A NONPARAMETRIC TEST FOR A SEMIPARAMETRIC MIXED ANCOVA MODEL FOR A NESTED DESIGN

Maricar C. Moreno
Master of Science (Statistics)

ABSTRACT

A nonparametric test for a postulated semiparametric mixed analysis of covariance model for a nested design is developed. In the nested design, the parametric part corresponds to the main treatment and nested effects while the nonparametric part corresponds to the fixed covariate. A hybrid backfitting algorithm was used to estimate the model and a bootstrap-based test was used to test the significance of the treatment effects. Simulation shows that the proposed test procedure is correctly-sized for models with random and mixed treatment effects. The test also performs better when the replicate size is small and more powerful than the ordinary analysis of covariance in the presence of misspecification error, nonnormality of errors, and dominating covariate effect.

Keywords: *bootstrap-based hypothesis testing, semiparametric model, analysis of covariance, nested design, backfitting, variance components*

1. INTRODUCTION

In most multifactor experiments, factorial designs are often used to properly measure the response variable. When factors are crossed, all possible combination of the independent variables are of concern which also allows examination of interaction among factors. However, when some factors cannot be crossed to another factor, and each of these factors appear at only one level of another factor, then factors are said to be nested. (Schielzeth and Nakagawa, 2013).

One of the most important types of experimental designs is the nested or hierarchical design wherein levels of one factor are similar but not identical for the different levels of one or more other factors. Nested design of experiment is widely used in industrial settings (Montgomery, 2013). Moreover, data with nested structures are very common in studies of evolution and ecology (Quinn & Keough 2002).

Analysis of covariance (ANCOVA) is a method used to correctly measure the treatment effects by setting aside the effect of covariates. (Montgomery, 2016; Searle, Casella, and McCulloch 1992). The problem of covariate in the traditional analysis of covariance is that the effect of covariate is not properly accounted. Due to these circumstances, properly estimating treatment effects to a response is of main interest. Recently, semiparametric ANCOVA model has been focused due to its greater flexibility to explain data (Jiuang, 2015). Alao (2016) postulated a semiparametric mixed ANCOVA model estimated by imbedding restricted maximum likelihood estimation and smoothing splines regression into backfitting algorithm.

With the presence of covariates in a design, there is no existing model yet that properly measures the treatment effects of a nested experimental design, which also specifies the relationship of the response variable and covariate component in a nonparametric form. In this study, we postulated a semiparametric mixed ANCOVA for a design with nested factor and tested the treatment effects after properly estimating the proposed model.

The aim of this study is to estimate nonparametrically the covariate effect to capture all the possible heterogeneity among the experimental units and to estimate the nesting effect, then set these aside to properly test for the significance of the treatment effects. Specifically, this paper aims to (1) postulate a semiparametric ANCOVA model for a nested design, (2) test and characterize the significance of the treatment effects/variance component based on a hybrid-backfitting method; and (3) evaluate the size and power of the test using simulation studies.

2. REVIEW OF RELATED LITERATURE

2.1 Semiparametric Analysis of Covariance Model

A semiparametric ANCOVA model is a model which specifies the relationship between the response variable and the covariate component in a nonparametric form. When the response variable is believed to be linearly related to the treatment effect but the effect of the covariate on the response is unknown, then a semiparametric ANCOVA model will be appropriate in this case. (Jiang, 1995)

Alao (2016) postulated a semiparametric mixed ANCOVA model with nonparametric part corresponding to the fixed covariate and the parametric part corresponding to the random effects. The model is given by:

$$Y_{ijk} = f(X_{ijk}) + \tau_j + \delta_k + (\tau\delta)_{jk} + \epsilon_{ijk}\ , \begin{cases} i = 1,2,\dots n\ observations \\ j = 1,2,\dots,p\ treatments \\ k = 1,2,\dots,q\ treatments \end{cases}$$

where

Y_{ijk} is the response variable of the $ijkl^{th}$ observation

X_{ijk} is fixed covariate score of the $ijkl^{th}$ observation

$f(X_{ijk})$ is the smoothed function of X_{ijk} (nonparametric part)

τ_j , δ_k are effects of treatment j and treatment k respectively

$(\tau\delta)_{jk}$ is the random interaction effect of the jk^{th} combination

ϵ_{ijk} is the error component

This model assumes that Y_{ijk} is a continuous response variable, *X* is a fixed and continuous covariate, the main effects τ_j and δ_k are random, the interaction $\tau\delta$ is random, the predictors are independent to each other; and the subjects are randomly chosen from a defined population and are assigned randomly to treatment groups. Two estimated procedures were proposed, one is a modified, iterative estimation procedure for semiparametric mixed ANCOVA model infusing REML and non-parametric regression (ARMS); and the next one is by incorporating a bootstrap approach into the backfitting framework of the first procedure (B-ARMS).

2.2 Nested Design

A nested model, or hierarchical model is an experimental design wherein the levels of one factor are similar but are not identical for different levels of another factor. When one factor is nested under the levels of another factor, then the design would be a two-stage nested design (Montgomery, 2013). An illustration where each level of a nested factor B is connected to only one unique level of the higher-level factor A is shown in Figure 1.

