PSYCHOMETRIKA—VOL. 52, NO. 4, 521–538
DECEMBER 1987

MEASUREMENT ERROR AND ANCOVA: FUNCTIONAL AND STRUCTURAL RELATIONSHIP APPROACHES

JEROEN G. W. RAAIJMAKERS

TNO INSTITUTE FOR PERCEPTION
SOESTERBERG, THE NETHERLANDS

JO P. M. PIETERS

RIJKS PSYCHOLOGISCHE DIENST
THE HAGUE, THE NETHERLANDS

This article discusses alternative procedures to the standard F-test for ANCOVA in case the covariate is measured with error. Both a functional and a structural relationship approach are described. Examples of both types of analysis are given for the simple two-group design. Several cases are discussed and special attention is given to issues of model identifiability. An approximate statistical test based on the functional relationship approach is described. On the basis of Monte Carlo simulation results it is concluded that this testing procedure is to be preferred to the conventional F-test of the ANCOVA null hypothesis. It is shown how the standard null hypothesis may be tested in a structural relationship approach. It is concluded that some knowledge of the reliability of the covariate is necessary in order to obtain meaningful results.

Key words: ANCOVA, measurement error, functional relationship, structural relationship.

Introduction

The analysis of covariance (ANCOVA) for a single-factor experiment is based on the following linear model for the dependent variable Y_{ij}:

$$Y_{ij} = \mu + \alpha_j + \beta(X_{ij} - \bar{X}) + \varepsilon_{i(j)}, \qquad i = 1, \ldots, n; j = 1, \ldots, c$$

where μ is the overall mean, α_j is the effect due to treatment condition j, β is the parameter of the regression of Y on X (the concomitant variable or covariate), and $\varepsilon_{i(j)}$ refers to the error component. It is important to note that X refers to the *observed* value on the covariate and is assumed to be measured without error. What will be the effect on the ANCOVA test results if the covariate measurements are in fact fallible? As has been recognized by many authors, the answer to this question depends on the exact assumptions made with respect to the linear model. If one assumes that the Y-variable is indeed linearly related to the observed value of the covariate, then the ANCOVA test results will be correct. If, on the other hand, it is assumed that Y is linearly related to the underlying true score on the covariate, the ANCOVA results will no longer be correct and the standard F-test will lead to biased results. A natural question to ask, then, is whether there are feasible alternatives to ANCOVA for such a case that do lead to valid results. In this paper we will present some results that are relevant to this issue.

Suppose the data are in accordance with the following model

$$Y_{ij} = \mu + \alpha_j + \beta(T_{ij} - \bar{T}) + \varepsilon_{i(j)}, \tag{1}$$

Requests for reprints should be addressed to Jeroen G. W. Raaijmakers, TNO Institute for Perception, Kampweg 5, 3769 DE Soesterberg, THE NETHERLANDS.

0033-3123/87/1200-8056$00.75/0
© 1987 The Psychometric Society

and

$$X_{ij} = T_{ij} + \delta_{i(j)}, \tag{2}$$

where T_{ij} denotes the true score on the covariate. It is assumed that the disturbance variables $\varepsilon_{i(j)}$ and $\delta_{i(j)}$ are normally distributed with expectation zero and variances equal to σ_ε^2 and σ_δ^2, respectively. This is similar to a well-known model that has been studied extensively in econometrics and mathematical statistics (see e.g., Kendall & Stuart, 1967) under the heading of linear *functional* and *structural* relationships. The distinction between a functional and a structural relationship is that in the former approach no distributional assumptions are made with respect to the T_{ij}. This implies that in the functional case all T_{ij} are parameters, that is, fixed unknown constants. In the structural case, these T_{ij} are independent, random variables, independent of the disturbance variables $\varepsilon_{i(j)}$ and $\delta_{i(j)}$. It is generally assumed that the T_{ij} are sampled from normal distributions with means μ_j and variances σ_j^2. Note that this distinction is similar to the difference between fixed effects and random effects models in ANOVA.

Specific cases of this approach to ANCOVA have been discussed by Lord (1960) and Sörbom (1978). The present paper, however, not only reviews this work but also examines a number of other cases. We will apply results that have been derived in the statistical and econometric literature on the errors-in-variables model to the present ANCOVA model. Procedures will be discussed for estimating the parameters for such models and methods for testing various hypotheses.

ANCOVA as a Functional Relationship

In the functional relationship approach, the true scores T_{ij} are not assumed to be a random sample from a particular parent distribution, but are considered to be fixed, unknown constants (i.e., parameters). For simplicity, we will restrict ourselves to the two group case with equal n's. In that case we may rewrite (1) and (2) as follows

$$y_{ij} = \mu + \alpha Z_j + \beta(T_{ij} - \bar{T}) + \varepsilon_{i(j)}, \qquad i = 1, \ldots, n; j = 1, 2$$

$$x_{ij} = T_{ij} + \delta_{i(j)}$$

where $Z_1 = 1$ and $Z_2 = -1$.

This implies that the model has $(2n + 5)$ parameters: $T_{11}, T_{21}, \ldots, T_{n1}, T_{12}, \ldots, T_{n2}$, μ, α, β, σ_ε^2, and σ_δ^2. The T_{ij}'s are usually referred to as *incidental* parameters, and the other five parameters are called *structural.* Incidental parameters are specific to individual observations, while the structural parameters are common to sets of observations. The presence of incidental parameters poses a problem in statistical estimation since the standard maximum likelihood (ML) theory of estimation does not apply in this case. Even if we make the assumption that the variance of the T_{ij}'s (and hence the variance of the X_{ij}'s) converges to a fixed value, the ML estimators for the structural parameters are not necessarily consistent. An example of this will be shown when we discuss the ML estimators for the parameters of the functional model for ANCOVA.

