AD 603 563

53p.
h - 3.00
mf - 0.50

ON INCREASING TREATMENT CONTRAST PRECISION AND THE ESTIMATION OF STRUCTURAL PARAMETERS IN COVARIANCE ANALYSIS

A Report

to

U. S. Army Biological Laboratories
Fort Detrick, Maryland

under

Contract DA 18-064-AMC-103(A)

Prepared by

William S. Mallios

Booz, Allen Applied Research, Inc.
4815 Rugby Avenue
Bethesda 14, Maryland

June 1964

BOOZ·ALLEN APPLIED RESEARCH INC.

WASHINGTON
CLEVELAND
CHICAGO

TABLE OF CONTENTS

SUMMARY

Analysis of covariance in randomized and balanced incomplete block designs is reconsidered in terms of structural regression. In practice, the covariable is usually uncontrolled and may follow a linear model as does the variable of interest. If the covariable and treatments are independent, then the covariable model contains no treatment effect, and the treatment contrast precision on the variable of interest is increased, not only asymptotically, but in the finite sample. If the covariable is affected by one more of the treatments, then the estimation of direct and indirect effects is considered. Finally, when structural parameters are underidentified, in which case direct and indirect effects are not estimable, an alternative estimation procedure is discussed.

1. INTRODUCTION

Consider the randomized block design model

$$y_{1ij} = \mu + \tau_i + \beta_j + \alpha y_{2ij} + \epsilon_{ij} \tag{1.1}$$

where τ_i (β_j) is the fixed differential effect of the i^{th} (j^{th}) treatment (block) on a response of interest, say y_{1ij}, which is taken from the $(i, j)^{th}$ experimental unit, $i = 1, \ldots, q$; $j = 1, \ldots, r$. μ is usually the base from which the differential effects are estimated, and ϵ_{ij} is the random model error. The covariable, y_{2ij}, is assumed independent of treatments and is included in model (1.1) to account for differences in the y_{2ij} from one experimental unit to the next in estimating treatment contrasts.

Consider, next, the following three conditions and/or assumptions which may accompany a covariance analysis:

1. The covariable is uncontrolled, and is measured with negligible measurement error.
2. The covariable, though independent of treatments, is, perhaps, dependent on blocks.
3. The covariable is unaffected by the variable of interest, though not conversely.

Under the circumstances, it may be reasonable to assume the following linear model for the covariable:

$$y_{2ij} = \mu_2 + \beta_{2j} + \epsilon_{2ij} \tag{1.2}$$

where β_{2ij} is the fixed differential effect of the j^{th} block on y_{2ij}, and ϵ_{2ij} is the random model error. For consistency of notation, model (1.1) is rewritten as

$$y_{1ij} = \mu_1 + \tau_i + \beta_{1j} + \alpha y_{2ij} + \epsilon_{1ij}.$$

When there is doubt regarding the independence of treatment and covariable, then a treatment effect, say τ_{2i}, should be included in the model for the covariable; i.e.,

$$y_{2ij} = \mu_2 + \tau_{2i} + \beta_{2j} + \epsilon_{2ij} \tag{1.3}$$

so that model (1.1) is now rewritten as

$$y_{1ij} = \mu_1 + \tau_{1i} + \beta_{1j} + \alpha y_{2ij} + \epsilon_{1ij}. \tag{1.4}$$

Substituting the expression for y_{2ij} in (1.3) into (1.4) yields

$$y_{1ij} = (\mu_1 + \alpha\mu_2) + (\tau_{1i} + \alpha\tau_{2i}) + (\beta_{1j} + \alpha\beta_{2j}) + (\epsilon_{1ij} + \alpha\epsilon_{2ij})$$

where $\tau_{1i} + \alpha\tau_{2i}$ is the overall treatment effect on y_1, τ_{1i} is the direct treatment effect on y_1, and $\alpha\tau_{2i}$ is the indirect treatment effect on y_1 or that treatment effect which is passed on to y_1 through y_2.

Regarding the model errors, ϵ_{1ij} and ϵ_{2ij}, the following assumptions are made:

$$E(\epsilon_{1ij}) = E(\epsilon_{2ij}) = 0$$

$$E(\epsilon_{1ij}^2) = \sigma_{\epsilon_1}^2, \qquad E(\epsilon_{2ij}^2) = \sigma_{\epsilon_2}^2$$

$$E(\epsilon_{1ij}\,\epsilon_{1i'j'}) = E(\epsilon_{2ij}\,\epsilon_{2i'j'}) = 0 \quad \text{for} \quad i \neq i' \quad \text{or} \quad j \neq j'$$

$$E(\epsilon_{1ij}\,\epsilon_{2ij}) = \sigma_{\epsilon_1 \epsilon_2}.$$

In Sections 2 and 3, we will illustrate the following. Assume $\sigma_{\epsilon_1 \epsilon_2} = 0$. Then

(1) if treatments and covariable are independent and if model (1.2) is adequate for the covariable, the treatment contrast precision on the variable of interest (y_1) is increased, not only asymptotically, but in the finite sample;

(2) if treatments affect the covariable and if model (1.3) is adequate for the covariable, then that estimated treatment effect on the variable of interest, which is obtained through the usual covariance analysis, is the direct treatment effect on y_1; and corresponding treatment contrasts also have increased precision in the finite sample.

If $\sigma_{\epsilon_1\epsilon_2} \neq 0$, which is considered in Section 4, it will be shown that statement (1) still holds. However, α, in the context of statement (1), and direct and indirect effects, in the context of statement (2), are not estimable through usual techniques due to underidentification. When parameters are underidentified, an alternative estimation procedure is discussed.

2. INCREASING TREATMENT CONTRAST PRECISION WHEN THE COVARIABLE IS UNCONTROLLED, INDEPENDENT OF TREATMENTS, AND $\sigma_{\epsilon_1 \epsilon_2} = 0$

2.1 THE STRUCTURAL AND REDUCED SYSTEMS, AND THE ESTIMATION OF TREATMENT EFFECTS WHEN α IS KNOWN

Consider those applications where the covariable is uncontrolled, independent of treatments, and where the system (described in Section 1)

$$y_{1ij} = \mu_1 + \tau_i + \beta_{1j} + \alpha y_{2ij} + \epsilon_{1ij} \tag{2.1.1}$$

$$y_{2ij} = \mu_2 + \beta_{2j} + \epsilon_{2ij} \tag{2.1.2}$$

is adequate. It is assumed that

$$\begin{bmatrix} \epsilon_{1ij} \\ \epsilon_{2ij} \end{bmatrix} : \text{i.i.d.} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\epsilon_1} & 0 \\ 0 & \sigma^2_{\epsilon_2} \end{bmatrix} \right) ; \tag{2.1.3}$$

i.e., $[\epsilon_{1ij}, \epsilon_{2ij}]$ is identically and independently distributed with expectation $[0, 0]$, and variance $\begin{bmatrix} \sigma^2_{\epsilon_1} & 0 \\ 0 & \sigma^2_{\epsilon_2} \end{bmatrix}$.

Models (2. 1. 1) and (2. 1. 2) describe a partial and relevant structu of the experimental unit and are thus termed the structural regression models or the structural system. It is inconsistent, logically though not necessarily statistically, to estimate the system's parameters separatel or model by model; for in doing so, y_2 is fixed in (2. 1. 1) and uncontrolled in (2. 1. 2); certainly y_2 cannot assume both roles simultaneously Moreover, if instead of (2. 1. 2), $y_{2ij} = \mu_2 + \tau_{2i} + \beta_{2j} + \alpha' y_{1ij} + \epsilon_{2ij}$, then separate least squares estimation leads to inconsistent estimates as is discussed by Haavelmo (1943).

Substituting (2. 1. 2) into (2. 1. 1) yields

$$y_{1ij} = (\mu_1 + \alpha\mu_2) + \tau_i + (\beta_{1j} + \alpha\beta_{2j}) + (\epsilon_{1ij} + \alpha\epsilon_{2ij}) \qquad (2.1.4)$$

which, along with (2. 1. 2), is termed the reduced system or the reduced regression models. In (2. 1. 4), not only is τ_i directly estimable, as opposed to, say, the direct block effect, β_{1j}, on y_1, but also, in the reduced system, y_2 no longer assumes a dual role. Separate estimation in this reduced system yields consistent estimates; however the errors of the reduced system, $\epsilon_{1ij} + \alpha\epsilon_{2ij}$ and ϵ_{2ij}, are correlated, which, concurrent with a nonexistent treatment effect in (2. 1. 2), implie an increased treatment contrast precision through joint estimation [Mallios (1961)], as will now be shown.

