A comparison of restricted maximum likelihood and method of moments variance estimation for small-sample split-plot experiments

by

Xuan Xu

B.S., China Agricultural University, 2008
M.S., China Agricultural University, 2010
Ph.D., Kansas State University, 2019

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2019

Approved by:

Major Professor
Dr. Christopher I. Vahl

Copyright

Abstract

Researchers may choose to perform an experiment using a split-plot design over more simple designs such as a completely randomized design or a randomized complete block design in order to conserve scarce resources. A split-plot becomes attractive when some treatment factors are more costly to apply to the experimental units or when it is difficult to change one factor from level to level. In such a case it may be more efficient to apply these costly treatments to a small set of larger experimental units (i.e. whole plots) and then apply the less costly treatments to more numerous smaller experimental units (i.e. subplots) nested within the larger ones. Because the subplot and whole-plot experimental units each have a corresponding variance component, the analysis of a split-plot study is more complicated. Making the split-plot analysis even more challenging, cost considerations may also lead to relatively small sample sizes for the whole-plot treatments. An unintended consequence is that some variance components in the split-plot design's model may be poorly estimated which in turn may have an unanticipated effect on the type I error rates for tests of the fixed effects.

As a motivating example, alfalfa yield data from a field study with a split-plot design with four randomized complete blocks at the whole-plot level serves as the basis for a simulation study to estimate the type I error rates of three fixed effects (whole-plot main effects, subplot main effects and whole-plot by subplot interaction). Several other scenarios where the number of blocks and the relative magnitudes of the variance components are varied are also explored. For each scenario, 10,000 data sets were randomly generated assuming normally distributed errors. Two linear mixed models were fit to each data set using the MIXED procedure in SAS; one method estimates the variance components via restricted maximum likelihood (REML) and the other by the method of moments (MoM) based on the type III sums of squares. The REML

models yielded inconsistent type I error rates for some tests of fixed effects compared to the MoM models but improved as the number of blocks increased. MoM models tended to hold their nominal type I error rates to within expected Monte Carlo error.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my gratitude to my advisor, Dr. Christopher I. Vahl, whose expertise, understanding and patience, engaged remarkably with my graduate experience. With his motivation and encouragement, I would not have considered a dual degree in Statistics. I appreciate his vast knowledge and skill not only in Statistics, but also in other areas, and his assistance in writing this report. I would like to thank other members of my committee, Dr. Perla Reyes, and Dr. Trevor Hefley for the assistance they provided at all levels of my research project. Appreciation also goes out to the office staff in the Department of Statistics for all the instances in which their assistance helped me along the way.

I would also like to thank my family for the support they provided me through my study and in particular, I must acknowledge my wife and daughters, without whose love and encouragement, I would not have finished this report.

Chapter 1 - Introduction

Background

Researchers planning an experiment may choose a split-plot design over a completely randomized design (CRD) or a randomized complete block design (RCBD) due to considerations of cost and efficiency (Jones and Nachtsheim, 2009). Split-plot designs were originally developed for agricultural experiments (Fischer, 1925) with the idea that it may be more economical to apply the costly treatments to larger but fewer experimental units (i.e. whole plots) and then apply the less costly treatments to smaller but more numerous experimental units (i.e. subplots) nested within the larger units. However, this makes the analysis of a split-plot experiment more complicated primarily because the split-plot has an extra set of random errors, one set for the whole plots and another set for the subplots. A basic assumption of the traditional analysis of a split-plot experiment is that each set of random errors has its own variance component which then must be estimated. A scarcity of resources that makes the split-plot attractive may also result in small sample sizes at the whole-plot level. As a result, some variance components may be poorly estimated. An expected effect of a small sample size is a reduction in the power of the tests for treatment effects, but there may also be an unanticipated effect on the type I error rates corresponding to tests related to the whole-plot treatments. The focus of this report is then to estimate the type I error rates for tests of fixed effects based on a real-life example of a split-plot experiment described in the next section.

Agronomic field study

A field study was conducted over three years beginning in 2015 at the Department of Agronomy Ashland Bottom Research Farm (39.13° N, 96.63° W) near Manhattan, Kansas. The experimental design was a split-plot with the whole-plot treatments assigned to whole plots in an

RCBD containing four blocks. The whole-plot treatments were two harvest intervals (factor HI) set at 28- and 35-days and the subplot treatments were three alfalfa varieties (factor V) which were Hi-Gest 360 (low lignin), Gunner (conventional) and RR Tonnica (roundup ready). The varieties were first planted on April 29th, 2015 (Xu, 2019). Several response variables were measured, but only yield (in metric tons per hectare) is considered here.

The MIXED procedure in SAS 9.4 (SAS Institute Inc., Cary, NC) was used to analyze yield data from 2015 to 2017 by fitting a separate linear mixed model for each year with factors HI, V and HI × V as fixed effects and Block and Block × HI as random effects with Block × HI representing the whole-plot error term. P-values for tests of fixed effects in each year are given in Table 1.1. The interaction HI × V was significant in years 2016 and 2017. The main effect of V was not significant in any year but showed a trend in 2017 ($p = 0.0623$). The main effect of HI was significant in all three years. Estimates of the variance components for block, whole-plot, and subplot errors in each year are given in Table 1.2. Estimated variances based on the alfalfa yield data in 2017 were selected one scenario in the simulation study.

The general form of the linear mixed model

To express the general form of the linear mixed model in mathematical notation, consider a split-plot experiment with whole-plot treatment factor WP with a levels assigned to whole-plots in an RCBD with b blocks of size a and subplot treatment factor SP with c levels assigned to subplots in an RCBD nested within each level of WP. One may express the experiment's model as

$$y_{ijk} = \mu + \alpha_i + b_k + w_{ik} + \gamma_j + (\alpha\gamma)_{ij} + \varepsilon_{ijk}, i = 1,2,\ldots,a, j = 1,2,\ldots,c, k = 1,2,\ldots,b$$

where y_{ijk} is the observed response on the subplot with the *j*th level of SP within the whole-plot in the *k*th block assigned to the *i*th level of WP, b_k denotes the effect of the *k*th block which is assumed to be distributed as $N(0,\sigma_B^2)$, w_{ik} indicates the whole plot error which is assumed to be distributed as $N(0,\sigma_w^2)$, and ε_{ijk} denotes the subplot error which is assumed to be distributed as $N(0,\sigma_\varepsilon^2)$. Additionally, it is assumed that all b_k, w_{ik}, and ε_{ijk} terms are mutually independent. The fixed effects in the model are μ, the overall mean; α_i, the effect of the *i*th level of WP; γ_j, the effect of the *j*th level of SP; and $(\alpha\gamma)_{ij}$, the effect of the interaction between the *i*th level of WP and the *j*th level of SP. From this model, it is easy to see that if one were to include a term in the model corresponding to Block × WP interaction, say $(\alpha b)_{ik}$, it would be completely confounded with the whole-plot error. Therefore, it is common in practice to include Block × WP as a random effect to serve as the denominator of the *F*-statistic for testing for an effect due to WP. The corresponding ANOVA table is given in Table 1.3. Table 1.4 gives the ANOVA shell (sources of variability and their degrees of freedom) along with their corresponding expected means squares. Since the field study has only four blocks and the whole-plot treatment has two levels, it clear to see that the whole-plot error term has only 3 degrees of freedom. It is also clear that there is no mean square in the ANOVA table that can be used as a direct estimator for σ_B^2.

The MoM estimator in the field study would be $\hat{\sigma}_B^2 = (MSE(WP) - MSE(SP))/3$. Unfortunately, this term is not guaranteed to be non-negative. The fixed effects are then estimated as a function of the data and the estimated variance components via generalized least squares. In contrast, REML estimates the variance components and fixed effects parameters in an iterative process. As implemented in SAS, variance components estimated to be negative are set to zero. For further detail, see Gbur et al. (2012), Littell et al. (2006), or Milliken and Johnson (2010).

Kenward-Roger Adjustment

Kenward-Roger (KR) adjustment to the denominator degrees of freedom is widely used to construct tests of the fixed effects in a linear mixed model via the MIXED procedure in SAS. A very high-level description is that the KR computes estimated denominator degrees of freedom by matching the first two moments of an *F*-distribution. (Kenward and Roger, 1997). In practice, this can result in error degrees of freedom much larger than expected, primarily when at least one variance component is estimated to be zero.

Preview

The remainder of this report is structured as follows. In Chapter 2, a simulation study generating split-plot experiments under a wide range of parameter settings where two different methods (REML vs MoM) for estimating the variance components is described. Results of the simulations with respect to type I error rates for tests on the fixed effects (WP Treatment, SP Treatment, and WP × SP Interaction), estimation of the variance components are summarized, and the denominator degrees of freedom of the fixed effects using the KR adjustment are summarized in Chapter 3 with further details given in Appendix A and the corresponding SAS code in Appendix B. Finally, discussion of the results and recommendations for practitioners along with suggestions for future research are given in Chapter 4.

Table 1.1 P-values from mixed model analysis on yield for the effect of harvest intervals, varieties and their interactions.

Effect	Year		
	2015	2016	2017
HI	0.0206*	0.0489*	0.0165*
V	0.3483	0.8430	0.0623
HI×V	0.9845	0.0386*	0.0016**

Table 1.2 Estimated variance components from mixed model analysis for alfalfa yield in 2015, 2016 and 2017.

Variance Parameter	**Year**		
	2015	2016	2017
Block ($\boldsymbol{\sigma_B^2}$)	0.1237	0.1602	0.0745
Whole-Plot ($\boldsymbol{\sigma_w^2}$)	0.07198	0.3432	0.3279
Subplot ($\boldsymbol{\sigma_\varepsilon^2}$)	0.1339	0.4253	0.2073

Table 1.3 Theoretical ANOVA table for split-plot design in randomized completely block.

Source	*df*	Sum of Squares	Mean Squares	F
Block	$b-1$	$SSBlock$	$SSBlock/(b-1)$	$MSBlock/MSE(WP)$
Whole Plot Trt	$a-1$	$SSWP$	$SSWP/(a-1)$	$MSWP/MSE(WP)$
Whole-Plot Error	$(a-1)(b-1)$	$SSE(WP)$	$SSE(WP)/(a-1)(b-1)$	
Subplot Trt	$c-1$	$SSSP$	$SSSP/(c-1)$	$MSSP/MSE(SP)$
Whole Plot Trt × Subplot Trt	$(a-1)(c-1)$	$SSWS$	$SSWS/(a-1)(c-1)$	$MSWS/MSE(SP)$
Subplot Error	$a(b-1)(c-1)$	$SSE(SP)$	$SSE(SP)/a(b-1)(c-1)$	
Total	$abc-1$	SST		

Table 1.4 Analysis of Variance Shell with Expected Mean Squares for the Split-plot Model in the Field Study.

Source	*df*	Expected Mean Square
Block	3	$\sigma_\varepsilon^2 + 3\sigma_w^2 + 6\sigma_B^2$
Whole plot	1	$\sigma_\varepsilon^2 + 3\sigma_w^2 + \phi^2(\alpha)$
Error(Whole plot)	3	$\sigma_\varepsilon^2 + 3\sigma_w^2$
Subplot	2	$\sigma_\varepsilon^2 + \phi^2(\gamma)$
Whole plot × Subplot	3	$\sigma_\varepsilon^2 + \phi^2(\alpha\gamma)$
Error(Subplot)	11	σ_ε^2

Chapter 2 - Materials and Methods

Simulation study

The primary purpose of the simulation is to estimate the effects of three factors on the type I error rate of the tests of fixed effects in a split-plot design similar to the agronomic example described in Chapter 1, i.e. a split-plot with the whole-plot treatments assigned to whole plots in an RCBD. The first factor considered here is the method for estimating the variance components and fixed effects parameters. For the sake of brevity, this work compares only two estimation methods for the variance components, namely REML (the default method used in the MIXED procedure in SAS) and the method of moments based on the type III sums of squares. In SAS these are respectively accomplished with "METHOD = REML" and "METHOD = Type3" options in the Proc MIXED statement. The second factor is the number of blocks in the whole-plot design. This work considers 4, 6 and 8 blocks. Larger block sizes were not considered as they were deemed impractical in a setting with limited resources. Note that because the whole-plot design is an RCBD, the number of blocks is also the number of replicates for each whole-plot treatment. The third and final factor is the particular settings of the variance components. Five different scenarios for the variance components were chosen and are listed in Table 2.1. The first scenario is based on the estimates from the 2017 agronomic field study with values of the variance components set as follows: $\sigma_B^2 = 0.0745$, $\sigma_W^2 = 0.3279$, and $\sigma_\varepsilon^2 = 0.2073$. The second scenario for the variance components keeps the relative magnitude of the three variances but inflated them to $\sigma_B^2 = 0.25$, $\sigma_W^2 = 4$, and $\sigma_\varepsilon^2 = 3.24$. For the third, fourth and fifth scenarios, one dominant variance component was set to 100, and the other two variances were set to 1.

