Research Article

Testing for Main Random Effects in Two-Way Random and Mixed Effects Models: Modifying the F Statistic

Trent Gaugler[1] and Michael G. Akritas[2]

[1] *Department of Statistics, Carnegie Mellon University, 5000 Forbes Avenue, 132B Baker Hall, Pittsburgh, PA 15213, USA*
[2] *Department of Statistics, Penn State University, 325 Thomas Building, University Park, PA 16802, USA*

Correspondence should be addressed to Trent Gaugler; tgaugler@andrew.cmu.edu

A procedure for testing the significance of the main random effect is proposed under a model which does not require the traditional assumptions of symmetry, homoscedasticity, and normality for the error term and random effects. To accommodate this level of model generality, and also unbalanced designs, suitable adjustments to the F-test are made. The extensive simulations performed under the random effects model, and the unrestricted and restricted versions of the mixed effects model, indicate that the classical F procedure is extremely liberal under heteroscedasticity and unbalancedness. The proposed test procedure performs well in all settings and is comparable to the classical F-test when the classical assumptions are met. An analysis of a dataset from the Mussel Watch Project is presented.

1. Introduction

Consider a two-factor mixed or random effects design, and let Y_{ijk} denote the kth observation at level i of the row factor and level j of the column factor. In the case of a mixed effects design we assume that the row factor is fixed.

The classical two-way model, compare [1–5], uses the decomposition

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}. \tag{1}$$

For the random effects model, the α_i's, β_j's, γ_{ij}'s, and ϵ_{ijk}'s are mutually independent, the α_i's are iid $N(0, \sigma_\alpha^2)$, the β_j's are iid $N(0, \sigma_\beta^2)$, the γ_{ij}'s are iid $N(0, \sigma_\gamma^2)$, and the ϵ_{ijk}'s are iid $N(0, \sigma_\epsilon^2)$.

For the mixed effects model, there are two common definitions of the effects. The *unrestricted* version of the model assumes

$$\sum_i \alpha_i = 0, \quad \beta_j \text{ are iid } N\left(0, \sigma_\beta^2\right), \quad \gamma_{ij} \text{ are iid } N\left(0, \sigma_\gamma^2\right), \tag{2}$$

and the β_j, γ_{ij} are all independent. The *restricted* version of the model keeps the above assumptions but requires that

$$\sum_i \gamma_{ij} = 0, \tag{3}$$

implying that $\gamma_{ij}, \gamma_{i'j}$ are correlated. Both versions assume the ϵ_{ijk} to be iid $N(0, \sigma_\epsilon^2)$ and independent from the other random effects. There is a dichotomy of opinion in the statistical literature as to which model should be used. Cornfield and Tukey [6], Scheffé [1], Winer [7], and Khuri et al. [5], among others, advocate the restricted version. Searle [2], Hocking [8], and Searle et al. [4] advocate the unrestricted version. While the statistic for testing for no main random effect differs in the two versions, both use what Scheffé [1, page 264] calls the *symmetry assumption*. Basically, this is the assumption of independence of the random main and interaction effects. This assumption was criticized in [9, 10] as unrealistic in most practical situations. More importantly, simulations demonstrated that the classical F-test for testing the significance of the main fixed effect is very sensitive to departures from the symmetry assumption even in balanced designs with homoscedastic errors. Additional simulations showed the classical F-test for testing the significance of the

interaction effect to also be very sensitive to unbalancedness and heteroscedasticity.

This paper deals with testing the significance of the main random effect in two-factor mixed and random effects designs. Simulations suggest that the classical F-test for this hypothesis is also very sensitive to departures from the symmetry assumption, as well as to unbalancedness and heteroscedasticity. New test procedures are presented under fully nonparametric models for the two-factor mixed and random effects designs. Unbalanced mixed and random effects designs are incorporated in the modeling by considering the sample sizes n_{ij} to be random. The models and test procedures pertain to both continuous and discrete data and allow heteroscedasticity in the error and interaction terms, as well as nonnormality. The interaction term is not assumed independent from the main effect (so the symmetry assumption is not made), though, under the random effects model, the two are shown to be uncorrelated.

The asymptotic theory of the test statistics is derived under the Neyman-Scott framework, in the sense that the number of levels of both factors is large but the group sizes can remain fixed. In such cases, test statistics are constructed by taking the difference (as opposed to the ratio) of the appropriate mean squares; compare [9–12] and references therein. In accordance to this literature, the proposed test statistics for testing for no main random effects are scaled versions of MSB-MSE* and MSB-MSAB, for the mixed and random effects designs, respectively, where the precise definition of these mean squares is given in Section 3.

The novelty of the proposed test statistics is twofold. First, the mean squares are based on unweighted cell averages to accommodate unbalanced designs. Second, in order to accommodate heteroscedasticity, MSE is replaced by a different linear combination of the cell sample variances, denoted by MSE*, which, under the null hypothesis, has the same expected value as MSB (ensuring zero mean value of the test statistic under the null hypothesis). A modified version of the MSE for accommodating heteroscedasticity was first used in Akritas and Papadatos [13] in the context of a high-dimensional one-way fixed effects design.

The rest of the paper is organised as follows. Section 2 reviews the fully nonparametric random and mixed effects models. In Section 3 we derive the test statistics and present their asymptotic distributions. Proofs of the propositions in Section 3 appear in the appendix. Section 4 discusses the estimation of the limiting null distributions and also presents results of simulations comparing our testing procedures to the usual F-tests. Finally, in Section 5 we present the analysis of a dataset from NOAA's NS&T Mussel Watch Program.

2. Fully Nonparametric Models

Here we review the fully nonparametric random and mixed effects models developed in [9, 10].

2.1. The Random Effects Model. Consider a two-way random effects design. The random levels of the row factor are obtained by random sampling from the population $\mathcal{A}_S$, while the random levels of the column factor are obtained by random sampling from the population $\mathcal{A}_T$. Let S denote a row level randomly selected from $\mathcal{A}_S$, and T denote a column level randomly selected from $\mathcal{A}_T$. Assume that the selections are independent. For each factor-level combination (S, T), observations Y_{STk}, $k = 1, \ldots, n_{ST}$, are generated from a distribution function, F_{ST}. F_{ST} can be any arbitrary distribution function, which depends only on the factor levels S, T.

Decompose Y_{STk} into a random mean and error term as follows:

$$
\begin{aligned}
Y_{STk} &= E\left(Y_{STk} \mid S, T\right) + \left[Y_{STk} - E\left(Y_{STk} \mid S, T\right)\right] \\
&= \mu_{ST} + \epsilon_{STk},
\end{aligned}
\tag{4}
$$

where μ_{ST} and ϵ_{STk} are defined implicitly in (4). The definitions of μ_{ST} and ϵ_{STk} imply

$$
\begin{aligned}
E\left(\epsilon_{STk}\right) &= 0, & E\left(\epsilon_{STk} \mid S\right) &= 0, \\
E\left(\epsilon_{STk} \mid T\right) &= 0, & E\left(\epsilon_{STk} \mid S, T\right) &= 0.
\end{aligned}
\tag{5}
$$

Conditioning on S, T and using the last relation in (5) it further follows that

$$
\begin{aligned}
E\left(\mu_{ST}\epsilon_{STk}\right) &= 0, & E\left(\mu_{ST}\epsilon_{STk} \mid S\right) &= 0, \\
E\left(\mu_{ST}\epsilon_{STk} \mid T\right) &= 0.
\end{aligned}
\tag{6}
$$

Next, decompose the random means as

$$
\mu_{ST} = \mu + \alpha_S + \beta_T + \gamma_{ST},
\tag{7}
$$

where the overall mean and the random effects α_S, β_T, and γ_{ST} are defined as

$$
\begin{aligned}
\mu &= E\left(Y_{STk}\right) = E\left(\mu_{ST}\right), \\
\alpha_S &= E\left(Y_{STk} \mid S\right) - \mu = E\left(\mu_{ST} \mid S\right) - \mu, \\
\beta_T &= E\left(Y_{STk} \mid T\right) - \mu = E\left(\mu_{ST} \mid T\right) - \mu, \\
\gamma_{ST} &= \mu_{ST} - \mu - \alpha_S - \beta_T.
\end{aligned}
\tag{8}
$$