From (2.1.3) it follows that

$$\left[\epsilon_{1ij} + \alpha\epsilon_{2ij}, \ \epsilon_{2ij}\right] \ : \ \text{i.i.d.} \quad \left[(0,0), \sum\right] \tag{2.1.5}$$

where

$$\sum = \begin{bmatrix} \sigma^2_{\epsilon_1} + \alpha^2\sigma^2_{\epsilon_2}, & \alpha\sigma^2_{\epsilon_2} \\ \alpha\sigma^2_{\epsilon_2}, & \sigma^2_{\epsilon_2} \end{bmatrix} = (\sigma_{hh'}). \tag{2.1.6}$$

$h, h' = 1, 2$. Denote $\sum^{-1}$ by $(\sigma^{hh'})$ and write the sample form of (2.1.4) and (2.1.2) as $y_{1ij} = m + t_i + b_j + e_{1ij}$ and $y_{2ij} = m_2 + b_{2j} + e_{2ij}$, respectively, where m, t_i, b_j and e_{1ij} correspond to $\mu_1 + \alpha\mu_2$, τ_i, $\beta_{ij} + \alpha\beta_{2j}$, and $\epsilon_{1ij} + \alpha\epsilon_{2ij}$.

For α known, best (Markoff) estimates, among the class of linear, unbiased estimators, are found by taking partials of

$$\sum_{h,h',i,j}^{2,2,q,r} e_{hij}\, e_{h'ij}\, \sigma^{hh'}, \tag{2.1.7}$$

equating these to zero, and utilizing the usual restrictions that

$$\sum_i \tau_i = \sum_j \beta_{hj} = 0. \tag{2.1.8}$$

The resultant estimates are

$$t_i = (\bar{y}_{1i} - \bar{y}_1) - \alpha(\bar{y}_{2i} - \bar{y}_2), \quad b_{2j} = \bar{y}_{2j} - \bar{y}_2 \qquad (2.1.9)$$

$$b_j = \bar{y}_{1j} - \bar{y}_1, \quad m_2 = \bar{y}_2, \quad m = \bar{y}_1,$$

where

$$\alpha = \sigma_{12}/\sigma_{22}, \quad \bar{y}_{hi} = \sum_j y_{hij}/r, \quad \bar{y}_{hj} = \sum_i y_{hij}/q,$$

and

$$\bar{y}_h = \sum_{ij} y_{hij}/qr.$$

Also, with $\sigma_{12} = \rho(\sigma_{11}\sigma_{22})^{1/2}$ in (2.1.6),

$$\text{var}(t_i - t_{i'}) = 2\sigma_{11}(1-\rho^2)/r = 2\sigma^2_{\epsilon_1}/r. \qquad (2.1.10)$$

If τ_{2i} is included in model (2.1.2) when, in fact, $\tau_{2i} = 0$, then τ_i is estimated by

$$t_i^{(o)} = (\bar{y}_{1i} - \bar{y}_1) = \tau_i + \bar{\epsilon}_{1i} = -\bar{\epsilon}_i + \alpha(\bar{\epsilon}_{2i} - \bar{\epsilon}_2) \qquad (2.1.11)$$

and $\text{var}(t_i^{(o)} - t_{i'}^{(o)}) = 2(\sigma^2_{\epsilon_1} + \alpha^2\sigma^2_{\epsilon_2})/r$, so that $\text{var}(t_i - t_{i'}) \leq \text{var}(t_i^{(o)} - t_{i'}^{(o)})$.

2.2 THE ESTIMATION OF TREATMENT EFFECTS WHEN α IS UNKNOWN

For $\sum$ and functions thereof unknown, various estimates of τ_i have been considered [Mallios (1961), Zellner (1962)], though in a slightly different context. Perhaps the most obvious estimate is that which maximizes the likelihood function, assuming normality in (2.1.3). This estimate requires iteration to convergence. Since in (2.1.9) α is unknown, only the t_i change in iteration which implies that the unbiased estimator

$$\sigma^*_{22} = \sum_{ij} (y_{2ij} - m_2 - b_{2j})^2/(qr - r) \tag{2.2.1}$$

remains unchanged in iteration, where $(qr - r)\ \sigma^*_{22}/qr$ is the maximum likelihood estimate of σ_{22}. Choosing an initial estimate of τ_i as $t_i^{(o)}$ in (2.1.11) and the other estimates as in (2.1.9), then the estimators

$$\sigma^{(o)}_{12} = \sum_{ij} (y_{1ij} - m - t_i^{(o)} - b_j)(y_{2ij} - m_2 - b_{2j})/(qr - q - r + 1) \tag{2.2.2}$$

and

$$\sigma^{(o)}_{11} = \sum_{ij} (y_{1ij} - m - t_i^{(o)} - b_j)^2/(qr - q - r + 1) \tag{2.2.3}$$

are consistent and unbiased for σ_{12} and σ_{11}. From $\sigma^{(o)}_{12}/\sigma^*_{22} = a^{(o)}$, a second estimate of τ_i, say $t_i^{(1)} = (\bar{y}_{1i} - \bar{y}_1) - a^{(o)}(\bar{y}_{2i} - \bar{y}_2)$, is obtained from which $\sigma_{12}^{(1)}$, $a^{(1)} = \sigma_{12}^{(1)}/\sigma_{22}^{*}$, and $\sigma_{11}^{(1)}$ are calculated. This process is continued until stable estimates, say t_i^{*}, $a^{*} = \sigma_{12}^{*}/\sigma_{22}^{*}$, and σ_{11}^{*}, are attained. The efficiency of the maximum likelihood estimator, $t_i^{*} - t_{i'}^{*}$, $i \neq i'$, is given by (2.1.10) and $\sigma_{\epsilon_1}^2 = \sigma_{11} - \alpha^2\sigma_{22}$ estimated by $\sigma_{\epsilon_1}^{*2} = \sigma_{11}^{*} - \sigma_{12}^{*2}/\sigma_{22}^{*}$.

Aside from $t_i^{*} - t_{i'}^{*}$, all estimates obtained during the iterative cycle (except for $t_i^{(o)} - t_{i'}^{(o)}$) are consistent and equally efficient; i.e., since $\sigma_{12}^{(o)}$, $\sigma_{11}^{(o)}$, and $\sigma_{22}^{*} = \sigma_{22}^{(o)}$ are consistent, the initial estimate, $t_i^{(1)} - t_{i'}^{(1)}$, is therefore consistent and efficient [Zellner (1962)]; and, since the initial estimate and the "final" estimate, $t_i^{*} - t_{i'}^{*}$, share the same large sample properties, so must all estimates obtained during the iterative cycle. While the exact finite sample variation of these estimates remains unresolved, Zellner (1964) in a somewhat different context, produced yet another $\tau_i - \tau_{i'}$ estimator and derived its finite sample properties. His approach is as follows: Assuming normality in (2.1.3) and including a treatment effect, say $\tau_{2i} = 0$, in (2.1.2), then it is seen that τ_{2i} is estimated by $t_{2i} = \bar{y}_{2i} - \bar{y}_2$ where $E(t_{2i}) = 0$; τ_i is then estimated by $t_i^{(o)}$ in (2.1.11); and the consistent, unbiased variance-covariance estimates $s_{11} = \sigma_{11}^{(o)}$

$$s_{12} = \sum_{ij} (y_{1ij} - m - t_i^{(o)} - b_j)(y_{2ij} - m_2 - t_{2i} - b_{2j})/(qr - q - r + 1), \quad (2.2.4)$$

and

$$s_{22} = \sum_{ij} (y_{2ij} - m_2 - t_{2i} - b_{2j})^2 / (qr - q - r + 1) \tag{2.2.5}$$

are distributed as Wishart variates independently of $\bar{y}_{1i} - \bar{y}_1$ and $\bar{y}_{2i} - \bar{y}_2$. If τ_i is reestimated by

$$\hat{\tau}_i = (\bar{y}_{1i} - \bar{y}_1) - (s_{12}/s_{22})(\bar{y}_{2i} - \bar{y}_2), \tag{2.2.6}$$

then, since s_{12} and s_{22} are consistent, the efficiency of $\hat{\tau}_i - \hat{\tau}_{i'}$ is given by (2.1.10); moreover, $\hat{\tau}_i - \hat{\tau}_{i'}$ is unbiased; i.e.,

$$E(\hat{\tau}_i) = E(\bar{y}_{1i} - \bar{y}_1) - E(s_{12}/s_{22})\, E(\bar{y}_{2i} - \bar{y}_2) = \tau_i$$

since s_{12}/s_{22} is distributed independently of $\bar{y}_{2i} - \bar{y}_2$, whose expectation is zero. Zellner then derived the exact second moment of $\hat{\tau}_i - \hat{\tau}_{i'}$, given by

$$\operatorname{var}(\hat{\tau}_i - \hat{\tau}_{i'}) = (2/r)\,\sigma_{11}(1 - \rho^2)(1 + \phi) = (2/r)\,\sigma^2_{\epsilon_1}(1 + \phi) \tag{2.2.7}$$

where $\phi = (N - 2)^{-1} \pi^{-1/2} \Gamma[(N + 1)/2] / \Gamma(N/2)$ and $N = qr - q - r + 1$, and showed that the exact distribution of $\hat{\tau}_i - \hat{\tau}_{i'}$ rapidly approaches normality, for $N > 10$. Note that when $\alpha = 0$, the exact second moment of a treatment contrast is $2\sigma^2_{\epsilon_1}/r$; comparing the latter with (2.2.7), it is seen that the inclusion of covariable in (2.1.1), when in fact y_1 is independent of y_2, i.e., $\alpha = 0$, produces a decrease in contrast precision. As such, it is assumed throughout that $|\alpha|$ is substantial.

