This is to certify that the

thesis entitled

THE EFFECTS ON THE WEIGHING COEFFICIENTS OF ERROR OF MEASUREMENT IN THE RANDOM PREDICTORS OF A QUANTAL RESPONSE TECHNIQUE

presented by

Robert Alan Carr

has been accepted towards fulfillment of the requirements for

Ph.D degree in Education

William H Schmidt
Major professor

Date 5/2/78

O-7639

THE EFFECTS ON THE [illegible]
OF ERRORS [illegible]
RANDOM PREDICTORS [illegible]

A DISSERTATION

Submitted to
Michigan State [illegible]
in partial fulfillment of [illegible]
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel [illegible]
and Educational Psychology

197[illegible]

THE EFFECTS ON THE WEIGHTING COEFFICIENTS
OF ERRORS OF MEASUREMENT IN THE
RANDOM PREDICTORS OF A QUANTAL RESPONSE TECHNIQUE

By

Robert Alan Carr

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel Services
and Educational Psychology

1978

this resear

fields. Th

of this res

criterion

factor varia

weighting co

set of weigh

used with

Th

effects that

factor varian

The second pr

of the weight

Q errors of

The

Weighting coef

level predi

ABSTRACT

THE EFFECTS ON THE WEIGHTING COEFFICIENTS OF ERRORS OF MEASUREMENT IN THE RANDOM PREDICTORS OF A QUANTAL RESPONSE TECHNIQUE

By

Robert Alan Carr

The random predictor quantal response model is examined in this research. Quantal response models are qualitative data analysis models. The general situation addressed by quantal response models of this research concern the relationship between a single qualitative criterion (quantal response) and one or more quantitative random predictor variables. This relationship is expressed in a series of weighting coefficients. For each category of the criterion there is a set of weighting coefficients with one weighting coefficient associated with each predictor variable.

The first problem of this research is to determine the effects that various levels of errors of measurement in the random predictor variables will have on the values of the weighting coefficients. The second problem is to describe a procedure for producing estimates of the weighting coefficients which would be produced if there were no errors of measurement in the random predictors.

The procedure used describes: a quantal response model and weighting coefficients based on the assumed existence of error-free (latent) predictors; a quantal response model and weighting

coefficien

and error

measuremen

the two q

efficient

value of t

observed p

and algeb

applicable

response m

variance s

would be t

would resu

response m

errors of

the predic

cases were

models, th

factor and

predictor

factor. F

data were

situations

and four g

coefficients based on the observed predictor counterparts, which contain errors of measurement, of these latent predictors; and two measurement models which provide two possible relationships between the two quantal response models. Then the value of a weighting coefficient based on the use of latent predictors is compared to the value of the corresponding weighting coefficient based on the use of observed predictors with a known level of error of measurement. General algebraic results of the comparison were sought which would be applicable across the universe of situations which define the quantal response models.

Then a set of estimation procedures called analysis of covariance structures procedures were examined to determine if they could be used to derive estimates of the weighting coefficients which would result if latent predictors were available for use in the quantal response model.

No generally applicable algebraic results of the effects of errors of measurement were discovered which apply to all possible random predictor quantal response models. Therefore, the two simplest cases were examined in detail. For one predictor quantal response models, the weighting coefficient based on the single observed predictor underestimates the weighting coefficient based on the latent predictor by a factor equal to the reliability of the observed predictor. For the two predictor quantal response model no simple results were found. However, the results for the universe of possible situations for two predictor quantal response models can be separated into four general categories and three special case categories of situatio

generic p

served

together

incremente

m produc

based on

uires d

approxima

earch al

teration

eighting

ions are

uter prog

situations, where all situations within a category will have the same generic pattern of comparison between weighting coefficients based on observed and latent predictors. The derivation of these categories, together with their descriptions and examples of situations, are documented in this research.

The analysis of covariance structures procedures as applied to produce maximum-likelihood estimates of the weighting coefficients based on the use of latent predictors are described. Since these procedures do not lead to explicitly solvable estimates, a numerical approximation procedure is needed to produce the estimates. This research also briefly describes a computer program, using numerical iteration procedures, which can produce the desired estimates of the weighting coefficients. Two simulated, non-randomly selected situations are included to illustrate the results of the use of the computer program.

for the or

was partic

this date

prompt rev

allowed th

A

thesis whi

Porter, Dr.

T

helped to k

was taken t

and Mother

B

than any ot

daily suppo

continued w

made. Her

the degree w

all the typ

ACKNOWLEDGEMENTS

During the final stages of writing the thesis and preparing for the oral examination Dr. William Schmidt, my committee chairman, was particularly helpful. I thank him for encouraging me to set an orals date earlier than I originally thought was feasible and for his prompt review and return of large volumes of draft material which allowed the early completion date to come to fruition on schedule.

Also, I appreciate the time and effort in reviewing the thesis which was spent by each member of my committee: Dr. Andrew Porter, Dr. Donald Freeman and Dr. Dennis Gilliland.

The care, concern, support and understanding of my family helped to keep me focused on the task over the nearly seven years it has taken to complete the degree. Thank you Mother and Dad Birchard and Mother and Dad Carr.

But most of all I wish to thank my wife, Jessie, who, more than any other individual, helped me to complete this degree. Her daily support and encouragement provided an atmosphere conducive to continued work even during the long periods where little progress was made. Her unselfish sacrifice of her own time allowed me to work on the degree whenever my time, energy and interest allowed. She also did all the typing for the drafts of the thesis. Her willingness to spend

long hours typing from handwritten drafts, especially just prior to the orals, helped to make an early completion of the oral examination possible. Thank you dear. Without your support I never would have finished.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Section

reliabil

problem a

less than

measures

not near

of measur

reliabili

problems

measureme

ever, whe

of one or

of analyt

questiona

De

of various

search.

of work do

of standar

qualitativ

CHAPTER 1

Section A: Introduction

In the social sciences one of the major problems concerns the reliability of various measures used in analytical procedures. The problem arises since the measures of interest in many situations are less than perfectly reliable. This problem with the reliability of measures also occurs in the physical sciences but in general it is not nearly as severe a problem as in the social sciences. When errors of measurement are present only to a very minimal extent, that is the reliabilities of all measures are near one, logic suggests that few problems are likely to be encountered if these minimal errors of measurement are ignored in using various analytical procedures. However, when more than very minimal errors of measurement are present in one or more measures of interest, the interpretation of results of analytical procedures based on the fallible measures becomes questionable.

Determining the effects of the use of such fallible measures in various analytical procedures has been the focus of previous research. In two presentations Cochran (1968, 1970) presents a review of work done on the effects of errors of measurement on a wide range of standard techniques of analysis including both quantitative and qualitative data analysis models. Porter (1971) provides a review of effects of

tical techn

which illu

efficients

errors of r

Assakul an

errors in c

Th

errors of r

model, a q

review some

response mo

for this re

effects of errors of measurement on four common quantitative statistical techniques. Wiley and Hornik (1973) provide a data example which illustrates the effects on interpretation of regression coefficients when fallible measures are used with no adjustment for errors of measurement. Bross (1954), Mote and Anderson (1965) and Assakul and Proctor (1967) provide discussions of the effect of errors in classification for three qualitative data analysis models.

This research extends the investigation of the effects of errors of measurement to include another qualitative data analysis model, a quantal response technique. The sections below will briefly review some of the research cited above, describe the quantal response model of interest for this research and present the problem for this research.

Section B:

Coo

which can a

when errors

not conside

in the pred

than one as

a situation

two fallibl

ables have

size relati

observed sc

opposite of

efficients

coefficient

tionship

based on ob

$\rho_1' < \rho_2'$

Al

meter valu

when the e

Wi

trates the

measures a

study of co

ficient in

Section B: Errors of measurement in quantitative data analysis models

Cochran (1968, pp. 656, 657) has shown some of the problems which can arise in the interpretation of regression coefficients when errors of measurement in the predictor variables are present but not considered. In this example the presence of errors of measurement in the predictors is indicated by a reliability coefficient of less than one associated with each predictor. Specifically, he discusses a situation with one dependent variable assumed to be error-free and two fallible predictor variables, i.e., each of the predictor variables have reliabilities less than one. In Cochran's example, the size relationship between estimated regression coefficients based on observed scores with no consideration of errors of measurement is the opposite of the size relationship between estimated regression coefficients based on latent scores. In this example the regression coefficients based on the latent scores, β_1 and β_2, have the relationship $\beta_1 > \beta_2$ while the corresponding regression coefficients based on observed scores, β_1' and β_2', have the relationship $\beta_1' < \beta_2'$.

Although Cochran's example is based on a specific set of parameter values it does illustrate the potential problem which can arise when the effects of errors of measurement are not considered.

Wiley and Hornik (1973) provide a data example which illustrates the potential for misinterpretation which exists when fallible measures are used in a regression analysis. The data came from a study of communication processes conducted in Central America. Sufficient information was collected to provide estimates of the true regression

individuall

predictor v

predictor w

reliable.

Con

variable th

i.e., a la

reliable pr

the regress

with no cor

from the es

For

had virtual

near zero)

tionship.

estimated s

errors of m

regression

tween the p

the estima

stronger tr

predictor t

more reliab

is, when er

which had a

regression coefficients. Each of two dependent variables were individually regressed on two fallible predictor variables. The two predictor variables were positively related to each other. One predictor was highly reliable while the other was considerably less reliable.

Considering the estimated true relationship, for one dependent variable the more reliable predictor had the stronger relationship (i.e., a larger estimated true regression coefficient) and the less reliable predictor had virtually no relationship. In this situation the regression coefficients estimated solely from the observed scores with no consideration of errors of measurement did not differ greatly from the estimated true regression coefficients.

For the second dependent variable the more reliable predictor had virtually no relationship (i.e., a true regression coefficient near zero) while the less reliable predictor had a very large relationship. In this situation, however, the regression coefficients estimated solely from the observed scores with no consideration of errors of measurement differed markedly from the estimated true regression coefficients. Because of the positive relationship between the predictors, not only did the errors of measurement attenuate the estimated relationship of the less reliable predictor (with the stronger true relationship) but also some of the relationship of this predictor to the dependent variable is spuriously attributed to the more reliable predictor (with virtually no true relationship). That is, when errors of measurement were not considered the predictor which had a high estimated true regression coefficient but low reliabilit

almost hal

the other

near zero

the first

was relati

Th

interpret

from the

measureme

considera

of the es

T

effects

search,

tted th

analysis

reliability produced an observed regression coefficient which was almost half the size of the true regression coefficient. However, the other predictor which had estimated true regression coefficient near zero but had high reliability and was positively correlated with the first predictor produced an observed regression coefficient which was relatively large.

Therefore in this example for the second dependent variable interpretations based on regression coefficients estimated solely from the use of observed scores with no consideration of errors of measurement in the predictors would lead to conclusions which are considerably different from the conclusions based on an examination of the estimated true regression coefficients.

These two examples provide an indication of the adverse effects of errors of measurements in regression analysis. Other research, as reviewed by Cochran (1968, 1970) and Porter (1971) has cited the problems which can occur with other quantitative data analysis models.

Section C:

The

restricted s

examines the

2 x 2 table

classified

association

be without

other of the

error. The

the populat

dimension a

population

In

each of the

positives 5

ference in

second dire

estimate of

Type I erro

the power o

If,

and false p

I errors of

Mo

lation whi

Section C: Errors of measurement in qualitative data analysis models

The problems associated with errors of measurement are not restricted solely to quantitative data analysis models. Bross (1954) examines the effects of errors in classification in one variable in a 2 × 2 table. In this case samples from each of two populations are classified into one of two categories on a second dimension. The association of any unit with a particular population is assumed to be without error but the classification of that unit into one or the other of the two categories on the second dimension is subject to error. The interest in this case is in the proportion of units from one population which are classified into one category on the second dimension as compared to the proportion of units from the second population classified into the same category on the second dimension.

In this case, if the proportions of false negatives from each of the two populations are equal and the proportions of false positives from each of the populations are also equal then the difference in the proportions of units assigned to one category of the second dimension when using the observed proportions is an underestimate of the difference based on the true proportions. Here the Type I errors of the test of significance will remain unchanged but the power of the test will be decreased.

If, however, the assumptions of the equality of false negatives and false positives across populations is not appropriate, the Type I errors of the test of significance are increased.

Mote and Anderson (1965) work with units from a single population which are classified into one of several categories on some dimensio

of the po

sample o

statisti

lation p

will hav

several

structed

tions re

in all c

fication

Here er

sidered

dimensi

dimensio

based o

based o

this co

test in

misclas

Hassk

to obse

availab

and An

dimension. When errors of classification are present the estimates of the population proportions in any one category based on a random sample of units from the population are biased. And standard statistical tests where the null hypotheses specify particular population proportions or relationships among the population proportions will have increased Type I errors. Mote and Anderson (1965) discuss several special case situations where statistical tests can be constructed which will have correct Type I errors. Each of these situations requires some specialized information which may not be available in all cases.

Assakul and Proctor (1967) examine the effects of misclassification in the $r \times c$ contingency table on the standard χ^2 test. Here errors of classification in each of the two dimensions are considered. If and only if the errors of misclassification in one dimension are independent of errors of misclassification in the second dimension then the null hypothesis for the χ^2 test of independence based on the true population proportions implies the null hypothesis based on the observed population proportions and vice-versa. Under this condition the Type I errors are unchanged but the power of the test in large samples is never increased and nearly always reduced by misclassification.

When the errors of misclassification are not independent, Assakul and Proctor show how to make an appropriate χ^2 test based on observed proportions when some very specialized information is available. In this case they follow the same procedure used by Mote and Anderson (1965) for one of their special case situations.

These three references indicate that errors of measurement can produce problems in quantitative data analysis models as well as qualitative data analysis models. A summary of the three references above with more detail, including algebraic formulations and major conclusions is found in Cochran (1968).

Section D: T

The

of the effec

nique for ra

qualitative

The

techniques c

criterion (q

variables.

the values c

occurrence c

model the re

weighting co

there is a s

associated w

of the weigh

strength of

Ther

That is, for

criterion i

ables is kn

the categor

discussion

1980, chap

And

estimates o

Section D: The data analysis model to be examined in this research

The research to be presented here extends the investigation of the effects of errors of measurement to a quantal response technique for random predictor variables. Quantal responses models are qualitative data analysis models.

The general situation which is addressed by quantal response techniques concerns the relationship between a single qualitative criterion (quantal response) and one or more quantitative predictor variables. The relationship of interest is the relationship between the values on the set of predictor variables and the probability of occurrence of a particular quantal response. In a quantal response model the relationship of interest is expressed in a series of weighting coefficients. For each category of the criterion variable there is a set of weighting coefficients with one weighting coefficient associated with each predictor variable. The sign and relative size of the weighting coefficient give an indication of the type and strength of the relationship.

These techniques can be employed as classification procedures. That is, for a particular subject whose classification on the criterion is not known but whose set of values on the predictor variables is known, these techniques provide information about which of the categories of the criterion is most probable. For a general discussion about the classification of observations see Anderson (1958, chapter 6) and Tatsuoka (1974).

Another use of quantal response techniques is to determine estimates of the relationship between the predictor variables and the

probabi
of the
predict
respons

types,
paramet

techniq
tween t
levels

Smith (
respons

either
tion of
chain e
tionshi
and the
and the
reasona
variabl
their m

When a

probability of occurrence of a category of the criterion. Estimates of the relationship, in the form of a weighting coefficient for each predictor variable, can be produced. It is this use of quantal response techniques that is of interest in this research.

Quantal response techniques fall into one of two general types, each with a model and associated procedures for estimating the parameters of the model.

The distinction between the two types of quantal response techniques depends on the type of relationship that is postulated between the predictor variables and the probability of occurrence of levels of the criterion.

Since McSweeney and Schmidt (1974) and Cornfield, Gordon and Smith (1960) both provide discussion about the two types of quantal response techniques, only a brief description will be given here.

The first type of quantal response model assumes that, either by the sampling procedure and/or by the theoretical consideration of the location of the predictor variables late in the causal chain ending with the criterion and the indirect mediational relationship of the predictors between other links in the causal chain and the criterion, a <u>functional</u> relationship between the predictors and the probability of occurrence of levels of the criterion seems reasonable. For this type of quantal response model the predictor variables are treated as fixed mathematical variables regardless of their method of selection.

The second type of quantal response model becomes appropriate when a functional relationship between the predictors and the probability

to be a reas

because the

links [in th

most likely

have a direc

as an indire

case a stat

are then tre

ables.

The

in the biolo

potency. A

determinatio

number of ex

with the dr

The criterio

polychotomo

to note her

and the dru

probability

The

forms of th

McSweeney a
Predictor Va

probability of occurrence of levels of the criterion does not seem to be a reasonable assumption. Because of sampling techniques and/or because the predictor variables "...can be thought of as intermediate links [in the causal chain] or as outcomes themselves then it is most likely true that the factors influencing the predictors will have a direct effect on the probability of [the criterion] as well as an indirect influence mediated through the predictors."[1] In this case a statistical relationship is assumed and the predictor variables are then treated as random variables rather than mathematical variables.

The first type of quantal response model has been employed in the biological sciences, particularly in assessments of drug potency. A simplistic prototypical experiment would involve the predetermination of a fixed number of drug dosage levels. A preset number of experimental animals at each dosage level would be injected with the drug and their response on some criterion would be noted. The criterion might be dichotomous (e.g. alive or dead) or polychotomous (e.g. alive, moribund or dead). The important thing to note here is that the dosage level is experimentally controllable and the drug dosage level is expected to have a direct effect on the probability of survival or non-survival.

The second type of quantal response model which will be the focus of this research will generally be more appropriate for social

---

[1]McSweeney and Schmidt; "Quantal Response Techniques for Random Predictor Variables," AERA presentation, 1974.

science ap

a hypothe

mastery o

is the cr

expressed

this case

of subjec

The choic

is not u

of subje

variable

an exper

plausibl

(e.g. mo

that af

the obs

could b

ables."

have a

Probab

predic

of int

Planta

respon

---

2 Ibid,

science applications. McSweeney and Schmidt (1974, pp. 5, 6) pose a hypothetical example of this second type of model. In the example, mastery of a learning task (with two levels, mastery and non-mastery) is the criterion of interest. The probability of mastery is to be expressed as a function of entry level knowledge of the student. "In this case, the data would be generated by classifying a random sample of subjects on the basis of their entry knowledge and their mastery. The choice of levels of entry knowledge of the subjects to be observed is not under the control of the experimenter and as such the number of subjects exhibiting X_k units of entry knowledge is a random variable (usually taking on only the values zero and one) rather than an experimenter-imposed-constraint. Furthermore, it would be plausible, logically, to postulate the existence of other variables (e.g. motivation, need for achievenemt, interest in subject matter) that affect both [entry level knowledge] and mastery. Consequently the observed relationship between the predictor and the criterion could be a result of the direct relationship of each to other variables."[2] Thus in this case the predictor variables are expected to have a statistical, as opposed to a functional, relationship with the probability of occurrence of levels of the criterion. Therefore the predictor variables are treated as random variables.

This second type of quantal response model which is the model of interest for this research will be called the Random Predictor Quantal Response Model to distinguish it from other uses of quantal response techniques not involved in this research.

---

[2]Ibid, p. 6.

Sectio

Respons

The cri

which b

ables a

assume

the qua

That is

unique

Howeve

error.

situati

the rel

qualita

Quantit

Respons

of the

Since

that i

estima

well th

genera

Section E: Presentation of the problem for this research

In the presentation of the model it is clear that the Quantal Response model contains both qualitative and quantitative variables. The criterion is a qualitative variable with two or more categories which have no necessary ordered relationship. The predictor variables are quantitative statistical variables which can conceivably assume any real value, positive, negative or zero. For this research the qualitative criterion variable will be assumed to be error-free. That is for any given unit the classification of that unit into one unique category of the criterion is accomplished without error. However, one or more of the predictor variables may be measured with error.

Therefore the impetus for this research is provided by a situation such as the following. There is an interest in determining the relationship between the occurrence of some category of a qualitative criterion variable and the true values on one or more quantitative predictor variables where the Random Predictor Quantal Response Model is the model of choice.

Since the relationship of interest involves the true values of the predictor variables, rather than the observed values, and since it is known for other statistical models (e.g. Linear Regression) that in the presence of errors of measurement the relationships estimated on the basis of observed scores do not always approximate well the relationships estimated on the basis of true scores, two general questions arise.

The

estimated re

variables co

the predicto

pending on a

of measureme

tion between

Sinc

jected based

the predicto

second quest

which will p

true scores

Thes

The first qu

of the resea

various leve

in the Rando

indicates th

Area two inv

coefficients

variable wer

Chap

Predictor Qu

Predictors.

will be prese

The first question is: How much variation is there in the estimated relationship based on observed scores of the predictor variables compared to the relationship based on the true scores of the predictor variables? The response to this question may vary depending on a variety of factors such as the extent to which errors of measurement are present in the predictor variables and the correlation between the predictor variables, among others.

Since some difference in estimated relationships can be expected based on research with other models and since true scores on the predictor variables are typically not directly measurable, the second question becomes: What estimation procedures can be developed which will provide the estimated relationship of interest based on true scores of the predictor variables?

These two questions provide the direction for this research. The first question provides the direction for the first major area of the research. Area one involves determining the effects of various levels of errors of measurement on the weighting coefficients in the Random Predictor Quantal Response Model. The second question indicates the direction of the second major area of the research. Area two involves developing techniques to estimate the weighting coefficients which would result if the true score for each predictor variable were available for use in the model.

Chapter 2 will provide a detailed presentation of the Random Predictor Quantal Response Model for observed predictors and for true predictors. In each case, the general model and two special cases will be presented along with various simplifying derivations and other intere

relates the

will be def

Chap

for the two

chapter 4 f

For

observed pre

to estimate

cedures ofte

the model a

Joreskog (1

cussions of

will be used

Chap

Program, us

estimates of

the use of t

Chap

major areas

study.

other interesting algebraic results. The measurement model which relates the true predictors with the associated observed predictors will be defined in this chapter.

Chapters 3 and 4 will present the results of the research for the two major areas identified above; chapter 3 for Area one and chapter 4 for Area two.

For chapter 4 an expanded measurement model will relate the observed predictors to the latent predictors. The task will then be to estimate latent parameters from the observed data. A set of procedures often used where errors of measurement are incorporated in the model are termed Analysis of Covariance Structures (ANCOVST). Joreskog (1970), Wiley, Schmidt and Bramble (1973) present discussions of ANCOVST procedures. Modifications of these procedures will be used in chapter 4.

Chapter 5 will contain a brief description of a computer program, using the methods described in Chapter 4, which can produce estimates of the latent weighting coefficients. An illustration of the use of the computer program will also be provided.

Chapter 6 will provide the summary of the results of both major areas of this research along with recommendations for further study.

Section

will be

of inter

dictor w

between

a partic

be prese

solely o

errors o

Predicto

identifi

efficien

on the u

Latent R

efficien

weightin

represen

the crite

ments of

CHAPTER 2

Section A: The Random Predictor Quantal Response Model - An Introduction

In this chapter the Random Predictor Quantal Response Model will be presented along with various algebraic derivations and results of interest. The weighting coefficients associated with each predictor which provide an indication of the conditional relationship between a given predictor and the probability of classification into a particular category of the criterion will be identified.

