How to estimate the measurement error variance associated with ancestry proportion estimates

Jasmin Divers[*], **David T. Redden**, **Raymond J. Carroll**, and **David B. Allison**

Abstract

To show how the variance of the measurement error (ME) associated with individual ancestry proportion estimates can be estimated, especially when the number of ancestral populations (*k*) is greater than 2.

We extend existing internal consistency measures to estimate the ME variance, and we compare these estimates with the ME variance estimated by use of the repeated measurement (RM) approach. Both approaches work by dividing the genotyped markers into subsets. We examine the effect of the number of subsets and of the allocation of markers to each subset on the performance of each approach. We used simulated data for all comparisons.

Independently of the value of *k*, the measures of internal reliability provided less biased and more precise estimates of the ME variance than did those obtained with the RM approach. Both methods tend to perform better when a large number of subsets of markers with similar sizes are considered.

Our results will facilitate the use of ME correction methods to address the ME problem in individual ancestry proportion estimates. Our method will improve the ability to control for type I error inflation and loss of power in association tests and other genomic research involving ancestry estimates.

Keywords

Population stratification; admixture; type I error inflation; reliability; Cronbach's alpha; measurement errors; measurement error variance

1. Introduction

Population stratification and admixture are concerns in genetic association studies. Failure to adequately control for them in genetic association tests may lead to inflated type I error and/ or loss of power. Structured association tests (SATs), in which individual admixture proportion estimates (IAPE) [1–4] or individual-level measures of genetic background [5;6] are computed and included in the model as covariates to control for confounding, are widely applied in genetic association tests. We have shown [7] that the SAT approach can be cast in the general linear model framework. If unmodified, the use of this framework implicitly assumes that the predictors are measured without error. We also showed in that study that controlling for admixture alone is not sufficient to control for spurious associations. Investigators must also account for the individual's true ancestry proportion as well as the product of the parental ancestries in the association test to completely remove the

---
[*]Address correspondence to Jasmin Divers, Section on Statistical Genetics and Bioinformatics, Center for Public Health Genomics, Department of Biostatistical Sciences, Division of Public Health Services, Wake Forest University Health Sciences, WC-2326, Medical Center Blvd. Winston-Salem, North Carolina, 27157, jdivers@wfubmc.edu.

confounding effect of admixture-induced linkage disequilibrium. We define an individual ancestry proportion, with respect to a specific ancestral population, as the proportion of that individual's ancestors who were members of that parental population in the generation before the first admixture event. This is in contrast to an individual's admixture, which is the proportion of the individual's genome that is inherited from a specific ancestral population. For example, two full siblings have the same ancestry. However, random variations that happen during meiosis may lead to different admixture proportions both at the global level and at the local level. Therefore, independently of the approach chosen to estimate the IAPEs, the resulting estimate should be seen as an imperfect or error-contaminated measurement of the true individual ancestry estimate. Existing methods and software for estimating admixture proportions such as STRUCTURE and FRAPPE provide an estimate of the standard error associated with the admixture proportion. Here, the purpose is to go one step further and compute an estimate of the ME associated with the IAPEs.

The introduction of error-contaminated covariates in a regression model can lead to type I error inflation and loss of power [8]. Divers et al. [9] and Padilla et al. [10] discussed how existing measurement correction methods can be tailored for application in SATs. These methods assume that an estimate of the ME variance is available. To support such availability, we previously showed how Cronbach's alpha, a measure of internal reliability, can be used to obtain an estimate of the measurement error variance. The methods that we used in these studies [9;10] assumed that k, the number of ancestral populations that intermated to create the admixed population, is equal to 2. In this case, the confounder can be represented by a single predictor in the model. We now extend this approach for estimating the ME variance when $k > 2$. We focus mainly on the case where $k = 3$ because it has direct application when controlling for admixture-induced confounding in genetic association involving, for example, Caribbean Hispanics. However, the methods that we discuss below are valid for any value of k. The properties of internal reliability measures like Cronbach's alpha are well known for univariate scales, but they are less well studied for multivariate scales. We also compare the estimate of the ME variance computed by using Cronbach's alpha to an estimate obtained by using the repeated measurement approach discussed in [8]. Both methods work by dividing the ancestry informative markers (AIMs) into p subsets. Individual proportion estimates are obtained for each subset and are then combined into an overall estimate for which the ME variance can be computed.

1.1 ME in individual admixture estimates

Given the distinction that we make between admixture and ancestry, it is easy to see that admixture is a function of true ancestry and random biological variations. The admixture estimates provided by existing software should be seen as error-contaminated measurements of true ancestry. These measurement errors occur for several reasons. (1) Only a subset of genetic markers with imperfectly known ancestral population allele frequencies is used to estimate the individual admixture proportions. (2) Imperfect historical knowledge about the admixed population can also lead to inaccurate estimates of individual admixture. For example, the number of ancestral populations that intermated to create the admixed population is not always well known. (3) Most AIMs are not perfectly ancestry informative. That is, they do not exhibit variants that are seen only in one ancestral population and not in the others. Therefore, when a variant is observed in an admixed individual, its ancestral origin cannot be inferred with certainty.

1.2 ME in genetic background measures

Measurement errors can also occur when principal component analysis (PCA) is used to control for population stratification and admixture. PCA does not necessarily yield admixture estimates. Instead, it identifies axes of variations that may correspond to

population substructure or other sources of variations, such as plate effects, genotyping error, or region of the genome with long-range linkage disequilibrium (LD). Even in cases where one or more axes of variation can be associated with population substructure, the inferred axes of variations may not be parallel to the true axes. In fact, Paul [11] showed that the inferred eigenvector is parallel to the true eigenvector if and only if the N corresponding

eigenvalue is between 1 and $1+ \sqrt{\dfrac{N}{M}}$ where N is the number of rows (samples) and M is the number of markers used in the PCA. This result assumes that $M \quad N$, which is not likely to be the case in a genome-wide association study (GWAS). We note that for a GWAS, principal components are usually computed on the transpose of the genetic data matrix in an effort to satisfy this requirement. Though enabling computation, it is not entirely clear that transposition of this matrix really solves the $M \quad N$ problem. We note that ideas similar to those described below can be applied to address the ME problem for principal components. However, we restrict our attention to the ME problem in ancestry estimates.