In comparing the estimators $\hat{\tau}_i - \hat{\tau}_{i'}$ and $t_i^{(1)} - t_{i'}^{(1)}$, there is reason to prefer the latter even though its finite sample variation is, at present, unknown. Note first that $\text{var}\, \sigma_{hh'}^{(o)} \leq \text{var}\ s_{hh'}$ where $\sigma_{11}^{(o)}$ is given in (2.2.3); $\sigma_{12}^{(o)}$ is given in (2.2.2); $\sigma_{22}^{(o)} = \sigma_{22}^{*}$ is given in (2.2.1); $s_{11} = \sigma_{11}^{(o)}$, s_{12} is given in (2.2.4); and s_{22} is given in (2.2.5); i.e.,

$$\text{var}\ \sigma_{11}^{(o)} = 2\,\sigma_{11}^{2}/(qr - q - r + 1) = \text{var}\ s_{11},$$

$$\text{var}\ \sigma_{12}^{(o)} = 2\sigma_{12}^{2}/(qr - q - r + 1) = \text{var}\ s_{12},$$

but

$$\text{var}\ \sigma_{22}^{(o)} = 2\sigma_{22}^{2}/(qr - r) < 2\sigma_{22}^{2}/(qr - q - r + 1) = \text{var}\ s_{22}.$$

Since $\hat{\tau}_i - \hat{\tau}_{i'}$ = function (s_{12}/s_{22}) while $t_i^{(1)} - t_{i'}^{(1)}$ = function $(\sigma_{12}^{(o)}/\sigma_{22}^{(o)})$, it is likely that $t_i^{(1)} - t_{i'}^{(1)}$ has greater precision in finite samples. Thus, in actual application, it is advisable to present both estimates.

2.3 TESTS OF SIGNIFICANCE, AND THE ANALOGY WITH THE USUAL COVARIANCE METHOD

In using $\hat{\tau}_i$ in estimating τ_i, the obvious estimate of $\sigma_{\epsilon_1}^{2} = \sigma_{11} - \sigma_{12}^{2}/\sigma_{22}$ is $s_{\epsilon_1}^{*2} = s_{11} - s_{12}^{2}/s_{22}$. However, the latter, found from a nonlinear combination of the $s_{hh'}$, is biased; making the adjustment $(qr - q - r + 1)\, s^{*2}_{\epsilon_1}/(qr - q - r) = s^2_{\epsilon_1}$, $s^2_{\epsilon_1}$ is unbiased for $\sigma^2_{\epsilon_1}$. Then the null hypothesis $\tau_1 = . . = \tau_q = 0$ is rejected if

$$\left[(qr - r - 1)\tilde{s}^2_{\epsilon_1} - (qr - q - r)s^2_{\epsilon_1}\right] \Big/ (q - 1)s^2_{\epsilon_1} > F_\alpha (q - 1, qr - q - r) \qquad (2.3.1)$$

where

$$(qr - r - 1)\tilde{s}^2_{\epsilon_1} = \sum_{ij} (y_{1ij} - m - b_j)^2$$

$$- \left[\sum_{ij} (y_{1ij} - m - b_j)(y_{2ij} - m_2 - b_{2j})\right]^2 \Big/ \sum_{ij} (y_{2ij} - m_2 - b_{2j})^2$$

is the estimate of $(qr - r - 1)\,\sigma^2_{\epsilon_1}$ with $\tau_i = 0$, and $F_\alpha (q - 1, qr - q - r)$ is the upper α critical value of the F distribution with $q - 1$ and $qr - q - r$ degrees of freedom. Note that if ϕ in (2.2.7) is negligible and $\hat{\tau}_i - \hat{\tau}_{i'}$ is sufficiently normal, then the test in (2.3.1) is nearly exact rather than asymptotic.

In the usual covariance analysis [Anderson and Bancroft (1952)], the estimate of τ_i is identically $\hat{\tau}_i$ in (2.2.6), while the adjusted treatment mean square is tested by (2.3.1). However, assuming the covariable fixed, the treatment contrast precision is

$$2\sigma^2_{\epsilon_1} \Big/ r + (\bar{y}_{2i} - \bar{y}_{2i'})^2 \; \sigma^2_{\epsilon_1} \Big/ \sum_{ij} (y_{2ij} - m_2 - t_{2i} - b_{2j})^2; \qquad (2.3.2)$$

in the limit and under (2. 1. 2), (2. 3. 2) reduces to (2. 1. 10), for then, the variance of the estimated slope is known and $\bar{y}_{2i} - \bar{y}_{2i'}$ becomes $E(\bar{y}_{2i}) - E(\bar{y}_{2i'}) = 0$. For uncontrolled covariables, (2. 3. 2) is an asymptotic result and is therefore comparable to (2. 1. 10), not to the finite sample result in (2. 2. 7). Hence, under (2. 1. 2) there results not only an increase in relative efficiency, but also, if Φ is negligible, this increase applies to finite samples.

Herein, one result is that τ_i is estimated by

$$\hat{\tau}_i = (\bar{y}_{1i} - \bar{y}_1) - (s_{12}/s_{22})(\bar{y}_{2i} - \bar{y}_2)$$

whether model (2. 1. 1) is considered alone (the usual covariance analysis) or whether the entire system in (2. 1. 1) and (2. 1. 2) is utilized. In the former case, the covariable is assumed fixed, and the estimated treatment effects are adjusted for differences in the y_2 between experimental units. When the entire system of (2. 1. 1) and (2. 1. 2) is applied, the adjustment is somewhat different. Under (2. 1. 2) the covariable is independent of treatments. However, $\bar{y}_{2i} - \bar{y}_2$, the estimate of τ_{2i}, is non-zero (with probability equal to 1) so that $\bar{y}_{2i} - \bar{y}_2$ becomes an estimable within sample bias. Since the errors of the reduced system in (2.1.4) and (2.1.2) are correlated (when $\alpha \neq 0$), $\bar{y}_{2i} - \bar{y}_2$ should be accounted for in the estimation of τ_i. Thus $\bar{y}_{1i} - \bar{y}_1$ is adjusted by an amount $-\alpha(\bar{y}_{2i} - \bar{y}_2)$ to produce the proper estimate of τ_i.

Example 2.1. Williams (1961, p. 119) describes an experiment on the effect of temperature on the maximum compressive strength of timber specimen. Material from ten trees was taken and a specimen from each tree was tested at each of five temperatures. The moisture content of each specimen was uncontrolled so that a covariance adjustment was made to the data. Williams' detailed analysis illustrates, in part, that the residual variability was reduced substantially through the introduction of the covariable and that temperature effects were highly significant.

If the trees (or blocks) in this example were fixed, then the structural system in (2.1.1) and (2.1.2) applies since moisture content is independent of temperature (moisture content is measured prior to the application of temperature) and follows the model in (2.1.2). However, the fact that trees are random will be disregarded in the same manner that least squares estimates are utilized in mixed and random models. As such, the treatment contrast precision in this example is, from (2.2.7),

$$(2/r)\,\sigma^2_{\epsilon_1}\,(1+\Phi) = (1/5)\,\sigma^2_{\epsilon_1}\,(1+\Phi)$$

where

$$\Phi = 34^{-1}\,\pi^{-1/2}\,\Gamma(18.5)/\,\Gamma(18) = 0.069.$$

Since Φ is negligible, the treatment contrast precision becomes $(1/5)\,\sigma^2_{\epsilon_1}$ rather than the usual result given by (2.3.2).

