Analysis of Unbalancedness in a One-Way Random Effects Model: A Simulation Appraisal

Yogasudha Veturi and Erica Dawson

April 24, 2014

1 Abstract

Unbalanced designs are commonplace in the real world; many factors ranging from inadequate resources for collecting measurements to missing data caused by death or sickness can lead to unbalancedness in the data. In this study, we designed a simulation to explore the role of unbalancedness in a one-way random effects model, which can potentially have an impact on the bias, inference, and particularly on the prediction of random effects. The influence of unbalancedness on biases of parameter estimates, type I error rates, power, and shrinkage in the best linear unbiased predictors, were examined under varying levels of overall sample sizes, number of groups, and between-unit variances. It was observed that unbalancedness did not have a role to play in influencing the bias of $\hat{\sigma}_a$ and $\hat{\mu}$ and in the estimates of their respective variances. However, extreme unbalancedness with a small number of units per group was observed to skew the type I error rates and consequently the power estimates. Unbalancedness was seen to greatly influence the extent of shrinkage in the best linear unbiased predictors; the greater the degree of unbalancedness, the larger was the extent of shrinkage observed.

2 Introduction

Unlike balanced designs, unbalanced designs have an unequal number of observations in each factorial combination or cell. Observational data are frequently unbalanced and many factors motivate the absence of measurements, e.g., death or sickness of patients, death of plants or animals, unexpected breakdowns in equipment, delays in arrival of materials, focused interest in a specific set of levels, cost considerations preventing full exploration of all levels, and the like. In most of these cases a few levels contain a lot more information than the rest. The extent of unbalancedness can range across a spectrum of levels and its interplay with various other factors (e.g., sample size and number of groups) can potentially influence all the subsequent inference. While the influence of unbalancedness on fixed effect models have been thoroughly examined along with recommendations of suitable methodologies being given to deal with such data (Henderson 1953, Shaw 1993), there is a need for an exhaustive examination of the influence of unbalancedness in random effects models at a much more fundamental level. Consequently, we designed a simulation study based on the maximum likelihood (ML) approach to provide a visualization of the interplay between unbalancedness and other critical factors (e.g. sample size, number of groups) on bias, inference and prediction in a basic one-way random effects model.

ML has long been used for ordinary ANOVA models and ML estimators are appealing in several ways: they are functions of sufficient statistics, consistent, asymptotically normal and efficient (Miller 1973, Harville 1977). For at least some unbalanced designs, there exist ML estimators that have smaller variance than Henderson estimators and ML estimates and information matrix can be obtained for any parameterization of the model (Harville 1977, Olsen, Seely, and Birkes 1976). However, estimation of variance components using ML for unbalanced data in random and mixed effects models is more complicated than it is for balanced data. In addition to the cumbersome algebra, there is no closed form to the ML equations in a random effects model and the ML estimates need to be obtained empirically using constrained optimization techniques (Lindstrom and Bates 1988; Pinheiro and Bates 1995). Also, ML estimators dont take into account the loss in degrees of freedom arising due to fixed effects and need a specific parametric form of the distribution (Harville 1977). However, since fixed effects were not used in this study, the loss in degrees of freedom was not considered to be an issue.

An even stronger case can be made about the influence of unbalancedness on prediction. BLUPs (Best Linear Unbiased Predictors) are estimates of random effects and are commonly used to estimate genetic merits in plant and animal breeding (Robinson 1991). BLUPs are best because they have the least mean squared error in the class of linear unbiased estimators. However, in order to achieve this, they need to undergo shrinkage, which results in accruing some bias, i.e., there is a bias-variance trade-off. The extent of shrinkage achieved in the estimation of the BLUP is directly affected by the sample size per group.

Given these considerations, it was decided to examine the influence of unbalancedness on parameter estimation (bias), hypothesis testing (type 1 errors and power under H_0: $\sigma_a^2 = 0$ vs. H_1: $\sigma_a^2 \neq 0$, and prediction (shrinkage in the BLUP) using a simple one-way random effects model, under varying levels of overall sample sizes, numbers of groups, and between-unit variances. There could be several real-life examples that could motivate such a study, e.g., examining whether there is a clinic effect on patient blood pressure when the same drug is administered to all patients, or examining whether there is a site effect on the harvest levels of wheat under the same treatment and so on. The results from this study could motivate such investigations with more complex models (Jennrich and Schluchter 1986).