1.3 Effect of Measurement errors in regression models

It is well known that including covariates measured with error in a model can lead to biased and inconsistent parameter estimates [8]. Measurement errors in the estimation of the individual ancestry proportion or genetic background means that residual confounding may still persist even after adjustments for population stratification and admixture are made. We have shown how ME correction methods [9;10] can be applied to help keep the type I error rate at its nominal level in genetic association tests. However, an estimate of the ME variance is required before the measurement correction methods can be applied. This estimate is usually derived by considering deviations from a 'gold standard' measurement, by considering previously collected information, or through repeated measurements. A gold standard measurement or a previous estimate of the ME variance associated with the individual ancestry proportions or individual genetic background measures is not likely to be available in genetic association studies. A straightforward repeated measurements approach consisting of genotyping different sets of AIMs on the same individuals for the sole purpose of ancestry proportion estimation may be not feasible or may be too expensive for this type of analysis. However, admixture proportion estimates computed on the autosomes or on any independent subsets of AIMs can be treated as repeated observations of the underlying true individual ancestry proportion. Once an estimate of the ME variance is available, the ME correction methods that we previously described can then be applied in most genetic association tests independently on the value of k.

The remainder of this article is organized as follows. In section 2.1 we present a short review of the reliability concept and how Cronbach's alpha in particular can be used to estimate the variance of the ME associated with individual ancestry estimates. In section 2.2., we present an extension of the Cronbach's alpha when the items have different weights. In section 2.3 we show how a measure of reliability can be obtained when the number of ancestral populations (k) is greater than 2. We describe our simulation procedure in section 3, show the result of these simulations in section 4, and then present our conclusions in the discussion section.

2. Method

All ME correction methods assume that an estimate of the ME variable is available. In this section, we show how this estimate can be computed for ancestry proportions that are used as control variables in SAT. We note that this variable can be either univariate, when the number of ancestral populations (k) is equal to 2, or multivariate when this number is greater

than 2. We begin by showing how the ME variance can be estimated in the first case and then extend these methods to cases where k is greater than 2.

2.1 The use of Cronbach's alpha to estimate measurement error variance when k=2

We showed previously how Cronbach's alpha [12], a measure of internal consistency, can be used to provide a lower bound of the reliability of individual admixture as a measure of the underlying individual ancestry. Cronbach's alpha estimates how well a set of items (or variables) measures an underlying unidimensional latent construct. Briefly, we suggested obtaining independent estimates of an individual admixture proportion. These estimates will serve as the items in the estimation of Cronbach's alpha. That is, they will serve as manifestations of the same underlying latent construct. For example, one can compute estimates at the autosomal level. Therefore, one would have 22 independent estimates of an individual's admixture proportion. Cronbach's alpha provides a measure of the reliability of their average as an estimate of the overall IAPE.

We expressed the individual ancestry proportion estimate in the classical true score model (CTM) [13–15] by writing the observed proportion as:

$$W_{ij}=X_i+U_{ij} \quad (1)$$

where W_{ij} is the admixture proportion estimated for the ith individual using markers selected on the jth subset, X_i is the individual's true but unobserved ancestry proportion, and U_{ij} is the ME associated with this estimate. The error term U_{ij} can conceptually be broken down into 2 components, $U_{ij}^{\mathbf{1}}$ and $U_{ij}^{\mathbf{2}}$, where $U_{ij}^{\mathbf{1}}$ represents the measurement errors whose sources are described in section 1.1 and $U_{ij}^{\mathbf{2}}$ summarizes all the biological variations that result in the difference in admixture between two full sibs. We will work with the aggregated variable U_{ij} and assume that

$$U_{ij} \sim N\left(0, \sigma_u^2\right) \quad (2)$$

Additional requirements of the CTM are that X_i and U_{ij} are independent and that σ_U^2 is constant. These assumptions lead to

$$Cov(W_{ij}, W_{il})=Cov(X_i, X_i)=Var(X_i)=\sigma_X^{\mathbf{2}} \quad (3)$$

The last equation implies that the admixture proportion estimates computed on the jth and l^{th} subsets are both measuring the same underlying latent variable, which in this case is the true ancestry proportion. The reliability of W as a measure of X is generally defined as the ratio:

$$\rho=\frac{\sigma_X^2}{\sigma_W^2}=\frac{\sigma_X^2}{\sigma_X^2+\sigma_U^2} \quad (4)$$

which can also be seen as the squared correlation between W and X. Cronbach's alpha [12] provides an estimate of the upper bound of the reliability measure with equality only under tau-equivalence. The constant variance assumption can be relaxed to allow for more informative subsets to have less measurement errors than subsets that are less informative. We address how Cronbach's alpha can be computed when the measurement error variance is not constant in section 2.2.

To compute Cronbach's alpha, let m be the total number of AIMs available for the study, divide m into p subsets and let m_j, $j=1,\ldots,p$ be the number of markers in the jth subset. The subsets can be an autosome or any combination of markers. The only requirement is that the subsets are mutually exclusive. One can then use existing software packages to obtain individual ancestry proportions on each subset, which we denote adx_i. Let V denote the observed variance-covariance matrix calculated from the admixture estimates obtained from each subset. A measure of the reliability of the adx_i's as an overall measure of individual ancestry is given by:

$$\alpha_{equal} = \frac{k-1}{k}\left(1 - \frac{\sum_{j=1}^{k}V_{jj}}{\sum_{l=1}^{k}\sum_{j=1}^{k}V_{lj}}\right) \quad (5)$$

where α_{equal} is the Cronbach's estimate obtained when all items are assigned the same weight. This relation holds as long as the overall individual admixture estimate can be seen as the sum of the individual admixture estimates computed on each subset [16]. The average can always be seen as a weighted sum. Once α is computed, an upper bound of the ME variance is given by:

$$\hat{\sigma}^2_{u,C} = (1-\alpha)S^2_w \quad (6)$$

where S^2_W is the sample variance of the estimated ancestry proportions and the letter C in the subscript denotes that this estimate is computed by using the Cronbach's alpha approach.