2.4 ON INCREASING TREATMENT CONTRAST PRECISION IN BALANCED INCOMPLETE BLOCK DESIGNS

Consider the incomplete block design model

$$y_{1ij} = \Delta_{ij}\,(\mu + \tau_i + \beta_j + \epsilon_{ij}) \qquad (2.4.1)$$

where $\Delta_{ij} = 0$ or 1; $i = 1, \ldots, q$; $j = 1, \ldots, r$. Let u denote the number of times a treatment is replicated; let v denote the number of plots per block, and let every treatment appear with every other treatment in the same block an equal number of times, say w. Thus, we have a balanced incomplete block (BIB) design and τ_i is estimated by

$$t'_i = \sum_j y_{1ij}/r - \sum_j \Delta_{ij} \sum_i y_{1ij}/rv. \qquad (2.4.2)$$

Denoting the variance of the model error in (2.4.1) by σ_{11}, the treatment contrast precision is

$$\mathrm{var}\,(t'_i - t'_{i'}) = 2\,\sigma_{11}/r\,E_f \qquad (2.4.3)$$

where the efficiency factor $E_f = [u(v - 1) + w]/uv$ and $E_f < 1$ [Anderson and Bancroft (1952)]. In (2.4.1) if $\Delta_{ij} = 1$ for all (i, j), then we have a randomized block (RB) design and the contrast precision is

$$2\,\sigma_{11}/r \;<\; 2\,\sigma_{11}/r\,E_f.$$

However, if blocks become heterogenous when the number of plots per block equals the number of treatments, then a comparison of the contrast precisions between the two designs is misleading; i. e., assuming heterogeneity within blocks, the estimated σ_{11} for the RB design becomes larger than the estimated σ_{11} for the BIB design.

It will now be shown that it is possible to utilize the BIB design (described by model (2. 4. 1)) and at the same time to achieve a treatment contrast precision which is nearly identical (if not greater than) the $2\,\sigma_{11}/r$ value for the RB design.

In all the design of the experiment, another response, say y_2, must be identified, where y_2 is independent of treatments and is highly correlated with y_1. Then, if y_2 follows the model

$$y_{2ij} = \Delta_{ij}\,(\mu_2 + \beta_{2j} + \epsilon_{2ij}) \qquad (2.4.4)$$

while

$$y_{1ij} = \Delta_{ij} (\mu_1 + \tau_i + \beta_{ij} + \alpha y_{2ij} + \epsilon_{2ij}), \tag{2.4.5}$$

we may apply the results of the previous sections. The reduced system corresponding to (2.4.5) and (2.4.4) is

$$y_{1ij} = \Delta_{ij} [(\mu_1 + \alpha\mu_2) + \tau_i + (\beta_{1j} + \alpha\beta_{2j}) + (\epsilon_{1ij} + \alpha\epsilon_{2ij})] \tag{2.4.6}$$

and (2.4.4). Letting $\mu = \mu_1 + \alpha\mu_2$, $\beta_j = \beta_{1j} + \alpha\beta_{2j}$, and $\epsilon_{ij} = \epsilon_{1ij} + \alpha\epsilon_{2ij}$, it is seen that the model in (2.4.6) is identical to the model in (2.4.1). Let $\text{var}(\epsilon_{1ij} + \alpha\epsilon_{2ij}, \epsilon_{2ij}) = (\sigma_{hh'})$, $h, h' = 1, 2$. Then, considering the entire reduced system, the Markoff estimate of τ_i is

$$t_i = \sum_j y_{1ij}/r - \sum_j \Delta_{ij} \sum_i y_{1ij}/rv$$

$$- (\sigma_{12}/\sigma_{22}) \left[\sum_j y_{2ij}/r - \sum_j \Delta_{ij} \sum_i y_{2ij}/rv\right]$$

and

$$\text{var}(t_i - t_{i'}) = (2/r)(\sigma_{11} - \sigma_{12}^2/\sigma_{22})/E_f$$

$$= (2\sigma_{11}/r)(1 - \rho^2)/E_f \tag{2.4.7}$$

The parameter σ_{12}/σ_{22} is replaced by an estimate, say s_{12}/s_{22}, precisely as in Section 2.2; i.e., including a treatment effect, say $\tau_{2i} = 0$, in (2.4.4), then the residual squares and cross products yield the estimates s_{22} and s_{12}. Thus τ_i is estimated by $\hat{\tau}_i$ where

$$\hat{\tau}_i = \sum_j y_{1ij}/r - \sum_j \Delta_{ij} \sum_i y_{1ij}/rv$$

$$- (s_{12}/s_{22})\left[\sum_j y_{2ij}/r - \sum_j \Delta_{ij} \sum_i y_{2ij}/rv\right] \tag{2.4.8}$$

and the exact second moment of $\hat{\tau}_i - \hat{\tau}_{i'}$ is

$$\text{var}\,(\hat{\tau}_i - \hat{\tau}_{i'}) = (2\sigma_{11}/r)(1-\rho^2)(1+\phi)/E_f \tag{2.4.9}$$

where

$$\phi = (qr-q-r-1)^{-1}\pi^{-1/2}\,\Gamma\,[(qr-q-r+2)/2] \Big/ \Gamma[(qr-q-r+1/2]$$

is the finite sample correction factor as in (2.2.7). Note that if ϕ is negligible, then (2.4.9) reduces to (2.4.7). Thus, if ρ^2 is sufficiently large, to the extent that $(1-\rho^2)$ and E_f cancel, then we have $(2/r)\,\sigma_{11}$, the contrast precision for a RB design.

3. THE QUESTION OF DIRECT AND INDIRECT EFFECTS WHEN THE TREATMENTS AFFECT THE COVARIABLE AND $\sigma_{\epsilon_1 \epsilon_2} = 0$

3.1 THE ESTIMATION OF DIRECT AND INDIRECT EFFECTS WHEN ALL TREATMENTS HAVE A POSSIBLE EFFECT ON THE COVARIABLE

Consider the structural system (described in Section 1)

$$y_{1ij} = \mu_1 + \tau_{1i} + \beta_{1j} + \alpha y_{2ij} + \epsilon_{1ij} \qquad (3.1.1)$$

$$y_{2ij} = \mu_2 + \tau_{2i} + \beta_{2j} + \epsilon_{2ij}. \qquad (3.1.2)$$

The reduced models are (3.1.2) and

$$y_{1ij} = (\mu_1 + \alpha\mu_2) + (\tau_{1i} + \alpha\tau_{2i}) + (\beta_{1j} + \alpha\beta_{2j}) + (\epsilon_{1ij} + \alpha\epsilon_{2ij}). \qquad (3.1.3)$$

Let the sample form of (3.1.2) and (3.1.3) be written as

$$y_{1ij} = m + t_i + b_j + e_{1ij} \qquad (3.1.4)$$

and

$$y_{2ij} = m_2 + t_{2i} + b_{2j} + e_{2ij} \qquad (3.1.5)$$

where m, t_i, b_j and e_{1ij} correspond to $\mu_1 + \alpha\mu_2$, $\tau_{1i} + \alpha\tau_{2i}$, $\beta_{1j} + \alpha\beta_{2j}$, and $\epsilon_{1ij} + \alpha\epsilon_{2ij}$, respectively. The assumptions regarding the ϵ's in (3.1.1) and (3.1.2) are identical to those in (2.1.3). Adding the restriction that $\sum_i \tau_{2i} = 0$ to those in (2.1.8), the Markoff estimates become

$$m = \bar{y}_1, \qquad t_i = (\bar{y}_{1i} - \bar{y}_1), \qquad b_j = (\bar{y}_{1j} - \bar{y}_1) \qquad (3.1.6)$$

$$m_2 = \bar{y}_2, \qquad t_{2i} = (\bar{y}_{2i} - \bar{y}_2), \qquad b_{2j} = (\bar{y}_{2j} - \bar{y}_2),$$

with $\mathrm{var}(t_i - t_{i'}) = (2/r)(\sigma^2_{\epsilon_1} + \alpha^2\sigma^2_{\epsilon_2})$ and $\mathrm{var}(t_{2i} - t_{2i'}) = 2\sigma^2_{\epsilon_2} \Big/ r$. Note that the estimates in (3.1.6) do not contain elements of the unknown $\sum$, a convenience which results when the design matrices of two or more linear, reduced models, with correlated errors, are identical [Mallios (1961)].

Since treatments may have direct and indirect effects on y_1, the parameters of the structural system need be estimated. τ_{2i} is estimated directly by t_{2i}; equating t_i to $\tau_{1i} + \alpha\tau_{2i}$ and substituting t_{2i} for τ_{2i}, there results q - 1 equations in q unknowns. But from the estimate of $\sum$, given by the $s_{hh'}$ of Section 2.2, we have three equations, s_{11}, s_{12}, s_{22}, in three unknowns $\sigma^2_{\epsilon_1}$, $\sigma^2_{\epsilon_2}$ and α. From the estimate of α, given by $a = s_{12}/s_{22}$, we can estimate τ_{1i}, the direct treatment effect on y_1, by $t_{1i} = t_i - at_{2i}$. But t_{1i} is precisely $\hat{\tau}_i$ in (2.2.6), so that the usual covariance technique provides proper tests and estimates of the direct treatment effect on y_1, under the structural system in (3.1.1) and (3.1.2); moreover, var $(t_{1i} - t_{1i'})$ is given in (2.2.7).