3 Methodology

The focus of our study was to investigate the influence of unbalancedness on bias, type I error, power, and prediction. The model of interest is the one-way ANOVA model with random effects, denoted $y_i = \mu + a_i + \epsilon_{ij}$, where $i = 1, ..., N$, $j = 1, ..., n_i$. We assume that $a_i \sim \mathrm{N}(0, \sigma_a^2)$, $\epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$, and $y_{ij} \sim \mathrm{N}(\mu, \sigma_a^2 + \sigma^2)$. Note, a_i and ϵ_{ij} are independent. y_i denotes that the weight of the jth subject in the ith group, and a_i denotes the random effect (e.g., drug clinics). Here, σ_a^2 and σ^2 denote the within-unit and between-unit variances in the response, respectively. For $\mathbf{y}_i = [y_{i1}, ...y_{in_i}]'$ the model has $y_{ij} \sim \mathrm{N}(\mu \mathbf{1}_{n_i}, \mathbf{V}_i)$, where

$$\mathbf{V}_i = \sigma^2 I_{n_i} + \sigma_a^2 J_{n_i}$$

and

$$\mathbf{V}_i^{-1} = \frac{1}{\sigma^2}\mathbf{I}_{n_i} - \frac{\sigma^2}{\sigma^2\left(\sigma^2 + n_i\sigma_a^2\right)}\mathbf{J}_{n_i}.$$

Thus,

$$f\left(\mathbf{y}_i|\mu,\sigma^2,\sigma_a^2\right) = (2\pi)^{-n_i/2}|\mathbf{V}_i|^{-1/2}\exp\left\{-\frac{1}{2}\left(\mathbf{y}_i - \mu\mathbf{1}_{n_i}\right)'\mathbf{V}_i^{-1}\left(\mathbf{y}_i - \mu\mathbf{1}_{n_i}\right)\right\}.$$

We ran simulations with assigned parameters of $\mu = 5$, $\sigma^2 = 2.5$, and $\sigma_a^2 = (0.1, 1, 2.5, 4, 5)$, Total sample size N = (100, 500), Number of groups = (2, 5, 10), and Degrees of unbalancedness = (Balanced, moderate and extreme) in R, to generate our data. The simulation was repeated a 1000 times to generate 1000 such datasets for each combination of cases (total of 90 combinations).
In order to investigate the influence of unbalancedness on bias, type I error, power, and prediction, we considered the following scenarios:

The three different levels of unbalancedness were investigated were arrived at thusly: balanced (when levels of our random effect term have the same number of observations), moderate (50% of the levels were sampled at twice the frequency of the other 50 %), and extreme (50 % of the levels were sampled at 4 times the frequency of the other 50 %).

3.1 Likelihood

Since our model of interest is the one-way ANOVA model with random effects, we have the following likelihood function:

$$L = \prod_{i=1}^{m}(2\pi)^{-n_i/2}|\mathbf{V}_i|^{-1/2}\exp\left\{-\frac{1}{2}\left(\mathbf{y}_i - \mu\mathbf{1}_{n_i}\right)'\mathbf{V}_i^{-1}\left(\mathbf{y}_i - \mu\mathbf{1}_{n_i}\right)\right\}.$$

By taking the logarithm of both sides of the likelihood function, we get that

$$\begin{aligned} l &= -\frac{1}{2}N\log 2\pi - \frac{1}{2}\sum_i \log(\sigma^2 + n_i\sigma_a^2) - \frac{1}{2}(N-m)\log\sigma^2 \\ &\quad -\frac{1}{2\sigma^2}\sum_i\sum_j(y_{ij}-\mu)^2 + \frac{\sigma_a^2}{2\sigma^2}\sum_i\frac{(y_{i\cdot} - n_i\mu)^2}{\sigma^2 + n_i\sigma_a^2} \end{aligned}$$

Define $\lambda_i = \sigma^2 + n_i\sigma_a^2$. Also, note $y_{ij} - \mu = y_{ij}$ as $y_{ij} - \bar{y}_{i\cdot} + \bar{y}_{i\cdot} - \mu$. This implies that the likelihood can be rewritten as

$$\begin{aligned} l &= -\frac{1}{2}N\log 2\pi - \frac{1}{2}\sum_i \log(\lambda_i) - \frac{1}{2}(N-m)\log\sigma^2 \\ &\quad -\frac{1}{2\sigma^2}\sum_i\sum_j(y_{ij}-\mu)^2 + \sum_i\frac{n_i(\bar{y}_{i\cdot} - n_i\mu)^2}{2\lambda_i} \end{aligned}$$

Figure 1: Simulation Parameters

σ^2 = 2.5, μ = 5, σ_a^2 =(0.1, 1, 2.5, 4, 5), 1000 Simulations

Overall N	Unbalancedness	No. of groups	n_i									
100	Balanced	2	50	50								
100	Moderate	2	64	36								
100	Extreme	2	90	10								
500	Balanced	2	250	250								
500	Moderate	2	343	157								
500	Extreme	2	420	80								
100	Balanced	5	20	20	20	20	20					
100	Moderate	5	18	25	13	29	15					
100	Extreme	5	38	37	7	7	11					
500	Balanced	5	100	100	100	100	100					
500	Moderate	5	132	138	75	74	81					
500	Extreme	5	185	192	45	34	44					
100	Balanced	10	10	10	10	10	10	10	10	10	10	10
100	Moderate	10	12	14	14	18	9	11	6	7	5	4
100	Extreme	10	16	23	14	15	23	3	2	1	3	
500	Balanced	10	50	50	50	50	50	50	50	50	50	50
500	Moderate	10	70	78	59	82	66	27	28	33	29	28
500	Extreme	10	80	93	98	82	95	11	12	10	11	8