In fact, two Random Predictor Quantal Response Models will be presented. The first model to be presented (Section B) is based solely on the use of observed predictors with no consideration of errors of measurement. This model will be called the Observed Random Predictor Quantal Response Model. The weighting coefficients identified from this model will be called the observed weighting coefficients. The second model to be presented (Section C) is based on the use of latent predictors. This model will be called the Latent Random Predictor Quantal Response Model. The weighting coefficients identified from this model will be called the true weighting coefficients. Although the true weighting coefficients represent the relationship of interest between the predictors and the criterion, seldom if ever will there be available direct measurements of the latent predictors. Thus, there will not be available estimates (

ments of th

Si

observed p

ship betwe

tence betw

weighting

predictors

the classi

needed for

Se

in this ch

subsequent

estimates of the true weighting coefficients based on direct measurements of the latent predictors.

Since the measures that are available in practice are for observed predictors only, it will be necessary to develop a relationship between the observed predictors and the latent predictors, hence between the observed weighting coefficients and the true weighting coefficients. The model which will relate the observed predictors to the latent predictors is a measurement model based on the classical true-score model. This model and its extensions as needed for this research will be presented in Section D below.

Section E will summarize the results of the work presented in this chapter and its relationship to the work to be presented in subsequent chapters.

Section B:

The

observed pr

mber of c

The Random

has been pre

Schmidt (19

in structure

the notation

needed for t

Let

Y = j, (j =

assigned a u

numbers 1,2,

does not cons

the numbering

tional conver

categories.

dictor variab

One c

ciples is to

category of t

i.e., Pr(Y =

provides the b

Let f

ditional distr

Section B: The Observed Random Predictor Quantal Response Model

The most general case of the quantal response model using observed predictors allows for a polychotomous criterion with J number of categories ($J \geq 2$) and p multiple predictors ($p \geq 1$). The Random Predictor Quantal Response Model using observed predictors has been presented for this most general case by McSweeney and Schmidt (1974, pp. 10-13). The model presented below is identical in structure to the model presented by McSweeney and Schmidt. Only the notational form has been changed to accommodate adjustments needed for this research.

Let Y be the criterion variable which takes on values $Y = j$, $(j = 1,2,\ldots,J)$ where each category of the criterion is assigned a unique value as an identifier chosen arbitrarily from the numbers $1,2,\ldots,J$. The quantal response model does not require and does not consider any ordering among the categories. Therefore, the numbering of the categories of the criterion is merely for notational convenience and need imply no ordered relationship among the categories. Let $\underset{\sim}{X}$ be the $p \times 1$ random vector of observed predictor variables where $\underset{\sim}{X}' = (X^1\ X^2\ \ldots\ X^p)$.

One of the traditional interests in quantal response techniques is to find the probability associated with membership in each category of the criterion given values for each predictor variable, i.e., $\Pr\{Y = j|\underset{\sim}{X}\}$ for $j = 1,2,\ldots,J$. It is this interest which provides the basis for the quantal response models.

Let $f(\underset{\sim}{X}|Y = j)$ for $j = 1,2,\ldots,J$ represent the J conditional distributions of the random vectors for predictor variables for the J

tural dis

are assume

matrices

Th

response

(2.1) Pri

where p

(5

A

sider a p

L

for the J categories of the criterion. Traditionally the conditional distributions of the predictor variables for the J categories are assumed to be p-variate normal with identical $p \times p$ covariance matrices Σ and mean vectors $\underset{\sim}{\mu}_X^{(j)}$ $(j = 1,2,\ldots,J)$.

Therefore the conditional probability of occurrence of response $Y = k$ $(k = 1,2,\ldots,J)$ can be expressed as

$$(2.1)\quad \Pr\{Y = k|\underset{\sim}{X}\} = P_k = 1/[1 + \sum_{\substack{j=1\\ j\neq k}}^{J} \frac{p_j}{p_k} \exp\{-\tfrac{1}{2}[(\underset{\sim}{X} - \underset{\sim}{\mu}_X^{(j)})'\Sigma^{-1}(\underset{\sim}{X} - \underset{\sim}{\mu}_X^{(j)})$$

$$- (\underset{\sim}{X} - \underset{\sim}{\mu}_X^{(k)})'\Sigma^{-1}(\underset{\sim}{X} - \underset{\sim}{\mu}_X^{(k)})]\}]$$

where p_j = the unconditional probability of occurrence of category j $(j = 1,2,\ldots,J)$,

$\underset{\sim}{\mu}_X^{(j)}$ = the $p \times 1$ mean vector of the distribution of observed predictor variables for level j of the criterion $(j = 1,2,\ldots,J)$, and

Σ = the $p \times p$ covariance matrix of each of the J conditional distributions of the observed predictor variables.

A simplification of this expression (2.1) is possible. Consider a portion of the exponent, as follows.

Let

$\hat{x}$

$\hat{x}$

$\hat{x}$

symmetr

$\hat{x} =$

$=$

$R =$

$=$

$k = 1, 2, \ldots$

$$Q = (\underset{\sim}{X} - \mu_X^{(j)})'\Sigma^{-1}(\underset{\sim}{X} - \mu_X^{(j)}) - (\underset{\sim}{X} - \mu_X^{(k)})'\Sigma^{-1}(\underset{\sim}{X} - \mu_X^{(k)})$$

$$Q = (\underset{\sim}{X}'\Sigma^{-1} - \mu_X^{(j)'}\Sigma^{-1})(\underset{\sim}{X} - \mu_X^{(j)}) - (\underset{\sim}{X}'\Sigma^{-1} - \mu_X^{(k)'}\Sigma^{-1})(\underset{\sim}{X} - \mu_X^{(k)})$$

$$Q = \underset{\sim}{X}'\Sigma^{-1}\underset{\sim}{X} - \mu_X^{(j)'}\Sigma^{-1}\underset{\sim}{X} - \underset{\sim}{X}'\Sigma^{-1}\mu_X^{(j)} + \mu_X^{(j)'}\Sigma^{-1}\mu_X^{(j)}$$

$$- \underset{\sim}{X}'\Sigma^{-1}\underset{\sim}{X} + \mu_X^{(k)'}\Sigma^{-1}\underset{\sim}{X} + \underset{\sim}{X}'\Sigma^{-1}\mu_X^{(k)} - \mu_X^{(k)'}\Sigma^{-1}\mu_X^{(k)} .$$

Note: Since $\underset{\sim}{X}'\Sigma^{-1}\mu_X^{(k)}$ is a 1×1 matrix and Σ is symmetric then $\underset{\sim}{X}'\Sigma^{-1}\mu_X^{(k)} = (\underset{\sim}{X}\Sigma^{-1}\mu_X^{(k)})' = \mu_X^{(k)'}\Sigma^{-1}\underset{\sim}{X}$.

Then

$$Q = -2\mu_X^{(j)'}\Sigma^{-1}\underset{\sim}{X} + 2\mu_X^{(k)'}\Sigma^{-1}\underset{\sim}{X} + \mu_X^{(j)'}\Sigma^{-1}\mu_X^{(j)} - \mu_X^{(k)}\Sigma^{-1}\mu_X^{(k)}$$

$$= 2[(\mu_X^{(k)} - \mu_X^{(j)})'\Sigma^{-1}\underset{\sim}{X}] + \mu_X^{(j)'}\Sigma^{-1}\mu_X^{(j)} - \mu_X^{(k)'}\Sigma^{-1}\mu_X^{(k)} .$$

Let

$$R = \frac{p_j}{p_k}\exp\{-\tfrac{1}{2}Q\} = \exp\{+ \mathrm{Ln}(\frac{p_j}{p_k}) - \tfrac{1}{2}Q\}$$

$$= \exp\{\mathrm{Ln}\,\frac{p_j}{p_k} + \frac{1}{2}[\mu_X^{(k)'}\Sigma^{-1}\mu_X^{(k)} - \mu_X^{(j)'}\Sigma^{-1}\mu_X^{(j)}] - (\mu_X^{(k)} - \mu_X^{(j)})'\Sigma^{-1}\underset{\sim}{X}\}.$$

Let $\underset{\sim}{\beta}_{k\cdot j} = \Sigma^{-1}(\mu_X^{(k)} - \mu_X^{(j)})$ and let

$$\alpha_{k\cdot j} = -\mathrm{Ln}(\frac{p_j}{p_k}) - \frac{1}{2}[\mu_X^{(k)'}\Sigma^{-1}\mu_X^{(k)} - \mu_X^{(j)'}\Sigma^{-1}\mu_X^{(j)}].$$

Therefore $R = \exp\{-(\alpha_{k\cdot j} + \underset{\sim}{\beta}'_{k\cdot j}\underset{\sim}{X})\}$.

Therefore the general model (2.1) can also be expressed (for $k = 1,2,\ldots,J$) as:

(2.2) Pr

where

and

I

of observ

are indic

that the

meters f

the lette

in the de

tion for

vector

of

variable

of a com

magnitud

icates

of the f

category

$$(2.2)\quad \Pr\{Y = k|\underset{\sim}{X}\} = P_k = 1/[1 + \sum_{\substack{j=1\\ j\neq k}}^{J} \exp\{-(\alpha_{k\cdot j} + \underset{\sim}{\beta}'_{k\cdot j}\underset{\sim}{X})\}]$$

where

$$\alpha_{k\cdot j} = -\mathrm{Ln}\left(\frac{p_j}{p_k}\right) - \frac{1}{2}[\underset{\sim}{\mu}_X^{(k)'}\Sigma^{-1}\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)'}\Sigma^{-1}\underset{\sim}{\mu}_X^{(j)}]$$

and

$$\underset{\sim}{\beta}_{k\cdot j} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)}) \quad \text{for } j \neq k,\ j,k = 1,2,\ldots,J.$$

In this formulation of the model (2.2) the $p \times 1$ vectors of observed weighting coefficients $\underset{\sim}{\beta}_{k\cdot j}$ $(j \neq k,\ j,k = 1,2,\ldots,J)$ are indicated. The $k\cdot j$ subscript notation is used to indicate that the value of the weighting coefficient is dependent upon parameters from two distinct categories j and k. The ordering of the letters in the subscript indicates the order of the subtraction in the definition of the weighting coefficient, i.e., $\underset{\sim}{\beta}_{k\cdot j} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)})$ or $\underset{\sim}{\beta}_{3\cdot 7} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(3)} - \underset{\sim}{\mu}_X^{(7)})$. An interpretation for $\underset{\sim}{\beta}_{k\cdot j}$ for some k,j $(j \neq k,\ j,k = 1,2,\ldots,J)$ associates vector $\underset{\sim}{\beta}_{k\cdot j}$ with category k. In this interpretation the components of $\underset{\sim}{\beta}_{k\cdot j}$ indicate the conditional weighting attached to each predictor variable in differentiating between categories k and j. The sign of a component indicates the direction of the weighting while the magnitude of a component indicates the strength of the weighting.

For each category k $(k = 1,2,\ldots,J)$ expression (2.2) indicates the existence of $J - 1$ vectors of weighting coefficients of the form $\underset{\sim}{\beta}_{k\cdot j}$ $(j \neq k,\ j.k = 1,2,\ldots,J)$ associated with category k. Therefore, for all J categories there will be

$J\cdot(J-1)$ vectors of weighting coefficients to consider. The values of these $J\cdot(J-1)$ vectors of weighting coefficients are not mutually independent. In fact, Appendix A.1 demonstrates that only a base set of $J-1$ vectors of weighting coefficients associated with some arbitrary category k of the criterion $(k = 1,2,\ldots,J)$ need to be considered. Each of the other $(J-1)$ vectors of weighting coefficients associated with any other category $k' \neq k$ are shown to be linear combinations of vectors from the base set of $J-1$ vectors associated with category k.

The interest for the first area of this research is in the relationship between the probability of occurrence of some category of the criterion variable and the values of the predictor variables. This relationship for observed predictors is given by the components of the vector of observed weighting coefficients.

Since there is interest in the inidividual components of the vector of weighting coefficients there will be some utility in deriving an expression for the individual weighting coefficient associated with some observed predictor X^q $(q = 1,2,\ldots,p)$.

Consider now some vector of observed weighting coefficients associated with some arbitrary category k $(k = 1,2,\ldots,J)$, $\underset{\sim}{\beta}_{k\cdot j}$ $(j \neq k,\ j,k = 1,2,\ldots,J)$, where $\underset{\sim}{\beta}_{k\cdot j} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)})$. Consider also some observed predictor X^q $(q = 1,2,\ldots,p)$ and the observed weighting coefficient component of vector $\underset{\sim}{\beta}_{k\cdot j}$ associated with predictor X^q, call it β_q, i.e. $\beta_q = (\underset{\sim}{\beta}_{k\cdot j})_q$.

The task now is to derive an expression for the single observed weighting coefficient for observed predictor X^q, i.e. β_q, from t

with c

Predic

(2.3)

observ

model

tion

set o

(2.4)

where

Since

(2.5)

from the vector of observed weighting coefficients $\beta_{k\cdot j}$ associated with category k $(j \neq k,\ j,k = 1,2,\ldots,J)$ of the Observed Random Predictor Quantal Response Model.

Therefore

$$(2.3)\quad \beta_q = (\beta_{k\cdot j})_q = (\Sigma^{-1}(\mu_X^{(k)} - \mu_X^{(j)}))_q = (\Sigma^{-1}\mu_X^{(k)})_q - (\Sigma^{-1}\mu_X^{(j)})_q .$$

The particular order in which the predictors from the set of observed predictors are considered is not important in the general model. Hence no generality is lost, and considerable help in notation is gained, by considering X^q to be the first predictor in the set of observed predictors, i.e. $q = 1$.

(2.4) Thus $\beta_1 = (\Sigma^{-1}\mu_X^{(k)})_1 - (\Sigma^{-1}\mu_X^{(j)})_1 .$

Consider $(\Sigma^{-1}\mu_X^{(k)})_1$ first.

<u>Note:</u>[1] $\Sigma^{-1} = \frac{1}{|\Sigma|}\Sigma^{c'}$ where $\Sigma^{c'} = [C_{ij}]$ $i,j = 1,\ldots,p$

where C_{ij} = the cofactor of element $(\Sigma)_{ij}$. Thus

$$(\Sigma^{-1}\mu_X^{(k)})_1 = (\frac{1}{|\Sigma|}\Sigma^{c'}\mu_X^{(k)})_1 = \frac{1}{|\Sigma|}(\Sigma^{c'}\mu_X^{(k)})_1 .$$

Since $\mu_X^{(k)}$ is a $p \times 1$ vector, $\mu_X^{(k)'} = [\mu_{X^1}^{(k)}\ \mu_{X^\ell}^{(k)}\ \ldots\ \mu_{X^2}^{(k)}\ \ldots\ \mu_{X^p}^{(k)}]$

$$(2.5)\quad (\Sigma^{-1}\mu_X^{(k)})_1 = \frac{1}{|\Sigma|}(\Sigma^{c'}\mu_X^{(k)})_1 = \frac{1}{|\Sigma|}\sum_{\ell=1}^{p} C_{\ell 1}\mu_{X^\ell}^{(k)}$$

$$= \frac{C_{11}}{|\Sigma|}[\mu_{X^1}^{(k)} + \sum_{\ell=2}^{p}\frac{C_{\ell 1}}{C_{11}}\mu_{X^\ell}^{(k)}].$$

---

[1] Davis, Philip J.; <u>The Mathematics of Matrices</u>; Ginn and Co., Boston, 1965, p. 182.

No

element (

No

first pred

and use th

the regres

(2.6) X

where b

regressio

Let

Following

where

Variance

served p

use of t

distingu

<u>Note</u>: $C_{\ell 1} = (-1)^{\ell+1}|M^{\Sigma}_{\ell 1}|$ where $M^{\Sigma}_{\ell 1}$ is the minor of element $(\Sigma)_{\ell 1}$.

Now consider the p predictors in cateogry k. Choose the first predictor X^1 as a dependent variable in a linear regression and use the remaining $p - 1$ predictors as independent variables in the regression. That is:

$$(2.6) \qquad X^1 = b^{(k)}_{1\cdot 0} + b^{(k)}_{1\cdot 2} X^2 + b^{(k)}_{1\cdot 3} X^3 + \ldots + b^{(k)}_{1\cdot p} X^p + \varepsilon$$

where $b^{(k)}_{1\cdot 0}$ is the constant of the regression and $b^{(k)}_{1\cdot m}$ is the regression coefficient associated with predictor X^m $(m = 2,\ldots,p)$.

Let

$$\underset{\sim}{B}^{(k)\prime} = [b^{(k)}_{1\cdot 2}\ b^{(k)}_{1\cdot 3} \ldots b^{(k)}_{1\cdot p}] \ .$$

Following a discussion from Bock (1975, pp. 136-138):

$$\underset{\sim}{B}^{(k)} = \Sigma^{-1}_{22}\, \underset{\sim}{\Sigma}_{21}$$

where

$$\Sigma = \left[\begin{array}{c|c} \sigma^2_{11} & \underset{\sim}{\Sigma}_{12} \\ \hline \underset{\sim}{\Sigma}_{21} & \Sigma_{22} \end{array}\right] \begin{array}{l} 1 \\ p-1 \end{array}$$
$$\qquad\qquad 1 \qquad p-1$$

Since the values of $\underset{\sim}{B}^{(k)}$ involve only elements of the covariance matrix Σ which is common to the distribution of the observed predictors in all categories, there is no need to continue the use of the superscript (k) to denote the set of coefficients or to distinguish individual regression coefficients.

g' = [b

where

(1.7)

where

ponent

the or

number

of [

accor

All i

this

numbe

In pr

eleme

Prec

col

Therefore let $\underset{\sim}{B}^{(k)} = \underset{\sim}{B} = \Sigma_{22}^{-1} \underset{\sim}{\Sigma}_{21}$ with $\underset{\sim}{B}' = [b_{1\cdot 2}\ b_{1\cdot 3}\ \ldots\ b_{1\cdot p}]$.

Consider some component of $\underset{\sim}{B}$, call it $b_{1\cdot m}$ $(m = 2,\ldots,p)$

$$b_{1\cdot m} = (\Sigma_{22}^{-1} \underset{\sim}{\Sigma}_{21})_m = (\frac{1}{|\Sigma_{22}|} \Sigma_{22}^{c'} \underset{\sim}{\Sigma}_{21})_m$$

where Σ_{22}^{c} is the matrix of cofactors of elements of Σ_{22}

$$= \frac{1}{|\Sigma_{22}|} (\Sigma_{22}^{c'} \underset{\sim}{\Sigma}_{21})_m$$

$$(2.7) \qquad b_{1\cdot m} = \frac{1}{|\Sigma_{22}|} \sum_{i=1}^{p-1} C_{i(m-1)}^{\Sigma_{22}} (\underset{\sim}{\Sigma}_{21})_i$$

where the subscript i here is associated with the renumbered components of Σ_{22} beginning with $(\Sigma_{22})_{11}$ which is actually σ_{22}^2 in the original numbering system. A dual notation system (element numbering system) for the elements of Σ_{22} and the other submatrices of Σ will be used. In one system the element will be numbered in accordance to its row and column location in the total matrix Σ. All individual elements identified as σ_{jj}^2 or σ_{jk} will be using this system. In the second numbering system the element will be numbered in accordance to its row and column location in the submatrix.

In equation (2.7), $C_{i(m-1)}^{\Sigma_{22}}$ is the cofactor of the element in row i and column $m - 1$ of matrix Σ_{22}, and $(\underset{\sim}{\Sigma}_{21})_i$ is the i^{th} element of vector $\underset{\sim}{\Sigma}_{21}$.

<u>Note 1</u>: $b_{1\cdot m}$ is the regression coefficient associated with predictor X^m in equation (2.6).

<u>Note 2</u>: Let M_{11}^{Σ} be the minor of the element in row one and column one of Σ. Then

$W^r_{11} =$

matrices

and row

W^r_{21}

p-1) x

$$M_{11}^{\Sigma} = \Sigma_{22} = \begin{bmatrix} \sigma_{22}^2 & \sigma_{23} & \cdots & \sigma_{2p} \\ \sigma_{32} & \sigma_{33}^2 & \cdots & \sigma_{3p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p2} & \sigma_{p3} & \cdots & \sigma_{pp}^2 \end{bmatrix} \quad \text{and} \quad |M_{11}^{\Sigma}| = |\Sigma_{22}| .$$

<u>Note 3</u>: $M_{\ell 1}^{\Sigma}$ and Σ_{22} are nearly identical $(p - 1) \times (p - 1)$ matrices, except that row one (1) of $M_{\ell 1}^{\Sigma}$ does not occur in Σ_{22} and row $(\ell - 1)$ of Σ_{22} does not occur in $M_{\ell 1}^{\Sigma}$.

$$\underset{(p-1)\times(p-1)}{M_{\ell 1}^{\Sigma}} = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ \ell-1 \\ \\ \vdots \\ p-1 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & p-1 \end{array} \\ \begin{bmatrix} \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1p} \\ \sigma_{22}^2 & \sigma_{23} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{\ell-1,2} & \sigma_{\ell-1,3} & \cdots & \sigma_{\ell-1,p} \\ \sigma_{\ell+1,2} & \sigma_{\ell+1,3} & \cdots & \sigma_{\ell+1,p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p2} & \sigma_{p3} & \cdots & \sigma_{pp}^2 \end{bmatrix} \end{array}$$

Σ_{22}

$p-1) \times ($

I

deleted f

Therefore

identical

of Σ_{22}

i.e.

(p

M_{ij} =

first, i

a and c

W_{21} b

$$\underset{(p-1)\times(p-1)}{\Sigma_{22}} = \begin{array}{c} \\ 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ \ell-2 \\ \ell-1 \\ \ell \\ \cdot \\ \cdot \\ \cdot \\ p-1 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & p-1 \end{array} \\ \left[\begin{array}{cccc} \sigma_{22}^2 & \sigma_{23} & \cdots & \sigma_{2p} \\ \sigma_{32} & \sigma_{33}^2 & \cdots & \sigma_{3p} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{\ell-1,2} & \sigma_{\ell-1,3} & \cdots & \sigma_{\ell-1,p} \\ \sigma_{\ell 2} & \sigma_{\ell 3} & \cdots & \sigma_{\ell p} \\ \sigma_{\ell+1,2} & \sigma_{\ell+1,3} & \cdots & \sigma_{\ell+1,p} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{p2} & \sigma_{p3} & & \sigma_{pp}^2 \end{array}\right] \end{array}$$

If row 1 were deleted from $M^{\Sigma}_{\ell 1}$ and row $(\ell - 1)$ were deleted from Σ_{22} the resulting matrices would be identical. Therefore the minor of any element in the first row of $M^{\Sigma}_{\ell 1}$ will be identical to the minor associated with the element in row $(\ell - 1)$ of Σ_{22} which has the same column subscript as the element in $M^{\Sigma}_{\ell 1}$, i.e. $\underset{(p-2)\times(p-2)}{M^{M^{\Sigma}_{\ell 1}}_{1i}} = \underset{(p-2)\times(p-2)}{M^{\Sigma_{22}}_{(\ell-1)i}}$ for $i = 1,\ldots,p-1$.