2.2 Estimating the ME variance for unequal weights

In Divers et al. [9], we suggested that one can divide the total number of markers into p subsets, compute individual admixture proportion estimates for each subset, and then average over these estimates to obtain an estimate of the overall ancestry proportion. Our simulations have shown that this approach works well. First, the Cronbach alpha estimates were very close to the true reliability values, which suggest that the CTM requirements were met. Second, the ancestry proportions obtained by using the proposed approach were highly correlated with ancestry proportions computed by using all the markers in a single estimate. These simulations assume that the same amount of information was available for each chromosome. However, we wonder whether it would be more appropriate to consider a weighted instead of an unweighted average. There are numerous reasons beyond the statistical argument to consider a weighted average. For example, if the subsets are made of markers selected from the same autosome, it should be expected that ancestry proportions estimated on longer chromosomes would be more accurate than those computed on shorter chromosomes. This is true because longer chromosomes are more likely to host AIMs that are effectively independent (i.e., more linkage groups) than are shorter ones. Consequently, longer chromosomes are expected to be better represented in most sampling plans that do not explicitly seek to select the same number of AIMs per chromosome. Also, if subset A contains makers that are more ancestry informative than those that are considered in subset B, the ancestry proportion estimates obtained with subset A will also tend to be more accurate than those computed from subset B assuming that both subsets contained approximately the same number of markers. In both cases, it makes sense to weight these subsets differently to capitalize upon their degree of accuracy. We should also note that two subsets with a different number of elements (cardinal) can contain markers with different ancestry informativeness contents such that the number of AIMs or the informativeness content of each marker alone is not sufficient to predict which subset will lead to more

accurate estimates of the individual ancestry proportions. The optimal weighting scheme will need to account for the total informativeness content of the subset rather than just its cardinality. In the following section, we show how appropriate weights can be determined and demonstrate how an estimate of the ME variance can be obtained when the subsets carry different weights.

2.2.1 Determining appropriate weights—Let m_j be the number of AIMs used to estimate individual ancestry estimates obtained on the jth subset. Similarly, let δ_{js} be the informativeness content of the sth marker in the jth subset. The informativeness content is often measured by the delta value, which is the absolute value of the allele frequency difference between two ancestral populations at a marker, or an entropy-based measure like the one described in Rosenberg et al. [17]. Assuming that the total number of AIMs available is divided into p subsets, a simple measure of the weight to assign to the jth subset of markers is given by the following equation.

$$\pi_j = \frac{\sum_{s=1}^{m_r} \delta_{js}}{\sum_{j=1}^{p}\sum_{s=1}^{m_s} \delta_{js}} \quad (7)$$

2.2.2 Weighted measures of reliability—We should note that the parallel measurement assumption [12;18] that underlies the derivation of Cronbach's alpha as a measure of reliability is no longer valid when the individual ancestry proportions (the items) carry different weights. However, various estimates of Cronbach's with unequal weights exist in the literature [19].

The simplest estimate is obtained by considering the weighted average of the W_{ij}'s. That is, $\overline{W}_i = \sum_{i=1}^{p} \pi_j W_{ij}$, where $\pi_j \geq 0$ and $\sum_{i=1}^{p} \pi_j = 1$. In this case, an estimate of Cronbach's alpha is given by

$$\alpha_{prop} = \frac{p}{p-1}\left(1 - \frac{{}^t\pi\, Diag(V)\pi}{{}^t\pi V \pi}\right) \quad (8)$$

Where $diag(V)$ is the diagonal matrix of the observed variances of W, and α_{prop} denotes the estimate of Cronbach's when the items are assigned different weights. Note that the weighted version of Cronbach's alpha may violate the underlying assumptions. For example, when different weights are assigned to different items, it does not make sense to assume that the variance of the ME is constant. However, Cronbach's alpha still provides useful results, which explains its appeal and is what makes it so widely used in practice. In fact, we are considering two possible estimates of the ME variance, because equation (6) yields two possible estimates: one using equation 5 (same weights) and another using equation 8 (different weights).

Another approach would be to consider Armor's theta [20], which is a special case of Cronbach's alpha where the weights π_i are chosen such that the reliability estimate is maximized. This maximization is realized by selecting the eigenvector corresponding to the largest eigenvalue of the correlation matrix computed from the individual ancestry proportion estimates. Armor showed that this estimate can be written as

$$\theta = \frac{p}{p-1}\left(1 - \frac{1}{\lambda_1}\right) \quad (9)$$

where λ_1 is the largest eigenvalue of the correlation matrix computed from the items, which in this case is the correlation matrix between the different sets of individual ancestry proportion estimates.

2.3 Estimating the ME variance by using the repeated measurement approach

The repeated measurement approach may be the most widely used approach to estimate the ME variance. Assume that the ME model can again be written as in equation (1). Under the assumption that the measurement errors are independent of the true ancestry proportions and $Var(U_j) = \sigma^2_{U,j}$, we have from equation (1) that

$$Var(W_{ij}) = \sigma^2_X + \sigma^2_{U,j}, \quad (10)$$

after relaxing the homoscedasticity assumption to allow the ME variance to vary with the informativeness content of the subset of markers that is used to estimate individual ancestry. There are several ways to estimate the ME variance by using repeated data. The simplest approach consists in estimating the measurement error variance from each subset and pooling the results into a single estimate. This is in fact the approach taken by Carroll et al. [8] (page 71, equation 4.3). However, they assume that the ME variance is constant in all subsets. Other estimates can be defined by relaxing the assumption made in equation (2) such that one now assumes that

$$Var(U_j) = \sigma^2_{uj} = \frac{\sigma^2_u}{\pi_j} \quad (11)$$

such that $Var(U_j) < Var(U_l)$ if subset j is more informative than subset l. Note that π_j is given by equation (7).

Let $\overline{W}_i = \sum_{j=1}^{p} \pi_j W_{ij}$ be the overall estimate of the ancestry proportion of the ith individual obtained by combining the p estimates (one for each subset) into a single estimate of ancestry proportions. As mentioned above, $\overline{W}_i$ is the estimate that will be used as a covariate in the association test. It is highly correlated with W_i, the estimate that would be obtained if all the AIMS were used to produce a single estimate. The value of using $\overline{W}_i$ instead of W_i is that unlike W_i, an estimate of the ME variance can be computed for $\overline{W}_i$. In fact, it can be shown that $\overline{W}_i$ is the minimum variance unbiased estimator of X_i. An estimate of the ME variance is given by

$$\hat{\sigma}^2_U = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p} \lambda_j (W_{ij} - \overline{W}_i)^2}{n(p-1)} \quad (12)$$

2.4 Reliability estimates for multivariate IAPE

Multivariate estimates of genetic background can arise with both individual ancestry proportions and principal components. If the admixed population is derived from $k > 2$ ancestral populations, the estimate of individual ancestry proportions can be represented by a vector with $k = k - 1$ elements because the k estimates sum to 1. Caribbean Hispanics, for example, exhibit various levels of Native American (mostly Taínos), European, and African ancestry [21]. Therefore, their ancestry proportion can be represented by a vector with 2 components. It should also be noted that a negative correlation is expected between these components because of the linear constraint. It is also not uncommon to use more than one principal component to achieve the appropriate type I error control in a GWAS. In fact, most investigators use the default setting of EIGENSTRAT, which suggests that the first 10 principal components are included as covariates in their association tests. The reliability of the sample eigenvector as an estimate of the true eigenvector in the population can also be a useful measure. Paul [11] provides a method to compute this reliability for the eigenvector corresponding to the largest eigenvalue. Kimmel et al. [22] show how a non-zero angle between true and estimated axes of variation can lead to type I error inflation in GWAS. Measurements errors in the estimation of the genetic background variable can occur in both cases.