Combining (4) and (7) we obtain

$$
Y_{STk} = \mu + \alpha_S + \beta_T + \gamma_{ST} + \epsilon_{STk},
\tag{9}
$$

where the random effects have zero means, are both marginally and conditionally on S and T uncorrelated from the error terms (this follows from (6)), and satisfy

$$
E\left(\alpha_S \mid T\right) = E\left(\beta_T \mid S\right) = E\left(\gamma_{ST} \mid S\right) = E\left(\gamma_{ST} \mid T\right) = 0,
$$

$$
E\left(\alpha_S\beta_T\right) = E\left(\alpha_S\gamma_{ST}\right) = E\left(\beta_T\gamma_{ST}\right) = 0,
\tag{10}
$$

as is easily seen using the independence of S, T. Note that even though the main and interaction effects are uncorrelated, in general they will not be independent. This is because the interaction term is typically related to the two main effects by a multiplicative relation rendering the interaction effect

nonnormal even in the presence of normal main effects. Therefore the symmetry assumption is typically violated in random effects designs.

Relation (9) with the associated properties (5), (6), and (10) is derived using only the assumption that S and T are independent. Thus, it is a nonparametric version of the classical random effects model, in the sense that it allows for arbitrary (rather than normal) distributions of the random effects and error term, the variance (as well as the entire distribution) of the error term can depend on the random factor levels, and the effects are uncorrelated, and uncorrelated to the error term, as opposed to the assumed independence of the classical model.

2.2. The Mixed Effects Model. Consider a design where the rows correspond to the fixed effects and the columns correspond to the random effects. The levels of the random effects are obtained by simple random sampling from a population $\mathscr{A}$. Let T denote a randomly selected element from $\mathscr{A}$, i a fixed effect level, and Y_{iT} an observation obtained from the cell (i, T). We place no restrictions on the form of the distribution function, F_{iT}, of Y_{iT} other than the obvious restriction that it has to depend on i and T. In particular, F_{iT} can be a discrete distribution function. Since T is obtained from a random selection, F_{iT}, its mean μ_{iT} and variance σ_{iT}^2 are all random. The nonparametric model consists of a decomposition of Y_{iT} in terms of μ_{iT} plus an error term, and further decomposition of μ_{iT} to define the main effects and interaction terms. In particular, write

$$
\begin{aligned}
Y_{iT} &= E\left(Y_{iT} \mid T\right) + \left(Y_{iT} - E\left(Y_{iT} \mid T\right)\right) \\
&= \mu_{iT} + \left(Y_{iT} - \mu_{iT}\right) = \mu_{iT} + \epsilon_{iT},
\end{aligned}
\tag{11}
$$

where the error term ϵ_{iT} is defined implicitly in the above relation. An easy consequence of the definition of ϵ_{iT} is

$$
E\left(\epsilon_{iT}\right) = E\left(\epsilon_{iT} \mid T\right) = 0, \qquad E\left(\epsilon_{iT}\mu_{iT}\right) = E\left(\epsilon_{iT}\mu_{iT} \mid T\right) = 0.
\tag{12}
$$

Next decompose μ_{iT} into main and interaction effects as

$$
\mu_{iT} = \mu + \alpha_i + \beta_T + \gamma_{iT},
\tag{13}
$$

where $\mu = (1/a) \sum_i E(\mu_{iT})$ is the overall mean and

$$
\alpha_i = E\left(\mu_{iT}\right) - \mu, \qquad \beta_T = \frac{1}{a}\sum_i \mu_{iT} - \mu,
$$
$$
\gamma_{iT} = \mu_{iT} - \mu - \alpha_i - \beta_T,
\tag{14}
$$

are the main and interaction effects. An immediate implication of (14) is

$$
\sum_i \alpha_i = 0, \qquad E\left(\beta_T\right) = 0, \qquad \sum_i \gamma_{iT} = 0, \qquad E\left(\gamma_{iT}\right) = 0.
\tag{15}
$$

Using (12) it further follows that

$$
\begin{aligned}
E\left(\epsilon_{iT}\beta_T\right) &= E\left(\epsilon_{iT}\beta_T \mid T\right) = 0, \\
E\left(\epsilon_{iT}\gamma_{iT}\right) &= E\left(\epsilon_{iT}\gamma_{iT} \mid T\right) = 0.
\end{aligned}
\tag{16}
$$

Combining (11) and (13) we have

$$
Y_{iT} = \mu + \alpha_i + \beta_T + \gamma_{iT} + \epsilon_{iT}.
\tag{17}
$$

While model (17) has the appearance of the usual mixed effects model, it does not require normality (or even continuity) of the observations, the variance of the random interaction effects can depend on i, and the random effects β_T and γ_{iT} are not uncorrelated without further assumptions (which we do not make) about constant variance of the random means and constant covariance for any two distinct random means.

Suppose now that there are a fixed levels, so $i = 1, \ldots, a$, and let T_j, $j = 1, \ldots, b$, denote the levels of the random effect. Recall that the T_j are obtained by simple random sampling from $\mathscr{A}$. Also let Y_{iT_jk}, $k = 1, \ldots, n_{ij}$, denote the replications (iid observations) in factor-level combination (i, T_j). For notational simplicity we write Y_{ijk} for Y_{iT_jk} and j for T_j. Thus, conditionally on j (or T_j), the Y_{ijk} are iid F_{ij}. From (17) we have

$$
Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk},
\tag{18}
$$

where the effects and error terms satisfy (14), (15), and (16). In addition, γ_{ij_1} and γ_{ij_2} are independent for $j_1 \neq j_2$, ϵ_{ijk}, $k = 1, \ldots, n_{ij}$, are iid conditionally on j, $\epsilon_{i_1 j_1 k_1}$, and $\epsilon_{i_2 j_2 k_2}$ are independent for $j_1 \neq j_2$, and $\epsilon_{i_1 jk_1}$ and $\epsilon_{i_2 jk_2}$ are uncorrelated for $i_1 \neq i_2$. These imply

$$
\begin{aligned}
E\left(\epsilon_{i_1 jk_1}\epsilon_{i_2 jk_2}\right) &= 0, \quad \text{for } k_1 \neq k_2, \\
E\left(\epsilon_{i_1 j_1 k_1}\epsilon_{i_2 j_2 k_2}\right) &= 0, \quad \text{for } j_1 \neq j_2, \\
E\left(\epsilon_{i_1 jk_1}\epsilon_{i_2 jk_2}\right) &= 0, \quad \text{for } i_1 \neq i_2, \\
E\left(\gamma_{ij_1}\gamma_{ij_2}\right) &= 0, \quad \text{for } j_1 \neq j_2.
\end{aligned}
\tag{19}
$$

We note that relation (14) implies that γ_{ij} is a function of the random level T_j. Since the random levels are iid, we have that the variances of γ_{ij}, $j = 1, \ldots, b$, are equal. Arguing similarly, we have that the variances of the ϵ_{ijk}, $j = 1, \ldots, b$, are equal. However the model allows heteroscedasticity across different levels of the fixed effect, that is, $\sigma^2(\gamma_{i_1 j_1}) \neq \sigma^2(\gamma_{i_2 j_2})$ and $\sigma^2(\epsilon_{i_1 j_1 k_1}) \neq \sigma^2(\epsilon_{i_2 j_2 k_2})$, for $i_1 \neq i_2$. Moreover, the conditional variances of the error term given different levels of the random effect can be different. Summarizing, we have

$$
\sigma^2\left(\gamma_{ij_1}\right) = \sigma^2\left(\gamma_{ij_2}\right), \qquad \sigma^2\left(\epsilon_{ij_1 k_1}\right) = \sigma^2\left(\epsilon_{ij_2 k_2}\right),
\tag{20}
$$

but the model allows

$$
\begin{aligned}
\sigma^2\left(\gamma_{i_1 j_1}\right) &\neq \sigma^2\left(\gamma_{i_2 j_2}\right), \quad \text{for } i_1 \neq i_2, \\
\sigma^2\left(\epsilon_{i_1 j_1 k_1}\right) &\neq \sigma^2\left(\epsilon_{i_2 j_2 k_2}\right), \quad \text{for } i_1 \neq i_2, \\
\operatorname{Var}\left(\epsilon_{ij_1 k_1} \mid T_{j_1}\right) &\neq \operatorname{Var}\left(\epsilon_{ij_2 k_2} \mid T_{j_2}\right).
\end{aligned}
\tag{21}
$$