3.2 Maximum Likelihood Equations and Their Solutions

Note, $\partial\lambda_i/\partial\sigma^2 = 1$ and $\partial\lambda_i/\partial\sigma_a^2 = n_i$. Let $l_\theta \equiv \partial logL/\partial\theta$. Then

$$l_\mu = \sum_i \frac{n_i(\bar{y}_{i\cdot} - n_i\mu)^2}{2\lambda_i}, \tag{1}$$

$$l_{\sigma^2} = \frac{-(N-m)}{2\sigma^2} - \frac{1}{2}\sum_i \frac{1}{\lambda_i} + \frac{\sum_i \sum_j (y_{ij}-\mu)^2}{2\sigma^4} + \sum_i \frac{n_i(\bar{y}_{i\cdot}-\mu)^2}{2\lambda_i^2}, \tag{2}$$

and

$$l_{\sigma_a^2} = -\frac{1}{2}\sum_i \frac{n_i}{\lambda_i} + \sum_i \frac{n_i^2(\bar{y}_{i\cdot}-\mu)^2}{\lambda_i^2} \tag{3}$$

In order to obtain the maximum likelihood (ML) equations, we must equate the above expressions to zero using $\dot{\mu}$, $\dot{\sigma}^2$, and $\dot{\lambda}_i = \dot{\sigma}^2 + n_i\sigma_a^2$ as the solutions. Thus,

$$\dot{\mu} = \frac{\sum_i \bar{y}_{i\cdot}/v\dot{a}r(\bar{y}_{i\cdot})}{\sum_i 1/v\dot{a}r(\bar{y}_{i\cdot})}, \tag{4}$$

where $v\dot{a}r(\bar{y}_{i\cdot}) = \sigma_a^2 + \sigma^2/n_i$. By equating the right-hand sides of (2) and (3) to zero we are able to derive $\dot{\sigma}^2$ and $\dot{\sigma}_a^2$, i.e.,

$$\frac{-(N-m)}{2\sigma^2} - \frac{1}{2}\sum_i \frac{1}{\lambda_i} + \frac{\sum_i \sum_j (y_{ij}-\mu)^2}{2\sigma^4} + \sum_i \frac{n_i(\bar{y}_{i\cdot}-\mu)^2}{2\lambda_i^2} = 0 \tag{5}$$

and

$$-\frac{1}{2}\sum_i \frac{n_i}{\lambda_i} + \sum_i \frac{n_i^2(\bar{y}_{i\cdot}-\mu)^2}{\lambda_i^2} = 0. \tag{6}$$

Since $\dot{\lambda} = \dot{\sigma}^2 + n_i\dot{\sigma}_a^2$ occurs in the denominators of the terms being summed, over i, in the above expressions, there is *no analytic solution* for the estimators of μ, σ^2, and σ_a^2.

3.3 Maximum Likelihood Estimators

We used Nelder and Mead optimization (Nelder and Mead 1965) to obtain maximum likelihood estimates for σ^2 and σ_a^2 instead of using the default lme4 package for mixed effect models. The solutions $\dot{\mu}$, $\dot{\sigma}^2$, and $\dot{\sigma}_a^2$ are MLEs of μ, σ^2, and σ_a^2, respectively, if the vector $(\dot{\mu}, \dot{\sigma}^2, \dot{\sigma}_a^2)$ is in the space of $(\mu, \sigma^2, \sigma_a^2)$. Equations (5), (6), and (7) must be solved numerically, using an iterative algorithm that handles non-linear equations. The following rule must be applied after solving the equations numerically:
if $\dot{\sigma}_a^2 \geq 0$,

$$\hat{\mu} = \dot{\mu}, \hat{\sigma}^2 = \dot{\sigma}^2, \text{ and } \hat{\sigma}_a^2 = \dot{\sigma}_a^2 \tag{7}$$

if $\dot{\sigma}_a^2 > 0$,

$$\hat{\mu} = \bar{y}_{i\cdot\cdot}, \hat{\sigma}^2 = \frac{1}{N}\sum_i \sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2, \text{ and } \hat{\sigma}_a^2 = 0 \tag{8}$$

.

3.4 Bias

When $\dot{\sigma}_a^2 > 0$, the expected value of $\hat{\sigma}^2$ is not an unbiased estimator of σ^2 and cannot be derived easily. Hence, the bias must be obtained empirically.