Thus, if the t_{2i} are significant while the $t_{1i} = \hat{\tau}_i$ are nonsignificant, then the treatment effect on y_1 is indirect; i.e., the treatment affects y_1 through the covariable and not directly. If both the t_{1i} and t_{2i} are significant, then there are both direct and indirect treatment effects on y_1.

Example 3.1. Anderson and Bancroft (1952, p. 302) discuss an experiment on the effect of fertilizer levels on the yield of sugar beets. The covariable is stand, which may be influenced by fertilizer though not by yield; and the field is divided into six homogeneous blocks. The analysis shows that treatment effects on yield, adjusted for stand, are not significant, while treatment effects on stand are highly significant. Disregarding the model in (3.1.2), the experiment certainly falls into the "uncertain class". However, stand is adequately predicted by blocks and treatments, as is hypothesized by model (3.1.2). Thus the analysis implies that there is an indirect rather than a direct treatment effect on yield; i.e., the treatment effect on yield is through the stand. Though there is little need to consider treatment contrast precision for direct effects on yield, the value of Φ is given by

$$(28)^{-1}\pi^{-1/2}\,\Gamma(15.5)/\Gamma(15) = 0.077.$$

Treatment contrast precision for direct treatment effect on stand is $2\sigma_{\epsilon_2}^{2}/6$ whose estimate is 318.5.

3.2 AN ILLUSTRATION WHEREIN THE TREATMENTS DEFINE A FACTORIAL EXPERIMENT AND THE COVARIABLE IS INDEPENDENT OF ONE FACTOR

Scheffe (1959, p. 217) presents data from an experimental piggery where six young pigs, three male and three female, were allotted to each of five pens. Three amounts of protein, say, f_1, f_2, and f_3, in increasing proportion, were given to one male and one female in each pen. The pigs were weighed individually each week for 16 weeks, and the growth rate (y_1) in pounds per week was calculated for each pig. The weight (y_2) of each pig at the beginning of the experiment is the measured covariable.

For purposes of adding to homogeneity within pens, the allotment of pigs to pens is by initial weight. Within each of the two groups of 15 male and 15 female pigs, the pigs are ordered from highest to lowest according to initial weight; i.e.,

$W_{M1} \geq W_{M2} \geq \ldots \geq W_{M,15}$ are the 15 male weights

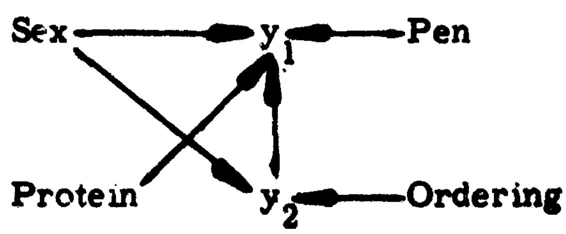
and

$W_{F1} \geq W_{F2} \geq \ldots \geq W_{F,15}$ are the 15 female weights.

The three heaviest males and the three heaviest females are assigned to the first pen, the next three males and the next three females are assigned to the second pen, etc.

Figure 3.1 presents a path diagram relating sex, protein, ordering, and pen to growth rate and initial weight. The arrow from y_2 to y_1 describes a possible effect of initial weight on growth rate. Sex may have a direct effect on y_1 and y_2, and hence an indirect effect on y_1. Also, protein may effect y_1, though certainly not y_2, and pen may have an effect on y_1.

Figure 3.1
A Path Diagram Relating Sex, Protein, Ordering, and Pen to Growth Rate (y_1) and Initial Weight (y_2)



The five ordered groups will, very likely, yield a significant source of variation for y_2. These groups can be looked upon as five fixed blocking effects on y_2 in the same way that pens are blocking effects on y_1. However, it must be assumed that the 30 selected pigs are representative of the population of pigs with respect to initial weight.

From Figure 3.1, the following linear structural models, in sample form, are hypothesized:

$$y_{1ijk} = m_1 + s_{1i} + f_j + (s_1 f)_{ij} + b_{1k} + ay_{2ijk} + e_{1ijk} \qquad (3.2.1)$$

$$y_{2ijk} = m_2 + s_{2i} + b_{2k} + e_{2ijk} \qquad (3.2.2)$$

where s_i : sex i = 1, 2; f_j : protein level, j = 1, 2, 3; b_{1k} : pen, k = 1, ... , 5; b_{2k}: ordering, k = 1, 2, ... , 5. The reduced models become (3. 2. 2) and

$$\begin{aligned} y_{1ijk} &= (m_1 + am_2) + (s_{1i} + as_{2i}) + f_j + (s_1 f)_{ij} \\ &\quad + (b_{1k} + ab_{2k}) + (e_{1ijk} + ae_{2ijk}) \qquad (3.2.3) \\ &= m + s_i + f_j + (s_1 f)_{ij} + b_k + e_{ijk}. \end{aligned}$$

under the restrictions that

$$\sum_i s_{hi} = \sum_j f_j = \sum_{ij} (s_i f)_{ij} = \sum_k b_{hk} = 0, h = 1, 2. \qquad (3.2.4)$$

The assumptions regarding $\epsilon_{1ijk} + \alpha\epsilon_{2ijk}$ and ϵ_{2ijk}, the population errors corresponding to e_{ijk} and e_{2ijk}, are that

$$\begin{bmatrix} \epsilon_{1ijk} + \alpha\epsilon_{2ijk} \\ \epsilon_{2ijk} \end{bmatrix} : \text{i.i.d.} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma \right)$$

where $\sum = (\sigma_{hh'})$ is given in (2. 1. 6).

Analogous to the derivation of $\hat{\tau}_i$ in (2. 2. 6), the fixed effects are estimated as follows. Include a protein effect, say f_{2j}, and a protein by sex interaction effect, say $(s_2f_2)_{ij}$, in (3. 2. 2), i. e.,

$$y_{2ijk} = m_2 + s_{2i} + f_{2j} + (s_2f_2)_{ij} + b_{2k} + \tilde{e}_{2ijk}$$

where $E(f_{2j}) = E(s_2f_2)_{ij} = 0$. Then $a = \hat{s}_{12}/\hat{s}_{22}$, where the $\sigma_{hh'}$ estimators are found from the sum of residual squares and cross products; e. g.,

$$\hat{s}_{12} = \sum_{ijk} e_{ijk}\tilde{e}_{2ijk}/20 .$$

Substituting $\hat{s}^{hh'}$ for $\sigma^{hh'}$ in

$$\sum_{ijk} (e_{ijk}^2 \sigma^{11} + e_{ijk} e_{2ijk} \sigma^{12} + e_{2ijk}^2 \sigma^{22})$$

and holding the $\hat{s}^{hh'}$ fixed in differentiation, then, under (3.2.4), the estimated fixed effects are

$$m_1 = \bar{y}_1 - a\bar{y}_2 ,$$

$$s_{1i} = (\bar{y}_{1i} - \bar{y}_1) - a(\bar{y}_{2i} - \bar{y}_2) ,$$

$$f_j = (\bar{y}_{1j} - \bar{y}_1) - a(\bar{y}_{2j} - \bar{y}_2) \tag{3.2.5}$$

$$(s_1 f)_{ij} = (\bar{y}_{1ij} - \bar{y}_1) - a(\bar{y}_{2ij} - \bar{y}_2) - s_i - f_i$$

$$b_{1k} = (\bar{y}_{1k} - \bar{y}_1) - a(\bar{y}_{2k} - \bar{y}_2)$$

$$m = \bar{y}_1 ,\ m_2 = \bar{y}_2 ,\ s_i = \bar{y}_{1i} - \bar{y}_1 ,\ s_{2i} = \bar{y}_{2i} - \bar{y}_2 \tag{3.2.6}$$

$$b_k = (\bar{y}_{1k} - \bar{y}_1) ,\ b_{2k} = (\bar{y}_{2k} - \bar{y}_2) .$$

The efficiency of $s_{1i} - s_{1i'}$ is $2\sigma^2_{\epsilon_1}/15$ while its exact second moment is $(2\sigma^2_{\epsilon_1}/15)(1+\phi)$ where ϕ is given in (2.2.7). Similarly, the efficiency of $f_j - f_{j'}$ is $2\sigma^2_{\epsilon_1}/10$ and its exact second moment is $(2\sigma^2_{\epsilon_1}/10)(1+\phi)$.