3. The Test Statistics and Their Asymptotic Distributions

Before proceeding, we first define our notation:

$$\overline{Y}_{ij\cdot} = \frac{1}{n_{ij}}\sum_{k=1}^{n_{ij}} Y_{ijk}, \qquad \widetilde{Y}_{i\cdot\cdot} = \frac{1}{b}\sum_{j=1}^{b}\overline{Y}_{ij\cdot}, \qquad \widetilde{Y}_{\cdot j\cdot} = \frac{1}{a}\sum_{i=1}^{a}\overline{Y}_{ij\cdot},$$

$$\widetilde{Y}_{\cdot\cdot\cdot} = \frac{1}{a}\sum_{i=1}^{a}\widetilde{Y}_{i\cdot\cdot} = \frac{1}{b}\sum_{j=1}^{b}\widetilde{Y}_{\cdot j\cdot}, \qquad \overline{\beta}_{\cdot} = \frac{1}{b}\sum_{j=1}^{b}\beta_j,$$

$$\overline{\gamma}_{i\cdot} = \frac{1}{b}\sum_{j=1}^{b}\gamma_{ij}, \qquad \overline{\epsilon}_{ij\cdot} = \frac{1}{n_{ij}}\sum_{k=1}^{n_{ij}}\epsilon_{ijk}, \qquad \widetilde{\epsilon}_{i\cdot\cdot} = \frac{1}{b}\sum_{j=1}^{b}\overline{\epsilon}_{ij\cdot},$$

$$\widetilde{\epsilon}_{\cdot j\cdot} = \frac{1}{a}\sum_{i=1}^{a}\overline{\epsilon}_{ij\cdot}, \qquad \widetilde{\epsilon}_{\cdot\cdot\cdot} = \frac{1}{b}\sum_{j=1}^{b}\widetilde{\epsilon}_{\cdot j\cdot} = \frac{1}{a}\sum_{i=1}^{a}\widetilde{\epsilon}_{i\cdot\cdot}.$$

$$(22)$$

The usual F statistic in the balanced case in the restricted version of the mixed model is given by

$$F = \frac{\text{MSB}}{\text{MSE}} = \frac{a^2 bn(n-1)\sum_j\left(\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdots}\right)^2}{(b-1)\sum_i\sum_j\sum_k\left(Y_{ijk} - \overline{Y}_{ij\cdot}\right)^2}, \quad (23)$$

while the usual F statistic in the balanced case in the random effects model is given by

$$F = \frac{\text{MSB}}{\text{MSAB}} = \frac{a(a-1)\sum_j\left(\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdots}\right)^2}{\sum_i\sum_j\left(\overline{Y}_{ij\cdot} - \overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot j\cdot} + \overline{Y}_{\cdots}\right)^2}. \quad (24)$$

In the unbalanced case there is no theoretically justified procedure, but Searle [14] does have a number of recommendations which are implemented in software packages.

The test statistics we propose for testing the hypothesis of no main random effect, $H_0(B) : \sigma^2(\beta_j) = 0$, $j = 1, \ldots, b$, are modeled after the usual F statistics given in (23) and (24), except that we take the difference rather that the ratio, as explained in the Introduction. Moreover, in order to accommodate unbalanced designs, $\overline{Y}_{i\cdot\cdot}$, $\overline{Y}_{\cdot j\cdot}$, and $\overline{Y}_{\cdots}$ are replaced by $\widetilde{Y}_{i\cdot\cdot}$, $\widetilde{Y}_{\cdot j\cdot}$, and $\widetilde{Y}_{\cdots}$, respectively. As mentioned in the Introduction, our asymptotic theory requires that the difference has mean value zero. While this is the case for the random effects design, the expected value of MSB-MSE is not zero for the mixed effects design with heteroscedastic errors. Therefore we replace MSE with a different linear combination of the cell sample variances:

$$\text{MSE}^* = \frac{1}{a^3 b}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{\cdot j}}{n_{ij}}\frac{1}{n_{ij}-1}\sum_{k=1}^{n_{ij}}\left(Y_{ijk} - \overline{Y}_{ij\cdot}\right)^2. \quad (25)$$

Thus, the proposed test statistics for the mixed effects and random effects models are, respectively,

$$S_{B,\text{mixed}} = \frac{1}{a(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}\left(\widetilde{Y}_{\cdot j\cdot} - \widetilde{Y}_{\cdots}\right)^2 - \text{MSE}^*, \quad (26)$$

$$S_{B,\text{rand}} = \frac{1}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}\left[(a-1)\left(\widetilde{Y}_{\cdot j\cdot} - \widetilde{Y}_{\cdots}\right)^2 - R_{ij}^2\right], \quad (27)$$

where $R_{ij} = \overline{Y}_{ij\cdot} - \widetilde{Y}_{\cdot j\cdot} - \widetilde{Y}_{i\cdot\cdot} + \widetilde{Y}_{\cdots}$.

Proposition 1. Let $S_{B,mixed}$ be as defined in (26) and let $S_{B,rand}$ be as defined in (27), with the observations Y_{ijk} generated according to the model of Section 2. Then,

(i) if the null hypothesis $H_0(B) : \sigma^2(\beta_j) = 0$, $j = 1, \ldots, b$ holds,

$$E\left(S_{B,mixed}\right) = E\left(S_{B,rand}\right) = 0, \quad (28)$$

(ii) if the null hypothesis $H_0(B) : \sigma^2(\beta_j) = 0$, $j = 1, \ldots, b$ does not hold,

$$E\left(S_{B,mixed}\right) > 0, \qquad E\left(S_{B,rand}\right) > 0. \quad (29)$$

The second item in Proposition 1 suggests that the null hypothesis should be rejected at level α whenever the test statistic takes a value larger than the $100(1-\alpha)$th percentile of the null distribution. The next proposition establishes asymptotic representations of $S_{B,\text{mixed}}$ and $S_{B,\text{rand}}$ that are useful for establishing their asymptotic distributions.

Proposition 2. Under $H_0(B)$,

(i)

$$\sqrt{b}S_{B,mixed} = \frac{1}{a\sqrt{b}}\sum_j n_{\cdot j}\left[\left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)^2 - E\left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)^2\right]$$

$$+ o_p(1), \quad (30)$$

(ii)

$$aS_{B,rand} = \frac{1}{ab}\sum_i\sum_{i_1 \neq i}\sum_j \gamma_{ij}\gamma_{i_1 j} + o_p(1). \quad (31)$$

Theorem 3. Let $s_b^2 = \text{Var}\left((1/\sqrt{b})\sum_j n_{ij}[(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots})^2 - E(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots})^2]\right)$. Then under $H_0(B)$

$$\sqrt{b}\frac{S_{B,mixed}}{s_b} \longrightarrow N(0,1), \quad (32)$$

in distribution as $a, b \to \infty$.