We focus on estimating the ME variance when individual ancestry proportion estimates are computed by using AIMs when $k > 2$. Previous papers on estimating reliability in the multivariate case have focused on relating s observed variables with t latent factors [23].

We propose the following ME model

$$W_{ij} = X_i + U_{ij}; j = 1, 2, .., p \quad (13)$$

where W_{ij} is a $(k, 1)$ vector of ancestry proportion estimates, X_i is the true but unobserved $(k, 1)$ vector of ancestry proportions, and U_{ij} is the $(k, 1)$ vector of Measurement errors associated with the estimation of W_{ij}. Note the we also assume that (1) $U_{ij} \sim MVN(0, \sum_j)$, (2) X_i and U_{ij} are independent, (3) U_{ij} is independent of U_{il} for any two subsets l and j, (4) and that $\sum_j = \left(\frac{1}{\pi_j}\right) \sum$, where π_j is defined similarly as in equation (7). Similarly to the univariate case, these assumptions lead to $Cov(W_{ij}, W_{il}) = Cov(X_i, X_i) = \sum_X$ except that $\sum_X$ is now a (k, k) positive definite matrix.

Equation (13) implies that

$$Cov(W_{ij}) = \sum_X + \sum_j, j = 1, 2, .., p \quad (14)$$

Similarly to the univariate case, the $(k, 1)$ vector

$$\overline{W}_i = \sum_{j=1}^{p} \pi_j W_{ij} \quad (15)$$

will be used as a covariate in the association test instead of W_i. As mentioned above, $\overline{W}_i$ is almost always perfectly correlated with W_i, and contrary to W_i an estimate of the ME variance can be computed for $\overline{W}_i$. Combining (14) and (15), and using the fact that under the assumptions of the CTM model, it can be shown that $Cov(W_i, W_j) = \sum_X$, we have

$$Cov(\overline{W}_i) = \sum_X + \sum_{j=1}^{p} \lambda_j^2 \sum_j \quad (16)$$

2.4.1 Estimating the ME covariance by using the repeated measurement approach—

There are a number of ways to estimate the ME covariance by using repeated data. However, we opted to use an estimate similar to (15) in which the variable W_{ij} is now a vector with k elements. The repeated measurement estimate of the ME covariance matrix is given by

$$\widehat{\sum}_{RM} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p} \pi_j (W_{ij} - \overline{W}_i)(W_{ij} - \overline{W}_i)^T}{n(p-1)} \quad (17)$$

2.4.2 Estimating the ME covariance by using the reliability approach—

Following Tarkkonen and Vehkalahti [23], we can write the reliability matrix as

$$\Omega = \sum_X \left(Cov(\overline{W}) \right)^{-1}. \quad (18)$$

This estimate cannot be computed based on equation (18) alone since the individual ancestry proportion are not directly observed. We know describe a procedure to estimate the reliability matrix. Consider the block matrix $C = Cov(W_1, W_2,...., W_p)^T$, which is a $(k\,p, k\,p)$ matrix, where the submatrix on the ith row and j^{th} column of $Cov(W_1, W_2,...., W_p)$ corresponds to a (k, k) matrix that is defined as:

$$C_{lj} = \begin{cases} \pi_j^2 n^{-1} \sum_{i=1}^{n} (W_{ij} - \overline{W}_i)(W_{ij} - \overline{W}_i)^T, l = j \\ \pi_l \pi_j n^{-1} \sum_{i=1}^{n} (W_{ij} - \overline{W}_i)(W_{ij} - \overline{W}_i)^T, l \neq j \end{cases}$$

where W_{ij} and W_{il} are the $(k, 1)$ vectors of ancestry proportion estimated using markers on the jth and l^{th} subsets.

An estimate of the reliability matrix is given by

$$\widehat{\Omega} = \frac{p}{p-1} \left(\left(\sum_{l \neq j}^{p} C_{lj} \right) \left(\sum_{j=1}^{p}\sum_{l=1}^{p} C_{lj} \right)^{-1} \right) \quad (19)$$

and an estimate of the ME covariance matrix is

$$\widehat{\sum}_{Rel} = (I_p - \widehat{\Omega})\widehat{C}ov(\overline{W}). \quad (20)$$

$\hat{C}ov(W)$ is the observed (p, p) covariance matrix of the overall estimate of the ancestry proportion computed in equation 19 for the n individuals in the sample.

In the case in which the estimated ME covariance matrix is not a positive definite matrix, a simple correction consists of replacing it by another matrix whose minimum eigenvalue is equal to a very small positive value. Let be the minimum eigenvalue of , where denotes the ME covariance matrix estimated by using either equation (17) or equation (20). If 0, then is not a positive definite matrix. However, the matrix $_{corrected} = + (0.01 -)I_{k-1}$, where I_{k-1} is the $(k-1) \times (k-1)$ identity matrix. The minimum eigenvalue of this new matrix is 0.01, which makes a positive definite matrix.

3. Simulation study

We present 2 sets of simulation studies. First, we assumed an admixed population derived from intermating between exactly 2 ancestral populations. In this case, an individual's ancestry proportion estimate can be represented by a scalar and the confounder is univariate. We also assumed that the total number of AIMs is divided into p subsets with m_j, $j = 1, 2, \ldots,$ p being the number of markers used to obtain the admixture estimate on the jth subset. Also,

let $M = \sum_{j=1}^{p} m_j$ is the total number of markers simulated. For simplicity, we set $p=22$ to mimic the more intuitive case in which each estimate is computed at the autosome level. We considered 2 cases: (1) $M=110$ and (2) $M=220$. In each case, we compared the estimate of the ME variance obtained with Cronbach's alpha to the estimate obtained when we used the repeated measurements approach. These results were compared by assuming (a) an equal allocation of the number of markers per subset and (b) an allocation proportional to the chromosome length. Weights under scenario (b) were determined by using the Marshfield sex-averaged chromosome lengths [24].

