding pseudo-Gaussian parametric tests are derived, and small-sample performances are investigated via a Monte-Carlo study.

**Key words and phrases :** Random effects, panel data, rank tests, local asymptotic normality.

## 1 Introduction.

### 1.1 Testing for random effects in panel data.

Panel data consist of a series of  $T$  observations made through time over a number  $n$  of cross-sectional items or experimental units. By combining cross-sectional and time-series features,

---

\*The research of Marc Hallin is supported by the Sonderforschungsbereich “Statistical modelling of nonlinear dynamic processes” (SFB 823) of the Deutsche Forschungsgemeinschaft, and by a Discovery Grant of the Australian Research Council. The research of Davy Paindaveine is supported by the Banque Nationale de Belgique. Marc Hallin and Davy Paindaveine are also members of ECORE, the association between CORE and ECARES.

panel data methods are able to identify and estimate effects that are not detectable via pure cross-sectional or time-series methods, while controlling for individual heterogeneity and taking into account dynamics; see Baltagi (2005) for background reading. Throughout, we consider the model

$$Y_{it} = \mu + \boldsymbol{\beta}' \mathbf{x}_{it} + \nu_{it} \quad i = 1, \dots, n, \quad t = 1, \dots, T, \quad (1.1)$$

with  $i$  denoting individuals or experimental units and  $t$  denoting time;  $Y_{it}$  is the observed response for individual  $i$  at time  $t$ ,  $\mathbf{x}_{it}$  is a vector of  $K$  nonstochastic exogenous regressors, and  $(\mu, \boldsymbol{\beta})'$  is a  $(K+1) \times 1$  unknown regression parameter. We will assume that the disturbances  $\nu_{it}$  follow a one-way error component model

$$\nu_{it} = u_i + \varepsilon_{it}, \quad (1.2)$$

where  $u_i$  denotes unobserved individual-specific random effect and  $\varepsilon_{it}$  is the error term. We assume that the  $u_i$ 's are i.i.d.  $(0, \sigma_u^2)$ , that the  $\varepsilon_{it}$ 's are i.i.d.  $(0, \sigma^2)$ , and that  $u_i$  and  $\varepsilon_{jt}$  are mutually independent for all  $i, j$ , and  $t$ . Note that the individual effects  $u_i$  are constant over time, and account for autocorrelation among the data, with  $\text{corr}(\nu_{it}, \nu_{is}) = \sigma_u^2 / (\sigma_u^2 + \sigma^2)$  for  $t \neq s$ . This random effect model is closer to repeated measurement than to time series models, as  $\text{corr}(\nu_{it}, \nu_{is})$  remains the same irrespective of the time lag  $t - s$ : for given  $i$ , the  $\nu_{it}$ 's thus are not mixing nor weakly dependent in any sense. Moreover, this autocorrelation is the same for all individuals.

Failing to take that autocorrelation structure into account when estimating the model, e.g. by ordinary least squares (OLS), may lead to seriously biased estimators and invalid testing procedures (Scott and Holt 1982; Moulton 1986). Therefore, it is important to be able to perform a preliminary test of the hypothesis that the variance component  $\sigma_u^2$  for individual effect is zero, that is,  $\mathcal{H}_0 : \sigma_u^2 = 0$ , versus the alternative  $\mathcal{H}_1 : \sigma_u^2 > 0$ .

A popular method for deriving locally optimal tests is the Lagrange Multiplier method based on Gaussian likelihoods. That approach has been considered by Breusch and Pagan (1980), who propose a test of  $\mathcal{H}_0$  based on the (“two-sided”) Lagrange Multiplier test statistic,

$$T_{\text{Breusch-Pagan}}^{(n,T)} := \frac{nT}{2(T-1)} \left[ \frac{\sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \tilde{\nu}_{it} \tilde{\nu}_{il}}{\sum_{i=1}^n \sum_{t=1}^T \tilde{\nu}_{it}^2} \right]^2,$$

where  $\tilde{\nu}_{it}$  denotes the OLS residuals obtained under  $\mathcal{H}_0$ . This statistic is asymptotically  $\chi_1^2$  (as both  $n \rightarrow \infty$  and  $T \rightarrow \infty$ ) under  $\mathcal{H}_0$ , and is easy to compute, as it only requires OLS residuals. Honda (1985) and King and Evans (1986) observed that a “one-sided” Lagrange Multiplier test, based on the asymptotically standard normal distribution of

$$T_{\text{Honda}}^{(n,T)} := \sqrt{\frac{nT}{2(T-1)}} \left[ \frac{\sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \tilde{\nu}_{it} \tilde{\nu}_{il}}{\sum_{i=1}^n \sum_{t=1}^T \tilde{\nu}_{it}^2} \right],$$

is sizeably more powerful in this inherently one-sided problem. Honda (1985) also established the robustness to nonnormal errors of the asymptotic normality (as  $n \rightarrow \infty$  and  $T \rightarrow \infty$ ) of  $T_{\text{Honda}}^{(n,T)}$ ; unfortunately, as shown by Moulton and Randolph (1989), the finite-sample performance of tests based on that normal approximation can be pretty poor, even in fairly large samples. This occurs if either there are many regressors or the regressors are highly correlated within groups. They suggest an alternative standardized Lagrange Multiplier statistic, whose asymptotic critical values are generally much closer to the exact critical values than those of the original one. An

alternative derivation of the Breusch and Pagan statistic has been obtained by Chesher (1984). Lagrange multiplier tests also have been derived by Hamerle (1990), Orme (1993), Jacqmin-Gadda and Commenges (1995), and Lin (1997) for generalized linear model versions of (1.1).

Small  $T$  values, however, are the rule rather than the exception in this context. In a fixed- $T$  setup, Wooldridge (2002) recently proposed the test statistic

$$T_{\text{Wooldridge}}^{(n)} := \left( \sum_{i=1}^n \sum_{t=2}^T \sum_{l=1}^{t-1} \tilde{\nu}_{it} \tilde{\nu}_{il} \right) / \left[ \sum_{i=1}^n \left( \sum_{t=2}^T \sum_{l=1}^{t-1} \tilde{\nu}_{it} \tilde{\nu}_{il} \right)^2 \right]^{1/2},$$

which is asymptotically standard normal (as  $n \rightarrow \infty$  under fixed  $T$ ). As pointed out by Wooldridge (2002, p. 265) himself, that statistic can be interpreted as a test statistic for the problem of detecting serial correlation among the  $\nu_{it}$ 's.

While none of the previous tests require specifying the distribution of random effects, they still are of a parametric nature, and require finite moments of order two. One way of avoiding such assumptions consists in basing the tests on statistics that are measurable with respect to invariant or distribution-free quantities such as ranks. The main objective of this paper is to construct, for the problem of testing the null hypothesis  $\mathcal{H}_0$  of no random effects, rank-based tests which are asymptotically (as  $n \rightarrow \infty$ , under fixed  $T$ ) distribution-free and hence remain (asymptotically) valid under extremely mild assumptions—in particular, in the absence of any moment assumptions—while retaining optimality (in the Le Cam sense) at correctly specified densities.

For instance, the normal-score or van der Waerden rank test rejecting the null hypothesis of no random effects for large values of

$$\tilde{T}_{\text{vdW}}^{*(n)}(\hat{\boldsymbol{\beta}}) := \frac{1}{2s_{\phi_1}^{(n)} \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \left\{ a \Phi^{-1} \left( \frac{R_{it}^{(n)}(\boldsymbol{\beta})}{N+1} \right) \Phi^{-1} \left( \frac{R_{il}^{(n)}(\boldsymbol{\beta})}{N+1} \right) - c_{\phi_1}^{(N)} \right\},$$

where  $\Phi$  is the standard normal distribution function,  $\hat{\boldsymbol{\beta}}$  an appropriate estimator of  $\boldsymbol{\beta}$ ,

$$N := nT, \quad c_{\phi_1}^{(N)} := \frac{a}{N(N-1)} \sum_{r=1}^N \sum_{s \neq r=1}^N \Phi^{-1} \left( \frac{r}{N+1} \right) \Phi^{-1} \left( \frac{s}{N+1} \right),$$

$R_{it}^{(n)}(\boldsymbol{\beta})$  stands for the rank of  $Z_{it} := Y_{it} - \mu - \boldsymbol{\beta}' \mathbf{x}_{it}$  among  $Z_{11}, \dots, Z_{nT}$  (this rank does not depend on  $\mu$ , which justifies the notation),  $a \approx 0.4549$  (see (2.2)) and  $s_{\phi_1}^{(n)}$  is defined in (4.2), is distribution-free under the hypothesis of no random effects, asymptotically optimal against Gaussian local alternatives, and asymptotically equivalent (as  $n \rightarrow \infty$ , under fixed  $T$ ) to the Honda test under Gaussian densities. Although asymptotic relative efficiency (ARE) comparisons between this van der Waerden test and Honda's classical one are somewhat unfair to van der Waerden (since  $T_{\text{Honda}}^{(n,T)}$  requires  $T \rightarrow \infty$  while  $\tilde{T}_{\text{vdW}}^{*(n)}$  does not), the AREs (see Section 4.4) of  $\tilde{T}_{\text{vdW}}^{*(n)}$ -based tests with respect to the  $T_{\text{Honda}}^{(n,T)}$ -based ones are strictly larger than one under a very broad range of densities. These theoretical findings are confirmed (Section 5) by finite-sample simulations.

## 1.2 Outline of the paper.

The paper is organized as follows. In Section 2.1, we collect the main assumptions needed in the sequel. Section 2.2 states the uniform local asymptotic normality (ULAN) result (for fixed

density  $f_1$  of the  $\varepsilon_{it}$ 's) on which our construction (Section 3.1) of locally and asymptotically optimal tests is based. Due to the mixture nature of likelihoods under random individual effects, establishing that ULAN result is particularly delicate. The special case of the pseudo-Gaussian tests (optimal under Gaussian densities but valid under non-Gaussian ones) is investigated in Section 3.2. In Section 4.1, we construct rank-based versions of the central sequences appearing in the ULAN result and, in Section 4.2, derive the corresponding rank-based tests for the absence of random effects. Some special cases (van der Warden and Wilcoxon scores) of the proposed rank-based statistics are considered in Section 4.3. Asymptotic relative efficiencies with respect to the pseudo-Gaussian tests are derived in Section 4.4. Section 5 provides some simulation results assessing the finite-sample performance of the various tests proposed. Finally, the appendix collects the proofs of ULAN and other technical results.

## 2 Uniform local asymptotic normality.

### 2.1 Notation and technical assumptions.

Denote by  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, \sigma_u^2; f_1, h_1}^{(n)}$  the probability distribution of the observation  $(\mathbf{Y}_1^{(n)'}, \mathbf{Y}_2^{(n)'}, \dots, \mathbf{Y}_n^{(n)'})'$ , where  $\mathbf{Y}_i^{(n)} := (Y_{i1}, \dots, Y_{iT})'$  is generated by

$$Y_{it} = \mu + \boldsymbol{\beta}' \mathbf{x}_{it} + u_i + \varepsilon_{it} = \boldsymbol{\theta}' \underline{\mathbf{x}}_{it} + u_i + \varepsilon_{it}, \quad i = 1, \dots, n, \quad t = 1, \dots, T,$$

where  $\mathbf{x}_{it} = (x_{it}^1, x_{it}^2, \dots, x_{it}^K)'$  is the  $K$ -tuple of explanatory variables for individual  $i$  at time  $t$ ,  $\underline{\mathbf{x}}_{it} := (1, \mathbf{x}_{it}')'$ , and  $\boldsymbol{\theta} := (\mu, \boldsymbol{\beta}')' \in \mathbb{R}^{K+1}$ ;  $\{u_i, i = 1, \dots, n\}$  is an unobservable i.i.d. sequence of individual random effects with mean zero, variance  $\sigma_u^2$ , and density  $u \mapsto h(u) := (1/\sigma_u)h_1(u/\sigma_u)$ ;  $\{\varepsilon_{it}, i = 1, \dots, n, t = 1, \dots, T\}$  is another unobservable i.i.d. sequence, with density  $\varepsilon \mapsto f(\varepsilon) := (1/\sigma)f_1(\varepsilon/\sigma)$ , for some scale parameter  $\sigma \in \mathbb{R}_0^+$  and  $f_1$  in the class of standardized densities

$$\mathcal{F}_0 := \left\{ f_1 : \int_{-\infty}^0 f_1(z) dz = 0.5 = \int_{-1}^1 f_1(z) dz \right\}.$$

Under  $f_1 \in \mathcal{F}_0$ , the  $\varepsilon_{it}$ 's therefore have median zero and median absolute deviation  $\sigma$ ; this standardization which, contrary to the usual one based on the mean and the standard error, avoids all moment assumptions, plays the role of an identification constraint, and has no impact on subsequent results. Finally, the individual effects  $u_i$  and the disturbances  $\varepsilon_{jt}$  are assumed to be mutually independent for all  $i, j$ , and  $t$ .

Clearly, whenever  $\sigma_u = 0$ , the probability density  $h$  of individual effects has no impact on the distribution of the observation; we emphasize that fact by writing  $\mathbb{P}_{\boldsymbol{\theta}; f_1}^{(n)}$  or  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; f_1}^{(n)}$  instead of  $\mathbb{P}_{\boldsymbol{\theta}; f_1, h_1}^{(n)}$  whenever  $\sigma_u = 0$ .

Our derivation of locally asymptotically optimal tests at density  $f_1$  will be based on the uniform local and asymptotic normality (ULAN) with respect to  $(\boldsymbol{\theta}', \sigma^2, \sigma_u^2)'$ , at  $(\boldsymbol{\theta}', \sigma^2, 0)'$ , of the families of distributions

$$\mathcal{P}_{f_1, h_1}^{(n)} := \left\{ \mathbb{P}_{\boldsymbol{\theta}, \sigma^2, \sigma_u^2; f_1, h_1}^{(n)} : \boldsymbol{\theta} \in \mathbb{R}^{K+1}, \sigma^2 > 0 \text{ and } \sigma_u^2 \geq 0 \right\}.$$

This ULAN property requires some technical assumptions on the innovation density  $f_1$ , the asymptotic behavior of the regressors, and the density function  $h$  of individual random effects.

ASSUMPTION (A). The density  $f_1$  is such that

(A1)  $f_1 \in \mathcal{F}_0$ ;

(A2)  $f_1(z) > 0$  for all  $z \in \mathbb{R}$ ;

(A3)  $z \mapsto f_1(z)$  is  $C^2$  on  $\mathbb{R}$ , with derivatives  $\dot{f}_1$  and  $\ddot{f}_1$ ; letting  $\phi_{f_1} := -\dot{f}_1/f_1$  and  $\psi_{f_1} := \ddot{f}_1/f_1$ ,

$$\mathcal{I}_\phi(f_1) := \int_{\mathbb{R}} \phi_{f_1}^2(z) f_1(z) dz < \infty, \quad \mathcal{I}_\psi(f_1) := \int_{\mathbb{R}} \psi_{f_1}^2(z) f_1(z) dz < \infty,$$

and

$$\mathcal{J}_\phi(f_1) := \int_{\mathbb{R}} z^2 \phi_{f_1}^2(z) f_1(z) dz < \infty,$$

which entails

$$\mathcal{I}_{\phi\psi}(f_1) := \int_{\mathbb{R}} \psi_{f_1}(z) \phi_{f_1}(z) f_1(z) dz < \infty \quad \text{and} \quad \mathcal{K}_{\phi\psi}(f_1) := \int_{\mathbb{R}} z \psi_{f_1}(z) \phi_{f_1}(z) f_1(z) dz < \infty.$$

The set of all densities satisfying Assumption (A) will be denoted as  $\mathcal{F}_A$ . It should be stressed that none of these assumptions requires the existence of any moment for the density  $f_1$ . They are satisfied, for instance, for all Student distributions with  $\nu > 0$  degrees of freedom, with standardized densities  $f_1(z) = f_{1,t\nu}(z) := C_\nu \sqrt{a_\nu} (1 + a_\nu z^2/\nu)^{-(1+\nu)/2}$ , for which

$$\mathcal{I}_\phi(f_1) = a_\nu \frac{(\nu+1)}{(\nu+3)}, \quad \mathcal{I}_\psi(f_1) = 2a_\nu^2 \left( \frac{\nu^4 + 8\nu^3 + 27\nu^2 + 40\nu + 20}{\nu^4 + 15\nu^3 + 71\nu^2 + 105\nu} \right),$$

$$\mathcal{J}_\phi(f_1) = 3 \frac{(\nu+1)}{(\nu+3)}, \quad \mathcal{I}_{\phi\psi}(f_1) = 0, \quad \text{and} \quad \mathcal{K}_{\phi\psi}(f_1) = 2a_\nu \frac{(\nu+1)(\nu+2)}{(\nu+3)(\nu+5)}. \quad (2.1)$$

The same remark holds for the logistic distribution, with standardized density  $f_1(z) = \ell_1(z) := \sqrt{b} \exp(-\sqrt{b}z) / (1 + \exp(-\sqrt{b}z))^2$ , for which  $\mathcal{I}_\phi(f_1) = b/3$ ,  $\mathcal{I}_\psi(f_1) = b^2/5$ ,  $\mathcal{J}_\phi(f_1) = (12 + \pi^2)/9$ ,  $\mathcal{I}_{\phi\psi}(f_1) = 0$ , and  $\mathcal{K}_{\phi\psi}(f_1) = b/2$ . The corresponding values for the Gaussian distribution, with standardized density  $f_1(z) = \phi_1(z) := \sqrt{a/2\pi} \exp(-az^2/2)$ , can be obtained by taking limits as  $\nu \rightarrow \infty$  of the Student information quantities in (2.1):

$$\mathcal{I}_\phi(f_1) = a \approx 0.4549, \quad \mathcal{I}_\psi(f_1) = 2a^2, \quad \mathcal{J}_\phi(f_1) = 3, \quad \mathcal{I}_{\phi\psi}(f_1) = 0, \quad \text{and} \quad \mathcal{K}_{\phi\psi}(f_1) = 2a. \quad (2.2)$$

The normalizing constants  $C_\nu$ ,  $a_\nu > 0$ ,  $a > 0$  and  $b > 0$  are such that  $f_1 \in \mathcal{F}_0$ .

