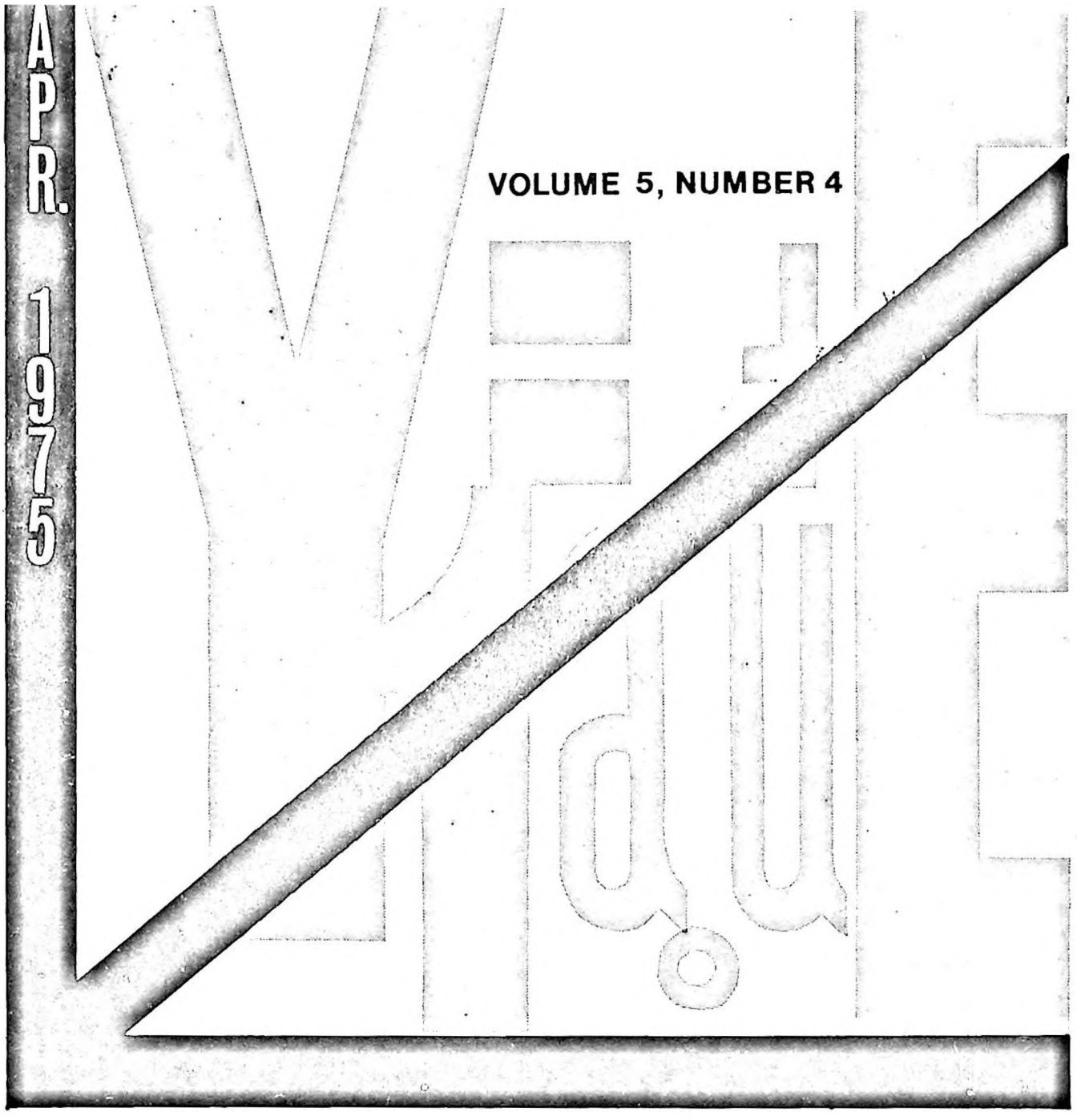
SPECIAL ISSUE
APR. 1975
VOLUME 5, NUMBER 4



MULTIPLE LINEAR REGRESSION VIEWPOINTS

A publication of the Special Interest Group
on Multiple Linear Regression
of The American Educational Research Association.

Chairman: James Bolding, Educational Foundations, University of Arkansas, Fayetteville, Arkansas 72701

Editor: Isadore Newman, Research and Design Consultant, The University of Akron, Akron, Ohio 44325

Secretary and Chairman-elect: John T. Pohlman, Testing Center, Southern Illinois University, Carbondale, Illinois 62901

Cover by: David G. Barr

Layout by: Edward Lasher

IF YOU ARE SUBMITTING A RESEARCH ATRICLE OTHER THAN NOTES OR COMMENTS, I WOULD LIKE TO SUGGEST THAT YOU USE THE FOLLOWING FORMAT, AS MUCH AS POSSIBLE:

TITLE
AUTHOR AND AFFILIATION
SINGLE-SPACED INDENTED ABSTRACT (THE REST OF THE MANUSCRIPT SHOULD BE DOUBLE-SPACED)
INTRODUCTION (PURPOSE--SHORT REVIEW OF LITERATURE, ETC.)
METHOD
RESULTS
DISCUSSION (CONCLUSION)
REFERENCES

ALL MANUSCRIPTS SHOULD BE SENT TO THE EDITOR AT THE ABOVE ADDRESS. (ALL MANUSCRIPTS SHOULD BE CAMERA-READY COPY)

TABLE OF CONTENTS

THE UNIVERSITY OF AKRON
AKRON, OHIO 44325

Department of Educational Foundations
College of Education
Phone: (216) 375-7773

February 24, 1975

ANNOUNCEMENT

Dear Sig Member:

The Multiple Linear Regression special interest group is scheduled to meet in Washington, Tuesday, April 1 from 8:45 a.m. to 10:15 a.m. The second meeting will be Wednesday, April 2 at 12:25 p.m. to 1:55 p.m.

As of yet, room assignment has not yet been given but should appear in the AERA program.

Respectfully,

Isadore Newman

Isadore Newman, Ph.D.
M.L.R.V.

IN/dt

P.S. We also expect to publish a post-convention issue of view points containing all papers presented at the special interest group meeting.

Multiple Linear Regression Viewpoints
Vol. 5, No. 4, 1975

A FURTHER DISCUSSION OF ISSUES RELATED TO CAUSAL INFERENCE

by

J. F. Byrne, Westinghouse Electric Corp., Philadelphia, Pa.
K. A. McNeil, Educational Monitoring Systems, Ann Arbor, Michigan

ABSTRACT

The role of causality in scientific investigation, and in particular as it relates to multiple regression needs to be further discussed in this journal. Statistical pragmatism (as measured by R^2) is seen as a necessary but not sufficient condition for ascertaining causality. A research strategy is described which will help the researcher define the level of causal inference to which he is entitled. Interpretations of R^2 will vary depending upon the level which the researcher has allowed.

In the November, 1974, issue of Viewpoints, Maola (1974b) criticised an earlier article by Byrne (1974). In our reading of the critique, we found that the author failed in making any of his points. We have attempted to make our comments comprehensible without the reading of the two articles. In any case, understanding the last part of this paper, starting with "Causality reconsidered," is not dependent upon the reading of the two papers.

An evaluation of Maola's criticisms

The third purpose of Maola's paper was, "to provide the reader with a brief encounter with the limitations of R^2 [Maola, 1974b, p. 2]." Maola did not indicate the limitations of R^2 as

Byrne used the concept. First, Byrne did not use a stepwise procedure as implied by Maola. Attributing variance accounted for to individual predictors is not a fruitful exercise, unless the predictors are orthogonal as in the Byrne example. Even then, discussing the total R^2 would seem to be a more fruitful procedure because the multivariate stance taken implies that it is realized that several predictors may be needed to account for variance. If the R^2 for a particular variable is low, then it is likely to represent a chance amount of variance. Figure 3 of the Byrne (1974, p. 19) paper substantiates this assertion.