Figure 1. Schematic illustration of nested design

		Factor A	
		A_1	A_2
Factor B	B_1	X	
	B_2	X	
	B_3		X
	B_4		X

The cell marked with X shows the combination of *AB* included in the design. The illustration above shows that each level of factor A is associated with a unique level of the level of factor B.

The difference of nested design to a factorial design is that in crossed design, at least one level of each factor is connected to more than one level of the other factor as illustrated below:

Figure 2. Schematic illustration of crossed design

		Factor A	
		A_1	A_2
Factor B	B_1	X	X
	B_2	X	X

3. METHODOLOGY

3.1 Postulated Semiparametric Mixed ANCOVA model for Nested Design

This study will focus on a semiparametric mixed ANCOVA model with a nested factor. The parametric part corresponds to the treatment effects and nested effect while the nonparametric part corresponds to the fixed covariate. In this postulated model, two factors cross with each other, and one factor is nested in one of them.

The model is defined as:

$$Y_{(jkl)i} = f\left(X_{(jkl)i}\right) + \tau_j + \delta_k + (\tau\delta)_{jk} + \gamma_{(k)l} + \epsilon_{(jkl)i}$$

$$\begin{cases} i = 1,2, \dots n \; observations \\ j = 1,2, \dots, p \; treatments \\ k = 1,2, \dots, q \; treatments \\ l = 1,2, \dots m \; treatments \end{cases}$$

where

$Y_{(jkl)i}$ is the response variable of the $ijkl^{th}$ observation

$X_{(jkl)i}$ is fixed covariate score of the $ijkl^{th}$ observation

$f\left(X_{(jkl)i}\right)$ is the smoothed function of $X_{(jkl)i}$ (nonparametric)

τ_j , δ_k are effects of treatment j and treatment k respectively

$(\tau\delta)_{jk}$ is the interaction effect of the jk^{th} combination

$\gamma_{(k)l}$ is the effect of the l^{th} treatment nested within the k^{th} treatment

$\epsilon_{(jkl)i}$ is the error component

The assumptions of the postulated model are as follows:

1. Y is a continuous response variable.
2. X is fixed and a continuous covariate.
3. τ_j , δ_k are effects used to address the effect of treatment levels.
4. $(\tau\delta)_{jk}$ is the interaction effect used to address the effect of jk^{th} treatment combinations.
5. $\gamma_{(k)l}$ is the nesting effect used to address the effect of l^{th} treatment nested within the k^{th} treatment levels.
6. The independent variable and the covariate are independent of each other.
7. The random effects and unobservable random errors are independent of each other.
8. The model follows a balanced design wherein p corresponds to the number of levels of factor j, q as the number of levels of factor k, and m as the number of levels of factor l within the k^{th} treatment

In this model, all levels of factor A appear simultaneously in all levels of factor B while the level of factor C can only be determined by knowing the level of factor B. The model is illustrated below:

Figure 4. Schematic illustration of a three-stage/factor nested design

		Factor A	
		A_1	A_2
Factor B	B_1	$\mathrm{C}_{(1)1}$ $Observations \begin{cases} y_{1111} \\ y_{1112} \end{cases}$	$\mathrm{C}_{(1)3}$ y_{2131} y_{2132}
	B_2	$\mathrm{C}_{(2)2}$ y_{1221} y_{1222}	$\mathrm{C}_{(2)4}$ y_{2241} y_{2242}

That is, if the level of factor B is B_1, then factor C would be $\mathrm{C}_{(1)1}$ and $\mathrm{C}_{(1)3}$, while if the higher-level factor B is B_2, then factor C would be $\mathrm{C}_{(2)2}$ and $\mathrm{C}_{(2)4}$. This design is widely used experiments in industrial settings where one factor cannot be associated to all of the settings in another factor, or can only be determined by identifying its high-level factor.

3.2 Estimation Procedure

The postulated model will be estimated through a hybrid backfitting algorithm by Alao (2016), called B-ARMS (Bootstrapped Analysis of Covariance via REML with splines) modified for nested models. The ARMS method alternately estimates the variance components for the random-effects model and the smooth function of the covariate. Using REML, variance component corresponding to the parametric part are first estimated by ignoring the effect of the nonparametric component which is the covariate. A nonparametric regression using the computed residuals after fitting the model without the nonparametric part will then be used to estimate the smooth function of the covariate. Using B-ARMS, estimates from ARMS are used and resampling with replacement of the residuals is applied. Note that for this study, a nesting specification is added in the model. Further, a confidence interval based on bootstrap percentiles is constructed and used to test the significance of the nested effect.

Below is the step-by-step procedure for the estimation of model components:

Algorithm using ARMS

The algorithm of ARMS below is applicable to models with random effects. Ordinary least squares regression was used in estimation of parameters instead of REML for fixed-effects model.