Let p_{rs} be the allele frequency of the sth marker $s = 1, 2, \ldots, M$ in the rth ancestral population $r = 1, 2$. The allele frequencies were chosen such that each marker was ancestry informative. In practice, a minimum absolute difference of 0.3 between allele frequencies in the 2 ancestral populations is required before a marker can qualify as an AIM. The true ancestry proportions, which is denoted by a_i, $i = 1, 2, \ldots, n$ were drawn from a beta distribution ($a_i \sim Beta(10, 40)$) such that the expected value was 0.2, which is close to estimates of the European genetic contribution to the African-American population [25]. We then computed the allele frequency at the sth marker for the ith individual as $q_{is} = a_i p_{1s} + (1 - a_i)p_{2s}$, which served as the parameter for the binomial distribution from which the individual genotype at that marker is drawn. Finally, we applied a maximum likelihood approach [26] to provide both a global and p IAPEs (one of each subset). The difference between the true ancestry proportions, which were drawn from the Beta distribution, and the global ancestry proportion estimates was computed and regarded as the ME variable. In practice, this difference is never observed because the true ancestry proportions are not known. However, computation of this difference allowed us to directly compute the ME variance and to evaluate the performance of each approach. The ME variance was estimated under 4 different scenarios. (1) The total number of markers was divided equally into 22 subsets of markers, and we applied both methods assuming the IAPEs obtained on each subset carried the same weight. (2) The total number of markers was again divided into 22 subsets, but marker allocation was done proportional to chromosome length. (3) The total number of markers was divided equally into 4 subsets. (4) The total number of markers was divided into 4 subsets with proportional allocation where 10% of markers were allocated to the first subset, 20% to the second, 30% to the third, and 40% to the last. We generated 10,000 data sets containing 1000 individuals each. The simulation results are summarized in Tables 1 to 4.

The purpose of the second simulation study was to evaluate the performance of the two ME variance estimation approaches when the confounder is multivariate, that is, when the

number of ancestral populations (k) is greater than 2. For simplicity, we focus on the case in which $k=3$. In this case, individual ancestry proportion estimates can be represented by a vector with 2 components. The simulation procedure when $k>3$ is similar to the one we just described above for the case in which $k=2$, except that the true ancestry proportions were drawn from a Dirichlet instead of a Beta distribution. Again, we assumed that the total number of AIMs (M) is divided into p subsets with m_j, $j = 1,2,\ldots, p$ being the number of markers used to obtain the ancestry estimate on the jth subset. The performance of the two ME variance estimation approaches was evaluated again by using the 4 scenarios described above.

Let p_{rs} be the allele frequency of the sth marker, $s = 1,2,\ldots, M$ in the rth ancestral population $r = 1,2,3$. Following Pfaff et al. [27], the allele frequency of the sth marker in the admixed population can be written as

$$p_s^{(adx)}=m_1 p_{1s}^{(1)}+m_2 p_{2s}^{(2)}+m_3 p_{3s}^{(3)} \quad (21)$$

where m_r is the genetic contribution of the rth ancestral population. Note that m_r is such that $\sum_{r=1}^{3} m_r = 1$. The genotype at each marker was simulated by two independent draws from a *Bernoulli* (p_s). The true ancestry proportions were drawn from a Dirichlet distribution with parameters (0.4, 0.3, 0.3), which is close to the ancestry proportions observed in Caribbean Hispanics such as Puerto Ricans [28]. We should note that since we generated markers to estimate the IAPE, the distribution of the ME variable was not known. The ME variance was computed for each simulation as the difference between the simulated true ancestry proportions and the overall estimate computed by combining all the subsets into a single data set.

4. Results

We note that both approaches for estimating the ME variance implicitly assume that the average IAPE computed over the p subsets is the control variable that is used in the association test to guard against spurious associations. We showed previously [9] that the correlation between the average estimates computed over the p subsets and the estimate obtained when all the markers are used to provide a single estimate was around 99% when we considered a real data set with over 6,000 individuals in which 1,312 AIMs based on the marker panel described in [29] were typed.

The simulation results for the case in which $k=2$ are given in Tables 1 to 4. These tables present the five-point summary of the estimate of the ME variance computed with each method. The distribution of these estimates was then compared with the true ME variance to evaluate their performance. These simulations showed that the estimates of the ME variance based on the internal reliability measures (i.e., Cronbach's alpha and Armor's theta in the univariate case) seemed to outperform those computed by using the repeated measurement approach. Considering the estimates based on the internal reliability approach, we see that the error variance estimate obtained with Armor's theta was always between the estimates computed by using Cronbach's alpha under the two weighting schemes. This result should be expected because Armor's theta corresponds to the maximum estimate of Cronbach's, which is obtained by considering all possible weightings of the IAPE computed for each subset. As can be seen in Tables 1 to 4, the estimate with Armor's theta was always closer to the estimate provided by Cronbach's alpha under equal allocation of markers. This result is in agreement with equation (6), because Armor's theta yields the maximal value of Cronbach's alpha when the weights are all equal. This result suggests that there might be an

advantage to weighting the items before computing Cronbach's alpha. The estimates based on the repeated measurement had less bias when the markers were allocated equally.

We also note that the performance of the ME variance estimated by using Cronbach's alpha varied with the total number of markers. That is, when we considered 110 AIMS, the allocation proportional to subset size worked better then when the markers were equally allocated to each subset. However, the reverse situation was observed when we considered 220 markers. Very few markers were assigned to the smaller subsets when only 110 markers were used. Therefore, the IAPE calculated from these subsets can have strong bias, which leads to higher variance estimates. The repeated measurements approach, as expected, was more affected by these biases than were the internal reliability measures. When we considered 220 AIMS, there seemed to be enough information such that not much was gained by allocating markers proportional to subset size. Finally, the number of subsets considered appeared to be a significant predictor of the performance of these methods. The repeated measurement approach with equal weights yielded acceptable results when 22 subsets were considered, but failed when only 4 subsets were used, with estimates that were more than twice larger than the true ME variance.