For each combination of block size and variance component scenario (a total of 15 combinations), 10,000 random datasets were generated in SAS where the whole-plot treatment factor (WP) had two levels and the subplot treatment factor (SP) has 3 levels with no difference in treatment means among all 6 treatment combinations. The errors for blocks, whole plots, and subplots were all assumed to be independent normal random variables with their assigned variances depending on the given scenario (Table 2.1). For each dataset, a linear mixed model with WP, SP and WP × SP as fixed effects and random effects Block and Block × WP was fit twice, once using REML and once using method of moments based on the type III sums of squares. The denominator degrees of freedom were adjusted using the DDFM = KR option in the model statement for all analyses. P-values for the tests of fixed effects and their denominator degrees of freedom as well as the estimates of the variance components were retained for each fitted model. Since the data were generated with no effects for WP, SP or WP × SP, each p-value smaller than the significance level of $\alpha = 0.05$ represents a type I error. An empirical type I error rate, $\hat{\alpha}$, was computed for each fixed effect as the number of significant p-values (i.e. $p \leq 0.05$) out of 10,000 for each test of a fixed effect. Confidence bands for each empirical type I error rate were computed based on a 95% confidence interval for a proportion, i.e. $\pm 1.96\sqrt{0.05 \times 0.95/10{,}000}$. These results are summarized in Chapter 3 and additional tables and figures are given in Appendix A. The SAS code for the entire simulation study is given in Appendix B.

Table 2.1 Settings of the variance components in the simulation study.

Scenario	σ_B^2	σ_w^2	σ_ε^2
1	0.0745	0.3279	0.2073
2	0.25	4	3.24
3	100	1	1
4	1	100	1
5	1	1	100

Chapter 3 - Results

Type I error rate

Type I error was calculated for each test of fixed effects under both REML and MoM where the denominator degrees of freedom were estimated using the Kenward-Roger method. MoM yielded estimated of type I error rates across all scenarios within the expected Monte Carlo error; however, the estimated type I error rates of three fixed effects under REML were inconsistent. For the empirical variance component from the field study, type I error rates for the whole plot effect were anti-conservative, lying outside the 95% confidence interval corresponding to 0.05 significance when block size was four while results for the other two fixed effects lay in the interval. For larger numbers of blocks, type I error rates for the whole plot effect were less liberal compared to four blocks. Type I error rates of the other two fixed effects stayed consistent across three block sizes (Fig. 3.1). When the magnitude of the variance component was inflated, patterns of type I error rates of three fixed effects were similar to results of empirical variance component across three block sizes. However, for inflated variance using six and eight blocks, type I error rates of the whole plot effect were more liberal compared to the empirical variance component (Fig. 3.1).

When one of the three variance components dominated the others, the estimation of type I error rates of three fixed effects showed different patterns across three scenarios. When block variance dominated, REML and MoM yielded a similar type I error rates, which all fell in the confidence interval regardless of block sizes (Fig. 3.2). When whole plot error dominated the variance components, REML yielded type I error rates similar to the empirical and inflated variance scenarios, type I error rates of whole plot effect were liberal (Fig. 3.3).

REML yielded the poorest estimation of type I error rates for three fixed effects with dominant residual variance (Fig. 3.4). Whole plot effect was too conservative and not sensitive to block sizes. Overestimation of type I error rates for the whole plot and subplot interaction and subplot effects occurred when residual variance dominated, however, increasing block sizes could alleviate the bias.

Histograms of estimated variance components

Histograms were used to evaluate the estimated variance components. MoM models produced good variance estimation of block and residual regardless of the block size (Fig. A.1 to A.15 and Table 3.1).

REML models improved the estimation of variance components with more blocks. When the variance of the block dominated estimated variance components, across three block sizes, REML and Type III models had a similar performance of estimating the three variance components. When the variance of whole plot error dominated, block variance estimates were more biased and residual variance estimates were less biased when block size was equal to four. When the variance of subplot dominated, block and whole plot error variance estimates were less biased with more blocks.

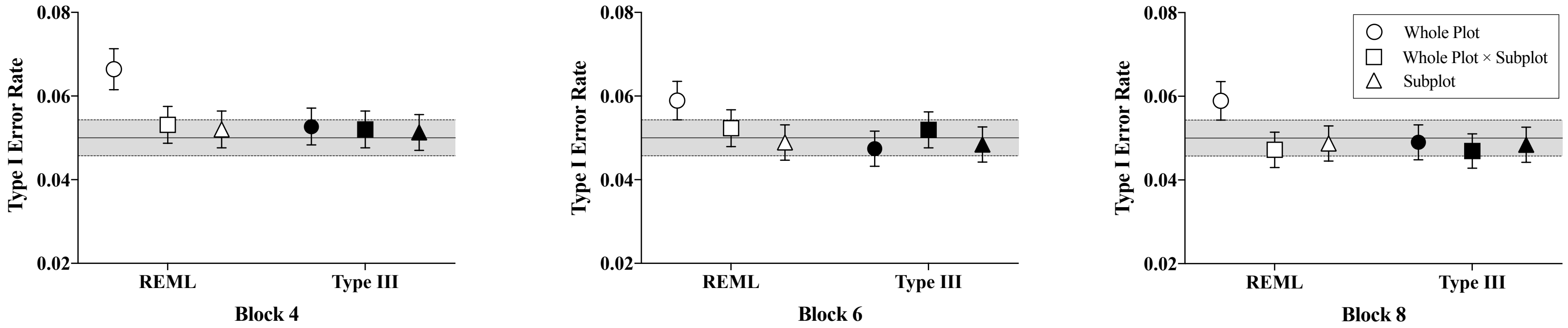
Frequency counts of denominator degrees of freedom

In SAS PROC MIXED, the Kenward-Roger method of adjusting the denominator degrees of freedom could not always fix the issue of poorly estimated variance components. Different proportions of denominator degrees of freedom (DDF) were inflated because KR with REML would return zero for all negative estimators.

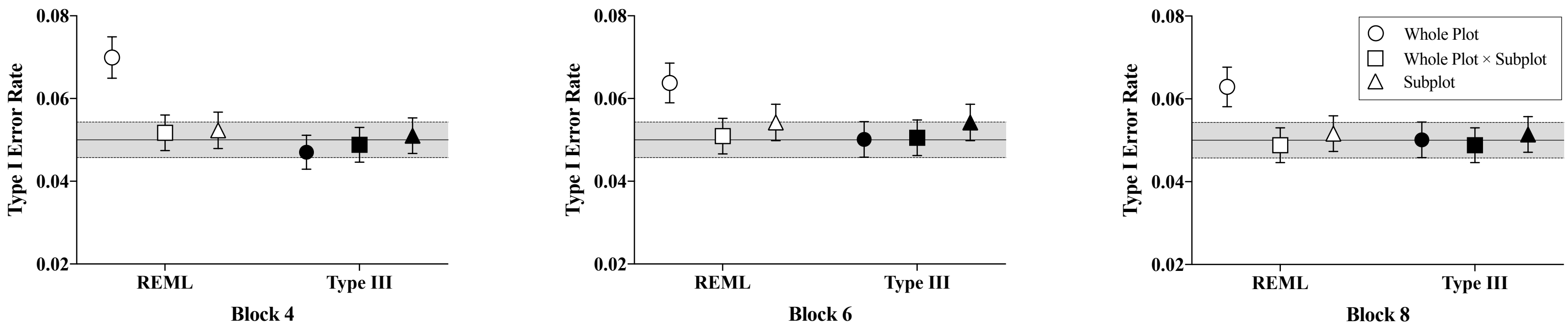
Frequency counts of DDF were used in the *F*-test of three fixed effects (Table 3.2 and 3.3). MoM models assigned correct DDF across simulation scenarios. Inversely, REML models assigned inflated DDF to parts of the simulations.

For empirical variance component (Table 3.2), approximately 51% DDF for the whole plot were overestimated under four blocks, while more 90% DDF for the other two fixed effects were normal. Increasing block sizes to six and eight, 61% and 64% DDF for the whole plot were correctly estimated, respectively. However, fewer correct estimations of DDF occurred under the inflated variance component scenario. It was observed that 42% of adjusted DDF under REML for the whole plot was correct when block size was four, 50% and 53% were correct for six and eight blocks, respectively.

For scenarios with one dominant variance component, the inconsistency of fixed effects DDF could be observed using REML across different scenarios (Table 3.3). When the block variance dominated, the estimation of DDF yielded the best result at 85%, 94% and 96% DDF for three fixed effects were correct at four, six and eight blocks, respectively. When whole plot error dominated, only 50% of DDF for whole plot effect was correct regardless of block sizes, while correct numbers of DDF for the other two fixed effect were 99%, 95% and 100% for four, six and eight blocks, respectively. Residual variance played an essential role in estimating DDF for three fixed effects (Table 3.3). When block size was four, 11%, 34% and 34% of DDF were correct for the whole plot, interaction, and subplot effects, respectively. Adding block sizes improved the correctness of DDF for three fixed effects slightly.



(a) Simulations with the realistic covariance parameters



(b) Simulations with the inflated realistic covariance parameters

Figure 3.1 Simulation with the realistic and inflated realistic variance parameters.

(a). True value of variance components is 0.0745 for block, 0.3279 for block × whole plot, and 0.2073 for residuals;

(b). True value of variance components is 0.25 for block, 4 for block × whole plot, and 3.24 for residuals;

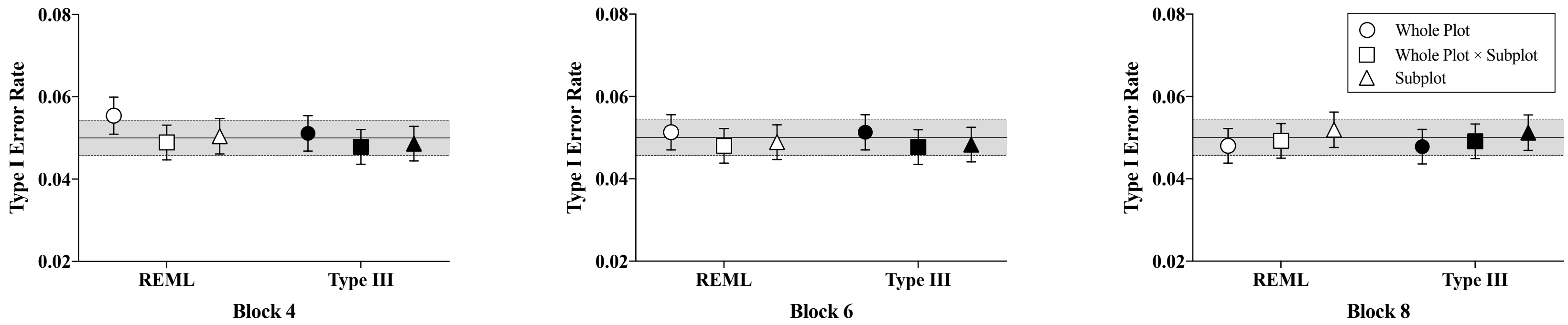


Figure 3.2 Simulations with variance components dominated by block effect.

True value of variance components is 100 for block, 1 for block × whole plot, and 1 for residuals.