Note that the estimators in (3.2.5) are precisely the estimates obtained through the usual covariance method which considers only the model in (3.2.1). By introducing (3.2.2), there results, in estimation, a subtraction of $a(\bar{y}_{2i} - \bar{y}_2)$ in the expression for s_{1i}, which yields the direct sex effect of y_1. The subtraction of comparable terms in the expressions for f_j and $(s_1f)_{ij}$ adjusts for within sample bias, and serves to increase contrast precision (as is discussed in Section 2.3).

The analysis, given in Table 3.1, illustrates the consequences of three approaches to this problem: (i) a model is hypothesized for the covariable, though it includes $f_{2j} + (s_2f_2)_{ij}$; hence the design matrices of the reduced models are identical and separate estimation is identical to joint estimation; (ii) only the model in (3.2.1) is considered; (iii) the structural system of (3.2.1) and (3.2.2) is hypothesized.

Under (i), the results are decidely conservative. Tests of effects, for each model separately, are shown in the unadjusted mean square columns of Table 3.1. Here, neither treatments nor pens have a significant effect on growth rate, while ordering has a highly significant effect on initial weight. Under approach (ii), the usual covariance method, the adjusted treatment mean square for growth rate is significant at the five percent level, and this is due to protein levels. The estimated slope $a = 0.088$, and the estimated variance of a is $(0.253)(442.933)^{-1}$. Finally, under (iii), statements regarding treatment effects on y_1 are identical to those under approach (ii). In addition, sex has a negligible effect, not only on growth rate, but on initial weight. And from the unadjusted mean square column for y_1, it is seen that the model for initial weight, given in (3.2.2), is adequate with $s_{2i} = 0$. Protein contrast precision is estimated by $2s^2_{\epsilon_1}/15 = 0.034$; however, $N = 20$ and $\phi = (18)^{-1}\pi^{-1/2}\Gamma(10.5)/\Gamma(10) = 0.097$ is negligible, so that the latter precision and the tests of significance apply to finite samples.

In Section 2.2 four treatment effect estimates were discussed: $t^{(o)}$, the unadjusted, unbiased conservation estimate; $t^{(1)}$, the efficient estimate based on one cycle of iteration; $\hat{\tau}$, the usual covariance estimate; and t^*, the maximum likelihood estimate. Table 3.2 presents the three protein means based on these four estimates, i.e., $f^{(o)} + \bar{y}_1$, $f^{(1)} + \bar{y}_1$, $f + \bar{y}_1$, $f^* + \bar{y}_1$ correspond to $t^{(o)}$, $t^{(1)}$, $\hat{\tau}$, and t^*, respectively.

Table 3.1
Analysis of Covariance of Growth Rate (y_1) and Initial Weight (y_2)

Source	d.f.	Unadjusted S.S.			Unadjusted M.S.		Adjusted S.S.	Adjusted M.S.
		y_1	$y_1 y_2$	y_2	y_1	y_2	y_1	y_1
Pens (Orderings)	4	4.851	39.905	605.867	1.212	151.466††		
Treatments	5	3.179	-0.765	59.900	0.635	11.980	3.714	0.742†
Sex	1	0.434	-3.730	32.034	0.434	32.034	0.969	0.969
Protein	2	2.268	-0.147	5.411	1.134	2.705	2.804	1.402†
Sex x Protein	2	0.476	3.112	22.455	0.238	11.227	1.011	0.505
Residual	20	8.314	39.367	442.933	0.415	22.146		
Regression	1						3.498	3.498††
Adjusted Residual	19						4.816	0.253

†† Sig. at 1% Level.
† Sig. at 5% Level.

Table 3.2

Adjusted and Unadjusted Protein Means

Protein Level	$\bar{y}_2$	$f^{(o)} + \bar{y}_1$	$f^{(1)} + \bar{y}_1$	$f + \bar{y}_1$	$f^* + \bar{y}_1$
1	39.8	9.649	9.678	9.675	10.061
2	40.7	9.288	9.154	9.235	9.134
3	39.8	8.976	8.707	9.002	8.717

4. THE ESTIMATION OF STRUCTURAL PARAMETERS WHEN $\sigma_{\epsilon_1\epsilon_2} \neq 0$

4.1 PRELIMINARY REMARKS

In Section 2.2, the reduced system

$$y_{1ij} = (\mu_1 + \alpha\mu_2) + \tau_i + (\beta_{1j} + \alpha\beta_{2j}) + (\epsilon_{1ij} + \alpha\epsilon_{2ij}) \qquad (4.1.1)$$

$$y_{2ij} = \mu_2 + \beta_{2j} + \epsilon_{2ij} \qquad (4.1.2)$$

which corresponds to the structural system in (2.1.1) and (2.1.2) was considered under the assumption that

$$\operatorname{var}\begin{bmatrix} \epsilon_{1ij} + \alpha\epsilon_{2ij} \\ \epsilon_{2ij} \end{bmatrix} = \begin{bmatrix} \sigma^2_{\epsilon_1} + \alpha^2\sigma^2_{\epsilon_2}, & \alpha\sigma^2_{\epsilon_2} \\ \alpha\,\sigma^2_{\epsilon_2}, & \sigma^2_{\epsilon_2} \end{bmatrix}. \qquad (4.1.3)$$

If $\sigma_{\epsilon_1\epsilon_2} \neq 0$, then

$$\operatorname{var}\begin{bmatrix} \epsilon_{1ij} + \alpha\epsilon_{2ij} \\ \epsilon_{2ij} \end{bmatrix} = \begin{bmatrix} \sigma^2_{\epsilon_1} + 2\alpha\sigma_{\epsilon_1\epsilon_2} + \alpha^2\sigma^2_{\epsilon_2}, & \sigma_{\epsilon_1\epsilon_2} + \alpha\sigma^2_{\epsilon_2} \\ \sigma_{\epsilon_1\epsilon_2} + \alpha\sigma^2_{\epsilon_2}, & \sigma^2_{\epsilon_2} \end{bmatrix} \qquad (4.1.4)$$

and the treatment contrast precision is still increased; i. e.,

$$\tau_i = \bar{y}_{1i} - \bar{y}_1 - (s_{12}/s_{22})(\bar{y}_{2i} - \bar{y}_2)$$

and

$$\text{var}\,(\hat{\tau}_i - \hat{\tau}_{i'}) = (2\,\sigma_{11}/r)(1 - \rho^2)(1 + \phi)$$

$$= (2\,\sigma^2_{\epsilon_1} \Big/ r)(1 + \phi) \quad \text{if} \quad \sigma_{\epsilon_1\epsilon_2} = 0\,,$$

where s_{12}/s_{22} is that estimate of σ_{12}/σ_{22} for which finite sample properties are available. However,

$$\sigma_{12}/\sigma_{22} = (\sigma_{\epsilon_1\epsilon_2} + \alpha\sigma^2_{\epsilon_2}) \Big/ \sigma^2_{\epsilon_2} \quad \text{if} \quad \sigma_{\epsilon_1\epsilon_2} \neq 0$$

$$\sigma_{12}/\sigma_{22} = \alpha \quad \text{if} \quad \sigma_{\epsilon_1\epsilon_2} = 0\,.$$

Moreover, within the realm of the data and assuming no further prior knowledge, α cannot be estimated by existing techniques due to under-identification; i. e., in model (4. 1. 1) there are $q + 2r$ unknown parameters $(\mu_1, \mu_2, \alpha, \tau_1, \ldots, \tau_{q-1}, \beta_{11}, \ldots, \beta_{1,r-1}, \beta_{21}, \ldots, \beta_{2,r-1})$, and in the covariance matrix of (4. 1. 4) there are an additional three parameters ($\sigma^2_{\epsilon_1}$, $\sigma^2_{\epsilon_2}$, and $\sigma_{\epsilon_1\epsilon_2}$), so that the total number of unknown parameters is $q + 2r + 3$; but from the reduced system of

(4.1.1) and (4.1.2) and from the estimate of the covariance matrix in (4.1.4), there are only $q+2r+2$ equations; and it is easily seen that the underidentified parameters are α, $\sigma^2_{\epsilon_1}$, and $\sigma_{\epsilon_1\epsilon_2}$.

The estimation of these underidentified parameters may be of importance for the following reasons.

(1) Since $\hat{\tau}_i$ = function (s_{12}/s_{22})

and since $\sigma_{12}/\sigma_{22} = (\sigma_{\epsilon_1\epsilon_2} + \alpha\sigma^2_{\epsilon_2}) \Big/ \sigma^2_{\epsilon_2}$,

it may be that $\sigma_{\epsilon_1\epsilon_2}$ $(\neq 0)$ is approximately equal to $\alpha\sigma^2_{\epsilon_2}$ $(\neq 0)$ but opposite in sign. In this case, σ_{12} is approximately zero, and from the sample, one might be led to the mistaken conclusion that y_1 is independent of y_2, or nearly so. Thus, if a resolvement of the structure of the experimental unit is the issue, then existing techniques should be applied and presented with caution.