3.5 Sampling Variance

Maximum likelihood estimators are consistent and asymptotically efficient (Casella and Berger). Hence, since (7) and (8) are derived by the maximum likelihood estimation, they are consistent and asymptotically efficient. This implies that

$$\mathrm{var}\begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \rightarrow \begin{bmatrix} \sum_i \frac{n_i}{\lambda_i} & 0 & 0 \\ 0 & \frac{N-M}{2\sigma^4} + \frac{1}{2}\sum_i \frac{1}{\lambda_i^2} & \frac{1}{2}\sum_i \frac{n_i}{\lambda_i^2} \\ 0 & \frac{1}{2}\sum_i \frac{n_i}{\lambda_i^2} & \frac{1}{2}\sum_i \frac{n_i^2}{\lambda_i^2} \end{bmatrix}^{-1}.$$

Thus,

$$\mathrm{var}(\hat{\mu}) \rightarrow (\sum_i \frac{n_i}{\lambda_i})^{-1} = (\sum_i \frac{n_i}{\sigma^2 + n_i\sigma_a^2})^{-1}$$

and

$$\mathrm{var}\begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \rightarrow \frac{2}{D}\begin{bmatrix} \sum_i \frac{n_i^2}{\lambda_i^2} & \sum_i \frac{n_i}{\lambda_i^2} \\ -\sum_i \frac{n_i}{\lambda_i^2} & \frac{N-M}{2\sigma^4} + \sum_i \frac{1}{\lambda_i^2} \end{bmatrix},$$

where

$$D = \frac{N-m}{\sigma^4}\sum_i \frac{n_i^2}{\lambda_i^2} + \sum_i \frac{1}{\lambda_i^2}\sum \frac{n_i^2}{\lambda_i^2} - (\sum_i \frac{n_i}{\lambda_i^2})^2.$$

3.6 Power and Type I Error

Our main hypothesis of interest is $H_0 : \sigma_a^2 = 0$ versus $H_1 : \sigma_a^2 \neq 0$.
In a random effects one-way ANOVA model, our main hypothesis of interest is $H_0 : \sigma_a^2 = 0$ versus $H_a : \sigma_a^2 \neq 0$. Using this hypothesis, type I error is calculated under the null model, i.e., $y_{ij} = \mu + \epsilon_{ij}$, where $i = 1, ..., N,\ j = 1, ..., n_i$. We assume that $\epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$, and $y_{ij} \sim \mathrm{N}(\mu, \sigma^2)$. Power is calculated using the alternative model, i.e., $y_i = \mu + a_i + \epsilon_{ij}$, where $i = 1, ..., N,\ j = 1, ..., n_i$. We assume that $a_i \sim \mathrm{N}(0, \sigma_a^2), \epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$, and $y_{ij} \sim \mathrm{N}(\mu, \sigma_a^2 + \sigma^2)$. Note, a_i and ϵ_{ij} are independent. Type I error rates and power were both calculated at varying levels of σ_a^2 for different sample sizes and levels of our random effect term. For balanced or unbalanced data a likelihood ratio test of $H_0 : \sigma_a^2 = 0$ is to reject H_0 when F = MSA/MSE $> F^{m-1}_{N-m,1-\alpha}$, where

$$MSA = \sum_i n_i(\bar{y}_{i.} - \bar{y}_{..})^2/(m-1)$$

and

$$MSE = \sum_i \sum_j n_i(\bar{y}_{i.} - \bar{y}_{..})^2/(N-m)$$

3.7 Prediction

In order to predict a_i, the random effect in term in our model of interest, use the fact that

$$\begin{aligned}
E[a_i|\mathbf{y}] &= E[a_i|\bar{y}_{i.}] \\
&= E[a_i] + \text{cov}(a_i, \bar{y}_{i.})[var(\bar{y}_{i.})]^{-1})(\bar{y}_{i.} - E[\bar{y}_{i.}]) \\
&= 0 + \sigma_a^2 \frac{1}{\sigma_a^2 + \sigma_2/n_i}(\bar{y}_{i.} - \mu) \\
&= \frac{\sigma_a^2}{\sigma_a^2 + \sigma_2/n_i}(\bar{y}_{i.} - \mu),
\end{aligned}$$

which is known as the best predictor of a_i, which is denoted by $\text{BP}(a_i)$. Note, the $\text{BP}(a_i)$ depends on unknown parameters μ, σ_a^2, and σ^2. By replacing the parameters with estimates, we obtain an estimated best predicted value, denoted by $\tilde{a}_i$.

$$\begin{aligned}
\tilde{a}_i &= \hat{\text{BP}}(a_i) = \hat{\text{E}}[a_i \bar{y}_{i.}] \\
&= \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_2/n_i}(\bar{y}_{i.} - \hat{\mu})
\end{aligned}$$