In the context of hypothesis testing, the only limitation of R^2 that we can think of is that it may be inflated due to the high ratio of variables to subjects. In this case, either a correction formula should be used (see Newman, 1973), or as Byrne (1974) admonished, the solution be replicated.

(The assertion by Newman (1974) that the R^2 will be higher if only extreme cases are investigated is correct, but should not be taken as a limitation of R^2. Rather, the researcher should be aware that a different population is being investigated, i.e., extreme scorers.)

The statement, "In most cases R^2 provides an overestimate of the predictor [Maola, 1974b, p. 2]," is, on the average, incorrect. The veracity of the statement depends upon a number of conditions, as depicted in Figure 1:

A) <u>The regression procedure used to obtain R^2</u>,

1) additional R^2 due to the included variable or,

2) the decrease in R^2 as a result of deleting a variable from the full model;

B) R^2 as an estimate of either,

1) unique variance accounted for by that predictor,

2) total variance accounted for by that predictor, or

3) total variance accounted for by that model; and

C) variance being estimated,

1) in the sample,

2) in the population, or

3) in the true (causal) functional accounting of the criterion.

Note that in very few cells of Figure 1 does the "R^2 overestimate the predictor." We suspect that Maola was referring to the A1B1C2 cell and thinking of only the first predictors that enter the stepwise solution. To the best of our analytical thinking, cell A2B1C2 produces a good estimate. Information to the contrary would be appreciated.

A fuller discussion of cell A2B1C3 may assist in the understanding of Figure 1. In this cell, a variable is deleted from the full regression model, and the R^2 loss is used as the estimate of the unique criterion variance that that variable accounts for in the true causative model. Now if all the true causers are in the model, the R^2 of the model will be 1.00 and the R^2 loss will be exact. Being highly unlikely that the true causers are in the model, some of the unique variance attributed to a variable will probably be correlated with other variables

B. R^2 Estimating	C. Variance being estimated	A. Regression Procedure	
		A1. Forward Stepwise	A2. Hypothesis Testing
B1. Unique criterion variance accounted for by the predictor	C1. Sample	All overestimate except the last which is correct	Exact
	C2. Population	Most overestimate, some may underestimate	Good Estimate
	C3. True causal model	Overestimate, unless suppressors operate	Overestimate, unless suppressors operate
B2. Total criterion variance accounted for by the predictor	C1. Sample	May be overestimated for first few variables, underestimated for last ones	Underestimates
	C2. Population	Underestimates	Underestimates
	C3. True causal model	Underestimates	Underestimates
B3. Total criterion variance accounted for by the model (The R^2 of the full model is referred to for this section)	C1. Sample	Exact	Exact
	C2. Population	Overestimates (could underestimate in unusual cases)	Overestimates (could underestimate in unusual cases)
	C3. True causal model	Underestimates	Underestimates

Figure 1. Adequacy of R^2 as an estimator of variance.

not already in the model. Hence the evaluation of "overestimate." Now if some of those other causers operate as suppressor variables, the R^2 loss will underestimate the unique variance accounted for. The bottom row of Figure 1 (B3) may also need some more explanation. Here the concern is how well the sample R^2 of the model estimates the R^2 in the sample, population, and the true causal model. If the stepwise procedure is allowed to take all its steps, the two procedures are equally effective in their estimates. To the extent that most stepwise applications are less well planned (more variables are included with less justification) than are the hypothesis testing applications, those stepwise applications will overestimate the population R^2 more so than will the hypothesis testing applications (unless suppressors are operating which the stepwise doesn't detect). Because of sample vagaries, either procedure could produce a sample R^2 which is lower than the population R^2 (B3C2).

The second objective of Maola (1974b) was to "provide the reader with the value of Byrne (1974) for presenting a rationale for reporting R^2 [p. 2]." Maola does succeed in telling us twice that Byrne did a good job.

The first objective is the one we are most concerned about--"an understanding of the semantic difference of prediction and causality [Maola, 1974b, p. 2]." Maola asserts that Byrne "deliberately or accidently misinterpreted causality as synonymous with prediction [p. 2]," although no reference is made to the source of that assertion. "The paper by Byrne (1974) <u>seems to imply</u> that if

$R^2 = 1.00$, then since 100 percent of the predictor variance is accounted for there is causality [Maola, 1974b, p. 1, emphasis ours]." We propose that the above conclusion was not ascribed to by Byrne, and was arrived at independently by Maola.

That prediction and causality were not equated should be clear from the following quote: "An assertion that causality is understood...can only be defended with a high R^2, low probability of chance occurrence, and replication [Byrne, 1974, p. 12]." The implication is that a high R^2 is only one of three necessary but not sufficient conditions for causality. "An R^2 of 1.00 means that 100% of the change in the phenomenon has been numerically connected to changes that take place among the variables and experimental manipulations of the scientist [Byrne 1974, p. 12]." This statement further leads us to interpret Byrne as saying R^2 is basically a numerical artifact.

Causality reconsidered

We propose that all researchers are attempting to find "the real cause." It is not the purpose of this paper to indulge in a deep philosophical discussion of causality, but simply to remind researchers that this is what they are searching for and that MLR can help them. (We hesitate to point out that Maola, in the pages preceeding the Byrne article, himself demonstrated "a method of using multiple regression for determining the effect of independent differences of providing treatment [Maola, 1974a, p. 10]."

Perhaps the consideration of causality at three levels will be

of some assistance--"the first being what is really causing the phenomenon, the second being what we think is causing the phenomenon, and the third being what our measurements and analysis of variance state is causing the phenomenon [Byrne, 1974, p. 11]."

It is ironic that the first level is reality, but most philosophers feel that we may never be sure that we have "the cause." Therefore, we either theorize about the cause (level two), or we go out into the real world and measure it (level three). Theory cannot exist separate from statistical verification, but statistical verification must be viewed as a necessary, but not sufficient condition for verification at level two theory. Some researchers would take a more empirical approach and eschew theory altogether. We ourselves probably lean in that direction. If we find a particular empirical fact (high R^2) occurring time and time again (replication) we would believe in it and readily use that information, even before we were able to fit it into a consistent theory. Have we not found the "cause" because we are not able to explain "why," and put our findings into a coherent context? If we are able to control the magnitude of the criterion, then we in some semantic way have found a cause. (See Ennis, (1973) for a very readable introduction to these notions.)

Perhaps using the word "cause" at the empirical level is a mistake. What we are investigating are possible causers--they will remain tenable until some competing explainer has eliminated them. The trouble is that there are an infinite number of completing explainers in the real world. Too often we view our research in

too short a time span and too narrow a spectrum. The research strategy described in Figure 2 may assist in placing the notion of causation in a better perspective. Note should be made that stepwise regression has a role in Stage 1, but that hypothesis testing regression is needed for Stages 2-5.

	Stage	Emphasis on Low Probability	Emphasis on High R^2
Static Variables	1. Data Snooping (Hypothesis Gen.)	No	Yes
	2. Hypothesis Testing	Yes	No
	3. Replication	No	Yes
Dynamic Variables	4. Manipulation (Dynamic Hypothesis Testing)	Yes	Yes
	5. Replication	Yes	Yes

Figure 2. Stages of a proposed research strategy and their relative emphasis on probability and R^2. (Adopted from McNeil, Kelly, and McNeil, 1975.)

Notice in Figure 2 that the concept of R^2 is not considered to be as important as is the concept of probability at certain stages. Research at Stage 1 would provide very little evidence for causal interpretations. But as the line of research proceeds successfully through the stages, more confidence is gained regarding causal interpretations. Notice that the last two stages are concerned with predictor variables which have been manipulated, whereas the first three stages involve data which has not been

manipulated. A researcher may gain some personal degree of belief in causers in the first three stages, but not until successful research has been performed at the last two stages should any public belief be expressed.