(2) If $\sigma_{\epsilon_1\epsilon_2} \neq 0$, then there is the very basic question, "Why are the errors of the structural system correlated?" It may be that the structural models are inadequate in that other important variables have been neglected, in which case $\sigma_{\epsilon_1\epsilon_2} \neq 0$. On the other hand, it may be that the

structural models are, in fact, adequate and that $\sigma_{\epsilon_1 \epsilon_2} \neq 0$ is due to the extraneous effects of the infinity of variables which can be measured from an experimental unit.

In what is to follow, an estimation technique is discussed whereby underidentified parameters are estimable. However, no pretense is made that the questions posed in (1) and (2) can, at present, be adequately resolved or even nearly so. All that is done is to suggest an approach which may be of some value in the actual consideration of (1) and (2).

4.2 AN EXTENSION OF THE "INSTRUMENTAL VARIABLE ESTIMATION TECHNIQUE"

Consider the functional relationship [Williams, (1961, Chapter 11)]

$$\eta = \gamma \xi \tag{4.2.1}$$

where γ is a parameter to be estimated, and η and ξ are measured with error by

$$y = \eta - \epsilon , \qquad x = \xi + \delta . \tag{4.2.2}$$

$y = \gamma x + (\epsilon + \gamma \delta)$ is obtained by substituting the expressions in (4.2.2) into (4.2.1). In a sample of, say, size n, the least squares estimate of γ may be inconsistent, since x is correlated with the error term $\epsilon + \gamma\delta$. As such, an alternative method of estimation is now discussed.

Assume

$$\begin{bmatrix} \epsilon_i \\ \\ \delta_i \end{bmatrix} : \text{i.i.d.} \left(\begin{bmatrix} 0 \\ \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\epsilon^2 & 0 \\ \\ 0 & \sigma_\delta^2 \end{bmatrix} \right) ,$$

$i = 1, \ldots, n$, so that $(\epsilon_i + \gamma\delta_i)$: i.i.d. $(0, \sigma_\epsilon^2 + \gamma^2\sigma_\delta^2)$. From the i^{th} experimental unit from which the responses y_i and x_i are drawn, let u_i denote another variable which will be termed an instrumental variable. Assume that the functional relationship in (4.2.1) is invariant under changes in extraneous (instrumental) variables and that the measurement errors are uncorrelated with the instrumental variable. Then in a regression of $\epsilon_i + \gamma\delta_i$ on u_i, say,

$$\epsilon_i + \gamma\delta_i = \alpha_o + \alpha_1 u_i + \Delta_i ,$$

$\alpha_o = \alpha_1 = 0$, where α_o, α_1, Δ are intercept, slope, and model error, respectively. In the sample, the estimate of α_1 is

$$a_i = \sum_i (\epsilon_i + \gamma\delta_i)\, u_i \Big/ \sum_i u_i^2 ,$$

assuming $\bar{u} = 0$. But $\alpha_1 = 0$, and equating a_1 to zero yields

$$\sum_i (\epsilon_i + \gamma\delta_i)\, u_i = 0 = \sum_i (y_i - \gamma x_i)\, u_i$$

whence

$$\hat{\gamma} = \sum_i y_i u_i \Big/ \sum_i x_i u_i \tag{4.2.3}$$

is a consistent estimate of γ. $\hat{\gamma}$ is termed the instrumental variable estimate of γ.

We will now show why the instrumental variable estimation technique cannot, in its present context, be applied to the estimation of parameters in a regression relationship where the independent variable is uncontrolled and measured with error. Consider the regression model

$$\eta = \gamma \xi + \epsilon^{+} \tag{4.2.4}$$

where the intercept is assumed zero, η and ξ are measured with error according to (4.2.2) and ϵ^{+} is the model error. Substituting the expressions in (4.2.2) into (4.2.4), we have

$$y = \gamma x + (\epsilon + \gamma\delta + \epsilon^{+}) . \tag{4.2.5}$$

Again, in a sample of size n, the usual least squares estimate of γ may be inconsistent since x and δ are correlated. In addition, the error term $\epsilon + \gamma\delta + \epsilon^{+}$ can hardly be assumed independent of extraneous variables since the model error ϵ^{+} is composed of variables such as u; i.e., the regression relationship is not invariant under changes in extraneous or instrumental variables.

To apply instrumental variable estimation to a regression relationship, we must find a variable u such that $\epsilon + \gamma\delta + \epsilon^+$ is independent of u. Such a variable u is available from two sources, a table of random numbers and disassociated experiments; i. e. $\epsilon + \gamma\delta + \epsilon^+$ is independent of u if the u_i are drawn from a table of random numbers; also, if the y_i and x_i in (4. 2. 5) correspond to, say, a biological experiment and the u_i are responses taken from an unrelated industrial experiment, then $\epsilon + \gamma\delta + \epsilon^+$ is again independent of u. Consequently, the instrumental variable estimation technique can be applied to a regression relationship (where the independent variable is subject to measurement error and is uncontrolled) if the u_i are properly chosen.

If the independent variable, ξ, in (4. 2. 4) is controlled and/or, if the measurement error $\delta = 0$, then the usual least squares estimate of γ, say $c = \sum_i x_i y_i \Big/ \sum_i x_i^2$, is the best estimate and

$$\text{var } c = \sigma^2 \Big/ \sum_i x_i^2 \qquad (4.2.6)$$

where $\sigma^2 = \sigma_\epsilon^2 + \sigma_{\epsilon^+}^2$. Regressing $\epsilon_i + \epsilon_i^+$ on u_i produces the estimate of γ as given in (4. 2. 3), where

$$\text{var } \hat{\gamma} = \sigma^2 \sum_i u_i^2 \Big/ \Big(\sum_i u_i x_i\Big)^2 . \qquad (4.2.7)$$

Comparing the result in (4.2.6) with that in (4.2.7), then we have the very obvious result that

$$\operatorname{var} \hat{\gamma} > \operatorname{var} c \qquad \text{where} \qquad u \neq x$$

$$= \operatorname{var} c \qquad \text{where} \qquad u = x, \quad \text{i.e.,}$$

$$\sigma^2 \sum u^2 \Big/ \left(\sum u x\right)^2 \geq \sigma^2 \Big/ \sum x^2$$

since

$$\left(\sum x u\right)^2 \Big/ \sum x^2 \sum u^2 \leq 1 .$$

The latter is a well known inequality.

Thus, we have shown that if least squares estimation (when applicable) is compared to instrumental variable estimation, the latter produces conservative estimates but has broader applicability.

4.3 AN EXAMPLE OF UNDERIDENTIFIED PARAMETERS AND THEIR ESTIMATES THROUGH INSTRUMENTAL VARIABLE ESTIMATION

In the previous section, the instrumental variable estimation technique was utilized in the estimation of parameters when there exis substantial measurement error. In this section, as in Sections 1, 2, and 3, we assume that measurement errors are negligible, and the instrumental variable estimation technique is applied for the expressed purpose of estimating underidentified parameters.

Consider the structural system

$$y_{1i} = \mu_1 + \alpha y_{2i} + \epsilon_{1i} \tag{4.3.1}$$

$$y_{2i} = \mu_2 + \epsilon_{2i} \tag{4.3.2}$$

$$= m_2 + e_{2i} \tag{4.3.3}$$

$i = 1, \ldots, n$, where μ_1 and μ_2 are population means; (4. 3. 3) is the sample form of (4. 3. 2); α is the rate of change in y_1 per unit change in y_2; and ϵ_{1i} and ϵ_{2i} are model errors. It is assumed that

$$\begin{bmatrix} \epsilon_{1i} \\ \epsilon_{2i} \end{bmatrix} : \text{i. i. d.} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma^2_{\epsilon_1} & \sigma_{\epsilon_1 \epsilon_2} \\ \sigma_{\epsilon_1 \epsilon_2} & \sigma^2_{\epsilon_2} \end{bmatrix} \right).$$

The reduced system corresponding to (4. 3. 1) and (4. 3. 2) is

$$y_{1ij} = (\mu_1 + \alpha \mu_2) + (\epsilon_{1i} + \alpha \epsilon_{2i}) \tag{4.3.4}$$

$$= \mu + \epsilon_i = m + e_i \tag{4.3.5}$$

and (4.3.2), and the covariance matrix of the reduced model errors, $(\epsilon_{1i} + \alpha\epsilon_{2i}, \epsilon_{2i})$, as given in (4.1.4).