Proof. Letting $T_b = (1/\sqrt{b})\sum_j n_{ij}[(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots})^2 - E(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots})^2]$, then, as stated above,

$$\sqrt{b}\frac{S_{B,\text{mixed}}}{s_b} \approx \frac{T_b}{s_b}. \quad (33)$$

The proof that $T_b/s_b \to N(0,1)$ in distribution follows from an application of the Lindeberg-Feller CLT. We first rewrite our statement in the notation of the Lindeberg-Feller CLT and then show that the Lindeberg condition holds after a straightforward application of the Dominated Convergence Theorem. Now let $Y_j^{a,b} = (1/\sqrt{b})n_{ij}[(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots})^2 - E(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots})^2]$, so that $T_b = \sum_j Y_j^{a,b}$. Also note that we can write

$$s_b^2 = \mathrm{Var}\,(T_b) = \frac{1}{b}\sum_j n_{ij}^2\,\mathrm{Var}\left(\left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)^2\right), \quad (34)$$

which is finite because we have assumed a finite fourth moment for the ϵ terms. Moreover, this implies that

$$\mathrm{Var}\left(\sqrt{b}Y_j^{a,b}\right) = n_{ij}^2\,\mathrm{Var}\left(\left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)^2\right), \quad (35)$$

which is again finite. Now in order to use the Lindeberg-Feller CLT, we need to verify the Lindeberg condition:

$$\forall \epsilon > 0, \quad \frac{1}{s_b^2}\sum_{j=1}^b E\left[\left(Y_j^{a,b}\right)^2 I\left\{\left|Y_j^{a,b}\right| > \epsilon s_b\right\}\right] \longrightarrow 0. \quad (36)$$

Note that

$$\frac{1}{s_b^2}\sum_{j=1}^b E\left[\left(Y_j^{a,b}\right)^2 I\left\{\left|Y_j^{a,b}\right| > \epsilon s_b\right\}\right]$$

$$= \frac{1}{bs_b^2}\sum_{j=1}^b E\left[\left(\sqrt{b}Y_j^{a,b}\right)^2 I\left\{\left|\sqrt{b}Y_j^{a,b}\right| > \sqrt{b}\epsilon s_b\right\}\right] \quad (37)$$

$$= \frac{1}{s_b^2} E\left[\left(\sqrt{b}Y_1^{a,b}\right)^2 I\left\{\left|\sqrt{b}Y_1^{a,b}\right| > \sqrt{b}\epsilon s_b\right\}\right].$$

Because $E(\sqrt{b}Y_1^{a,b}) = 0$ and $\mathrm{Var}(\sqrt{b}Y_1^{a,b})$ is finite, the Dominated Convergence Theorem implies that $E[(\sqrt{b}Y_1^{a,b})^2 I\{|\sqrt{b}Y_1^{a,b}| > \sqrt{b}\epsilon s_b\}] \to 0$, and hence the Lindeberg condition is verified. This completes the proof of the theorem. □

Theorem 4. *Define the operator $H_b : g \to H_b g$, for $g \in L_2([0,1], \mathcal{B}_{[0,1]}, \lambda_{[0,1]})$, where $\mathcal{B}_{[0,1]}$ is the class of Borel subsets of $[0,1]$ and $\lambda_{[0,1]}$ is the Lebesgue measure on $\mathcal{B}_{[0,1]}$, by*

$$(H_b g)(s) = \int h_b(s,t)\,g(t)\,dt, \quad (38)$$

where $h_b(s,t) = (1/b)\sum_{j=1}^b \gamma_{sj}\gamma_{tj}$. Let $\lambda_\nu^b, f_\nu^b, \nu = 1,2,\ldots,$ be the eigenvalues and eigenfunctions of the integral equation

$$(H_b f)(s) = \lambda f(s), \quad (39)$$

and assume that the eigenvalues satisfy

$$\sum_{\nu=1}^\infty \left(\sup_b \lambda_\nu^b\right)^2 < \infty. \quad (40)$$

Then, if γ_{ij} and ϵ_{ij} have finite fourth moments and the covariances $\mathrm{Cov}\,(\gamma_{ij}, \gamma_{ij'})$, $\mathrm{Cov}\,(\gamma_{ij}^2, \overline{\epsilon}_{ij\cdot})$, $\mathrm{Cov}\,(\gamma_{ij}^2, \gamma_{ij'}^2)$, and

$\mathrm{Cov}\,(\epsilon_{ij}^2, \epsilon_{ij'}^2)$ *are bounded uniformly in j, j', the asymptotic null distribution of $S_{B,\,rand}$, as $a \to \infty$, $b \to \infty$, is the same as the asymptotic distribution of $\sum_{\nu=1}^\infty \lambda_\nu^b(Z_\nu^2 - 1)$, as $b \to \infty$, where $Z_\nu, \nu = 1,2,\ldots,$ are independent standard normal random variables.*

Proof. The proof proceeds by first conditioning on the levels of the column factor $\mathbf{T} = (T_1, \ldots, T_b)$ and then applying the arguments used in the proof of Theorem 3.3 of [9]. Indeed, the proof of the aforementioned theorem uses an expansion similar to (31) of Proposition 2 (with the indices i, j interchanged) which is modeled as a U-statistic with kernel depending on a. The same arguments establish the limiting distribution of the test statistic conditionally on $\mathbf{T} = (T_1, \ldots, T_b)$. As $b \to \infty$, the strong law of large numbers guarantees that the kernel $h_b(s,t)$ converges to the same limit, and thus the conditional limiting distribution does not depend on $\mathbf{T} = (T_1, \ldots, T_b)$. Thus, the test statistic converges to the same distribution unconditionally. □

4. Test Procedures and Simulation Results

In testing for the random main effect in the mixed effects model, we only need to estimate the variance term s_b^2 in Theorem 3. According to (30) of Proposition 2 we need to estimate the variance of

$$\frac{1}{\sqrt{b}}\sum_j n_{ij}\left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)^2. \quad (41)$$

It is easy to see that

$$\frac{1}{\sqrt{b}}\sum_j n_{ij}\left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)^2 - \frac{1}{\sqrt{b}}\sum_j n_{ij}\widetilde{\epsilon}_{\cdot j\cdot}^2 \longrightarrow 0, \quad (42)$$

in probability. Thus, we need to estimate the variance of the independent terms $\widetilde{\epsilon}_{\cdot j\cdot}^2$. By the Central Limit Theorem, as $a \to \infty$, $\widetilde{\epsilon}_{\cdot j\cdot}$ is approximately Normally distributed with mean zero and variance $(1/a^2)\sum_i(\sigma_{ij}^2/n_{ij})$. Therefore, the approximate variance of $\widetilde{\epsilon}_{\cdot j\cdot}^2$ is $2((1/a^2)\sum_i(\sigma_{ij}^2/n_{ij}))^2$. As a result, we estimate s_b^2 by

$$\widehat{s}_b^2 = \frac{2}{a^6 b}\sum_j\left(n_{\cdot j}\sum_i\frac{s_{ij}^2}{n_{ij}}\right)^2. \quad (43)$$

Then, according to Proposition 1 and Theorem 3, $H_0(B)$ is rejected at level α when

$$b^{1/2}\frac{S_{B,\,\mathrm{mixed}}}{\widehat{s}_b} > z_\alpha, \quad (44)$$

where z_α is the $(1-\alpha)100$th percentile of the standard normal distribution.

Testing for main effects in the random effects model involves estimation of the asymptotic distribution of Theorem 4. Koltchinskii and Giné [15] show that the infinite spectrum of a Hilbert-Schmidt operator H can

be approximated by the finite spectrum of the empirical matrix version of the operator. We will use this result to approximate the spectrum of the operator H_b in Theorem 4 by its empirical version, $\widetilde{H}_a$, defined below. To this end, the following representation of $S_{B,\mathrm{rand}}$ as a U-statistic is useful:

$$aS_{B,\mathrm{rand}} = \frac{1}{ab}\sum_{i_1}\sum_{i_2 \neq i_1}\sum_j \left(\overline{Y}_{i_1 j \cdot} - \overline{Y}_{i_1 \cdot \cdot}\right)\left(\overline{Y}_{i_2 j \cdot} - \overline{Y}_{i_2 \cdot \cdot}\right) = \frac{a-1}{b}U_a, \tag{45}$$

where the U-statistic U_a is defined by the kernel

$$h\left(S_{i_1}, S_{i_2}\right) = \sum_j \left(\overline{Y}_{i_1 j \cdot} - \overline{Y}_{i_1 \cdot \cdot}\right)\left(\overline{Y}_{i_2 j \cdot} - \overline{Y}_{i_2 \cdot \cdot}\right). \tag{46}$$

Under the null hypothesis $H_0(B)$, the empirical version, $\widetilde{H}_a$, of the operator H_b is the $a \times a$ matrix with entries $\widetilde{H}_{i,i'} = (1/a)h(S_i, S_{i'})$. The a eigenvalues $\lambda_1^a, \ldots, \lambda_a^a$ of $\widetilde{H}_a$ are then used to estimate the asymptotic distribution of $S_{B,\mathrm{rand}}$ as $\sum_{i=1}^a \lambda_i^a(Z_i^2 - 1)$, where the Z_i^2 are randomly sampled from the χ_1^2 distribution. To approximate the estimated distribution, we randomly generate a large number B of sets $Z_1^2, \ldots, Z_a^2$, and use the formula above to obtain B realizations from the distribution. Using this approximate distribution, for a level α test we reject $H_0(B)$ when $S_{B,\mathrm{rand}}$ is greater than the $(1 - \alpha)100$th percentile.