ASSUMPTION (B). Let  $\mathbf{C}^{(n)} := \frac{1}{n} \sum_{i=1}^n \sum_{t=1}^T \mathbf{x}_{it} \mathbf{x}'_{it}$ . Denoting by  $\mathbf{D}^{(n)}$  the diagonal matrix with elements  $(\mathbf{C}^{(n)})_{11}, \dots, (\mathbf{C}^{(n)})_{KK}$ , define  $\mathbf{R}^{(n)} := (\mathbf{D}^{(n)})^{-1/2} \mathbf{C}^{(n)} (\mathbf{D}^{(n)})^{-1/2}$ .

(B1) The limit  $\lim_{n \rightarrow \infty} \mathbf{R}^{(n)} =: \mathbf{R}$  exists, is positive definite, and therefore factorizes into  $\mathbf{R} = (\mathbf{K}\mathbf{K}')^{-1}$  for some full-rank  $K \times K$  matrix  $\mathbf{K}$ . Letting  $\mathbf{K}^{(n)} := (\mathbf{D}^{(n)})^{-1/2} \mathbf{K}$ , note that  $\mathbf{K}^{(n)}$  is also of full rank.

(B2) The classical Noether conditions hold: letting  $\bar{x}_k^{(n)} := \frac{1}{nT} \sum_{i,t} (\mathbf{x}_{it})_k$ ,

$$\lim_{n \rightarrow \infty} \frac{\max_{1 \leq i \leq n} \left( (\mathbf{x}_{it})_k - \bar{x}_k^{(n)} \right)^2}{\sum_{i=1}^n \sum_{t=1}^T \left( (\mathbf{x}_{it})_k - \bar{x}_k^{(n)} \right)^2} = 0, \quad k = 1, \dots, K, \quad t = 1, \dots, T.$$

Note that the Noether conditions also imply that

$$\lim_{n \rightarrow \infty} \frac{\max_{1 \leq i \leq n} (\mathbf{x}_{it})_k^2}{\sum_{i=1}^n \sum_{t=1}^T (\mathbf{x}_{it})_k^2} = 0, \quad k = 1, \dots, K, \quad t = 1, \dots, T.$$

ASSUMPTION (C).

$$(C1) \quad \int_{\mathbb{R}} u h_1(u) du = 0 \text{ and } \int_{\mathbb{R}} u^2 h_1(u) du = 1.$$

(C2) Letting  $\mathbf{z} := (z_1, \dots, z_T)'$ ,  $K_{\mathbf{z}}(u, y) := \prod_{t=1}^T f_1(z_t - yu)$ , and, for  $y > 0$ ,  $\ddot{K}_{\mathbf{z}}(u, y) := \partial^2 K_{\mathbf{z}}(u, y) / \partial y^2$ , the Fisher information associated with  $\sigma_u$

$$y \mapsto \mathcal{I}_{\phi\psi}(f_1, y) := \begin{cases} \frac{1}{y^2} \int_{\mathbb{R}^T} \frac{\left[ \int_{w=0}^y \int_{\mathbb{R}} \ddot{K}_{\mathbf{z}}(u, w) h_1(u) dudw \right]^2}{\int_{\mathbb{R}} \prod_{t=1}^T f_1(z_t - yu) h_1(u) du} d\mathbf{z} & \text{if } y > 0 \\ T\mathcal{I}_{\psi}(f_1) + 2T(T-1)\mathcal{I}_{\phi}^2(f_1) & \text{if } y = 0, \end{cases} \quad (2.3)$$

is continuous from the right at  $y = 0$ .

Assumption (C2) actually is an assumption involving the couple of densities  $(f_1, h_1)$ . For all  $f_1 \in \mathcal{F}_A$ , let

$$\mathcal{F}_{C|f_1} := \left\{ h_1 \mid h_1 \text{ and } (f_1, h_1) \text{ satisfy Assumptions (C1) and (C2), respectively} \right\}.$$

## 2.2 ULAN.

As mentioned above, we first construct tests that are optimal at correctly specified densities, in the sense of Le Cam's asymptotic theory of statistical experiments; these tests settle the optimality bounds our rank-based tests are to compete with. In this section we establish the *uniform local asymptotic normality* (ULAN) result (with respect to intercept, regression coefficient, scale parameter  $\sigma^2$ , and the parameter of interest  $\sigma_u^2$ , for fixed density  $f_1$ ) on which optimality will be based.

Letting  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$ , consider sequences of *local alternatives* of the form  $\boldsymbol{\vartheta} + n^{-\frac{1}{2}} \boldsymbol{\nu}^{(n)} \boldsymbol{\tau}^{(n)}$ , where

$$\boldsymbol{\nu}^{(n)} := \begin{pmatrix} \boldsymbol{\nu}_1^{(n)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{pmatrix}, \quad \text{with } \boldsymbol{\nu}_1^{(n)} := \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{K}^{(n)} \end{pmatrix},$$

and  $\boldsymbol{\tau}^{(n)} := (\tau_1^{(n)}, \tau_2^{(n)'}, \tau_3^{(n)}, \tau_4^{(n)'})' \in \mathbb{R}^{K+2} \times \mathbb{R}_0^+$  (indeed,  $\tau_4^{(n)}$  is intrinsically nonnegative) is such that  $\sup_n \boldsymbol{\tau}^{(n)'} \boldsymbol{\tau}^{(n)} < \infty$ . Defining

$$Z_{it} = Z_{it}(\boldsymbol{\theta}, \sigma^2) := \sigma^{-1} (Y_{it} - \mu - \boldsymbol{\beta}' \mathbf{x}_{it}), \quad i = 1, \dots, n, \quad t = 1, \dots, T,$$

note that, under  $P_{\boldsymbol{\theta}, \sigma^2, 0; f_1}^{(n)}$ ,  $\sigma Z_{it}(\boldsymbol{\theta}, \sigma^2)$  coincides with  $\varepsilon_{it}$ . We have the following result (see the Appendix for a proof).

**Proposition 2.1** *Let Assumptions (B) and (C) hold. Fix  $f_1 \in \mathcal{F}_A$  and  $h_1 \in \mathcal{F}_{C|f_1}$ . Then, the family  $\mathcal{P}_{f_1, h_1}^{(n)}$  is ULAN (for  $n \rightarrow \infty$  with fixed  $T$ ) at any  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$ , with central sequence*

$$\Delta_{f_1}^{(n)}(\boldsymbol{\vartheta}) := \begin{pmatrix} \Delta_{f_1;1}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;2}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;3}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;4}^{(n)}(\boldsymbol{\vartheta}) \end{pmatrix} := \begin{pmatrix} \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \phi_{f_1}(Z_{it}) \\ \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \phi_{f_1}(Z_{it})(\mathbf{K}^{(n)})' \mathbf{x}_{it} \\ \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (Z_{it}\phi_{f_1}(Z_{it}) - 1) \\ \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \left\{ \sum_{t=1}^T \psi_{f_1}(Z_{it}) + \sum_{t=1}^T \sum_{l \neq t=1}^T \phi_{f_1}(Z_{it})\phi_{f_1}(Z_{il}) \right\} \end{pmatrix} \quad (2.4)$$

and full-rank information matrix

$$\mathbf{\Gamma}_{f_1}(\boldsymbol{\vartheta}) := \begin{pmatrix} \Gamma_{f_1;11}(\boldsymbol{\vartheta}) & \mathbf{0} & \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \Gamma_{f_1;14}(\boldsymbol{\vartheta}) \\ \mathbf{0} & \Gamma_{f_1;22}(\boldsymbol{\vartheta}) & \mathbf{0} & \mathbf{0} \\ \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \mathbf{0} & \Gamma_{f_1;33}(\boldsymbol{\vartheta}) & \Gamma_{f_1;34}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;14}(\boldsymbol{\vartheta}) & \mathbf{0} & \Gamma_{f_1;34}(\boldsymbol{\vartheta}) & \Gamma_{f_1;44}(\boldsymbol{\vartheta}) \end{pmatrix}, \quad (2.5)$$

where

$$\Gamma_{f_1;11}(\boldsymbol{\vartheta}) := \frac{T}{\sigma^2} \mathcal{I}_\phi(f_1), \quad \Gamma_{f_1;22}(\boldsymbol{\vartheta}) := \frac{1}{\sigma^2} \mathcal{I}_\phi(f_1) \mathbf{I}_K, \quad \Gamma_{f_1;33}(\boldsymbol{\vartheta}) := \frac{T}{4\sigma^4} (\mathcal{J}_\phi(f_1) - 1),$$

$$\Gamma_{f_1;13}(\boldsymbol{\vartheta}) := \frac{T}{2\sigma^3} \mathcal{K}_{\phi\phi}(f_1), \quad \Gamma_{f_1;14}(\boldsymbol{\vartheta}) := \frac{T}{2\sigma^3} \mathcal{I}_{\phi\psi}(f_1), \quad \Gamma_{f_1;34}(\boldsymbol{\vartheta}) := \frac{T}{4\sigma^4} \mathcal{K}_{\phi\psi}(f_1)$$

and

$$\Gamma_{f_1;44}(\boldsymbol{\vartheta}) := \frac{T}{4\sigma^4} \left\{ \mathcal{I}_\psi(f_1) + 2(T-1)\mathcal{I}_\phi^2(f_1) \right\}.$$

More precisely, for any  $\boldsymbol{\vartheta}^{(n)} = (\boldsymbol{\mu}^{(n)}, \boldsymbol{\beta}^{(n)'}, \sigma^{2(n)}, 0)'$  such that  $\boldsymbol{\mu}^{(n)} - \boldsymbol{\mu} = O(n^{-1/2})$ ,  $(\mathbf{K}^{(n)})^{-1}(\boldsymbol{\beta}^{(n)} - \boldsymbol{\beta}) = O(n^{-1/2})$  and  $\sigma^{2(n)} - \sigma^2 = O(n^{-1/2})$ , and any bounded sequence  $\boldsymbol{\tau}^{(n)} \in \mathbb{R}^{K+2} \times \mathbb{R}_0^+$ , we have

$$\begin{aligned} \Lambda_{\boldsymbol{\vartheta}^{(n)} + n^{-\frac{1}{2}} \boldsymbol{\nu}^{(n)} \boldsymbol{\tau}^{(n)} / \boldsymbol{\vartheta}^{(n)}; f_1, h_1}^{(n)} &:= \log \left( d\mathbb{P}_{\boldsymbol{\vartheta}^{(n)} + n^{-\frac{1}{2}} \boldsymbol{\nu}^{(n)} \boldsymbol{\tau}^{(n)}; f_1, h_1}^{(n)} / d\mathbb{P}_{\boldsymbol{\vartheta}^{(n)}; f_1}^{(n)} \right) \\ &= \boldsymbol{\tau}^{(n)'} \Delta_{f_1}^{(n)}(\boldsymbol{\vartheta}^{(n)}) - \frac{1}{2} \boldsymbol{\tau}^{(n)'} \mathbf{\Gamma}_{f_1}(\boldsymbol{\vartheta}) \boldsymbol{\tau}^{(n)} + o_{\mathbb{P}}(1) \end{aligned}$$

and

$$\Delta_{f_1}^{(n)}(\boldsymbol{\vartheta}^{(n)}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{f_1}(\boldsymbol{\vartheta})), \quad \boldsymbol{\vartheta} = (\boldsymbol{\mu}, \boldsymbol{\beta}', \sigma^2, 0)',$$

under  $\mathbb{P}_{\boldsymbol{\vartheta}^{(n)}; f_1}^{(n)}$ , as  $n \rightarrow \infty$  with fixed  $T$ .

Note the non-diagonal form of the information matrix  $\mathbf{\Gamma}_{f_1}(\boldsymbol{\vartheta})$ , which implies that the intercept parameter  $\boldsymbol{\mu}$  and the scale parameters  $\sigma^2$  and  $\sigma_u^2$  are not mutually information-orthogonal. The notation  $\Delta_{f_1}^{(n)}(\boldsymbol{\vartheta})$  and  $\mathbf{\Gamma}_{f_1}(\boldsymbol{\vartheta})$  emphasizes the fact that the central sequence and the information matrix do not depend on  $h_1$ .

The Gaussian versions of (2.4) and (2.5) are

$$\Delta_{\mathcal{N}}^{(n)}(\boldsymbol{\vartheta}) = \begin{pmatrix} \frac{a}{\sigma\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T Z_{it} \\ \frac{a}{\sigma\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T Z_{it} (\mathbf{K}^{(n)})' \mathbf{x}_{it} \\ \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (a Z_{it}^2 - 1) \\ \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \left\{ \sum_{t=1}^T (a^2 Z_{it}^2 - a) + a^2 \sum_{t=1}^T \sum_{l \neq t=1}^T Z_{it} Z_{il} \right\} \end{pmatrix}$$

and

$$\Gamma_{\mathcal{N}}(\boldsymbol{\vartheta}) = \begin{pmatrix} a\sigma^{-2}T & \mathbf{0} & 0 & 0 \\ \mathbf{0} & a\sigma^{-2}\mathbf{I}_K & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & \frac{1}{2}\sigma^{-4}T & \frac{1}{2}a\sigma^{-4}T \\ 0 & \mathbf{0} & \frac{1}{2}a\sigma^{-4}T & \frac{1}{2}a^2\sigma^{-4}T^2 \end{pmatrix},$$

respectively.

### 3 Optimal parametric tests and pseudo-Gaussian tests.

We are interested in testing the null hypothesis

$$\mathcal{H}_0^{(n)} := \bigcup_{g_1 \in \mathcal{F}_0} \mathcal{H}_{g_1}^{(n)} := \bigcup_{g_1 \in \mathcal{F}_0} \bigcup_{\boldsymbol{\theta} \in \mathbb{R}^{K+1}} \bigcup_{\sigma^2 > 0} \{P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}\},$$

optimality being sought against alternatives of the form

$$\bigcup_{\boldsymbol{\theta} \in \mathbb{R}^{K+1}} \bigcup_{\sigma^2 > 0} \bigcup_{\sigma_u^2 > 0} \bigcup_{h_1 \in \mathcal{F}_{C|f_1}} \{P_{\boldsymbol{\theta}, \sigma^2, \sigma_u^2; f_1, h_1}^{(n)}\}$$

for some fixed ‘‘target’’ density  $f_1 \in \mathcal{F}_A$ . The parameters  $\mu, \boldsymbol{\beta}$ , and  $\sigma^2$  thus are nuisance parameters, while  $\sigma_u^2$  is the parameter of interest. The asymptotic covariances, in (2.5), between the  $\mu$ -part  $\Delta_{f_1;1}^{(n)}(\boldsymbol{\vartheta})$ , the  $\sigma^2$ -part  $\Delta_{f_1;3}^{(n)}(\boldsymbol{\vartheta})$  and the  $\sigma_u^2$ -part  $\Delta_{f_1;4}^{(n)}(\boldsymbol{\vartheta})$  of (2.4) are not zero. This implies, via Le Cam’s Third Lemma, that a local perturbation of  $\mu$  or  $\sigma^2$  has the same asymptotic impact on  $\Delta_{f_1;4}^{(n)}(\boldsymbol{\vartheta})$  as a local perturbation of  $\sigma_u^2$ : hence, the cost of not knowing the actual values of  $\mu$  or  $\sigma^2$  when performing inference on  $\sigma_u^2$ , in general (the Gaussian case, as far as  $\mu$  is concerned, is an exception), will be strictly positive.