The ideal strategy would be to perform successful research starting at Stage 1, and to work up to Stage 5 one stage at a time. An unsuccessful research finding at any point would force one to reenter the process at Stage 1. In reality, most researchers short circuit this process by starting at a stage other than Stage 1.

Some consternating thoughts on which we would like some feedback

1. Variables which are not predictive in a static situation may be predictive when manipulated.
2. Variables which are predictive in a static situation may not be predictive when manipulated.
3. Variables which are manipulated may affect not only the criterion, but other predictor variables, such that the criterion is higher or lower than would be predicted from the static situation.
4. Some kinds of subjects may not change on the criterion as a result of manipulating a predictor variable (subject by causer interaction?).
5. Criterion scores may increase as a result of manipulation, but the construct hasn't increased (enter concepts of practice effect, experimenter bias).

6. Criterion scores may increase as a result of manipulation--but not enough to satisfy the researcher. And since the manipulated variable has been manipulated to its limit, other causers must be found to obtain the desired degree on the criterion.

REFERENCES

Byrne, J.F. The use of regression equations to demonstrate causality. Multiple Linear Regression Viewpoints, 1974, 5(1), 11-22.

Ennis, R.H. On Causality. Educational Researcher, 1973, 26, 4-11.

Maola, J.F. The multiple regression approach for analyzing differential treatment effects--the reversed gestalt model. Multiple Linear Regression Viewpoints, 1974, 5(1), 7-10.

Maola, J.F. Causality and prediction, similar but not synonymous. Multiple Linear Regression Viewpoints, 1974, 5(2), 1-2.

McNeil, K.A., Kelly, F.J. & McNeil, J.T. Testing Research Hypotheses Using Multiple Regression. Carbondale, Illinois: Southern Illinois University Press, 1975.

Newman, I. Variations between shrinkage estimation formulas and the appropriateness of their interpretation. Multiple Linear Regression Viewpoints, 1973, 4(2), 45-48.

Newman, I. Dinner with Earl Jennings. Multiple Linear Regression Viewpoints, 1974, 5(1), 40-43.

Multiple Linear Regression Viewpoints
Vol. 5, No. 4, 1975

A COMMENT ON POHLMANN'S ALGORITHM FOR SUBSET SELECTION IN MULTIPLE REGRESSION ANALYSIS

by

MaryAnn Ross and George P. McCabe
Purdue University

Abstract

Pohlmann's algorithm for incorporating cost criteria into the variable selection problem in multiple regression is examined. It is pointed out that this algorithm has the property that the choice of the optimal subset can be artificially changed by the addition of another variable. An example is included to illustrate this property.

Pohlmann (4) has presented an algorithm which incorporates cost information into the selection of a subset of variables in multiple regression analysis. Other approaches to this problem are given in (2) and (3). We suggest that Pohlmann's proposed method of reducing the costs and losses due to lack of predictability to a common scale of measurement may not be the most desirable.

The loss function to be considered is

$$L_I = k_1(c_1 \times cost_I) + k_2(c_2 \times (1-R_I^2)).$$

The weighting coefficients, c_1 and c_2, suggested are

$$c_1 = \left(\sum_{I=1}^{J} cost_I\right)^{-1}$$

and

$$c_2 = \left(\sum_{I=1}^{J} (1-R_I^2)\right)^{-1}$$

where

J = the total number of subsets under consideration.

Using the above definitions, the revised costs and losses due to lack of predictability and in turn the loss function for any particular subset are dependent on the total number of subsets under consideration. In other words, if an additional variable is added to or deleted from the original set of predictor variables, the loss function for the original subsets can be changed. This may, in fact, lead to the choice of a different subset from the original collection as the optimal subset. Ideally, the relative loss corresponding to any given subset, as compared to the loss for another given subset should not be dependent on the cost and lack of predictability of extraneous variables.

An Example

The following example uses data presented by Hald (1) with arbitrarily assigned cost values. In Table 1 the data are analyzed as suggested by Pohlmann. The chosen subset in this case is that containing predictor variates 3 and 4. Now let us assume that we can measure the dependent variable precisely for $100 per observation, but further, that $100 is the maximum amount per observation that we are willing to spend. Table 2 contains the analysis for this problem. It can be seen that the chosen subset in this case is that containing variables 1, 3 and 4. Even though the added variable is not a feasible choice because of its high cost, the preferred subset according to this method has changed to another subset of the original set.

An Alternative Method of Standardizing Cost and Validity Data

In order to overcome the above problem it is suggested that the weighting coefficients in the loss function might be defined as:

$$c_1 = (c_{max})^{-1}$$

where c_{max} is the maximum cost which the investigator is prepared to spend, and

$$c_2 = 1.0.$$

In this way, the revised values of both cost and loss due to lack of predictability are between 0.0 and 1.0 and the combined loss value for any subset will not depend on the number of subsets under consideration.

REFERENCES

1. Hald, A., Statistical Theory with Engineering Applications. New York: Wiley, 1952.

2. Lindley, D. V., The choice of variables in multiple regression. Journal of the Royal Statistical Society, Series B, 30, 1968, 31-35.

3. McCabe, G. P. and Ross, M. A., A stepwise algorithm for selecting regression variables using cost criteria. To appear in "Proceedings, Computer Science and Statistics: Eighth Annual Symposium on the Interface", 1975.

4. Pohlmann, J. T., Incorporating cost information into the selection of variables in multiple regression analysis. Viewpoints, 1973, 4, 18-26.

TABLE 1. ANALYSIS OF HALD DATA.

	Cost A	Revised Cost B	$1-R^2$ C	Revised Loss D	B+D E	B+3D F
1	5	.052	.47	.118	.170	.406
2	5	.052	.33	.083	.135	.301
3	1	.010	.72	.181	.191	.553
4	1	.010	.32	.080	.090	.250
1,2	10	.104	.02	.005	.109	.119
1,3	6	.062	.45	.113	.175	.401
1,4	6	.062	.03	.008	.070	.086
2,3	6	.062	.15	.038	.100	.176
2,4	6	.062	.32	.080	.142	.302
3,4	2	.021	.06	.015	.036	.066*
1,2,3	11	.115	.02	.005	.120	.130
1,2,4	11	.115	.02	.005	.120	.130
1,3,4	7	.073	.02	.005	.078	.088
2,3,4	7	.073	.03	.008	.081	.097
1,2,3,4	12	.125	.02	.005	.130	.140
NULL	0	.0	1.00	.251	.251	.753
TOTAL	96	.998	3.98	1.000		

TABLE 2. ANALYSIS OF HALD DATA WITH ADDITIONAL VARIABLE.