Step 1: Fit the nested part of the model using REML method with nesting specification by ignoring $f\left(X_{(jkl)i}\right)$ and the parametric part $\tau_j + \delta_k + (\tau\delta)_{jk}$ so that the model would be $Y_{(jkl)i} = \gamma_{(k)l} + \epsilon_{(jkl)i}$. This will contain estimates of the components which we are of interest.

Step 2: Compute the partial residuals $p_{(jkl)i} = Y_{(jkl)i} - \hat{Y}_{(jkl)i}$. These partial residuals now contain information on $f\left(X_{(jkl)i}\right)$ and $\tau_j + \delta_k + (\tau\delta)_{jk}$ which will then be used to estimate $f\left(X_{(jkl)i}\right)$ and $\tau_j + \delta_k + (\tau\delta)_{jk}$ consecutively.

Step 3: Using smoothing spline, estimate $f\left(X_{(jkl)i}\right)$ nonparametrically, $p_{(jkl)i} = f\left(X_{(jkl)i}\right) + error$.

Step 4: Compute the new partial residuals $p^*_{(jkl)i} = p_{(jkl)i} - \hat{f}\left(X_{(jkl)i}\right)$. These partial residuals contain information on $\tau_j + \delta_k + (\tau\delta)_{jk}$ and thus, will be used to estimate the main effects.

Step 5: Estimate the parametric part containing the main effects using REML. The random model would be $p^*_{(jkl)i} = \tau_j + \delta_k + (\tau\delta)_{jk} + error$.

Step 6: Compute for the new partial residuals $e_{(jkl)i} = Y_{(jkl)i} - \hat{f}\left(X_{(jkl)i}\right) - (\hat{\tau}_j + \hat{\delta}_k + \widehat{(\tau\delta)}_{jk})$. These partial residuals contain information on the nesting effect and then be used to estimate the component $\gamma_{(k)l}$ in Step 1.

Step 7: Repeat steps 1 to 6 until convergence, i.e. the change in the new estimates from the previous estimates do not change more than the tolerance level (0.001)

The nested part of the design is estimated first since it poses the limitation on the structure of the model. Further, the nested factor limits the constraints of the lower-level factor. In many cases, the nested effects contribute first in the model and so the covariates are estimated last or at a latter stage in the hierarchy.

Algorithm using B-ARMS:

Step 1: Generate the initial estimates and residuals $\emptyset_{(jkl)i} = Y_{(jkl)i} - \hat{Y}_{(jkl)i}$ by fitting the model (2) using ARMS. R bootstrap samples of residuals in Step 2 below are obtained using these residuals, while the initial estimates are used to compute for the new values of dependent variables in Step 3 below.

Step 2: Generate new set of residuals ϵ^* by obtaining samples of $\emptyset_{(jkl)i}$ from Step 1 with replacement from {1, 2, …, n}. These residuals will be used to compute the new values of Y in Step 3.

Step 3: Compute for the new values of the dependent variable,

$Y^*_{(ijk)l} = \hat{f}\left(X_{(jkl)i}\right) + \hat{\tau}_j + \hat{\delta}_k + \widehat{(\tau\delta)}_{jk} + \hat{\gamma}_{(k)l} + \epsilon^*_{(jkl)i}$, where $\hat{f}\left(X_{(jkl)i}\right)$, $\hat{\tau}_j$, $\hat{\delta}_k$, $\widehat{(\tau\delta)}_{jk}$ and $\hat{\gamma}_{k(l)}$ are estimates from Step 1 and new residuals $\emptyset^*_{(jkl)i}$ from Step 2.

Step 4: Fit the model from the pseudo data in Step 3 using ARMS. This contains the estimates of the components from the dependent variable $Y^*_{(jkl)i}$.

In order to test the significance of the nesting effect given the postulated model, an empirical distribution of the test statistic will be obtained to test if the l^{th} nested factor has no effect on the response by obtaining an empirical distribution of the test statistic by repeating Steps 2 to 4 of B-ARMS algorithm R times then constructing a 95% bootstrap confidence interval using the empirical distribution and reject the null hypothesis if the interval contains 0.

For models with fixed effects, the hypotheses being tested at 0.05 level of significance are:

Ho: $\gamma_{(k)l} = 0,\quad$ $the\ l^{th}\ treatment\ is\ not\ significant\ on\ the\ response$

Ha: $\gamma_{(k)l} \neq 0,\quad$ $the\ l^{th}\ treatment\ is\ significant\ on\ the\ response$

For models with random effects, variance components of the nested factor are tested at 0.05 level of significance. The hypotheses being tested are:

Ho: $\sigma^2_{(k)l} = 0,\quad$ $the\ l^{th}\ treatment\ is\ not\ significant\ on\ the\ response$

Ha: $\sigma^2_{(k)l} \neq 0,\quad$ $the\ l^{th}\ treatment\ is\ significant\ on\ the\ response$

To evaluate the test, power and size of the test will be computed. The power of the test will be computed by constructing the confidence interval 200 times and the proportion of confidence interval that does not contain 0, or the rejected null hypothesis is the power of the test. For size of the test, data will be generated without the treatment effects and then confidence interval will be constructed using the steps in 1 to 4 in Section 3.2. The proportion of rejected null hypothesis is the size of the test.