For estimation in a fixed effect model, i.e., $y_i = \mu + \alpha_i + \epsilon_{ij}$, where $i = 1, ..., N$, $j = 1, ..., n_i$. We assume that $\epsilon_{ij} \ \sim \text{N}(0, \sigma^2)$, and $y_{ij} \ \sim \text{N}(\mu + \alpha, \sigma_a^2 + \sigma^2)$,

$$\hat{\alpha}_i = (\bar{y}_{i.} - \bar{y}_{..}).$$

Both $\tilde{a}_i$ and $\hat{\alpha}_i$ are based on $\bar{y}_{i.} - \bar{y}_{..}$, $\tilde{a}_i$ is smaller compared to $\hat{a}_i$. $\tilde{a}_i$ is always smaller than $\hat{\alpha}_i$, but how small depends on the size of $\hat{\sigma}_a^2$ and $\hat{\sigma}^2/n_i$. If $\hat{\sigma}_a^2$ is large with respect to the estimate of $\hat{\sigma}^2/n_i$, then $\tilde{a}_i$ and $\hat{\alpha}_i$ are similar.

4 Results

4.1 Bias

Very little bias was observed in the estimate of μ (under the balanced design, $\hat{\mu}$ is an unbiased estimate of μ) and the bias further dropped to zero when the number of groups increased. Unbalancedness was not observed to have a special effect on the bias of $\hat{\mu}$ (Fig. 2 and 3). Likewise, the bias in the estimate of variance of $\hat{\mu}$ was also close to zero. However, for $\hat{\sigma}_a$, big biases were observed for both the means and variances, especially when the number of groups was small (Fig. 4 and 5). Since $\hat{\sigma}_a^2$ is a biased estimator of σ_a^2, the bias (even with 10 groups) did not reduce to zero (for both means and variances), although there was a drastic reduction in the bias as the number of groups increased to 10.

4.2 Type I Errors

The type I error rates were estimated under a chosen α of 0.05. The means and medians of the type I error rates were centered at $\alpha = 0.05$ in all scenarios except for one: when there was extreme unbalancedness with 10 groups under a sample size of 100 (Fig. 6; top, right). A comparison with the one-way fixed effect model was made to see if it performed better; indeed α was centered at 0.05 even for this scenario under the fixed effects model (Fig 7; top, right). Standard errors were based on p-values obtained from permutation tests (1 p-value per test, performed 100 times) on each of the 1000 simulated samples.

4.3 Power

As expected, a clear influence of sample size was observed on power. Under a sample size of 500, 100% power was reached more quickly (at a smaller σ_a^2 of 1) as compared to a sample size of 100. However, a bigger impact on power was due to the number of groups. When the number of groups was equal to 2, a maximum power of just 90% could be achieved even when the total sample size was as large as 500. However, the difference in power curves between 5 and 10 groups was minimal (Fig 8). The effect of unbalancedness on power could be observed in the scenario when sample size was 100; power was observed to be lower. Very similar power curves were observed under the one-way fixed effects model as well, except when there was extreme unbalancedness with 10 groups under a sample size of 100 (Fig 9; top, right). This was the same scenario for which the median type I error rate deviated from $\alpha = 0.05$ under the random effects model. Consequently, the power curve obtained from this scenario was treated with skepticism.

4.4 Shrinkage

There was a distinct impact of unbalancedness on shrinkage of the BLUP. This effect was more pronounced under a smaller overall sample size (Fig 10). Fig. 10 examines the extent of shrinkage for a model with 10 groups. Extent of shrinkage is influenced by a combination of n_i, σ^2, and σ_a^2. When σ_a^2 is estimated to be small with respect to σ^2/n_i, the extent of shrinkage can be substantial (Searle 2012), which was corroborated by our simulation. When σ_a^2 was larger than σ^2 the extent of shrinkage decreased considerably for the balanced and moderate unbalancedness cases, although there still remained extensive shrinkage for the extreme unbalancedness case. Again, overall sample size had a major role to play in determining the extent of shrinkage. The influence of unbalancedness on shrinkage was vastly muted under a sample size of 500, especially when σ_a^2 was larger than σ^2.