#### 3.1 Optimal parametric tests.

ULAN and the convergence to of local experiments to the Gaussian shifts

$$\begin{pmatrix} \Delta_1 \\ \Delta_3 \\ \Delta_4 \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} \Gamma_{f_1;11}(\boldsymbol{\vartheta}) & \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \Gamma_{f_1;14}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \Gamma_{f_1;33}(\boldsymbol{\vartheta}) & \Gamma_{f_1;34}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;14}(\boldsymbol{\vartheta}) & \Gamma_{f_1;34}(\boldsymbol{\vartheta}) & \Gamma_{f_1;44}(\boldsymbol{\vartheta}) \end{pmatrix} \begin{pmatrix} \tau_1 \\ \tau_3 \\ \tau_4 \end{pmatrix}, \begin{pmatrix} \Gamma_{f_1;11}(\boldsymbol{\vartheta}) & \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \Gamma_{f_1;14}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \Gamma_{f_1;33}(\boldsymbol{\vartheta}) & \Gamma_{f_1;34}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;14}(\boldsymbol{\vartheta}) & \Gamma_{f_1;34}(\boldsymbol{\vartheta}) & \Gamma_{f_1;44}(\boldsymbol{\vartheta}) \end{pmatrix} \right) \quad (3.6)$$

imply that locally optimal inference on  $\sigma_u^2$ , in the presence of unspecified  $\mu$  and  $\sigma^2$ , should be based on the residual of the regression (in (3.6)), of  $\Delta_4$  with respect to  $(\Delta_1, \Delta_3)'$ , computed at  $\Delta_{f_1;4}^{(n)}(\boldsymbol{\vartheta})$  and  $(\Delta_{f_1;1}^{(n)}(\boldsymbol{\vartheta}), \Delta_{f_1;3}^{(n)}(\boldsymbol{\vartheta}))'$ . That residual, called the *efficient central sequence* for  $\sigma_u^2$ , takes the form

$$\begin{aligned}\Delta_{f_1;4}^{*(n)}(\boldsymbol{\vartheta}) &:= \Delta_{f_1;4}^{(n)}(\boldsymbol{\vartheta}) - (\Gamma_{f_1;14}(\boldsymbol{\vartheta}), \Gamma_{f_1;34}(\boldsymbol{\vartheta})) \begin{pmatrix} \Gamma_{f_1;11}(\boldsymbol{\vartheta}) & \Gamma_{f_1;13}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \Gamma_{f_1;33}(\boldsymbol{\vartheta}) \end{pmatrix}^{-1} \begin{pmatrix} \Delta_{f_1;1}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{f_1;3}^{(n)}(\boldsymbol{\vartheta}) \end{pmatrix} \\ &= \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \phi_{f_1}(Z_{it})\phi_{f_1}(Z_{il});\end{aligned}\tag{3.7}$$

under  $\mathbb{P}_{\boldsymbol{\vartheta};f_1}^{(n)}$ , with  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$ , it is asymptotically normal, with mean zero and variance

$$\begin{aligned}\Gamma_{f_1;44}^*(\boldsymbol{\vartheta}) &:= \Gamma_{f_1;44}(\boldsymbol{\vartheta}) - (\Gamma_{f_1;14}(\boldsymbol{\vartheta}), \Gamma_{f_1;34}(\boldsymbol{\vartheta})) \begin{pmatrix} \Gamma_{f_1;11}(\boldsymbol{\vartheta}) & \Gamma_{f_1;13}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \Gamma_{f_1;33}(\boldsymbol{\vartheta}) \end{pmatrix}^{-1} \begin{pmatrix} \Gamma_{f_1;14}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1;34}(\boldsymbol{\vartheta}) \end{pmatrix} \\ &= \frac{T(T-1)}{2\sigma^4} \mathcal{I}_{\phi}^2(f_1) =: \Gamma_{f_1;44}^*(\sigma^2).\end{aligned}\tag{3.8}$$

Next, recall that a sequence  $\hat{\gamma}^{(n)}$  of estimators of  $\gamma$ , defined over a sequence of experiments  $\{\mathbb{P}_{\gamma}^{(n)} \mid \gamma \in \Gamma\}$ , is  $n^{1/2}(\nu^{(n)})^{-1}$ -consistent and asymptotically discrete if, under  $\mathbb{P}_{\gamma}^{(n)}$ , as  $n \rightarrow \infty$ ,

$$(D1) \quad n^{1/2}(\nu^{(n)})^{-1}(\hat{\gamma}^{(n)} - \gamma) = O_{\mathbb{P}}(1),$$

(D2) the number of possible values of  $\hat{\gamma}^{(n)}$  in balls with  $O(n^{-1/2}\nu^{(n)})$  radius centered at  $\gamma$  is bounded as  $n \rightarrow \infty$ .

Note that any  $n^{1/2}(\nu^{(n)})^{-1}$ -consistent estimator  $\hat{\gamma}^{(n)}$  is easily turned into a locally discrete one  $\hat{\gamma}_{\#}^{(n)}$  by letting  $\hat{\gamma}_{\#}^{(n)} := c^{-1} \text{sgn}(\hat{\gamma}^{(n)}) n^{-1/2}\nu^{(n)} \lceil n^{1/2}(\nu^{(n)})^{-1}c|\hat{\gamma}^{(n)}| \rceil$ , where  $\lceil z \rceil$  stands for the smallest integer larger than or equal to  $z$ , for some arbitrary positive constant  $c$  that does not depend on  $n$ . Subscripts  $\#$  in the sequel indicate such discretization. However, this discretization is a purely technical requirement, with little practical implications under fixed  $n$ , as the constant  $c$  can be chosen arbitrarily large.

Classical results on ULAN families (see, e.g., Chapter 11 of Le Cam 1986) then show that, for any fixed  $f_1 \in \mathcal{F}_A$ , locally uniformly asymptotically optimal (*most stringent*) tests of  $\mathcal{H}_{f_1}^{(n)} = \bigcup_{\boldsymbol{\theta}} \bigcup_{\sigma^2} \{\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; f_1}^{(n)}\}$  can be based on  $\Delta_{f_1;4}^{*(n)}(\hat{\boldsymbol{\vartheta}}_{\#})$ , hence on  $T_{f_1}^{*(n)}(\hat{\boldsymbol{\vartheta}}_{\#})$ , with

$$T_{f_1}^{*(n)}(\boldsymbol{\vartheta}) := (\Gamma_{f_1;44}^*(\boldsymbol{\vartheta}))^{-1/2} \Delta_{f_1;4}^{*(n)}(\boldsymbol{\vartheta}),\tag{3.9}$$

and  $\hat{\boldsymbol{\vartheta}}_{\#} := (\hat{\boldsymbol{\theta}}_{\#}^{(n)'}, \hat{\sigma}_{\#}^{2(n)}, 0)'$ , where  $\hat{\boldsymbol{\theta}}_{\#}^{(n)}$  and  $\hat{\sigma}_{\#}^{2(n)}$  are estimators satisfying (D1) and (D2) under  $\mathbb{P}_{\boldsymbol{\vartheta};f_1}^{(n)}$ , with  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$ . More precisely, we have the following result.

**Proposition 3.1** *Let Assumptions (B) and (C) hold, assume that  $\hat{\boldsymbol{\vartheta}}_{\#}$  satisfies (D1) and (D2), and fix  $f_1 \in \mathcal{F}_A$ . Then,*

- (i) for any  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$ ,  $T_{f_1}^{*(n)}(\hat{\boldsymbol{\vartheta}}_{\#}) = T_{f_1}^{*(n)}(\boldsymbol{\vartheta}) + o_{\mathbb{P}}(1)$  is asymptotically normal, with mean zero and mean  $(\Gamma_{f_1;44}^*(\boldsymbol{\vartheta}))^{1/2} \tau_4$  under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; f_1}^{(n)}$  and  $\mathbb{P}_{\boldsymbol{\vartheta} + n^{-1/2}\nu^{(n)}\boldsymbol{\tau}; f_1, h_1}^{(n)}$  ( $h_1 \in \mathcal{F}_{C|f_1}$ ), respectively, and variance one under both;

(ii) the sequence of tests  $\phi_{f_1}^{(n)}$  rejecting the null hypothesis  $\mathcal{H}_{f_1}^{(n)}$  as soon as  $T_{f_1}^{*(n)}(\hat{\boldsymbol{\vartheta}}_{\#})$  exceeds the  $(1 - \alpha)$ -quantile of the standard normal distribution, is locally asymptotically most stringent, at asymptotic probability level  $\alpha$ , for  $\mathcal{H}_{f_1}^{(n)}$  against alternatives of the form  $\bigcup_{\boldsymbol{\theta}} \bigcup_{\sigma^2} \bigcup_{\sigma_u^2 > 0} \bigcup_{h_1 \in \mathcal{F}_{C|f_1}} \{P_{\boldsymbol{\theta}, \sigma^2, \sigma_u^2; f_1, h_1}^{(n)}\}$ .

An important advantage of the proposed test statistic  $T_{f_1}^{*(n)}(\hat{\boldsymbol{\vartheta}}_{\#})$  is that it does not require specifying the (standardized) density  $h_1$  of random effects; on the other hand, the validity of the test in principle is limited to the standardized disturbance density  $f_1$ . The Gaussian versions of (3.7), (3.8) and (3.9), obtained for  $f_1 = \phi_1$ , are given by

$$\Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\vartheta}) = \frac{a^2}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T Z_{it}Z_{il}, \quad \Gamma_{\mathcal{N},44}^*(\boldsymbol{\vartheta}) = \frac{a^2 T(T-1)}{2\sigma^4},$$

and

$$T_{\mathcal{N}}^{*(n)}(\boldsymbol{\vartheta}) = \frac{a}{\sqrt{2nT(T-1)}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T Z_{it}Z_{il}, \quad (3.10)$$

respectively.

### 3.2 Pseudo-Gaussian tests.

The central sequence  $\Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\vartheta})$  allows for asymptotically optimal tests under  $f_1 = \phi_1$ , hence for efficient detection of random effects under Gaussian assumptions on the disturbance. Relaxing that Gaussian assumption is of course highly desirable. Let us show that this is indeed possible, and that the test based on (3.10) can be used as a *pseudo-Gaussian* test.

Define  $\mu_k(g_1) := \int_{\mathbb{R}} z^k g_1(z) dz$ , for  $k \in \mathbb{N}$  and consider the class  $\mathcal{F}_A^2$  of all densities  $g_1 \in \mathcal{F}_A$  such that  $\mu_2(g_1) < \infty$ . Since the intercept  $\mu$ , the regression coefficients  $\boldsymbol{\beta}$ , and the scale parameter  $\sigma^2$  under the null hypothesis remain unspecified, some care has to be taken with the asymptotic impact of estimating  $\mu$ ,  $\boldsymbol{\beta}$ , and  $\sigma^2$  under unspecified density  $g_1$ . That impact on  $\Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\vartheta})$  can be obtained, via Le Cam's Third Lemma, from the asymptotic behavior, under  $P_{\boldsymbol{\vartheta};g_1}^{(n)}$  ( $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$ ), of

$$\begin{pmatrix} \Delta_{g_1;1}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{g_1;2}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{g_1;3}^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\vartheta}) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \phi_{g_1}(Z_{it}) \\ \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \phi_{g_1}(Z_{it})(\mathbf{K}^{(n)})' \mathbf{x}_{it} \\ \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (Z_{it}\phi_{g_1}(Z_{it}) - 1) \\ \frac{a^2}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T Z_{it}Z_{il} \end{pmatrix},$$

which is asymptotically normal, with mean zero and block-diagonal covariance matrix

$$\begin{pmatrix} \frac{T}{\sigma^2} \mathcal{I}_{\phi}(g_1) & \mathbf{0} & \frac{T}{2\sigma^3} \mathcal{K}_{\phi\phi}(g_1) & 0 \\ \mathbf{0} & \frac{1}{\sigma^2} \mathcal{I}_{\phi}(g_1) \mathbf{I}_K & \mathbf{0} & \mathbf{0} \\ \frac{T}{2\sigma^3} \mathcal{K}_{\phi\phi}(g_1) & \mathbf{0} & \frac{T}{4\sigma^4} (\mathcal{J}_{\phi}(g_1) - 1) & 0 \\ 0 & \mathbf{0} & 0 & a^4 \frac{T(T-1)}{2\sigma^4} (\mu_2(g_1))^2 \end{pmatrix}.$$

It follows that  $\Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\vartheta})$ , under  $P_{\boldsymbol{\vartheta};g_1}^{(n)}$ , has the desirable property of being asymptotically uncorrelated with  $\Delta_{g_1;1}^{(n)}(\boldsymbol{\vartheta})$ ,  $\Delta_{g_1;2}^{(n)}(\boldsymbol{\vartheta})$ , and  $\Delta_{g_1;3}^{(n)}(\boldsymbol{\vartheta})$ —hence, asymptotically insensitive (in distribution), under  $P_{\boldsymbol{\vartheta};g_1}^{(n)}$  ( $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$ ) and contiguous alternatives, to local perturbations of  $\mu$ ,  $\boldsymbol{\beta}$ , and  $\sigma^2$ .

This equivalence in distribution is not sufficient if  $\mu$ ,  $\boldsymbol{\beta}$ , and  $\sigma^2$  in  $T_{\mathcal{N}}^{*(n)}$  are to be replaced with estimators. A stronger asymptotic equivalence in probability result, however, can be established, based on the following asymptotic linearity result (which only considers perturbations of  $\boldsymbol{\theta} = (\mu, \boldsymbol{\beta}')'$ , since the linearity of Gaussian scores allows for controlling the non-specification of  $\sigma^2$  through the continuous mapping theorem).

**Proposition 3.2** *Let Assumptions (B) and (C) hold and fix  $\boldsymbol{\theta} \in \mathbb{R}^{K+1}$ ,  $\sigma^2 > 0$ , and  $g_1 \in \mathcal{F}_A^2$ . Then, for any bounded sequence  $\mathbf{t}^{(n)} \in \mathbb{R}^{K+1}$ ,*

$$\Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\theta} + n^{-1/2}\boldsymbol{\nu}_1^{(n)}\mathbf{t}^{(n)}, \sigma^2, 0) - \Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\theta}, \sigma^2, 0) = o_P(1), \quad (3.11)$$

under  $P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ , as  $n \rightarrow \infty$  with fixed  $T$ .

The following proposition then is an immediate corollary to Proposition 3.2 and Lemma 4.4 in Kreiss (1987).

**Proposition 3.3** *Let Assumptions (B) and (C) hold, assume that  $\hat{\boldsymbol{\theta}}_{\#}$  satisfies (D1) and (D2), and fix  $\boldsymbol{\theta} \in \mathbb{R}^{K+1}$ ,  $\sigma^2 > 0$ , and  $g_1 \in \mathcal{F}_A^2$ . Then  $\Delta_{\mathcal{N};4}^{*(n)}(\hat{\boldsymbol{\theta}}_{\#}^{(n)}, \sigma^2, 0) - \Delta_{\mathcal{N};4}^{*(n)}(\boldsymbol{\theta}, \sigma^2, 0) = o_P(1)$  under  $P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ , as  $n \rightarrow \infty$  with fixed  $T$ .*

Under  $P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ ,  $m_2^{(n)}(\boldsymbol{\theta}) := (nT)^{-1} \sum_{i=1}^n \sum_{t=1}^T (Y_{it} - \mu - \boldsymbol{\beta}'\mathbf{x}_{it})^2$  converges in probability to  $\sigma^2\mu_2(g_1)$ . In practice,  $m_2^{(n)}(\boldsymbol{\theta})$  cannot be computed from the observations, and  $(Y_{it} - \hat{\mu}_{\#}^{(n)} - \hat{\boldsymbol{\beta}}_{\#}^{(n)'}\mathbf{x}_{it})$  is to be substituted for  $(Y_{it} - \mu - \boldsymbol{\beta}'\mathbf{x}_{it})$  in  $m_2^{(n)}(\boldsymbol{\theta})$  yielding  $m_2^{(n)} := m_2^{(n)}(\hat{\boldsymbol{\theta}}_{\#}^{(n)})$ . Since the asymptotic covariance, under  $P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ ,  $g_1 \in \mathcal{F}_A^2$ , of  $\sqrt{nm_2^{(n)}(\boldsymbol{\theta})}$  with  $(\Delta_{g_1;1}^{(n)}(\boldsymbol{\vartheta}), \Delta_{g_1;2}^{(n)'}(\boldsymbol{\vartheta}))'$  is finite, a routine application of Le Cam's third Lemma yields  $m_2^{(n)}(\hat{\boldsymbol{\theta}}_{\#}^{(n)}) - m_2^{(n)}(\boldsymbol{\theta}) = o_P(1)$  under  $P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ ,  $g_1 \in \mathcal{F}_A^2$ . Defining

$$T_{\phi_1}^{*(n)}(\boldsymbol{\theta}) := \frac{1}{\sqrt{2nT(T-1)}} \frac{1}{m_2^{(n)}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T (Y_{it} - \mu - \boldsymbol{\beta}'\mathbf{x}_{it})(Y_{il} - \mu - \boldsymbol{\beta}'\mathbf{x}_{il}), \quad (3.12)$$

we then have the following result.