	Cost A	Revised Cost B	$1-R^2$ C	Revised Loss D	B+D E	B+3D F
5	100	.510	.00	.000	.510	.510
1	5	.026	.47	.118	.144	.380
2	5	.026	.33	.083	.109	.275
3	1	.005	.72	.181	.186	.548
4	1	.005	.32	.080	.085	.245
1,2	10	.051	.02	.005	.056	.066
1,3	6	.031	.45	.113	.144	.370
1,4	6	.031	.03	.008	.039	.055
2,3	6	.031	.15	.038	.069	.145
2,4	6	.031	.32	.080	.111	.271
3,4	2	.010	.06	.015	.025	.055
1,2,3	11	.056	.02	.005	.061	.071
1,2,4	11	.056	.02	.005	.061	.071
1,3,4	7	.036	.02	.005	.041	.051*
2,3,4	7	.036	.03	.008	.044	.060
1,2,3,4	12	.061	.02	.005	.066	.076
NULL	0	.000	1.00	.251	.251	.753
TOTAL	196	1.002	3.98	1.000		

Multiple Linear Regression Viewpoints
Vol. 5, No. 4, 1975

THE USE OF REGRESSION ANALYSIS IN RODENT MATERNAL BEHAVIOR: MULTIPLE REGRESSION'S ANSWER TO THE AGE CONFOUND

by

Joel Burkholder, David Kapusinski & Robert Deitchman
Psychology Department
The University of Akron

It has been a popular convention in developmental psychology to consider development as a continuous process in which the organism is observed to change across a given interval of time. While this conceptual scheme has had much favor (Wholwill, 1970), the researcher investigating developmental change often fails to consider the importance of this convention by applying ineffective statistical procedures. Some examples of this kind are also found in the animal literature concerned with maternal behavior and early experience (Ader and Deitchman, 1970; Ader and Deitchman, 1972; Grota and Ader, 1969; Priestnall, 1973; Wehmer, 1965). In most instances the researcher either; a) sums across all repeat observations, or more importantly, b) arbitrarily collapses sets of repeated observations based on an a-priori decision. In turn, the data is then cast into a conventional analysis of variance for repeated measures design which yeilds its particular main effects, interactions and error terms. Likewise, in terms of analysis of variance, each main effect, whether a treatment or repeat, and interaction is considered unique and uncorrelated with the other

main effects and interactions. However, if we conceptually assume a continuous developmental process that interacts with the particular manipulation in question then this procedure ignores the possibility that a correlation exists between the treatments and repeated observations that cannot be considered unique. Simply stated, variance associated with a repeat or age Main effect or age by treatment interaction is conceptually meaningless. Secondly, if we apply a categorical procedure (i.e., a conventional ANOVA design) on continuous data we are also violating our assumption of continuity of development. This would also be the case if we applied a trend analysis procedure.

In applying a conventional ANOVA design to developmental data we are asking a research question which may not correspond with the statistical question being asked. For example, assume the experimenter wants to test the following research question:

> Research Hypothesis 1A: A significant overall difference between treatment "A" mothers and treatment "B" mothers exists, taking into consideration repeated measurements.

If we apply a conventional ANOVA for repeated measures design the question becomes:

> ANOVA Hypothesis 1B: A significant difference exists between treatment "A" mothers and treatment "B" mothers over and above repeats and treatment by repeats interaction.

It is quite apparent that the experimenter has asked for one answer but received a completely different one.

The researcher may also hypothesize that the two treatments will progressively change the mothers' maternal response to their offspring across time. Let us suppose that mothers "A" will sustain a high level of mother-offspring relationship while mothers "B" become progressively altered yielding a decline in the mother-offspring relationship. The experimenter may ask the following question:

> Research Hypothesis 2A: That a progressive difference in the mother-offspring relationship exists between treatment groups across repeated measures.

What the experimenter really tests with a conventional ANOVA design is:

> ANOVA Hypothesis 2B: That a difference exists common to the treatments by repeated measures interaction over and above variance associated with the treatment and repeats main effects.

This second ANOVA hypothesis indicates several logical errors that are in violation of the convention that development is a continuous process. Experimenters often falsely assume that the variance that composed the repeat or treatment main effects will also be involved in the interaction term common to both; this is not the case. In an analysis of variance design each main effect and interaction is orthoginal to each other. If this be the case, then our convention of developmental continuity is being violated. Variance uniquely associated with a repeats main effect and developmental change in this case would be confounded.

However, if we were to apply a multiple regression approach (McNeil, Kelly and McNeil, 1973), we could avoid these problems by analyzing directly the functional differences between our two treatments while taking repeated measures into account. In terms of the two research hypotheses mentioned above, a regression hypothesis could be formed that would more closely reflect the researcher's original intention:

> Regression Hypothesis 1C: That a significant difference exists between treatment "A" and "B" mothers across continous repeated observations
>
> and,
>
> Regression Hypothesis 2C: That a significant difference exists between the two treatment conditions across continuous repeat observations (i.e., slopes) over and above what is accounted for by initial group differences (i.e., intercepts).

The following experiment (Kapusinski, 1972) was analyzed using a conventional analysis of variance for repeated measures (BMDO8V) and Multiple Linear Regression (LINEAR) for purposes of comparing the two techniques.

Method

Thirty female C31 mice were randomly assigned to two community cages (15 x 10 x 6 inch) until the experimental treatment was initiated. Males were placed into these cages to mate, and then were removed after one week. One group of females was randomly assigned to the experimental condition and subjected to handling

until parturition. "Handling" consisted of taking the females from their home cages and individually placing them into small closed containers for 15 minutes in the morning and evening. Control subjects were not handled. About 2½ weeks into the handling procedure, the visibly pregnant females from both treatment and control groups were moved to individual cages where bedding and nesting material was provided.

At parturition, the offspring litters were culled (to 3 males and 3 females) and fostered to mothers of the same prenatal condition or cross fostered to mothers of the opposite condition. However, for purposes of comparison, only the fostered experimental and control groups are analyzed here. Maternal behavior categories devised by Deitchman (1968) were made on an Esterline Angus 20 pen recorder twice daily (at 7-10 A.M. and 7-10 P.M.) for the first 21 days postpartum. To simplify further, only two behavioral categories (i.e., nesting behavior and neonatal grooming) will be reported.

Description of Variables

$Y_{1,2}$ = the criterion variables of nesting behavior and neonatal grooming.

X_1 = 1 if experimental group of handled mothers fostered with pups from other handled mothers; zero otherwise.

X_2 = 1 if control group of nonhandled mothers fostered with pups from other nonhandled mothers; zero otherwise. .

X_3 = continucus variable of days ranging from 1 to 21.

$X_4 = (X_1 * X_3)$ days for X_1

$X_5 = (X_2 * X_3)$ days for X_2

$X_6 = (X_1 * X_2)$

Where:

$Y_{1,2}$ = criterion score described

a_0-a_6 = partial regression weights

U = Unit Vector

X_1 - X_6 = variables previously defined

E = $(Y - \hat{Y})$ error

Research Hypotheses

Research Hypothesis 1: That a significant difference exists between experimental and control mothers across days.

Full Model $Y_{1,2} = a_0U + a_1X_1 + a_2X_4 + a_3X_2 + a_4X_5 + E$

Restricted Model 1. $Y_{1,2} = a_0U + E$

Research Hypothesis 2: That a significant difference exists between the slopes of the experimental and control groups over and above what is accounted for by the intercepts.

Restriction $(a_1 = a_3)$

Restricted Model 2: $Y_{1,2} = a_0U + a_1X_1 + a_3X_2 + E$

Research Hypothesis 3: That a significant difference exists between the intercepts of the experimental and control groups over and above what is accounted for by slope differences.

Restriction $(a_2 = a_4)$

Restricted Model 3: $Y_{1,2} = a_0 U + a_2 X_4 + a_4 X_5 + E$

Note that these last two hypotheses could use a similiar set of restricted models for testing the slopes and intercepts respectively:

Restriction $(a_1 = a_3 = a_5)$ where $X_1 + X_2 = X_3$

Restricted Model 2a: $Y_{1,2} = a_0 U + a_1 X_1 + a_3 X_2 + a_5 X_3 + E$

and,

Restriction $(a_2 = a_4 = a_5)$

Restricted Model 3a: $Y_{1,2} = a_0 U + a_2 X_3 + a_4 X_5 + a_5 X_3 + E$

Results and Discussion

A conventional ANOVA for repeated measures suggested no differences between experimental and control mothers for the treatment or repeat main effects on nesting behavior ($F= .26$, $df= 1,8$ $p=NS$; $F= 1.04$, $df=20,160$ $p>.1$, respectively). However, a significant treatments by repeats interaction was indicated ($F= 2.06$ $df= 20, 160$ $p<.005$). The experimenter, at this point, is faced with a problem; both main effects indicated no differences, yet the repeats by treatment interaction suggests differences of one kind or another are present across time over and above all other factors. For the developmentalist attempting to investigate developmental change, both overall mean differences or interactions unique to

to these overall mean differences may actually be of little real importance if the particular behavior under study is epigentic in nature.