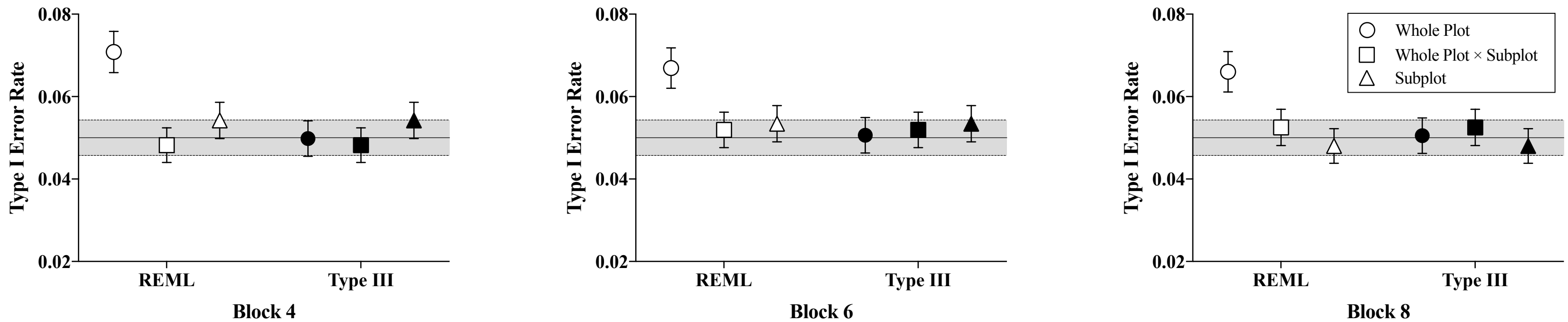


Figure 3.3 Simulations with variance components dominated by whole plot error.

True value of variance components is 1 for block, 100 for block × whole plot, and 1 for residuals.

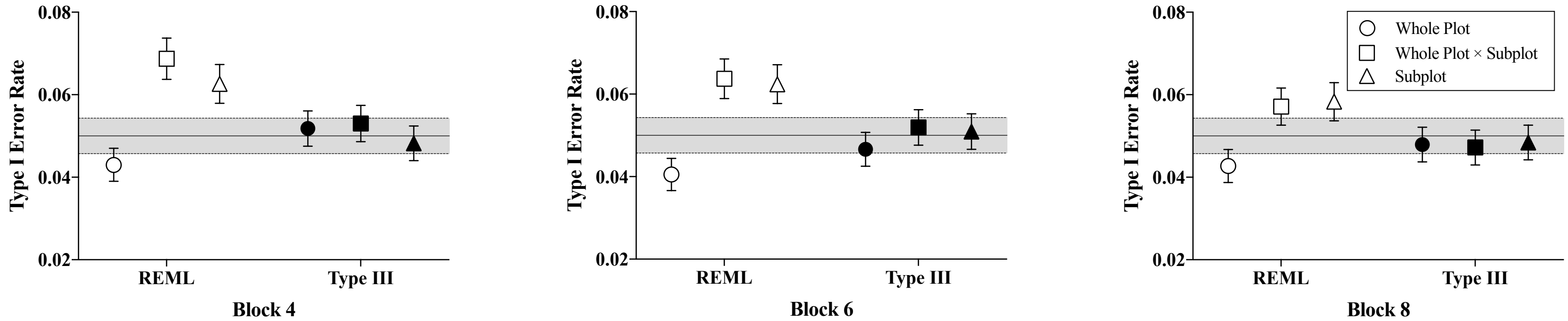


Figure 3.4 Simulations with variance components dominated by subplot error.

True value of variance components is 1 for block, 1 for block × whole plot, and 100 for residuals.

Table 3.1 Variance parameter estimation of different simulation scenarios.

		Variance Parameter Estimate		
Block Size	Variance	Block	Block × WP	Residual
	TRUE	0.0745	0.3279	0.2073
4	REML	0.1384	0.2623	0.2052
	Type III	0.0712	0.3272	0.2074
6	REML	0.1192	0.2801	0.2059
	Type III	0.0722	0.3267	0.2064
8	REML	0.1093	0.2927	0.2071
	Type III	0.0708	0.3309	0.2072
	TRUE	0.25	4	3.24
4	REML	1.2485	3.0568	3.1918
	Type III	0.2618	3.9985	3.2369
6	REML	1.0420	3.2033	3.2143
	Type III	0.2873	3.9451	3.2273
8	REML	0.8804	3.3712	3.2278
	Type III	0.2253	4.0219	3.2323
	TRUE	100	1	1
4	REML	101.3127	1.0329	0.9840
	Type III	101.3230	1.0082	0.9991
6	REML	99.4537	0.9995	0.9971
	Type III	99.4574	0.9912	1.0022
8	REML	99.9711	1.0029	0.9977
	Type III	99.9729	0.9991	0.9999
	TRUE	1	100	1
4	REML	22.1591	78.5182	0.9921
	Type III	1.5128	99.1704	0.9921
6	REML	17.5205	83.0848	0.9966
	Type III	1.0953	99.5122	0.9966
8	REML	15.5979	85.7916	0.9983
	Type III	1.4212	99.9706	0.9983
	TRUE	1	1	100
4	REML	5.4560	6.4849	89.3672
	Type III	0.9571	0.6042	99.7474
6	REML	4.4079	5.5592	92.0358
	Type III	1.0985	1.0999	99.8050
8	REML	3.6877	4.7753	93.5182
	Type III	0.8551	0.9888	100.1377

Table 3.2 Frequency counts of fixed effects DDF of realistic and inflated realistic variance components.

Block Size	Method	Whole Plot		Whole Plot × Subplot		Subplot	
		DDF	Counts	DDF	Counts	DDF	Counts
		Var(blk) = 0.0745, Var(blk*WP) = 0.3279, Var(Residual) = 0.2073					
4	REML	3	5120	12	9051	12	9051
		6	3931	15	835	15	835
		15	835	18	114	18	114
		18	114				
	Type III	3	10000	12	10000	12	10000
6	REML	5	6080	20	9709	20	9709
		10	3629	25	277	25	277
		25	277	30	14	30	14
		30	14				
	Type III	5	10000	20	10000	20	10000
8	REML	7	6394	28	9896	28	9896
		14	3502	35	102	35	102
		35	102	42	2	42	2
		42	2				
	Type III	7	10000	28	10000	28	10000
		Var(blk) = 0.25, Var(blk*WP) = 4, Var(Residual) = 3.24					
4	REML	3	4218	12	8820	12	8820
		6	4584	15	970	15	970
		15	970	18	228	18	228
		18	228				
	Type III	3	10000	12	10000	12	10000
6	REML	5	5005	20	9485	20	9485
		10	4480	25	475	25	475
		25	475	30	40	30	40
		30	40				
	Type III	5	10000	20	10000	20	10000
8	REML	7	5253	28	9780	28	9780
		14	4527	35	207	35	207
		35	207	42	13	42	13
		42	13				
	Type III	7	10000	28	10000	28	10000

Table 3.3 Frequency counts of fixed effects DDF with one dominant variance component

Block Size	Method	Whole Plot		Whole Plot × Subplot		Subplot	
		DDF	Counts	DDF	Counts	DDF	Counts
		Var(blk) = 100, Var(blk*WP) = 1, Var(Residual) = 1					
4	REML	3	8525	12	8536	12	8536
		6	11	15	1464	15	1464
		15	1464				
	Type III	3	10000	12	10000	12	10000
6	REML	5	9354	20	9355	20	9355
		10	1	25	645	25	645
		25	645				
	Type III	5	10000	20	10000	20	10000
8	REML	7	9653	28	9653	28	9653
		35	347	35	347	35	347
	Type III	7	10000	28	10000	28	10000
		Var(blk) = 1, Var(blk*WP) = 100, Var(Residual) = 1					
4	REML	3	5057	12	9998	12	9998
		6	4941	15	2	15	2
		15	2				
	Type III	3	10000	12	10000	12	10000
6	REML	5	5005	20	9485	20	9485
		10	4480	25	475	25	475
		25	475	30	40	30	40
		30	40				
	Type III	5	10000	20	10000	20	10000
8	REML	7	4828	28	10000	28	10000
		14	5172				
	Type III	7	10000	28	10000	28	10000
		Var(blk) = 1, Var(blk*WP) = 1, Var(Residual) = 100					
4	REML	3	1142	12	3443	12	3443
		6	2301	15	2607	15	2607
		15	2607	18	3950	18	3950
		18	3950				
	Type III	3	10000	12	10000	12	10000
6	REML	5	1344	20	3799	20	3799
		10	2455	25	2554	25	2554
		25	3647	30	3647	30	3647
		30	3647				
	Type III	5	10000	20	10000	20	10000
8	REML	7	1397	28	3898	28	3898
		14	2501	35	2544	35	2544
		35	2544	42	3558	42	3558
		42	3558				
	Type III	7	10000	28	10000	28	10000

Chapter 4 - Discussion and Conclusions

We conducted the simulation study by using REML and the method of moments based on type III sums of squares. For studies with balanced data, REML estimators should be similar to the ANOVA estimators given adequate sample sizes (Corbeil and Searle, 1976). However, REML variance component estimates produced inflated test statistics and hence type I error rates of the fixed effect approximately 1.5 times the normal significance of 0.05. REML might not be an appropriate procedure when negative or zero variance component estimates occur, especially in conjunction with the Kenward-Roger adjustment to the denominator degrees of freedom. REML with the NOBOUND option fixed the issue of messy denominator degrees of freedom of the fixed effect (results not shown here), which returned similar results to Type III MoM models. One should also note that REML has the advantage of being able to deal with complicated designs that cannot be solved using Type III methods.

When residual variance dominated other variance components, REML not only inflated type I error rates for tests of SP and WP × SP effects, but also deflated type I error rates for the WP effect. While type I error rates for the other two fixed effects were usually normal in other scenarios, type I error rates of whole plot effect were generally inflated except for the case of a dominant subplot variance component.

The simulation study performed here also showed that negative estimates of variance component could affect inference on the fixed effects in mixed models as noted by Stroup and Littell (2002). Type I error rates of fixed effects that were affected were also associated with strange estimates of the denominator degrees of freedom of those effects, which were inflated due to setting the estimate to zero by REML.

This study has considered one configuration of balanced data from a split-plot design with three random effect variance components. However, the implications of negative variance component estimates in different scenarios indicate uncertainty with respect to the behavior of statistical tests under other conditions. For a more complex model or unbalanced data, several issues remain to be addressed.

REML solves the bias asymptotically for Gaussian linear mixed models, and for balanced data with normality. REML also is known to be identical to ANOVA estimators with optimal minimum variance properties (Searle et al., 1992). Other than REML, future simulations might be based on minimum variance quadratic unbiased estimators (MIVQUE), or Bayes estimators as alternative methods for estimating variance components (Milliken and Johnson, 2010; Searle et al., 1992).

References

Corbeil, R.R., Searle, S.R., 1976. Restricted maximum likelihood (reml) estimation of variance components in the mixed model. Technometrics. doi:10.1080/00401706.1976.10489397

Gbur, E.E., Stroup, W.W., McCarter, K.S., Durham, S., Young, L.J., Christman, M., West, M., Kramer, M., 2012. Analysis of Generalized Linear Mixed Models in the Agricultural and Natural Resources Sciences. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI. doi:10.2134/2012.generalized-linear-mixed-models.frontmatter

Fischer, R.A., 1925. Statistical methods for research workers, 1925. Edinburgh Oliver Boyd.

Jones, B., Nachtsheim, C.J., 2009. Split-plot designs: What, why, and how. J. Qual. Technol. doi:10.1080/00224065.2009.11917790

Kenward, M.G., Roger, J.H., 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O., 2006. SAS for mixed models. SAS institute Cary, NC.