<u>Note 4</u>: For any symmetric matrix $M_{ij} = M'_{ji}$. Therefore $|M_{ij}| = |M_{ji}|$.

Thus in category k, for any predictor X^m other than the first, i.e. $m = 2,\ldots,p$, the cofactor C_{m1} of the element in row m and column 1 of Σ is: $C_{m1} = (-1)^{m+1}|M^{\Sigma}_{m1}|$ and expanding $|M^{\Sigma}_{m1}|$ by cofactors of the first row of M^{Σ}_{m1}

$$C_{m1} = (-1)^{m+1}\left[\sum_{i=1}^{p-1} (-1)^{i+1} \cdot |M^{M^{\Sigma}_{m1}}_{1i}| \cdot \sigma_{(i+1)1}\right]$$

where i

$\bar{W}_{gl}$ and

; which

(1.8)

b_i

(1.9)

where

for m

colu

Com

where $i = 1,\ldots,p-1$ represents the renumbered row subscripts for M_{ml}^{Σ} and $\sigma_{(i+1)1}$ is the element in row $(i+1)$ and column 1 of Σ which is also the i^{th} element of the vector $\underset{\sim}{\Sigma}_{21}$.

Therefore

$$(2.8)\qquad \frac{C_{ml}}{C_{11}} = \frac{(-1)^{m+1} \cdot (-1) \sum_{i=1}^{p-1} (-1)^{i} \cdot |M_{1i}^{M_{ml}^{\Sigma}}| \cdot \sigma_{(i+1)1}}{|M_{11}^{\Sigma}|}\ .$$

From equation (2.7):

$$b_{1\cdot m} = \frac{1}{|\Sigma_{22}|} \sum_{i=1}^{p-1} C_{i(m-1)}^{\Sigma_{22}} (\underset{\sim}{\Sigma}_{21})_i$$

$$= \frac{1}{|\Sigma_{22}|} \sum_{i=1}^{p-1} [(-1)^{i+m-1} \cdot |M_{i(m-1)}^{\Sigma_{22}}| \cdot \sigma_{(i+1)1}]$$

$$(2.9)\qquad b_{1\cdot m} = \frac{(-1)^{m-1}}{|\Sigma_{22}|} \sum_{i=1}^{p-1} (-1)^{i} \cdot |M_{i(m-1)}^{\Sigma_{22}}| \cdot \sigma_{(i+1)1}$$

where $i = 1,\ldots,p-1$ represents the renumbered column subscripts for matrix Σ_{22} and $\sigma_{(i+1)1}$ is the element in row $(i+1)$ and column 1 of Σ which is also the i^{th} element of the vector $\underset{\sim}{\Sigma}_{21}$.

Compare equations (2.8) and (2.9).

$$(-1)^{m+1} = (-1)^{2}(-1)^{m-1} = (-1)^{m-1}$$

$|M_{11}^{\Sigma}| = |\Sigma_{22}|$ from Note 2 above, and

$|M_{i(m-1)}^{\Sigma 22}| = |M_{(m-1)i}^{\Sigma_{22}}| = |M_{1i}^{M_{ml}^{\Sigma}}|$ from Note 4 above and Note 3 above respectively.

Therefore,

(2.10)

where b

with pre

factors

predicto

(2.11)

This

where

represe

categor

(2.12)

from eq

(2.10) $$b_{1\cdot m} = -\frac{C_{m1}}{C_{11}} .$$

Thus equation (2.5) becomes:

$$(\Sigma^{-1}\underset{\sim}{\mu}_X^{(k)})_1 = \frac{C_{11}}{|\Sigma|}[\mu_{X^1}^{(k)} - \sum_{\ell=2}^{p} b_{1\cdot\ell}\,\mu_{X^\ell}^{(k)}]$$

where $b_{1\cdot\ell}$ $(\ell = 2,\dots,p)$ is the regression coefficient associated with predictor X^ℓ when predictor X^ℓ and the other $p - 2$ predictors are regressed on the first predictor as in equation (2.6).

In general for any predictor X^q in category k, since any predictor can be put first in the ordering of the set of predictors:

(2.11) $$(\Sigma^{-1}\underset{\sim}{\mu}_X^{(k)})_q = \frac{C_{qq}}{|\Sigma|}[\mu_{X^q}^{(k)} - \sum_{\substack{\ell=1\\ \ell\neq q}}^{p} b_{q\cdot\ell}\,\mu_{X^\ell}^{(k)}] .$$

<u>Note</u>: $$\mu_{X^q}^{(k)} = b_{q\cdot 0}^{(k)} + b_{q\cdot 1}\mu_{X^1}^{(k)} + \dots + b_{q\cdot q-1}\mu_{X^{q-1}}^{(k)}$$

$$+ b_{q\cdot q+1}\mu_{X^{q+1}}^{(k)} + \dots + b_{q\cdot p}\mu_{X^p}^{(k)} .$$

Thus $$b_{q\cdot 0}^{(k)} = \mu_{X^q}^{(k)} - \sum_{\substack{\ell=1\\ \ell\neq q}}^{p} b_{q\cdot\ell}\,\mu_{X^\ell}^{(k)}$$

where $b_{q\cdot 0}^{(k)}$ is the constant in the regression equation above and represents the intercept of the regression hyperplane of X^q for category k data.

Therefore

(2.12) $$(\Sigma^{-1}\underset{\sim}{\mu}_X^{(k)})_q = \frac{C_{qq}}{|\Sigma|}\, b_{q\cdot 0}^{(k)} .$$

In a similar manner for category j $(j \neq k,\ j,k = 1,2,\dots,J)$, from equation (2.11):

(2.23)

A

becomes

(2.24)

U

$\xi_q = (\xi_k$

(2.25)

factor, .

category

regressi

factor a

(2.26)

from equ

$$(2.13)\qquad (\Sigma^{-1}\underset{\sim}{\mu}_X^{(j)})_q = \frac{C_{qq}}{|\Sigma|}\,[\mu^{(j)}_{X_q} - \sum_{\substack{\ell=1\\ \ell\neq q}}^{p} b_{q\cdot\ell}\,\mu_{X_\ell}].$$

And similar to the work for equation (2.12), equation (2.13) becomes

$$(2.14)\qquad (\Sigma^{-1}\underset{\sim}{\mu}_X^{(j)})_q = \frac{C_{qq}}{|\Sigma|}\, b^{(j)}_{q\cdot 0}\ .$$

Using the results (2.12) and (2.14), equation (2.3) $\beta_q = (\underset{\sim}{\beta}_{k\cdot j})_q = (\Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)}))_q$ becomes

$$(2.15)\qquad \beta_q = \frac{C_{qq}}{|\Sigma|}\,(b^{(k)}_{q\cdot 0} - b^{(j)}_{q\cdot 0})\ .$$

Thus, the observed weighting coefficient for any given predictor, in any vector of weighting coefficients associated with any category of the criterion is the difference in intercepts between two regression hyperplanes, in the form of (2.6), multiplied by a scale factor associated with the chosen predictor.

Additional formulations of $\beta_q = (\beta_{k\cdot j})_q$ can be produced.

<u>Note</u>: $|\Sigma| = \sum_{\ell=1}^{p} \sigma_{q\ell} C_{q\ell}$ for some $q = 1,2,\ldots,p$

$$= C_{qq}\,[\sigma^2_{qq} + \sum_{\substack{\ell=1\\ \ell\neq q}}^{p} \sigma_{q\ell}\,\frac{C_{q\ell}}{C_{qq}}]$$

$$(2.16)\qquad |\Sigma| = C_{qq}\,[\sigma^2_{qq} - \sum_{\substack{\ell=1\\ \ell\neq q}}^{p} b_{q\cdot\ell}\,\sigma_{q\cdot\ell}]$$

from equation (2.10) generalized. Therefore:

(2.17a)

from equa

(2.17b)

from equa

E

two linea

I

and to e

above se

i.e., (2

mental

for sing

(2.17a)
$$\beta_q = \frac{b_{q\cdot 0}^{(k)} - b_{q\cdot 0}^{(j)}}{\sigma_{qq}^2 - \sum_{\substack{\ell=1 \\ \ell\neq q}}^{p} b_{q\cdot\ell}\sigma_{q\cdot\ell}}$$

from equations (2.15) and (2.16) or

(2.17b)
$$\beta_q = \frac{(\mu_{x^q}^{(k)} - \mu_{x^q}^{(j)}) - \sum_{\substack{\ell=1 \\ \ell\neq q}}^{p} b_{q\cdot\ell}(\mu_{x^\ell}^{(k)} - \mu_{x^\ell}^{(j)})}{\sigma_{qq}^2 = \sum_{\substack{\ell=1 \\ \ell\neq q}}^{p} b_{q\cdot\ell}\sigma_{q\ell}}$$

from equations (2.11), (2.13), (2.3), and (2.16).

Equation (2.17b) expresses $\beta_q = (\underset{\sim}{\beta}_{k\cdot j})_q$ as the ratio of two linear combinations each of the form

$$\theta_q - \sum_{\substack{\ell=1 \\ \ell\neq q}}^{p} b_{q\cdot\ell}\theta_\ell \ .$$

Although there are other ways to express the general model and to express the single weighting coefficients, those presented above seem to have the greatest utility for the work which follows, i.e., (2.2), for the general case of the Observed Random Predictor Quantal Response Model and (2.15) and (2.17b) for the expressions for single weighting coefficients.

Section C

T

of a gen

factor.

Random P

velopmen

This mos

with J

P = j (

arbitrar

Let T

where T

tional c

covarian

and mea

for the

model f

of T

variati

model a

used as

paramet

Section C: The Latent Predictor Quantal Response Model

The development of this model depends on the assumed existence of a general unobservable entity, a latent measurement for each predictor. Then the development of the most general case of the Latent Random Predictor Quantal Response Model follows easily from the development of the Observed Random Predictor Quantal Response Model. This most general case model allows for a polychotomous criterion with J categories $(J \geq 2)$ and p multiple predictors $(p \geq 1)$.

Let Y be the criterion variable which takes on values $Y = j$ $(j = 1,2,\ldots,J)$ where each category of the criterion is arbitrarily assigned a unique value from the numbers $1,2,\ldots,J$. Let $\underset{\sim}{T}$ be the $p \times 1$ random vector of latent predictor variables where $\underset{\sim}{T}' = (T^1\ T^2\ \ldots\ T^p)$.

Assuming that for each category of the criterion the conditional distribution of $\underset{\sim}{T}$ is p-variate normal with identical $p \times p$ covariance matrices, Φ, assumed homogeneous across all categories, and mean vectors $\underset{\sim}{\mu}_T^{(j)}$ $(j = 1,2,\ldots,J)$, the derivation of a model for the latent predictors exactly parallels the derivation of the model for observed predictors with $\underset{\sim}{T}$ in place of $\underset{\sim}{X}$, Φ in place of Σ and $\underset{\sim}{\mu}_T^{(j)}$ in place of $\underset{\sim}{\mu}_X^{(j)}$ $(j = 1,2,\ldots,J)$. One other variation in notation will be made to differentiate between this model and the observed predictor model. An asterisk (*) will be used as a superscript for some parameters to indicate that the parameter is associated with the latent predictor model.

for the La

(2.18) Pr

=

where p_j

(j)

and

T

can be us

simplific

notation.

(2.19) P

where

Using the above replacements in the derivation of expression (2.1) produces for some category k $(k = 1,2,\ldots,J)$ an expression for the Latent Random Predictor Quantal Response Model.

$$(2.18)\quad \Pr\{Y = k|\underset{\sim}{T}\} = P_k^*$$

$$= 1/[1 + \sum_{\substack{j=1\\j\neq k}}^{J} \frac{p_j}{p_k} \exp\{-\tfrac{1}{2}[(\underset{\sim}{T} - \underset{\sim}{\mu}_T^{(j)})'\Phi^{-1}(\underset{\sim}{T} - \underset{\sim}{\mu}_T^{(j)})$$

$$- (\underset{\sim}{T} - \underset{\sim}{\mu}_T^{(k)})'\Phi^{-1}(\underset{\sim}{T} - \underset{\sim}{\mu}_T^{(k)})]\}]$$

where p_j = the unconditional probability of occurrence of category j $(j = 1,2,\ldots,J)$,

$\underset{\sim}{\mu}_T^{(j)}$ = the $p \times 1$ mean vector of the distribution of latent predictor variables for level j of the criterion $(j = 1,2,\ldots,J)$

and Φ = the $p \times p$ covariance matrix of each of the J conditional distributions of the latent predictor variables.

The argument which produced expression (2.2) from (2.1) can be used to produce from expression (2.18) the following simplification, by merely replacing $\underset{\sim}{X}$ by $\underset{\sim}{T}$ and adjusting other notation to indicate that latent parameters are involved.

$$(2.19)\quad \Pr\{Y = k|\underset{\sim}{T}\} = P_k^* = 1/[1 + \sum_{\substack{j=1\\j\neq k}}^{J} \exp -(\alpha_{k\cdot j}^* + \beta_{k\cdot j}^{*'} \underset{\sim}{T})\}]$$

where

$$\alpha_{k\cdot j}^* = -\ln(\frac{p_j}{p_k}) - \frac{1}{2}[\underset{\sim}{\mu}_T^{(k)'}\Phi^{-1}\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(j)'}\Phi^{-1}\underset{\sim}{\mu}_T^{(j)}]$$

and

for j

Respons

efficie

script

interpr

in Sect

to cons

cients

All oth

vectors

model h

express

2.17b,

of true

associa

with pr

the exp

2.20,

and

$$\underset{\sim}{\beta}^{*}_{k\cdot j} = \Phi^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(j)})$$

for $j \neq k$, $j,k = 1,2,\ldots,J$.

In this formulation of the Latent Random Predictor Quantal Response Model (2.19) the $p \times 1$ vectors of latent weighting coefficients $\underset{\sim}{\beta}^{*}_{k\cdot j}$ $(j \neq k,\ j,k = 1,2,\ldots,J)$ are indicated. The subscript notation carries a parallel interpretation to that of the interpretation for the observed weighting coefficients given above in Section B of this chapter.

As with the observed weighting coefficients it is necessary to consider only a base set of $J - 1$ vectors of latent coefficients associated with some arbitrary category k $(k = 1,2,\ldots,J)$. All other latent weighting coefficients are linear combinations of vectors from the base set. See Appendix A.2 for proof.

Two other results of interest from the observed predictor model have direct parallels in terms of latent predictors. Two expressions for the individual weighting coefficient β_q, (2.15) and (2.17b), become: for some category k $(k = 1,2,\ldots,J)$, some vector of true weighting coefficients $\underset{\sim}{\beta}^{*}_{k\cdot j}$ $(j \neq k,\ j,k = 1,2,\ldots,J)$ associated with category k and some individual latent predictor T^q $(q = 1,2,\ldots,p)$, the single weighting coefficient associated with predictor T^q, call it β^{*}_q where $\beta^{*}_q = (\underset{\sim}{\beta}^{*}_{k\cdot j})_q$ has the following expressions.

$$(2.20) \qquad \beta^{*}_q = \frac{C^{*}_{qq}}{|\Phi|}\,(b^{(k)*}_{q\cdot 0} - b^{(j)*}_{q\cdot 0})$$

and,

(2.21)

where th

from the

scripted

Predict

for the

utility

and,

$$(2.21)\quad \beta_q^* = \frac{(\mu_{T^q}^{(k)} - \mu_{T^q}^{(j)}) - \sum_{\substack{\ell=1 \\ \ell \neq q}}^{p} b_{q\cdot\ell}^* (\mu_{T^\ell}^{(k)} - \mu_{T^\ell}^{(j)})}{\sigma_{qq}^{2*} - \sum_{\substack{\ell=1 \\ \ell \neq q}}^{p} b_{q\cdot\ell}^* \sigma_{q\ell}^*}$$

where the * as superscript indicates the parameter is a parameter from the latent predictor model which corresponds to the non-superscripted parameter from the observed predictor model.

Expression (2.19) for the general case of the Latent Random Predictor Quantal Response Model and expressions (2.20) and (2.21) for the single weighting coefficients appear to have the most utility for the work which follows.

Section D

I

been deve

the model

ictors a

the param

introduce

models.

T

the class

above X

is the p

variate e

the p ×

tion

(1.2)

where

and

and Nov

case are

Lord, F

ures,

Section D: The Measurement Model

In the two previous sections two quantal response models have been developed. Although there is an obvious parallelism between the models as evidenced by a comparison of (2.2) for observed predictors and (2.19) for latent predictors there is no link between the parameters of the models. The purpose of this section is to introduce a measurement model which will provide the link between the models.

The basic measurement model to be used in this research is the classical true score model. Following the notation introduced above $\underset{\sim}{X}$ is the $p \times 1$ random vector of observed predictors and $\underset{\sim}{T}$ is the $p \times 1$ random vector of true predictors. Using a multivariate extension of information presented in Lord and Novick (1974)[2], the $p \times 1$ error random vector $\underset{\sim}{E}$ is defined by the linear relation

(2.22) $$\underset{\sim}{X} = \underset{\sim}{T} + \underset{\sim}{E}$$

where $\underset{\sim}{X}' = [X^1 \; X^2 \; \ldots \; X^p]$

$\underset{\sim}{T}' = [T^1 \; T^2 \; \ldots \; T^p]$

and $\underset{\sim}{E}' = [e^1 \; e^2 \; \ldots \; e^p]$.

The assumptions of this classical model again taken from Lord and Novick (1974) with appropriate extensions to the multivariate case are:

---

[2] Lord, F.M. and Novick, M.R.; Statistical Theories of Mental Test Scores, Addison Wesley, Reading, 1974, p. 56.

(2.22a)

that is,

jects is

(2.22b)

where

populat

correla

factor

ρ^2 = d
$_{PP}$

(2.23c)

That is

scores

cov (

between

true p

zero

from e

(2.24a

for an

(2.23a) $$\mathcal{E}\,\underset{\sim}{E} = \underset{\sim}{0}$$

that is, the expected value of the errors in the population of subjects is zero.

(2.23b) $$\text{Var}\,(\underset{\sim}{E}) = \Psi^2$$

where Ψ^2 is a diagonal matrix of error variances. That is, in the population of subjects the errors between any two predictors are uncorrelated $(\rho e^i e^j = 0,\ i \neq j)$ and the error variances of any predictor i $(i = 1,2,\ldots,p)$ is given by $\sigma^2_{e^i}$. Therefore $\underset{p\times p}{\Psi^2} = \text{diag}\{\sigma^2_{e^1}\ \sigma^2_{e^2}\ \ldots\ \sigma^2_{e^p}\}$.

(2.23c) $$\text{Cov}\,(\underset{\sim}{E},\underset{\sim}{T}) = \underset{p\times p}{[0]}$$

That is, in the population of subjects the covariance between error scores and true predictor scores on the same predictor is zero $(\text{cov}\,(T^i,\ e^i) = 0$ for any i $(i = 1,2,\ldots,p)$ and the covariance between error scores on some predictor i $(i = 1,2,\ldots,p)$ and the true predictor scores for any other predictor $j,\ j \neq i$ is also zero $(\text{cov}\,(T^j,\ e^i) = 0)$. Several important results come immediately from expressions (2.22) through (2.23c) above.

(2.24a) $$\mu_X^{(k)} = \mu_T^{(k)}$$

for any category k $(k = 1,2,\ldots,J)$.

To demonstrate this, consider any category k $(k = 1,2,\ldots,J)$

$$\mu_X^{(k)} \equiv \mathcal{E}_k(\underset{\sim}{X}) = \mathcal{E}_k(\underset{\sim}{T} + \underset{\sim}{E}) = \mathcal{E}_k(\underset{\sim}{T}) + \mathcal{E}_k(\underset{\sim}{E}) \ .$$

S

(1.14b)

for any

k = 1,2

[illegible]

With ea

must be

Respons

Model

For use

Since $\mathcal{E}_k(\underset{\sim}{E}) = \underset{\sim}{0}$ by (2.23a),

$$\underset{\sim}{\mu}_X^{(k)} = \mathcal{E}_k(\underset{\sim}{T}) + \underset{\sim}{0} = \mathcal{E}_k(\underset{\sim}{T}) \equiv \underset{\sim}{\mu}_T^{(k)} .$$

(2.24b) $$\Sigma = \Phi + \Psi^2$$

for any category k $(k = 1,2,\ldots,J)$.

To demonstrate this result, consider any category k $(k = 1,2,\ldots,J)$

$$\Sigma \equiv \mathrm{Var}(\underset{\sim}{X}) = \mathcal{E}_k(\underset{\sim}{X}\,\underset{\sim}{X}') - \mathcal{E}_k(\underset{\sim}{X})\mathcal{E}_k(\underset{\sim}{X}')$$

$$= \mathcal{E}_k((\underset{\sim}{T} + \underset{\sim}{E})(\underset{\sim}{T} + \underset{\sim}{E})') - \mathcal{E}_k(\underset{\sim}{T} + \underset{\sim}{E})\mathcal{E}_k((\underset{\sim}{T} + \underset{\sim}{E})')$$

$$= \mathcal{E}_k(\underset{\sim\sim}{TT}' + \underset{\sim\sim}{ET}' + \underset{\sim\sim}{TE}' + \underset{\sim\sim}{EE}') - \mathcal{E}_k(\underset{\sim}{T})\mathcal{E}_k(\underset{\sim}{T}')$$

$$= \underbrace{\mathcal{E}_k(\underset{\sim\sim}{TT}') - \mathcal{E}_k(\underset{\sim}{T})\mathcal{E}_k(\underset{\sim}{T}')}_{} + \underbrace{\mathcal{E}_k(\underset{\sim\sim}{EE}')}_{} + \underbrace{\mathcal{E}_k(\underset{\sim\sim}{ET}')}_{} + \underbrace{\mathcal{E}_k(\underset{\sim\sim}{TE}')}_{}$$

$$= \mathrm{Var}(\underset{\sim}{T}) + \mathrm{Var}(\underset{\sim}{E}) + \underset{p\times p}{[0]} + \underset{p\times p}{[0]}$$

$$= \mathrm{Var}(\underset{\sim}{T}) + \mathrm{Var}(\underset{\sim}{E})$$

$$\Sigma = \Phi + \Psi^2 .$$

<u>Note</u>: $\mathcal{E}_k(\underset{\sim}{E}\underset{\sim}{T}') = \mathcal{E}_k(\underset{\sim}{T}\underset{\sim}{E}') = [0]$ where $[0]$ is a $p \times p$ matrix with each element a zero. This follows from (2.23c).

Expressions (2.22) through (2.24b) provide the basic relationship between the elements of the Observed Random Predictor Quantal Response Model (2.2) and the Latent Random Predictor Quantal Response Model (2.19).