We now present simulation results to compare our test procedures for the main random effect in the mixed and random effects models to the usual F-tests given by (23) and (24). Note that these statistics are only valid for balanced designs. In the unbalanced designs, the datasets used to obtain the achieved size and power are written out to files and input to SAS for the F-test to be carried out in SAS PROC MIXED. The achieved size and power are calculated from the P values given from SAS.

Using the decomposition in relation (18), we generate our data for the mixed effects model as follows.

(1) Set μ equal to any constant.

(2) Generate a vector of main random effects T_j, $j = 1, \ldots, b$ from the standard normal distribution.

(3) Independently generate a vector of fixed "functions" λ_i, $i = 1, \ldots, a$, from the standard exponential distribution and set $\gamma_i^* = \lambda_i - \overline{\lambda}$.

(4) Generate Y_{ijk} from a normal distribution with variance σ_i^2 and mean $\mu + \alpha_i + T_j + T_j * \gamma_i^*$, where the last two terms are the main random and interaction effects, respectively.

Similarly, using the decomposition in relation (9), we generate our data for the random effects model as follows.

(1) Set μ equal to any constant.

(2) Generate a vector of main row random effects $\widetilde{S}_i$, $i = 1, \ldots, a$ from the standard normal distribution.

(3) Independently generate a vector of main column random effects T_j, $j = 1, \ldots, b$ from the standard normal distribution, and set $T_j^* = \tau T_j$, where τ is used to set the degree of deviation from the null hypothesis.

(4) Set the interaction effects $\gamma_{ij} = S_i T_j$.

(5) Generate Y_{ijk} from a normal distribution with variance σ_i^2 and mean $\mu + S_i + T_j^* + S_i T_j$.

In Table 1, each simulation used $a = b = 20$. For the unbalanced setups, rows $i = 1, \ldots, 15$ had 4 observations per cell, whereas rows $i = 16, \ldots, 20$ had 2 observations per cell. For the heteroscedastic setups, the first 15 rows all had variance 1, while the remaining rows had variance 5; that is, $\sigma_1^2 = \cdots = \sigma_{15}^2 = 1$, $\sigma_{16}^2 = \cdots = \sigma_{20}^2 = 5$.

5. Data Analysis

One real-world application for our methodology can be found through the National Oceanic and Atmospheric Administration's National Status and Trends Program. In 1986, this division undertook a very large scale project to monitor the levels of numerous chemical contaminants and organic chemical constituents in marine sediment and bivalve (mollusk) tissue samples. This project (http://ccma.nos.noaa.gov/about/coast/nsandt/musselwatch.aspx), dubbed the Mussel Watch Project, is still on-going and there are no apparent plans to discontinue it in the near future. There are currently over 300 coastal sites at which sediment and bivalve samples are collected and analysed for the project. Each site is categorised as being within a certain estuarine drainage area (EDA). For our data analysis, we chose to analyse the Mercury concentrations in tissue samples of the *Crassostrea virginica*, or Eastern Oyster. There was a high degree of missingness of Mercury concentrations in the dataset, so we used imputations to generate these measurements. Because of how we performed our imputations, only EDAs that did not have missing measurements for two consecutive years and EDAs that had measurements in 1986 were included in the analysis. Once we had a complete imputed dataset, we employed our methodology to test for effects due to EDA.

5.1. Imputations. Because of the $1/(n_{ij} - 1)$ coefficient that is used in the statistic $S_{B,\mathrm{mixed}}$, each cell needs to have at least 2 observations. To facilitate the imputations, we identified 195 instances of 2 consecutive measurements falling into the following groupings: 1986-1987, 1988-1989, 1990-1991, 1992-1993, 1994-1995, 1996-1997, 1998-1999, 2000-2001, 2002-2003, and 2004-2005. If any cell had more than one observation, they were averaged. We used these 195 pairs to build the regression model $\widehat{CV}_{t,j} = 0.046424 + 0.65852CV_{t-1,j}$, where we predict a value of Mercury concentration at time t from the Mercury concentration at time $t - 1$. For this model, the P value for the test of $H_0 : \beta_1 = 0$ versus $H_A : \beta_1 \neq 0$ was 0.000 and the $R^2 = 36.8\%$. When fitting the regression model in Minitab, we also saved the residuals from the model and the actual imputations that we used were $\widehat{CV}_{t,j} + R^*$, where R^* is a randomly selected residual from the fitting of the model. We separately treat imputations for cells with only

TABLE 1: Comparison of achieved type I error rates and power for $S_{B,\text{mixed}}$, $S_{B,\text{rand}}$, and F with $N = 1000$ tests for the random and mixed effects models.

	Random effects model								Mixed effects model							
	Balanced				Unbalanced				Balanced				Unbalanced			
	$\alpha = .05$		$\alpha = .10$		$\alpha = .05$		$\alpha = .10$		$\alpha = .05$		$\alpha = .10$		$\alpha = .05$		$\alpha = .10$	
	S_B	F	S_B	F	S_B	F	S_B	F	S_B	F	S_B	F	S_B	F	S_B	F
Homo.																
H_0	.054	.220	.111	.247	.049	.167	.092	.198	.065	.047	.112	.099	.063	.045	.109	.086
A_1	.053	.229	.123	.265	.069	NA	.140	NA	.227	.187	.318	.296	.184	.138	.260	.228
A_2	.125	.337	.221	.367	.113	NA	.201	NA	.761	.719	.820	.804	.542	.489	.634	.638
A_3	.236	.532	.369	.563	.234	NA	.360	NA	.975	.963	.982	.982	.888	.863	.920	.917
Hetero.																
H_0	.045	.150	.089	.182	.015	.230	.042	.303	.070	.056	.106	.100	.062	.935	.107	.952
A_1	.045	.160	.088	.199	.021	NA	.053	NA	.077	.060	.129	.121	.074	NA	.124	NA
A_2	.114	.273	.190	.316	.047	NA	.094	NA	.151	.129	.221	.215	.104	NA	.152	NA
A_3	.195	.430	.318	.467	.117	NA	.215	NA	.270	.235	.353	.341	.131	NA	.189	NA

one observation and imputations for empty cells. When cells are empty, we use the regression equation

$$\widehat{CV}_{t,j} = 0.046424 + 0.65852 CV_{t-1,j} + R^*, \qquad (47)$$

twice to impute two values of Hg concentration at time t. When a cell has only one observation, we use a different technique.

(1) If the EDA has any cells with more than one observation in any year, we calculate residuals for each year with more than one observation as $e_{ijk} = Y_{ijk} - \overline{Y}_{ij\cdot}$, $k = 1,\dots,n_{ij}$. Then for the EDA and year of interest, we impute a second observation as $Y_{ij2} = Y_{ij1} + e_{ijk}^*$, where e_{ijk}^* denotes a randomly selected residual.