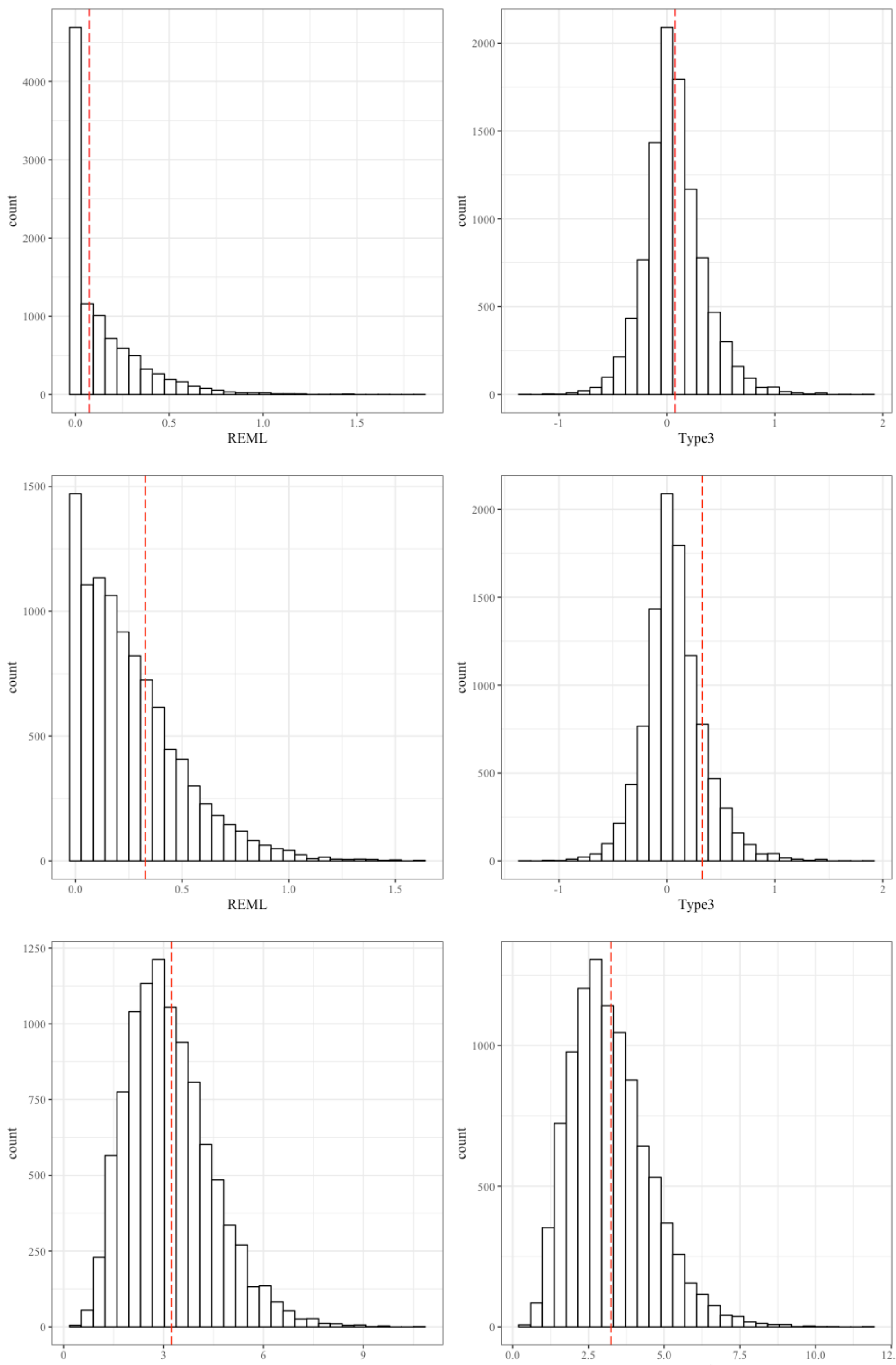
Milliken, G., Johnson, D., 2010. Analysis of Messy Data Volume 1, Analysis of Messy Data Volume 1. doi:10.1201/ebk1584883340

Searle, S.R., Casella, G., McCulloch, C.E., 1992. Variance Components, Book.

Stroup, W., Littell, R., 2002. Impact of variance component estimates on fixed effect inference in unbalanced linear mixed models. Conf. Appl. Stat. Agric. doi:10.4148/2475-7772.1198

Xu, X., 2019. Low lignin alfalfa cutting management study (Doctoral dissertation).

Appendix A - Supplementary Figures



(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.1 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components. (4 blocks, real variances)

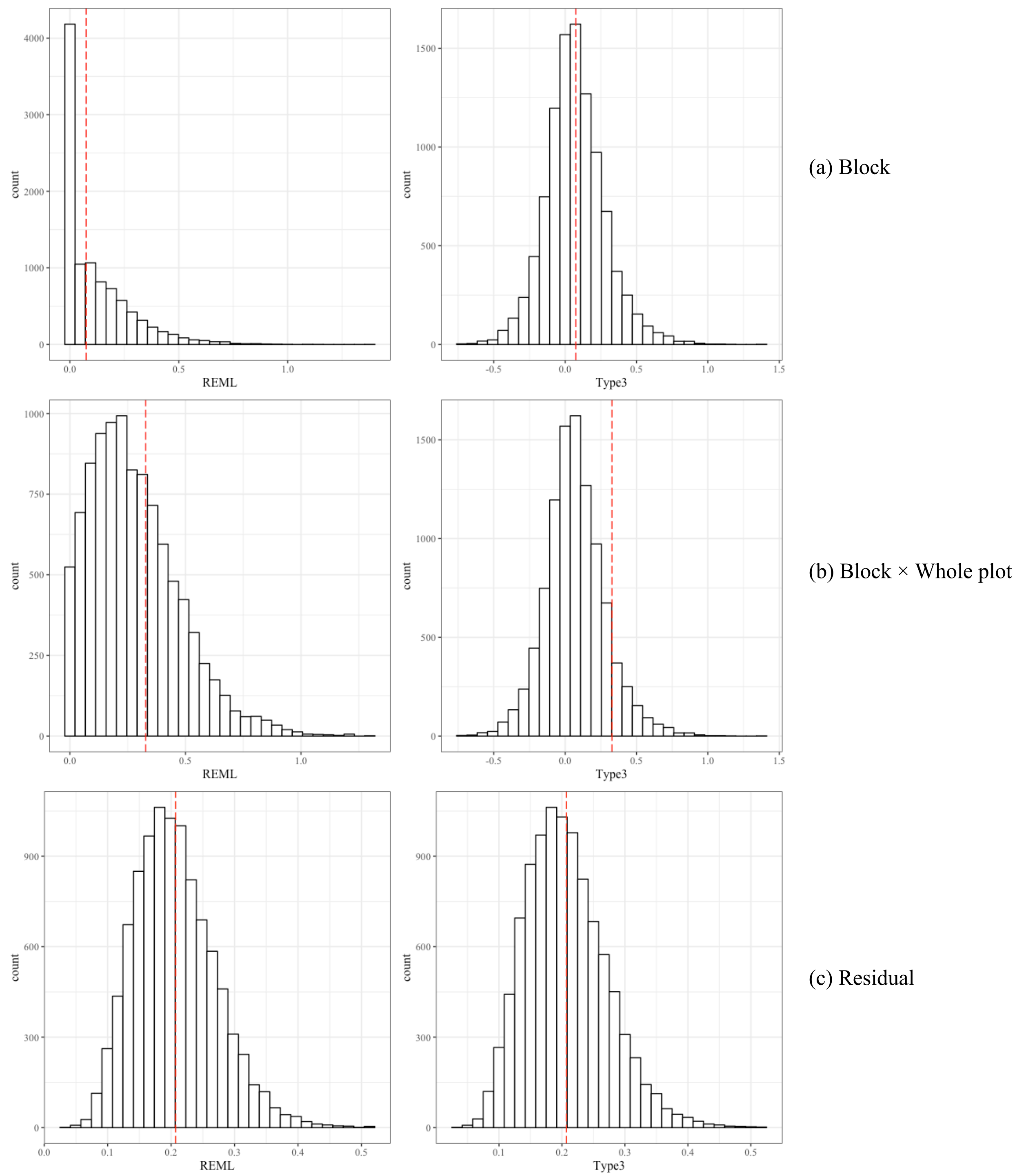


Figure A.2. Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (6 blocks, real variances)

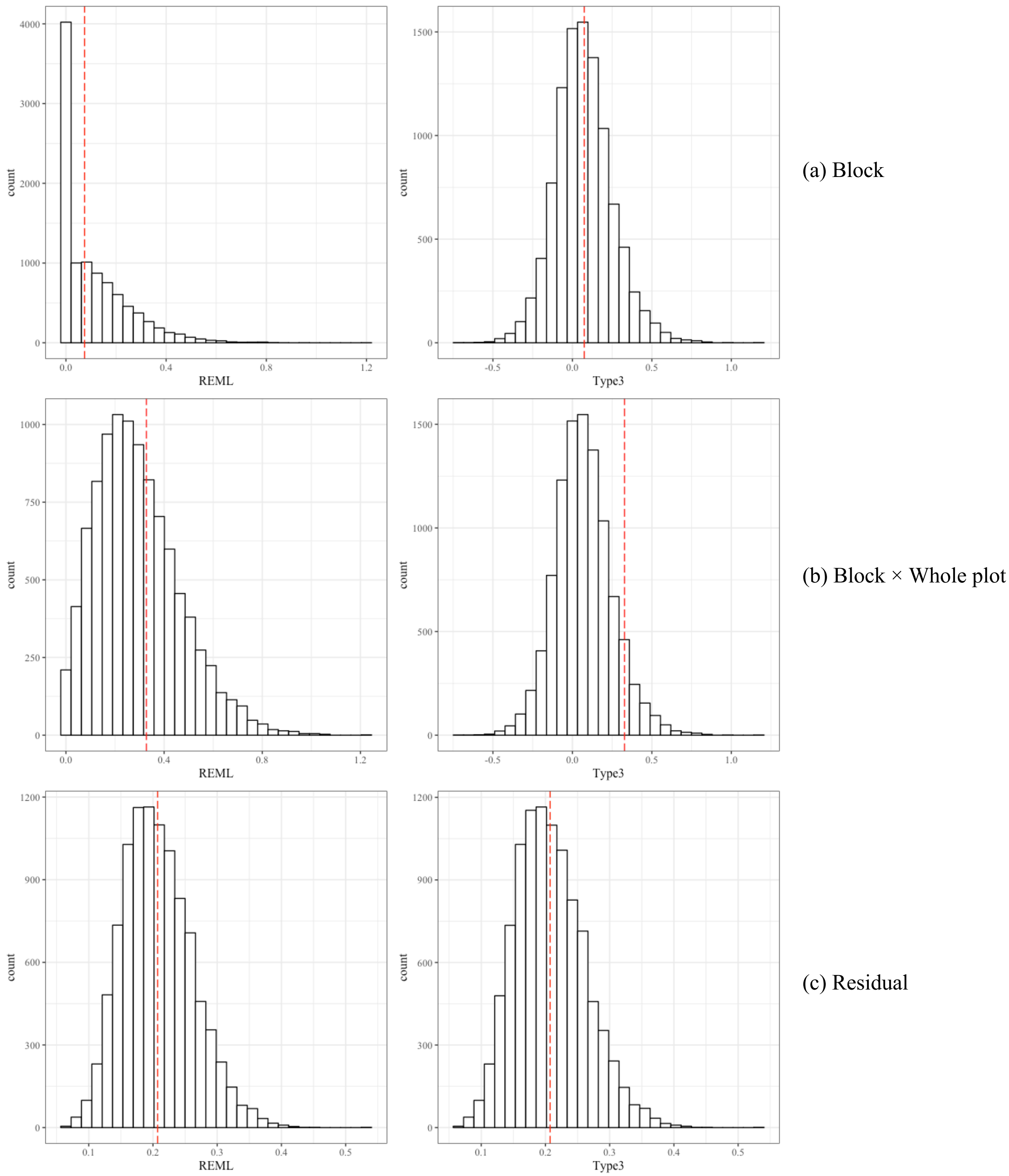


Figure A.3 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (8 blocks, real variances).