The measurement model as given by (2.22) will be sufficient for use with the research for Area one in determining the effects of unreliabi

this work

paralleli

ing obser

be suffic

veloping

the basic

ficient.

Area two

T

vides for

predictor

replicate

presented

Factor f

for mult

all of w

various

recorded

cluded i

with m

(11a)

where

unreliability of the predictors on the weighting coefficients. In this work to be presented in Chapter 3, the direct one-to-one parallelism between an unobserved true predictor and the corresponding observed predictor with the connection provided by (2.22) will be sufficient. However, for the Area two work which involves developing estimation procedures for the latent weighting coefficients, the basic measurement model represented by (2.22) will not be sufficient. An expansion of the basic model will be necessary. The Area two work will be presented in Chapter 4.

The expansion of the basic model which will be needed provides for the use of replicate observed measurements for each of the predictor variables. A more detailed discussion of the need for replicate measures and their use in the estimation process will be presented in Chapter 4.

The expanded model assumes the existence of a single true predictor for each construct to be considered as a predictor but allows for multiple observed measurements to be recorded for each predictor, all of which provide information about the predictor. Since the various observed measurements for a given predictor may not be recorded using the same scale of measurement a scale factor is included in the model.

The model relating some latent predictor T^j $(j = 1,2,...,p)$ with m number of replicate observed measures becomes

(2.25a) $$\underset{m\times 1}{\underset{\sim}{X}^j} = \underset{m\times 1}{\underset{\sim}{\Lambda}^j} T^j + \underset{m\times 1}{\underset{\sim}{E}^j}$$

where $\underset{\sim}{\Lambda}^j$ is an $m \times 1$ vector of scale factors which relates the latent p

an m

assumed

allowan

lar ch

scaling

referen

to be c

theoret

cedure

procedu

chosen

factor

the sca

scale c

For thi

theoret

first c

fr

scale

arbitra

extensi

have

chs

latent predictor T^j to the non-error portion of $\underset{\sim}{X}^j$, and $\underset{\sim}{E}^j$ is an $m \times 1$ vector of errors for predictor j.

For a given latent predictor each observed replication is assumed to provide a measure of this latent predictor including allowance for error and for the scale of measurement of that particular observed replication. Before a value can be assigned to a scaling factor it is necessary to assume the presence of some master reference scale of measurement for each predictor. For most variables to be considered, this assumption is not typically operationally or theoretically feasible. It is thus necessary to resort to a procedure commonly used in analysis of covariance structures (ANCOVST) procedures. In this procedure one of the observed replications is chosen to provide the reference scale and the value of the scaling factor of that replication is arbitrarily set to 1. By doing this the scales of all the other replications can be referenced to the scale of the chosen replication rather than to some absolute scale. For this technique the choice of the reference replication is theoretically immaterial and is typically chosen to be either the first or last observed replication for convenience. Thus the form of $\underset{\sim}{\Lambda}^j$ from (2.24a) becomes $\underset{\sim}{\Lambda}^{j'} = [1 \quad \lambda_2 \quad \ldots \quad \lambda_m]$ where λ_1, the scale factor associated with the first observed replication is arbitrarily set to 1, i.e., $\lambda_1 = 1$.

To produce the general model in matrix terms, which is the extension of (2.22) let any true predictor T^j $(j = 1,2,\ldots,p)$ have K_j observed replications where $K_j \geq 1$. That is, there are K_j observed predictors associated with true predictor T^j.

measureme

T

(1.25b)

with j =

where X

T

(

and

be writt

(1.26)

where

Let $V = \sum_{j=1}^{p} K_j$ that is, V is the total number of observed measurements associated with the p true predictors.

The model for a single observed measurement can be written

$$X_i^j = \lambda_i^j T^j + e_i^j \qquad (2.25b)$$

with $j = 1,2,\ldots,p$ and $i = 1,2,\ldots,K_j$

where X_i^j is the i^{th} replication of true predictor T^j,

T^j is the j^{th} true predictor,

e_i^j is the error for the i^{th} replication of predictor j,

and λ_i^j is the scaling factor which relates the true predictor T^j to the i^{th} observed replication of T^j. (Note: $\lambda_1^j = 1$ for every j).

In matrix terms the general model for the p predictors can be written as

$$\underset{V\times 1}{\underset{\sim}{X}} = \underset{V\times p}{\Lambda}\ \underset{p\times 1}{\underset{\sim}{T}} + \underset{V\times 1}{\underset{\sim}{E}} \qquad (2.26)$$

where

$$\underset{1\times V}{\underset{\sim}{X}'} = [X_1^1\ X_2^1 \ldots X_{k_1}^1 \mid X_1^2\ X_2^2 \ldots X_{k_2}^2 \mid \ldots \mid X_1^p\ X_2^p \ldots X_{k_p}^p]\ ,$$

$$\underset{1\times p}{\underset{\sim}{T}'} = [T^1\ T^2 \ldots T^p],$$

$$\underset{V\times p}{\Lambda} = \left[\begin{array}{ccccc}
1 & 0 & 0 & \cdots & 0 \\
\lambda_2^1 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
\lambda_{k_1}^1 & 0 & 0 & \cdots & 0 \\
\hline
0 & 1 & 0 & \cdots & 0 \\
0 & \lambda_2^2 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
0 & \lambda_{k_2}^2 & 0 & \cdots & 0 \\
\hline
 & & \vdots & & \\
\hline
0 & 0 & 0 & \cdots & 1 \\
0 & 0 & 0 & \cdots & \lambda_2^p \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & \lambda_{k_p}^p
\end{array}\right] \quad \text{and} \quad \underset{V\times 1}{\underset{\sim}{E}} = \left[\begin{array}{c}
e_1^1 \\
e_2^1 \\
\vdots \\
e_{k_1}^1 \\
\hline
e_1^2 \\
e_2^2 \\
\vdots \\
e_{k_2}^2 \\
\hline
\vdots \\
\hline
e_1^p \\
e_2^p \\
\vdots \\
e_{k_p}^p
\end{array}\right]$$

The assumptions about elements of the basic model (2.23a), (2.23b) and (2.23c) all apply to the extended model with the appropriate adjustments in notation to accommodate the increased number of observed and error parameters.

Therefore, results comparable to (2.24a) and (2.24b) can be produced.

(2.27a) $$\underset{V\times 1}{\underset{\sim}{\mu}}{}_X^{(k)} = \underset{V\times p}{\underset{\sim}{\Lambda}} \; \underset{p\times 1}{\underset{\sim}{\mu}}{}_T^{(k)}$$

for any category k $(k = 1,2,\ldots,J)$.

To demonstrate this, consider any category k (k = 1,2,...,J)

$$\underset{\sim}{\mu}_X^{(k)} \equiv E_k(\underset{\sim}{X}) = E_k(\Lambda\underset{\sim}{T} + \underset{\sim}{E}) = E_k(\Lambda\underset{\sim}{T}) + E_k(\underset{\sim}{E})$$

$$= E_k(\Lambda\underset{\sim}{T}) = \Lambda E_k(\underset{\sim}{T}) = \Lambda\underset{\sim}{\mu}_T^{(k)}.$$

(2.27b) $$\underset{V\times V}{\Sigma} = \underset{V\times p}{\Lambda}\ \underset{p\times p}{\Phi}\ \underset{p\times V}{\Lambda'} + \underset{V\times V}{\Psi^2}$$

for any cateogry k (k = 1,2,...,J).

To demonstrate this, consider any category k (k = 1,2,...,J)

$$\Sigma \equiv \mathrm{Var}(\underset{\sim}{X}) = \mathrm{Var}(\Lambda\underset{\sim}{T} + \underset{\sim}{E})$$

$$= E_k(\Lambda\underset{\sim}{T} + \underset{\sim}{E})(\Lambda\underset{\sim}{T} + \underset{\sim}{E})' - E_k(\Lambda\underset{\sim}{T} + \underset{\sim}{E})E_k(\Lambda\underset{\sim}{T} + \underset{\sim}{E})'$$

$$= E_k(\Lambda\underset{\sim\sim}{TT'}\Lambda' + \underset{\sim\sim}{ET'}\Lambda' + \Lambda\underset{\sim\sim}{TE'} + \underset{\sim\sim}{EE'}) - E_k(\Lambda\underset{\sim}{T})E_k(\Lambda\underset{\sim}{T})'$$

$$= \Lambda[E_k(\underset{\sim\sim}{TT'})]\Lambda' - \Lambda[E_k(\underset{\sim}{T})E_k(\underset{\sim}{T}')]\Lambda' + E_k(\underset{\sim\sim}{EE'})$$

$$+ [E_k(\underset{\sim\sim}{ET'})]\Lambda' + \Lambda[E_k(\underset{\sim\sim}{TE'})]$$

$$= \Lambda[E_k(\underset{\sim\sim}{TT'}) - E_k(\underset{\sim}{T})E_k(\underset{\sim}{T}')]\Lambda' + E_k(\underset{\sim\sim}{EE'}) + \underset{V\times V}{[0]} + \underset{V\times V}{[0]}$$

$$= \Lambda[\mathrm{Var}(\underset{\sim}{T})]\Lambda' + \mathrm{Var}(\underset{\sim}{E})$$

$$\Sigma = \Lambda\Phi\Lambda' + \Psi^2$$

where

$$\underset{V\times V}{\Psi^2} = \mathrm{diag}\left\{\sigma^2_{e_1^1}\ \sigma^2_{e_2^1} \cdots \sigma^2_{e_{k_1}^1} \,\Big|\, \sigma^2_{e_1^2}\ \sigma^2_{e_2^2} \cdots \sigma^2_{e_{k_2}^2} \,\Big|\, \cdots \,\Big|\, \sigma^2_{e_1^p} \cdots \sigma^2_{e_{k_p}^p}\right\}.$$

For both of these derivations Λ is a matrix of constants with respect to the expectation across the subjects in the population of category k (k = 1,2,...,J).

of t

impo

redu

mode

meas

only

meas

the w

suffi

Since some similar notation is used for matrices from each of the two measurement models which have different definitions it is important to clearly specify the measurement model being used. This redundancy of notation will pose no problem since each measurement model will be used in distinct and different situations. The basic measurement model, given by (2.22), will be used in Chapter 3 and only the early exploratory stages of Chapter 4, while the extended measurement model, given by (2.26) will be used for the majority of the work in Chapter 4 once the basic model has been shown to be insufficient for use in the estimation of latent weighting coefficients.

Section E: Summary

In this chapter the models which are needed for the research to be presented in the following chapters have been defined and developed.

The Observed Random Predictor Quantal Response Model which indicates the existence and the form of the observed weighting coefficients is given by (2.2).

The Latent Random Predictor Quantal Response Model which indicates the existence and the form of the latent weighting coefficients is given by (2.19).

The basic measurement model which relates the components of the two quantal response models is given by (2.22). Through this relationship the Area one research to be presented in Chapter 3, will examine the relative values of the individual observed and latent weighting coefficients associated with a predictor for the whole range of possible situations.

The Area two research, to be presented in Chapter 4, will demonstrate that the basic measurement model given by (2.22) is not sufficient to allow the estimation of the latent weighting coefficients. The research will demonstrate that a model in the form of the extended measurement model, given by (2.26), is needed.

In chapter 4 the measurement model will be used to give two reformulations of the Observed Random Predictor Quantal Response Model (2.2) in terms of parameters from the Latent Random Predictor Quantal Response Model (2.19) and parameters describing errors of measurement. Each of these reformulations will be examined to dete

clem

asso

determine whether estimates exist for the latent weighting coefficients. If estimates do exist then the estimation procedure associated with the reformulation will be described.

CHAPTER 3

Section A: Introduction and Approach to the Problem

For an analysis which involves the use of random predictor variables in a quantal response model to determine the relationship of interest between the predictor variables and the probability of occurrence of the categories of a qualitative criterion variable, the relationships of interest are given by the latent weighting coefficients. The vector of latent weighting coefficients identified in expression (2.19) from Chapter 2, provides the relationships between error-free predictors and the criterion. However, in practice, the measurements of the predictors contain errors of measurement. Thus information will be available for observed predictors and not the true predictors. Therefore, for a given situation the relationship which can be found using available data and available quantal response techniques is that provided by the observed weighting coefficients, identified in expression (2.2) from Chapter 2, although the relationship of interest is provided by the latent weighting coefficients (2.19).

The work to be presented in this chapter examines the effects of errors of measurement in the random predictors of a quantal response technique by examining the relationship between the observed weighting coefficients (from (2.2)) and the latent weighting coefficients (from

(2.19)) for the universe of situations. The precise information needed to identify a "situation" will be provided in Section C below. For situations where the observed weighting coefficients have nearly identical values to the latent weighting coefficients the effects of errors of measurement are considered to be small. In these situations the observed weighting coefficeints are acceptable estimators of the latent weighting coefficients. For other situations, however, the observed weighting coefficients may provide values which sufficiently overestimate or underestimate the values of the latent weighting coefficients so as to make the observed weighting coefficients poor estimators of the latent weighting coefficients.

The concern for this chapter then becomes how accurate an estimate of the latent weighting coefficient is provided by the observed weighting coefficient. To research this question, the ratio of the observed weighting coefficient to the latent weighting coefficient for a single predictor will be formed. Using the notation introduced in Chapter 2 for the single weighting coefficients this ratio becomes β_q/β_q^*, where $\beta_q = (\underset{\sim}{\beta}_{k\cdot j})_q$ and $\beta_q^* = (\underset{\sim}{\beta}_{k\cdot j}^*)_q$ represent the observed and latent weighting coefficients respectively for some predictor q $(q = 1,2,\ldots,p)$ from the observed and latent vectors of weighting coefficients associated with category k of the criterion $(k = 1,2,\ldots,J)$. The use of this ratio requires that $\beta_q^* \neq 0$. Situations where $\beta_q^* = 0$ will be examined separately.

The value of this ratio β_q/β_q^* will be examined for each predictor for generally applicable results as well as situation specific results. One of three categories of results are possible for a given

situation and a given predictor:

$\beta_q/\beta_q^* = 1$ i.e. the observed weighting coefficient has the same value as the latent weighting coefficient.

$\beta_q/\beta_q^* < 1$ i.e. the observed weighting coefficient has a value less than the latent weighting coefficient (an underestimate of the latent coefficient).

and $\beta_q/\beta_q^* > 1$ i.e. the observed weighting coefficient has a value greater than the latent weighting coefficient (an overestimate of the latent coefficeint).

The results which follow will consider β_q/β_q^* for each predictor q $(q = 1,2,\ldots,p)$ in each of the $J - 1$ vectors of weighting coefficients which comprise the base set.

In all cases the prime interest will be to identify those situations which correspond to each of the three categories of results identified above, with the prime interest on those situations where the observed weighting coefficient is equal to the latent weighting coefficient, i.e. $\beta_q/\beta_q^* = 1$.

The degree of error of measurement for an observed predictor is given by the reliability coefficient associated with that predictor. The reliability coefficient is the ratio of the true score variance of a predictor to the observed score variance, i.e. $\rho_{ii} = \sigma^2_{T^i}/\sigma^2_{X^i}$.

Since the search for generally applicable effects of errors of measurement for the most general case of the quantal response model with J categories $(J \geq 2)$ and p predictor $(p \geq 1)$ has proven fruitless, and since the algebraic manipulation required for the above approach has proven nearly intractable for more than two predictors

($p > 2$), the research to be presented in this chapter is based on two special cases of the general quantal response models. The two special cases of the general quantal response models to be examined are one predictor models (i.e., $p = 1$) in Section B and two predictor models (i.e., $p = 2$) in Section C. Most of the work on two predictor models will be done for models with a dichotomous criterion ($J = 2$). The two-predictor, polychotomous criterion model will be shown to represent a simple extension of the two-predictor, dichotomous criterion model.

Section B: One Predictor Models (p = 1)

The first case to be examined is that with a polychotomous criterion ($J \geq 2$) and one predictor ($p = 1$). For this case the Observed Random Predictor Quantal Response Model (2.2) becomes for some category k $(k = 1,2,\ldots,J)$

$$\text{(3.1)} \qquad \Pr\{Y = k|X\} = P_k = \frac{1}{1 + \sum\limits_{\substack{j=1 \\ j \neq k}}^{J} \exp\{-(\alpha_{k\cdot j} + \beta_{k\cdot j}X)\}}$$

where

$$\alpha_{k\cdot j} = -\text{Ln}\left(\frac{p_j}{p_k}\right) - \frac{1}{2}\left[\frac{(\mu_X^{(k)})^2 - (\mu_X^{(j)})^2}{\sigma_X^2}\right]$$

and

$$\beta_{k\cdot j} = (\mu_X^{(k)} - \mu_X^{(j)})/\sigma_X^2$$

with $\mu_X^{(i)}$ and σ_X^2, the mean and the variance, respectively, of the distribution of the single observed predictor X for category i.

For this special case the general latent predictor model (2.19) becomes for that same category k identified above:

$$\text{(3.2)} \qquad \Pr\{Y = k|T\} = P_k^* = \frac{1}{1 + \sum\limits_{\substack{j=1 \\ j \neq k}}^{J} \exp\{-(\alpha_{k\cdot j}^* + \beta_{k\cdot j}^* T)\}}$$

where

$$\alpha_{k\cdot j}^* = -\text{Ln}\left(\frac{p_j}{p_k}\right) - \frac{1}{2}\left[\frac{(\mu_T^{(k)})^2 - (\mu_T^{(j)})^2}{\sigma_T^2}\right]$$

and

$$\beta_{k\cdot j}^* = (\mu_T^{(k)} - \mu_T^{(j)})/\sigma_T^2$$

wit

dis

for

sinc

$\varepsilon_{k \cdot j}$

(3.3)

where

serve

the l

relial

closer

corres

observ

corres

reliab

that

that

with $\mu_T^{(i)}$ and σ_T^2, the mean and the variance, respectively, for the distribution of the single latent predictor T for category i.

Using the expression for $\beta_{k \cdot j}$ from (3.1) and the expression for $\beta^*_{k \cdot j}$ from (3.2) for some $j \neq k$, $j,k = 1,2,\ldots,J$

$$\frac{\beta_{k \cdot j}}{\beta^*_{k \cdot j}} = \frac{(\mu_X^{(k)} - \mu_X^{(j)})/\sigma_X^2}{(\mu_T^{(k)} - \mu_T^{(j)})/\sigma_T^2} = \frac{\sigma_T^2}{\sigma_X^2}$$

since $\mu_X^{(k)} = \mu_T^{(k)}$ and $\mu_X^{(j)} = \mu_T^{(j)}$ from (2.24a).

Thus for any category k and any weighting coefficients $\beta_{k \cdot j}$ and $\beta^*_{k \cdot j}$

$$\text{(3.3)} \qquad \frac{\beta_{k \cdot j}}{\beta^*_{k \cdot j}} = \frac{\sigma_T^2}{\sigma_X^2} = \rho_{XX} \quad (j \neq k,\ j.k = 1,2,\ldots,J)$$

where $\rho_{XX} = \sigma_T^2/\sigma_X^2$ is the reliability of the observed predictor X.

Therefore for all one predictor models the value of the observed weighting coefficient will be an underestimate of the value of the latent (true) weighting coefficient by a factor equal to the reliability of the predictor. The more reliable the predictor the closer the values of the observed weighting coefficient will be to the corresponding latent weighting coefficient. However, the values of the observed weighting coefficient will be identical to the values of the corresponding latent weighting coefficient only for a perfectly reliable predictor, i.e. $\rho_{XX} = 1$.

If $\beta^*_{k \cdot j} = 0$ for some $j \neq k$, $j,k = 1,2,\ldots,J$, this implies that $\mu_T^{(k)} - \mu_T^{(j)} = 0$ but then $\mu_T^{(k)} - \mu_T^{(j)} = \mu_X^{(k)} - \mu_X^{(j)}$ requires that $\mu_X^{(k)} - \mu_X^{(j)} = 0$ and thus $\beta_{k \cdot j} = 0$ and conversely.

Therefore in any one predictor model if $\beta_{k \cdot j} = 0$ for any $j \neq k$, $j,k = 1,2,\ldots,J$ then $\beta^*_{k \cdot j} = 0$ and the observed weighting coefficient provides an exact estimate of the latent weighting coefficient.

Section C: Two Predictor Models (p = 2)

The polychotomous criterion ($J \geq 2$) two-predictor ($p = 2$) models for both observed and latent predictors have the same appearance as the general case models given by (2.2) and (2.19). The specialization to two predictors is obvious only when the precise form of the vectors and matrices of the models are examined. The parallel structure of the two models (2.2) and (2.19) allows the identification of the two predictor case to proceed for each model simultaneously. The identification of the vectors and matrices from the Observed Random Predictor Quantal Response Model (2.2) will be presented below on the left with the corresponding vectors and matrices from the Latent Random Predictor Quantal Response Model (2.19) on the right.

The vectors of predictors become:

$$\underset{\sim}{X} = \begin{bmatrix} X^1 \\ X^2 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{T} = \begin{bmatrix} T^1 \\ T^2 \end{bmatrix} .$$

The vectors of predictor means for some category i ($i = 1,2,\ldots,J$) become:

$$\underset{\sim}{\mu}_X^{(i)} = \begin{bmatrix} \mu_{X^1}^{(i)} \\ \mu_{X^2}^{(i)} \end{bmatrix} \quad \text{and} \quad \underset{\sim}{\mu}_T^{(i)} = \begin{bmatrix} \mu_{T^1}^{(i)} \\ \mu_{T^2}^{(i)} \end{bmatrix} .$$

The matrices of predictor variances and covariances, assumed homogeneous across all categories become:

$$\Sigma = \begin{bmatrix} \sigma^2_{X^1} & \sigma_{X^1X^2} \\ \sigma_{X^2X^1} & \sigma^2_{X^2} \end{bmatrix} \quad \text{and} \quad \Phi = \begin{bmatrix} \sigma^2_{T^1} & \sigma_{T^1T^2} \\ \sigma_{T^2T^1} & \sigma^2_{T^2} \end{bmatrix}$$

And the vectors of weighting coefficeints for some category k are:

$$\underset{\sim}{\beta}_{k\cdot j} = \begin{bmatrix} \beta_{k\cdot j(X^1)} \\ \beta_{k\cdot j(X^2)} \end{bmatrix} \quad \text{and} \quad \underset{\sim}{\beta}^*_{k\cdot j} = \begin{bmatrix} \beta_{k\cdot j(T^1)} \\ \beta_{k\cdot j(T^2)} \end{bmatrix}$$

where $j \neq k$, $j,k = 1,2,\ldots,J$.

The approach to this speical case ($p = 2$) will proceed by first examining the simplest two predictor model. This simplest model involves two categories of the criterion ($J = 2$) and the two predictors. Results for more complex models involving more than two categories of the criterion ($J > 2$) and two predictors will be shown to be simple extensions of the results for the simplest two predictor model.