We observed similar results when we considered the case in which the number of ancestral populations was equal to 3. Because the IAPEs must sum to 1, the ME covariance matrix can be represented by a (2,2) matrix. Comparisons between each ME covariance estimation approach and the true ME covariance are presented in Tables 5 and 6. These tables show the error covariance matrix estimated after 10,000 replications. The elements of each covariance matrix are shown in bold. We present the standard error associated with the estimation of each element in italic below this element. We focus on the case in which the IAPEs were computed with 110 AIMs in Table 5. Table 6 is very similar to Table 5, except that it presents the results for the case in which 220 AIMs were used to estimate the IAPEs. Each table also compares the effect of equal vs. allocation proportional to subset size as well as the effect of dividing the total number of AIMs in 22 vs. 4 subsets. Both tables confirm the observations made in the univariate case. That is, (1) the reliability approach leads to more accurate and less biased estimates of the true covariance matrix than the repeated measurement approach, and (2) the allocation of markers proportionally to the number of markers in the subset appears to be beneficial only when 110 AIMs are considered. Table 7 presents the Frobenius norm, the L-2 norm for a matrix, between the estimated covariance matrices and the true covariance matrix is the 4 scenarios shown in Tables 5 and 6. This table confirms that the internal reliability approach seemed to perform better in the context of estimating the ME associated with the estimation of individual ancestry proportions.

5. Discussion

Population stratification and admixture-induced linkage disequilibrium remain a concern in genetic association tests. These tests are often conducted under the assumption that the measure of genetic background that is used to control for the confounding effect caused by population stratification and admixture is obtained without errors. This assumption is not likely to hold, however. Consequently, the type I error and the power of these studies are not likely to remain at their nominal levels. We presented here a simple procedure that can be applied in conjunction with well known ME correction methods to help to account for these Measurement errors.

We show that a generalization of Cronbach's alpha can be used to estimate the ME covariance matrix. This estimate performed better than the one obtained by using the repeated measurement approach. We note that repeated measurements are the most widely used approach to estimating the ME variance in ME correction efforts. The reliability

approach relies on several assumptions. First, it is assumed that the estimate computed on each subset is measuring the same underlying latent variable, which in this case is the true individual ancestry proportion. This assumption is likely to hold when the subsets are created so that they each yield an estimate of the overall ancestry proportion. We note that this assumption is likely violated, however, in efforts to estimate local ancestry because of local variations in the ancestry estimate. We did not evaluate the robustness of this approach under this type of violation and do not advocate its application in efforts to estimate the ME associated with local ancestry estimates. Second, we note that Carroll et al. (page 3, ref #8) for example, referred to the error contaminated variables W and the ME variable U as conditional distributions, where the conditioning is done on the true unobserved variable (X). In this case, $U_{ij}|X_i$ and $U_{il}|X_i$ will be independent assuming that the subsets of AIMS used to estimate the individual admixture proportions are disjoint and they are selected far enough from each other such that admixture induced LD is the only source of correlation. By conditioning on the true ancestry, the admixture induced LD is eliminated. We ran a simulation study to evaluate the magnitude of this correlation. We only consider the univariate case for simplicity and apply the same simulation procedure described in the manuscript in section 3. Briefly, we draw the true individual ancestry (X_i) from a beta distribution, we simulate 2 sets (M = 5,10) of ancestry informative markers conditional on the underlying ancestry proportion and compute admixture proportion estimate for each set (W_{i1} and W_{i2}), we then obtained the ME variables (U_{i1} and U_{i2}) as the difference between W_{i1} and W_{i2} and X_i and compute the correlation between them. We chose to presentation the correlations instead of the covariances simply for ease of interpretation. The average correlation after a 1000 iterations was 8×10^{-4} with a standard error of 5×10^{-3} for M=5 and 2×10^{-3} with a standard error of 3×10^{-3} for M=10. The simulation showed that these correlations are very small, which implies that the assumption of independence is not too strong to invalidate the proposed methods. The correlation between W_{ij} and W_{il} is also always positive, which means that the covariance between these two variables is also always positive.

Third, a closer look at equation 12 shows that the repeated measurement estimate of the ME variance does not make use of these assumptions. Therefore, the resulting estimate should not be affected if these assumptions do not hold. However, the estimates based on the reliability approach may be biased and equations (3) and (16) would have to be modified to account for the correlation in the Measurement errors.

We considered two possible ways of partitioning the AIMs into subsets. The first partition consisted of dividing the AIMs into 22 subsets. This partition can be seen as a natural way of dividing the data when an investigator seeks to obtain an estimate for each chromosome. In the second partition, we divided the AIMs into 4 subsets. We chose this partition to evaluate whether the second method would perform better when fewer subsets with a larger number of AIMs were considered. We observed an advantage of considering fewer but larger subsets only in the case in which 110 markers were used.

In conclusion, our results offer information that can be used to enhance the estimation of the degree of ME or, conversely, the reliability of individual ancestry proportion estimates. The estimate of ME variance can in turn be incorporated into other analyses, such as genetic association tests in candidate gene and genome-wide association studies.

Acknowledgments

References

1. Pritchard JK, Stephens M, Donnelly PJ. Correcting for population stratification in linkage disequilibrium mapping studies. American Journal of Human Genetics. 1999; 65:A101.

2. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. American Journal of Human Genetics. 1999; 65:220–228. [PubMed: 10364535]

3. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. American Journal of Human Genetics. 2000; 67:170–181. [PubMed: 10827107]

4. Pritchard JK, Donnelly P. Case-control studies of association in structured or admixed populations. Theoretical Population Biology. 2001; 60:227–237. [PubMed: 11855957]

5. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–909. [PubMed: 16862161]

6. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. Plos Genetics. 2006; 2:2074–2093.

7. Redden D, Divers J, Vaughan L, Tiwari H, et al. Regional admixture mapping and structured association testing: conceptual unification and an extensible general linear model. Plos Genetics. 2006; 2:1254–1264.

8. Carroll, RJ.; Rupper, D.; Stefanski, LA.; Crainiceanu, CM. ME in nonlinear models a modern perspective. 2. Boca Raton, FL: Chapman & Hall/CRC; 2006.

9. Divers J, Vaughan LK, Padilla MA, Fernandez JR, et al. Correcting for ME in individual ancestry estimates in structured association tests. Genetics. 2007; 176:1823–1833. [PubMed: 17507670]

10. Padilla MA, Divers J, Vaughan LK, Allison DB, et al. Multiple imputation to correct for ME in admixture estimates in genetic structured association testing 1. Hum Hered. 2009; 68:65–72. [PubMed: 19339787]

11. Paul D. Asymptotics of sample eigenstruture for a large dimensional spiked covariance model. Statistica Sinica. 2007; 17:1617–1642.

12. Cronbach L. Coefficient alpha and the internal structure of tests. Psychometrika. 1951; 16:297–334.

13. Allen, MJ.; Yen, WM. Introduction to measurement theory. Monterey, CA: Brooks/Cole Pub. Co; 1979.

14. Novick MR, Lewis C. Coefficient alpha and the reliability of composite measurements. Psychometrika. 1967; 32:1–13. [PubMed: 5232569]

15. Crocker, LM.; Algina, J. Introduction to classical and modern test theory. New York, NY: Holt, Rinehart, and Winston; 1986.