A multiple regression analysis of the overall differences between the two conditions on nesting behavior, taking repeats into consideration, suggested no differences (Model 1; $F= 1.66$ $df= 4,124$ $p<.25>.1$). Further, a comparison of the two respective slopes (over and above initial group identity) indicated that the mothers nesting behavior progressively changed in relation to their offspring (Model 2; $F= 3.26$ $df= 2,124$ $p<.05>.025$). Intercepts were found to have nominal differences (Model 3; $F= 2.35$ $df= 2,124$ $p<.25>.1$). It becomes evident that the regression approach yeilds information which is more readily interpretable than the conventional analysis of variance. By using a regression approach, variance correlated with observed ongoing development need not be partialled out of the analysis; instead, the developmentalist is able to identify the differences between the two treatments progressively changed throughout the observational interval. One may also recognize that variance associated with a repeats main effect in a conventional ANOVA design is artifactual in nature, because the variance associated with individual differences, observational method used, and the process of development are confounded. In studies concerned with developmental differences or change, it cannot be assumed

that variance associated with repeated observations can be meaningfully separated from observed changes in behavior.

In the second behavioral category i.e. neonatal grooming, the treatment main effect suggested a very nominal trending toward significance ($F= 2.09$ $df=1,8$ $p<.25>.1$) with no differences indicated for the repeats main effect ($F=.43$ $df=20,160$ $p=NS$) or treatment by repeats interaction ($F= .67$ $df=20,160$ $p=NS$). In this case regression analysis did not suggest an overall difference between the two treatments with repeated observations considered (Model 1; $F= 1.11$ $df=4,124$ $p<.25$). A test of the slopes also yeilded no differences (Model 2; $F= .39$ $df=2,124$ $p=NS$) while the test of intercepts did (Model 3; $F=3.54$ $df=1,124$ $p<.1>.05$), suggesting that the initial differences between the two conditions were accounting for most of the variance. For neonatal grooming, the conventional analysis of variance treatment main effect only partially estimated the differences obtained by the regression test of the intercepts.

The impetus of this paper has been directed mainly toward a logical, and meaningful, way to analyze continuous developmental change. The second of the two examples given indicated that both ANOVA and regression accounted for some of the information in a similiar fashion. But the main advantage of regression over ANOVA was in the first example where the regression indicated a functional

change, a finding not collaborated by comparing mean differences using the conventional analysis of variance approach. One may suggest a trend analysis be the logical choice instead of a conventional ANOVA design. This would be an improvement, however, trend analysis utilizes categorical data making it less than optimal if we are to abide by our convention of developmental continuity. Multiple regression, in this case, becomes our most effective tool for dealing with experimental situations having continous developmental data.

A twofold recommendation to the developmental psychologist using repeated measurements across a given interval of time would follow; 1) that the investigator be aware of how to best reflect the research question in the statistical analysis, and 2) to entertain the use of mulitple regression as a principle method of preserving the continuity of developmental changes.

References

Ader, R., and Deitchman, R., Prenatal Maternal X-irradiation: Maternal and Offspring Effects, Journal of Comparative and Physiological Psychology, 1972, 78, 202-207.

Ader, R., and Deitchman, R., Effects of Prenatal Maternal Handling on the Maturation of Rhythmic Processes, Journal of Comparative and Physiological Psychology, 1970, 71, 492-496.

Deitchman, R., Observational Analysis of Maternal Behavior in Irradiated and NonIrradiated Wistar Rats and its Relation to Later Behavior of the Pups, Unpublished Doctoral Dissertation, University of Tennessee, 1968.

Grota, L.J., and Ader, R., Continuous Recording of Maternal Behavior in Rattus Norvegicus, Animal Behaviour, 1969, 17, 722-729

Kapusinski, D., An Observational Analysis of Maternal Behavior in Handled and NonHandled Mice and its Relation to Later Behavior of the Pups, Unpublished Masters Thesis, The University of Akron, 1972.

McNeil, K., Kelly, F.J., and McNeil, J., Testing Research Hypotheses Using Multiple Linear Regression, In Press, Copyright 1973.

Priestnall, R., Effects of Handling on Maternal Behavior in the Mouse (Mus Musculus): An Observational Study, Animal Behaviour, 1973, 21, 383-386.

Wehmer, F., Effects of Prior Experience with Objects on Maternal Behaviors in the Rat, Journal of Comparative and Physiological Psychology, 1965, 60, 294-296.

Wholwill, J.F., The Age Variable in Psychological Research, Psychological Review, 1970, 77, 49-64.

Multiple Linear Regression Viewpoints
Vol. 5, No. 4, 1975

ANOTHER VIEWPOINT OF ISSUES IN MULTIPLE LINEAR REGRESSION

by

Keith A. McNeil, Educational Monitoring Systems
Judy T. McNeil, High/Scope Educational Research Foundation
Francis J. Kelly, Southern Illinois University

ABSTRACT

In a recent issue of Viewpoints, Houston and Ohlson (1974) presented five issues in the teaching of MLR which they felt should be covered in an MLR course, but were not covered in the texts by Ward and Jennings (1973) and Kelly, Beggs, McNeil, Eichelberger, and Lyon (1969). (A major revision of the latter text, McNeil, Kelly, and McNeil [in press] does not cover these topics either.)

We disagree with Houston and Ohlson regarding basic courses in MLR on the basis that they (as well as many other statisticians) fail to make distinctions between behavioral science researchers and statistics majors. Indeed, past teaching has led to inadequate understanding of hypothesis testing and to the simple minded research approach which Snow (1974) so well criticized. To be sure, statistics majors need matrix algebra, orthogonal coding, etc., but our experience with non-statistics majors is that overly mathematized basic design courses lead to blind learning of algorithms with little comprehension of how MLR and other statistical procedures can be used in research decision making. We would be in favor of an AERA pre-session wherein the issues raised by Houston and Ohlson (and criticized in detail below) are more deeply discussed--from an hypothesis-testing point of view.

Issue 1--MLR as a generalized procedure

Since many authors have shown how ANOVA procedures can be tested with MLR, it is not clear how "ANOVA procedures are more sensitive and efficient" (Houston and Ohlson, 1974, p.42). If one is producing the answers by hand, ANOVA may be easier. Certain canned computer programs can be used more efficiently, but that is a computer interface concern. Furthermore, most canned ANOVA programs do not answer some/all of the questions the behavioral scientist has in mind. Even we have been known to run canned ANOVA programs. But we knew that the programs existed, that they would answer the specific questions we had, that it would take us less time to use these programs than to use MLR, and most importantly we already knew how to run those programs. You see, we had the (mis)fortune of learning computer programming and learning about a good statistical package (Veldman, 1967) before we learned multiple linear regression.

We would not necessarily recommend that a behavioral scientist who already knows MLR and how to run an MLR computer program should also learn to run an ANOVA program. He can answer all his "ANOVA questions' with the one tool, MLR. But we disagree little with Houston and Ohlson (1974, p.42) on the following point: "We recommend that the behavioral scientist develop expertise in both approaches and, armed with knowledge of their interrelationships, allow the research questions to dictate which approach should be undertaken." Unfortunately, most ANOVA procedures are testing non-directional questions which do not produce definitive conclusions

(McNeil et al., in press). Perhaps the only justification for learning ANOVA is to be able to read some journals.

Research Hypothesis I: A course in MLR coupled with a research design course will generate better researchers than will a course in ANOVA coupled with a research design course.

Issue 2--Use of matrix algebra in MLR

We can agree that most behavioral scientists should develop skills in other multivariate procedures, but the question is to what depth. Is it not possible that the conceptual value of canonical analysis can be appreciated without knowing how to invert a matrix? Understanding the questions that a statistical technique can answer, and being able to properly interpret computer output seem to circumvent the necessity of knowing matrix algebra.