**Proposition 3.4** *Let Assumptions (B) and (C) hold and assume that  $\hat{\boldsymbol{\theta}}_{\#}$  satisfies (D1) and (D2). Then,*

(i) *for any  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$  and  $g_1 \in \mathcal{F}_A^2$ ,  $T_{\phi_1}^{*(n)}(\hat{\boldsymbol{\theta}}_{\#}^{(n)}) = T_{\phi_1}^{*(n)}(\boldsymbol{\theta}) + o_P(1)$  is asymptotically normal, with mean zero and mean*

$$\frac{\sqrt{T(T-1)}}{\sqrt{2}\sigma^2\mu_2(g_1)}\tau_4$$

under  $P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$  and  $P_{\boldsymbol{\vartheta} + n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}; g_1, h_1}^{(n)}$  ( $h_1 \in \mathcal{F}_{C|g_1}$ ), respectively, and variance one under both;

(ii) the sequence of tests  $\phi_N^{*(n)}$  rejecting the null hypothesis  $\mathcal{H}_A^{(n)2} := \bigcup_{g_1 \in \mathcal{F}_A^2} \bigcup_{\boldsymbol{\theta}} \bigcup_{\sigma^2} \{P_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}\}$  as soon as  $T_{\phi_1}^{*(n)}(\hat{\boldsymbol{\theta}}_{\#}^{(n)})$  exceeds the  $(1 - \alpha)$ -quantile of the standard normal distribution is locally asymptotically most stringent, at asymptotic probability level  $\alpha$ , for  $\mathcal{H}_A^{(n)2}$  against alternatives of the form  $\bigcup_{\boldsymbol{\theta}} \bigcup_{\sigma^2} \bigcup_{\sigma_u^2 > 0} \bigcup_{h_1 \in \mathcal{F}_{C|\phi_1}} \{P_{\boldsymbol{\theta}, \sigma^2, \sigma_u^2; \phi_1, h_1}^{(n)}\}$ .

The test statistic  $T_{\phi_1}^{*(n)}(\hat{\boldsymbol{\theta}}_{\#}^{(n)})$  thus defines a pseudo-Gaussian test, that is, a test which is optimal under Gaussian assumptions but remains valid under a much broader class of densities.

## 4 Optimal rank tests.

### 4.1 Rank-based versions of central sequences.

A serious drawback of the parametric tests of Section 3.1 is that their validity is restricted, in general, to the correctly specified density  $f_1$  they have been constructed for. As for the pseudo-Gaussian tests of Section 3.2, they still require finite moments of order two; moreover, for fixed  $n$ , the quality of the Gaussian approximation they are based on is likely to depend on the actual underlying density.

Since a correct specification of the actual density  $g_1$  in practice is highly unrealistic, the problem has to be considered from a semiparametric point of view, where  $g_1$  plays the role of a nuisance. A general result by Hallin and Werker (2003) suggests that, in such context, semiparametrically efficient (at selected  $g_1 = f_1$ ) tests can be obtained by conditioning the  $f_1$ -central sequence on the maximal invariant  $\sigma$ -field associated with some appropriate generating group.

More precisely, note that the null hypothesis  $\mathcal{H}_{\boldsymbol{\beta}}^{(n)} := \bigcup_{g_1 \in \mathcal{F}_0} \bigcup_{\mu \in \mathbb{R}} \bigcup_{\sigma^2 > 0} \{P_{\mu, \boldsymbol{\beta}, \sigma^2, 0; g_1}^{(n)}\}$  is invariant under the action of the group  $\mathcal{G}_{\boldsymbol{\beta}}^{(nT), \circ}$  of all transformations  $\mathcal{G}_l$  of  $\mathbb{R}^{nT}$  such that

$$\mathcal{G}_l(y_{11}, \dots, y_{nT}) := (\boldsymbol{\beta}' \mathbf{x}_{11} + l(y_{11} - \boldsymbol{\beta}' \mathbf{x}_{11}), \dots, \boldsymbol{\beta}' \mathbf{x}_{nT} + l(y_{nT} - \boldsymbol{\beta}' \mathbf{x}_{11})),$$

where  $z \mapsto l(z)$  is continuous and monotone increasing and  $\lim_{z \rightarrow \pm\infty} l(z) = \pm\infty$  (the *order-preserving group*); note that the  $\mathcal{G}_{\boldsymbol{\beta}}^{(nT), \circ}$  is a generating group for  $\mathcal{H}_{\boldsymbol{\beta}}^{(n)}$ , and has maximal invariant the vector of ranks  $(R_{11}^{(n)}(\boldsymbol{\beta}), \dots, R_{nT}^{(n)}(\boldsymbol{\beta}))$ , where  $R_{it}^{(n)}(\boldsymbol{\beta})$  denotes the rank of  $Y_{it} - \mu - \boldsymbol{\beta}' \mathbf{x}_{it}$  among  $Y_{11} - \mu - \boldsymbol{\beta}' \mathbf{x}_{11}, \dots, Y_{nT} - \mu - \boldsymbol{\beta}' \mathbf{x}_{nT}$  (the  $R_{it}^{(n)}(\boldsymbol{\beta})$ 's do not depend on  $\mu$ , which justifies the notation). The idea of considering tests that are measurable with respect to those ranks looks quite natural and appealing, since such tests would be distribution-free (or asymptotically so, since the unspecified  $\boldsymbol{\beta}$  will have to be replaced with some estimator) and hopefully semiparametrically efficient. As we shall show, this is indeed the case, and semiparametric efficiency moreover turns out to match parametric efficiency at the selected reference density  $f_1$ .

The rank-based version  $\Delta_{f_1;4}^{*(n)}(\boldsymbol{\vartheta})$  of the  $\sigma_u^2$ -efficient central sequence  $\Delta_{f_1;4}^{*(n)}(\boldsymbol{\vartheta})$  we plan to base our tests on is

$$\Delta_{f_1;4}^{*(n)}(\boldsymbol{\beta}, \sigma^2) := \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \left\{ \varphi_{f_1} \left( \frac{R_{it}^{(n)}(\boldsymbol{\beta})}{N+1} \right) \varphi_{f_1} \left( \frac{R_{il}^{(n)}(\boldsymbol{\beta})}{N+1} \right) - c_{f_1}^{(n)} \right\}, \quad (4.1)$$

with  $c_{f_1}^{(n)} := \frac{1}{N(N-1)} \sum_{r=1}^N \sum_{s \neq r=1}^N \varphi_{f_1} \left( \frac{r}{N+1} \right) \varphi_{f_1} \left( \frac{s}{N+1} \right)$  and  $\varphi_{f_1} := \phi_{f_1} \circ F_1^{-1}$ .

Beyond its role in the derivation of the asymptotic distribution of  $\Delta_{f_1;4}^{*(n)}(\boldsymbol{\beta}, \sigma^2)$ , the fol-

lowing asymptotic representation result shows that  $\underline{\Delta}_{f_1;4}^{*(n)}(\boldsymbol{\beta}, \sigma^2)$  is indeed another version of the efficient central sequence  $\underline{\Delta}_{f_1;4}^{*(n)}(\boldsymbol{\vartheta})$  (compare (3.7) with  $\underline{\Delta}_{f_1, f_1;4}^{*(n)}(\boldsymbol{\vartheta})$  defined in (4.3) below). Writing  $\sum_{1 \leq r_1 \neq \dots \neq r_q \leq N}$  for a sum running over the  $(N!/(N-q)!)$  ordered  $q$ -tuples of distinct integers in  $\{1, \dots, N\}$ , let

$$s_{f_1}^{2(n)} := \gamma_1 \sum_{r=1}^N \sum_{s \neq r=1}^N \varphi_{f_1}^2\left(\frac{r}{N+1}\right) \varphi_{f_1}^2\left(\frac{s}{N+1}\right) + \gamma_2 \sum_{1 \leq r \neq s \neq v \leq N} \varphi_{f_1}^2\left(\frac{r}{N+1}\right) \varphi_{f_1}\left(\frac{s}{N+1}\right) \varphi_{f_1}\left(\frac{v}{N+1}\right) \\ + \gamma_3 \sum_{1 \leq r \neq s \neq v \neq w \leq N} \varphi_{f_1}\left(\frac{r}{N+1}\right) \varphi_{f_1}\left(\frac{s}{N+1}\right) \varphi_{f_1}\left(\frac{v}{N+1}\right) \varphi_{f_1}\left(\frac{w}{N+1}\right) - \gamma_4 \left(c_{f_1}^{(n)}\right)^2, \quad (4.2)$$

where

$$\gamma_1 := \frac{T(T-1)}{2N(N-1)}, \quad \gamma_2 := \frac{7T(T-1)(T-2)}{3N(N-1)(N-2)}, \quad \gamma_3 := \frac{T(T-1)(nT^2 - (n+4)T + 6)}{4N(N-1)(N-2)(N-3)},$$

and

$$\gamma_4 := \frac{T(T-1)(3nT^2 - (3n-16)T - 32)}{12}.$$

Defining the *cross-information coefficient*

$$\mathcal{I}_\phi(f_1, g_1) := \int_0^1 \phi_{f_1}(F_1^{-1}(u)) \phi_{g_1}(G_1^{-1}(u)) du,$$

and denoting by  $\widetilde{\mathcal{F}}_A$  the class of all densities  $f_1 \in \mathcal{F}_A$  such that  $\phi_{f_1}$  can be expressed as the difference of two monotone increasing functions, we have the following result.

**Proposition 4.1** Fix  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, 0)'$  (with  $\boldsymbol{\theta} \in \mathbb{R}^{K+1}$  and  $\sigma^2 > 0$ ),  $f_1 \in \widetilde{\mathcal{F}}_A$ , and  $g_1 \in \mathcal{F}_0$ . Then,

(i) under  $\mathbb{P}_{\boldsymbol{\vartheta}; g_1}^{(n)}$ , as  $n \rightarrow \infty$  with fixed  $T$ ,

$$\underline{\Delta}_{f_1;4}^{*(n)}(\boldsymbol{\beta}, \sigma^2) = \mathbb{E}_{\boldsymbol{\vartheta}; g_1}^{(n)}[\underline{\Delta}_{f_1;4}^{(n)}(\boldsymbol{\vartheta}) | R_{11}^{(n)}(\boldsymbol{\beta}), \dots, R_{nT}^{(n)}(\boldsymbol{\beta})] + o_{L^2}(1) = \underline{\Delta}_{f_1, g_1;4}^{*(n)}(\boldsymbol{\vartheta}) + o_{L^2}(1),$$

with (denoting by  $G_1$  the distribution function associated with  $g_1$ )

$$\underline{\Delta}_{f_1, g_1;4}^{*(n)}(\boldsymbol{\vartheta}) := \frac{1}{2\sigma^2\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \varphi_{f_1}(G_1(Z_{it})) \varphi_{f_1}(G_1(Z_{il})); \quad (4.3)$$

(ii) under  $\mathbb{P}_{\boldsymbol{\vartheta}; g_1}^{(n)}$ ,  $\underline{\Delta}_{f_1;4}^{*(n)}(\boldsymbol{\beta}, \sigma^2)$  has mean zero and variance  $\Gamma_{f_1;44}^{*(n)}(\sigma^2) := \sigma^{-4} s_{f_1}^{2(n)} = \Gamma_{f_1;44}^*(\sigma^2) + o(1)$  as  $n \rightarrow \infty$  with fixed  $T$ , where  $\Gamma_{f_1;44}^*(\sigma^2)$  was defined in (3.8);

(iii)  $\underline{\Delta}_{f_1, g_1;4}^{*(n)}(\boldsymbol{\vartheta})$  is asymptotically normal with mean zero and mean  $(T(T-1)/2\sigma^4) \mathcal{I}_\phi^2(f_1, g_1) \tau_4$  under  $\mathbb{P}_{\boldsymbol{\vartheta}; g_1}^{(n)}$  and  $\mathbb{P}_{\boldsymbol{\vartheta} + n^{-1/2} \boldsymbol{\nu}^{(n)}; g_1, h_1}^{(n)}$ , respectively, and variance  $\Gamma_{f_1;44}^*(\sigma^2)$  under both (the claim under  $\mathbb{P}_{\boldsymbol{\vartheta} + n^{-1/2} \boldsymbol{\nu}^{(n)}; g_1, h_1}^{(n)}$  further requires  $g_1 \in \mathcal{F}_A$  and  $h_1 \in \mathcal{F}_{C|g_1}$ ).

## 4.2 Optimal rank tests.

The parameters  $\mu$ ,  $\boldsymbol{\beta}$ , and  $\sigma^2$  remain unspecified under the null. Since only  $\boldsymbol{\beta}$  has an influence on the ranks, a consistent estimator  $\hat{\boldsymbol{\beta}}^{(n)}$  has to be substituted for the actual  $\boldsymbol{\beta}$  value, yielding *aligned ranks*  $R_{it}^{(n)}(\hat{\boldsymbol{\beta}}^{(n)})$ . The effect of this alignment procedure is taken care of in a similar way as in Section 3, via the asymptotic linearity results of Propositions 4.2 and 4.3 below.

**Proposition 4.2** *Let Assumptions (B) and (C) hold and fix  $\mu \in \mathbb{R}$ ,  $\boldsymbol{\beta} \in \mathbb{R}^K$ ,  $\sigma^2 > 0$ ,  $f_1 \in \tilde{\mathcal{F}}_A$ , and  $g_1 \in \mathcal{F}_A$ . Then, for any bounded sequence  $\mathbf{b}^{(n)} \in \mathbb{R}^K$ ,*

$$\underline{\Delta}_{f_1;4}^{*(n)}(\mu, \boldsymbol{\beta} + n^{-1/2} \mathbf{K}^{(n)} \mathbf{b}^{(n)}, \sigma^2, 0) - \underline{\Delta}_{f_1;4}^{*(n)}(\mu, \boldsymbol{\beta}, \sigma^2, 0) = o_{\mathbb{P}}(1), \quad (4.4)$$

under  $\mathbb{P}_{\mu, \boldsymbol{\beta}, \sigma^2, 0; g_1}^{(n)}$ , as  $n \rightarrow \infty$  with fixed  $T$ .

The following proposition then is an immediate corollary of Proposition 4.2 and Lemma 4.4 in Kreiss (1987).

**Proposition 4.3** *Let Assumptions (B) and (C) hold, assume that  $\hat{\boldsymbol{\beta}}_{\#}$  satisfies (D1) and (D2), and fix  $\boldsymbol{\theta} \in \mathbb{R}^{K+1}$ ,  $\sigma^2 > 0$ ,  $f_1 \in \tilde{\mathcal{F}}_A$ , and  $g_1 \in \mathcal{F}_A$ . Then*

$$\underline{\Delta}_{f_1;4}^{*(n)}(\mu, \hat{\boldsymbol{\beta}}_{\#}^{(n)}, \sigma^2, 0) - \underline{\Delta}_{f_1;4}^{*(n)}(\mu, \boldsymbol{\beta}, \sigma^2, 0) = o_{\mathbb{P}}(1),$$

under  $\mathbb{P}_{\mu, \boldsymbol{\beta}, \sigma^2, 0; g_1}^{(n)}$ , as  $n \rightarrow \infty$  with fixed  $T$ .

Local asymptotic optimality under density  $f_1$  is achieved by the test based on

$$\begin{aligned} \underline{T}_{f_1}^{*(n)}(\boldsymbol{\theta}) &:= \left( \Gamma_{f_1;44}^{*(n)}(\sigma^2) \right)^{-1/2} \underline{\Delta}_{f_1;4}^{*(n)}(\boldsymbol{\theta}) \\ &= \frac{1}{2s_{f_1}^{(n)} \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \left\{ \varphi_{f_1} \left( \frac{R_{it}^{(n)}(\boldsymbol{\beta})}{N+1} \right) \varphi_{f_1} \left( \frac{R_{il}^{(n)}(\boldsymbol{\beta})}{N+1} \right) - c_{f_1}^{(n)} \right\} =: \underline{T}_{f_1}^{*(n)}(\boldsymbol{\beta}). \end{aligned} \quad (4.5)$$

More precisely, we have the following result.

**Proposition 4.4** *Let Assumptions (B) and (C) hold, assume that  $\hat{\boldsymbol{\beta}}_{\#}$  satisfies (D1) and (D2), and fix  $f_1 \in \tilde{\mathcal{F}}_A$ . Then*

(i) *for any  $g_1 \in \mathcal{F}_A$ ,  $\underline{T}_{f_1}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#})$  is asymptotically normal with mean zero and mean*

$$\frac{T(T-1) \mathcal{I}_{\phi}^2(f_1, g_1)}{2\sigma^4 (\Gamma_{f_1;44}^*(\sigma^2))^{1/2}} \tau_4$$

under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$  and  $\mathbb{P}_{\boldsymbol{\theta} + n^{-1/2} \boldsymbol{\nu}^{(n)} \boldsymbol{\tau}; g_1, h_1}^{(n)}$  ( $h_1 \in \mathcal{F}_{C|g_1}$ ), respectively, and variance one under both.

(ii) the sequence of tests  $\phi_{f_1}^{*(n)}$  rejecting the null hypothesis

$$\mathcal{H}_A^{(n)} := \bigcup_{g_1 \in \mathcal{F}_A} \bigcup_{\boldsymbol{\theta}} \bigcup_{\sigma^2} \left\{ \mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)} \right\}$$

as soon as  $T_{\tilde{f}_1}^{(n)}(\hat{\boldsymbol{\beta}}_{\#})$  exceeds the  $(1 - \alpha)$ -quantile of the standard normal distribution is locally asymptotically most stringent, at asymptotic probability level  $\alpha$ , for  $\mathcal{H}_A^{(n)}$  against alternatives of the form  $\bigcup_{\boldsymbol{\theta}} \bigcup_{\sigma^2} \bigcup_{\sigma_u^2 > 0} \bigcup_{h_1 \in \mathcal{F}_{C|f_1}} \left\{ \mathbb{P}_{\boldsymbol{\theta}, \sigma^2, \sigma_u^2; f_1, h_1}^{(n)} \right\}$ .