16. Miller, Michael B. Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. Structural Equation Modeling. 1995; 2:255–273.

17. Rosenberg NA, Li LM, Ward R, Pritchard JK. Informativeness of genetic markers for inference of ancestry. American Journal of Human Genetics. 2003; 73:1402–1422. [PubMed: 14631557]

18. Kuder GW, Richardson MW. The theory of the estimation of test reliability. Psychometrika. 1937; 2:151–160.

19. Greene VL, Carmines EG. Assessing the Reliability of Linear Composites. Sociological Methodology. 1980; 11:160–175.

20. Armor DJ. Theta Reliability and Factor Scaling. Sociological Methodology. 1974; 5:17–50.

21. Tang H, Choudhry S, Mei R, Morgan M, et al. Recent Genetic Selection in the Ancestral Admixture of Puerto Ricans. The American Journal of Human Genetics. 2007; 81:626–633.

22. Kimmel G, Jordan MI, Halperin E, Shamir R, et al. A Randomization Test for Controlling Population Stratification in Whole-Genome Association Studies. The American Journal of Human Genetics. 2007; 81:895–905.

23. Tarkkonen L, Vehkalahti K. Measurement errors in multivariate measurement scales. Journal of Multivariate Analysis. 2005; 96:172–189.

24. Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, et al. A high-resolution recombination map of the human genome. Nat Genet. 2002; 31:241–247. [PubMed: 12053178]

25. Parra EJ, Marcini A, Akey L, Martinson J, et al. Estimating African American admixture proportions by use of population-specific alleles. American Journal of Human Genetics. 1998; 63:1839–1851. [PubMed: 9837836]

26. Tang H, Peng J, Wang P, Risch NJ. Estimation of individual admixture: Analytical and study design considerations. Genetic Epidemiology. 2005; 28:289–301. [PubMed: 15712363]

27. Pfaff CL, Barnholtz-Sloan J, Wagner JK, Long JC. Information on ancestry from genetic markers. Genetic Epidemiology. 2004; 26:305–315. [PubMed: 15095390]

28. Bonilla C, Shriver MD, Parra EJ, Jones A, et al. Ancestral proportions and their association with skin pigmentation and bone mineral density in Puerto Rican women from New York city. Human Genetics. 2004; 115:57–68. [PubMed: 15118905]

29. Smith MW, Patterson N, Lautenberger JA, Truelove AL, et al. A high-density admixture map for disease gene discovery in African Americans. American Journal of Human Genetics. 2004; 74:1001–1013. [PubMed: 15088270]

Appendix

A.- Table of abbreviations

ME	Measurement error
RM	Repeated measurement
SATs	Structured association tests
IAPE	Individual admixture proportion estimates
AIMs	Ancestry informative markers
PCA	Principal component analysis
Linkage disequilibrium	LD
GWAS	Genome wide association studies
CTM	Classical true score model

B.- Proof that equation (12) is unbiased

Let

$$W_{ij} = X_i + U_{ij} \quad (21)$$

$$\text{Assume that} \begin{cases} E(U_{ij}) = 0 \; \forall j \\ Var(U_{ij}) = \sigma_j^2 \end{cases}$$

Let π_j be a set of weights such that $\pi_j \geq 0$ and $\sum_{j=1}^{p} \pi_j = 1$ where p is the number of subsets of AIMs.

Define

$$\overline{W}_i = \sum_{j=1}^{p} W_{ij} \quad (22)$$

Let

$$\sigma_j^2 = \frac{\sigma^2}{\pi_j} \quad (23)$$

Therefore, $\overline{W_i}$ is minimum variance unbiased linear estimator $E(X_i)$. From equation (1) we have $E(W_{ij}) = E(X_i)$

Note that from (1) and (2), we also have

$$\overline{W}_i = X + \overline{U}_i \quad (24)$$

where $\overline{U}_i = \sum_{j=1}^{p} U_{ij}$

Subtracting (3) from (1) we have

$$W_{ij} - \overline{W}_i = U_{ij} - \overline{U}_i \quad (25)$$

which leads to

$$\sum_{j=1}^{p} \pi_j (W_{ij} - \overline{W}_i)^2 = \sum_{j=1}^{p} \pi_j (U_{ij} - \overline{U}_i)^2 = \sigma^2 \sum_{j=1}^{p} \frac{(U_{ij} - \overline{U}_i)^2}{\sigma_j^2} \quad (26)$$

$$E \left\{ \sum_{j=1}^{p} \pi_j (W_{ij} - \overline{W}_i)^2 \right\} = \sigma^2 E \left\{ \sum_{j=1}^{p} \frac{(U_{ij} - \overline{U}_i)^2}{\sigma_j^2} \right\} \quad (27)$$

$$E \left\{ \sum_{j=1}^{p} \pi_j (W_{ij} - \overline{W}_i)^2 \right\} = \sigma^2 (p - 1) \quad (28)$$

Finally,

$$E(\hat{\sigma}^2) = \frac{1}{n(p-1)} \sum_{i=1}^{n} E \left\{ \sum_{j=1}^{p} \pi_j (W_{ij} - \overline{W}_i)^2 \right\} = \sigma^2.$$

The same approach can be taken to show that equation 17 also provides an unbiased estimate of the ME variance when k>2.

Table 1

Five-point summary and standard error of the distribution of the measurement error variance associated with ancestry proportion estimates for *K*=2 when 110 AIMS divided into 22 subsets are used.

Estimates	Minimum	Q1	Q2	Mean	St. Error	Q3	Maximum
True error variance	1.988	2.596	2.742	2.755	0.225	2.901	3.771
Cronbach's alpha (prop)	2.222	2.602	2.701	2.706	0.150	2.804	3.314
Cronbach's alpha (equal)	1.931	2.275	2.363	2.370	0.138	2.460	2.879
Armor's theta	1.816	2.170	2.252	2.257	0.127	2.340	2.765
Repeated measurements (prop)	2.616	3.247	3.418	3.435	0.268	3.608	4.715
Repeated measurements (equal)	2.236	2.785	2.909	2.918	0.192	3.044	3.785

Note: All numbers in the table should be divided by 1000 to get back to the original scale.