Two Category, Two Predictor Models ($J = 2$, $p = 2$)

I. Simplify the notation

In order to simplify the appearance of the algebra below several notational adjustments to the general models will be made. The two observed predictors will be denoted as x and y, i.e. $\underset{\sim}{X}' = [x\ y]$. The two latent predictors will be denoted as ξ and η, i.e. $\underset{\sim}{T}' = [\xi\ \eta]$ where $x = \xi + e_x$ and $y = \eta + e_y$, i.e. $\underset{\sim}{X} = \underset{\sim}{T} + \underset{\sim}{E}$. The two categories of the criterion will be identified by the numerals

0 and 1, rather than 1 and 2, so that the category identification is consistent with the notation used for the dichotomous criterion model in both McSweeney and Schmidt (1974) and Cornfield, Gordon and Smith (1960). Also let p_1 and p_o the unconditional probabilities of occurrence of category 1 and 0 respectively be $p_1 = p$ and $p_o = 1 - p_1 = q$.

Since for the dichotomous case $\Pr\{Y = 1|\underset{\sim}{X}\} + \Pr\{Y = 0|\underset{\sim}{X}\} = 1$ and $\Pr\{Y = 1|\underset{\sim}{T}\} + \Pr\{Y = 0|\underset{\sim}{T}\} = 1$, it will be sufficient to work with the expressions of the observed and latent predictor models associated with $\Pr\{Y = 1|\underset{\sim}{X}\}$ and $\Pr\{Y = 1|\underset{\sim}{T}\}$ respectively. Associated with each of these model expressions is a single weighting coefficient, $\underset{\sim}{\beta}_{1\cdot o}$ and $\underset{\sim}{\beta}^*_{1\cdot o}$ each with two components. Since the proofs in Appendices A.1 and A.2 indicated $\underset{\sim}{\beta}_{1\cdot o} = -\underset{\sim}{\beta}_{o\cdot 1}$ and $\underset{\sim}{\beta}^*_{1\cdot o} = -\underset{\sim}{\beta}^*_{o\cdot 1}$, that is for each model there is only one distinct weighting coefficient, let $\underset{\sim}{\beta} = \underset{\sim}{\beta}_{1\cdot o}$ and $\underset{\sim}{\beta}^* = \underset{\sim}{\beta}^*_{1\cdot o}$.

Therefore the dichotomous criterion (J = 2), two predictor (p = 2) observed predictor model can be expressed as:

$$(3.4) \qquad \Pr\{Y = 1|\underset{\sim}{X}\} = P_1 = \frac{1}{1 + \exp\{-(\alpha_{1\cdot o} + \underset{\sim}{\beta}'\underset{\sim}{X})\}}$$

where $\underset{\sim}{X} = \begin{bmatrix} x \\ y \end{bmatrix}$

$$\alpha_{1\cdot o} = -\mathrm{Ln}\left(\frac{q}{p}\right) - \frac{1}{2}\left[\underset{\sim}{\mu}_X^{(1)\prime}\Sigma^{-1}\underset{\sim}{\mu}_X^{(1)} - \underset{\sim}{\mu}_X^{(o)\prime}\Sigma^{-1}\underset{\sim}{\mu}_X^{(o)}\right]$$

and

$$\underset{\sim}{\beta} = \begin{bmatrix} \beta_x \\ \beta_y \end{bmatrix} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(1)} - \underset{\sim}{\mu}_X^{(o)}),$$

and the latent predictor model can be expressed as:

$$\text{(3.5)} \quad \Pr\{Y = 1 | \underset{\sim}{T}\} = P_1^* = \frac{1}{1 + \exp\{-(\alpha_{1\cdot o}^* + \underset{\sim}{\beta}^{*'}\underset{\sim}{T})\}}$$

where $\underset{\sim}{T} = \begin{bmatrix} \xi \\ \eta \end{bmatrix}$

$$\alpha_{1\cdot o}^* = -\text{Ln}\left(\frac{q}{p}\right) - \frac{1}{2}\left[\underset{\sim}{\mu}_T^{(1)'}\Phi^{-1}\underset{\sim}{\mu}_T^{(1)} - \underset{\sim}{\mu}_T^{(o)'}\Phi^{-1}\underset{\sim}{\mu}_T^{(o)}\right]$$

and

$$\underset{\sim}{\beta}^* = \begin{bmatrix} \beta_\xi^* \\ \beta_\eta^* \end{bmatrix} = \Phi^{-1}(\underset{\sim}{\mu}_T^{(1)} - \underset{\sim}{\mu}_T^{(o)}).$$

II. <u>Derive Expressions for β_x/β_ξ^* and β_y/β_η^* in terms of latent and error parameters</u>

In order to study the effects of errors of measurement on the weighting coefficients the ratios β_x/β_ξ^* and β_y/β_η^* will be examined. The formulas presented previously for single predictor weighting coefficients will prove of limited usefulness for this task, therefore the first necessity will be to derive expressions and conditions for existence for β_x, β_y, β_ξ^* and β_η^*. From these expressions the desired ratios can be formed.

Consider first

$$\underset{\sim}{\beta}^* = \begin{bmatrix} \beta_\xi^* \\ \beta_\eta^* \end{bmatrix} = \Phi^{-1}(\underset{\sim}{\mu}_T^{(1)} - \underset{\sim}{\mu}_T^{(o)})$$

$$\Phi = \begin{bmatrix} \sigma_\xi^2 & \sigma_{\xi\eta} \\ \sigma_{\eta\xi} & \sigma_\eta^2 \end{bmatrix} \quad \text{and} \quad \Phi^{-1} = \frac{1}{1-\rho_{\xi\eta}^2}\begin{bmatrix} \frac{1}{\sigma_\xi^2} & \frac{-\rho_{\xi\eta}}{\sigma_\xi\sigma_\eta} \\ \frac{-\rho_{\xi\eta}}{\sigma_\xi\sigma_\eta} & \frac{1}{\sigma_\eta^2} \end{bmatrix}$$

where $\rho_{\xi\eta}$ is the correlation between the latent predictors ξ and η. Note: Φ^{-1} exist only if $1 - \rho_{\xi\eta}^2 \neq 0$ which implies $\rho_{\xi\eta} \neq \pm 1$.

Let $\underset{\sim}{a}_T = \begin{bmatrix} a_\xi \\ a_\eta \end{bmatrix} = \underset{\sim}{\mu}_T^{(1)} - \underset{\sim}{\mu}_T^{(o)} = \begin{bmatrix} \mu_\xi^{(1)} - \mu_\xi^{(o)} \\ \mu_\eta^{(1)} - \eta_\eta^{(o)} \end{bmatrix}$.

Therefore

$$\underset{\sim}{\beta}^* = \begin{bmatrix} \beta_\xi^* \\ \beta_\eta^* \end{bmatrix} = \Phi^{-1}\underset{\sim}{a}_T = \frac{1}{1 - \rho_{\xi\eta}^2} \begin{bmatrix} \frac{1}{\sigma_\xi^2} & \frac{-\rho_{\xi\eta}}{\sigma_\xi\sigma_\eta} \\ \frac{-\rho_{\xi\eta}}{\sigma_\xi\sigma_\eta} & \frac{1}{\sigma_\eta^2} \end{bmatrix} \begin{bmatrix} a_\xi \\ a_\eta \end{bmatrix}$$

$$\underset{\sim}{\beta}^* = \frac{1}{1 - \rho_{\xi\eta}^2} \begin{bmatrix} \frac{a_\xi}{\sigma_\xi^2} - \frac{a_\eta\,\rho_{\xi\eta}}{\sigma_\xi\sigma_\eta} \\ \frac{a_\eta}{\sigma_\eta^2} - \frac{a_\xi\,\rho_{\xi\eta}}{\sigma_\xi\sigma_\eta} \end{bmatrix}$$

and

$$\text{(3.6a)} \qquad \beta_\xi^* = \frac{1}{1 - \rho_{\xi\eta}^2}\left[\frac{a_\xi}{\sigma_\xi^2} - \frac{a_\eta\,\rho_{\xi\eta}}{\sigma_\xi\sigma_\eta}\right]$$

$$= \frac{1}{1 - \rho_{\xi\eta}^2} \quad \frac{a_\xi}{\sigma_\xi^2}\left[1 - \frac{a_\eta\,\sigma_\xi\,\rho_{\xi\eta}}{\sigma_\eta a_\xi}\right] \text{ for } a_\xi \neq 0.$$

Thus

$$\text{(3.6b)} \qquad \beta_\xi^* = \frac{b_\xi^*}{1 - \rho_{\xi\eta}^2}\left[1 - d_\xi\,\rho_{\xi\eta}\right] \quad \text{where} \quad b_\xi^* = \frac{a_\xi}{\sigma_\xi^2} = \frac{\mu_\xi^{(1)} - \mu_\eta^{(o)}}{\sigma_\xi^2}$$

$$\text{and} \qquad d_\xi = \frac{a_\eta/\sigma_\eta}{a_\xi/\sigma_\xi} = \frac{(\mu_\eta^{(1)} - \mu_\eta^{(o)})/\sigma_\eta}{(\mu_\xi^{(1)} - \mu_\xi^{(o)})/\sigma_\xi} \quad .$$

In this formulation $b_\xi^* = (\mu_\xi^{(1)} - \mu_\xi^{(o)})/\sigma_\xi^2$ has the form of a latent weighting coefficient for ξ from a single predictor model (see (3.2) for an example). In this formulation d_ξ represents the ratio of category mean differences for the two latent predictors, where each mean difference is in standard units (i.e. divided by the standard deviation of the distribution of the latent predictors). Therefore, a large positive d_ξ value indicates a larger standard unit mean difference between category 1 and category 0 for latent predictor η than for latent predictor ξ. Other values of d_ξ would carry appropriate corresponding interpretations in terms of ratios of standard unit mean differences between categories.

And

$$\text{(3.7a)} \qquad \beta^*_\eta = \frac{1}{1-\rho^2_{\xi\eta}} \left[\frac{a_\eta}{\sigma^2_\eta} - \frac{a_\xi \, \rho_{\xi\eta}}{\sigma_\xi \sigma_\eta} \right]$$

$$= \frac{1}{1-\rho^2_{\xi\eta}} \cdot \frac{a_\eta}{\sigma^2_\eta} \left[1 - \frac{a_\xi \, \sigma_\eta \, \rho_{\xi\eta}}{\sigma_\xi \, a_\eta} \right] \quad \text{for } a_\eta \neq 0.$$

$$\text{(3.7b)} \qquad \beta^*_\eta = \frac{b^*_\eta}{1-\rho^2_{\xi\eta}} \left[1 - d_\eta \, \rho_{\xi\eta} \right]$$

where $b^*_\eta = \dfrac{a_\eta}{\sigma^2_\eta} = \dfrac{\mu^{(1)}_\eta - \mu^{(o)}_\eta}{\sigma^2_\eta}$ and $d_\eta = \dfrac{a_\xi/\sigma_\xi}{a_\eta/\sigma_\eta}$. Since $d_\xi = \dfrac{a_\eta/\sigma_\eta}{a_\xi/\sigma_\xi}$ as defined above for β^*_ξ, $d_\eta = \dfrac{1}{d_\xi}$ for $d_\xi \neq 0$.

Consider now the expressions for β_x and β_y from the observed predictor model.

$$\beta = \begin{bmatrix} \beta_x \\ \beta_y \end{bmatrix} = \Sigma^{-1}(\mu^{(1)}_X - \mu^{(o)}_X)$$

$$\Sigma = \begin{bmatrix} \sigma^2_x & \sigma_{xy} \\ \sigma_{yx} & \sigma^2_y \end{bmatrix} \quad \text{and} \quad \Sigma^{-1} = \frac{1}{1-\rho^2_{xy}} \begin{bmatrix} \dfrac{1}{\sigma^2_x} & \dfrac{-\rho_{xy}}{\sigma_x\sigma_y} \\ \dfrac{-\rho_{xy}}{\sigma_x\sigma_y} & \dfrac{1}{\sigma^2_y} \end{bmatrix}$$

where ρ_{xy} is the correlation between the observed predictors x and y.

Note: Σ^{-1} exists only if $1 - \rho_{xy}^2 \neq 0$ which implies $\rho_{xy} \neq \pm 1$.

Let $$\underset{\sim}{a}_X = \begin{bmatrix} a_x \\ a_y \end{bmatrix} = \underset{\sim}{\mu}_X^{(1)} - \underset{\sim}{\mu}_X^{(o)} = \begin{bmatrix} \mu_x^{(1)} - \mu_x^{(o)} \\ \mu_y^{(1)} - \mu_y^{(o)} \end{bmatrix} .$$

Thus $$\underset{\sim}{\beta} = \begin{bmatrix} \beta_x \\ \beta_y \end{bmatrix} = \Sigma^{-1} \underset{\sim}{a}_X = \frac{1}{1-\rho_{xy}^2} \begin{bmatrix} \frac{1}{\sigma_x^2} & \frac{-\rho_{xy}}{\sigma_x\sigma_y} \\ \frac{-\rho_{xy}}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix}$$

$$\underset{\sim}{\beta} = \frac{1}{1-\rho_{xy}^2} \begin{bmatrix} \frac{a_x}{\sigma_x^2} - \frac{a_y\,\rho_{xy}}{\sigma_x\sigma_y} \\ \frac{a_y}{\sigma_y^2} - \frac{a_x\,\rho_{xy}}{\sigma_x\sigma_y} \end{bmatrix}$$

Before proceeding further with finding expressions for β_x and β_y from the observed predictor model consider some of the important relationships between expressions involving observed predictor model parameters and the corresponding expressions involving latent predictor model parameters.

(3.8) $\underset{\sim}{a}_X = \underset{\sim}{\mu}_X^{(1)} - \underset{\sim}{\mu}_X^{(o)} = \underset{\sim}{\mu}_T^{(1)} - \underset{\sim}{\mu}_T^{(o)} = \underset{\sim}{a}_T$ therefore $a_x = a_\xi$ and $a_y = a_\eta$.

(3.9a) $\rho_{xx} = \sigma_\xi^2/\sigma_x^2$ therefore $\sigma_x^2 = \sigma_\xi^2/\rho_{xx}$

(3.9b) $\rho_{yy} = \sigma_\eta^2/\sigma_y^2$ therefore $\sigma_y^2 = \sigma_\eta^2/\rho_{yy}$

(3.10) $$\rho_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 \sigma_y^2}} = \frac{\sigma_{\xi\eta}}{\sqrt{\frac{\sigma_\xi^2}{\rho_{xx}} \cdot \frac{\sigma_\eta^2}{\rho_{yy}}}} = \frac{\sigma_{\xi\eta}}{\sqrt{\sigma_\xi^2 \sigma_\eta}} \sqrt{\rho_{xx}\rho_{yy}} = \rho_{\xi\eta} \sqrt{\rho_{xx}\rho_{yy}}$$

i.e. $\rho_{xy} = \rho_{\xi\eta} \sqrt{\rho_{xx}\rho_{yy}}$.

Since the ratios β_x/β_ξ^* and β_y/β_η^* are the ultimate expressions of interest for the research which is to follow, expressions (3.8) - (3.10) above will be used to express β_x and β_y in terms of parameters from the latent predictor model.

The expression for β_x using parameters from the observed predictor model is:

(3.11a)
$$\beta_x = \frac{1}{1 - \rho_{xy}^2} \left[\frac{a_x}{\sigma_x^2} - \frac{a_y \rho_{xy}}{\sigma_x \sigma_y} \right]$$

$$\beta_x = \frac{1}{1 - \rho_{xy}^2} \cdot \frac{a_x}{\sigma_x^2} \left[1 - \frac{a_y \sigma_x \sigma_{xy}}{\sigma_y a_x} \right] \quad \text{for} \quad a_x \neq 0 .$$

And using expressions (3.8) - (3.10) β_x becomes:

$$\beta_x = \frac{1}{1 - \rho_{\xi\eta}^2 \rho_{xx} \rho_{yy}} \cdot \frac{a_\xi}{\sigma_\xi^2/\rho_{xx}} \cdot \left[1 - \frac{a_\eta \cdot \sigma_\xi/\sqrt{\rho_{xx}} \cdot \rho_{\xi\eta}\sqrt{\rho_{xx}\rho_{yy}}}{\sigma_\eta/\sqrt{\rho_{yy}} \cdot a_\xi} \right]$$

(3.11b) $$\beta_x = \frac{b_\xi^* \rho_{xx}}{1 - \rho_{\xi\eta}^2 \rho_{xx}\rho_{yy}} \left[1 - d_\xi \rho_{\xi\eta} \rho_{yy} \right]$$

where $$b_\xi^* = \frac{a_\xi}{\sigma_\xi^2} = \frac{\mu_\xi^{(1)} - \mu_\xi^{(o)}}{\sigma_\xi^2}$$

and $$d_\xi = \frac{a_\eta/\sigma_\eta}{a_\xi/\sigma_\xi} = \frac{(\mu_\eta^{(1)} - \mu_\eta^{(o)})/\sigma_\eta}{(\mu_\xi^{(1)} - \mu_\xi^{(o)})/\sigma_\xi} .$$

Therefore using the ratio of (3.11b) and (3.6b), β_x/β_ξ^* will exist if $\beta_\xi^* \neq 0$, and

$$\beta_x/\beta_\xi^* = \frac{b_\xi^*\rho_{xx}}{1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}}\left[1 - d_\xi\rho_{\xi\eta}\rho_{yy}\right] \Big/ \frac{b_\xi^*}{1 - \rho_{\xi\eta}^2}\left[1 - d_\xi\rho_{\xi\eta}\right]$$

(3.12)

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d_\xi\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d_\xi\rho_{\xi\eta})} .$$

This expression for β_x/β_ξ^* (3.12) will exist if:

1.) $\rho_{\xi\eta} \neq \pm 1$ (Needed for Φ^{-1} to exist.)

Note: $\rho_{xy} = \rho_{\xi\eta}\sqrt{\rho_{xx}\rho_{yy}}$ therefore $|\rho_{xy}| \leq |\rho_{\xi\eta}|$ therefore if $\rho_{\xi\eta} \neq \pm 1$ then $\rho_{xy} \neq \pm 1$ and Σ^{-1} will exist.

2.) $a_\xi \neq 0$

That is, $a_\xi = \mu_\xi^{(1)} - \mu_\xi^{(o)} \neq 0 \Rightarrow \mu_\xi^{(1)} \neq \mu_\xi^{(o)}$. This is needed only to guarantee the existence of the specific formulation for β_x/β_ξ^* being used. A variation of the expression β_x/β_ξ^* for $a_\xi = 0$ will be examined below.

3.) $1 - d_\xi\rho_{\xi\eta} \neq 0 \Rightarrow d_\xi\rho_{\xi\eta} \neq 1$

This is needed to guarantee that $\beta_\xi^* \neq 0$. The second requirement, $a_\xi^* \neq 0$, guarantees that $b_\xi^* \neq 0$ thus $\beta_\xi^* \neq 0$. When $\beta_\xi^* = 0$, β_x will be examined briefly below.

Expressions for β_y can also be produced using arguments similar to those for β_x above:

The expression for β_y using parameters from the observed predictor model is:

(3.13a) $$\beta_y = \frac{1}{1 - \rho_{xy}^2} \cdot \frac{a_y}{\sigma_y^2} \cdot \left[1 - \frac{a_x \sigma_y \rho_{xy}}{\sigma_x a_y} \right] \quad \text{for } a_y \neq 0.$$

And using expressions (3.8) - (3.10) β_y becomes:

(3.13b) $$\beta_y = \frac{b_\eta^* \rho_{yy}}{1 - \rho_{\xi\eta}^2 \rho_{xx} \rho_{yy}} \left[1 - d_\eta \rho_{\xi\eta} \rho_{xx} \right]$$

where $b_\eta^* = \frac{a_\eta}{\sigma_\eta^2} = \frac{\mu_\eta^{(1)} - \mu_\eta^{(1)}}{\sigma_\eta^2}$ and $d_\eta = \frac{a_\xi / \sigma_\xi}{a_\eta / \sigma_\eta}$ with $d_\eta = \frac{1}{d_\xi}$ for $d_\xi \neq 0$.

Therefore, using the ratio of (3.13b) and (3.7b), β_y / β_η^* will exist if $\beta_\eta^* \neq 0$, and

(3.14) $$\beta_y / \beta_\eta^* = \frac{(1 - \rho_{\xi\eta}^2) \rho_{yy} (1 - d_\eta \rho_{\xi\eta} \rho_{xx})}{(1 - \rho_{\xi\eta}^2 \rho_{xx} \rho_{yy})(1 - d_\eta \rho_{\xi\eta})} .$$

This expression for β_y / β_η^* (3.14) will exist if:

1.) $\rho_{\xi\eta} \neq \pm 1$ (Needed for Φ^{-1} to exist. Also, see Note with condition 1 for existence of β_x / β_ξ^* above.)

2.) $a_\eta \neq 0$

That is, $a_\eta = \mu_\eta^{(1)} - \mu_\eta^{(o)} \neq 0 \Rightarrow \mu_\eta^{(1)} \neq \mu_\eta^{(o)}$. This is needed only to guarantee the existence of the specific formulation for β_y / β_η^* being used. A variation of the expression of β_y / β_η^* will be examined for $a_\eta = 0$.

Not

for

the

clo

str

the

ε_x/

the

res

whi

of

is

pro

giv

val

(i.e

3.) $1 - d_{\eta}\rho_{\xi\eta} \neq 0 \Rightarrow d_{\eta}\rho_{\xi\eta} \neq 1$

This along with the second requirement above is needed to guarantee that $\beta_{\eta}^{*} \neq 0$.

Note since $\sigma_{\xi}^{2} > 0$ then $\rho_{xx} > 0$, and $\sigma_{\eta}^{2} > 0$ then $\rho_{yy} > 0$. Therefore the effective ranges for ρ_{xx}, ρ_{yy} and $\rho_{\xi\eta}$ are:

$$0 < \rho_{xx} \leq 1$$

$$0 < \rho_{yy} \leq 1$$

$-1 < \rho_{\xi\eta} < +1$ (with one possible exception for one of the ratios depending upon the value of d_{ξ} or d_{η}).

Expressions (3.12) for β_x/β_{ξ}^{*} and (3.14) for β_y/β_{η}^{*} will be the primary expressions of interest for the work below. However, a close examination of expressions (3.12) and (3.14) shows an identical structure for each expression. Because of this identical structure, the expression for β_y/β_{η}^{*} (3.14) can be found from the expression for β_x/β_{ξ}^{*} (3.12) by merely interchanging the x's and y's as well as the ξ's and η's in the notation for β_x/β_{ξ}^{*}. Therefore any algebraic result derived for β_x/β_{ξ}^{*} will have a corresponding result for β_y/β_{η}^{*} which can be simply stated, rather than derived, using this property of interchangeability of x and y (and ξ and η as well). It is important to note that the values of $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_{ξ} which produce a given value of β_x/β_{ξ}^{*} say R (i.e., $\beta_x/\beta_{\xi}^{*} = R$ for the given $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d) will not in general also produce a value of R for β_y/β_{η}^{*}. That is, in general, for a given situation (i.e., a specific set of values for $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_{ξ}) the values of β_x/β_ξ^* and β_y/β_η^* will not be identical. However, by the use of the property of interchangeability of x and y it is possible to identify a different situation (i.e., different values for $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_ξ) where $\beta_y/\beta_\eta^* = R$. If we let $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_ξ represent the situation where $\beta_x/\beta_\xi^* = R$ and $\rho'_{\xi\eta}$, ρ'_{xx}, ρ'_{yy} and d'_ξ represent the generally different situation where $\beta_y/\beta_\eta^* = R$ then Appendix B.1 demonstrates that the two situations have the following relationship:

(3.15a) $$\rho'_{\xi\eta} = \rho_{\xi\eta}$$

(3.15b) $$\rho'_{xx} = \rho_{yy}$$

(3.15c) $$\rho'_{yy} = \rho_{xx}$$

(3.15d) $$d'_\eta = d_\xi \ .$$

Therefore, it will be necessary to examine only β_x/β_ξ^* in detail across the universe of situations (i.e., values of $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_ξ). Corresponding results for β_y/β_η^* can be obtained through the use of expressions (3.15a) - (3.15d). The prime in the notation will rarely be used unless the interchanging of x's and y's becomes ambiguous without its use.