### 4.3 The van der Waerden and Wilcoxon test statistics.

The statistic  $T_{\tilde{f}_1}^{(n)}(\hat{\boldsymbol{\beta}}_{\#})$  is providing a general form for the optimal rank tests of the null hypothesis of absence of random effects. Important particular cases are the Wilcoxon (logistic scores) and van der Waerden (normal scores) test statistics, which are optimal at logistic and normal densities, respectively.

The van der Waerden tests use the standard normal reference density  $f_1 = \phi_1$ . One easily obtains that  $\phi_{f_1}(F_1^{-1}(u)) = a^{1/2}\Phi^{-1}(u)$ , where  $\Phi$  denotes the standard normal distribution function. The resulting rank test statistic is then

$$T_{\tilde{\text{vdW}}}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#}) := T_{\tilde{\phi}_1}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#}) = \frac{1}{2s_{\phi_1}^{(n)}\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \left\{ a\Phi^{-1}\left(\frac{R_{it}^{(n)}(\hat{\boldsymbol{\beta}}_{\#})}{N+1}\right) \Phi^{-1}\left(\frac{R_{il}^{(n)}(\hat{\boldsymbol{\beta}}_{\#})}{N+1}\right) - c_{\phi_1}^{(n)} \right\},$$

with

$$c_{\phi_1}^{(n)} = \frac{a}{N(N-1)} \sum_{r=1}^N \sum_{s \neq r=1}^N \Phi^{-1}\left(\frac{r}{N+1}\right) \Phi^{-1}\left(\frac{s}{N+1}\right).$$

In the Wilcoxon case,  $f_1(z) = \ell_1(z) := \sqrt{b} \exp(-\sqrt{b}z)/(1 + \exp(-\sqrt{b}z))^2$  is a standardized logistic density. One easily checks that  $\phi_{f_1}(F_1^{-1}(u)) = b^{1/2}u$ . Therefore, the Wilcoxon test statistic takes the form

$$T_{\tilde{\text{W}}}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#}) := T_{\tilde{\ell}_1}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#}) = \frac{1}{2s_{\ell_1}^{(n)}\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T \left\{ b \left(\frac{R_{it}^{(n)}(\hat{\boldsymbol{\beta}}_{\#})}{N+1}\right) \left(\frac{R_{il}^{(n)}(\hat{\boldsymbol{\beta}}_{\#})}{N+1}\right) - c_{\ell_1}^{(n)} \right\},$$

with

$$c_{\ell_1}^{(n)} = \frac{b}{N(N-1)} \sum_{r=1}^N \sum_{s \neq r=1}^N \left(\frac{r}{N+1}\right) \left(\frac{s}{N+1}\right).$$

It is worth noting that the scale factors  $a$  (for van der Waerden) and  $b$  (for Wilcoxon) disappear in the final expression of the test statistics, due to the (exact) standardization by  $s_{\phi_1}^{(n)}$  and  $s_{\ell_1}^{(n)}$ , respectively. This confirms that our choice of the median of absolute deviations as a scale parameter in the definition of  $\mathcal{F}_0$  has no impact on the results.

### 4.4 Asymptotic relative efficiencies.

Propositions 3.4 and 4.4 allow for computing ARE values for the rank tests based on  $T_{\tilde{f}_1}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#})$  with respect to the pseudo-Gaussian tests based on  $T_{\mathcal{N}}^{*(n)}(\hat{\boldsymbol{\theta}})$ ; these AREs as usual are obtained as ratios of the squares of the shifts of these test statistics under local alternatives.

**Proposition 4.5** Let  $f_1 \in \mathcal{F}_A$ . Then, the asymptotic relative efficiencies, under  $g_1 \in \mathcal{F}_A^2$ , of the rank tests based on  $T_{f_1}^{*(n)}(\hat{\beta}_\#)$  with respect to the pseudo-Gaussian tests based on  $T_{\mathcal{N}}^{*(n)}(\hat{\theta})$  are

$$\text{ARE}_{g_1}(T_{f_1}^{*(n)}(\hat{\theta})/T_{\mathcal{N}}^{*(n)}(\hat{\theta})) = \left\{ \mu_2(g_1) \mathcal{I}_\phi^2(f_1, g_1) / \mathcal{I}_\phi(f_1) \right\}^2. \quad (4.6)$$

Strictly speaking, however, the interpretation of (4.6) as an ARE value only holds for individual effect standardized densities  $h_1 \in \mathcal{F}_{C|g_1}$ .

Numerical values of the ARE values of Proposition 4.5, under  $t_3, t_5, t_8, t_{10}, t_{20}$ , normal, and logistic densities, are displayed in Table 1. These values all are good, particularly so under heavy tails (see the Student density with 3 degrees of freedom). Also, the AREs of the van der Waerden tests are uniformly larger than or equal to one for all distributions considered in Table 1, and are equal to one in the Gaussian case only. Since the AREs in (4.6) are the squares of those obtained in univariate location problems, this is a general result: the Chernoff-Savage (1958) property extends to the problem considered in this paper, showing that our van der Waerden rank tests, from the Pitman point of view, uniformly dominate their pseudo-Gaussian parametric competitors.

		actual density $g_1$						
		$f_{1,t_3}$	$f_{1,t_5}$	$f_{1,t_8}$	$f_{1,t_{10}}$	$f_{1,t_{20}}$	$\phi_1$	$\ell_1$
scores	$f_{1,t_3}$	4.0000	1.4923	1.0498	0.9498	0.7917	0.6718	1.0851
	$f_{1,t_5}$	3.8202	1.5625	1.1667	1.0783	0.9401	0.8370	1.1857
	$f_{1,t_8}$	3.5284	1.5319	1.1901	1.1158	1.0029	0.9232	1.1993
	$f_{1,t_{10}}$	3.3960	1.5061	1.1870	1.1187	1.0168	0.9476	1.1933
	$f_{1,t_{20}}$	3.0768	1.4272	1.1597	1.1053	1.0292	0.9851	1.1610
	$\phi_1$	2.6873	1.3079	1.0986	1.0601	1.0139	1.0000	1.0966
	$\ell_1$	3.6091	1.5405	1.1869	1.1101	0.9936	0.9119	1.2026

Table 1: AREs, under Student (3, 5, 8, 10, and 20 degrees of freedom), normal and logistic densities, of various rank tests (based on Student, van der Waerden, and Wilcoxon scores), with respect to the pseudo-Gaussian tests.

## 5 Simulations.

In this section, we conduct a Monte Carlo experiment to investigate the finite-sample behavior of our rank tests under a variety of error distributions. More precisely, we considered the model

$$Y_{it} = \mu + \beta_1 x_{it}^1 + \beta_2 x_{it}^2 + u_i + \varepsilon_{it}, \quad i = 1, \dots, n = 100, \quad t = 1, \dots, T = 5, \quad (5.7)$$

where

- $\mu = \beta_1 = \beta_2 = 1$ ;
- the  $x_{it}^1$ 's are i.i.d. uniform over  $(0, 1)$  and the  $x_{it}^2$ 's are obtained from i.i.d.  $\nu_{it}$  uniformly distributed over  $(-0.5, 0.5)$  via  $x_{it}^2 = 0.1t + 0.5x_{i,t-1}^2 + \nu_{it}$  and  $x_{i0}^2 = 5 + 10\nu_{i0}$  (see Nerlove 1971);
- the  $u_i$ 's are i.i.d. Gaussian with mean zero and variance  $\sigma_u^2 = 0$  (null hypothesis), 0.2, 0.3, 0.4, or 0.5 (increasingly distant alternatives);

- the  $\varepsilon_{it}$ 's are i.i.d. Student with 1, 3, and 8 degrees of freedom, Gaussian, or logistic (more precisely, they are i.i.d. with density  $f_{1,t_1}$ ,  $f_{1,t_3}$ ,  $f_{1,t_8}$ ,  $\phi_1$ , or  $\ell_1$ ).

For each combination of  $\sigma_u^2$  and a distribution for the  $\varepsilon_{it}$ 's, we generated  $M = 2,500$  independent samples from (5.7), with the same values of the regressors  $x_{it}^j$  over replications. For each replication, the following six tests were performed at nominal probability level  $\alpha = 5\%$ : the pseudo-Gaussian test based on  $T_{\phi_1}^{*(n)}(\hat{\boldsymbol{\theta}}_{\#}^{(n)})$ , three  $t_\nu$ -score tests based on  $T_{f_{1,t_\nu}}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#}^{(n)})$  ( $\nu = 1, 3, 8$ ), the van der Waerden test based on  $T_{\text{vdW}}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#}^{(n)})$ , and the Wilcoxon test based on  $T_{\text{W}}^{*(n)}(\hat{\boldsymbol{\beta}}_{\#}^{(n)})$ . The estimator  $\hat{\boldsymbol{\theta}}_{\#}^{(n)} = (\hat{\boldsymbol{\mu}}_{\#}^{(n)}, \hat{\boldsymbol{\beta}}_{\#}^{(n)})'$  used is the least absolute deviations (LAD) estimator. Rejection frequencies are reported in Table 2.

These simulations show that the pseudo-Gaussian test, though resisting non-Gaussian densities with finite second-order moments, are collapsing under the heavy-tailed  $t_1$  distribution. In sharp contrast with this, all proposed rank-based tests appear to meet the 5% probability level constraint. Empirical power rankings are essentially consistent with the corresponding ARE values from Table 1. For instance, under Gaussian densities, the powers of the  $t_\nu$ -score rank tests are increasing with  $\nu$  as expected, whereas the asymptotic optimality of the same tests under the corresponding Student distribution with  $\nu$  degrees of freedom is confirmed.

## 6 Appendix.

### 6.1 Proof of Proposition 2.1.

The proof consists in checking that Swensen's (1985) sufficient conditions (1.2)-(1.7) hold. Conditions (1.3)-(1.7) (to be checked, for ULAN, under sequences  $\boldsymbol{\vartheta}^{(n)}$ ) follow more or less routinely from the assumptions made, and the only delicate one actually is condition (1.2). This condition is a direct consequence (see Swensen's Lemma 2) of the quadratic mean differentiability, at any  $(\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, 0)$ , of

$$(\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, \sigma_u^2) \mapsto \underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, \sigma_u^2; f_1}^{1/2}(\mathbf{y}) := \left\{ \frac{1}{\sigma^T} \int_{\mathbb{R}} \prod_{t=1}^T f_1 \left( \frac{1}{\sigma} (y_t - \boldsymbol{\mu} - \boldsymbol{\beta}' \mathbf{x}_t - \sigma_u u) \right) h(u) du \right\}^{1/2}$$

with  $\mathbf{y} := (y_1, \dots, y_T)' \in \mathbb{R}^T$  and  $\mathbf{x}_t := (x_t^1, x_t^2, \dots, x_t^K)' \in \mathbb{R}^K$ ,  $t = 1, \dots, T$ . The technical difficulty lies in the fact that  $\underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, \sigma_u^2; f_1}$  has the form of a mixture density. Quadratic mean differentiability is established in the following lemma.

**Lemma 6.1** *Let Assumptions (B) and (C) hold and fix  $f_1 \in \mathcal{F}_A$ . Define, for  $\mathbf{y} \in \mathbb{R}^T$ ,*

$$\begin{aligned} D_{\boldsymbol{\mu}} \underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) &:= \frac{1}{2\sigma} \underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \left[ \sum_{t=1}^T \phi_{f_1} \left( \frac{y_t - \boldsymbol{\mu} - \boldsymbol{\beta}' \mathbf{x}_t}{\sigma} \right) \right], \\ D_{\boldsymbol{\beta}} \underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) &:= \frac{1}{2\sigma} \underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \left[ \sum_{t=1}^T \phi_{f_1} \left( \frac{y_t - \boldsymbol{\mu} - \boldsymbol{\beta}' \mathbf{x}_t}{\sigma} \right) (\mathbf{K}^{(n)})' \mathbf{x}_t \right], \\ D_{\sigma^2} \underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) &:= \frac{1}{4\sigma^2} \underline{f}_{\boldsymbol{\mu}, \boldsymbol{\beta}, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \left[ \sum_{t=1}^T \left( \left( \frac{y_t - \boldsymbol{\mu} - \boldsymbol{\beta}' \mathbf{x}_t}{\sigma} \right) \phi_{f_1} \left( \frac{y_t - \boldsymbol{\mu} - \boldsymbol{\beta}' \mathbf{x}_t}{\sigma} \right) - 1 \right) \right], \end{aligned}$$

		$\sigma_u$				
Actual density	Test	0	0.2	0.3	0.4	0.5
$\phi_1$	$\mathcal{N}$	0.0444	0.3268	0.7740	0.9768	0.9996
	vdW	0.0488	0.3416	0.7820	0.9812	0.9996
	W	0.0500	0.3144	0.7544	0.9708	0.9996
	$t_1$	0.0572	0.1396	0.3176	0.5668	0.8256
	$t_3$	0.0504	0.2596	0.6528	0.9372	0.9976
	$t_8$	0.0500	0.3216	0.7628	0.9740	0.9996
$\ell_1$	$\mathcal{N}$	0.0556	0.0952	0.2216	0.4124	0.6732
	vdW	0.0552	0.1108	0.2428	0.4448	0.7068
	W	0.0552	0.1164	0.2600	0.4520	0.7260
	$t_1$	0.0492	0.0936	0.1500	0.2456	0.4184
	$t_3$	0.0544	0.1224	0.2496	0.4196	0.6832
	$t_8$	0.0556	0.1188	0.2572	0.4508	0.7236
$f_{1,t_1}$	$\mathcal{N}$	0.0176	0.0192	0.0196	0.0184	0.0192
	vdW	0.0536	0.0884	0.1616	0.2668	0.4344
	W	0.0540	0.1180	0.2148	0.3772	0.6104
	$t_1$	0.0536	0.1648	0.3380	0.5844	0.8300
	$t_3$	0.0568	0.1388	0.2744	0.4920	0.7300
	$t_8$	0.0536	0.1108	0.2052	0.3588	0.5808
$f_{1,t_3}$	$\mathcal{N}$	0.0480	0.1176	0.2596	0.5224	0.7552
	vdW	0.0576	0.1896	0.4332	0.7472	0.9408
	W	0.0572	0.2176	0.4928	0.8228	0.9704
	$t_1$	0.0592	0.1648	0.3712	0.6440	0.8636
	$t_3$	0.0580	0.2184	0.5248	0.8356	0.9760
	$t_8$	0.0576	0.2148	0.4880	0.8148	0.9668
$f_{1,t_8}$	$\mathcal{N}$	0.0512	0.2360	0.5960	0.9040	0.9920
	vdW	0.0556	0.2636	0.6412	0.9272	0.9948
	W	0.0560	0.2828	0.6548	0.9388	0.9956
	$t_1$	0.0468	0.1520	0.3340	0.6012	0.8420
	$t_3$	0.0560	0.2592	0.6120	0.9132	0.9908
	$t_8$	0.0552	0.2820	0.6580	0.9384	0.9960

Table 2: Rejection frequencies (out of  $M = 2,500$  replications), for  $\sigma_u = 0$  (null hypothesis) 0.2, 0.3, 0.4, 0.5 (alternative hypotheses), with Gaussian ( $\phi_1$ ), logistic ( $\ell_1$ ),  $t_\nu$  with  $\nu = 1, 3, 8$  ( $f_{1,t_1}, f_{1,t_3}, f_{1,t_8}$ ) disturbances, of the pseudo-Gaussian test ( $\mathcal{N}$ ), and the van der Waerden (vdW), Wilcoxon (W), and  $t_\nu$ -score with  $\nu = 1, 3, 8$  ( $t_1, t_3, t_8$ ) rank tests; the sample size is 500 ( $n = 100$  and  $T = 5$ ).

and

$$D_{\sigma_u^2} \underline{f}_{\mu, \beta, \sigma^2, \sigma_u^2; f_1}^{1/2}(\mathbf{y})|_{\sigma_u^2=0} := \frac{1}{4\sigma^2} \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{y_t - \mu - \beta' \mathbf{x}_t}{\sigma} \right) + \sum_{t=1}^T \sum_{l=1, l \neq t}^T \phi_{f_1} \left( \frac{y_t - \mu - \beta' \mathbf{x}_t}{\sigma} \right) \phi_{f_1} \left( \frac{y_l - \mu - \beta' \mathbf{x}_l}{\sigma} \right) \right].$$