It should be noted that under the null hypothesis $\alpha = 0$, the present model reduces to the famous linear functional relationship problem discussed by, for example, Kendall and Stuart (1967) and Anderson (1976, 1984). In this case the likelihood function is given by

$$L = (4\pi^2\sigma_\delta^2\sigma_\varepsilon^2)^{-n} \exp\{-(2\sigma_\delta^2)^{-1}\sum\sum(X_{ij} - T_{ij})^2 - (2\sigma_\varepsilon^2)^{-1}\sum\sum(Y_{ij} - \mu - \beta(T_{ij} - \bar{T}))^2\}.$$

It can be shown that this likelihood function has no maximum (as a function of the parameters). To illustrate this fact, suppose we let $T_{ij} = x_{ij}$. It is easy to see that with this substitution $L \to \infty$ as σ_δ^2 approaches zero. Since the likelihood function has no maximum, maximum likelihood estimators do not exist (see also Anderson, 1984; Anderson & Rubin, 1956). Similar results can be obtained for the present ANCOVA model in case $\alpha \neq 0$. Estimation methods other than the ML-method do exist for this case, for example, Geary's method of using product cumulants (see Kendall & Stuart, 1967). This method however becomes useless as the distribution of the true scores T_{ij} approaches a normal distribution (in that case all cumulants of order ≥ 3 are zero and the equation system used in estimating the parameters becomes unsolvable). Hence, it is to be expected that such a method will not be very useful in practice.

The Assumption of Equal Variances

In order to obtain more meaningful results, we must either impose some restriction on the model or obtain additional information such as knowledge concerning the error variances or the reliability of the covariate. The latter case will be discussed in the next section. The most common identifying restriction that is made in this situation is that the error variances of X and Y are equal, that is, $\sigma_\varepsilon^2 = \sigma_\delta^2$. Note that the more general assumption $\sigma_\varepsilon^2 = \lambda\sigma_\delta^2$ (with λ a known constant), is identical to the present restriction provided that we rescale the observations and the resulting estimates for μ, α and β accordingly.

With this assumption, the following ML-estimators are obtained:

$$\hat{T}_{ij} = \frac{X_{ij} + \hat{\beta}(Y_{ij} - \bar{Y} - \hat{\alpha}Z_j + \hat{\beta}\bar{X})}{1 + \hat{\beta}^2}, \qquad i = 1, \ldots, n; j = 1, 2$$

where $Z_1 = 1$ and $Z_2 = -1$,

$$\hat{\mu} = \bar{Y},$$

$$\hat{\alpha} = (\bar{Y}_1 - \bar{Y}) - \hat{\beta}(\bar{X}_1 - \bar{X}),$$

$$\hat{\sigma}_\varepsilon^2 = \frac{\hat{\beta}^2 W_{xx} - 2\hat{\beta} W_{xy} + W_{yy}}{2N(1 + \hat{\beta}^2)},$$

and

$$\hat{\beta} = \frac{W_{yy} - W_{xx} + \{(W_{yy} - W_{xx})^2 + 4W_{xy}^2\}^{1/2}}{2W_{xy}},$$

provided that $W_{xy} \neq 0$. In these equations W_{xx}, W_{yy}, and W_{xy} denote the pooled within-groups sums of squares and cross-products. N denotes the total number of observations, that is, $N = 2n$.

What can be said about the properties of these estimators? First of all, let us consider the question of the consistency of these estimators. Let us assume that the pooled within-groups variance of T,

$$\sum\sum \frac{(T_{ij} - \bar{T}_j)}{N},$$

converges to a fixed value S_T^2. In that case the sample pooled within-groups variances and covariances converge in probability to their expectations:

$$\frac{W_{xx}}{N} \rightarrow S_T^2 + \sigma_\varepsilon^2,$$

$$\frac{W_{xy}}{N} \rightarrow \beta S_T^2,$$

$$\frac{W_{yy}}{N} \rightarrow \beta^2 S_T^2 + \sigma_\varepsilon^2.$$

Upon insertion of these results into the equation for $\hat{\beta}$, we obtain the result that $\hat{\beta} \rightarrow \beta$, and hence, $\hat{\beta}$ is a consistent estimator for β. The estimator for the error variance σ_ε^2, however, is not consistent but converges to $\sigma_\varepsilon^2/2$. This illustrates the above mentioned fact that in the presence of incidental parameters ML-estimators are not always consistent. In this case, however, the inconsistency can be easily remedied by using $2\hat{\sigma}_\varepsilon^2$ as an estimator for σ_ε^2. In the remainder of this paper we will denote the estimator $2\hat{\sigma}_\varepsilon^2$ by $\hat{\sigma}^2$.

On comparing these estimators with those of the conventional ANCOVA model, we note a number of similarities. First, both $\hat{\beta}$ and b_W (the regression coefficient in ANCOVA) are calculated from the pooled within-groups variance-covariance matrix. Hence, these estimates are not sensitive to the differences between groups. Next, the estimation equations for both $\hat{\mu}$ and $\hat{\alpha}$ are identical to those of the conventional ANCOVA model, provided that $\hat{\beta}$ is substituted for b_W. Therefore, the present analysis corresponds to a conventional ANCOVA analysis provided that in the estimation of the slope of the regression lines the measurement error in the covariate is taken into account. We will make use of this correspondence in the construction of a test statistic for the hypothesis $\alpha = 0$.

Since we are using the ML-method, the most natural test statistic would seem to be the traditional chi-square test based on the likelihood ratio statistic (λ). This approach, however, breaks down due to the presence of incidental parameters. The reason for this is that it is based on the assumption that the number of parameters does not change with sample size. This condition is obviously violated in this case.

This conclusion was verified by Monte-Carlo simulation of the present functional ANCOVA model (see Table 1). These results are based on 5000 simulations of a two-group design with one covariate, using (1) and (2) with $\alpha = 0$. Random number generation was performed with the DEC-Fortran function RAN, which has been shown to be a satisfactory pseudo-random number generator (see Edgell, 1979). As a further check, the results were divided in successive blocks of 1000 simulations each. The resulting distributions of the test statistics to be considered in this paper were then tested for equality. No systematic deviations were observed. Normally distributed values for the error variables were generated using the method described in Box and Muller (1958). A frequency distribution of the obtained values of the test statistic was formed using equally probable class intervals, that is, the class limits were determined from the percentile points of the appropriate chi-square distribution. In addition to the mean covariate difference between the two groups, we also varied n, ρ, the reliability of the covariate, and β. Within each group, the true covariate scores were uniformly distributed. These results show that $-2 \log \lambda$ does not approximate a chi-square distribution, not even when the sample size is quite large.