(2) If the EDA of interest has at most one observation in each cell, we identify the closest neighboring EDA (denoted by level j^*) and calculate residuals for each year with more than one observation in the neighboring EDA as $e_{ij^*k} = Y_{ij^*k} - \overline{Y}_{ij^*\cdot}$, $k = 1,\dots,n_{ij^*}$. Then for the EDA and year of interest, we impute a second observation as $Y_{ij2} = Y_{ij1} + e_{ij^*k}^*$, where $e_{ij^*k}^*$ denotes a randomly selected residual.

After performing the imputations, there are a total of $N = 1808$ observations of Hg concentration in CV tissue samples. The 20 years included are 1986–2005, and there are 39 EDAs included in the dataset. The data are rather unbalanced, with most EDAs having only two observations per year, while others such as the Galveston Bay EDA have as many as six observations per year. Moreover, the observations do not seem to be homoscedastic. These points are illustrated in Figure 1, where we show scatterplots of Hg concentrations for six EDAs after imputation.

5.2. The Analysis. In our analysis of this dataset under the mixed effects model, we take the years as the fixed effect and the EDAs as the random effect. We can think of the EDAs as

a random effect because we are only analyzing a very small subset of a much larger group of EDAs and the location effect is not of specific interest. Some may view the time effect as being of specific interest, while others may not. Whether one views the time effect as a fixed or random factor would precipitate analysis using a mixed or random effects model.

We want to model the Hg concentration with an interaction effect between year and location, so we use the model in relation (18). Because we are dealing with compositional data, we employ the common practice of analyzing the data after taking a log transformation; compare [16].

When we apply our proposed test procedure based on the test statistic S_C, we obtain a P value of 0.0588. When we apply the F-test computed from SAS PROC MIXED, we obtain a P value of 0.0016.

The P value for the test using S_C is simply calculated as $1 - \Phi(\cdot)$, where Φ is the distribution function of the standard normal random variable. The P value for the SAS F-test is calculated as $1 - F_{\nu_1,\nu_2}(\cdot)$, where ν_1 and ν_2 denote the numerator and denominator degrees of freedom. In the test for the random interaction effect, SAS uses $\nu_1 = 722$, $\nu_2 = 1028$. The simulation results shown in Table 1 suggest that the significance of the interaction effect indicated by the F-test cannot be trusted. Indeed, according to these simulation results, we expect the F-test to detect interaction in almost any unbalanced and heteroscedastic dataset. On the other hand, the simulation results suggest that the marginal significance indicated by the P value obtained from the proposed test procedure can be trusted. According to this we recommend that the interaction effect be included in the modeling of this dataset.

6. Summary

The classical F statistic for testing the significance of main random effects in two-factor mixed and random effects model is shown to be very sensitive to violations of the assumptions under which it is derived, in particular those of symmetry, homoscedasticity, and balancedness. Two new
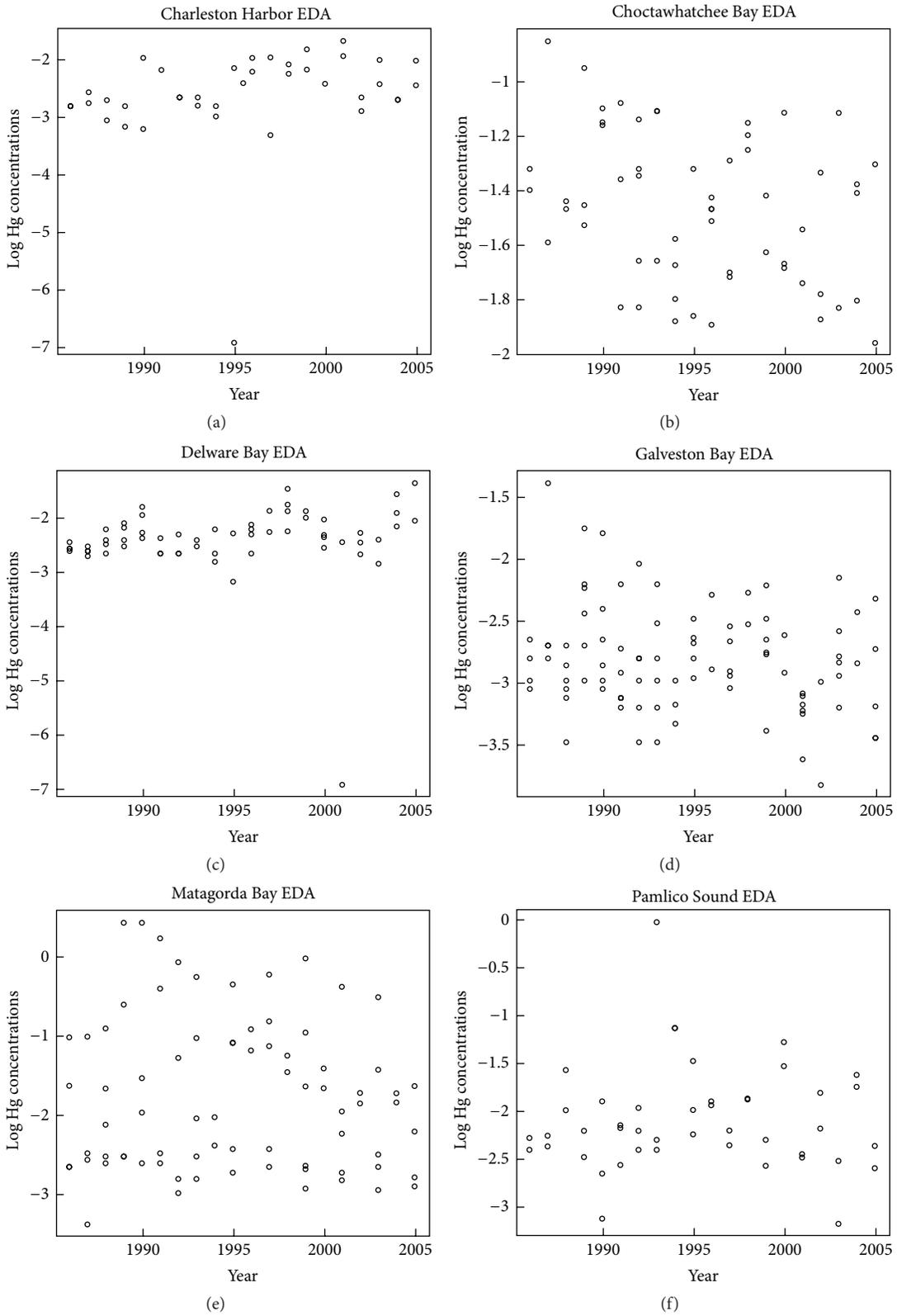
Figure 1: Scatterplot of Mercury concentrations in Eastern Oysters by year after imputation for six EDAs.

test procedures for testing the hypothesis of no main random effects are proposed under fully nonparametric modeling of the mixed and random effects designs. The test statistics are defined as differences (as opposed to ratios) of suitably defined mean squares, and their asymptotic theory is derived as the number of levels tends to infinity. Simulations suggest that the new procedures perform reasonably well in situations where the F statistic breaks down and have comparable performance to it when the assumptions it requires are met. The practical value of the proposed procedures is demonstrated with the analysis of a real dataset.

Appendices

A. Proof of Proposition 1

We first present the proof for the statistic under the mixed effects model, and then for the statistic under the random effects model.