3.3 Simulation

In order to conduct the estimation and hypothesis testing of the postulated model, data are simulated under different scenarios. The experimental design has 3 factors, two of which are crossed to each other, and the third factor is nested to the other factor. Models with all fixed, all random effects, and mixed effects are considered.

For the nonparametric part, a linear and nonlinear function is introduced to the model. The latter to impose heterogeneity and to verify the predictive ability of the postulated method when nonlinearity is present in the model.

Furthermore, a constant *m* and *w* are also imposed in the model such that the data generating process will become $Y_{(ijk)l} = f\left(X_{(ijk)l}\right) + \mathrm{m} * \tau_j + m * \delta_k + (\tau\delta)_{jk} + \gamma_{(k)l} + w * \epsilon_{(ijk)l}$.

The constant *m* is introduced to determine the effect of the covariate and factors, while the constant *w* is imposed to induce misspecification error. To simulate contamination in the covariate, a constant $\beta = 2$ is introduced. Further, scenarios where the nonparametric part of the model is linear or nonlinear are also observed. Functional form of $f\left(X_{(ijk)l}\right) = \beta X_{(ijk)l}$ indicates a linear form, while $f\left(X_{(ijk)l}\right) = ExpX_{(ijk)l}$ manifests nonlinearity of covariates.

In addition, all the settings were simultaneously applied to different levels of percentage differences between the values assigned to levels of $\tau_j, \delta_k,\ and\ \gamma_{k(l)}$. The values of the parameters $\tau_j, \delta_k, and\ \gamma_{k(l)}$ in fixed effects model is summarized in Appendix. For the random-effect models where $\tau_j, \delta_k, (\tau\delta)_{jk}, and\ \gamma_{k(l)}$ are random factors, all are assumed to be normally distributed.

Table 2 shows the variables considered for the different scenarios and the distributions of models with random factors:

Table 2. Variables and cases for simulation

Variable	Cases
1. Distribution of $X_{(ijk)l}$	Normal (8,2)
2. Functional form of $f\left(X_{(ijk)l}\right)$	$\beta X_{(ijk)l}$ $Exp(\beta X_{(ijk)l})$
3. Value of β	1,2
4. Distribution of τ_j	Normal (μ_j, σ_j)
5. Distribution of δ_k	Normal (μ_k, σ_k)
6. Distribution of interaction, $(\tau\delta)_{jk}$	Normal (μ_{jk}, σ_{jk})
7. Distribution of nesting effect, $\gamma_{(k)l}$	Normal $(\mu_{(k)l}, \sigma_{(k)l})$
8. Constant *m, w*	1,2
9. Distribution of error term, $\epsilon_{(ijk)l}$	Normal (0,1.5) Cauchy (0,1.5)
10. Number of treatments *p,q*	2 levels for p and q
11. Number of *l or* nested levels	1 level of nesting
12. Replicate size	2 and 5 replicates

4. RESULTS AND DISCUSSION

The power and size were compared with the parametric nested ANCOVA test to evaluate the proposed nonparametric test for a semiparametric mixed analysis of covariance for a nested design. The significance of the estimates of the nesting effect were being tested for the model with fixed effects, meanwhile, the significance of the variance components of the model with random effects were being tested.

4.1 Overall Size of the Test

The test is considered correctly sized if the estimated probability of rejecting a true null hypothesis is less than or equal to the level of significance set in the study, that is, 5%. In order to compute for the size, treatment effects in the proposed and baseline models were removed in the simulation of data and significance of the nesting effect were then tested.

Based from the results, fixed effects model is not correctly sized. In fact, most of the time, the proposed test rejects the null hypothesis. The baseline test is also not correctly sized at 0.05 level of significance but are much smaller than the proposed test.

Table 3. Average size of the proposed and baseline tests under fixed effects model

Scenario	Proposed	Baseline
Data without contamination	0.9744	0.1019
Data with contamination	0.9884	0.0888
Dominating covariate effect	0.9844	0.0922
Minimal covariate effect	0.9784	0.0984
Without misspecification error	0.9831	0.0981
With misspecification error	0.9797	0.0925
Normal errors	1.0000	0.0969
Non-normal errors	0.9628	0.0938
Linear covariate	0.9631	0.0922
Non-linear covariate	0.9997	0.0984
Replicate =2	0.9906	0.0928
Replicate = 5	0.9722	0.0978

For random effects model, all scenarios in the proposed test procedure are correctly sized except in the case when replicates are set to 2.

Table 4. Average size of the proposed and baseline tests in different scenarios on random effects model

Scenario	Proposed	Baseline
Data without contamination	0.0413	0.0116
Data with contamination	0.0363	0.0084
Dominating covariate effect	0.0375	0.0113
Minimal covariate effect	0.0400	0.0088
Without misspecification error	0.0375	0.0094
With misspecification error	0.0400	0.0106
Normal errors	0.0428	0.0116
Non-normal errors	0.0347	0.0084
Linear covariate	0.0425	0.0091
Non-linear covariate	0.0350	0.0109
Replicate =2	0.0725	0.0050
Replicate = 5	0.0191	0.0009

Looking at mixed effects model wherein one factor is random (τ), the other factor is fixed (δ), and the nesting effect (γ) is random, the size of the proposed test is correctly sized except for the scenario when the replicate is only 2.