In order to obtain a meaningful test statistic we evidently have to take a different approach. A possibly fruitful angle to attack this problem is provided by a reconsideration of the test statistic in the ordinary ANCOVA model. It can be shown that in this model the estimator $\hat{\alpha}$ for the treatment effect, is normally distributed with mean α and variance $var(\hat{\alpha})$. By standard methods an estimate $v\hat{a}r(\hat{\alpha})$ of $var(\hat{\alpha})$ may be obtained from the sample results. Hence,

$$t = \frac{\hat{\alpha} - \alpha}{\{v\hat{a}r(\hat{\alpha})\}^{1/2}}$$

follows a t-distribution with $(N - 3)$ degrees of freedom ($N - 3$ since 3 degrees of freedom

TABLE 1

Goodness-of-Fit (χ^2, df=9) of Likelihood Ratio Statistic to Chi-Square Distribution (df=1)

		$\bar{T}_1=\bar{T}_2=0$		$\bar{T}_1=-10,\ \bar{T}_2=10$	
n	ρ	β=1	β=5	β=1	β=5
	.5	2502	2546	4260	2191
10	.7	2492	2535	3046	2135
	.9	2519	2538	2451	2087
	.5	1475	1399	3654	1676
50	.7	1478	1411	2306	1599
	.9	1466	1418	1654	1565
	.5	1363	1397	3484	1624
100	.7	1357	1374	2372	1559
	.9	1355	1378	1670	1479

are used up in the estimation of μ, α, and β). The important result for our purposes is that in the case of the null hypothesis $\alpha = 0$, t^2 is equivalent to the usual F-statistic in ANCOVA.

It happens to be the case that in the functional ANCOVA model, $\hat{\alpha}$ is asymptotically also normally distributed as $N \to \infty$ (assuming the pooled within-groups variance of T converges to a fixed value S_T^2). Hence, we will consider the test statistic

$$t = \frac{\hat{\alpha}}{\{v\hat{a}r(\hat{\alpha})\}^{1/2}}$$

for testing the null hypothesis $\alpha = 0$. The problem is thus reduced to finding a reasonable estimate for $var(\hat{\alpha})$. A first-order approximation (to the order N^{-1}) of this variance is

$$var(\hat{\alpha}) = \frac{\sigma_\varepsilon^2(1 + \beta^2)}{N} + var(\hat{\beta})\,\frac{(\mu_1 - \mu_2)^2}{4} + O(N^{-1}).$$

Furthermore, it can be shown (see also Robertson, 1974) that

$$var(\hat{\beta}) = \frac{\sigma_\varepsilon^2\{(1 + \beta^2)S_T^2 + \sigma_\varepsilon^2\}}{NS_T^4} + O(N^{-1}). \tag{3}$$

Unfortunately, these formulas are large-sample approximations that are not very good with small samples and/or large error variances (as was observed from Monte-Carlo simulations). This is probably related to the fact that the exact distribution of $\hat{\beta}$ has some peculiar characteristics (e.g., infinite moments, see Anderson & Sawa, 1982). In unfavorable circumstances, these formulas severely underestimate the variances obtained from Monte-Carlo simulations. It turns out, however, that a simple correction for bias reduces many of these problems considerably.

It may be shown (see e.g., Robertson, 1974) that the expected value of $\hat{\beta}$ is approxi-

TABLE 2

Goodness-of-Fit (χ^2, df=9) for Proposed t-Statistic in Comparison to Conventional F-Test

		t-statistic				conventional F-test			
		$\bar{T}_1=\bar{T}_2$		$\bar{T}_1\neq\bar{T}_2$		$\bar{T}_1=\bar{T}_2$		$\bar{T}_1\neq\bar{T}_2$	
n	ρ	β=1	β=5	β=1	β=5	β=1	β=5	β=1	β=5
	.5	13.6	30.6	555.1	822.3	54.1	463.9	16975.4	40158.1
10	.7	12.9	20.0	210.6	339.7	33.7	123.6	6574.5	20921.2
	.9	17.6	23.5	47.4	91.2	6.7	13.5	771.1	2621.6
	.5	11.1	15.2	151.9	173.9	60.9	602.3	44940.0	45000.0
50	.7	10.0	16.4	45.2	72.9	40.4	195.7	43263.0	45000.0
	.9	11.5	15.9	12.1	17.5	8.6	21.4	18554.9	35816.3
	.5	7.7	15.9	94.5	136.9	79.4	623.5	45000.0	45000.0
100	.7	6.5	15.0	44.9	65.6	45.6	206.8	44980.0	45000.0
	.9	7.6	14.9	23.5	32.8	11.6	39.2	37251.6	44740.5

Note: Maximum value of chi-square is 45000

mately equal to

$$E(\hat{\beta}) = \beta\left\{1 + \frac{\sigma_\varepsilon^2((1+\beta^2)S_T^2 + \sigma_\varepsilon^2)}{N(1+\beta^2)S_T^4}\right\} + O(N^{-1}).$$

Hence, this formula may be used to obtain an approximately unbiased estimate $\hat{\beta}_c$ for β:

$$\hat{\beta}_c = \frac{\hat{\beta}}{C},$$

where C is given by

$$C = 1 + \frac{\hat{\sigma}^2\{(1+\hat{\beta}^2)\hat{S}_T^2 + \hat{\sigma}^2\}}{N(1+\hat{\beta}^2)\hat{S}_T^4},$$

with

$$\hat{S}_T^2 = \frac{W_{xx}}{N} - \hat{\sigma}^2.$$

Simulation results show that the variance of $\hat{\beta}_c$ is well approximated (even with relatively small sample sizes) by (3). A corrected estimate for α is then given by:

$$\hat{\alpha}_c = (\bar{Y}_1 - \bar{Y}) - \hat{\beta}_c(\bar{X}_1 - \bar{X}). \tag{4}$$

Hence, we conclude that $\hat{\beta}_c$ and $\hat{\alpha}_c$ are approximately normally distributed with mean β and α, respectively, and variances $var(\hat{\beta})$ and $var(\hat{\alpha})$ as given above. As a final step, sample estimates have to be inserted into these formulas to obtain estimated variances for $\hat{\alpha}_c$ and $\hat{\beta}_c$. It turns out that the best approximation is provided by using $\hat{\beta}_c$, $\hat{S}_T^2$ and $N\hat{\sigma}^2/(N-3)$ in these formulas as estimators for, respectively, β, S_T^2 and σ_ε^2. Table 2 gives some results showing how well the resulting test statistic for the hypothesis $\alpha = 0$ approaches a t-

TABLE 3

Numerical Example of the Proposed Test Procedure

Pooled within-group covariance matrix:

	Y	X		
Y	51.025		$\bar{y}_1=-1.55$	$\bar{y}_2=1.25$
X	24.750	14.600	$\bar{x}_1=5.00$	$\bar{x}_2=8.00$
			$n_1=10$	$n_2=10$