Then, as  $t, s, v,$  and  $r \rightarrow 0,$

$$(i) \int \left\{ \underline{f}_{\mu+t, \beta+s, \sigma^2+v, r^2; f_1}^{1/2}(\mathbf{y}) - \underline{f}_{\mu+t, \beta+s, \sigma^2+v, 0; f_1}^{1/2}(\mathbf{y}) - r^2 D_{\sigma_u^2} \underline{f}_{\mu+t, \beta+s, \sigma^2+v, \sigma_u^2; f_1}^{1/2}(\mathbf{y})|_{\sigma_u^2=0} \right\}^2 d\mathbf{y} = o(r^4),$$

$$(ii) \int \left\{ \underline{f}_{\mu+t, \beta+s, \sigma^2+v, 0; f_1}^{1/2}(\mathbf{y}) - \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) - (t, \mathbf{s}', v) \begin{pmatrix} D_{\mu} \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \\ D_{\beta} \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \\ D_{\sigma^2} \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \end{pmatrix} \right\}^2 d\mathbf{y} = o(\|(t, \mathbf{s}', v)\|^2),$$

$$(iii) \int \left\{ D_{\sigma_u^2} \underline{f}_{\mu+t, \beta+s, \sigma^2+v, \sigma_u^2; f_1}^{1/2}(\mathbf{y})|_{\sigma_u^2=0} - D_{\sigma_u^2} \underline{f}_{\mu, \beta, \sigma^2, \sigma_u^2; f_1}^{1/2}(\mathbf{y})|_{\sigma_u^2=0} \right\}^2 d\mathbf{y} = o(1), \text{ and}$$

$$(iv) \int \left\{ \underline{f}_{\mu+t, \beta+s, \sigma^2+v, r^2; f_1}^{1/2}(\mathbf{y}) - \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) - (t, \mathbf{s}', v, r^2) \begin{pmatrix} D_{\mu} \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \\ D_{\beta} \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \\ D_{\sigma^2} \underline{f}_{\mu, \beta, \sigma^2, 0; f_1}^{1/2}(\mathbf{y}) \\ D_{\sigma_u^2} \underline{f}_{\mu, \beta, \sigma^2, \sigma_u^2; f_1}^{1/2}(\mathbf{y})|_{\sigma_u^2=0} \end{pmatrix} \right\}^2 d\mathbf{y} = o(\|(t, \mathbf{s}', v, r^2)\|^2).$$

**Proof.** (i) Letting  $z_t := (y_t - (\mu + t) - (\beta + \mathbf{s}')' \mathbf{x}_t) / (\sigma^2 + v)^{1/2}$  and  $\mathbf{z} := (z_1, z_2, \dots, z_T)'$ , the left-hand side in (i) takes the form

$$\int_{\mathbb{R}^T} \left[ \left( \int_{\mathbb{R}} \prod_{t=1}^T f_1 \left( z_t - \frac{r}{(\sigma^2 + v)^{1/2}} u \right) h(u) du \right)^{1/2} - \left( \prod_{t=1}^T f_1(z_t) \right)^{1/2} - \frac{r^2}{4(\sigma^2 + v)} \left( \prod_{t=1}^T f_1(z_t) \right)^{1/2} \left( \sum_{t=1}^T \psi_{f_1}(z_t) + \sum_{t=1}^T \sum_{l=1, l \neq t}^T \phi_{f_1}(z_t) \phi_{f_1}(z_l) \right) \right]^2 d\mathbf{z}.$$

In order to prove (i), it is thus sufficient to establish differentiability in quadratic mean with respect to  $(r/(\sigma^2 + v)^{1/2})^2$ . This quadratic mean differentiability property, however, is somewhat nonstandard, as it involves the second-order derivatives of the product  $\left[ \prod_{t=1}^T f_1(z_t) \right]$ . As in Akharif and Hallin (2003), the proof is decomposed into three parts.

(a) With the above notation,  $y^2 \mapsto L_{\mathbf{z}}(y) := \int_{\mathbb{R}} K_{\mathbf{z}}(u, y) h(u) du$ , where  $K_{\mathbf{z}}(u, y) := \prod_{t=1}^T f_1(z_t - yu)$ , is absolutely continuous at  $y = 0$ , with a.e. derivative

$$\begin{aligned} \Psi_{\mathbf{z}}(y) &:= \frac{1}{2y} \int_{w=0}^y \int_{\mathbb{R}} \left\{ \sum_{t=1}^T \ddot{f}_1(z_t - wu) \left[ \prod_{\substack{l=1 \\ l \neq t}}^T f_1(z_l - wu) \right] \right. \\ &\quad \left. + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \dot{f}_1(z_t - wu) \dot{f}_1(z_l - wu) \left[ \prod_{\substack{m=1 \\ m \neq t, m \neq l}}^T f_1(z_m - wu) \right] \right\} u^2 h(u) du dw. \quad (6.1) \end{aligned}$$

We obtain

$$\begin{aligned}
L(y) - L(0) &= \int_{u=-\infty}^{\infty} [K(u, y) - K(u, 0)]h(u) du \\
&= \int_{u=-\infty}^{\infty} \int_{a=0}^y \dot{K}(u, a) da h(u) du \\
&= \int_{u=-\infty}^{\infty} \int_{a=0}^y [\dot{K}(u, a) - \dot{K}(u, 0)]da h(u) du + \int_{u=-\infty}^{\infty} \int_{a=0}^y \dot{K}(u, 0)da h(u) du \\
&= \int_{u=-\infty}^{\infty} \int_{a=0}^y \int_{w=0}^a \ddot{K}(u, w)dw da h(u) du, \tag{6.2}
\end{aligned}$$

where

$$\begin{aligned}
\ddot{K}(u, w) &:= \sum_{t=1}^T \ddot{f}_1(z_t - wu) \left[ \prod_{\substack{l=1 \\ l \neq t}}^T f_1(z_l - wu) \right] u^2 \\
&\quad + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \dot{f}_1(z_t - wu) \dot{f}_1(z_l - wu) \left[ \prod_{\substack{m=1 \\ m \neq t, m \neq l}}^T f_1(z_m - wu) \right] u^2.
\end{aligned}$$

The value (6.1) of the a.e. derivative for  $y > 0$  follows. At  $y = 0$ , the right derivative is defined as the limit, as  $y \downarrow 0$ , of  $[L(y) - L(0)]/y^2$ , for which (6.2) yields  $0/0$ . Applying L'Hospital's rule we have

$$\lim_{y \downarrow 0} [L(y) - L(0)]/y^2 = \frac{1}{2} \left\{ \sum_{t=1}^T \psi_{f_1}(z_t) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1}(z_t) \phi_{f_1}(z_l) \right\} \left[ \prod_{m=1}^T f_1(z_m) \right].$$

(b) It follows that  $y^2 \mapsto s_{\mathbf{z}}(y) := [L_{\mathbf{z}}(y)]^{\frac{1}{2}}$  is also absolutely continuous in a neighborhood of  $y = 0$ , with a.e. derivative

$$\dot{s}_{\mathbf{z}}(y) = \frac{1}{4y} \int_{w=0}^y \int_{\mathbb{R}} \ddot{K}(u, w)h(u) du / \left[ \int_{\mathbb{R}} K(u, y)h(u) du \right]^{\frac{1}{2}} dw. \tag{6.3}$$

L'Hospital's rule at  $y = 0$  yields

$$\dot{s}_{\mathbf{z}}(0) = \frac{1}{4} \left( \prod_{l=1}^T f_1(z_l) \right)^{-\frac{1}{2}} \left\{ \sum_{t=1}^T \ddot{f}_1(z_t) \left[ \prod_{\substack{l=1 \\ l \neq t}}^T f_1(z_l) \right] + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \dot{f}_1(z_t) \dot{f}_1(z_l) \left[ \prod_{\substack{m=1 \\ m \neq t, m \neq l}}^T f_1(z_m) \right] \right\}.$$

Hence, for all  $\mathbf{z}$ ,

$$\lim_{y \rightarrow 0} \left\{ \frac{1}{y^2} [s_{\mathbf{z}}(y) - s_{\mathbf{z}}(0)] \right\} = \dot{s}_{\mathbf{z}}(0). \tag{6.4}$$

(c) The partial quadratic mean differentiability property to be proved takes the form

$$\lim_{y \rightarrow 0} \int_{\mathbb{R}^T} \left\{ \frac{1}{y^2} [s_{\mathbf{z}}(y) - s_{\mathbf{z}}(0)] - \frac{1}{4} \dot{s}_{\mathbf{z}}(0) \right\}^2 d\mathbf{z} = 0. \tag{6.5}$$

From (b) above,

$$\left\{ \frac{1}{y^2} [s_{\mathbf{z}}(y) - s_{\mathbf{z}}(0)] \right\}^2 = \left( \frac{1}{y^2} \right)^2 \left( \int_{\lambda=0}^{y^2} \dot{s}_{\mathbf{z}}(\sqrt{\lambda}) d\lambda \right)^2 \leq \frac{1}{y^2} \int_{\lambda=0}^{y^2} [\dot{s}_{\mathbf{z}}(\sqrt{\lambda})]^2 d\lambda,$$

for all  $\mathbf{z}$ . Fubini's theorem and (6.3) yields

$$\begin{aligned} \int_{\mathbb{R}^T} \left\{ \frac{1}{y^2} [s_{\mathbf{z}}(y) - s_{\mathbf{z}}(0)] \right\}^2 d\mathbf{z} &\leq \frac{1}{y^2} \int_{\lambda=0}^{y^2} \int_{\mathbb{R}^T} [\dot{s}_{\mathbf{z}}(\sqrt{\lambda})]^2 d\mathbf{z} d\lambda \\ &= \frac{1}{16y^2} \int_{\lambda=0}^{y^2} \mathcal{I}_{\psi\phi}(f_1; \sqrt{\lambda}) d\lambda, \end{aligned} \quad (6.6)$$

with  $\mathcal{I}_{\psi\phi}(f_1; y)$  defined in (2.3). The continuity assumption in (B3) implies that this latter quantity converges, as  $y \rightarrow 0$ , to

$$\frac{1}{16} \mathcal{I}_{\psi\phi}(f_1; 0) = \int_{\mathbb{R}^T} [\dot{s}_{\mathbf{z}}(0)]^2 d\mathbf{z},$$

which together with (6.6), entails that

$$\limsup_{y \rightarrow 0} \int_{\mathbb{R}^T} \left\{ \frac{1}{y^2} [s_{\mathbf{z}}(y) - s_{\mathbf{z}}(0)] \right\}^2 d\mathbf{z} \leq \int_{\mathbb{R}^T} [\dot{s}_{\mathbf{z}}(0)]^2 d\mathbf{z}. \quad (6.7)$$

In view of Theorem V.I.3 of Hájek and Šidák (1967) [also in Hájek, Šidák and Sen (1999)], (6.4) and (6.7) jointly imply (6.5). This completes the proof of part (i) of the lemma.

(ii) The problem here reduces to the classical case of linear models considered by Swensen (1985).

(iii) First note that, as  $t, \mathbf{s} \rightarrow 0$ ,

$$\int_{\mathbb{R}^T} \left\{ D_{\sigma_u^2} f_{\mu+t, \boldsymbol{\beta}+\mathbf{s}, \sigma_u^2, \sigma_u^2; f_1}(\mathbf{y})|_{\sigma_u^2=0} - D_{\sigma_u^2} f_{\mu, \boldsymbol{\beta}, \sigma_u^2, \sigma_u^2; f_1}(\mathbf{y})|_{\sigma_u^2=0} \right\}^2 d\mathbf{y} = o(1).$$

For the perturbation of  $\sigma^2$ , letting  $z_t := y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t$  for  $t = 1, \dots, T$ , we have

$$\begin{aligned} T &:= \int_{\mathbb{R}^T} \left\{ D_{\sigma_u^2} f_{\mu, \boldsymbol{\beta}, \sigma^2+h, \sigma_u^2; f_1}(\mathbf{y})|_{\sigma_u^2=0} - D_{\sigma_u^2} f_{\mu, \boldsymbol{\beta}, \sigma^2, \sigma_u^2; f_1}(\mathbf{y})|_{\sigma_u^2=0} \right\}^2 d\mathbf{y} \\ &= \int_{\mathbb{R}^T} \left\{ \frac{1}{4} \frac{1}{(\sigma^2+h)^{T+4/4}} \left[ \prod_{t=1}^T f_1^{1/2} \left( \frac{y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t}{(\sigma^2+h)^{1/2}} \right) \right] \right. \\ &\quad \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t}{(\sigma^2+h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t}{(\sigma^2+h)^{1/2}} \right) \phi_{f_1} \left( \frac{y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t}{(\sigma^2+h)^{1/2}} \right) \right] \\ &\quad \left. - \frac{1}{4} \frac{1}{(\sigma^2)^{T+4/4}} \left[ \prod_{t=1}^T f_1^{1/2} \left( \frac{y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t}{\sigma} \right) \right] \right. \\ &\quad \left. \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t}{\sigma} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{y_t - \mu - \boldsymbol{\beta}'\mathbf{x}_t}{\sigma} \right) \phi_{f_1} \left( \frac{y_t - \mu - x_l \beta}{\sigma} \right) \right] \right\}^2 d\mathbf{y} \\ &= \int_{\mathbb{R}^T} \left\{ \frac{1}{4} \frac{1}{(\sigma^2+h)^{T+4/4}} \left[ \prod_{t=1}^T f_1^{1/2} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) \right] \right. \\ &\quad \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) \right] \\ &\quad \left. - \frac{1}{4} \frac{1}{(\sigma^2)^{T+4/4}} \left[ \prod_{t=1}^T f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{\sigma} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{\sigma} \right) \phi_{f_1} \left( \frac{z_l}{\sigma} \right) \right] \right\}^2 d\mathbf{z} \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^T} \left\{ \frac{1}{4} \left[ \frac{1}{(\sigma^2 + h)} \prod_{t=1}^T \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \frac{1}{\sigma^2} \prod_{t=1}^T \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \right. \\
&\quad \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) \right] \\
&\quad + \frac{1}{4} \frac{1}{\sigma^2} \left[ \prod_{t=1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \left\{ \sum_{t=1}^T \left[ \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \psi_{f_1} \left( \frac{z_t}{\sigma} \right) \right] \right. \\
&\quad \left. \left. + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \left[ \phi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) - \phi_{f_1} \left( \frac{z_t}{\sigma} \right) \phi_{f_1} \left( \frac{z_l}{\sigma} \right) \right] \right\}^2 d\mathbf{z} \\
&= \int_{\mathbb{R}^T} \left\{ \frac{1}{4} \left[ \frac{1}{(\sigma^2 + h)} - \frac{1}{\sigma^2} \right] \prod_{t=1}^T \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \right. \\
&\quad \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) \right] \\
&\quad + \frac{1}{4} \left[ \frac{1}{\sigma^2} \prod_{t=1}^T \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \frac{1}{\sigma^2} \prod_{t=1}^T \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \\
&\quad \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) \right] \\
&\quad + \frac{1}{4\sigma^2} \left[ \prod_{t=1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \left\{ \sum_{t=1}^T \left[ \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \psi_{f_1} \left( \frac{z_t}{\sigma} \right) \right] \right. \\
&\quad \left. + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \left[ \phi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) - \phi_{f_1} \left( \frac{z_t}{\sigma} \right) \phi_{f_1} \left( \frac{z_l}{\sigma} \right) \right] \right\}^2 d\mathbf{z} \\
&\leq C(T_1 + T_2 + T_3),
\end{aligned}$$

where

$$\begin{aligned}
T_1 &:= \int_{\mathbb{R}^T} \left\{ \frac{1}{4} \left[ \frac{1}{(\sigma^2 + h)} - \frac{1}{\sigma^2} \right] \left[ \prod_{t=1}^T \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \right] \right. \\
&\quad \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) \right] \left. \right\}^2 d\mathbf{z},
\end{aligned}$$

$$\begin{aligned}
T_2 &:= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \left[ \prod_{t=1}^T \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \prod_{t=1}^T \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \right. \\
&\quad \times \left[ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) \right] \left. \right\}^2 d\mathbf{z},
\end{aligned}$$

and

$$\begin{aligned}
T_3 &:= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \left[ \prod_{t=1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \left\{ \sum_{t=1}^T \left[ \psi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) - \psi_{f_1} \left( \frac{z_t}{\sigma} \right) \right] \right. \right. \\
&\quad \left. \left. + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \left[ \phi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2+h)^{1/2}} \right) - \phi_{f_1} \left( \frac{z_t}{\sigma} \right) \phi_{f_1} \left( \frac{z_l}{\sigma} \right) \right] \right\} \right\}^2 d\mathbf{z}.
\end{aligned}$$