Table 2

Five-point summary and standard error of the distribution of the measurement error variance associated with ancestry proportion estimates for $K=2$ when 110 AIMS divided into divided into 4 subsets are used

Estimates	Minimum	Q1	Q2	Mean	St. Error	Q3	Maximum
True error variance	2.025	2.594	2.740	2.751	0.224	2.893	3.772
Cronbach's alpha (prop)	2.390	2.768	2.855	2.861	0.136	2.950	3.543
Cronbach's alpha (equal)	2.553	2.989	3.099	3.109	0.176	3.222	3.885
Armor's theta	2.537	2.933	3.035	3.042	0.161	3.147	3.691
Repeated measurements (prop)	2.967	4.116	4.629	4.908	1.152	5.405	14.210
Repeated measurements (equal)	3.654	5.143	5.821	6.089	1.330	6.757	16.540

Note: All numbers in the table should be divided by 1000 to get back to the original scale.

Table 3

Five-point summary and standard error of the distribution of the measurement error variance associated with ancestry proportion estimates for *K*=2 when 220 AIMS divided into 22 subsets are used.

Estimates	Minimum	Q1	Q2	Mean	St. Error	Q3	Maximum
True error variance	1.057	1.302	1.361	1.363	0.090	1.421	1.773
Cronbach's alpha (prop)	1.312	1.527	1.582	1.587	0.086	1.643	2.019
Cronbach's alpha (equal)	1.156	1.302	1.340	1.341	0.057	1.379	1.585
Armor's theta	1.129	1.268	1.303	1.304	0.054	1.340	1.549
Repeated measurements (prop)	1.493	1.790	1.878	1.891	0.143	1.981	2.564
Repeated measurements (equal)	1.290	1.501	1.551	1.556	0.079	1.606	1.951

Note: All numbers in the table should be divided by 1000 to get back to the original scale.

Table 4

Five-point summary and standard error of the distribution of the measurement error variance associated with ancestry proportion estimates for K=2 when 220 AIMS divided into 4 subsets

Estimates	Minimum	Q1	Q2	Mean	St. Error	Q3	Maximum
True error variance	1.083	1.301	1.359	1.362	0.090	1.421	1.767
Cronbach's alpha (prop)	1.704	1.917	1.964	1.966	0.073	2.014	2.241
Cronbach's alpha (equal)	1.595	1.786	1.828	1.830	0.066	1.874	2.088
Armor's theta	1.591	1.783	1.825	1.828	0.066	1.871	2.087
Repeated measurements (prop)	1.999	2.649	3.062	3.277	0.888	3.670	12.260
Repeated measurements (equal)	1.982	2.698	3.101	3.284	0.813	3.679	9.135

Note: All numbers in the table should be divided by 1000 to get back to the original scale.

Table 5

Performance of the repeated measurement approach and multivariate reliability when $k=3$ and 110 AIMs are used to estimate the individual ancestry proportions

Number of subsets	True measurement error covariance		Repeated Measurement				Multivariate Reliability			
			Allocation equal		Allocation proportional		Allocation equal		Allocation proportional	
22	**0.0083**	**−0.0044**	**0.0036**	**−0.0020**	**0.0050**	**−0.0027**	**0.0061**	**−0.0034**	**0.0073**	**−0.0040**
	0.0007	*0.0006*	*0.0005*	*0.0007*	*0.0024*	*0.0018*	*0.0001*	*0.0002*	*0.0002*	*0.0002*
	−0.0044	**0.0047**	**−0.0020**	**0.0030**	**−0.0027**	**0.0055**	**−0.0034**	**0.0051**	**−0.0040**	**0.0064**
	0.0006	*0.0006*	*0.0007*	*0.0007*	*0.0018*	*0.0019*	*0.0002*	*0.0002*	*0.0002*	*0.0002*
4	**0.0083**	**−0.0044**	**0.0097**	**−0.0051**	**0.0032**	**−0.0017**	**0.0076**	**−0.0040**	**0.0077**	**−0.0040**
	0.0007	*0.0006*	*0.0013*	*0.0011*	*0.0002*	*0.0002*	*0.0005*	*0.0005*	*0.0003*	*0.0003*
	−0.0044	**0.0047**	**−0.0051**	**0.0066**	**−0.0017**	**0.0020**	**−0.0040**	**0.0050**	**−0.0040**	**0.0048**
	0.0006	*0.0006*	*0.0011*	*0.0012*	*0.0002*	*0.0002*	*0.0005*	*0.0005*	*0.0003*	*0.0002*

Table 6

Performance of the repeated measurement approach and multivariate reliability when $k=3$ and 220 AIMs are used to estimate the individual ancestry proportions

Number of subsets	True measurement error covariance		Repeated Measurement				Multivariate Reliability			
			Allocation equal		Allocation proportional		Allocation equal		Allocation proportional	
22	**0.0050**	**−0.0028**	**0.0024**	**−0.0013**	**0.0042**	**−0.0023**	**0.0041**	**−0.0022**	**0.0046**	**−0.0025**
	0.0005	*0.0005*	*0.0003*	*0.0004*	*0.0006*	*0.0008*	*0.0001*	*0.0001*	*0.0002*	*0.0002*
	−0.0028	**0.0029**	**−0.0013**	**0.0017**	**−0.0023**	**0.0036**	**−0.0022**	**0.0029**	**−0.0025**	**0.0035**
	0.0005	*0.0005*	*0.0004*	*0.0004*	*0.0008*	*0.0008*	*0.0001*	*0.0001*	*0.0002*	*0.0002*
4	**0.0050**	**−0.0027**	**0.0019**	**−0.0010**	**0.0036**	**−0.0019**	**0.0047**	**−0.0025**	**0.0045**	**−0.0023**
	0.0006	*0.0005*	*0.0001*	*0.0001*	*0.0006*	*0.0005*	*0.0002*	*0.0002*	*0.0002*	*0.0002*
	−0.0027	**0.0029**	**−0.0010**	**0.0011**	**−0.0027**	**0.0029**	**−0.0025**	**0.0028**	**−0.0023**	**0.0026**
	0.0005	*0.0005*	*0.0001*	*0.0001*	*0.0005*	*0.0005*	*0.0001*	*0.0001*	*0.0001*	*0.0001*

Table 7

Frobrenius norm between the estimated covariance using each approach and true error covariance

Number of subsets	Subset	Repeated Measurement		Multivariate Reliability	
		Equal	Proportional	Equal	Proportional
220	22	0.0036	0.0013	0.0012	0.0008
	4	0.0044	0.0017	0.0005	0.0009
110	22	0.0061	0.0041	0.0026	0.0020
	4	0.0025	0.0069	0.0009	0.0007