5 Discussion

This study was conducted to investigate the influence of unbalancedness in a one-way random effects model on four components: a) Bias b) Type I errors c) Power d) Shrinkage. A simulation study was designed to this effect under two overall sample sizes (100 and 500), three numbers of groups (2, 5, and 10), and five levels of between-unit variance or σ_a^2 (0.1, 1, 2.5, 4, 5). Extent of unbalancedness was varied under three levels 1) Balanced (no unbalancedness) 2) Moderate and 3) Extreme. Unbalancedness did not seem to influence bias in $\hat{\mu}$, $\hat{var}(\hat{\mu})$, $\hat{\sigma}^2$ and $\hat{var}(\hat{\sigma}_a^2)$. Type I errors were centered at 0.05, which was the chosen alpha. However, when there was extreme unbalancedness and an inadequate number of units in a few groups, the null distribution of p-values for H_0: $\sigma_a^2 = 0$ was no longer as expected, resulting in a skewed type I error rate (Fig 5). Consequently, the estimated power in this scenario was questionable (Fig 7). For this scenario, a fixed effects model could be a better choice to test for a difference in means between the groups. However, such a choice is not straightforward; the larger objectives of the study and the interpretation of the random effect term under question need to be borne in mind before making this choice. If interest lay in prediction or in estimating variance components, it would never be appropriate to fit a fixed effects model. The reason behind good power with a fixed effects model is that the number of random effect levels sampled is too small to not cause a significant difference between means test. It is likely that the power with the fixed effects model will drop as the number of levels of the random effect term increases well beyond 10.

Finally, unbalancedness had a key role to play in determining the extent of shrinkage (Fig 9). The more the unbalancedness, the greater was the shrinkage, especially when the sample size was smaller. However, influence of shrinkage was governed more by the interplay between n_i, σ^2, and σ_a^2.

The number of groups in the model influenced both, the bias in estimates of $\hat{\sigma}_a^2$ and $\hat{var}(\hat{\sigma}_a^2)$. Bias decreased as the number of groups increased, irrespective of the degree of unbalancedness or sample size. Power also increased with the number of groups. A model with just two groups was observed to be having markedly low power and large biases in both $\hat{\sigma}_a^2$ and $\hat{var}(\hat{\sigma}_a^2)$ as compared to a model with 5 or 10 groups. The power was as high as 90% even under a moderate σ_a^2 of 1, with just 5 groups and a smaller overall sample size of 100. Additionally, as expected, the overall sample size had a marked role in the power and extent of shrinkage but did not have an impact in reducing the bias in $\hat{\sigma}_a^2$ and $\hat{var}(\hat{\sigma}_a^2)$. Based on this study, we could infer that biases in the estimates of $\hat{\sigma}_a^2$ and $\hat{var}(\hat{\sigma}_a^2)$ can be handled primarily by increasing the number of groups.

Influence of unbalancedness on mixed effect models with multi-way interactions or non-orthogonal predictor variables where covariance structures need to be modeled (e.g. repeated measures data) would be an interesting extension of this study. Additionally, examining how unbalancedness affects binary response and count data responses is also worthy of further examination.

Figure 2: Bias in $\hat{\mu}$ (one-way random effects model)

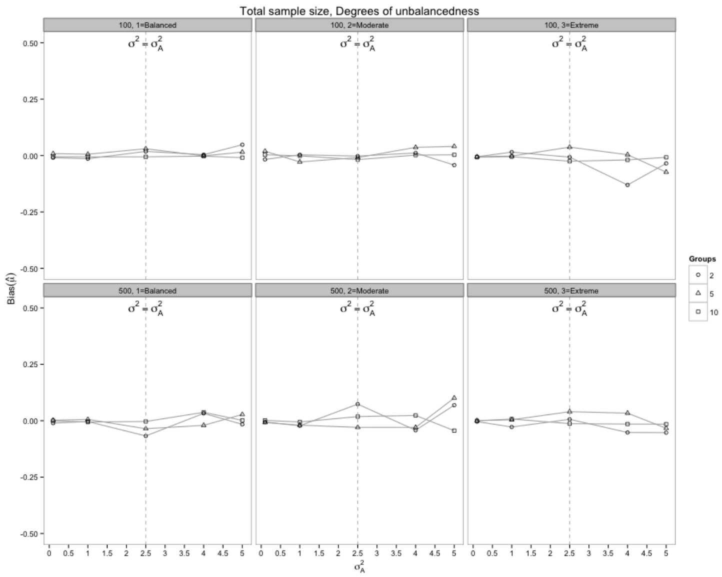

Figure 3: Bias in $v\hat{a}r(\hat{\mu})$ (one-way random effects model)

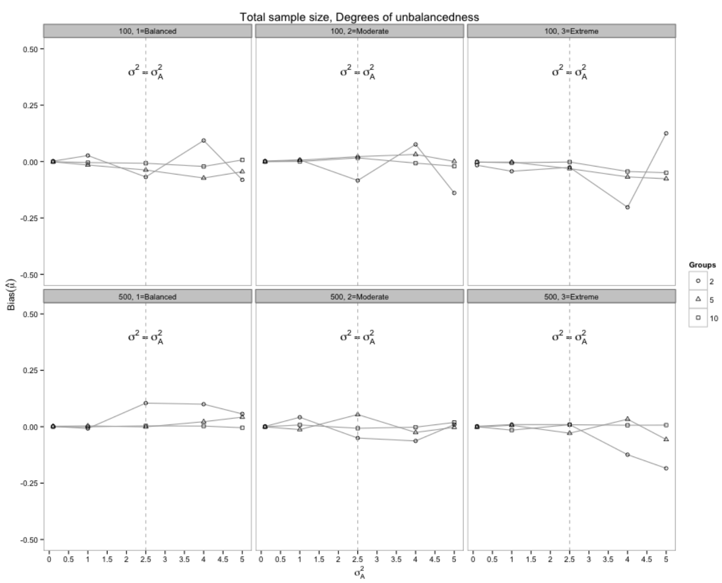

Figure 4: Bias in $\hat{\sigma}_a^2$ (one-way random effects model)