One of the authors (K. McNeil) taught MLR for six years to some 22 classes without using matrix algebra, and he felt that the students had a good grasp of the notions of MLR. Many of these students went on to a multivariate course (where matrix algebra was used), but we suspect that a heavy dose of matrix algebra in the MLR course would have turned many of them off. We fail to see how, to the non-mathematical expert, matrix algebra is elegant, and how scalar algebra is involved and messy.

Research Hypothesis II: The matrix approach to MLR will be rated as more "involved and messy" than will the vector approach by more than 75% of the members of SIG-MLR.

Issue 3--Redundant models in MLR

As believers in the Law of Parsimony, we can't quarrel too much with the concern of keeping redundant vectors in regression models. But linearly independent vectors don't bother us too much because (1) we use an iterative computer program, (2) linearly dependent vectors become useful in making certain restrictions--and in determining the number of linearly independent vectors in certain complex models, and (3) we abhor rules of thumb. Once the concept of linear dependency is grasped, one should be able to determine the number of linearly independent pieces of information in the model (the degree of parsimony). The availability of all of the group membership vectors allows for a restriction to involve any one of them. For instance, given these two non-orthogonal research questions: $\mu_A > \frac{\mu_B + \mu_C}{2}$; $\mu_A > \mu_B$ (using group C to help estimate within variability) the approach proposed by Houston and Ohlson (1974) would leave the behavioral scientist in limbo, we suspect.

We don't know what alternative procedure Houston and Ohlson (1974) are suggesting. If they are suggesting that linearly dependent vectors be omitted, we would not be too concerned. But we would strongly argue against using the (1,0,-1) coding scheme (Lewis & Mouw, 1973) or an effect scheme (see Kerlinger and Pedhazur, 1973, for details). There is absolutely no conceptual value in the (1,0,-1) scheme, and the effects scheme requires all orthogonal effects to be coded (even though those effects may not be of interest to the behavioral scientist). At any rate, the

notion of linearly independent vectors is used implicitly in ascertaining the number of effect vectors one must use.
Research Hypothesis III: The issue of redundant models in MLR is not an overpowering issue with SIG-MLR members.

Issue 4--Orthogonal coding

The defense of orthogonal coding presented by Houston and Ohlson is that it is easier when computing without a computer. True, but most behavioral scientists have access to a computer (particularly if they are going to be interested in multivariate questions). We don't know of any researchers who calculate MLR problems without a computer. Orthogonal coding has several bad aspects. Additional questions are tested by the behavioral scientist when in fact he may not have wanted to test those questions. (He may have been interested in only a few of the orthogonal questions.) The rule of thumb that the one-way F be significant before post hoc comparisons can be made has never made any sense to some behavioral scientists who have specific hypothesis in mind. By the way, does any post hoc procedure allow for a directional interpretation, or do the statistics simply say that there is a difference between these two groups, or combination of groups? An anecdote for thought: an ANOVA trained person recently used the McNeil et al. (in press) text in her class. Since the section on orthogonal coding is nonexistant, she decided to develop the material for her class. Unfortunately, according to her, she developed incorrect models for the class. The point is that the

(1,0) vectors can be used to test any of the post hoc comparisons (Williams, 1974; Kerlinger & Pedhazur, 1973), and therefore special vectors do not need to be used. In fact, even well-trained statisticians can get confused by such special vectors.

Research Hypothesis IV: The research hypothesis being tested is more clearly understood by the typical behavioral scientist when (1,0) vectors are used than when orthogonal coding vectors are used.

Issue 5--Data analysis and MLR

The writers concur with Houston and Ohlson (1974), "that MLR will be relied on by the behavioral scientist more as a data analytic tool, often suggesting ideas for subsequent research, rather than as a procedure for analyzing a classical design." The kind of research strategy depicted in McNeil et al., (in press) and discussed in the Byrne and McNeil (1975) article in this same issue certainly calls for such activity.

We repeat that we do feel that MLR will be used by more and more behavioral scientists and that the technique will enable them to more appropriately answer their specific research questions. We are therefore concerned that behavioral scientists develop a conceptual understanding of MLR as it is applied to hypothesis testing. We contend that the first four issues raised by Houston and Ohlson are more concerned with mathematics and statistics per se than with the application of MLR to a research strategy. We would rather ignore these issues in an introductory course and concentrate on understandings that will facilitate the answering

of research questions.

To moderate and clarify the preceeding, we believe one should be aware of two audiences (1) applied statisticians and (2) educational researchers. We believe MLR Viewpoints has admirably served the first audience, but to a large degree has slighted the second audience. We may have passed the need to show solely how MLR answers traditional ANOVA questions (and satisfy statisticians) and need now to present MLR models which cope with the complex problems associated with molar problems presented by Snow (1974).

REFERENCES

Byrne, J.F. and McNeil, K.A. A further discussion of issues related to causal inference. Multiple Linear Regression Viewpoints, 1975, 5(4).

Houston, S.R. and Ohlson, E.L. Issues in teaching multiple linear regression for behavioral research. Multiple Linear Regression Viewpoints, 1974, 5(2), 41-44.

Kelly, F.J., Beggs, D.L., McNeil, K.A., Eichelberger, T. and Lyon, J. Research Design in the Behavioral Sciences: Multiple Regression Approach. Carbondale: Southern Illinois University Press, 1969.

Kerlinger, F.N. and Pedhazur, E.J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart, and Winston, 1973.

Lewis, E.L. and Mouw, J.T. The use of contrast coding to simplify anova and ancova procedures in multiple linear regression. Multiple Linear Regression Viewpoints, 1973, 4(2), 27-44.

McNeil, K.A., Kelly, F.J. and McNeil, J.T. Testing Research Hypotheses Using Multiple Linear Regression. Carbondale: Southern Illinois University Press, in press.

Snow, R.E. Representative and quasi-representative designs for research on teaching. Review of Educational Research, 1974, 44, 265-291.

Veldman, D.J. FORTRAN Programming for the Behavioral Sciences. New York: Holt, Rinehart, and Winston, 1967.

Ward, J.H. and Jennings, E.E. Introduction to Linear Models. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Williams, J.D. Regression Analysis in Educational Research. New York: MSS Information Corporation, 1974.

Multiple Linear Regression Viewpoints
Vol. 5, No. 4, 1975

A DISCUSSION OF COMPONENT ANALYSIS: ITS INTENDED PURPOSE, STRENGTHS, AND WEAKNESSES

Isadore Newman
The University of Akron

and

Carole Newman
on Leave from Akron Public Schools

Presented at the 1975 American Educational Research Association;
Multiple Linear Regression Special Interest Group

* We would like to acknowledge and publicly thank Dr. Keith McNeil for his critical reading of this manuscript.

The purpose of this paper is to examine a statistical procedure which is referred to as commonality analysis, elements analysis, or component analysis. This procedure was developed in the late 60's to aid researchers in explaining and interpreting the results of statistical analysis in which the predictor variables are not independent (nonorthogonal).

Frequently investigators are interested in having their research applied to practical (natural) situations. For this reason, one has to delineate the relationships between the predictor variables and the criterion. If the variables are interrelated, such as intelligence, socio-economic status, race, etc., one has difficulty in accurately estimating the relative importance of each predictor variable to the criterion. Darlington (1968), Mood (1969, 1971), McNeil, Kelly, McNeil (1975) Kerlinger and Padhazur (1973) clearly delineate the various aspects of this problem.

Ex post facto research has greatly suffered from the limitation that the results are subject to various interpretations. This is true because one cannot state the relative importance of the predictor variables to the criterion.