Successive steps in calculating test statistic:

$\hat{\beta} = 1.977$	$\hat{\beta}_c = 1.961$
$\hat{\alpha} = 1.566$	$\hat{\alpha}_c = 1.542$
$\hat{\sigma}_\varepsilon^2 = 1.042$	$2N\hat{\sigma}_\varepsilon^2/(N-3) = 2.451$
$\hat{\sigma}^2 = 2\hat{\sigma}_\varepsilon^2 = 2.084$	$\hat{var}(\hat{\beta}) = 0.049$
$\hat{S}_T^2 = 12.516$	$\hat{var}(\hat{\alpha}) = 0.705$

$t(17) = 1.836$

distribution with $(N - 3)$ degrees of freedom. Note that the approximation becomes better as σ_ε^2 decreases and as sample size increases. Although in some cases the approximation cannot be said to be very good, it should be noted that even in such cases the present test statistic is still always quite superior to the usual F-test. Hence, we may conclude that *the present test statistic is uniformly superior to the traditional F-test.* Of particular interest is the fact that the new test statistic is also superior when the true covariate means are equal. Carroll, Gallo, and Gleser (1985) showed that under this condition the least squares estimate of β has a smaller limiting variance than the ML estimate. However, this is probably not true for the corrected ML-estimate. Table 3 gives a numerical example in order to illustrate the necessary calculations.

Extensions of the Basic Model

Similar procedures can be developed for a number of extensions of the above model. The major problem with that model is the rather restrictive assumption concerning the error variances. The assumption of equality of σ_ε^2 and σ_δ^2 (or the equivalent assumption that the ratio of these variances is known) may not be realistic in many applications. This is especially so since σ_ε^2 will usually consist of two components, only one of which contributes to σ_δ^2. These two components are the measurement error and the error in the equation, that is, the deviation of the error-free dependent variable score from the value predicted on the basis of the functional relationship. σ_ε^2 measures the combined effect of these two sources of variation, while σ_δ^2 consists of measurement error only. However, there is no way out of this predicament unless we have some additional information that allows us to identify the error variances separately.

In practice, if additional information is available, it will usually be of a kind that enables the reliability of the covariate measurements to be determined or estimated. One such instance was analyzed (in a not widely known paper) by Lord (1960). This analysis (which is consistent with the general approach followed in this paper) assumes that two *parallel* measurements of the covariate are available. In effect, this assumption implies that replicated observations are available concerning the T_{ij}-scores. This of course allows σ_δ^2 to be estimated from the replicated observations, and hence a separate consistent estimator for σ_ε^2 may be obtained (see Lord). DeGracie and Fuller (1972) and Stroud (1972) present similar analyses assuming that the error variance of the covariate is either known or has been estimated independently. However, it should be noted that although in this case consistent estimators for the parameters may be obtained, these are *not* ML-estimators. The likelihood function for this situation still remains unbounded, for similar reasons as discussed above. The same difficulty arises in all other cases in which additional information is available that allows σ_δ^2 to be estimated. However, consistent estimators may be derived, based on the pooled within-groups variance-covariance matrix. As in the model discussed in the previous section, the resulting estimator for β may then be corrected for bias and an approximate t-test may be constructed for the hypothesis $\alpha = 0$.

It should be noted that the assumption of *parallel* measurements is not necessary. Consistent estimation is also possible when the "true" covariate is measured through two socalled *congeneric* tests. In this case a second covariate, Z, is available that is known to be correlated with the true score of X, but is independent of the errors of X and Y. Such a variable Z is usually referred to as an *instrumental variable* and its use in the estimation of the parameters of functional relationships has been studied in the statistical and econometric literature (see e.g., Kendall & Stuart, 1967; Moran, 1971).

As an example, let us consider the case that the reliability of the covariate, ρ_{xx}, is known. In that case, σ_δ^2 may be estimated as

$$\hat{\sigma}_\delta^2 = \frac{(1 - \rho_{xx})W_{xx}}{N}.$$

Hence, $\hat{S}_T^2$ is given by

$$\hat{S}_T^2 = \frac{\rho W_{xx}}{N}.$$

Consistent moment estimators for the remaining parameters may now be obtained as follows. The estimator for β is defined as:

$$\hat{\beta} = \frac{W_{xy}}{N\hat{S}_T^2}.$$

Estimators for α and σ_ε^2 are given by

$$\hat{\alpha} = (\bar{Y}_1 - \bar{Y}) - \hat{\beta}(\bar{X}_1 - \bar{X}),$$

$$\hat{\sigma}_\varepsilon^2 = \frac{W_{yy}}{N} - \hat{\beta}^2\hat{S}_T^2.$$

Using standard methods, it may be shown that the expectation and variance of β are approximately equal to

$$E(\hat{\beta}) = \beta\left\{1 + \frac{2S_T^2\sigma_\delta^2}{N(S_T^2 + \sigma_\delta^2)^2}\right\} + O(N^{-1}),$$

$$var(\hat{\beta}) = \frac{S_T^2(\sigma_\varepsilon^2 + \beta^2\sigma_\delta^2) + \sigma_\varepsilon^2\sigma_\delta^2}{NS_T^4} - \frac{2\beta^2\sigma_\delta^4}{N(S_T^2 + \sigma_\delta^2)^2} + O(N^{-1}).$$

As in the previous case, an approximately unbiased estimator $\hat{\beta}_c$ for β is:

$$\hat{\beta}_c = \frac{\hat{\beta}}{1 + \dfrac{2\hat{S}_T^2\hat{\sigma}_\delta^2}{N(\hat{S}_T^2 + \hat{\sigma}_\delta^2)^2}}.$$

The corrected estimator $\hat{\alpha}_c$ for α is defined as before, see (4). The variance of $\hat{\alpha}_c$ is equal to

$$var(\hat{\alpha}) = \frac{\sigma_\varepsilon^2 + \beta^2\sigma_\delta^2}{N} + \frac{var(\hat{\beta})(\mu_1 - \mu_2)^2}{4} + O(N^{-1}).$$

As before, these formulas are large sample approximations. Using similar arguments as in the previous case, a t-test may be constructed for the null hypothesis $\alpha = 0$. As an illustration, Table 4 gives the necessary calculations when this procedure is applied to the numerical example given by Lord (1960). Since the number of observations is in this case different in the two groups, it is easiest to test the hypothesis $\alpha_1 - \alpha_2 = 0$ instead of $\alpha = 0$. The approximate variance of $\hat{\alpha}_1 - \hat{\alpha}_2$ is then given by:

$$var(\hat{\alpha}_1 - \hat{\alpha}_2) = \frac{(\sigma_\varepsilon^2 + \beta^2\sigma_\delta^2)N}{n_1 n_2} + var(\hat{\beta})(\mu_1 - \mu_2)^2 + O(N^{-1}).$$

One of the assumptions of ANCOVA is the equality of the within-groups regression coefficients. In the ordinary ANCOVA model this assumption may be tested by comparing the residual sum of squares about the within-groups regression lines based on the pooled estimate for the within-groups regression coefficient with the residual sum of squares obtained by using a separate regression coefficient for each group (see e.g., Winer, 1971, p. 772-773).