Table 5. Average size of the proposed and baseline tests in different scenarios on mixed effects model

Scenario	Proposed	Baseline
Data without contamination	0.0269	0.0069
Data with contamination	0.0316	0.0084
Dominating covariate effect	0.0306	0.0044
Minimal covariate effect	0.0278	0.0109
Without misspecification error	0.0297	0.0088
With misspecification error	0.0288	0.0066
Normal errors	0.0300	0.0025
Non-normal errors	0.0284	0.0128
Linear covariate	0.0309	0.0106
Non-linear covariate	0.0275	0.0047
Replicate =2	0.0547	0.0066
Replicate = 5	0.0038	0.0088

4.2 Overall Power of the Test

The proposed test is evaluated based on its power. The higher the power, the higher the probability that the test arrived at a correct decision. To further evaluate the power of the proposed test, differences between the levels of each treatment were considered. The levels of percentage differences of the values between levels of each treatment effects taken into account were 10%, 30%, 50% and 100%.

For the fixed effects and random effects model, the test is powerful in detecting the significance of the nesting effect in all scenarios compared to the baseline test. Notably, as the level of differences in treatment increases, power also increases.

Table 6. Average power of the proposed and baseline tests in different percentage differences in treatment levels on fixed effects model

Differences in Treatment Levels	Proposed	Baseline
10%	0.9928	0.1847
30%	0.9948	0.3691
50%	0.9969	0.4427
100%	0.9978	0.5092

Table 7. Average power of the proposed and baseline tests in different percentage differences in treatment levels on random effects model

Differences in Treatment Levels	Proposed	Baseline
10%	0.6420	0.4403
30%	0.6831	0.5038
50%	0.7217	0.5597
100%	0.7931	0.6650

For the case of the mixed effects model where one factor is random (τ), the second factor is fixed (δ), and the nesting factor is random (γ), power is relatively low compared to the model with all random factors. However, the proposed test procedure has greater power compared to the baseline test across all the levels of differences in the treatments.

Table 8. Average power of the proposed and baseline tests in different percentage differences in treatment levels on mixed effects model

Differences in Treatment Levels	Proposed	Baseline
10%	0.3778	0.2750
30%	0.3997	0.3106
50%	0.4252	0.3441
100%	0.4663	0.4043

4.3 Effect of the degree of contribution of covariate

To further evaluate the proposed test procedure, a constant *m* is introduced to determine the effect of the contribution of the covariate. This is simulated by setting the value of m=2 when the role of the covariate is minimal, and m=1 when the covariate is dominating.

Table 9 shows that the power of proposed test procedure for fixed effects model is generally much higher than the computed power of the baseline test. However, the baseline test procedure posted relatively higher power when the effect of the covariate is minimal compared to its power when the role of the covariate is dominating.

Table 9. Average power of the proposed and baseline tests from the data with minimal and dominating covariate effects on fixed effects model

Fixed by Fixed	Dominating Covariate		Minimal Covariate	
	Proposed	Baseline	Proposed	Baseline
10%	0.9938	0.1738	0.9919	0.1956
30%	0.9953	0.3563	0.9944	0.3819
50%	0.9975	0.4347	0.9963	0.4506
100%	0.9984	0.5075	0.9972	0.5109

When the treatment effects are all random as shown in Table 10, the proposed test procedure is powerful. Notably, as the percentage difference in the levels of treatment effects increases, the power of both the proposed test and baseline test also increases. However, the baseline test registered higher power when the effect of the covariate is dominant. On the other hand, the proposed test procedure resulted to higher power when the covariate is dominating (vs minimal covariate effect) when the level of percentage difference is at 100%.

Table 10. Average power of the proposed and baseline tests from the data with minimal and dominating covariate effect on random effects model

Random by Random	**Dominating Covariate**		**Minimal Covariate**	
	Proposed	**Baseline**	**Proposed**	**Baseline**
10%	0.6388	0.4444	0.6453	0.4363
30%	0.6825	0.5116	0.6838	0.4959
50%	0.7206	0.5647	0.7228	0.5547
100%	0.7988	0.6691	0.7875	0.6609

The test for mixed effects model had much lower power compared to the fixed and random effects, however, comparing with the baseline test for mixed effects model, the former performs better.

Table 11. Average power of the proposed and baseline tests from the data with minimal and dominating covariate effect on mixed effects model

Random by Fixed	**Dominating Covariate**		**Minimal Covariate**	
	Proposed	**Baseline**	**Proposed**	**Baseline**
10%	0.3813	0.2834	0.3744	0.2665
30%	0.3981	0.3134	0.4013	0.3078
50%	0.4275	0.3444	0.4228	0.3438
100%	0.4672	0.4013	0.4654	0.4074

4.4 Effect of misspecification error

To induce misspecification, a constant w is imposed in the model, particularly to the error term to further evaluate the performance of the proposed test procedure. A value of $w=2$ is multiplied to the error term to simulate misspecification while $w=1$ to simulate no misspecification error.