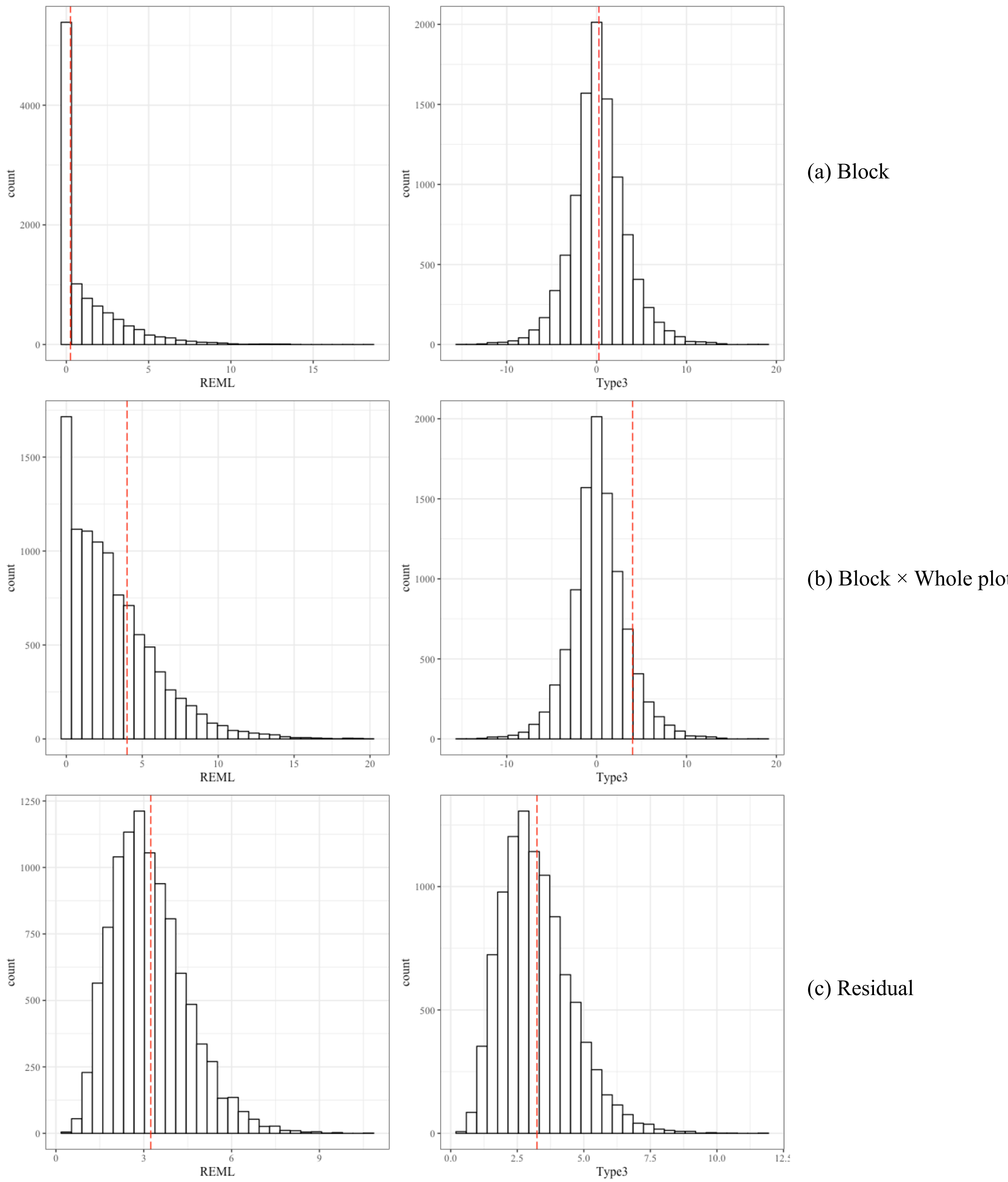
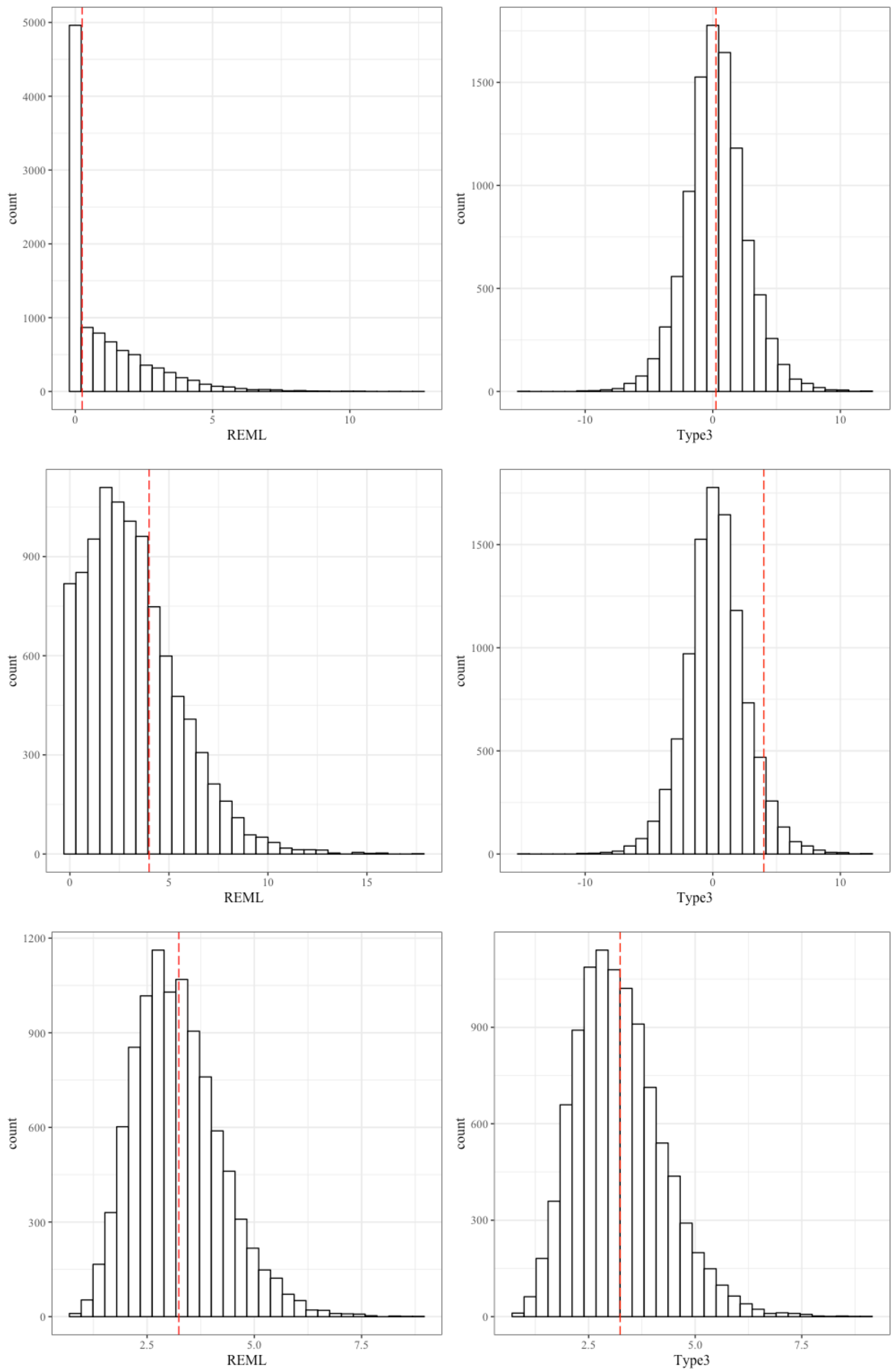


Figure A.4 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (4 blocks, inflated variances).

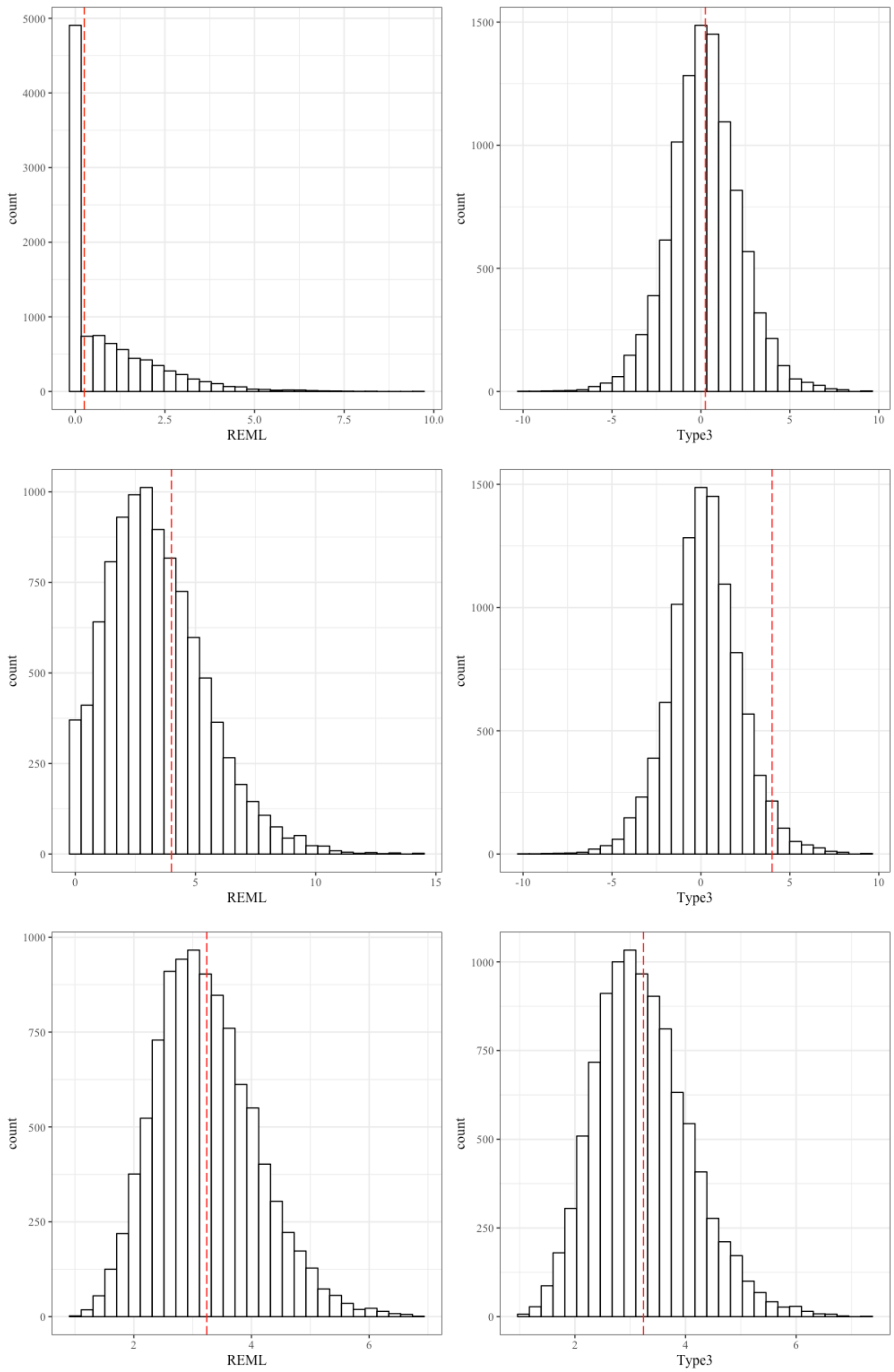


(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.5 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (6 blocks, inflated variances).

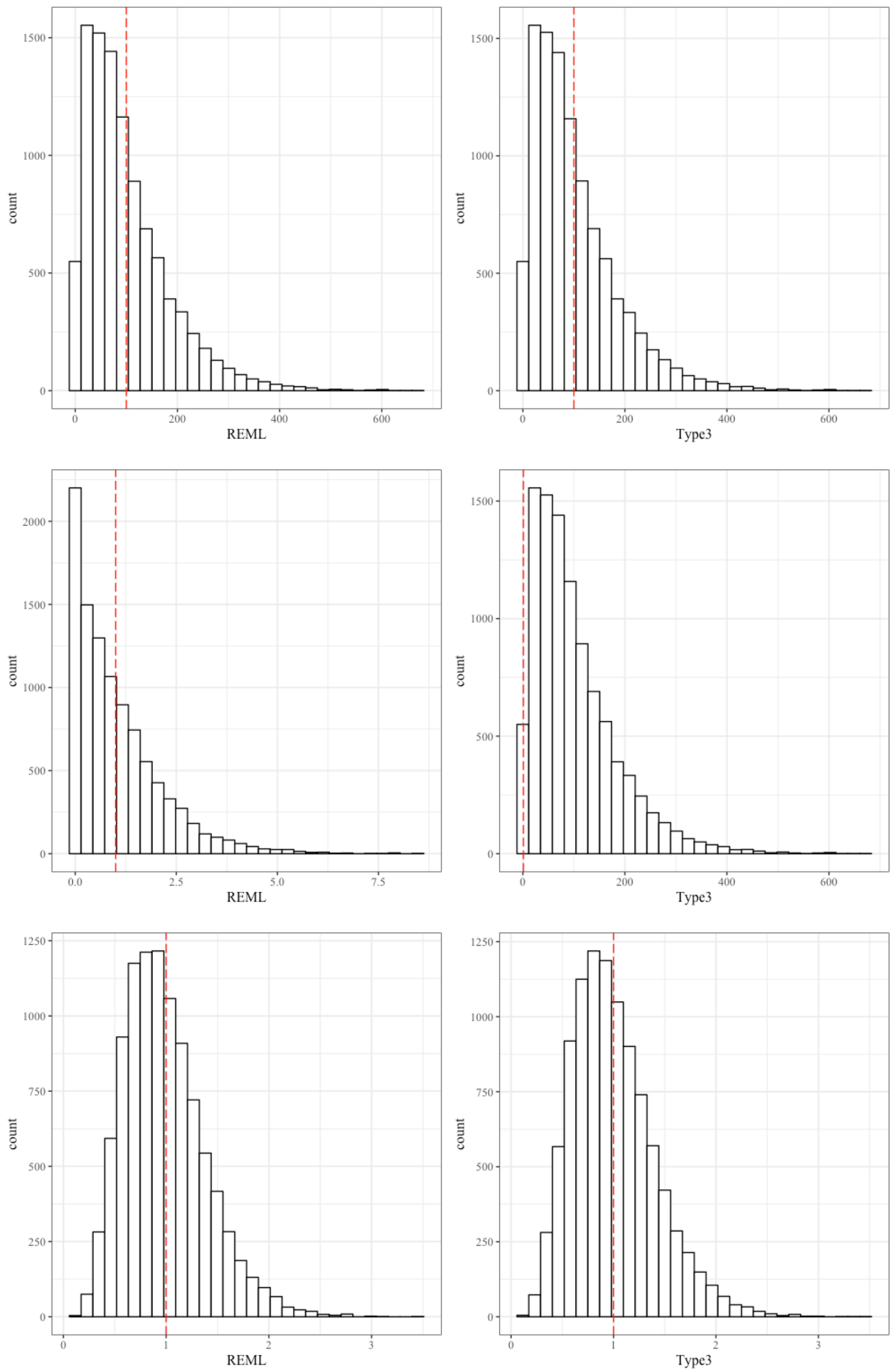


(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.6 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (8 blocks, inflated variances).