Component analysis (Cp) is a procedure which divides the proportion of variance accounted for into common and unique variance. The unique variance (Uq) is the proportion of variance attributed to a particular variable when entered last in the regression equation. It is what Botenberg and Ward (1963) and McNeil, et al. (1975)

call the proportion of variance attributed to a particular variable, above and beyond the variance accounted for by the other independent variables in the equation (analysis of covariance, semipartial correlation). Therefore, the unique variance accounted for is represented by a full model which contains all the independent variables tested against a restricted model in which all the predictor variables are presented, except for the one(s) for which the unique variance is to be estimated. Unique variance is symbolically represented as:

$$Uq(5) = R^2_{(f)}\ y_1 \cdot 123456 - R^2_{(r)}\ y_1 \cdot 12346$$

where: $Uq(5)$ = the unique variance of y_1 that can be attributed specifically to variable 5, independent of what is being measured by variables 1, 2, 3, 4, and 6.

$R^2_{(f)}$ = R^2 of the full model

$R^2_{(r)}$ = R^2 of the restricted model

Common variance (Cv) may be conceptually thought of as the degree the overlap of correlated variables are predictive of the criterion. It must be independent of unique and other common variance.

In an example with three predictor variables, there are three sets of unique variance $\left[Uq_{(1)}, Uq_{(2)}, Uq_{(3)}\right]$, three sets of second order commonality $\left[Cv_{(1,2)}, Cv_{(1,3)}, Cv_{(2,3)}\right]$, and one third order commonality variance $\left[Cv_{(1,2,3)}\right]$

The number of independent components in a component analysis procedure can be determined by the equation:

$$2^N - 1$$

where: N = number of predictor variables

Therefore, if one had four predictor variables, the number of components would equal:

$$2^4-1 = 15$$

Variance estimates would have to be calculated for the analyses. Since there are four predictor variables, there will be four components of unique variance (Uq), six components of second order common variance, four components of third order common variance, and one component of fourth order common variance.

The number of second, third and fourth order variance can be determined by the following formula:

$$NCn = \frac{N!}{n!(N-n)!}$$

where: N = number of predictor variables
n = number of variables taken at a time
NCn = number of combinations of N objects, taking n number at a time, independent of order

In the following example there are four predictor variables (N = 4). The number of second order commonalities (n = 2) can be determined by:

$$NCn = \frac{N!}{n!(N-n)!} = \frac{4x3x2x1}{2x1(2x1)} = 6$$

The number of third order commonalities (n=3) can be determined by:

$$NCn = \frac{N!}{n!(N-n)!} = \frac{4x3x2x1}{3x2x1(1)} = 4$$

The number of fourth order commonalities (n=4) is obviously 1.

The fifteen components are:

Uq(1)	C(1,2)	C(1,2,3)
Uq(2)	C(1,3)	C(1,2,4)
Uq(3)	C(1,4)	C(2,3,4)
Uq(4)	C(2,3)	C(1,3,4)
	C(2,4)	
	C(3,4)	C(1,2,3,4)

In this manner, one can separate the components of variance into mutually exclusive components that are estimates of unique and common variance. These components are additive and when summed will equal the total proportion of variance accounted for by the R^2_f of the full model.

Mood (1969) developed a rule for determining the R^2_s necessary for calculating unique and common components of variance. The rule is:

Develop products of the variables being considered.

For example, if one is interested in the $Uq(X_1)$ in an example with four predictor variables (X_1, X_2, X_3, X_4), first subtract that variable of interest (X_1) from one, multiplied by a -1, and the other variables in the equation

$$\text{rule: } -1(1-X_1)\ X_2,\ X_3,\ X_4 =$$
$$-X_2X_3X_4 + X_1X_2X_3X_4$$

Next, take the variables that are a product of the expansion and calculate the R^2_s that is indicated by each set (separated by + and - signs)

$$Uq(X_1) = -R^2y\cdot 234 + R^2y\cdot 1234$$

Similarly, the unique variance attributed to X_2 would be obtained by:

rule: $-1(1-X_2)X_1X_3X_4 =$

$-X_1X_3X_4 + X_1X_2X_3X_4 =$

$\therefore \quad Uq(X_2) = -R^2y\cdot134 + R^2y\cdot1234$

$Uq(X_3)$:

rule: $-1(1-X_3)X_1X_2X_4$

$-X_1X_2X_4 + X_1X_2X_3X_4$

$\therefore \quad Uq(X_3) = -R^2y\cdot124 + R^2y\cdot1234$

$Uq(X_4)$:

rule: $-1(1-X_4)X_1X_2X_3$

$-X_1X_2X_3 + X_1X_2X_3X_4$

$\therefore \quad Uq(X_4) = -R^2y\cdot123 + R^2y\cdot1234$

In this example there are six second order, four third order, and one fourth order commonalities. The rule for calculating these is basically the same as it is for the unique components. The variables of specific interest are placed in parentheses and subtracted from one. The other variables are multiplied by the product of the variable(s) in the parentheses and a -1.

Two examples of the second order, two of third order, and the fourth order commonalities will be presented.

C(1,2):

rule: $-1(1-X_1)(1-X_2)X_3X_4 =$

$-(X_3X_4-X_1X_3X_4 - X_2X_3X_4 + X_1X_2X_3X_4) =$

$-X_3X_4 + X_1X_3X_4 + X_2X_3X_4 - X_1X_2X_3X_4$

$\therefore \quad C(1,2) = -R^2y\cdot34 + R^2y\cdot134 + R^2y\cdot234 - R^2y\cdot1234$

C(2,3):

rule: $-1(1-X_2)(1-X_3)X_1X_4 =$

$-X_1X_4 + X_1X_3X_4 + X_1X_2X_4 - X_1X_2X_3X_4$

$\therefore$ $C(2,3) = -R^2y\cdot 14 + R^2y\cdot 134 + R^2y\cdot 124 - R^2y\cdot 1234$

C(1,2,3):

rule: $-1(1-X_1)(1-X_2)(1-X_3)X_4 =$

$-X_4 + X_1X_4 + X_2X_4 - X_1X_2X_4 + X_3X_4 - X_1X_3X_4$

$-X_2X_3X_4 + X_1X_2X_3X_4$

$\therefore$ $C(1,2,3) = -R^2y\cdot 4 + R^2y\cdot 14 + R^2y\cdot 24 - R^2y\cdot 124$

$+R^2y\cdot 34 - R^2y\cdot 134 - R^2y\cdot 234 + R^2y\cdot 1234$

C(1,2,4):

rule: $-1(1-X_1)(1-X_2)(1-X_4)X_3 =$

$-X_3 + X_1X_3 + X_2X_3 - X_1X_2X_3 + X_3X_4 - X_1X_3X_4$

$-X_2X_3X_4 + X_1X_2X_3X_4$

$\therefore$ $C(1,2,4) = -R^2y\cdot 3 + R^2y\cdot 13 + R^2y\cdot 23 - R^2y\cdot 123$

$+ R^2y\cdot 34 - R^2y\cdot 134 - R^2y\cdot 234 + R^2y\cdot 1234$

C(1,2,3,4):

rule: $-1(1-X_1)(1-X_2)(1-X_3)(1-X_4) =$

$-1 + X_1 + X_2 - X_1X_2 + X_3 - X_1X_3 - X_2X_3 + X_1X_2X_3$

$+ X_4 + X_1X_4 - X_2X_4 + X_1X_2X_4 - X_3X_4 + X_1X_3X_4$

$+ X_2X_3X_4 - X_1X_2X_3X_4$

$\therefore$ $C(1,2,3,4) = R^2y\cdot 1 + R^2y\cdot 2 - R^2y\cdot 12 + R^2y\cdot 3 - R^2y\cdot 13$

$- R^2y\cdot 23 + R^2y\cdot 123 + R^2y\cdot 4 + R^2y\cdot 14$

$- R^2y\cdot 24 + R^2y\cdot 124 - R^2y\cdot 34 + R^2y\cdot 134$

$+ R^2y\cdot 234 - R^2y\cdot 1234$

(Please note, when a one is by itself in the expansion, it is ignored in determining which R^2s should be calculated.)