Table 12 shows that the proposed test procedure for fixed effects model is generally powerful with or without misspecification error. It can be observed that the power is higher when there is no misspecification error, both for the proposed and baseline test procedure.

Table 12. Average power of the proposed and baseline tests from the data with and without misspecification error on fixed effects model

Fixed by Fixed	**Without Misspecification Error**		**With Misspecification Error**	
	Proposed	Baseline	Proposed	Baseline
10%	0.9953	0.2306	0.9903	0.1388
30%	0.9966	0.4225	0.9931	0.3156
50%	0.9975	0.4766	0.9963	0.4088
100%	0.9991	0.5281	0.9966	0.4903

Similarly, the power of the test decreases in the presence of misspecification error when all the treatment effects are random. Although the power of the test from the data with misspecification is still at par compared to the test from the data without misspecification error.

Table 13. Average power of the proposed and baseline tests from the data with and without misspecification error on random effects model

Random by Random	**Without Misspecification Error**		**With Misspecification Error**	
	Proposed	Baseline	Proposed	Baseline
10%	0.6794	0.4931	0.6047	0.3875
30%	0.7150	0.5463	0.6513	0.4613
50%	0.7481	0.5953	0.6953	0.5241
100%	0.8056	0.6916	0.7806	0.6384

Table 14 shows that the test for the mixed effects model resulted to much lower power compared to the fixed and random effects model. Similarly, power is relatively higher when there is no misspecification error. The proposed test procedure also has higher power compared to the baseline test.

Table 14. Average power of the proposed and baseline tests from the data with and without misspecification error on mixed effects model

Random by Fixed	**Without Misspecification Error**		**With Misspecification Error**	
	Proposed	**Baseline**	**Proposed**	**Baseline**
10%	0.4128	0.3180	0.3428	0.2319
30%	0.4259	0.3488	0.3734	0.2725
50%	0.4484	0.3781	0.4019	0.3100
100%	0.4807	0.4266	0.4519	0.3820

4.5 Effect of type of error

The power of test is evaluated when the error term is normally distributed and when the error term is non-normally distributed, i.e. Cauchy distributed.

Table 15 shows that when the error term is normally distributed, the proposed test procedure always detects significance of the nesting effect for fixed effects model. The baseline test also has high power which shows that when the treatment effects are all fixed, the null hypothesis is rejected most of the time.

For non-normal errors, both the proposed test procedure and the baseline test are powerful and increases as the percentage differences between the levels of treatment effects increases.

Table 15. Average power of the proposed and baseline tests from the data with normal and non-normal errors on fixed effects model

Fixed by Fixed	**Normal Errors**		**Non-Normal Errors**	
	Proposed	Baseline	Proposed	Baseline
10%	1.0000	0.2541	0.9856	0.1153
30%	1.0000	0.4966	0.9897	0.2416
50%	1.0000	0.5400	0.9938	0.3453
100%	1.0000	0.5459	0.9956	0.4725

Power is also higher when the error are normal both for the proposed test procedure and the baseline test for random effects model. As percentage differences between the levels of treatment effect increases, the power also increases.

Table 16. Average power of the proposed and baseline tests from the data with normal and non-normal errors on random effects model

Random by Random	Normal Errors		Non-Normal Errors	
	Proposed	Baseline	Proposed	Baseline
10%	0.7600	0.5844	0.5241	0.2963
30%	0.7775	0.6338	0.5888	0.3738
50%	0.7975	0.6794	0.6459	0.4400
100%	0.8366	0.7447	0.7497	0.5853

For mixed effects model, the proposed test resulted to higher power when the error is normally distributed except for the case when the percentage difference between the treatment levels is at 100%. The proposed test also resulted to higher power compared to the baseline test.

Table 17. Average power of the proposed and baseline tests from the data with normal and non-normal errors on mixed effects model

Random by Fixed	Normal Errors		Non-Normal Errors	
	Proposed	**Baseline**	**Proposed**	**Baseline**
10%	0.4631	0.3771	0.2925	0.1728
30%	0.4759	0.4147	0.3234	0.2066
50%	0.4881	0.4425	0.3622	0.2456
100%	0.5053	0.4734	0.4273	0.3352

4.6 Effect of nonlinearity of covariates

The power of the proposed and baseline test procedure is also evaluated when the nonparametric part of the model is a linear and nonlinear function. The latter is to impose heterogeneity in the postulated model and to evaluate the effect in the power of the test when there is a presence of nonlinearity.

Table 18 shows that the proposed test procedure detects significance in the nesting effect 100% of the time when the covariate is nonlinear for fixed effects model. In contrast, power is very low when the covariate is nonlinear for the baseline test.