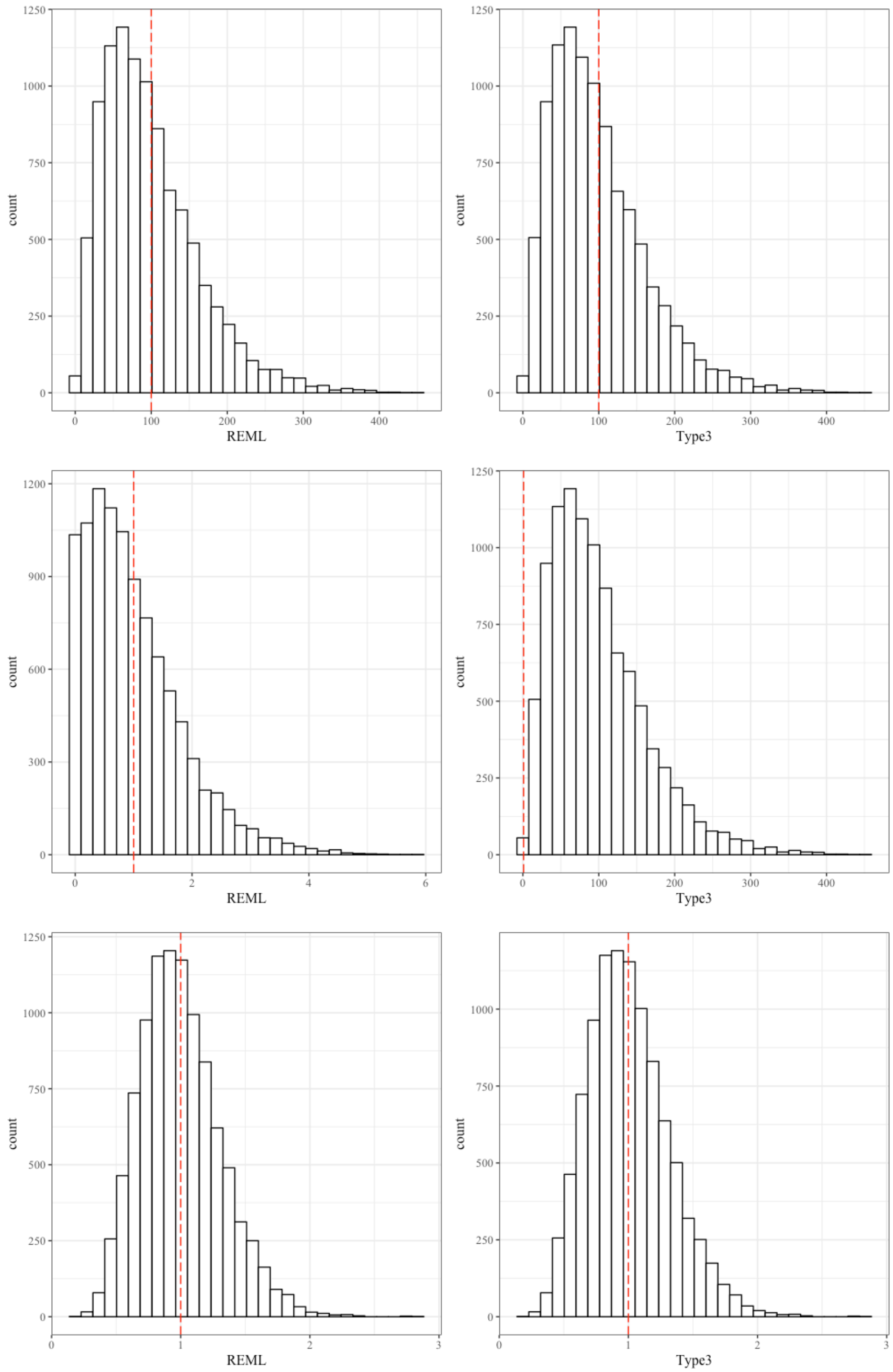


(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.7 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (4 blocks, block dominated).

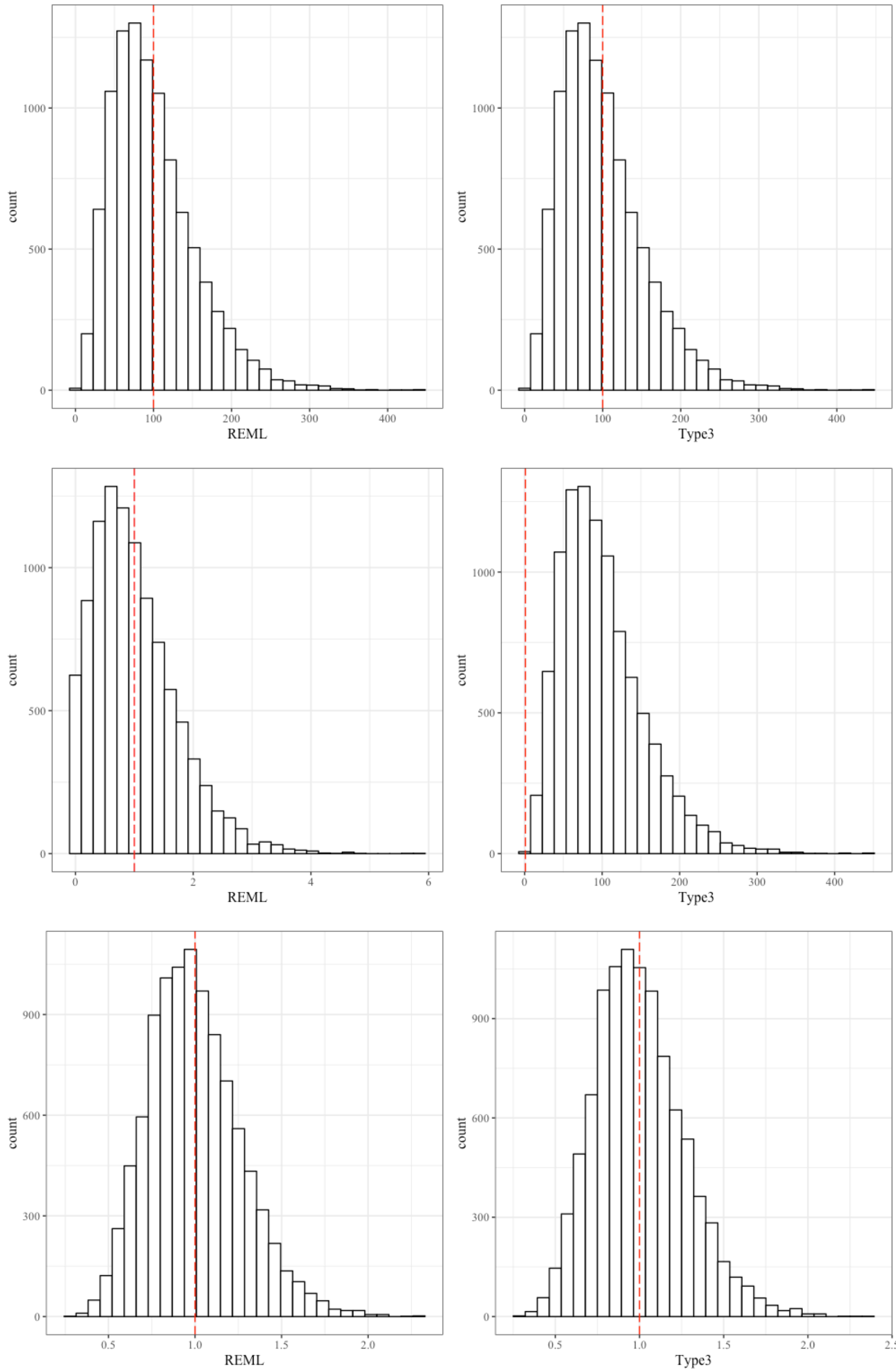


(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.8 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (6 blocks, block dominated).



(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.9 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (8 blocks, block dominated).