Proof for $S_{B,\mathrm{mixed}}$. Using relations (15) and (18) we write

$$S_{B,\mathrm{mixed}} = \frac{1}{a(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}\left[\left(\beta_j - \overline{\beta}_\cdot\right) + \left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)\right]^2 - \mathrm{MSE}^*.$$
(A.1)

The conditional expectation of the first part given the vector $\mathbf{T}$ of random column levels is

$$
\begin{aligned}
&E\left[\frac{1}{a(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\left[\left(\beta_j - \overline{\beta}_\cdot\right) + \left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)\right]^2 \mid \mathbf{T}\right] \\
&= \frac{1}{a(b-1)}\sum_{i}\sum_{j} n_{ij}\left[\left(\beta_j - \overline{\beta}_\cdot\right)^2 + \left(1 - \frac{1}{b}\right)\mathrm{Var}\left(\widetilde{\epsilon}_{\cdot j\cdot} \mid T_j\right)\right] \\
&= \frac{1}{a(b-1)} \\
&\quad \times \sum_{i}\sum_{j} n_{ij}\left[\left(\beta_j - \overline{\beta}_\cdot\right)^2 + \left(1 - \frac{1}{b}\right)\frac{1}{a^2}\sum_{i_1}\frac{\mathrm{Var}\left(\epsilon_{i_1 jk} \mid T_j\right)}{n_{i_1 j}}\right].
\end{aligned}
$$
(A.2)

Therefore, under the null hypothesis that $\beta_j = 0$, a.s., for all j, the expected value of $S_{B,\mathrm{mixed}}$ is zero conditionally on $\mathbf{T}$, and thus also unconditionally. Moreover, it is clear that under the alternative the expected value of $S_{B,\mathrm{mixed}}$ is positive.

Proof for $S_{B,\mathrm{rand}}$. Using relation (9), we write

$$S_{B,\mathrm{rand}} = \frac{1}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}(a-1)$$

$$\times\left(\left(\beta_j - \overline{\beta}_\cdot\right) + \left(\overline{\gamma}_{\cdot j} - \widetilde{\gamma}_{\cdot\cdot}\right) + \left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)\right)^2$$

$$-\frac{1}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\left(\gamma_{ij} - \overline{\gamma}_{\cdot j} - \overline{\gamma}_{i\cdot} + \widetilde{\gamma}_{\cdot\cdot}\right)\right.$$

$$\left. + \left(\overline{\epsilon}_{ij\cdot} - \widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{i\cdot\cdot} + \widetilde{\epsilon}_{\cdots}\right)\right)^2.$$
(A.3)

Thus, the proof of this part of the proposition will follow if we show

$$E\left[\frac{1}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}(a-1)\left(\overline{\gamma}_{\cdot j} - \widetilde{\gamma}_{\cdot\cdot}\right)^2\right]$$
$$= E\left[\frac{1}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\gamma_{ij} - \overline{\gamma}_{\cdot j} - \overline{\gamma}_{i\cdot} + \widetilde{\gamma}_{\cdot\cdot}\right)^2\right],$$
(A.4)

$$E\left[\frac{1}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}(a-1)\left(\widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{\cdots}\right)^2\right]$$
$$= E\left[\frac{1}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{\epsilon}_{ij\cdot} - \widetilde{\epsilon}_{\cdot j\cdot} - \widetilde{\epsilon}_{i\cdot\cdot} + \widetilde{\epsilon}_{\cdots}\right)^2\right].$$
(A.5)

Relation (A.4) follows from

$$
\begin{aligned}
E\left[(a-1)\left(\overline{\gamma}_{\cdot j} - \widetilde{\gamma}_{\cdot\cdot}\right)^2\right] &= E\left\{E\left[(a-1)\left(\overline{\gamma}_{\cdot j} - \widetilde{\gamma}_{\cdot\cdot}\right)^2 \mid \mathbf{T}\right]\right\} \\
&= \frac{(a-1)(b-1)}{ab}E\left[\mathrm{Var}\left(\gamma_{ij} \mid T_j\right)\right],
\end{aligned}
$$
$$
\begin{aligned}
E\left[\left(\gamma_{ij} - \overline{\gamma}_{\cdot j} - \overline{\gamma}_{i\cdot} + \widetilde{\gamma}_{\cdot\cdot}\right)^2\right] \\
= E\left[\left(\gamma_{ij} - \overline{\gamma}_{i\cdot}\right)^2 - 2\left(\gamma_{ij} - \overline{\gamma}_{i\cdot}\right)\left(\overline{\gamma}_{\cdot j} - \widetilde{\gamma}_{\cdot\cdot}\right) + \left(\overline{\gamma}_{\cdot j} - \widetilde{\gamma}_{\cdot\cdot}\right)^2\right] \\
= E\left[\left(1 - \frac{2}{a}\right)\left(\gamma_{ij} - \overline{\gamma}_{i\cdot}\right)^2 + \frac{1}{a^2}\sum_{i_1}\left(\gamma_{i_1 j} - \overline{\gamma}_{i_1\cdot}\right)^2\right] \\
= \frac{a-1}{a}E\left[\left(\gamma_{ij} - \overline{\gamma}_{i\cdot}\right)^2\right] = \frac{(a-1)(b-1)}{ab}E\left[\mathrm{Var}\left(\gamma_{ij} \mid T_j\right)\right],
\end{aligned}
$$
(A.6)

where the second equality follows from the fact that $E(\gamma_{i_1 j} - \overline{\gamma}_{i_1\cdot})(\gamma_{i_2 j} - \overline{\gamma}_{i_2\cdot}) = 0$ for $i_1 \neq i_2$. Relation (A.5) follows similarly.

B. Proof of Proposition 2

We first show the proof for the statistic under the mixed effects model, and then for the statistic under the random effects model.

Proof for $S_{B,\mathrm{mixed}}$. Because we have shown above in the proof of Proposition 1 that, under $H_0(B)$, the statistic has mean value zero, it suffices to show that with data generated

as in Section 2 and effects defined in (18), the following convergence holds:

$$
\frac{1}{a^2\sqrt{b}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}\left[\frac{1}{n_{ij}-1}\left(\epsilon_{ijk}-\bar{\epsilon}_{ij\cdot}\right)^2\right.
$$
$$
\left.-E\left(\frac{1}{n_{ij}-1}\left(\epsilon_{ijk}-\bar{\epsilon}_{ij\cdot}\right)^2\right)\right]\longrightarrow 0,
$$

(B.1)

in probability as $a,b \rightarrow \infty$. Now, first note that the above expression is centered, so we only need to show that its variance tends to zero as a,b tend to ∞. This is easily done using the facts, given in relation (19), that the ϵ terms are uncorrelated for different levels of the fixed effect, independent for different levels of the random effect, and, conditionally on the level of the random effect, iid for $k = 1,\ldots,n_{ij}$. Note that, using these relations,

$$
\mathrm{Var}\left[\frac{1}{a^2\sqrt{b}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}\left[\frac{\left(\epsilon_{ijk}-\bar{\epsilon}_{ij\cdot}\right)^2}{n_{ij}-1}\right.\right.
$$
$$
\left.\left.-E\left(\frac{\left(\epsilon_{ijk}-\bar{\epsilon}_{ij\cdot}\right)^2}{n_{ij}-1}\right)\right]\mid \vec{S},\vec{T}\right]
$$
$$
=\frac{1}{a^4b}\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\,\mathrm{Var}\left[\frac{1}{n_{ij}-1}\left(\epsilon_{ijk}-\bar{\epsilon}_{ij\cdot}\right)^2\right.
$$
$$
\left.-E\left(\frac{1}{n_{ij}-1}\left(\epsilon_{ijk}-\bar{\epsilon}_{ij\cdot}\right)^2\right)\mid \vec{S},\vec{T}\right].
$$

(B.2)

Because we assume a finite fourth moment for the ϵ terms and a finite covariance for any two distinct squared ϵ terms, this conditional variance tends to zero. Therefore, by the Dominated Convergence Theorem, the unconditional variance also tends to zero, thus proving the proposition.