Table 18. Average power of the proposed and baseline tests from the data with linear and nonlinear covariate on fixed effects model

Fixed by Fixed	Linear Covariate		Nonlinear Covariate	
	Proposed	Baseline	Proposed	Baseline
10%	0.9859	0.2713	0.9997	0.0981
30%	0.9897	0.6416	1.0000	0.0966
50%	0.9938	0.7888	1.0000	0.0966
100%	0.9956	0.9194	1.0000	0.0991

For the random effects model, the test is powerful when the covariate is linear. The proposed test also had higher power compared to the baseline test.

Table 19. Average power of the proposed and baseline tests from the data with linear and nonlinear covariate on random effects model

Random by Random	**Linear Covariate**		**Nonlinear Covariate**	
	Proposed	Baseline	Proposed	Baseline
10%	0.8353	0.6372	0.4488	0.2434
30%	0.8706	0.7003	0.4956	0.3072
50%	0.9013	0.7541	0.5422	0.3653
100%	0.9484	0.8559	0.6378	0.4741

Table 20 shows that power is very low when the covariate is nonlinear, both for the proposed test procedure and the baseline test for mixed effects model. Notably, power is the same across the different percentage differences in the levels of treatment effects for the proposed test, except at 100% percentage difference. When the covariate is linear, the proposed test procedure registered higher power compared to the baseline.

Table 20. Average power of the proposed and baseline tests from the data with linear and nonlinear covariate on mixed effects model

Random by Fixed	**Linear Covariate**		**Nonlinear Covariate**	
	Proposed	**Baseline**	**Proposed**	**Baseline**
10%	0.7291	0.5431	0.0266	0.0069
30%	0.7728	0.6138	0.0266	0.0075
50%	0.8238	0.6816	0.0266	0.0066
100%	0.9051	0.8021	0.0275	0.0066

4.7 Effect of replicate size

Table 21 shows that when the replicate size is 2, the proposed test procedure is relatively powerful. The baseline test, however, posted higher power when the number of replicate is five.

Table 21. Average power of the proposed and baseline tests from the data with two and five replicates on fixed effects model

Fixed by Fixed	**Replicate = 2**		**Replicate = 5**	
	Proposed	Baseline	Proposed	Baseline
10%	0.9963	0.1303	0.9894	0.2391
30%	0.9981	0.3263	0.9916	0.4119
50%	0.9981	0.4256	0.9956	0.4597
100%	0.9981	0.5022	0.9975	0.5163

Table 22. Average power of the proposed and baseline tests from the data with two and five replicates on random effects model

Random by Random	**Replicate = 2**		**Replicate = 5**	
	Proposed	Baseline	Proposed	Baseline
10%	0.6425	0.4181	0.6416	0.4625
30%	0.6841	0.4863	0.6822	0.5213
50%	0.7253	0.5484	0.7181	0.5709
100%	0.8009	0.6653	0.7853	0.6647

Table 23. Average power of the proposed and baseline tests from the data with two and five replicates on mixed effects model

Random by Fixed	**Replicate = 2**		**Replicate = 5**	
	Proposed	**Baseline**	**Proposed**	**Baseline**
10%	0.3766	0.2418	0.3791	0.3081
30%	0.3959	0.2825	0.4034	0.3388
50%	0.4222	0.3219	0.4281	0.3663
100%	0.4666	0.3863	0.4660	0.4224

5. SUMMARY AND RECOMMENDATION

The results of the simulation show that the proposed test is correctly sized for random and mixed effects model. Size is also comparable to the baseline tests except for fixed effects model where the test is found to be incorrectly sized.

The proposed test is also found to be more powerful than the baseline test across the different percentage differences in treatment levels compared to the baseline test procedure. In the presence of contamination of the covariate, dominating covariate effect, misspecification error and non-normal errors, the proposed test procedure performs better than the baseline test.

However, in terms of linear covariate, the fixed effects model detects nesting effect 100% of the time, while the baseline test had much lower power (less than 1%). Notably, the proposed test is more powerful when the replicate size is small (r=2), than when the replicate size is large (n=5).

References

Alao V. (2016). Estimation of Semiparametric Mixed Analysis of Covariance Model. *University of the Philippines Diliman*

Buja, A., Hastie T., and Tibshirani, R.(1989), Linear Smoothers and Additive Models*, The annals of Statistics*, 17(2):453-510

Burdick, R.K., Borror, C.M., and Montgomery, D.C. et al (2005). *Design and Analysis of Gauge R&R Studies: Making Decisions with Confidence Intervals in Random and Mixed ANOVA models.* ASA-SIAM Series on Statistics and Applied Probability SIAM, Philadelphia,ASA, Alexandria, VA,

Efron, B.(1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1-26

Hastie, T.J., and Tibshirani, R.J.(1986) *Generalized Additive Models*, Statistical Science, 1(3): 297-310

Gentle, J., Härdle, W., Mori, Y., 2012 (2004), *Hand Book of Computational Statistics: Concepts and Methods,* Heidelberg: Springer

Good, P., and Lunneborg, Cliff. (2006). Limitations of the Analysis of Variance. *Journal of Modern Applied Statistical Methods,* Vol 5., No. 1, 41-43

Jiang, Yan (1995). Semiparametric ANCOVA using shape restrictions. *University of Georgia Theses and Dissertions* accessed 5 May 2018, https://athenaeum.libs.uga.edu/handle/10724/22408

Keppel, G. and Wickens, T.D. (2004*). Design and Analysis: A Researcher's Handbook. 4thed.* New York: Prentice Hall

Milliken,G. and Johnson, E. (2009). *Analysis of Messy Data Volume 1 Designed Experiments 2nd Ed.* New York: Chapman & Hall/CRC

Montgomery, D.C. (2013). *Design and Analysis of Experiments*. 8th ed. New York: John Wiley & Sons, Inc.