As in the example, when there are four predictor variables, there will be fifteen components. One can easily see the horendous number of R^2s that have to be calculated for just four predictor variables in the full model. However, in using multiple regression, the investigator frequently has many more than four predictor variables. Therefore, the number of components can easily become impractical to handle. (This problem will be discussed later.)

For further details in how to calculate component analysis, see Mood (1969, 1971), Kerlinger (1973) and Houston and Bolding (1975).

One of the advantages of traditional factorial design is that in a four factorial design (ABCD factors) one can easily ascertain the independent variance accounted for by each factor in predicting the criterion. This is done by simply dividing the sum of squares for that particular factor by the total sum of squares. This procedure would also be used to obtain the independent proportion of variance accounted for by the interactions. However, one of the underlying assumptions of the traditional factorial design is that the factors are orthogonal. It also assumes among other things, equal number of subjects or proportionality of cells.

If the components in the analysis are independent of each other, that is the A,B,C,D, AXB, AXC, BXC, BXB, etc. are independent, then the F-tests for each of the fifteen components are independent. Therefore, the probabilities associated with each of the F's are theoretically accurate and no adjustments for non-independent F's are required.

When dealing with research in natural settings, one rarely has the control necessary to obtain data to meet these assumptions; however, when possible, one should try.

If one has a variety of nonorthogonal predictor variables and a variety of F-tests are used to determine if any one or set of these predictor variables are significant, then one is violating the underlying assumption of independence. Therefore, the probability associated with the F-test is inappropriate. That is, one would actually find more significant F's than is indicated by the probability associated with that specific F. Component analysis divides the sum of variances into independent partitions. Therefore, the F of any of these partitions is independent.

Similarly, if one considers the predictor variables as factors in a factorial design, then component analysis produces a set of components that are representationally like the components of the traditional factorial analysis. For example, in a two factorial design, the components are:

A main effect
B main effect
A x B interaction

In a regression equation in which two predictor variables are predicting Y a component analysis would produce:

$Uq(X_1)$

$Uq(X_2)$

$C(X_1,X_2)$

With all techniques, one must be aware of the limitations so that the technique can be employed most efficiently. The following are some of the limitations one should be sensitive to

when using component analysis:

1. An integral part of component analysis is the concept of Uq. Uq is operationally defined as:

> variance accounted for by a variable when entered last in a multiple regression equation.

Therefore, the Uq depends upon and is affected by the variables that are already under investigation. Even though the Uq is independent, in the set of variables for that sample, the variable is not independent.

2. As the number of predictor variables increases, the number of components generated increases drastically. So, if one has a large number of predictor variables, it may become impractical to calculate component analysis.

3. As the number of predictor variables increase, the number of higher order commonality components also increase. Just as it is difficult to interpret higher than third order interactions in traditional analysis of variance, it is also difficult to interpret higher than third order commonalities.

4. In examining some of the formuli for calculating the commonality components, one becomes sensitive to the possibility that some of the components can easily account for a negative proportion of variance. When this situation is encountered, it becomes very difficult to interpret or make conceptual sense out of the analysis.

5. With any non-manipulative research technique, "causation" cannot be assumed. A causal relationship can only be assumed in situations that have a true experimental design, i.e. a situation in which the experimenter has total control of the independent variable. Since

one of the major purposes for calculating component analysis is to attempt to improve the explanation of ex post facto research designs, this can lead one to mistakenly believe that the Uq accounted for by an independent variable with a criterion is of a causal nature.

6. Mood (1971) stated an important limitation one should consider. The unique variance (Uq) accounted for by an independent variable can change radically from situation to situation. However, the Uq attributed to a factor that the variable is a part of is not likely to change. Therefore, Mood suggests that one should group the variables based on the underlying concept they seem to be measuring. This would produce a more stable estimate. This group process will also have a side benefit of reducing the total number of predictor variables which will make the component analysis much more manageable. However, if one uses the procedure suggested by Mood, the weighting of each variable becomes a problem. Do the factors account for the same 100 percent of the proportion of variance accounted for when each variable is used separately? If not, one is loosing possibly significant information. Finally, it is difficult to decide on which variables should go together. Quite often, variables that look as if they are measuring the same underlying construct, are not.

Factor multiple regression is a procedure that may circumvent some of these problems. [Massy, (1965), Duff, Houston and Bloom (1971), Connett, Houston and Shaw (1972), Newman (1972).] It is a method that enables one to empirically determine the factors with which the variables are associated. If one calculates the factor

scores for each factor, and uses an orthogonal rotation, such as Varimax, then by definition, the predictor set of variables would be orthogonal. Then one can easily determine the relative importance of each factor by examining its beta weight.

Duff et al. (1971) found that principle component factor analysis with Varimax rotation and an eigen value of one as a factor cutoff, produced empirically determined factors which were very similar to the factors they subjectively determined. These subjective factors were formed by selecting subsets of their predictor variables which seemed to be measuring the same underlying constructs. (This is similar to what was suggested by Mood, 1971.) The advantages of using empirically determined factors as predictor variables in a regression equation are discussed by Connett et al. (1972) and some limitations of this procedure are discussed by Newman (1972).

It is the authors' opinion that the factor regression approach may be more appropriate than component analysis where one is interested in determining the unique variance accounted for, especially when the number of predictor variables is relatively large and there are a minimum of ten subjects for every variable. However, if one is interested in the commonality, the factor regression procedure is not appropriate. In this case, if one has a large number of variables and subjects, it is possible to use factor analysis with oblique rotation. This procedure will condense the large number of variables into factors which can be used as a new set of predictor variables. Since these factors may be oblique (correlated) one may then wish to do a component analysis

which will yield estimates of the unique and common variance attributed to the factors. Obviously, the oblique solutions lack many of the desirable characteristics which make the orthogonal solution easier to interpret. However, there are times that a researcher may be interested in the common proportion of variance attributed to factors which are theoretically and empirically related.

BIBLIOGRAPHY

Bottenberg, R. and Ward, J. Applied Multiple Linear Regression U.S. Department of Commerce, Office of Technical Services, AD413128, 1963.

Connett, W.; Houston, S.; and Shaw, D. The Use of Factor Regression in Data Analysis. Multiple Linear Regression Viewpoints, March, 1972, 46-49.

Duff, W.; Houston, S.; and Bloom, S. A Regression/Principle Components Analysis of School Outputs. Multiple Linear Regression Viewpoints, May, 1971, 5-18.

Houston, S. and Bolding, J. Part, Partial and Multiple Correlation In Commonality Analysis of Multiple Linear Regression Models. Multiple Linear Regression Viewpoints, January, 1975, 36-40.

Kerlinger, F.N. and Pedhazur, E. J. Multiple Regression In Behavioral Research, New York: Holt, Rinehart and Winston Inc., 1973.

Massy, W. F. Principal Components Regression In Exploratory Statistical Research. Journal of American Statistical Association, March, 1965, 234-256.

McNeil, K.; Kelly, F. J.; McNeil, J. Testing Research Hypotheses Using Multiple Linear Regression. In Press

Mood, A.M. Macro-Analysis of the American Educational System. Operations Research, 1969, 17, 770-784.

Mood, A.M. Partitioning Variance in Multiple Regression Analyses as a Tool for Developing Learning Models. A.E.R.J., March, 1971, 191-202.

Newman, I. Some Further Considerations of Using Factor Regression Analysis. Multiple Linear Regression Viewpoints, October, 1972, 39-41.

"A publication of the Multiple Linear Regression Special Interest Group of the American Educational Research Association, published primarily to facilitate communication, authorship, creativity and exchange of ideas among the members of the group and others in the field. As such it is not sponsored by the American Educational Research Association nor necessarily bound by the Association's regulations.

TABLE OF CONTENTS