The Markoff estimates of $(\mu_1 + \alpha\mu_2) = \mu$, μ_2, $\sigma^2_{\epsilon_1} + 2\alpha\sigma_{\epsilon_1\epsilon_2} + \alpha^2\sigma^2_{\epsilon_2}$, $\sigma_{\epsilon_1\epsilon_2} + \alpha\sigma^2_{\epsilon_2}$, and $\sigma^2_{\epsilon_2}$ are, respectively,

$$\sum_i y_{1i} \Big/ n,\ \sum_i y_{2i}\ n,\ \sum_i (y_{1i} - \bar{y}_1) \Big/ (n-1),\ \sum_i (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2) \Big/ (n-1),$$

and $\sum_i (y_{2i} - \bar{y}_2)^2 \Big/ (n-1)$. Thus, μ_1, α, $\sigma_{\epsilon_1\epsilon_2}$, and $\sigma^2_{\epsilon_1}$ are under-identified, and the instrumental variable estimation technique is applied.

The n x 1 vectors $\underline{u}_1 = (u_{1i})$ and $\underline{u}_2 = (u_{2i})$ are drawn, say, from a table of random numbers. Let $U(n \times 2) = (\underline{u}_1, \underline{u}_2)$ and $\bar{u}_1 = \bar{u}_2 = 0$. Then in the regression model

$$\epsilon_{1i} = \alpha_o + \alpha_1 u_{1i} + \alpha_2 u_{2i} + \Delta_i, \qquad (4.3.6)$$

$\alpha_o = \alpha_1 = \alpha_2 = 0$, where ϵ_{1i} is the model error in (4.3.1) and Δ_i is the model error in (4.3.6). Let $\underline{\epsilon}_1 (n \times 1) = (\epsilon_{1i})$. Then the least squares estimates of α_1 and α_2, say

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \left[U'U\right]^{-1} U' \underline{\epsilon}_1,$$

are equated to zero, so that

$$[U' U]^{-1} U' (\underline{y}_1 - X\underline{\theta}_1) = 0$$

where $\underline{y}_1$ (n x 1) = (y_{1i}), X (n x 2) = $(\underline{1}, \underline{y}_2)$, $\underline{1}$ (n x 1) = (1), $\underline{y}_2$ (n x 1) = (y_{2i}), and $\underline{\theta}_1' = (\mu_1, \alpha)$. Thus,

$$\hat{\underline{\theta}}_1 = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\alpha} \end{pmatrix} = (U' X)^{-1} U' \underline{y}_1 \qquad (4.3.7)$$

$$\text{var } \hat{\underline{\theta}}_1 = (U' X)^{-1} (U' U)^{-1} (U' X)'^{-1} \sigma^2_{\epsilon_1} ,$$

and the estimate of $\sigma^2_{\epsilon_1}$ is

$$\hat{\sigma}^2_{\epsilon_1} = \sum_i e^2_{1i} \Big/ (n - 2) ,$$

where $e_{1i} = y_{1i} - \hat{\mu}_1 - \hat{\alpha} y_{2i}$.

The estimate of μ_2 is $\hat{\mu}_2 = \bar{y}_2$ so that $\sigma^2_{\epsilon_2}$ is estimated by $\hat{\sigma}^2_{\epsilon_2} = \sum_i e^2_{2i} \Big/ (n - 1)$, where $e_{2i} = y_{2i} - \bar{y}_2$. Thus $\sigma_{\epsilon_1 \epsilon_2}$ is estimated by $\hat{\sigma}_{\epsilon_1 \epsilon_2} = \sum_i e_{1i} e_{2i} \Big/ (n - 2)$.

A significant departure of $\hat{\rho} = \hat{\sigma}_{\epsilon_1 \epsilon_2} \Big/ \hat{\sigma}_{\epsilon_1} \hat{\sigma}_{\epsilon_2}$ from zero would imply the rejection of the hypothesis that $\sigma_{\epsilon_1 \epsilon_2} = 0$. Utilizing an approximate test, the hypothesis $\sigma_{\epsilon_1 \epsilon_2} = 0$ is rejected if

$$t = [(n - 3) \hat{\rho}^2 / (1 - \hat{\rho}^2)]^{1/2} > t_\alpha \; [(n - 3) \text{ d. f.}] ,$$

where t_α is the upper critical value of Student's t distribution with $n - 3$ degrees of freedom.

4.4 THE ESTIMATION OF DIRECT AND INDIRECT EFFECTS, AS DISCUSSED IN SECTION 3, WHEN $\sigma_{\epsilon_1 \epsilon_2} \neq 0$

Consider the structural system in (3.1.1) and (3.1.2). The corresponding reduced system, in sample form, is given by (3.1.4) and (3.1.5). Since direct and indirect effects on y_1 are underidentified, when $\sigma_{\epsilon_1 \epsilon_2} \neq 0$, we again apply the instrumental variable estimation technique.

Let the sample form of (3.1.1) be written as

$$y_{1ij} = m_1 + t_{1i} + b_{1j} + a y_{2ij} + \hat{\epsilon}_{1ij} . \tag{4.4.1}$$

Then

$$\hat{\underline{\epsilon}}_1 = \underline{y}_1 - X^+ \hat{\underline{\theta}}_1^+ \tag{4.4.2}$$

where $\hat{\underline{\epsilon}}_1 \, (qr \times 1) = (\hat{\epsilon}_{1ij})$, $\underline{y}_1 \, (qr \times 1) = (y_{1ij})$, $X^+ \, [qr \times (q+r+2)]$ is the design matrix corresponding to the model in (3.1.1), and $\hat{\underline{\theta}}_1^+ = (m_1, t_{11}, \ldots, t_{1q}, b_{11}, \ldots, b_{1q}, a)'$. Since X^+ is singular, choose a basis of X^+, say $X \, [qr \times (q + r)]$. Then $\underline{\theta}_1 \, [(q + r) \times 1]$ is the corresponding vector of non-redundant parameters, and (4.4.2) is rewritten as

$$\hat{\underline{\epsilon}}_1 = \underline{y}_1 - X\hat{\underline{\theta}} . \qquad (4.4.3)$$

Select q + r vectors, say $\underline{u}_1, \ldots, \underline{u}_\ell, \ldots, \underline{u}_{q+r}$, such that ϵ_1 is independent of u_ℓ; and let

$$U\,[\,qr \times (q+r)\,] = (\underline{u}_1, \ldots, \underline{u}_\ell, \ldots, \underline{u}_{q+r}) .$$

Then in the model

$$\epsilon_1 = \alpha_o + \alpha_1 u_1 + \ldots + \alpha_\ell u_\ell + \ldots + \alpha_{q+r} u_{q+r} + \Delta ,$$

the α's are zero. The least squares estimate of $\underline{\alpha} = (\alpha_\ell)$ is

$$\underline{a} = (U' U)^{-1} U' \underline{\epsilon}_1 .$$

Equating $\underline{a}$ to $\underline{0}$, $\hat{\underline{\theta}}$ in (4.4.3) is given by

$$\hat{\underline{\theta}} = (U' X)^{-1} U' \underline{y}_1$$

and

$$\text{var}\ \hat{\underline{\theta}} = (U' X)^{-1} (U' U)^{-1} (U' X)'^{-1} \sigma^2_{\epsilon_1} = V \sigma^2_{\epsilon_1} .$$

Let $\underline{z}$ and $V_{\underline{z}}$ denote that portion of $\hat{\underline{\theta}}$ and V corresponding to the non-redundant direct treatment effect estimates on y_1. Then the hypothesis $\tau_{11} = \ldots = \tau_{1q} = 0$ is rejected if

$$(J\underline{z})' (J V_{\underline{z}} J')^{-1} (J\underline{z})/(q-1)\hat{\sigma}^2_{\epsilon_1} > F_\alpha$$

where F_α is the upper α critical value of the F distribution with q - 1 and qr - q - r degrees of freedom, and

$$J = \begin{bmatrix} 1, & -1 & & & & \\ & 1, & -1 & & 0 & \\ 0 & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & 1, & -1 \end{bmatrix}.$$

REFERENCES

1. Anderson, R. L., and Bancroft, T. A., (1952). Statistical Theory in Research, McGraw-Hill, New York.

2. Haavelmo, T., (1943). The Statistical Implications of a System of Simultaneous Equations, Econometrika, 11, 1-12.

3. Mallios, W. S., (1961). On Linear Regression Systems, Proceedings of the Seventh Conference on the Design of Experiments in Army Research Development, and Testing, U.S. Army Research Office, Durham, N. C.

4. Scheffe, H., (1959). The Analysis of Variance, Wiley, New York.

5. Williams, E. J., (1959). Regression Analysis, Wiley, New York.

6. Zellner, A., (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, J. Amer. Stat. Assoc., 57, 348-68.

7. Zellner, A., (1963). Estimators for Seemingly Unrelated Regression Equation: Some Exact Finite Sample Results, J. Amer. Stat. Assoc., 58, 977-993.