Clearly,  $T_1 = O([\sigma^2 + h]^{-1} - \sigma^{-2})^2 (\mathcal{I}_\psi(f_1) + \mathcal{I}_\phi^2(f_1))$ , which implies that  $T_1 = o(1)$ , as  $h \rightarrow 0$ . Turning to  $T_2$ , we have

$$\begin{aligned}
T_2 &= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \left\{ \sum_{t=1}^T \left[ \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) - \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \right. \right. \\
&\quad \times \prod_{k=1}^{t-1} \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2+h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\} \\
&\quad \times \left\{ \sum_{t=1}^T \psi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) + \sum_{t=1}^T \sum_{\substack{l=1 \\ l \neq t}}^T \phi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2+h)^{1/2}} \right) \right\} \right\}^2 d\mathbf{z} \\
&= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \left\{ \sum_{t=1}^T \sum_{l=1}^T \left[ \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) - \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \right. \right. \\
&\quad \times \psi_{f_1} \left( \frac{z_l}{(\sigma^2+h)^{1/2}} \right) \prod_{k=1}^{t-1} \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2+h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\} \\
&\quad + \frac{1}{4\sigma^2} \sum_{t=1}^T \sum_{l=1}^T \sum_{\substack{m=1 \\ m \neq l}}^T \left[ \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) - \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \phi_{f_1} \left( \frac{z_l}{(\sigma^2+h)^{1/2}} \right) \\
&\quad \times \phi_{f_1} \left( \frac{z_m}{(\sigma^2+h)^{1/2}} \right) \prod_{k=1}^{t-1} \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2+h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\} \right\}^2 d\mathbf{z} \\
&\leq C_1 (T_2^1 + T_2^2),
\end{aligned}$$

where

$$\begin{aligned}
T_2^1 &:= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \left\{ \sum_{t=1}^T \sum_{l=1}^T \left[ \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) - \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \right. \right. \\
&\quad \times \psi_{f_1} \left( \frac{z_l}{(\sigma^2+h)^{1/2}} \right) \prod_{k=1}^{t-1} \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2+h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\} \right\}^2 d\mathbf{z}
\end{aligned}$$

and

$$\begin{aligned}
T_2^2 &:= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \sum_{t=1}^T \sum_{l=1}^T \sum_{\substack{m=1 \\ m \neq l}}^T \left[ \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) - \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \phi_{f_1} \left( \frac{z_l}{(\sigma^2+h)^{1/2}} \right) \right. \\
&\quad \times \phi_{f_1} \left( \frac{z_m}{(\sigma^2+h)^{1/2}} \right) \prod_{k=1}^{t-1} \frac{1}{(\sigma^2+h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2+h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\} \right\}^2 d\mathbf{z}.
\end{aligned}$$

Then, in order to prove that  $T_2 = o(1)$ , it is clearly sufficient to show that  $T_2^1$  and  $T_2^2$  are  $o(1)$  as  $h \rightarrow 0$ . We start with  $T_2^1$ , which is bounded by  $B(T_2^{11} + T_2^{12} + T_2^{13})$ , where  $B$  is some positive constant,

$$\begin{aligned} T_2^{11} &:= \int_{\mathbb{R}^T} \left\{ \sum_{t=1}^T \left[ \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \psi_{f_1} \left( \frac{z_t}{\sigma} \right) \right] \right. \\ &\quad \times \prod_{k=1}^{t-1} \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2 + h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\}^2 dz \\ &\leq B_1 \int_{\mathbb{R}} \left\{ e^{\frac{1}{2} \left[ u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2} \right]} f_1^{1/2} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) \psi_{f_1} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) \right. \\ &\quad \left. - e^{\frac{1}{2} u} f_1^{1/2} (e^u) \psi_{f_1} (e^u) \right\}^2 du, \quad (6.8) \end{aligned}$$

$$\begin{aligned} T_2^{12} &:= \int_{\mathbb{R}} \left\{ \sum_{t=1}^T \left[ \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \psi_{f_1} \left( \frac{z_t}{\sigma} \right) \right] \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right. \\ &\quad \times \prod_{k=1}^{t-1} \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2 + h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\}^2 dz \\ &\leq B_2 \int_{\mathbb{R}} \left\{ \psi_{f_1} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) - \psi_{f_1} (e^u) \right\}^2 e^u f_1 (e^u) du \quad (6.9) \end{aligned}$$

and

$$\begin{aligned} T_2^{13} &:= \int_{\mathbb{R}^T} \left\{ 2 \sum_{t=2}^T \sum_{l=1}^{t-1} \left[ \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \frac{1}{\sigma^{1/2}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \right. \\ &\quad \times \psi_{f_1} \left( \frac{z_l}{(\sigma^2 + h)^{1/2}} \right) \prod_{k=1}^{t-1} \frac{1}{(\sigma^2 + h)^{1/4}} f_1^{1/2} \left( \frac{z_k}{(\sigma^2 + h)^{1/2}} \right) \prod_{k=t-1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_k}{\sigma} \right) \left. \right\}^2 dz \\ &\leq B_3 \int_{\mathbb{R}} \left\{ e^{\frac{1}{2} \left[ u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2} \right]} f_1^{1/2} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) - e^{\frac{1}{2} u} f_1^{1/2} (e^u) \right\}^2 du \times \mathcal{I}_\psi(f_1). \quad (6.10) \end{aligned}$$

Since  $e^{\frac{1}{2} u} f_1^{1/2} (e^u)$ ,  $e^{\frac{1}{2} u} f_1^{1/2} (e^u) \psi_{f_1} (e^u)$  and  $\psi_{f_1} (e^u)$  are square integrable, quadratic mean continuity implies that the integrals in (6.8), (6.9), and (6.10) are  $o(1)$  as  $h \rightarrow 0$ .

Similarly, it easily shown that

$$\begin{aligned} T_2^2 &\leq C \left[ \int_{\mathbb{R}} \left\{ e^{\frac{1}{2} \left[ u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2} \right]} f_1^{1/2} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) \phi_{f_1} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) - e^{\frac{1}{2} u} f_1^{1/2} (e^u) \psi_{f_1} (e^u) \right\}^2 du \right. \\ &\quad \times \mathcal{I}_\phi(f_1) + \int_{\mathbb{R}} \left\{ \phi_{f_1} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) - \phi_{f_1} (e^u) \right\}^2 e^u f_1 (e^u) du \times \mathcal{I}_\phi(f_1) \\ &\quad \left. + \int_{\mathbb{R}} \left\{ e^{\frac{1}{2} \left[ u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2} \right]} f_1^{1/2} \left( e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right) - e^{\frac{1}{2} u} f_1^{1/2} (e^u) \right\}^2 du \times \mathcal{I}_\phi^2(f_1) \right] = o(1), \text{ as } h \rightarrow 0, \end{aligned}$$

since  $e^{\frac{1}{2} u} f_1^{1/2} (e^u)$ ,  $e^{\frac{1}{2} u} f_1^{1/2} (e^u) \psi_{f_1} (e^u)$  and  $\phi_{f_1} (e^u)$  are square integrable.

As for  $T_3$ , note that  $T_3 \leq D(T_3^1 + T_3^2)$  where

$$\begin{aligned} T_3^1 &:= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \left[ \prod_{t=1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \left\{ \sum_{t=1}^T \left[ \psi_{f_1} \left( \frac{z_t}{(\sigma^2 + h)^{1/2}} \right) - \psi_{f_1} \left( \frac{z_t}{\sigma} \right) \right] \right\}^2 \right. \\ &\quad \left. \right\}^2 dz \\ &\leq D_1 \int_{\mathbb{R}} \left( \psi_{f_1} \left\{ e^{u - \ln \left( 1 + \frac{h}{\sigma^2} \right)^{1/2}} \right\} - \psi_{f_1} (e^u) \right)^2 e^u f_1 (e^u) du = o(1), \text{ as } h \rightarrow 0, \end{aligned}$$

and

$$\begin{aligned}
T_3^2 &:= \int_{\mathbb{R}^T} \left\{ \frac{1}{4\sigma^2} \left[ \prod_{t=1}^T \frac{1}{(\sigma^2)^{1/4}} f_1^{1/2} \left( \frac{z_t}{\sigma} \right) \right] \left\{ \sum_{t=1}^T \sum_{1=l \neq t}^T \left[ \phi_{f_1} \left( \frac{z_t}{(\sigma^2+h)^{1/2}} \right) \phi_{f_1} \left( \frac{z_l}{(\sigma^2+h)^{1/2}} \right) \right. \right. \right. \\
&\quad \left. \left. \left. - \phi_{f_1} \left( \frac{z_t}{\sigma} \right) \phi_{f_1} \left( \frac{z_l}{\sigma} \right) \right] \right\} \right\}^2 d\mathbf{z} \\
&\leq D_2 \int_{\mathbb{R}} \left\{ \phi_{f_1} \left( e^{u - \ln(1 + \frac{h}{\sigma^2})^{1/2}} \right) - \phi_{f_1}(e^u) \right\}^2 e^u f_1(e^u) du = o(1), \text{ as } h \rightarrow 0,
\end{aligned}$$

since  $\phi_{f_1}(e^u)$  and  $\psi_{f_1}(e^u)$  are square integrable. This completes the proof of Lemma 6.1, hence of Proposition 2.1.  $\square$

## 6.2 Proofs of Propositions 3.2, 4.1, and 4.2.

**Proof of Proposition 4.1.** The proof of Part (i) follows along the same lines as in Section 4.1 of Hallin, Ingenbleek, and Puri (1985), and therefore is omitted. Part (ii) of the proposition is obtained by direct computation. As for Part (iii), under  $P_{\boldsymbol{\vartheta};g_1}^{(n)}$ , the result straightforwardly follows from classical central limit theorems. On the other hand, it is easy to see that, still under  $P_{\boldsymbol{\vartheta};g_1,h_1}^{(n)}$ ,  $\Delta_{f_1,g_1;4}^{*(n)}(\boldsymbol{\vartheta})$  and the log-likelihood  $\Lambda_{\boldsymbol{\vartheta} + n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau};g_1}^{(n)}$  are jointly multinormal; the desired result then follows from a routine application of Le Cam's Third Lemma.  $\square$

**Proofs of Propositions 3.2 and 4.2.** Since (letting  $K(u) := (a^{1/2}G_1^{-1}(u))^2$ ,  $u \in (0,1)$ ) the proof of Proposition 3.2 is essentially similar to that of Proposition 4.2, we safely can focus on the latter. Throughout this proof, we let  $\boldsymbol{\theta}^{(n)} := \boldsymbol{\theta} + \boldsymbol{\nu}_1^{(n)} \mathbf{t}^{(n)}$ . Accordingly, let  $Z_{it}^0 := \sigma^{-1}(Y_{it} - \underline{x}'_{it} \boldsymbol{\theta})$  and  $Z_{it}^n := \sigma^{-1}(Y_{it} - \underline{x}'_{it} \boldsymbol{\theta}^{(n)})$ .

Proposition 4.1 implies that  $\Delta_{f_1;4}^{*(n)}(\boldsymbol{\theta}, \sigma^2, 0) - \Delta_{f_1,g_1;4}^{*(n)}(\boldsymbol{\theta}, \sigma^2, 0)$  is  $o_P(1)$  under  $P_{\boldsymbol{\theta},\sigma^2,0;g_1}^{(n)}$ . Similarly,  $\Delta_{f_1;4}^{*(n)}(\boldsymbol{\theta}^{(n)}, \sigma^2, 0) - \Delta_{f_1,g_1;4}^{*(n)}(\boldsymbol{\theta}^{(n)}, \sigma^2, 0)$  is  $o_P(1)$  under  $P_{\boldsymbol{\theta}^{(n)},\sigma^2,0;g_1}^{(n)}$  – hence, from contiguity, also under  $P_{\boldsymbol{\vartheta};g_1}^{(n)}$ . Consequently, the left hand side in (4.4) is asymptotically equivalent, under  $P_{\boldsymbol{\theta},\sigma^2,0;g_1}^{(n)}$ , to

$$\begin{aligned}
C^{(n)}(\boldsymbol{\theta}, \sigma^2) &:= \Delta_{f_1,g_1;4}^{*(n)}(\boldsymbol{\theta}^{(n)}, \sigma^2, 0) - \Delta_{f_1,g_1;4}^{*(n)}(\boldsymbol{\theta}, \sigma^2, 0) \\
&= n^{-1/2} \frac{1}{2\sigma^2} \sum_{i=1}^n \left\{ \sum_{t=1}^T \sum_{l \neq t=1}^T \phi_{f_1} \left( F_1^{-1}(G_1(Z_{it}^n)) \right) \phi_{f_1} \left( F_1^{-1}(G_1(Z_{il}^n)) \right) \right. \\
&\quad \left. - \sum_{t=1}^T \sum_{l \neq t=1}^T \phi_{f_1} \left( F_1^{-1}(G_1(Z_{it}^0)) \right) \phi_{f_1} \left( F_1^{-1}(G_1(Z_{il}^0)) \right) \right\}.
\end{aligned}$$

and we need only to prove that  $C^{(n)}(\boldsymbol{\theta}, \sigma^2)$  is  $o_P(1)$ .

Let  $K_{f_1}(u) := \phi_{f_1}(F_1^{-1}(u))$  and consider the following truncation: for all  $\ell \in \mathbb{N}_0$ , define

$$\begin{aligned}
K_{f_1}^{(\ell)}(u) &:= K_{f_1} \left( \frac{2}{\ell} \right) \ell \left( u - \frac{1}{\ell} \right) I_{([\frac{1}{\ell} < u \leq \frac{2}{\ell}])} + K_{f_1}(u) I_{([\frac{2}{\ell} < u \leq 1 - \frac{2}{\ell}])} \\
&\quad + K_{f_1} \left( 1 - \frac{2}{\ell} \right) \ell \left( \left( 1 - \frac{1}{\ell} \right) - u \right) I_{([1 - \frac{2}{\ell} < u \leq 1 - \frac{1}{\ell}])}.
\end{aligned}$$

Continuity of  $u \mapsto K_{f_1}(u)$  implies the continuity of  $u \mapsto K_{f_1}^{(\ell)}(u)$  on the interval  $(0,1)$ . It follows that the truncated scores  $K_{f_1}^{(\ell)}$  are bounded for all  $\ell \in \mathbb{N}_0$ . Moreover, since we have

assumed that  $K_{f_1}$  is monotone increasing, there exists some  $L$  such that  $|K_{f_1}^{(\ell)}(u)| \leq |K_{f_1}(u)|$  for all  $u \in (0, 1)$  and all  $\ell \geq L$ .

One shows easily that  $C^{(n)}(\boldsymbol{\theta}, \sigma^2)$  decomposes into  $D_1^{(n;\ell)} + D_2^{(n;\ell)} - R_1^{(n;\ell)} + R_2^{(n;\ell)}$ , where, denoting by  $E_0$  and  $\text{Var}_0$  expectation and variance, respectively, under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ ,

$$\begin{aligned} D_1^{(n;\ell)} &:= \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) \\ &\quad - \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \\ &\quad - \frac{n^{-1/2}}{2\sigma^2} E_0 \left[ \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) \right], \end{aligned}$$

$$D_2^{(n;\ell)} := \frac{n^{-1/2}}{2\sigma^2} E_0 \left[ \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) \right],$$

$$\begin{aligned} R_1^{(n;\ell)} &:= \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}(G_1(Z_{it}^0)) K_{f_1}(G_1(Z_{il}^0)) \\ &\quad - \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)), \end{aligned}$$

and

$$\begin{aligned} R_2^{(n;\ell)} &:= \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}(G_1(Z_{it}^n)) K_{f_1}(G_1(Z_{il}^n)) \\ &\quad - \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)). \end{aligned}$$

We prove that  $C^{(n)}(\boldsymbol{\theta}, \sigma^2) = o_{\mathbb{P}}(1)$  by establishing that  $D_1^{(n;\ell)}$  and  $D_2^{(n;\ell)}$  are  $o_{\mathbb{P}}(1)$  under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ ,  $n \rightarrow \infty$ , for fixed  $\ell$  and that  $R_1^{(n;\ell)}$  and  $R_2^{(n;\ell)}$  are  $o_{\mathbb{P}}(1)$  under the same sequence of distributions, as  $\ell \rightarrow \infty$ , uniformly in  $n$ . For the sake of convenience, these three results are treated separately (Lemmas 6.2, 6.3 and 6.4)

Decompose  $D_1^{(n;\ell)}$  into  $D_{1,1}^{(n;\ell)} + D_{1,2}^{(n;\ell)} - E_0[D_{1,1}^{(n;\ell)}]$ , where

$$\begin{aligned} D_{1,1}^{(n;\ell)} &:= \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) \\ &\quad - \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \end{aligned}$$

and

$$\begin{aligned} D_{1,2}^{(n;\ell)} &:= \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) \\ &\quad - \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)). \end{aligned}$$

(taking into account the independence between  $Z_{il}^0$  and  $(Z_{it}^0, Z_{it}^n)$  under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ ). We then have the following.