A similar approach may be followed in the present case. The ANCOVA model defined in (1) and (2) is based on the assumption of a single, common, regression coefficient β. A test for the equality of the within-groups regression coefficients may be obtained, starting from a model with separate regression coefficients, that is (1) with β_j substituted for β. Under the assumption that $\sigma_\varepsilon^2 = \sigma_\delta^2$, β_j may be estimated from the sums of squares and crossproducts in group j. It may be shown that a ML-estimator for β_j is given by

$$\hat{\beta}_j = \frac{W_{j,yy} - W_{j,xx} + \{(W_{j,yy} - W_{j,xx})^2 + 4W_{j,xy}^2\}^{1/2}}{2W_{j,xy}},$$

where $W_{j,..}$ is the sum of squares or crossproducts in group j. As before, unbiased estimates for the β_j's may be obtained as follows:

$$\hat{\beta}_{j,c} = \frac{\hat{\beta}_j}{1 + \dfrac{\hat{\sigma}^2\{(1 + \hat{\beta}_j^2)\hat{S}_{j,T}^2 + \hat{\sigma}^2\}}{n(1 + \hat{\beta}_j^2)\hat{S}_{j,T}^4}}.$$

TABLE 4

Application of Test Procedure to Lord's Example

	group 1	group 2
Data:		
number of cases	119	93
dependent variable, Y		
mean	1.40	1.57
s.d.	0.75	0.61
covariate, X		
mean	4.07	5.34
s.d.	2.30	1.97
reliability of X	0.80	0.73
Computed values:		
pooled reliability, ρ		0.7735
error variance of X, $\hat{\sigma}^2_\delta$		1.058
true score variance, $\hat{S}^2_T$		3.614
slope estimate, $\hat{\beta}$		0.241
error variance of Y, $\hat{\sigma}^2_\varepsilon$		0.269
corrected slope estimate, $\hat{\beta}_c$		0.241
estimate of group difference, $\hat{\alpha}_1-\hat{\alpha}_2$		0.136
variance of $\hat{\beta}_c$		0.000506
variance of $\hat{\alpha}_1-\hat{\alpha}_2$		0.007146
test statistic, t(209)		1.609

In this equation $\hat{\sigma}^2$ and $\hat{S}^2_{j,T}$ are given by

$$\hat{\sigma}^2 = \sum \frac{\hat{\beta}_j^2 W_{j,xx} - 2\hat{\beta}_j W_{j,xy} + W_{j,yy}}{N(1 + \hat{\beta}_j^2)},$$

and

$$\hat{S}^2_{j,T} = \frac{W_{j,xx}}{n} - \hat{\sigma}^2.$$

In large samples, $\hat{\beta}_{j,c}$ will be approximately normally distributed with mean β_j and variance $var(\hat{\beta}_j)$,

$$var(\hat{\beta}_j) \simeq \frac{\sigma_\varepsilon^2\{(1+\beta_j^2)S_{j,T}^2 + \sigma_\varepsilon^2\}}{nS_{j,T}^4},$$

where sample estimates have to be substituted for population parameters, using $N\hat{\sigma}^2/(N - 4)$ as an estimator for σ_ε^2. The null hypothesis $\beta_1 = \beta_2$ may be tested with the statistic

$$t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\left(\dfrac{2S^2}{n}\right)^{1/2}},$$

where

$$S^2 = \frac{v\hat{a}r(\hat{\beta}_1) + v\hat{a}r(\hat{\beta}_2)}{2}.$$

This statistic is approximately t-distributed with $(N - 4)$ degrees of freedom. It should be noted that if the assumption of equal β's has to be rejected, application of the ANCOVA model should be strongly discouraged, since the results will generally not be meaningful. In that case, a comparison between groups depends on the value of the covariate at which the comparison is made (see e.g., Tatsuoka, 1971).

ANCOVA as a Structural Relationship

In the structural relationship approach, it is usually assumed that the true scores T_{ij} are randomly sampled from normal distributions with means μ_j and variances σ_j^2. Given this assumption, the data in each group follow a bivariate normal distribution. It is then possible to derive ML-estimates for the parameters of the model by maximizing the likelihood function. Since these ML-estimates may be obtained from the LISREL-program (Jöreskog & Sörbom, 1981), we will not present the likelihood function nor any of the equations that can be derived for the parameter estimates. For a description of the LISREL model and examples of its use, we refer to Lomax (1982, 1983) and Jöreskog and Sörbom.

The Case of Equal Variances

In this section we will discuss the analysis of the structural ANCOVA model using the LISREL approach. Using the LISREL terminology, the basic equations of the structural ANCOVA model, (1) and (2), may be rewritten as follows:

$$\begin{bmatrix}\eta_{ij}\\ T_{ij}\end{bmatrix} = \begin{bmatrix}0 & \beta\\ 0 & 0\end{bmatrix}\begin{bmatrix}\eta_{ij}\\ T_{ij}\end{bmatrix} + \begin{bmatrix}\mu + \alpha_j - \beta\bar{T}\\ \mu_j\end{bmatrix}[1] + \begin{bmatrix}0\\ T_{ij} - \mu_j\end{bmatrix},$$

$$\begin{bmatrix}Y_{ij}\\ X_{ij}\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\eta_{ij}\\ T_{ij}\end{bmatrix} + \begin{bmatrix}\varepsilon_{i(j)}\\ \delta_{i(j)}\end{bmatrix},$$

Since in each group the equations contain constant intercept terms, equal to, respectively, $\mu + \alpha_j - \beta\bar{T}$ and μ_j, we have the structured means version of the LISREL model. This implies that the LISREL specification should contain an X-variable which is identical to 1. In this case, one should analyze the raw moment matrix instead of the covariance or correlation matrix. In addition, the fixed-x option of LISREL should be used (see Jöreskog & Sörbom, 1981).