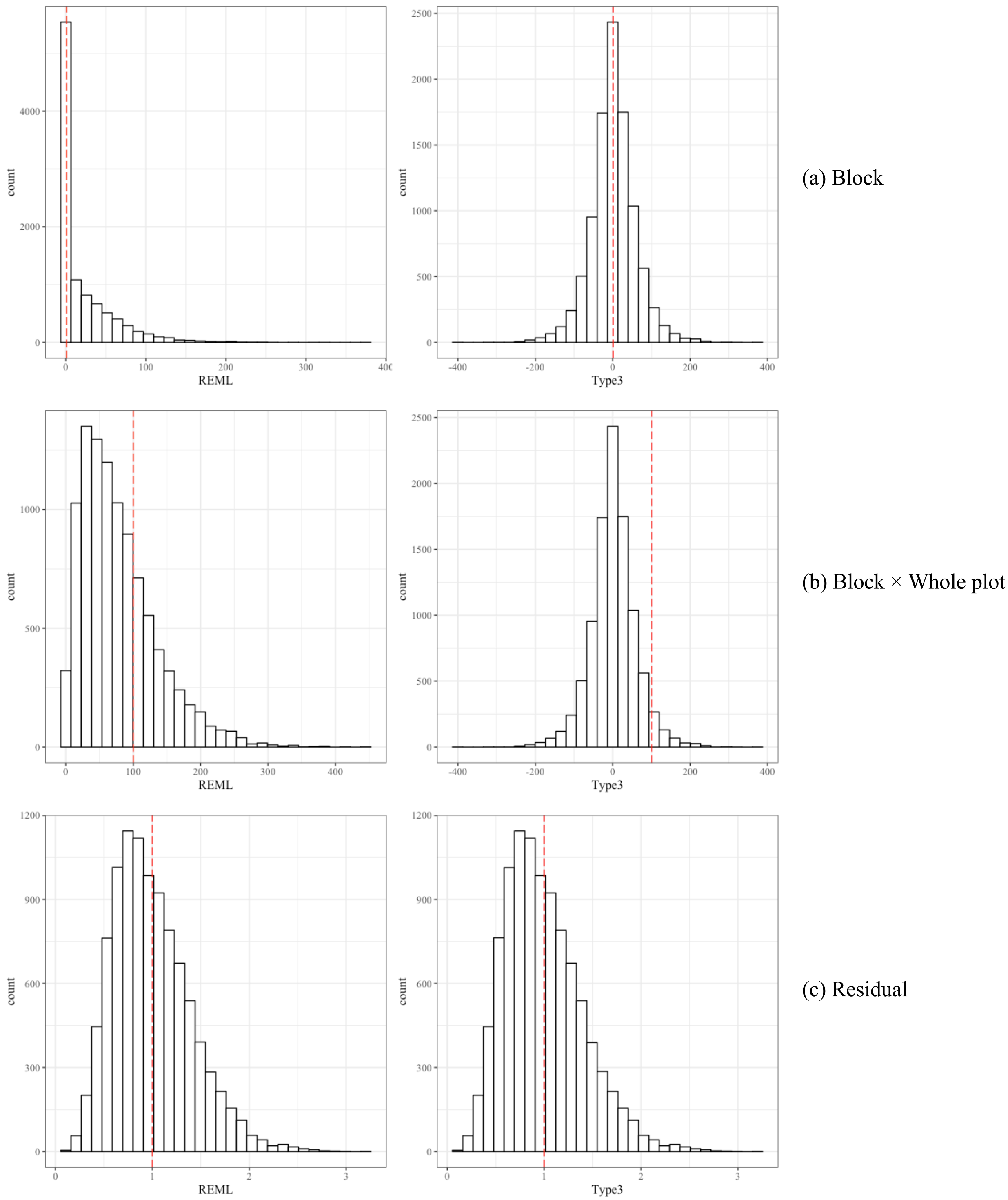
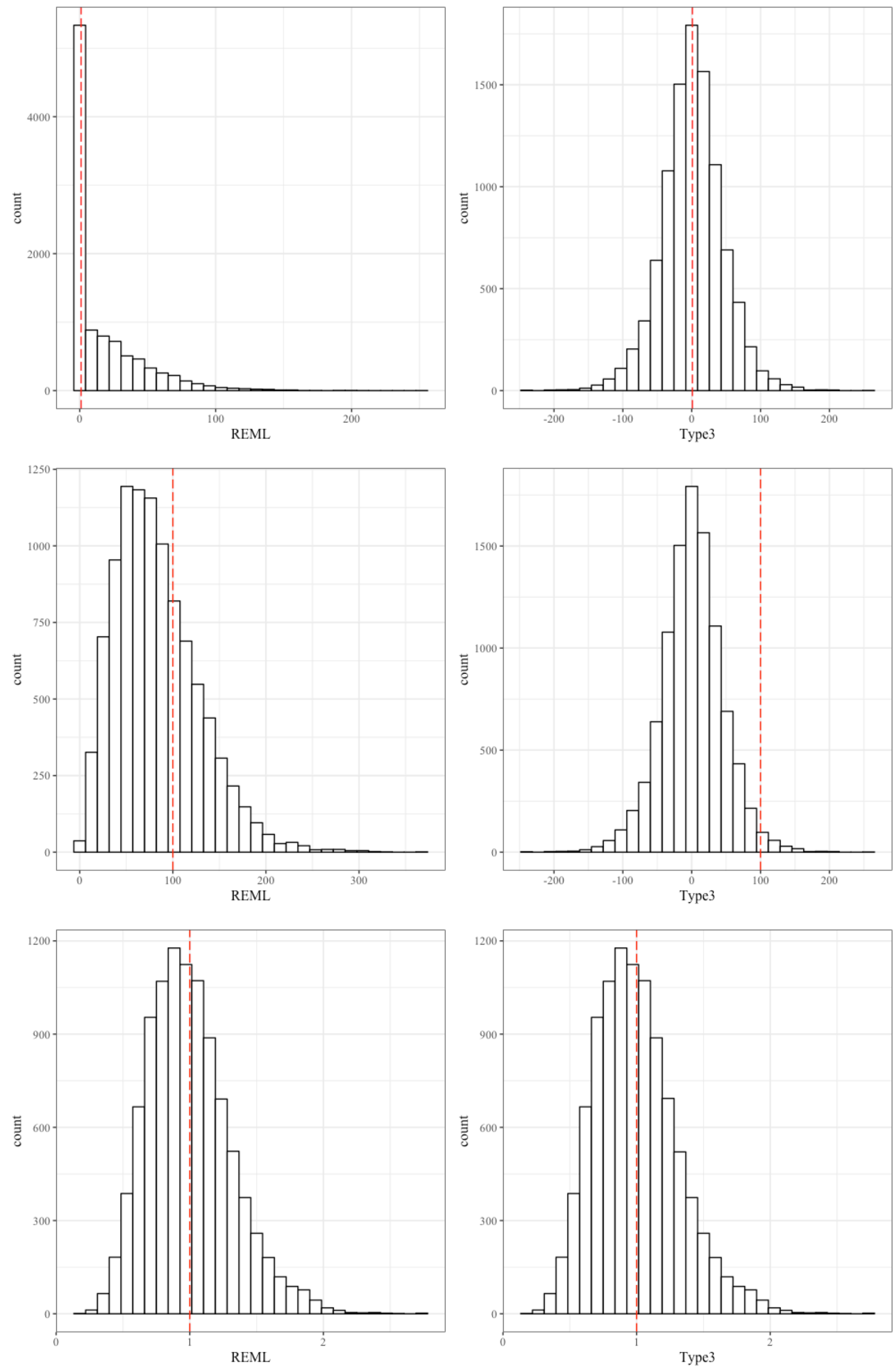


Figure A.10 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (4 blocks, Blk×WP dominated).



(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.11 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (6 blocks, Blk×WP dominated).

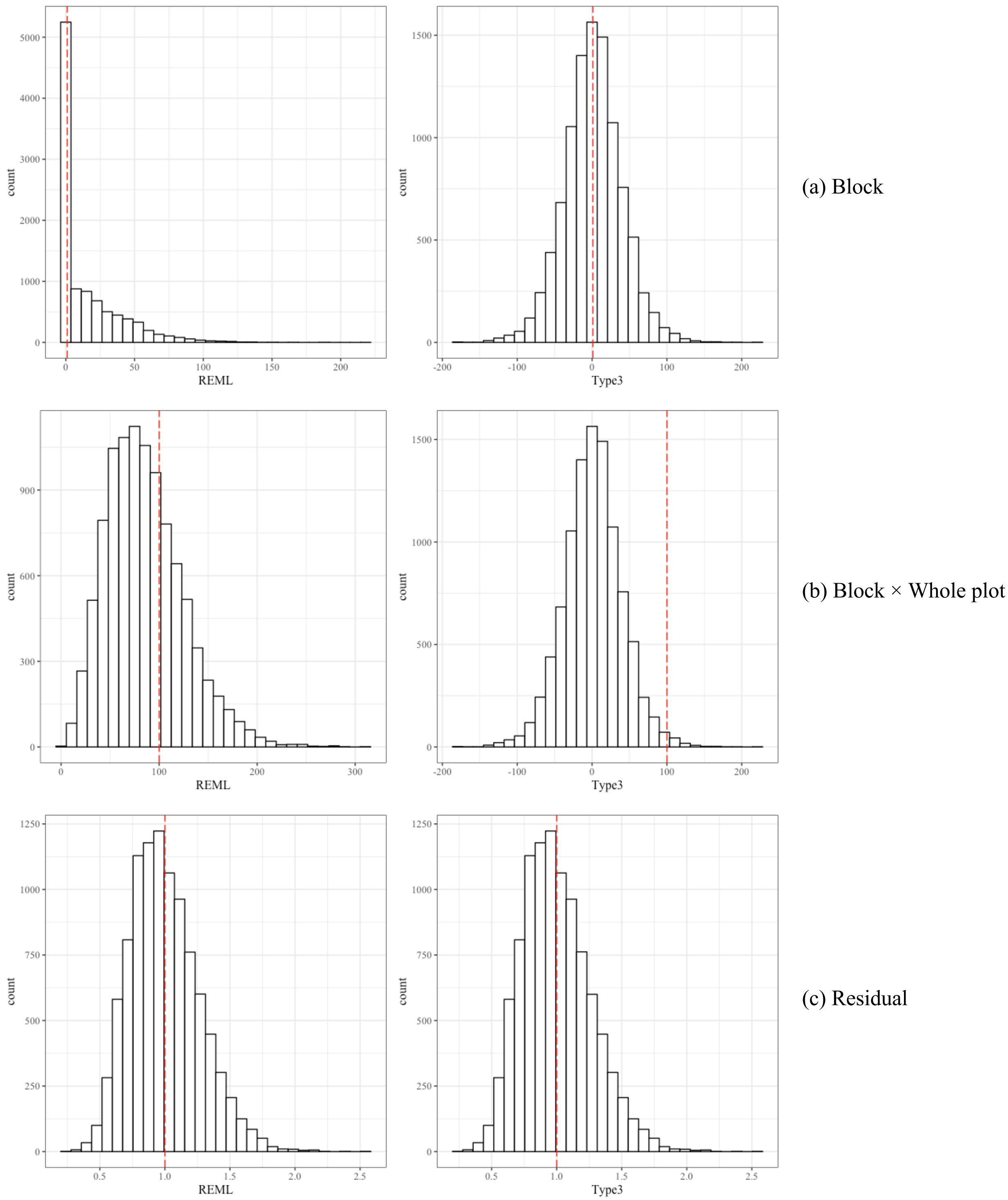


Figure A.12 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (8 blocks, Blk×WP dominated).