**Lemma 6.2** *For any fixed  $\ell$ ,*

$$(i) \quad \mathbb{E}_0 \left[ \left( D_{1,1}^{(n;\ell)} - \mathbb{E}_0[D_{1,1}^{(n;\ell)}] \right)^2 \right] = o(1), \text{ as } n \rightarrow \infty;$$

$$(ii) \quad \mathbb{E}_0 \left[ \left( D_{1,2}^{(n;\ell)} \right)^2 \right] = o(1), \text{ as } n \rightarrow \infty;$$

$$(iii) \quad D_1^{(n;\ell)} = o_{\mathbb{P}}(1), \text{ as } n \rightarrow \infty, \text{ under } \mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}.$$

**Lemma 6.3** *For any fixed  $\ell$ ,  $D_2^{(n;\ell)} = o(1)$ , as  $n \rightarrow \infty$ , under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ .*

**Lemma 6.4** *As  $\ell \rightarrow \infty$ , uniformly in  $n$ ,*

$$(a) \quad R_1^{(n;\ell)} \text{ is } o_{\mathbb{P}}(1) \text{ under } \mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)},$$

$$(b) \quad R_2^{(n;\ell)} \text{ is } o_{\mathbb{P}}(1) \text{ under } \mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)} \text{ for } n \text{ sufficiently large.}$$

**Proof of Lemma 6.2** Let us begin with the second part of Lemma 6.2.

(ii) First note that  $D_{1,2}^{(n;\ell)} := (n^{-1/2}/2\sigma^2) \sum_{i=1}^n \mathbb{T}_i^{(n;\ell)}$ , where, for  $i = 1, \dots, n$ ,

$$\begin{aligned} \mathbb{T}_i^{(n;\ell)} &:= \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) \left[ K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right] \\ &= 2 \sum_{t=1}^{T-1} \sum_{l=t+1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) \left[ K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right]. \end{aligned}$$

are i.i.d. Using the independence, for  $t \neq l$ , between  $Z_{1t}^0$  and  $(Z_{1l}^0, Z_{1l}^n)$  and the boundedness

of  $K_{f_1}^{(\ell)}$ , we have that

$$\begin{aligned}
\mathbb{E}_0 \left[ \left( D_{1,2}^{(n;\ell)} \right)^2 \right] &= \frac{n^{-1}}{4\sigma^4} \sum_{i,j=1}^n \mathbb{E}_0 \left[ T_i^{(n;\ell)} T_j^{(n;\ell)} \right] = \frac{n^{-1}}{4\sigma^4} \sum_{i=1}^n \mathbb{E}_0 \left[ \left( T_i^{(n;\ell)} \right)^2 \right] \\
&= \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \left\{ \sum_{t=1}^{T-1} \sum_{l=t+1}^T \mathbb{E}_0 \left[ \left( K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) \right)^2 \right] \mathbb{E}_0 \left[ \left( K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right)^2 \right] \right. \\
&\quad \left. + \sum_{t=1}^{T-1} \sum_{l \neq k=t+1}^T \mathbb{E}_0 \left[ \left( K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) \right)^2 \right] \right. \\
&\quad \left. \times \mathbb{E}_0 \left[ \left( K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right) \left( K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0)) \right) \right] \right\} \\
&\leq C_1 \max_{1 \leq k \leq T} \left\{ \mathbb{E}_0 \left[ \left( K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0)) \right)^2 \right] \right\}.
\end{aligned}$$

Then, it only remains to be shown that  $\mathbb{E}_0 \left[ \left( K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0)) \right)^2 \right] = o(1)$ , as  $n \rightarrow \infty$ , uniformly in  $k$ . Continuity of  $K_{f_1}^{(\ell)} \circ G_1$  and the fact that  $|Z_{il}^n - Z_{il}^0|$  is  $o_P(1)$  together imply that  $K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) = o_P(1)$ , under  $P_{\theta, \sigma^2, 0; g_1}^{(n)}$ , as  $n \rightarrow \infty$ . Moreover, since  $K_{f_1}^{(\ell)}$  is bounded, this convergence to zero also holds in quadratic mean.

(i) Letting  $T_i^{(n;\ell)} := \sum_{t=1}^{T-1} \sum_{l=t+1}^T \left[ K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) \right] K_{f_1}^{(\ell)}(G_1(Z_{il}^n))$ , we have

$$D_{1,1}^{(n;\ell)} - \mathbb{E}_0[D_{1,1}^{(n;\ell)}] = \frac{n^{-1/2}}{\sigma^2} \sum_{i=1}^n \left[ T_i^{(n;\ell)} - \mathbb{E}_0[T_i^{(n;\ell)}] \right],$$

and then

$$\begin{aligned}
\mathbb{E}_0 \left[ \left( D_{1,1}^{(n;\ell)} - \mathbb{E}_0[D_{1,1}^{(n;\ell)}] \right)^2 \right] &= \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \mathbb{E}_0 \left[ \left( T_i^{(n;\ell)} - \mathbb{E}_0[T_i^{(n;\ell)}] \right)^2 \right] = \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \text{Var}_0 \left[ T_i^{(n;\ell)} \right] \\
&\leq \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \mathbb{E}_0 \left[ \left( T_i^{(n;\ell)} \right)^2 \right] = \frac{1}{\sigma^4} \mathbb{E}_0 \left[ \left( T_1^{(n;\ell)} \right)^2 \right],
\end{aligned}$$

and, it only remains to show that  $\mathbb{E}_0\left[\left(\mathbb{T}_1^{(n;\ell)}\right)^2\right] = o(1)$ , as  $n \rightarrow \infty$ :

$$\begin{aligned}
\mathbb{E}_0\left[\left(\mathbb{T}_1^{(n;\ell)}\right)^2\right] &= \sum_{t=1}^{T-1} \mathbb{E}_0\left[\left(K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0))\right)^2\right] \mathbb{E}_0\left[\left(\sum_{l=t+1}^T K_{f_1}^{(\ell)}(G_1(Z_{il}^n))\right)^2\right] \\
&\quad + 2 \sum_{t=2}^{T-1} \sum_{k=1}^{t-1} \sum_{l=t+1}^T \sum_{s=k+1}^T \mathbb{E}_0\left[\left(K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0))\right)\right. \\
&\quad \quad \quad \left.\times \left(K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0))\right) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) K_{f_1}^{(\ell)}(G_1(Z_{is}^n))\right] \\
&= \sum_{t=1}^{T-1} \mathbb{E}_0\left[\left(K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0))\right)^2\right] \mathbb{E}_0\left[\left(\sum_{l=t+1}^T K_{f_1}^{(\ell)}(G_1(Z_{il}^n))\right)^2\right] \\
&\quad + 2 \sum_{t=2}^{T-1} \sum_{k=1}^{t-1} \sum_{l=t+1}^T \mathbb{E}_0\left[\left(K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0))\right)\right. \\
&\quad \quad \quad \left.\times \left(K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0))\right) \left(K_{f_1}^{(\ell)}(G_1(Z_{il}^n))\right)^2\right] \\
&\quad + 2 \sum_{t=2}^{T-1} \sum_{k=1}^{t-1} \sum_{l=t+1}^T \mathbb{E}_0\left[\left(K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0))\right) K_{f_1}^{(\ell)}(G_1(Z_{it}^n))\right. \\
&\quad \quad \quad \left.\times \left(K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0))\right) K_{f_1}^{(\ell)}(G_1(Z_{il}^n))\right] \\
&\quad + 2 \sum_{t=2}^{T-1} \sum_{k=1}^{t-1} \sum_{l=t+1}^T \sum_{s \neq l, s, t=k+1}^T \mathbb{E}_0\left[\left(K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0))\right)\right. \\
&\quad \quad \quad \left.\times \left(K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0))\right) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)) K_{f_1}^{(\ell)}(G_1(Z_{is}^n))\right] \\
&\leq C_2 \max_{1 \leq k \leq T} \left\{ \mathbb{E}_0\left[\left(K_{f_1}^{(\ell)}(G_1(Z_{ik}^n)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0))\right)^2\right] \right\} = o(1).
\end{aligned}$$

(iii) trivially follows from (i)-(ii) and the fact that convergence in quadratic mean implies convergence in probability.  $\square$

**Proof of Lemma 6.3.** Letting

$$B_1^{(n;\ell)} := \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)),$$

one can show that, under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ , as  $n \rightarrow \infty$ ,

$$B_1^{(n;\ell)} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{T(T-1)}{2\sigma^4} \left(\mathbb{E}[(K_{f_1}^{(\ell)}(U))^2]\right)^2\right),$$

where  $U$  stands for a random variable uniformly distributed over  $(0, 1)$ . Defining

$$B_2^{(n;\ell)} := \frac{n^{-1/2}}{2\sigma^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{l \neq t=1}^T K_{f_1}^{(\ell)}(G_1(Z_{it}^n)) K_{f_1}^{(\ell)}(G_1(Z_{il}^n)),$$

it follows from ULAN that, under  $\mathbb{P}_{\boldsymbol{\theta}, \sigma^2, 0; g_1}^{(n)}$ , as  $n \rightarrow \infty$ ,

$$B_2^{(n;\ell)} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{T(T-1)}{2\sigma^4} \left(\mathbb{E}[(K_{f_1}^{(\ell)}(U))^2]\right)^2\right). \quad (6.11)$$

Since  $D_1^{(n;\ell)} = B_2^{(n;\ell)} - B_1^{(n;\ell)} - E_0[B_2^{(n;\ell)}] = o_P(1)$ , we have that

$$B_2^{(n;\ell)} - E_0[B_2^{(n;\ell)}] \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{T(T-1)}{2\sigma^4} \left(E[(K_{f_1}^{(\ell)}(U))^2]\right)^2\right) \quad (6.12)$$

as  $n \rightarrow \infty$ , under  $P_{\theta, \sigma^2, 0; g_1}^{(n)}$ . From (6.11) and (6.12), it follows that  $D_2^{(n;\ell)} = E_0[B_2^{(n;\ell)}]$  is  $o(1)$  as  $n \rightarrow \infty$ .  $\square$

**Proof of Lemma 6.4** (a) We have that

$$\begin{aligned} & E_0[(R_1^{(n;\ell)})^2] \\ &= \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n E_0 \left[ \left( \sum_{t=1}^{T-1} \sum_{l=t+1}^T \left[ K_{f_1}(G_1(Z_{it}^0)) K_{f_1}(G_1(Z_{il}^0)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right] \right)^2 \right] \\ &= \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \sum_{t=1}^{T-1} \sum_{l=t+1}^T E_0 \left[ \left( K_{f_1}(G_1(Z_{it}^0)) K_{f_1}(G_1(Z_{il}^0)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right)^2 \right] \\ &\quad + \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \sum_{t=1}^{T-1} \sum_{t+1 \leq l \neq k \leq T} E_0 \left[ \left( K_{f_1}(G_1(Z_{it}^0)) K_{f_1}(G_1(Z_{il}^0)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right) \right. \\ &\quad \quad \left. \times \left( K_{f_1}(G_1(Z_{it}^0)) K_{f_1}(G_1(Z_{ik}^0)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{ik}^0)) \right) \right] \\ &\quad + \frac{2n^{-1}}{\sigma^4} \sum_{i=1}^n \sum_{t=2}^{T-1} \sum_{k=1}^{t-1} \sum_{l=t+1}^T \sum_{s=k+1}^T E_0 \left[ \left( K_{f_1}(G_1(Z_{it}^0)) K_{f_1}(G_1(Z_{il}^0)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right) \right. \\ &\quad \quad \left. \times \left( K_{f_1}(G_1(Z_{ik}^0)) K_{f_1}(G_1(Z_{is}^0)) - K_{f_1}^{(\ell)}(G_1(Z_{ik}^0)) K_{f_1}^{(\ell)}(G_1(Z_{is}^0)) \right) \right] \\ &=: A_1 + A_2 + A_3. \end{aligned}$$

Using the fact that  $E_0[K_{f_1}^{(\ell)}(G_1(Z_{it}^0))] = o(1)$  as  $\ell \rightarrow \infty$ , uniformly in  $n$  and for all  $t$  and  $i$ , it follows that  $A_2$  and  $A_3$  are  $o(1)$  as  $\ell \rightarrow \infty$ , uniformly in  $n$  and it only remains to show that  $A_1$  is  $o(1)$  as  $\ell \rightarrow \infty$ , uniformly in  $n$ .

Now, we have that

$$\begin{aligned} A_1 &= \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \sum_{t=1}^{T-1} \sum_{l=t+1}^T E_0 \left[ \left( K_{f_1}(G_1(Z_{it}^0)) K_{f_1}(G_1(Z_{il}^0)) - K_{f_1}^{(\ell)}(G_1(Z_{it}^0)) K_{f_1}^{(\ell)}(G_1(Z_{il}^0)) \right)^2 \right] \\ &= \frac{n^{-1}}{\sigma^4} \sum_{i=1}^n \sum_{t=1}^{T-1} \sum_{l=t+1}^T \int_0^1 \int_0^1 \left( K_{f_1}(u) K_{f_1}(v) - K_{f_1}^{(\ell)}(u) K_{f_1}^{(\ell)}(v) \right)^2 du dv \\ &= \frac{T(T-1)}{2\sigma^4} \int_0^1 \int_0^1 \left( K_{f_1}(u) K_{f_1}(v) - K_{f_1}^{(\ell)}(u) K_{f_1}^{(\ell)}(v) \right)^2 du dv. \end{aligned} \quad (6.13)$$

Since  $K_{f_1}^{(\ell)}(u) K_{f_1}^{(\ell)}(v)$  converges to  $K_{f_1}(u) K_{f_1}(v)$  for all  $(u, v) \in (0, 1) \times (0, 1)$ , and since  $|K_{f_1}^{(\ell)}(u)|$  is bounded by  $|K_{f_1}(u)|$  for all  $\ell \geq L$ , the integrand in (6.13) is bounded (uniformly in  $\ell$ ) by  $4|K_{f_1}(u)|^2 |K_{f_1}(v)|^2$ , which is integrable on  $(0, 1) \times (0, 1)$ . The Lebesgue dominated convergence theorem implies that  $A_1 = o(1)$  as  $\ell \rightarrow \infty$ , uniformly in  $n$ .

(b) The claim here is the same as in (a), except that  $Z_{it}^n$  replaces  $Z_{it}^0$ . Accordingly, (b) holds under  $P_{\theta^{(n)}, \sigma_n^2, 0; g_1}^{(n)}$ . That is also holds under  $P_{\theta, \sigma^2, 0; g_1}^{(n)}$  follows from Lemma 3.5 in Jurečková (1969).  $\square$

## References

- [1] Akharif, A., and Hallin, M. (2003). Efficient detection of random coefficients in autoregressive models. *The Annals of Statistics* **31**, 675-704.
- [2] Baltagi, B.H. (2005). *Econometric Analysis of Panel Data*. J. Wiley, Chichester.
- [3] Bickel, P.J., Klaassen, C.A.J., Ritov, Y., and Wellner, J.A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Univ. Press, Baltimore, MD.
- [4] Breusch, T.S., and Pagan, A.R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* **47**, 239-253.
- [5] Cassart, D., Hallin, M., and Paindaveine, D. (2007). A class of optimal tests for symmetry based on Edgeworth approximations. Submitted.
- [6] Chesher, A.D. (1984). Testing for neglected heterogeneity. *Econometrica* **52**, 865-872.
- [7] Hájek, J., and Šidák, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.
- [8] Hájek, J., Šidák, Z., and Sen, P.K. (1999). *Theory of Rank Tests*. 2nd ed. Academic Press, New York.
- [9] Hallin, M., Ingenbleek, J.-F., and Puri, M.L. (1985). Linear serial rank tests for randomness against ARMA alternatives. *The Annals of Statistics* **13**, 1156-1181.
- [10] Hallin, M., and Werker, B. (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9**, 137-165.
- [11] Hamerle, A. (1990). On a simple test for neglected heterogeneity in panel studies. *Biometrics* **46**, 193-199.
- [12] Honda, Y. (1985). Testing the error components model with non-Normal disturbances. *Review of Economic Studies* **52**, 681-690.
- [13] Jacqmin-Gadda, H., and Commenges, D. (1995). Tests of homogeneity for generalized linear models. *Journal of the American Statistical Association* **90**, 1237-1246.
- [14] Jurečková, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *The Annals of Mathematical Statistics* **40**, 1889-1900.
- [15] King, M.L., and Evans, M.A. (1986). Testing for block effects in regression models based on survey data. *Journal of the American Statistical Association* **81**, 677-679.
- [16] Kreiss, J.P. (1987). On adaptive estimation in stationary ARMA processes. *Annals of Statistics* **15**, 112-133.
- [17] Le Cam, L.M. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer-Verlag, New York.
- [18] Lin, X. (1997). Variance component testing in generalised linear models with random effects. *Biometrika* **84**, 309-326.

- [19] Moulton, B.R. (1986). Random group effects and the precision of regression estimates. *Journal fo Econometrics* **32**, 385-397.
- [20] Moulton, B.R. and Randolph, W.C. (1989). Alternative tests of the error components model. *Econometrica* **57**, 685-693.
- [21] Nerlove, M. (1971). Further evidence on the estimation of dynamic economic relations from a time-series of cross-sections. *Econometrica* **39**, 359-382.
- [22] Orme, C. (1993). A comment on "A simple test for neglected heterogeneity in panel studies". *Biometrics* **49**, 665-667.
- [23] Puri, M.L., and Sen, P.K. (1985), *Nonparametric Methods in general Linear Models*. J. Wiley, New York.
- [24] Scott, A.J., and Holt, D. (1982). The effects of Two-stage sampling on ordinary least squares methods. *Journal of the American Statistical Association* **77**, 848-854.
- [25] Swensen, A.R. (1985), The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *Journal of Multivariate Analysis* **16**, 54-70.
- [26] Wooldridge, J.M. (2002), *Econometric Analysis of Cross-Section and Panel Data*. MIT Press, Cambridge, Mass.