Although it is not entirely clear why, it turns out that the ML-estimates do not behave very well unless it is assumed that $\sigma_\varepsilon^2 = \sigma_\delta^2$ (except when the model fits the data perfectly). Perhaps this should come as no surprise since ML-estimators do not exist in the corresponding functional model. We will return to this problem later on in this paper.

Under the assumption that $\sigma_\varepsilon^2 = \sigma_\delta^2$, the model contains $3g + 2$ free parameters: β, σ_ε^2, and for each of the g groups $\mu + \alpha_j - \beta\bar{T}$, μ_j, and σ_j^2. The estimates of σ_ε^2 and σ_j^2 are given by the variance-covariance matrices of ε and ζ. The null hypothesis $\alpha_j = 0$ may be tested by constraining the intercept parameters $\mu + \alpha_j - \beta\bar{T}$ to be equal. The appropriate likelihood ratio (LR) statistic for testing this hypothesis is formed by subtracting the chi-square value reported by LISREL under the unconstrained model from the chi-square value obtained in the restricted case. In large samples this statistic follows a chi-square distribution with $(g - 1)$ degrees of freedom.

It should be noted, however, that the chi-square values reported by the LISREL program are not quite correct. It turns out that the reported chi-square values should be multiplied by a factor $n/(n - 1)$, assuming equal n's in each group. (No simple correction can be applied if the groups contain unequal numbers of observations, since the LISREL program does not report chi-square values for each group). The reason for this discrepancy is that the LISREL model is based on the assumption that the observed covariance matrix follows a Wishart distribution. The structural ANCOVA model, however, is based on the assumption that the observations in each group follow a multivariate normal distribution. Although these two assumptions are closely related, they are not equivalent. In particular, the likelihood functions are slightly different. This difference implies that the chi-square values in each group have to be corrected. This is generally the case if intercept terms are included in the model and the raw moment matrix is analyzed (see Jöreskog, 1973, p. 93).

An example of the LISREL approach to the analysis of a linear structural ANCOVA model is presented in Table 5. This table includes the appropriate parameter estimates and the corrected chi-square values. The example which will be discussed, is the same as was used to illustrate the functional solution (see Table 3). These data were not obtained by Monte Carlo simulation but by inserting particular values for the parameters into the theoretical equations for the moments. Application of the LISREL program to these data enables a direct examination of the quality of these estimates. The data were generated in such a way that the observed statistics would correspond almost perfectly to a structural ANCOVA model. There are only two minor deviations from the model as presented above: (a) the two groups were slightly different with respect to the parameter β, and (b) σ_δ^2 and σ_ε^2 were also given different (but nearly equal) values. The parameter values were: $\mu = 0.0$, $\alpha = 1.6$, $\beta_1 = 2.1$, $\beta_2 = 1.9$, $\mu_1 = 5.0$, $\mu_2 = 8.0$, $\sigma_1^2 = 10.0$, $\sigma_2^2 = 15.0$, $\sigma_\varepsilon^2 = 1.9$, $\sigma_\delta^2 = 2.1$ and $n = 10$ in each group. The resulting moment matrices are given in Table 5.

Note that the parameter estimates closely resemble the corresponding functional estimates. Thus, the latter values may be used as initial estimates for the structural parameters. Initial values for the two remaining parameters, σ_1^2 and σ_2^2, may be obtained from the following formula:

$$\hat{\sigma}_j^2 = \frac{s_{x(j)}^2 + 2\hat{\beta}s_{xy(j)} + \hat{\beta}^2 s_{y(j)}^2 - \hat{\sigma}_\varepsilon^2(1 + \hat{\beta}^2)}{(1 + \hat{\beta}^2)^2}.$$

These initial parameter estimates are quite useful since the LISREL program may arrive at an incorrect solution when supplied with bad initial estimates. Incorrect solutions are not infrequent in LISREL and are characterized by negative variance estimates. The LISREL program may generate such solutions because it does not restrict the variance estimates to nonnegative values. It turns out however that in many cases the likelihood function has several (local) maxima, only one of which is in the correct parameter space.

TABLE 5

Numerical Example of the Analysis of Covariance as a Linear Structural Relationship

Observed moment matrices (input for LISREL):

	group 1 (n=10)			group 2 (n=10)		
	Y	X	1	Y	X	1
Y	48.403			57.613		
X	13.250	37.100		38.500	81.100	
1	-1.550	5.000	1.000	1.250	8.000	1.000

Maximum likelihood solution for parameters:

	H_0: α=0	H_1: $\alpha \neq 0$
$\hat{\mu}$	-0.150	-0.150
$\hat{\alpha}$	0.0	1.566
$\hat{\beta}$	1.874	1.977
$\hat{\mu}_1$	5.586	5.000
$\hat{\mu}_2$	7.414	8.000
$\hat{\sigma}_1^2$	11.830	10.986
$\hat{\sigma}_2^2$	15.190	14.055
$\hat{\sigma}_\varepsilon^2$	2.555	2.084
LISREL χ^2 (corrected)	4.29 df=3	0.22 df=2

Test of H_0: χ^2 = 4.07, df=1

For example, when supplied with bad initial estimates, LISREL obtained an incorrect solution (under the alternative hypothesis) for the numerical data given above with a chi-square value of 0.37.

Extensions and Further Tests of the Structural Model

Our solution to the fallible covariate problem in the structural case (as well as in the functional case) is based on the assumption $\sigma_\varepsilon^2 = \sigma_\delta^2$. As mentioned in the discussion of the functional model, this is a very restrictive assumption that is difficult to defend. It is not known to what extent violation of this assumption biases the results of the analysis.

In particular, we do not know whether this leads to a substantial bias in the likelihood-ratio test for the ANCOVA null hypothesis. Although the present case seems quite similar to the functional case, there are some important differences. In the functional model, the parameters σ_ε^2 and σ_δ^2 are not both identifiable (unless additional information is present). In the structural model, on the other hand, these parameters cannot be said to be unidentifiable. The reason for this is that these parameters are estimated correctly when the observed covariance matrix (or moment matrix) fits the model *exactly*, that is, if a perfect solution is possible. Problems arise however as soon as the observed matrix deviates slightly (and nonsignificantly) from the predicted structure.