Opsomer, J.D.(2000). Asymptotic Properties of Backfitting Estimators. *Journal of Multivariate Analysis* 73(2):166-179. The Stata Journal, 9(1):86-136

Paparoditis E. and Politis, D. (2005). Bootstrap hypothesis testing in regression models. *Statistics and Probability Letters,* Vol 74 Issue 4, 356-365

Quinn, G.P. & Keough, M.J. (2002) *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, Cambridge.

Ryan, T.P. (2007) *Modern Experimental Design.* John Wiley & Sons, Chichester

Scheffé H.(1959). *The Analysis of Variance*. John Wiley & Sons, Inc.

Schielzeth, H. and Nakagawa S. (2013). Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution,* 4, 14- 24

Searle, S. R. G., Casella, and C.E. McCulloch (1992). *Variance Components*. New Jersey: John Wiley & Sons, Inc.

Simonoff, J. S. 1996. *Smoothing Methods in Statistics*. New York: Springer-Verlag, Inc.

Smith, J.R. & Beverly, J. M. The use of and analysis of staggered nested factorial designs. *Journal of Quality Technology.* 1981 – Issue 3

Speckman, P. (1998). Kernel smoothing in partial linear models. *J.R. Statistic. Soc. B,* 50, 413-436

Tan W. Y and Cheng S. S. (1984). On testing variance components in three-stages unbalanced nested random effects models. *The Indian Journal of Statistics Series B,* Vol 46. No. 2, 188-200

Wang, Y. 2011. Smoothing Splines: Methods and Applications. New York: Chapman and Hall/CRC.

Xu, L., Mei B., Chen, R.R et. al (2014). Parametric bootstrap tests for unbalanced nested design under heteroscedasticity. *Journal of Statistical Computationand Simulation.* Vol 84, No. 9, 2059-2070

Young, S.G. and Bowman, A.W (1995). Non-parametric Analysis of covariance. *Biometrics,* 51, 920-931

Zhang, S., Paul, J., et. al (2014). Empirical comparison of four baseline covariate adjustment methods in analysis of continuous outcomes in randomized controlled trials. *Clin Epidemiol*. doi: 10.2147/CLEP.S56554. eCollection 2014.

APPENDIX
Appendix A

Table A1. Parameters used in different percentage differences for fixed effects model

Percentage Difference	τ		δ		λ			
	Level 1	**Level 2**	**Level 1**	**Level 2**	**L11**	**L12**	**L21**	**L22**
0%	2	2	5	5	7	7	7	7
10%	2	2.2	5	5.5	7	7.7	8.47	9.31
30%	2	2.6	5	6.5	7	9.1	11.83	15.379
50%	2	3	5	7.5	7	10.5	15.75	23.625
100%	2	4	5	10	7	14	21	31.5

Table A2. Parameters used in different percentage differences for random effects model

Percentage Difference	τ		δ		λ			
	Level 1	**Level 2**	**Level 1**	**Level 2**	**L11**	**L12**	**L21**	**L22**
0%	None	None	None	None	None	None	None	None
10%	(0,2)	(0,2.2)	(0,5)	(0,5)	(50,20)	(50,22)	(50,24.2)	(50,26.62)
30%	(0,2)	(0,2.6)	(0,5)	(0,5.5)	(50,20)	(50,26)	(50,33.8)	(50,43.94)
50%	(0,2)	(0,2.3)	(0,5)	(0,6.5)	(50,20)	(50,30)	(50,45)	(50,67.5)
100%	(0,2)	(0,4)	(0,5)	(0,10)	(50,20)	(50,40)	(50,80)	(50,160)

Table A3. Parameters used in different percentage differences for mixed effects model

Percentage Difference	τ (random)		δ (fixed)		λ(random)			
	Level 1	**Level 2**	**Level 1**	**Level 2**	**L11**	**L12**	**L21**	**L22**
0%	None	None	2	2	None	None	None	None
10%	(8,2)	(8,2.2)	2	2.2	(50,20)	(50,22)	(50,24.2)	(50,26.62)
30%	(8,2)	(8,2.6)	2	2.6	(50,20)	(50,26)	(50,33.8)	(50,43.94)
50%	(8,2)	(8,2.3)	2	3	(50,20)	(50,30)	(50,45)	(50,67.5)
100%	(8,2)	(8,4)	2	4	(50,20)	(50,40)	(50,80)	(50,160)