Since only values of β_x/β_ξ^* need be examined in detail and since $d_\xi = 1/d_\eta$ let $d = d_\xi = 1/d_\eta$ be used where no ambiguity will result. Appendix B.1 also demonstrates that only values of $d \geq 0$ need be considered in the examination of β_x/β_ξ^*. The expression for β_x/β_ξ^* with $d < 0$ is the reflection through the line $\rho_{\xi\eta} = 0$ of the expression for β_x/β_ξ^* with $|d| > 0$. That is, for $d < 0$, there will exist values d'' $(d'' > 0)$ and $\rho''_{\xi\eta}$ which produce the same value of β_x/β^*_ξ as d and $\rho_{\xi\eta}$. These values d'' and $\rho''_{\xi\eta}$ have the following relationship to d and $\rho_{\xi\eta}$. (See Appendix B.1 for details.)

(3.16a) $\quad d'' = -d \quad$ (since $d < 0$, $d'' > 0$), and

(3.16b) $\quad \rho''_{\xi\eta} = -\rho_{\xi\eta}$.

Thus in the work which follows only β_x/β^*_ξ for $d \geq 0$ will be examined since results for other situations for β_x/β^*_ξ and all situations for β_y/β^*_η can then be derived using expressions (3.15a) - (3.15d) or (3.16a) and (3.16b). Results for special case situations, i.e., $d_\eta = 0$ (i.e., $d = d_\xi$ undefined), $d = d_\xi = 0$, $\rho_{\xi\eta} = 0$, $\rho_{xx} = \rho_{yy} = 1$, $\rho_{xx} = 1$ and $\rho_{yy} < 1$, and $\rho_{xx} < 1$ and $\rho_{yy} = 1$, are presented in Appendix B.2.

III. <u>Presentation of the Approach to the Examination of</u> β_x/β^*_ξ

As indicated in section A of this chapter the interest for this area of the research is to determine for what situations β_x is an overestimate, an exact estimate or an underestimate of β^*_ξ. To pursue this question comparison of the ratio β_x/β^*_ξ to 1 is to be examined. There is an algebraic expression which will aid in this examination. Let

(3.17) $$Q = \rho^2_{\xi\eta}\rho_{xx}(1 - \rho_{yy}) - d\rho_{\xi\eta}(1 - \rho_{xx}\rho_{yy}) + (1 - \rho_{xx})$$

for $0 < \rho_{yy} < 1$, $0 < \rho_{xx} \leq 1$, $-1 < \rho_{\xi\eta} < +1$ and any d such that $d\rho_{\xi\eta} \neq 1$.

To see how Q can aid in the search for relationships between β_x/β^*_ξ and one consider:

a) $$\frac{\beta_x}{\beta_\xi^*} = 1 \Leftrightarrow \frac{\beta_x}{\beta_\xi^*} = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})} = 1$$

$$\Leftrightarrow (1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy}) = (1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})$$

$$\Leftrightarrow \rho_{xx} - \rho_{\xi\eta}^2\rho_{xx} - d\rho_{\xi\eta}\rho_{xx}\rho_{yy} + d\rho_{\xi\eta}^3\rho_{xx}\rho_{yy}$$

$$= 1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy} - d\rho_{\xi\eta} + d\rho_{\xi\eta}^3\rho_{xx}\rho_{yy}$$

$$\Leftrightarrow 0 = \rho_{\xi\eta}^2\rho_{xx} - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy} - d\rho_{\xi\eta} + d\rho_{\xi\eta}\rho_{xx}\rho_{yy} + 1 - \rho_{xx}$$

$$\Leftrightarrow 0 = \rho_{\xi\eta}^2\rho_{xx}(1 - \rho_{yy}) - d\rho_{\xi\eta}(1 - \rho_{xx}\rho_{yy}) + (1 - \rho_{xx}).$$

Therefore

$$\frac{\beta_x}{\beta_\xi^*} = 1 \Leftrightarrow 0 = Q .$$

b) $$\frac{\beta_x}{\beta_\xi^*} < 1 \Leftrightarrow \frac{\beta_x}{\beta_\xi^*} = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})} < 1$$

1) if $1 - d\rho_{\xi\eta} > 0 \Leftrightarrow d\rho_{\xi\eta} < 1$ and since $0 < 1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy} \leq 1$

then

$$\frac{\beta_x}{\beta_\xi^*} < 1 \Leftrightarrow (1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})$$

$$< (1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta}).$$

Using algebra from a) above with appropriate attention for the inequality yields:

$$\frac{\beta_x}{\beta_\xi^*} < 1 \Leftrightarrow 0 < Q.$$

2) if $1 - d\rho_{\xi\eta} < 0 \Leftrightarrow d\rho_{\xi\eta} > 1$ then

$$\frac{\beta_x}{\beta_\xi^*} < 1 \Leftrightarrow (1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})$$

$$> (1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta}).$$

Using algebra from a) above with appropriate attention for the inequality yields:

$$\frac{\beta_x}{\beta_\xi^*} < 1 \Leftrightarrow 0 > Q .$$

c) $$\frac{\beta_x}{\beta_\xi^*} > 1 \Leftrightarrow \frac{\beta_x}{\beta_\xi^*} = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})} > 1$$

1) If $1 - d\rho_{\xi\eta} > 0 \Leftrightarrow d\rho_{\xi\eta} < 1$ and since

$0 < 1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy} \leq 1$

$$\frac{\beta_x}{\beta_\xi^*} > 1 \Leftrightarrow (1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})$$

$$> (1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta}) .$$

Using algebra from a) above with appropriate attention for the inequality yields:

$$\frac{\beta_x}{\beta_\xi^*} > 1 \Leftrightarrow 0 > Q.$$

2) if $1 - d\rho_{\xi\eta} < 0 \Leftrightarrow d\rho_{\xi\eta} > 1$

$$\frac{\beta_x}{\beta_\xi^*} > 1 \Leftrightarrow (1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})$$

$$< (1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta}).$$

Using algebra from a) above with appropriate attention for inequality yields:

$$\frac{\beta_x}{\beta_\xi^*} > 1 \Leftrightarrow 0 < Q.$$

Therefore, combining results from a), b) and c) above produces:

If $d\rho_{\xi\eta} > 1$ then

(3.18a) $Q < 0 \Leftrightarrow \beta_x/\beta_\xi^* < 1$ i.e., β_x underestimates β_ξ^*,

(3.18b) $Q = 0 \Leftrightarrow \beta_x/\beta_\xi^* = 1$ i.e., β_x exactly estimates β_ξ^*,

(3.18c) $Q > 0 \Leftrightarrow \beta_x/\beta_\xi^* > 1$ i.e., β_x overestimates β_ξ^*.

If $d\rho_{\xi\eta} < 1$ then

(3.19a) $Q < 0 \Leftrightarrow \beta_x/\beta_\xi^* > 1$ i.e., β_x overestimates β_ξ^*,

(3.19b) $Q = 0 \Leftrightarrow \beta_x/\beta_\xi^* = 1$ i.e., β_x exactly estimates β_ξ^*,

(3.19c) $Q > 0 \Leftrightarrow \beta_x/\beta_\xi^* < 1$ i.e., β_x underestimates β_ξ^*.

If $d\rho_{\xi\eta} = 1$, then β_x/β_ξ^* is undefined since $\beta_\xi^* = 0$.

Note: It is also possible to consider $d\rho_{\xi\eta}$ as the ratio of two slopes. The numerator of the ratio represents the slope of the pooled within categories regression line of ξ on η. The denominator of the ratio represents the slope of the line joining the midpoints of the joint distributions of ξ and η between the two categories. For more information about this interpretation see Appendix B.9.

Thus the examination for relationships between β_x/β_ξ^* and one can be pursued by examining the relationship between Q and zero. The questions now are, for what values of $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d will Q be less than zero, equal to zero, and greater than zero. The approach to answering these questions will be to consider three of the four variables ($\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d) as fixed, thus Q can be considered solely as a function of the fourth non-fixed variable for the given combination of the three fixed variables.

Although any one of the four variables could be selected as the non-fixed variable, the most interesting and useful information has come from fixing ρ_{xx}, ρ_{yy} and d and examining Q (and hence β_x/β_ξ^* as well) as $\rho_{\xi\eta}$ varies from -1 to +1 for various combinations of ρ_{xx}, ρ_{yy} and d. Following this approach Q is clearly a quadratic function in $\rho_{\xi\eta}$.

IV. The Search for Categories of Distribution of β_x/β_ξ^* as a Function of $\rho_{\xi\eta}$

Consider expression (3.17) for Q as a function of $\rho_{\xi\eta}$ for fixed values of ρ_{xx}, ρ_{yy} and d;

$$(3.20)\quad Q = \rho_{\xi\eta}^2\rho_{xx}(1 - \rho_{yy}) - d\rho_{\xi\eta}(1 - \rho_{xx}\rho_{yy}) + (1 - \rho_{xx}).$$

Let $Q = a_x\rho_{\xi\eta}^2 + b_x\rho\xi\eta + c_x$ where $a_x = \rho_{xx}(1 - \rho_{yy})$, $b_x = -d(1 - \rho_{xx}\rho_{yy})$ and $c_x = (1 - \rho_{xx})$.

Expression (3.20) clearly illustrates that Q is a quadratic function of $\rho_{\xi\eta}$. As a quadratic function of $\rho_{\xi\eta}$, Q will possess two roots call them $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ which are defined as:

$$(3.21a)\quad \rho_{\xi\eta}^{-(x)} = \frac{d(1 - \rho_{xx}\rho_{yy}) - \sqrt{d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}}{2\rho_{xx}(1 - \rho_{yy})}$$

$$(3.21b)\quad \rho_{\xi\eta}^{+(x)} = \frac{d(1 - \rho_{xx}\rho_{yy}) + \sqrt{d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}}{2\rho_{xx}(1 - \rho_{yy})}.$$

Both roots, $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ will exist with $\rho_{\xi\eta}^{-(x)} \leq \rho_{\xi\eta}^{+(x)}$ if $|d| \geq \sqrt{\dfrac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}}$. (See Appendix B.4 for details.) Since the existence of the roots $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ as real numbers is important, the quantity on the right of the existence expression will be used frequently. As an abbreviation in notation, let

$$(3.22) \qquad \sqrt{x} \equiv \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} .$$

Here the square root sign indicates that the quantity involved is a square root and the x indicates that the expression is related to β_x/β_ξ^*.

Since $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ represent possible values for the correlation between the two true predictors, $\rho_{\xi\eta}$ where $-1 < \rho_{\xi\eta} < +1$, the existence of $\rho_{\xi\eta}^{-(x)}$ or $\rho_{\xi\eta}^{+(x)}$ in the interval from -1 to +1 is as important as their existence as real numbers.

Therefore, from Appendix B.4: $\rho_{\xi\eta}^{-(x)}$ will exist with $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$, for $0 < \rho_{yy} < 1$,

(3.23a) for $0 < \rho_{xx} \leq \dfrac{1}{2 - \rho_{yy}}$ if $1 < d$

(3.23b) for $\dfrac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ if $\sqrt{x} \leq d$

or if $-1 < d \leq -\sqrt{x}$, and $\rho_{\xi\eta}^{+(x)}$ will exist with $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$ for $0 < \rho_{yy} < 1$,

(3.24a) for $0 < \rho_{xx} \leq \dfrac{1}{2 - \rho_{yy}}$ if $d < -1$

(3.24b) for $\frac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ if $\sqrt{x} \leq d < 1$

or if $d \leq - \sqrt{x}$.

Note 1: When $\rho_{yy} = 1$, $0 < \beta_x/\beta_\xi^* \leq 1$ for all $\rho_{\xi\eta}$, ρ_{xx}, d with equality only when $\rho_{xx} = 1$ too. (See Appendix B.2 for proof.)

Note 2: When $\rho_{xx} = 1$,

if $0 \leq d < 1$ then $\rho_{\xi\eta}^{-(x)} = 0$, $\rho_{\xi\eta}^{+(x)} = d$

if $d \geq 1$ then $\rho_{\xi\eta}^{-(x)} = 0$, $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$.

Note 3: $\sqrt{x} \leq 1$ with equality if and only if $\rho_{xx} = \frac{1}{2 - \rho_{yy}}$.

(See Appendix B.3 for proof.)

Prior to identifying general categories of β_x/β_ξ^* it will be worthwhile to examine the relationship of Q to 0 for various combinations of situations since as noted above in subsection III the relationship of Q to 0 provides some direct information about the relationship between β_x/β_ξ^* and 1. Much of the derivation for the results which follow has been developed in Appendix B.4.

If $|d| \geq \sqrt{x}$ then

(3.25a) $Q < 0$ for $\max(-1, \rho_{\xi\eta}^{-(x)}) < \rho_{\xi\eta} < \min(+1, \rho_{\xi\eta}^{+(x)})$

[By B.4.3a]

(3.25b) $Q = 0$ for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ provided that $\rho_{\xi\eta}^{-(x)} \in (-1,+1)$

[By B.4.3.b]

or for $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$ provided that $\rho_{\xi\eta}^{+(x)} \in (-1,+1)$,

[By B.4.3b]

(3.25c) $Q > 0$ for $-1 < \rho_{\xi\eta} < \max(-1, \rho_{\xi\eta}^{-(x)})$ [By B.4.3c]

or for $\min(+1, \rho_{\xi\eta}^{+(x)}) < \rho_{\xi\eta} < +1.$ [By B.4.3c]

(3.25d) If $|d| < \sqrt{x}$ then $Q > 0$ for all $\rho_{\xi\eta} \in (-1,+1)$.

[By B.4.4]

Now combine results (3.23a-b) or (3.24a-b), (3.25a-d) and (3.18a-c) or (3.19a-c) to derive general categories of distributions of β_x/β_ξ^* versus $\rho_{\xi\eta}$. Examination of these expressions will produce three general categories of distributions.

For exploration for the first general category, let $d > 1$ and $0 < \rho_{\xi\eta} < 1$. Therefore $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ by (3.23a) but (3.24a) and (3.24b) indicate that $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$. Using results (3.25a-d) and considering all possible values of $\rho_{\xi\eta}$, $\rho_{\xi\eta} \in (-1, +1)$ produces:

(3.26a) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$ then $Q > 0$

(3.26b) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ then $Q = 0$

(3.26c) and for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1$ then $Q < 0$.

(3.27) Note: For $d > 1$, $\rho_{\xi\eta}^{-(x)} < \frac{1}{d}$. (See Appendix B.5 for proof.)

Therefore since $d > 1$ (hence $\frac{1}{d} < 1$) combine the results from above and from (3.18a-c) and (3.19a-c) to produce information about β_x/β_ξ^* for values of $\rho_{\xi\eta} \in (-1, +1)$.

The following information will be presented for values of $\rho_{\xi\eta}$ which cover the whole interval from minus one to plus one.

a) Consider $d > 1$ and $\rho_{\xi\eta}$ such that $1/d < \rho_{\xi\eta} < +1$.

Therefore $d\rho_{\xi\eta} > 1$.

By (3.27) $\rho_{\xi\eta}^{-(x)} < 1/d$. Hence $\rho_{\xi\eta}^{-(x)} < 1/d < \rho_{\xi\eta} < +1$.

By (3.26c) for values of $\rho_{\xi\eta} > \rho_{\xi\eta}^{-(x)}$, $Q < 0$.

By (3.18a) when $d\rho_{\xi\eta} > 1$ and $Q < 0$, $\beta_x/\beta_\xi^* < 1$.

(3.28a) Therefore, when $d > 1$ and $1/d < \rho_{\xi\eta} < +1$ then $\beta_x/\beta_\xi^* < 1$.

b) Consider $d > 1$ and $\rho_{\xi\eta}$ such that $\rho_{\xi\eta} = 1/d$. Therefore $d\rho_{\xi\eta} = 1$. When $d\rho_{\xi\eta} = 1$, $\beta_\xi^* = 0$.

(3.28b) Therefore, when $d > 1$ and $\rho_{\xi\eta} = 1/d$, β_x/β_ξ^* is not defined. In this case $|\beta_x|$ is an overestimate of β_ξ^* unless β_x is also zero.

c) Consider $d > 1$ and $\rho_{\xi\eta}$ such that $-1 < \rho_{\xi\eta} < 1/d$ (where $1/d < 1$). Therefore $d\rho_{\xi\eta} < 1$.

By (3.27) $\rho_{\xi\eta}^{-(x)} < 1/d$. Therefore there are three subintervals of values for $\rho_{\xi\eta}$ here which must be examined.

I) For $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1/d$, then $Q < 0$ by (3.26c). By (3.19a) when $d\rho_{\xi\eta} < 1$ and $Q < 0$, $\beta_x/\beta_\xi^* < 1$, that is $|\beta_x|$ is an overestimate of $|\beta_\xi^*|$ for correlations in the interval $(\rho_{\xi\eta}^{-(x)}, 1/d)$.

(3.28c) Therefore, when $d > 1$ and $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1/d$, then $\beta_x/\beta_\xi^* > 1$.

II) For $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$, then $Q = 0$ by (3.26b). Thus by (3.19b), $\beta_x/\beta_\xi^* = 1$, that is when $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$, $\beta_x = \beta_\xi^*$.

(3.28d) Therefore, when $d > 1$ and $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ (where $\rho_{\xi\eta}^{-(x)} < 1/d$) then $\beta_x/\beta_\xi^* = 1$.

III) For $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$, then $Q > 0$ by (3.26a). Thus by (3.19c) $\beta_x/\beta_\xi^* < 1$.

(3.28c) Therefore, when $d > 1$ and $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$ (where

$\rho_{\xi\eta}^{-(x)} < 1/d)$ then $\beta_x/\beta_\xi^* < 1$.

To determine the relationship between β_x/β_ξ^* and zero for the range of values for $\rho_{\xi\eta}$, apply results (B.6.4a-d) from Appendix B.6 for $d > 1$. Since $0 < \rho_{yy} < 1$, then $1/d < 1/d\rho_{yy}$. Therefore, the Appendix B.6 results produce:

(3.29a) for $-1 < \rho_{\xi\eta} < 1/d$ then $\beta_x/\beta_\xi^* > 0$ [from (B.6.4a)],

(3.29b) for $\rho_{\xi\eta} = 1/d$ then β_x/β_ξ^* is undefined [from (B.6.4c)],

(3.29c) for $1/d < \rho_{\xi\eta} < \min(1, 1/d\rho_{yy})$ then $\beta_x/\beta_\xi^* < 0$ [from (B.6.4b)],

(3.29d) for $\rho_{\xi\eta} = 1/d\rho_{yy}$ then $\beta_x/\beta_\xi^* = 0$ [from (B.6.4d)],

and

(3.29e) for $\min(1, 1/d\rho_{yy}) < \rho_{\xi\eta} < 1$ then $\beta_x/\beta_\xi^* > 0$ [from (B.6.4a)].

Combining results (3.28a-e) with corresponding results from (3.29a-e) yields general category one (G.C.I.) of distributions for β_x/β_ξ^*.

General category One (G.C.I) of distributions for β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ has the following form as values of $\rho_{\xi\eta}$ vary across the interval $(-1, +1)$.

For $d > 1$, any $\rho_{xx}, \rho_{yy} \neq 1$ and

(3.30a) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$, $0 < \beta_x/\beta_\xi^* < 1$ [from (3.29a) and (3.28e)]

(3.30b) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$, $\beta_x/\beta_\xi^* = 1$ [from (3.28d)]

(3.30c) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \frac{1}{d}$, $\beta_x/\beta_\xi^* > 1$ [from (3.28c)]

(3.30d) for $\rho_{\xi\eta} = 1/d$, β_x/β_ξ^* is undefined since $\beta_\xi^* = 0$ [from (3.28b) or (3.29b)]

(3.30e) and for $1/d < \rho_{\xi\eta} < 1$, $\beta_x/\beta_\xi^* < 1$ [from (3.28a)].

Result (3.30e) can be further specialized as follows:

(3.30f) for $1/d < \rho_{\xi\eta} < \min(+1, 1/d\rho_{yy})$, $\beta_x/\beta_\xi^* < 0$ [from (3.29c)],

(3.30g) for $\rho_{\xi\eta} = 1/d\rho_{yy}$, $\beta_x/\beta_\xi^* = 0$ [from (3.29d)],

(3.30h) for $\min(+1, 1/d\rho_{yy}) < \rho_{\xi\eta} < +1$, $0 < \beta_x/\beta_\xi^* < 1$ [from (3.29e) and (3.28a)].

Note that G.C.I for β_x/β_ξ^* could actually be considered as having two subcategories depending on the behavior of β_x/β_ξ^* when $1/d < \rho_{\xi\eta} < 1$.

a) If $d > 1$ is also sufficiently large enough so that $d\rho_{yy} > 1$ $(1/d\rho_{yy} < 1)$ then (3.30f) becomes

(3.30j) for $1/d < \rho_{\xi\eta} < 1/d\rho_{yy}$, $\beta_x/\beta_\xi^* < 0$,

(3.30g) becomes

(3.30k) for $\rho_{\xi\eta} = 1/d\rho_{yy}$, $\beta_x/\beta_\xi^* = 0$,

and (3.30h) becomes

(3.30l) for $1/d\rho_{yy} < \rho_{\xi\eta} < 1$, $0 < \beta_x/\beta_\xi^* < 1$.

b) However, if $d > 1$ but $d\rho_{yy} < 1$ then $1/d\rho_{yy} > 1$ and (3.30f) becomes

(3.30m) for $1/d < \rho_{\xi\eta} < +1$, $\beta_x/\beta_\xi^* < 0$,

and (3.30g) and (3.30h) are not applicable.

Since the prime interest in examining β_x/β_ξ^* is in relationship to one, and since when $d > 1$ for $1/d < \rho_{\xi\eta} < 1$, $\beta_x/\beta_\xi^* < 1$, there is only academic interest in differentiating between the two subcategories of G.C.I identified above. Therefore, G.C. I will be considered as a single category of distributions with regard to the relationship of β_x/β_ξ^* to one.