In order to demonstrate this point, we generated a number of moment matrices which violated the assumption $\sigma_\varepsilon^2 = \sigma_\delta^2$ in varying degrees. In each case, σ_ε^2 was set equal to 2.0, while σ_δ^2 was varied between 2.0 and 17.0. The remaining parameters (except for β_1 and β_2) were equal to those used to generate the data in Table 5. When the regression coefficients in the two groups, β_1 and β_2, were set equal to each other (in which case a perfect solution is possible), the correct solution was always obtained and all the parameter estimates were equal to the true values. However, when β_1 and β_2 were given slightly different values, strange and unexpected results were obtained. Note that none of these datamatrices violate the structural ANCOVA model with $\sigma_\varepsilon^2 \neq \sigma_\delta^2$ to any significant degree. Table 6 gives the most important results of this analysis. In this table, the estimates for σ_ε^2 and σ_δ^2 are given as well as the chi-square values from the LR-test for the hypothesis $\sigma_\varepsilon^2 = \sigma_\delta^2$. These parameter estimates were not obtained with the LISREL program, but with a general purpose minimization program that allows upper and lower limits on the parameter values (James & Roos, 1975). This program was used because in these cases the LISREL estimates from σ_ε^2 and σ_δ^2 were often outside the admissible parameter space (LISREL does not restrict the parameter estimates to values within the admissible parameter space). For these parameter values, the results are very unstable and strongly dependent on small differences in β_1 and β_2.

The lefthand part of Table 6 gives the results for $\beta_1 = 1.9$ and $\beta_2 = 2.1$. In this case σ_ε^2 was always estimated as 0.0 (the lower bound), while σ_δ^2 was always overestimated in a

TABLE 6

ML-Estimates and Test for Equality of σ_ε^2 and σ_δ^2

	$\beta_1=1.9$, $\beta_2=2.1$			$\beta_1=2.1$, $\beta_2=1.9$		
σ_δ^2	$\hat\sigma_\varepsilon^2$	$\hat\sigma_\delta^2$	χ^2	$\hat\sigma_\varepsilon^2$	$\hat\sigma_\delta^2$	χ^2
2	0.0	2.499	0.09	8.879	0.0	0.160
3	0.0	3.499	0.10	11.605	0.0	0.122
5	0.0	5.499	0.11	14.436	0.804	0.069
7	0.0	7.499	0.13	15.059	2.515	0.041
9	0.0	9.499	0.14	15.487	4.312	0.026
11	0.0	11.499	0.15	15.801	6.159	0.016
13	0.0	13.499	0.17	16.043	8.039	0.009
15	0.0	15.499	0.19	16.235	9.943	0.005
17	0.0	17.499	0.20	16.391	11.864	0.003

systematic way: The estimated value for σ_δ^2 was equal to the true value plus a constant (0.499). More importantly, the LR-test does not seem to be very sensitive to changes in σ_δ^2 (although the chi-square values become slightly larger as σ_δ^2 deviates more and more from σ_ε^2). In all cases, the value of this statistic is quite small and never leads to rejection of the hypothesis $\sigma_\varepsilon^2 = \sigma_\delta^2$. The righthand part of Table 6 gives the corresponding results for $\beta_1 = 2.1$ and $\beta_2 = 1.9$. Although the regression coefficients have not been changed very much, the pattern of the parameter estimates is completely different from the previous case. In this case, σ_ε^2 is grossly overestimated while σ_δ^2 is severely underestimated. Moreover, the chi-square values of the LR-test *decrease* with increasing differences between σ_ε^2 and σ_δ^2. Hence, we may conclude that the hypothesis $\sigma_\varepsilon^2 = \sigma_\delta^2$ is not testable and that the separate estimation of σ_ε^2 and σ_δ^2 leads to unsatisfactory results.

More satisfactory results can only be obtained if additional information is available that allows the identification of both σ_ε^2 and σ_δ^2. Suppose for example, that we know (or have information that permits the estimation of) the reliability of the covariate, ρ_{xx}. In that case, σ_δ^2 might be fixed at $(1 - \rho_{xx})W_{xx}/N$, as in the corresponding functional case. This allows σ_ε^2 to be estimated. Although this does not correspond to the conventional ANCOVA model, knowledge regarding ρ_{xx} allows one to estimate separate error variances within each group. It is usually assumed that the measurement error σ_δ^2 is equal in all groups. If there is reason to suspect that this assumption is not correct, separate reliabilities should be used for the estimation of these variances.

As an example of such an analysis with different reliabilities, we reanalyzed Lord's numerical example (Lord, 1960). Three types of analysis were performed, the results of which are given in Table 7. Model I is the type of analysis we have just described, adapted to this situation, that is σ_δ^2 in each group is set equal to $(1 - \rho_j)W_{j,xx}/n_j$. Model II is the correct analysis given the assumption of unequal measurement errors. In this analysis σ_δ^2 was set equal to $\sigma_j^2(1 - \rho)/\rho$, where σ_j^2 is the true score variance in group j. Since the LISREL program does not allow such a restriction on the parameters, these estimates were obtained by direct minimization of the appropriate likelihood function using a general-purpose minimization routine. Another type of solution was presented by Sörbom (1978). Sörbom used the information concerning the reliabilities in a different way. Instead of fixing or restraining σ_δ^2, Sörbom created an artificial second covariate which was constructed in such a way that the two covariates were parallel measurements with a correlation equal to the observed reliability. In doing so, Sörbom followed the original approach taken by Lord (1960) who analyzed these data in a similar way. Model III gives the estimates for this analysis obtained with the LISREL program (for reasons unclear to us, the results deviate somewhat from those reported by Sörbom, 1978).

On comparing the results, it is evident that there are only minor differences in this case between the three approaches. However, if one does not have easy access to a general-purpose minimization routine and prefers to use the LISREL program, it is in our opinion advisable to use the Model I type of analysis instead of the Model III or Sörbom type of analysis. The behavior of the likelihood function may depend on the assumption of independent parallel measurements (which are in fact not available) and this might affect properties of the estimates such as their standard errors.