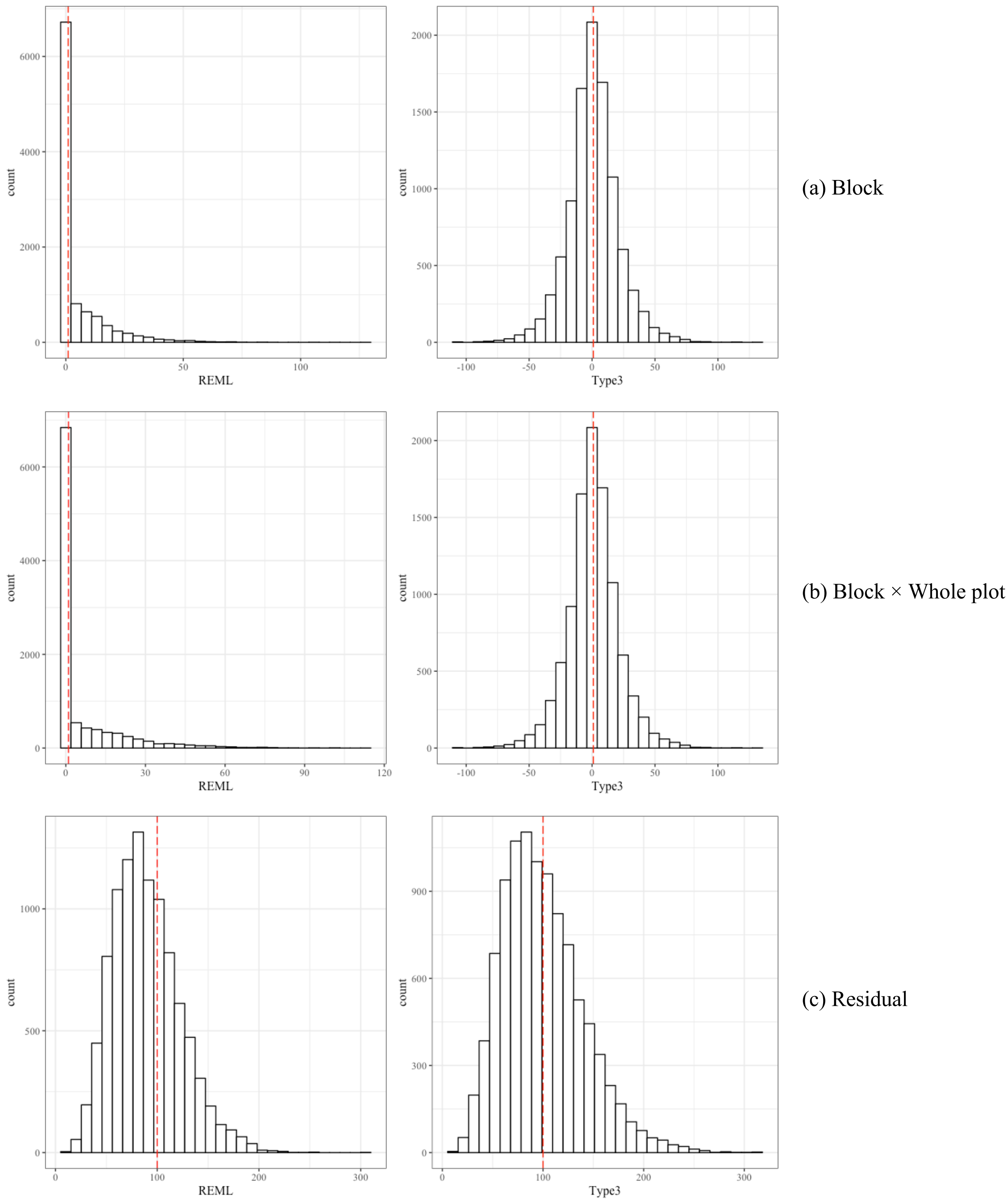
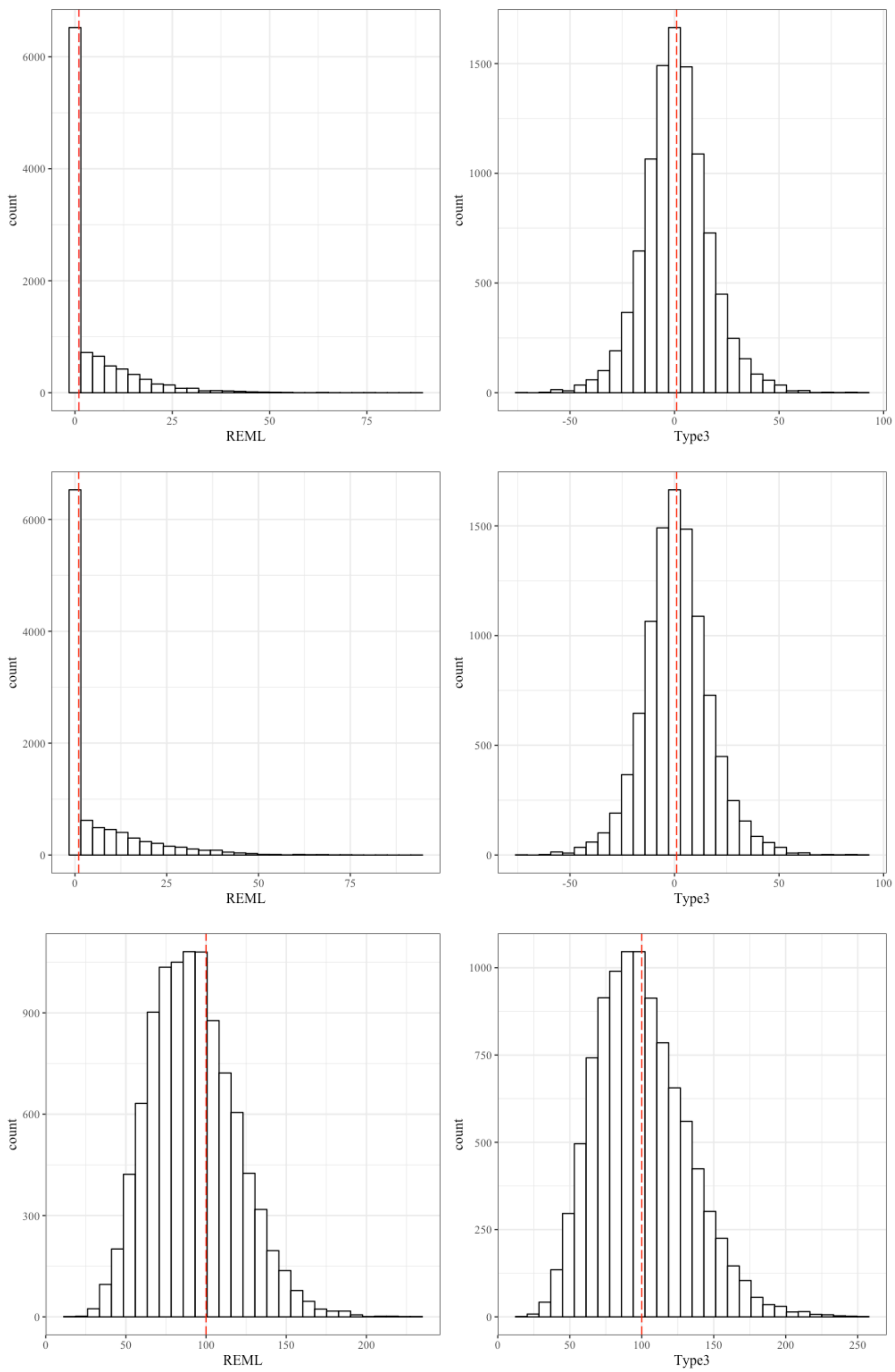


Figure A.13 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (4 blocks, residual dominated).

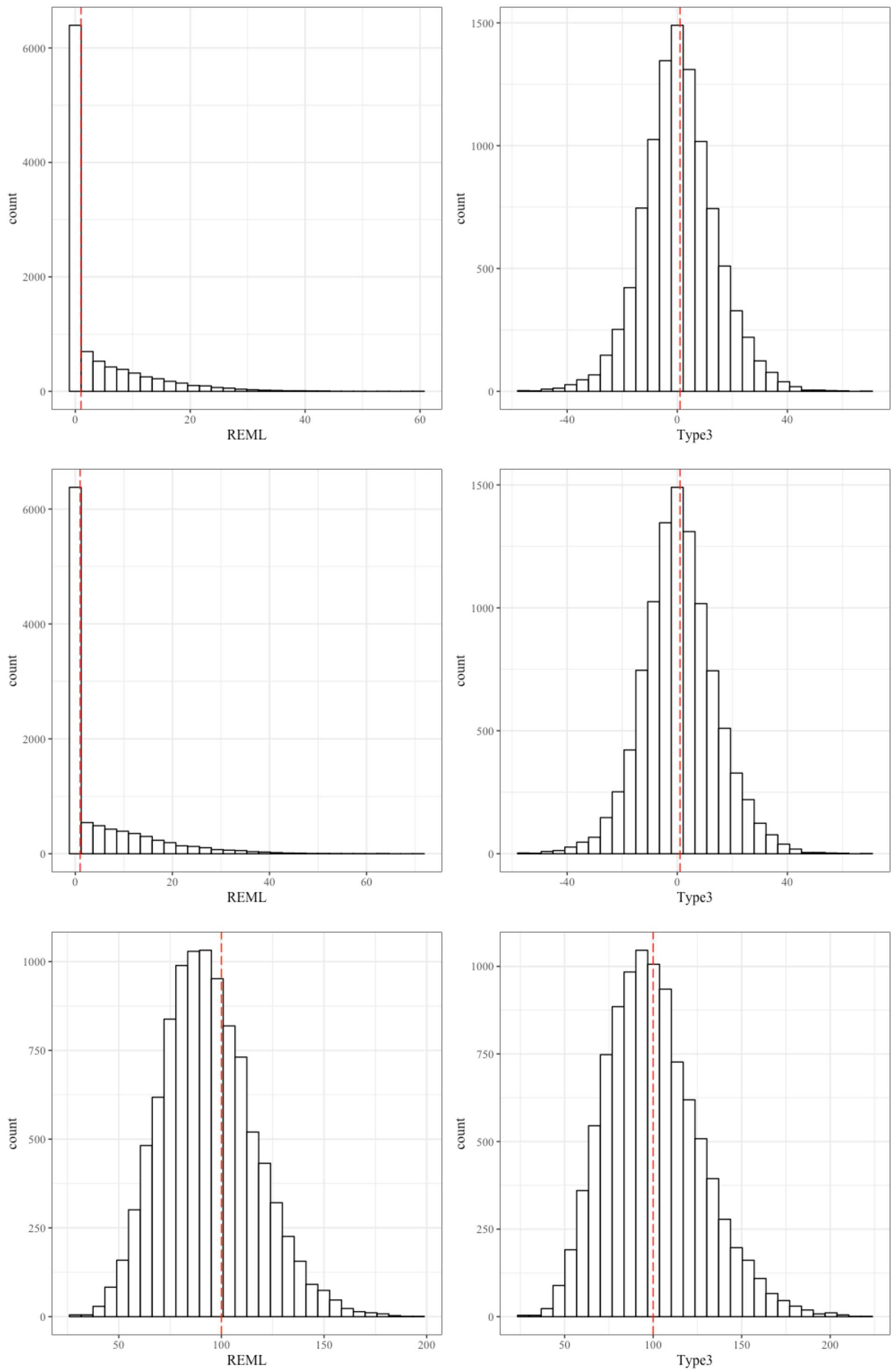


(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.14 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (6 blocks, residual dominated).



(a) Block

(b) Block × Whole plot

(c) Residual

Figure A.15 Histograms of the average estimates of variance components for block, block × whole plot, and residuals. Red lines indicate true values of variance components (8 blocks, residual dominated).

Appendix B - SAS Code

```
proc import datafile = "C:\Users\161-agrlab\Downloads\XVSC_17.csv"
                out = XVSC
                dbms = csv
                replace;
run;
/*---simulation noise sturcture from the XVSC data, IML is used in order to
accomodate the negative variance component---*/
%let std_bk=10;
%let std_bel=1;
%let std_resid=1;
/*---Sim 1---*/
data sim1;
  call streaminit(10101);
    do sim = 1 to 1000;
          do Block = 1 to 4;
           b = rand('normal')*&std_bk;
           do EL = 1 to 2;
              w = rand('normal')*&std_bel;
              do V = 1 to 3;
               e = rand('normal')*&std_resid;
               output;
              end;
            end;
      end;
    end;
run;
data sim1;
  set sim1;
  y = b + w + e;
```

```
  keep sim block EL V y;
run;
ods Exclude all;
/*---REML---*/
proc mixed data=sim1 method=reml;
       by sim;
       class Block EL V;
       model y = EL|V / ddfm=KR;
       random Block Block*EL;
       ods output tests3=t3sim1r covparms=cpsim1r;
run;
proc sort data=cpsim1r;
       by covparm;
run;
proc means data=cpsim1r;
       by covparm;
       var estimate;
run;
data t3sim1r;
       set t3sim1r;
       if probf < 0.05 then sig = 1;
       else sig = 0;
proc sort data=t3sim1r;
       by effect;
proc means data=t3sim1r mean;
       by effect;
       var sig;
run;
/*---Type3---*/
proc mixed data=sim1 method=type3;
       by sim;
```

```
    class Block EL V;
    model y = EL|V / ddfm=KR;
    random Block Block*EL;
    ods output tests3=t3sim1t covparms=cpsim1t;
run;
proc sort data=cpsim1t;
    by covparm;
run;
proc means data=cpsim1t;
    by covparm;
    var estimate;
run;
data t3sim1t;
    set t3sim1t;
    if probf < 0.05 then sig = 1;
    else sig = 0;
proc sort data=t3sim1t;
    by effect;
proc means data=t3sim1t mean;
    by effect;
    var sig;
run;
ods Exclude all;
/*---Sim 2 to Sim 10---*/

/*---Combining results from Sim 1 to 10---*/
ods select all;
/*---merge---*/
/*---REML---*/
data cpsim_tr;
set cpsim1r cpsim2r cpsim3r cpsim4r cpsim5r cpsim6r cpsim7r cpsim8r cpsim9r cpsim0r;
```

```
proc sort data=cpsim_tr;
        by covparm;
run;
proc means data=cpsim_tr;
        by covparm;
        var estimate;
title 'cpsim REML';
run;
data t3sim_tr;
        set t3sim1r t3sim2r t3sim3r t3sim4r t3sim5r t3sim6r t3sim7r t3sim8r t3sim9r t3sim0r;
        if probf < 0.05 then sig = 1;
        else sig = 0;
proc sort data=t3sim_tr;
        by effect;
proc means data=t3sim_tr mean;
        by effect;
        var sig;
title 't3sim REML';
run;
/*---end of REML---*/
/*---merge---*/
/*---Type3---*/
data cpsim_tt;
set cpsim1t cpsim2t cpsim3t cpsim4t cpsim5t cpsim6t cpsim7t cpsim8t cpsim9t cpsim0t;
proc sort data=cpsim_tt;
        by covparm;
run;
proc means data=cpsim_tt;
        by covparm;
        var estimate;
title 'cpsim Type3';
```

```
run;
data t3sim_tt;
	set t3sim1t t3sim2t t3sim3t t3sim4t t3sim5t t3sim6t t3sim7t t3sim8t t3sim9t t3sim0t;
	if probf < 0.05 then sig = 1;
	else sig = 0;
proc sort data=t3sim_tt;
	by effect;
proc means data=t3sim_tt mean;
	by effect;
	var sig;
title 't3sim Type3';
run;
/*---end of type3---*/
```