G.C. I for β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ covers all values of ρ_{xx}, $\rho_{yy} \neq 1$ and $d > 1$. Therefore other general categories will involve values of d where $0 \leq d \leq 1$.

For exploration of the second general category, let $\sqrt{x} \leq d < 1$, $0 < \rho_{yy} < 1$ and $\frac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$. Therefore $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ by (3.23b) and $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$ by (3.24b). Using results (3.25a-c) and considering all possible values of $\rho_{\xi\eta}$, $\rho_{\xi\eta} \in (-1, +1)$ produces:

(3.31a) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$ then $Q > 0$ [from (3.25c)],

(3.31b) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$ then $Q < 0$ [from (3.25a)],

(3.31c) for $\rho_{\xi\eta}^{+(x)} < \rho_{\xi\eta} < 1$ then $Q > 0$ [from (3.25c)],

(3.31d) and for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ or $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$ then $Q = 0$ [from (3.25b)].

Note, since $\sqrt{x} \leq d < 1$ and $\rho_{\xi\eta} < +1$ then $d\rho_{\xi\eta} < 1$. Therefore by (B.6.4a), $\beta_x/\beta_\xi^* > 0$.

Thus combining the results (3.31a-d) with (3.19a-c) to produce information about β_x/β_ξ^* for values of $\rho_{\xi\eta} \in (-1, +1)$, yields general category two (G.C. II) of distributions for β_x/β_ξ^*.

When $\sqrt{x} \leq d < 1$ and $0 < \rho_{yy} < 1$

(3.32a) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$ then $0 < \beta_x/\beta_\xi^* < 1$ [from (B.6.4a), (3.31a) and (3.19c)],

(3.32b) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$ then $\beta_x/\beta_\xi^* > 1$ [from (3.31b) and (3.19a)].

(3.32c) for $\rho_{\xi\eta}^{+(x)} < \rho_{\xi\eta} < +1$ then $0 < \beta_x/\beta_\xi^* < 1$ [from (B.6.4a), (3.31c) and (3.19c)],

(3.32d) and for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ or $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$ then $\beta_x/\beta_\xi^* = 1$ [from (3.31d) and (3.19b)].

For exploration of the third general category, let $0 \leq d < \sqrt{x}$, $0 < \rho_{yy} < 1$ and consider any ρ_{xx}. Then neither $\rho_{\xi\eta}^{-(x)}$ nor $\rho_{\xi\eta}^{+(x)}$ will exist [see (B.4.2) from Appendix B.4]. By result (3.25d) $Q > 0$ for all $\rho_{\xi\eta} \in (-1, +1)$. Since $d < \sqrt{x}$ and $\sqrt{x} \leq 1$ [by (B.3.2) from Appendix B.3] then $d < 1$ and $d\rho_{\xi\eta} < 1$ also.

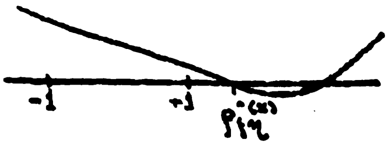
Therefore, since $d\rho_{\xi\eta} < 1$ and $Q > 0$ for $\rho_{\xi\eta}$ (-1, +1) applying (B.6.4a) and (3.19c) produces the following result:

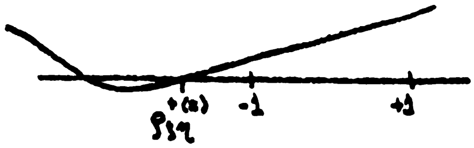
(3.33a) When $0 \leq d < \sqrt{x}$, and $0 < \rho_{yy} < 1$, for any ρ_{xx} and any $\rho_{\xi\eta} \in (-1, +1)$, then $0 < \beta_x/\beta_\xi^* < 1$.

This is general category three (G.C. III) of distributions of β_x/β_ξ^* .

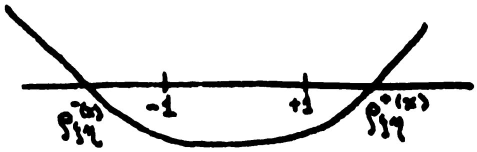
One other set of distributions also fall into G.C. III. Let $\sqrt{x} \leq d < 1$, $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$, and $0 < \rho_{yy} < 1$. Then $\rho_{\xi\eta}^{-(x)} \notin (-1, +1)$ and $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$.

If $\rho_{\xi\eta}^{-(x)} > +1$ or $\rho_{\xi\eta}^{+(x)} < -1$, then $Q > 0$ for $\rho_{\xi\eta} \in (-1, +1)$.





If $\rho_{\xi\eta}^{-(x)} < -1$ and $\rho_{\xi\eta}^{+(x)} > +1$ then $Q < 0$ for $\rho_{\xi\eta} \in (-1, +1)$.



It will be sufficient to show that for $d \geq \sqrt{x}$, $\rho_{\xi\eta}^{-(x)} \geq 0$ and thus $Q > 0$ for $\rho_{\xi\eta} \in (-1, +1)$.

Let $d \geq \sqrt{x}$ therefore $\rho_{\xi\eta}^{-(x)}$ will exist but may not exist in the interval $(-1, +1)$. Therefore the question is, for what values of ρ_{xx}, $\rho_{yy} \neq 1$ is $\rho_{\xi\eta}^{-(x)} \geq 0$?

$$\rho_{\xi\eta}^{-(x)} \geq 0 \Leftrightarrow \frac{d(1-\rho_{xx}\rho_{yy}) - \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})} \geq 0$$

since $2\rho_{xx}(1 - \rho_{yy}) > 0$ for $\rho_{yy} \neq 1$,

$$\rho_{\xi\eta}^{-(x)} \geq 0 \Leftrightarrow d(1 - \rho_{xx}\rho_{yy}) - \sqrt{d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})} \geq 0$$

$$\Leftrightarrow d(1 - \rho_{xx}\rho_{yy}) \geq \sqrt{d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}$$

$$\Leftrightarrow d^2(1 - \rho_{xx}\rho_{yy})^2 \geq d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})$$

since $d(1 - \rho_{xx}\rho_{yy}) > 0$ for $d \geq \sqrt{x}$ and $\rho_{yy} \neq 1$.

$$\rho_{\xi\eta}^{-(x)} \geq 0 \Leftrightarrow 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}) \geq 0.$$

But $0 < \rho_{xx} \leq 1$ and $0 < \rho_{yy} < 1$ by definition,

(3.34) Therefore, $\rho_{\xi\eta}^{-(x)} \geq 0$ for all values of ρ_{xx} and ρ_{yy} ($\rho_{yy} \neq 1$) when $d \geq \sqrt{x}$.

Since $\rho_{\xi\eta}^{-(x)} \geq 0$ when $d \geq \sqrt{x}$, then $Q > 0$ for $\rho_{\xi\eta} \in (-1, +1)$. For $\sqrt{x} \leq d < 1$, $d\rho_{\xi\eta} < 1$. Therefore, since $d\rho_{\xi\eta} < 1$ and $Q > 0$ for $\rho_{\xi\eta} \in (-1, +1)$, applying (B.6.4a) and (3.19c) produces the following result:

(3.33b) When $\sqrt{x} \leq d < 1$, $0 < \rho_{yy} < 1$, and $0 < \rho_{xx} \leq \dfrac{1}{2 - \rho_{yy}}$, for any $\rho_{\xi\eta} \in (-1, +1)$, then $0 < \beta_x/\beta_\xi^* < 1$.

The three general categories of distributions of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ include all values of the parameters d, ρ_{xx}, and ρ_{yy} except $d = 1$ and $\rho_{yy} = 1$. G.C. I includes values of $d > 1$, all values of ρ_{xx} and all values of ρ_{yy} except $\rho_{yy} = 1$. G.C. II includes values of d such that $\sqrt{x} \leq d < 1$, values of ρ_{xx} such that $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ and all values of ρ_{yy} except $\rho_{yy} = 1$. G.C. III includes values of d such that $\sqrt{x} \leq d < 1$, values of ρ_{xx} such that $0 < \rho_{xx} \leq 1/2 - \rho_{xx}$ and all values of ρ_{yy} except $\rho_{yy} = 1$. G.C. III also includes all values of d such that

$0 \leq d < \sqrt{x}$ for all values of ρ_{xx} and all values of $\rho_{yy} \neq 1$.

When $\rho_{yy} = 1$, β_x/β_ξ^* was examined in Appendix B.2. When $\rho_{yy} = 1$, then $0 < \beta_x/\beta_\xi^* < 1$ for all values of $\rho_{\xi\eta} \in (-1, +1)$ and $0 < \rho_{xx} < 1$ [by (B.2.8) from Appendix B.2]. When $\rho_{yy} = \rho_{xx} = 1$, β_x/β_ξ^* was examined in Appendix B.2. When $\rho_{yy} = \rho_{xx} = 1$, then $\beta_x/\beta_\xi^* = 1$ for all values of $\rho_{\xi\eta} \in (-1, +1)$ [by Appendix B.2, Section D].

The situation when $d = 1$ will be shown to represent a slight variation of G.C. III for $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$ with $\rho_{yy} \neq 1$ and to represent a middle ground between G.C. I and G.C. II for $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ with $\rho_{yy} \neq 1$. To examine the situation where $d = 1$ first determine the conditions for existence of $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$.

Let $d = 1$, then $\rho_{\xi\eta}^{-(x)}$ will exist and $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$, for $0 < \rho_{yy} < 1$ and for $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ by (3.23b). For $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$, $\rho_{\xi\eta}^{-(x)} \notin (-1, +1)$ by (3.23a). Referring to (3.24a) and (3.24b), $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$ for any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$).

Consider $d = 1$, $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$ and $\rho_{yy} \neq 1$. Since $d = 1$, then $d \geq \sqrt{x}$ (by (B.3.2) in Appendix B.3) and $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ exist but $\rho_{\xi\eta}^{-(x)} \notin (-1, +1)$ and $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$. The same argumet presented for $\sqrt{x} \leq d < 1$ when $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$, $\rho_{yy} \neq 1$ in general category three is completely applicable here hence

$$(3.35a) \qquad 0 < \beta_x/\beta_\xi^* < 1.$$

The variation of G.C. III which results when $d = 1$ is not obvious yet. It will be identified below.

Consider $d = 1$, $1/2 - \rho_{yy} < \rho_{xx} \leq 1$, $\rho_{yy} \neq 1$. Since $d = 1$, then $d \geq \sqrt{x}$ and both $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ exist but $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ by (3.23b) while $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$ by (3.24b). Since $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$, $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$ and $\rho_{\xi\eta}^{-(x)} \leq \rho_{\xi\eta}^{+(x)}$ then $\rho_{\xi\eta}^{+(x)} \geq +1$. Using results from (3.25a-c) and considering values of $\rho_{\xi\eta} \in (-1, +1)$ yields:

(3.36a) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$ then $Q > 0$ [from (3.25c)],

(3.36b) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ then $Q = 0$ [from (3.25b)],

and

(3.36c) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < +1$ then $Q < 0$ [from (3.25a)].

When $d = 1$, then $d\rho_{\xi\eta} < 1$ for $\rho_{\xi\eta} \in (-1, +1)$. Therefore applying results (B.6.4a) and (3.19a-c) to (3.36a-c) yields

(3.35b) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$ then $0 < \beta_x/\beta_\xi^* < 1$ [from (B.6.4a) and (3.19c)]

(3.35c) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ then $\beta_x/\beta_\xi^* = 1$ [from (3.19b)],

(3.35d) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1$ then $\beta_x/\beta_\xi^* > 1$ [from (3.19a)].

Some similarities to both G.C. I and G.C. II are obvious. More direct comparisons and contrasts require additional work to be presented below.

To continue to add information about the three general categories identified above as well as the situation when $d = 1$. The work which follows will examine the limiting case of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ as $\rho_{\xi\eta}$ is allowed to approach various values of interest.

For all three general categories as well as the situation where $d = 1$, the limiting case of β_x/β_ξ^* will be considered for values of $\rho_{\xi\eta}$ in an arbitrarily small neighborhood of negative one (-1). Since, by definition, $\rho_{\xi\eta} > -1$ the only values of $\rho_{\xi\eta}$ which can be included in the arbitrarily small neighborhood of negative one are values which are greater than negative one. The notation which will be used here is: the value of β_x/β_ξ^* will be examined as $\rho_{\xi\eta} \to -1^+$. The notation indicates that the value of β_x/β_ξ^* is to be examined for values of $\rho_{\xi\eta}$ which are greater than negative one (indicated by the + as a superscript) but which are arbitrarily close to negative one (indicated by the $\to$).

The value of β_x/β_ξ^* will also be examined as $\rho_{\xi\eta} \to +1^-$ for each subcategory of G.C. I, for G.C. II and G.C. III combined and for $d = 1$. [$\rho_{\xi\eta} \to +1^-$ indicates that the values of $\rho_{\xi\eta}$ which are to be considered are those values which are less than +1 (indicated by the - as a superscript) but which are arbitrarily close to +1].

For case G.C. I only, the values of β_x/β_ξ^* will be examined as $\rho_{\xi\eta} \to 1/d^-$ and as $\rho_{\xi\eta} \to 1/d^+$. (Recall: for G.C. I, $d > 1$, thus $1/d < 1$, and β_x/β_ξ^* is not defined for $\rho_{\xi\eta} = 1/d$ (i.e. $d\rho_{\xi\eta} = 1$) since $\beta_\xi^* = 0$).

The approach to the work on limits will be to determine the limits of β_x and β_ξ^* separately first and then consider the limit of β_x/β_ξ^* based on the work for the separate limits.

Therefore consider

$$\beta_x = \frac{b_\xi^* \rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})}{1 - \rho_{\xi\eta}^2 \rho_{xx}\rho_{yy}} \qquad \text{[from (3.11b) with } d = d_\xi\text{]},$$

and

$$\beta^*_\xi = \frac{b^*_\xi(1 - d\rho_{\xi\eta})}{1 - \rho^2_{\xi\eta}} \qquad \text{[from (3.6b) with } d = d_\xi\text{]}.$$

For general notation let $\beta_{x_{lim}} = \lim_{\rho_{\xi\eta} \to q} \beta_x$ and $\beta^*_{\xi lim} = \lim_{\rho_{\xi\eta} \to q} \beta^*_\xi$.

As $\rho_{\xi\eta} \to -1^+$, $\beta_x \to \beta_{x_{lim}} = \dfrac{b^*_\xi \rho_{xx}(1 + d\rho_{yy})}{1 - \rho_{xx}\rho_{yy}}$.

Since $d \geq 0$, $\beta_{x_{lim}} > 0$ if $b^*_\xi > 0$

$\beta_{x_{lim}} < 0$ if $b^*_\xi < 0$.

As $\rho_{\xi\eta} \to -1^+$, $1 - d\rho_{\xi\eta} \to 1 + d > 0$ and $1 - \rho^2_{\xi\eta} \to 0^+$.

Therefore $\beta^*_\xi \to +\infty$ if $b^*_\xi > 0$

$\beta^*_\xi \to -\infty$ if $b^*_\xi < 0$.

Therefore for any $d \geq 0$, ρ_{xx} and ρ_{yy} (i.e. any of the three general categories as well as $d = 1$), as

(3.37) $\qquad \rho_{\xi\eta} \to -1^+$, $\beta_x/\beta^*_\xi \to 0^+$.

Consider the subcategory of G.C. I with $d > 1/\rho_{yy}$ i.e. $d\rho_{yy} > 1$. (See expressions (3.30j-1) for the behavior of β_x/β^*_ξ when $1/d < \rho_{\xi\eta} < 1$ and $d < 1/\rho_{yy}$).

As $\rho_{\xi\eta} \to +1^-$, $\beta_x \to \beta_{x_{lim}} = \dfrac{b^*_\xi \rho_{xx}(1 - d\rho_{yy})}{1 - \rho_{xx}\rho_{yy}}$.

Since $d\rho_{yy} > 1 \Leftrightarrow 1 - d\rho_{yy} < 0$, then

$\beta_{x_{lim}} > 0$ if $b^*_\xi < 0$

$\beta_{x_{lim}} < 0$ if $b^*_\xi > 0$.

As $\rho_{\xi\eta} \to +1^-$, $1 - d\rho_{\xi\eta} \to 1 - d < 0$ (since $d > 1/\rho_{yy} > 1$) and $1 - \rho_{\xi\eta}^2 \to 0^+$. Therefore

$$\beta_\xi^* \to +\infty \quad \text{if} \quad b_\xi^* < 0$$

$$\beta_\xi^* \to -\infty \quad \text{if} \quad b_\xi^* > 0 \ .$$

Therefore, for G.C. I when $d\rho_{yy} > 1$

(3.38) as $\rho_{\xi\eta} \to +1^-$, $\beta_x/\beta_\xi^* \to 0^+$.

Consider the subcategory of G.C. I with $1 < d < 1/\rho_{yy}$ i.e. $d\rho_{yy} < 1$. (See expression (3.30m) for the behavior of β_x/β_ξ^* when $1/d < \rho_{\xi\eta} < 1$ and $d < 1/\rho_{yy}$).

As $\rho_{\xi\eta} \to +1^-$, $\beta_x \to \beta_{x_{lim}} = \dfrac{b_\xi^* \rho_{xx}(1 - d\rho_{yy})}{1 - \rho_{xx}\rho_{yy}}$.

Since $d\rho_{yy} < 1 \Leftrightarrow 1 - d\rho_{yy} > 0$, then

$$\beta_{x_{lim}} > 0 \quad \text{if} \quad b_\xi^* > 0$$

$$\beta_{x_{lim}} < 0 \quad \text{if} \quad b_\xi^* < 0 \ .$$

As $\rho_{\xi\eta} \to +1^-$, $1 - d\rho_{\xi\eta} \to 1 - d < 0$ (since $d > 1$), and $1 - \rho_{\xi\eta}^2 \to 0^+$. Therefore

$$\beta_\xi^* \to +\infty \quad \text{if} \quad b_\xi^* < 0$$

$$\beta_\xi^* \to -\infty \quad \text{if} \quad b_\xi^* > 0 \ .$$

Therefore, for G.C. I when $d\rho_{yy} < 1$,

(3.39) as $\rho_{\xi\eta} \to +1^-$, $\beta_x/\beta_\xi^* \to 0^-$.

Consider either G.C. II or G.C. III. In each category $d < 1$.

As $\rho_{\xi\eta} \to +1^-$, $\beta_X \to \beta_{X_{lim}} = \dfrac{b_\xi^* \rho_{xx}(1 - d\rho_{yy})}{1 - \rho_{xx}\rho_{yy}}$.

Since $d < 1$, $d\rho_{yy} < 1 \Leftrightarrow 1 - d\rho_{yy} > 0$, then

$$\beta_{X_{lim}} > 0 \quad \text{if} \quad b_\xi^* > 0$$

$$\beta_{X_{lim}} < 0 \quad \text{if} \quad b_\xi^* < 0 \ .$$

As $\rho_{\xi\eta} \to +1^-$, $1 - d\rho_{\xi\eta} \to 1 - d > 0$ (since $d < 1$) and $1 - \rho_{\xi\eta}^2 \to 0^+$.

Therefore,

$$\beta_\xi^* \to +\infty \quad \text{if} \quad b_\xi^* > 0$$

$$\beta_\xi^* \to -\infty \quad \text{if} \quad b_\xi^* < 0.$$

Therefore, for either G.C. II or G.C. III (i.e., $d < 1$),

(3.40) as $\rho_{\xi\eta} \to +1^-$, $\beta_X/\beta_\xi^* \to 0^+$.

Consider the situation where $d = 1$.

As $\rho_{\xi\eta} \to +1^-$, $\beta_X \to \beta_{X_{lim}} = \dfrac{b_\xi^* \rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}}$.

$$\beta_{X_{lim}} > 0 \quad \text{if} \quad b_\xi^* > 0$$

$$\beta_{X_{lim}} < 0 \quad \text{if} \quad b_\xi^* < 0.$$

As $\rho_{\xi\eta} \to +1^-$, $\beta_\xi^* = \dfrac{b_\xi^*(1 - \rho_{\xi\eta})}{(1 - \rho_{\xi\eta}^2)} \to \beta_{\xi_{lim}}^* = \dfrac{b_\xi^*}{2}$ by L'Hopital's Rule.[1]

---

[1]Thomas, Goerge B., Calculus and Analytic Geometry, Addison-Wesley Publishing Co., Reading, 1968, pg. 651.

Therefore, when $d = 1$

(3.41a) as $\rho_{\xi\eta} \to +1^-$, $\beta_x/\beta_\xi^* \to \dfrac{\beta_{x_{lim}}}{\beta^*_{\xi lim}} = \dfrac{2\rho_{xx}(1 - \rho_{yy})}{(1 - \rho_{xx}\rho_{yy})}$.

Note

(3.41b) $\dfrac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} \leq 1 \Leftrightarrow 0 < \rho_{xx} \leq \dfrac{1}{2 - \rho_{yy}}$ [See (B.3.4a) and (B.3.4b) from Appendix B.3 for proof].

$\dfrac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} > 1 \Leftrightarrow \dfrac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ [See (B.3.4c) from Appendix B.3 for proof].

Consider G.C. I, where $d > 1$ and for $\rho_{\xi\eta} = 1/d$ (i.e. $d\rho_{\xi\eta} = 1$) β_x/β_ξ^* is not defined. For the arguments below consider d as some fixed value such that $d > 1$.

As $\rho_{\xi\eta} \to 1/d^-$, i.e. $d\rho_{\xi\eta} \to 1^-$,

$$\beta_x \to \beta_{x_{lim}} = \frac{b_\xi^*\rho_{xx}(1 - \rho_{yy})}{1 - \dfrac{\rho_{xx}\rho_{yy}}{d^2}} .$$

Since $d > 1$, $1 - \dfrac{\rho_{xx}\rho_{yy}}{d^2} > 0$, therefore as $\rho_{\xi\eta} \to 1/d^-$,

$$\beta_{x_{lim}} > 0 \quad \text{if} \quad b_\xi^* > 0$$

$$\beta_{x_{lim}} < 0 \quad \text{if} \quad b_\xi^* < 0.$$

Since $d > 1$, $1 - \dfrac{1}{d^2} > 0$, thus as

$$\rho_{\xi\eta} \to 1/d^- \quad \text{i.e.} \quad d\rho_{\xi\eta} \to 1^- \quad \beta_\xi^* \to \beta^*_{\xi lim} \to \begin{cases} 0^+ & \text{if} \quad b_\xi^* > 0 \\ 0^- & \text{if} \quad b_\xi^* < 0 . \end{cases}$$

Therefore for G.C. I

(3.42) as $\rho_{\xi\eta} \to 1/d^-$, $\beta_x/\beta_\xi^* \to +\infty$.