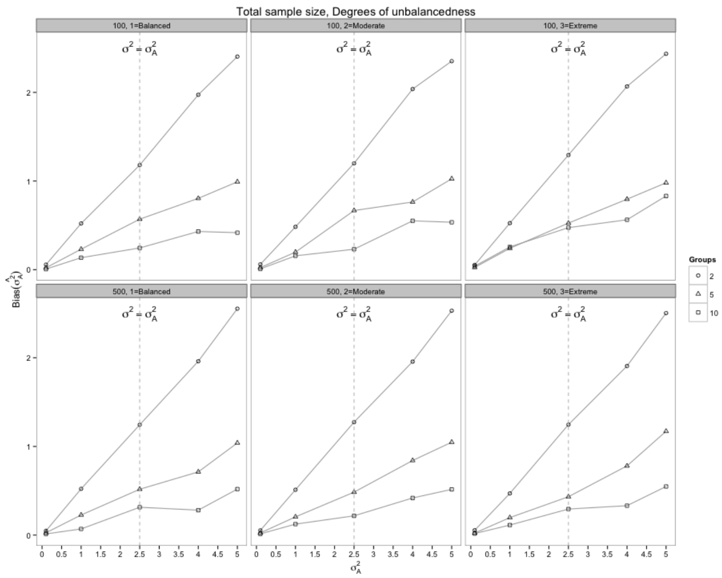

Figure 5: Bias in $v\hat{a}r(\hat{\sigma}_a^2)$ (one-way random effects model)

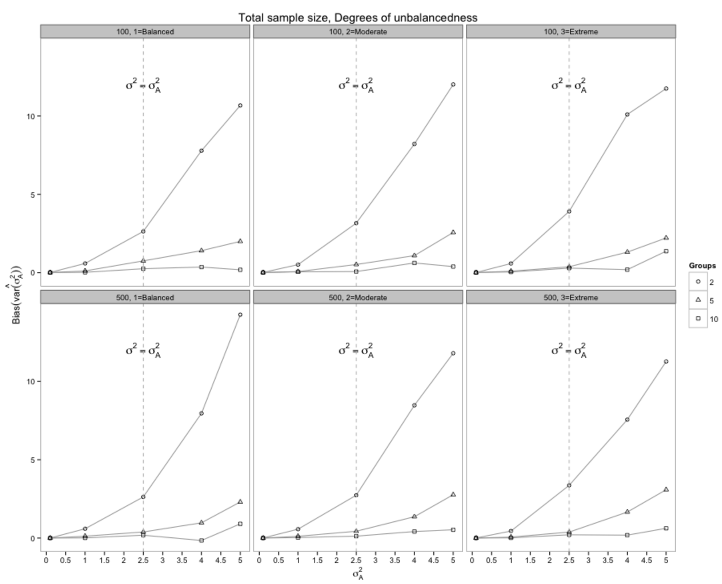

Figure 6: Type I Error Rates (one-way random effects model)

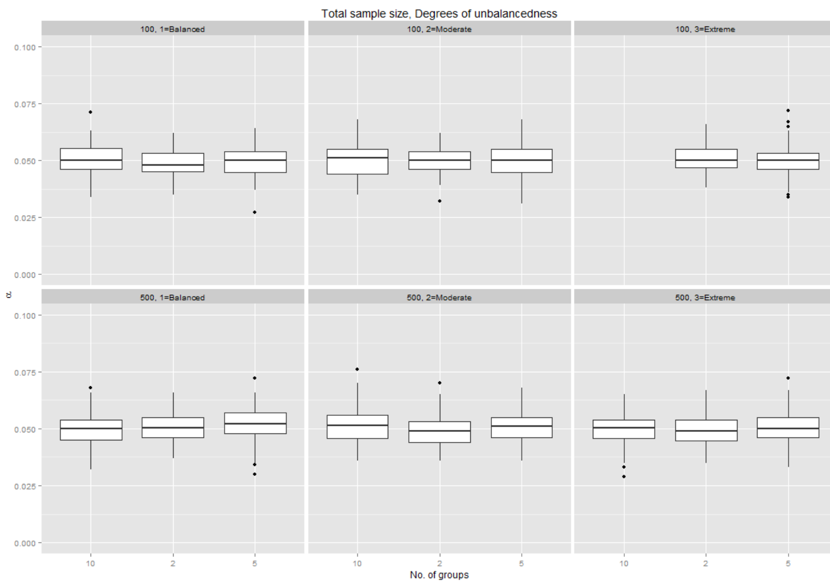

Figure 7: Type I error rates (one-way fixed effects model)