Finally, a likelihood-ratio test for the assumption of equal within-groups regression coefficients can be obtained in a straightforward manner with the LISREL program. However, contrary to the assertion of Sörbom (1978), it is not possible to test the null hypothesis $\alpha_j = 0$ with the LISREL program if the assumption of equal regression coefficients is not tenable. The reason for this is that in the model underlying the LISREL approach the test for $\alpha_j = 0$ is not independent of interval scale transformations of the covariate measurements. In the LISREL approach to ANCOVA, the model may be

TABLE 7

Analysis of Lord's Numerical Example According to Structural Equation Model with Different Reliabilities in each group (see text for explanation)

	model I	model II	model III
$\hat{\mu}$	1.466	1.466	1.466
$\hat{\alpha}$	0.069	0.068	0.069
$\hat{\beta}$	0.243	0.241	0.242
$\hat{\mu}_1$	4.070	4.070	4.070
$\hat{\mu}_2$	5.340	5.340	5.340
$\hat{\sigma}_1^2$	4.326	4.196	4.282
$\hat{\sigma}_2^2$	2.580	2.792	2.662
$\hat{\sigma}_{\varepsilon 1}^2$	0.273	0.275	0.271
$\hat{\sigma}_{\varepsilon 2}^2$	0.249	0.250	0.256
$\hat{\sigma}_{\delta 1}^2$	1.058	1.049	1.033
$\hat{\sigma}_{\delta 2}^2$	1.058	1.046	1.084
χ^2 test for α=0	2.67 (*)	2.64	2.92 (*)

Note: (*) uncorrected value from LISREL program

written as

$$y_{ij} = \mu_j + \beta_j T_{ij} + \varepsilon_{i(j)},$$

where

$$\mu_j = \mu + \alpha_j - \beta_j \bar{T}.$$

The ANCOVA null hypothesis is tested in LISREL by comparing the μ_j's. The problem now is that these μ_j's may be different even though $\alpha_j = 0$ (for all j). Moreover, scale transformations of T affect the outcome of this likelihood-ratio test. However, as we have already mentioned earlier, it is best not to proceed with an ANCOVA type of analysis when the assumption of equal regression coefficients has been rejected.

Conclusions

In this paper we have examined the application of a functional and a structural relationship approach to the analysis of covariance with a fallible covariate. Examples of both types of analysis have been given for a simple two-group design. Several cases have been discussed and we have given special attention to issues of model identifiability. An approximate statistical test based on the functional relationship approach has been constructed. On the basis of our simulation results it may be concluded that this testing procedure is to be preferred to the conventional F-test of the ANCOVA null hypothesis. Further analysis of this approach to the problem of fallible covariates is obviously desirable, especially regarding its extension to more complicated ANCOVA designs. If one is willing to assume a normal distribution for the covariate scores in each group, the ANCOVA model may be formulated as a structural relationship problem. In this case, an analysis based on the LISREL methodology should be performed. It would seem that in both the functional and the structural case, knowledge of the reliability of the covariate is desirable. In most cases, one should try to obtain parallel measurements of the covariate. With such additional information, arbitrary assumptions concerning the error variances can be avoided.

The results that we have presented with respect to the structural relationship approach may also have a wider significance. We have demonstrated that such a model may lead to problematic results when it is applied to a situation where the corresponding functional relationship would be unidentifiable. Uncritical use of a standard program such as LISREL may lead to invalid results. Unfortunately, we have not been able to determine the exact reason for these strange results. Based on the standard LISREL approach, one would probably not have been alerted to these problems. However, this implies that there may be other situations where the standard LISREL approach may not applicable. Hence, further analysis of the structural relationship model, especially with respect to the question of identifiability, would seem to be called for.

Finally, we wish to reiterate that the present approach should only be used when one assumes that the dependent variable is linearly related to the true score on the covariate. If the dependent variable is assumed to be related to the *observed* score on the covariate, the standard ANCOVA results will be valid. However, we doubt whether such an assumption is in fact reasonable in most applications of ANCOVA.

References

Anderson, T. W. (1976). Estimation of linear functional relationships: Approximate distributions and connections with simultaneous equations in econometrics. *Journal of the Royal Statistical Society, Series B, 38*, 1–36.

Anderson, T. W. (1984). Estimating linear statistical relationships. *The Annals of Statistics*, 12, 1–45.

Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 5*, 111–150.

Anderson, T. W., & Sawa, T. (1982). Exact and approximate distributions of the maximum likelihood estimator of a slope coefficient. *Journal of the Royal Statistical Society, Series B, 44*, 52–62.

Box, G. E. P., & Muller, M. E. (1958). A note on the generation of random normal deviates. *Annals of Mathematical Statistics, 29*, 610–611.

Carroll, R. J., Gallo, P., & Gleser, L. J. (1985). Comparison of least squares and errors-in-variables regression with special reference to randomized analysis of covariance. *Journal of the American Statistical Association, 80*, 929–932.

DeGracie, J. S., & Fuller, W. A. (1972). Estimation of the slope and analysis of covariance when the concomitant variable is measured with error. *Journal of the American Statistical Association, 67*, 930–937.

Edgell, S. E. (1979). A statistical check of the DECsystem-10 FORTRAN pseudorandom number generator. *Behavior Research Methods & Instrumentation, 11*, 529–530.

James, F., & Roos, M. (1975). MINUIT, a system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications, 10*, 343–367.

Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences.* New York: Seminar Press.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares methods.* Chicago: International Educational Services.

Kendall, M. G., & Stuart, A. (1967). *The advanced theory of statistics. Vol. 2: Inference and relationship* (2nd ed). London: Griffin.

Lomax, R. G. (1982). A guide to LISREL-type structural equation modeling. *Behavior Research Methods & Instrumentation, 14*, 1–8.

Lomax, R. G. (1983). A guide to multiple-sample structural equation modeling. *Behavior Research Methods & Instrumentation, 15*, 580–584.

Lord, F. M. (1960). Large-sample covariance analysis when the control variable is fallible. *Journal of the American Statistical Association, 55*, 307–321.

Moran, P. A. P. (1971). Estimating structural and functional relationships. *Journal of Multivariate Analysis, 1*, 232–255.

Robertson, C. A. (1974). Large-sample theory for the linear structural relation. *Biometrika, 61*, 353–359.

Sörbom, D. (1978). An alternative to the methodology for analysis of covariance. *Psychometrika, 43*, 381–396.

Stroud, T. W. F. (1972). Comparing conditional means and variances in a regression model with measurement errors of known variances. *Journal of the American Statistical Association, 67*, 407–414.

Tatsuoka, M. M. (1971). *Multivariate analysis: Techniques for educational and psychological research.* New York: Wiley.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed). New York: McGraw-Hill.

Manuscript received 12/16/85
Final version received 11/19/86