Proof for $S_{B,\mathrm{rand}}$. First, note that after some straightforward algebra we can write

$$
aS_{B,\mathrm{rand}}
$$
$$
=\frac{1}{ab}\sum_{i}\sum_{i_1\neq i}\sum_{j}\left[\left(\gamma_{ij}-\bar{\gamma}_{i\cdot}\right)\left(\gamma_{i_1j}-\bar{\gamma}_{i_1\cdot}\right)\right.
$$
$$
+\left(\bar{\epsilon}_{ij\cdot}-\bar{\epsilon}_{i\cdot\cdot}\right)\left(\bar{\epsilon}_{i_1j\cdot}-\bar{\epsilon}_{i_1\cdot\cdot}\right)\Big]
$$
$$
+\frac{1}{ab}\sum_{i}\sum_{i_1\neq i}\sum_{j}\left[\frac{2(a-1)}{a}\left(\gamma_{ij}-\bar{\gamma}_{i\cdot}\right)\left(\bar{\epsilon}_{i_1j\cdot}-\bar{\epsilon}_{i_1\cdot\cdot}\right)\right.
$$
$$
\left.+\frac{2}{ab}\left(\gamma_{i_1j}-\bar{\gamma}_{i_1\cdot}\right)\left(\bar{\epsilon}_{ij\cdot}-\bar{\epsilon}_{i\cdot\cdot}\right)\right].
$$

(B.3)

To complete the proof, we claim that $(1/ab)\sum_i\sum_{i_1\neq i}\sum_j(\bar{\epsilon}_{ij\cdot}-\bar{\epsilon}_{i\cdot\cdot})(\bar{\epsilon}_{i_1j\cdot}-\bar{\epsilon}_{i_1\cdot\cdot})$ converges to zero after showing that

the leading term of this product tends to zero. We proceed similarly for the last two terms in $S_{B,\mathrm{rand}}$. Finally, we show that $(1/ab)\sum_i\sum_{i_1\neq i}\sum_j\gamma_{ij}\bar{\gamma}_{i_1\cdot}$ tends to zero to finish the proof. Note that each of these terms has expected value zero, so to demonstrate these facts, we only need to show that the variances tend to zero. To this end, note that

(1) $\mathrm{Var}\left(\dfrac{1}{ab}\sum_i\sum_{i_1\neq i}\sum_j\bar{\epsilon}_{ij\cdot}\bar{\epsilon}_{i_1j\cdot}\right)$

$$
=\frac{1}{a^2b^2}\sum_{i_1}\sum_{i_2\neq i_1}\sum_{j_1}\sum_{i_3}\sum_{i_4\neq i_3}\sum_{j_2}E\left(\bar{\epsilon}_{i_1j_1\cdot}\bar{\epsilon}_{i_2j_1\cdot}\bar{\epsilon}_{i_3j_2\cdot}\bar{\epsilon}_{i_4j_2\cdot}\right)
$$
$$
=\frac{1}{a^2b^2}\sum_{i_1}\sum_{i_2\neq i_1}\sum_j E\left(\bar{\epsilon}_{i_1j\cdot}^2\bar{\epsilon}_{i_2j\cdot}^2\right)\longrightarrow 0,
$$

(2) $\mathrm{Var}\left(\dfrac{2(a-1)}{a^2b}\sum_i\sum_{i_1\neq i}\sum_j\gamma_{ij}\bar{\epsilon}_{i_1j\cdot}\right)$

$$
=\frac{4(a-1)^2}{a^4b^2}\sum_{i_1}\sum_{i_2\neq i_1}\sum_{j_1}\sum_{i_3}\sum_{i_4\neq i_3}\sum_{j_2}E\left(\gamma_{i_1j_1}\bar{\epsilon}_{i_2j_1\cdot}\gamma_{i_3j_2}\bar{\epsilon}_{i_4j_2\cdot}\right)
$$
$$
=\frac{4(a-1)^2}{a^4b^2}\sum_{i_1}\sum_{i_2\neq i_1}\sum_j E\left(\gamma_{i_1j}^2\bar{\epsilon}_{i_2j\cdot}^2\right)\longrightarrow 0,
$$

(3) $\mathrm{Var}\left(\dfrac{1}{ab}\sum_i\sum_{i_1\neq i}\sum_j\gamma_{ij}\bar{\gamma}_{i_1\cdot}\right)$

$$
=\mathrm{Var}\left(\frac{1}{ab^2}\sum_i\sum_{i_1\neq i}\sum_j\gamma_{ij}\gamma_{i_1j}\right)
$$
$$
+\mathrm{Var}\left(\frac{1}{ab^2}\sum_i\sum_{i_1\neq i}\sum_j\sum_{j_1\neq j}\gamma_{ij}\gamma_{i_1j_1}\right)
$$
$$
=\frac{1}{a^2b^4}\sum_{i_1}\sum_{i_2\neq i_1}\sum_{j_1}\sum_{i_3}\sum_{i_4\neq i_3}\sum_{j_2}E\left(\gamma_{i_1j_1}\gamma_{i_2j_1}\gamma_{i_3j_2}\gamma_{i_4j_2}\right)
$$
$$
+\frac{1}{a^2b^4}
$$
$$
\times\sum_{i_1}\sum_{i_2\neq i_1}\sum_{j_1}\sum_{j_2\neq j_1}\sum_{i_3}\sum_{i_4\neq i_3}\sum_{j_3}\sum_{j_4\neq j_3}E\left(\gamma_{i_1j_1}\gamma_{i_2j_2}\gamma_{i_3j_3}\gamma_{i_4j_4}\right)
$$
$$
=\frac{1}{a^2b^4}\sum_{i_1}\sum_{i_2\neq i_1}\sum_j E\left(\gamma_{i_1j}^2\gamma_{i_2j}^2\right)
$$
$$
+\frac{1}{a^2b^4}\sum_{i_1}\sum_{i_2\neq i_1}\sum_{j_1}\sum_{j_2\neq j_1}E\left(\gamma_{i_1j_1}^2\gamma_{i_2j_2}^2\right)\longrightarrow 0.
$$

(B.4)

This completes the proof of the proposition.

Acknowledgments

References

[1] H. Scheffé, *The Analysis of Variance*, John Wiley & Sons, New York, NY, USA, 1959.

[2] S. R. Searle, *Linear Models*, John Wiley & Sons, New York, NY, USA, 1971.

[3] S. F. Arnold, *The Theory of Linear Models and Multivariate Analysis*, John Wiley & Sons, New York, NY, USA, 1981.

[4] S. R. Searle, G. Casella, and C. E. McCulloch, *Variance Components*, John Wiley & Sons, New York, NY, USA, 1992.

[5] A. I. Khuri, T. Mathew, and B. K. Sinha, *Statistical Tests for Mixed Linear Models*, John Wiley & Sons, New York, NY, USA, 1998.

[6] J. Cornfield and J. W. Tukey, "Average values of mean squares in factorials," *Annals of Mathematical Statistics*, vol. 27, pp. 907–949, 1956.

[7] B. Winer, *Statistical Principles in Experimental Design*, McGraw-Hill, New York, NY, USA, 1971.

[8] R. R. Hocking, "A discussion of the two-way mixed model," *The American Statistician*, vol. 27, pp. 148–152, 1973.

[9] T. Gaugler and M. G. Akritas, "Mixed effects designs: the symmetry assumption and missing data," *Journal of the American Statistical Association*, vol. 107, no. 499, pp. 1230–1238, 2012.

[10] T. Gaugler and M. G. Akritas, "Testing for interaction in two-way random and mixed effects models: the fully nonparametric approach," *Biometrics*, vol. 67, no. 4, pp. 1314–1320, 2011.

[11] A. Bathke, "The ANOVA F test can still be used in some balanced designs with unequal variances and nonnormal data," *Journal of Statistical Planning and Inference*, vol. 126, no. 2, pp. 413–422, 2004.

[12] A. K. Gupta, S. W. Harrar, and Y. Fujikoshi, "Asymptotics for testing hypothesis in some multivariate variance components model under non-normality," *Journal of Multivariate Analysis*, vol. 97, no. 1, pp. 148–178, 2006.

[13] M. G. Akritas and N. Papadatos, "Heteroscedastic one-way ANOVA and lack-of-fit tests," *Journal of the American Statistical Association*, vol. 99, no. 466, pp. 368–382, 2004.

[14] S. R. Searle, *Linear Models for Unbalanced Data*, John Wiley & Sons, New York, NY, USA, 1987.

[15] V. Koltchinskii and E. Giné, "Random matrix approximation of spectra of integral operators," *Bernoulli*, vol. 6, no. 1, pp. 113–167, 2000.

[16] J. Aitchison, *The Statistical Analysis of Compositional Data*, Chapman & Hall, London, UK, 1986.