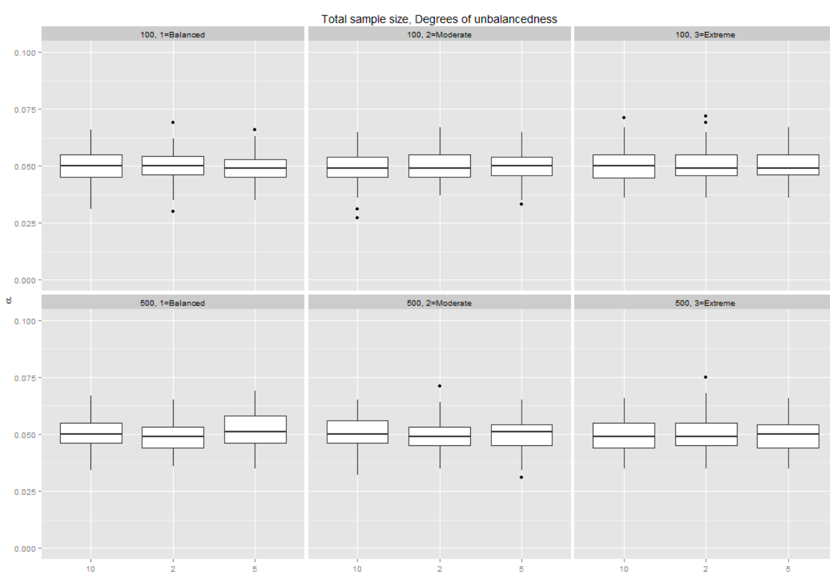

Figure 8: Power (one-way random effects model)

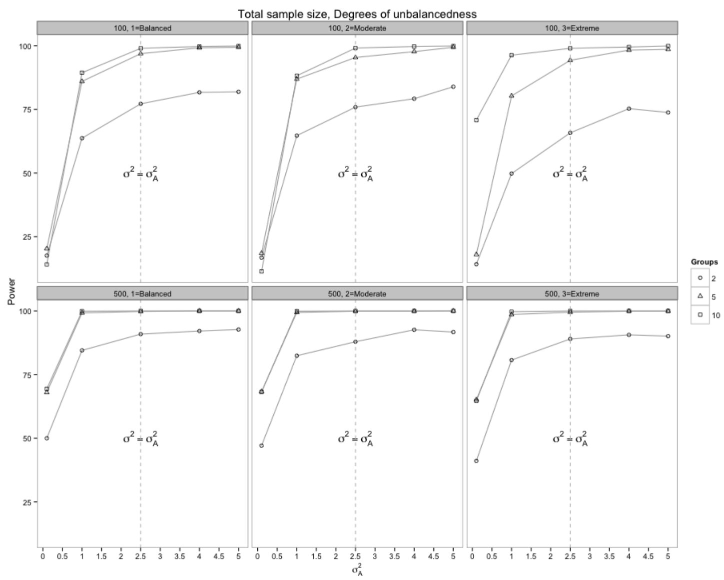

Figure 9: Power (one-way fixed effects model)

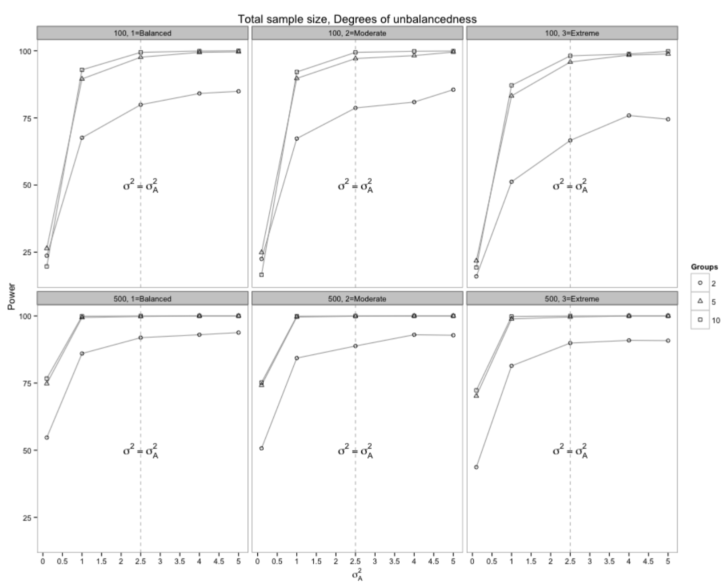

Figure 10: Shrinkage with BLUP for 10 groups (one-way random effects model)