As $\rho_{\xi\eta} \to 1/d^+$ i.e. $d\rho_{\xi\eta} \to 1^+$,

$$\beta_X \to \beta_{X_{lim}} = \frac{b_\xi^* \rho_{xx}(1 - \rho_{yy})}{1 - \frac{\rho_{xx}\rho_{yy}}{d^2}} .$$

Since $d > 1$, $\frac{1 - \rho_{xx}\rho_{yy}}{d^2} > 0$, therefore as $\rho_{\xi\eta} \to 1/d^+$,

$$\beta_{X_{lim}} > 0 \quad \text{if} \quad b_\xi^* > 0$$

$$\beta_{X_{lim}} < 0 \quad \text{if} \quad b_\xi^* < 0.$$

As $\rho_{\xi\eta} \to 1/d^+$ i.e. $d\rho_{\xi\eta} \to 1^+$,

$$\beta_\xi^* \to \beta_{\xi lim}^* \to \begin{cases} 0^+ & \text{if} \quad b_\xi^* < 0 \\ 0^- & \text{if} \quad b_\xi^* > 0 . \end{cases}$$

Therefore for G.C. I

(3.43) as $\rho_{\xi\eta} \to 1/d^+$, $\beta_x/\beta_\xi^* \to -\infty$.

Now combining the results on limits (3.37) through (3.43) with the results on the relationship of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ to zero and one for each of the three general categories as well as the situation when $d = 1$ [(3.30a-e), (3.30j-l), (3.30m) for G.C. I; (3.32a-d) for G.C. II; (3.33a-b) for G.C. III and (3.35a-d) for $d = 1$], it is possible to describe more fully the characteristics of each general category and to produce for each general category a generic graph which represents the general shape of all distributions in the category. The information summarized below is presented as $\rho_{\xi\eta}$ ranges from near -1, through 0 and finally to near +1.

G.C. I, i.e. $d > 1$, any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$):

Subcategory a) $d > 1/\rho_{yy}$ i.e. $d\rho_{yy} > 1$ and $1/d < 1/d\rho_{yy} < 1$,

(3.44a) as $\rho_{\xi\eta} \to -1^+$, $\beta_x/\beta_\xi^* \to 0^+$ [by (3.37)],

(3.44b) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$, $0 < \beta_x/\beta_\xi^* < 1$ [by (3.30a)],

(3.44c) for $\rho_{\xi\eta} = 0$, $\beta_x/\beta_\xi^* = \rho_{xx}$ [by section C Appendix B.3],

(3.44d) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$, $\beta_x/\beta_\xi^* = 1$ [by (3.30b)],

(3.44e) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1/d$, $\beta_x/\beta_\xi^* > 1$ [by (3.30c)],

(3.44f) as $\rho_{\xi\eta} \to 1/d^-$, $\beta_x/\beta_\xi^* \to +\infty$ [by (3.42)],

(3.44g) for $\rho_{\xi\eta} = 1/d$, β_x/β_ξ^* is undefined [by (3.30d)],

(3.44h) as $\rho_{\xi\eta} \to 1/d^+$, $\beta_x/\beta_\xi^* \to -\infty$ [by (3.43)],

(3.44i) for $1/d < \rho_{\xi\eta} < 1/d\rho_{yy}$, $\beta_x/\beta_\xi^* < 0$ [by (3.30j)],

(3.44j) for $\rho_{\xi\eta} = 1/d\rho_{yy}$, $\beta_x/\beta_\xi^* = 0$ [by (3.30k)],

(3.44k) for $1/d\rho_{yy} < \rho_{\xi\eta} < 1$, $0 < \beta_x/\beta_\xi^* < 1$ [by (3.30l)],

and

(3.44l) as $\rho_{\xi\eta} \to +1^-$, $\beta_x/\beta_\xi^* \to 0^+$ [by (3.38)].

Note for G.C.I: since $d > 1$, $\rho_{\xi\eta}^{-(x)} \geq 0$ [by (3.34)].

Thus the generic graph of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ for G.C. I subcategory a) has the following general shape:

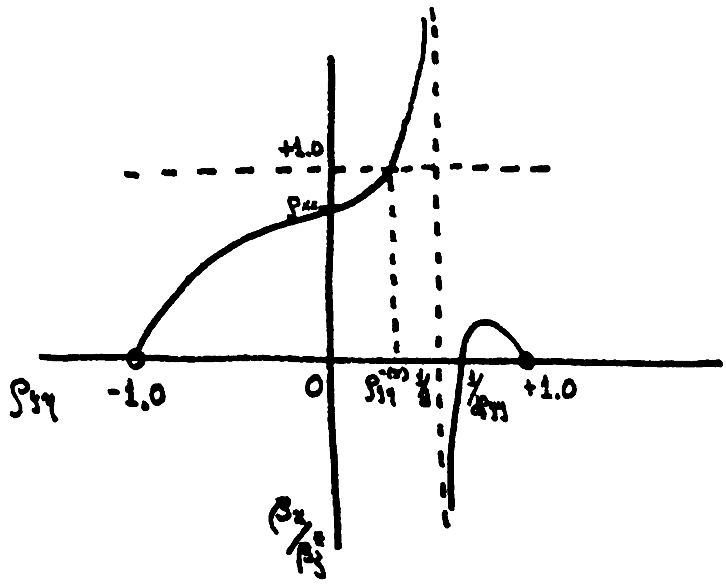


Figure 3.1a

Subcategory b) $1 < d < \frac{1}{\rho_{yy}}$, i.e. $d\rho_{yy} < 1$ and $1/d < 1 < 1/d\rho_{yy}$. Since subcategory b) differs from subcategory a) only when $1/d < \rho_{\xi\eta} < 1$, expressions (3.44a-h) for subcategory a) also apply for subcategory b). Therefore all that is needed to finish specifying subcategory b) is:

(3.44m) for $1/d < \rho_{\xi\eta} < 1$, $\beta_x/\beta_\xi^* < 0$ [by (3.30m)],

and

(3.44n) as $\rho_{\xi\eta} \to +1^-$, $\beta_x/\beta_\xi^* \to 0^-$ [by (3.39)].

Thus the generic graph of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ for G.C. I subcategory b) has the following general shape:

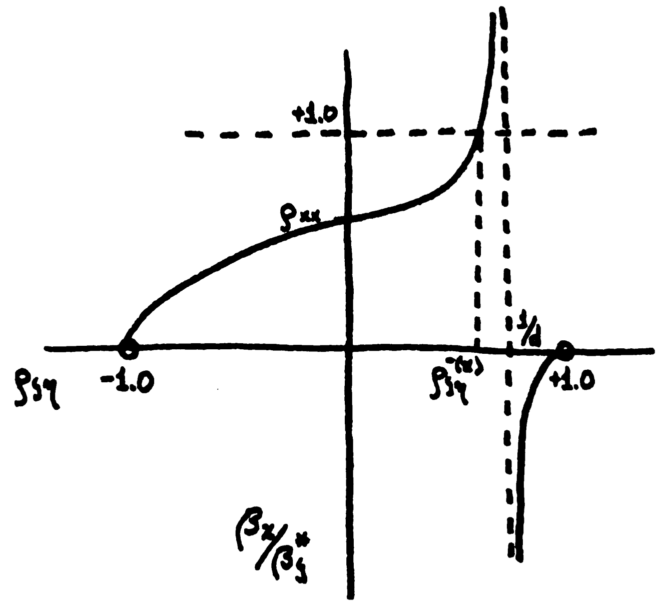


Figure 3.1b

G.C. II, i.e. $\sqrt{x} \leq d < 1$, $1/2 - \rho_{yy} < \rho_{xx} \leq 1$, $0 < \rho_{yy} < 1$ (i.e. $\rho_{yy} \neq 1$):

(3.45a) as $\rho_{\xi\eta} \to -1^{+}$, $\beta_x/\beta_\xi^* \to 0^{+}$ [by (3.37)],

(3.45b) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$, $0 < \beta_x/\beta_\xi^* < 1$ [by (3.32a)],

(3.45c) for $\rho_{\xi\eta} = 0$, $\beta_x/\beta_\xi^* = \rho_{xx}$ [by section C Appendix B.2],

(3.45d) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$, $\beta_x/\beta_\xi^* = 1$ [by (3.32d)],

(3.45e) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$, $\beta_x/\beta_\xi^* > 1$ [by (3.32b)],

(3.45f) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$, $\beta_x/\beta_\xi^* = 1$ [by (3.32d)],

(3.45g) for $\rho_{\xi\eta}^{+(x)} < \rho_{\xi\eta} < 1$, $0 < \beta_x/\beta_\xi^* < 1$ [by (3.32c)],

and

(3.45h) as $\rho_{\xi\eta} \to +1^{-}$, $\beta_x/\beta_\xi^* \to 0^{+}$ [by (3.40)].

Thus the generic graph of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ for G.C. II has the following general shape:

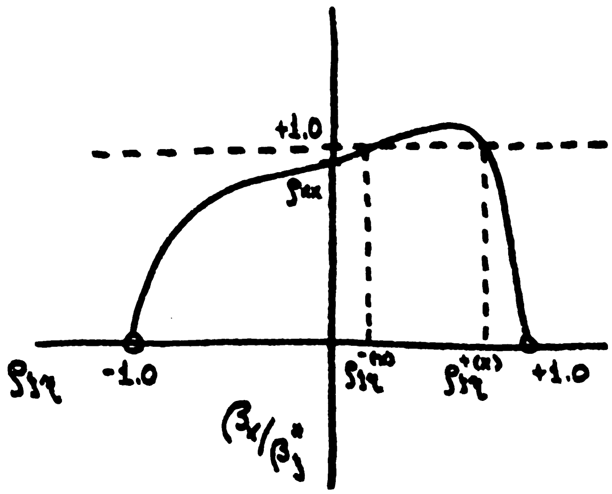


Figure 3.2

G.C. III, i.e. $0 \leq d < \sqrt{x}$, any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$) or $\sqrt{x} \leq d < 1$, $0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}}$, $0 < \rho_{yy} < 1$ (i.e. $\rho_{yy} \neq 1$):

(3.46a) as $\rho_{\xi\eta} \to -1^+$, $\beta_x/\beta_\xi^* \to 0^+$ [by (3.37)],

(3.46b) for $-1 < \rho_{\xi\eta} < +1$, $0 < \beta_x/\beta_\xi^* < 1$ [by (3.33a) or (3.33b)] ,

(3.46c) for $\rho_{\xi\eta} = 0$, $\beta_x/\beta_\xi^* = \rho_{xx}$ [by section C Appendix B.2],

and

(3.46d) as $\rho_{\xi\eta} \to +1^-$, $\beta_x/\beta_\xi^* \to 0^+$ [by (3.40)].

Thus the generic graph of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ for G.C. III has the following general shape:

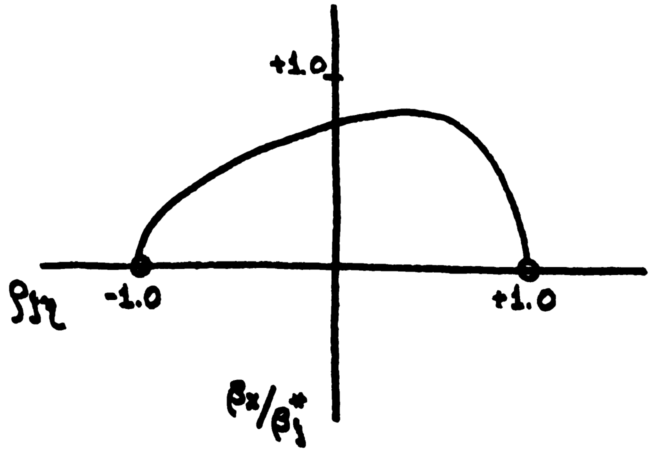


Figure 3.3

When $d = 1$ and $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$ with $\rho_{yy} \neq 1$, then:

(3.47a) as $\rho_{\xi\eta} \to -1^+$, $\beta_x/\beta_\xi^* \to 0^+$ [by (3.37)],

(3.47b) for $-1 < \rho_{\xi\eta} < +1$, $0 < \beta_x/\beta_\xi^* < 1$ [by (3.35a)],

(3.47c) for $\rho_{\xi\eta} = 0$, $\beta_x/\beta_\xi^* = \rho_{xx}$ [by section C Appendix B.2],

(3.47d) as $\rho_{\xi\eta} \to +1^-$, $\beta_x/\beta_\xi^* \to \dfrac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} \leq 1$ [by (3.41a) and (3.41b)] .

Note that (3.47a) through (3.47c) above are identical to (3.46a) through (3.46c) for G.C. III. The only difference, the variation, occurs as $\rho_{\xi\eta}$ approaches one, (3.47d) versus (3.46d).

Thus the generic graph of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ for $d = 1$, $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$ with $\rho_{yy} \neq 1$, which is somewhat similar to the generic graph for G.C. III, has the following general shape:

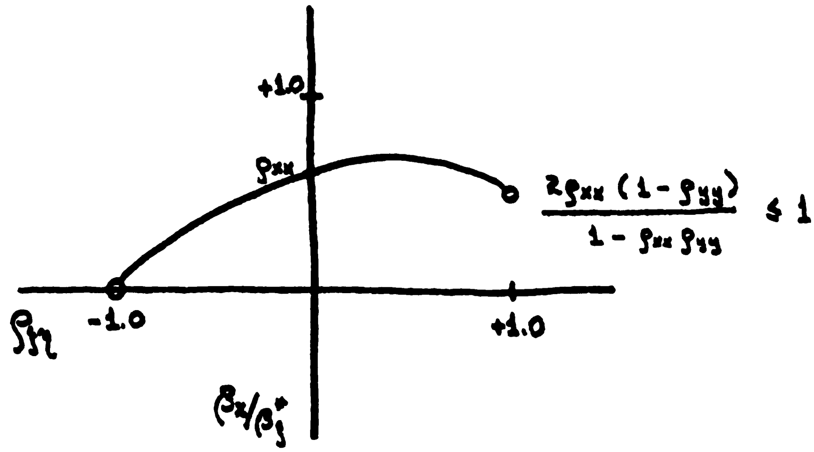


Figure 3.4a

When $d = 1$ and $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ with $\rho_{yy} \neq 1$, then:

(3.48a) as $\rho_{\xi\eta} \to -1^{+}$, $\beta_x/\beta^*_\xi \to 0^{+}$ [by (3.37)],

(3.48b) for $-1 < \rho_{\xi\eta} < \rho^{-(x)}_{\xi\eta}$, $0 < \beta_x/\beta^*_\xi < 1$ [by (3.35b)],

(3.48c) for $\rho_{\xi\eta} = 0$, $\beta_x/\beta^*_\xi = \rho_{xx}$ [by section C Appendix B.2],

(3.48d) for $\rho_{\xi\eta} = \rho^{-(x)}_{\xi\eta}$, $\beta_x/\beta^*_\xi = 1$ [by (3.35c)],

(3.48e) for $\rho^{-(x)}_{\xi\eta} < \rho_{\xi\eta} < 1$, $\beta_x/\beta^*_\xi > 1$ [by (3.35d)],

and

(3.48f) as $\rho_{\xi\eta} \to +1^{-}$, $\beta_x/\beta^*_\xi \to \dfrac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} > 1$

[by (3.41a) and (3.41b)] .

Note that (3.48a) through (3.48d) above are identical to (3.44a) through (3.44d) of G.C. I and to (3.45a) through (3.45d) of G.C. II. Like both G.C. I and G.C. II $\beta_x/\beta^*_\xi > 1$, for the situation above with $d = 1$, when $\rho_{\xi\eta} > \rho^{-(x)}_{\xi\eta}$. But unlike G.C. II β_x/β^*_ξ

never gets infinitely large and never gets negative for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1$, and unlike G.C. III β_x/β_ξ^* does not approach zero as $\rho_{\xi\eta}$ approaches +1.

Thus the generic graph of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ for $d = 1$, $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ with $\rho_{yy} \neq 1$ has the following general shape:

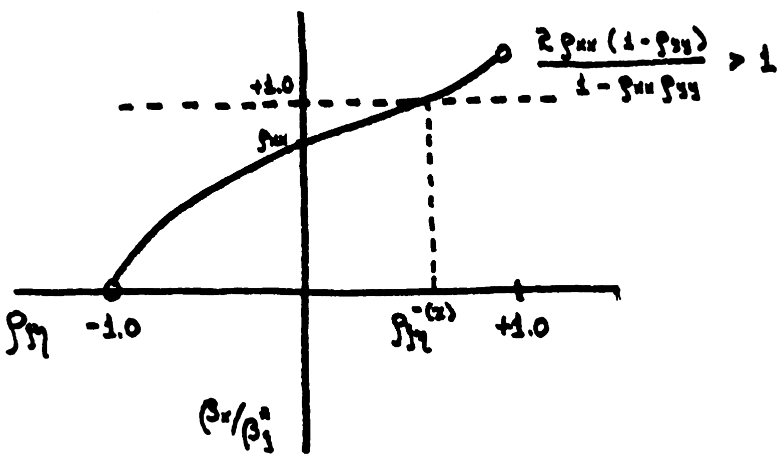


Figure 3.4b

Examples of graphs in each of the three general categories for fixed values of d, ρ_{xx} and ρ_{yy} will be presented below combined with the work on general categories of distributions for β_y/β_η^* as a function of $\rho_{\xi\eta}$.

V. The Search for Categories of Distributions of β_y/β_η^* as a Function of $\rho_{\xi\eta}$

A question which arises immediately relates to the need for this section based on the results demonstrated in Appendix B.1 for the property of interchangeability of x and y. In Appendix B.1 it is shown that given any situation (i.e., given values of $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy}

and d_η) where the value of β_y/β_η^* is needed, that value of β_y/β_η^* is identical to the value of β_x/β_ξ^* for another situation with the relationship between the two situations provided by (3.15a-d). Hence it is necessary to derive detailed results for β_x/β_ξ^* only.

The important thing to note about that result is that in general the two situations, the one of interest for β_y/β_η^* and the adjusted one for β_x/β_ξ^*, will be different situations. That is the values $\rho_{\xi\eta}$, ρ_{yy}, ρ_{xx} and d_η used to get a value for β_y/β_η^* will in general not be identical to the values $\rho'_{\xi\eta}$, ρ'_{yy}, ρ'_{xx} and d'_ξ (as related to $\rho_{\xi\eta}$, ρ_{yy}, ρ_{xx} and d_η by (3.15a-d) used to get the same value for β_x/β_ξ^*.

Therefore since one of the interests of this research is to examine the ratios β_x/β_ξ^* and β_y/β_η^* for the same situations, i.e., values of $\rho_{\xi\eta}$, ρ_{yy}, ρ_{xx} and $d = d_\xi = 1/d_\eta$, it will be necessary to state categories of distributions and important algebraic results for β_y/β_η^* which are analogous to those derived for β_x/β_ξ^*.

The properties (3.15a-d), i.e., the property of interchangeability of x and y will simplify the work for β_y/β_η^* considerably. It will be necessary to consider the major results for β_x/β_ξ^* and apply the property of interchangeability to arrive at analogous conclusions for β_y/β_η^*. In simple terms applying the property of interchangeability to a result for β_x/β_ξ^* involves replacing every x with a y, every y with an x, every ξ and an η and every η with a ξ. Therefore $\rho_{\xi\eta}$ is replaced by $\rho_{\eta\xi}$ but since there is no variable ordering in a correlation coefficient $\rho_{\eta\xi} = \rho_{\xi\eta}$, ρ_{xx} is replaced by ρ_{yy}, ρ_{yy} is replaced by ρ_{xx} and $d = d_\xi$ is replaced by

$d_\eta = 1/d_\xi = 1/d$ (i.e., d is replaced by $1/d$ provided $d \neq 0$). Therefore for all applications of the property of interchangeability consider $d > 0$. The situation when $d = 0$ will be examined separately.

The guiding interest in this phase of the research is to compare β_x/β_ξ^* and now β_y/β_η^* to one. That is, to see for what situations is β_y an overestimate ($\beta_y/\beta_\eta^* > 1$), an underestimate ($\beta_y/\beta_\eta^* < 1$) or an exact estimate ($\beta_y/\beta_\eta^* = 1$) of β_η^*.

Combining results (3.18b) or (3.19b) with (B.4.2) produces: there will exist a value of $\rho_{\xi\eta} \in (-1, +1)$ such that $\beta_x/\beta_\xi^* = 1$

$$\text{if} \quad |d| \geq \sqrt{\frac{4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}{(1-\rho_{xx}\rho_{yy})^2}} \equiv \sqrt{x} \;.$$

The corresponding result for β_y/β_η^* becomes by the property of interchangeability:

there will exist a value of $\rho_{\xi\eta} \in (-1, +1)$ such that $\beta_y/\beta_\eta^* = 1$

$$\text{(3.49)} \qquad \text{if} \quad \left|\frac{1}{d}\right| \geq \sqrt{\frac{4\rho_{yy}(1-\rho_{xx})(1-\rho_{yy})}{(1-\rho_{yy}\rho_{xx})^2}} \equiv \sqrt{(y)} \quad \text{for } d \neq 0.$$

Since it would be generally more useful to consider values of $|d|$ rather than $|1/d|$ (3.49) becomes: there will exist a value of $\rho_{\xi\eta} \in (-1, +1)$ such that $\beta_y/\beta_\eta^* = 1$

$$\text{(3.50)} \qquad \text{if} \quad |d| \leq \sqrt{\frac{(1-\rho_{xx}\rho_{yy})^2}{4\rho_{yy}(1-\rho_{xx})(1-\rho_{yy})}} \equiv \sqrt{y} \quad \text{for } d \neq 0,$$

$$\rho_{xx} \neq 1,\ \rho_{yy} \neq 1.$$

Note that $\sqrt{(y)}$ and $\sqrt{y}$ are used as abbreviations in notation. Here the square root sign indicates that the quantity involved is a square root and the y indicates that the expression is related to work on β_y/β_η^*. $\sqrt{(y)}$ is a temporary symbol to be used only for the immediate results of the application of the property of interchangeability to $\sqrt{x}$ (defined by (3.22)). The permanent statement of results will involve $\sqrt{y}$ as will be seen below.

Since $\sqrt{x} \leq 1$ for all ρ_{xx}, ρ_{yy} (except $\rho_{xx}\rho_{yy} = 1$) with equality for $\rho_{xx} = 1/2 - \rho_{yy}$. Then

$$\sqrt{(y)} \leq 1 \quad \text{for all } \rho_{yy}, \rho_{xx} \quad (\text{except } \rho_{yy}\rho_{xx} = 1)$$

with equality for $\rho_{yy} = 1/2 - \rho_{xx}$, which is equivalent to

$$(3.51) \qquad \sqrt{y} \geq 1 \quad \text{for all } \rho_{xx}, \rho_{yy} \quad (\text{except } \rho_{xx} = 1, \rho_{yy} = 1)$$

with equality for $\rho_{xx} = \dfrac{2\rho_{yy} - 1}{\rho_{yy}}$, since $\dfrac{1}{\sqrt{(y)}} = \sqrt{y}$ (by definition in (3.50)) and

$$\rho_{yy} = \frac{1}{2 - \rho_{yy}} \Leftrightarrow \rho_{xx} = \frac{2\rho_{yy} - 1}{\rho_{yy}}$$

(by the second result in (B.3.3b) from Appendix B.3).

For β_x/β_ξ^* if $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ or $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$ then $\beta_x/\beta_\xi^* = 1$. Let $\rho_{\xi\eta}^{-(x)}$ be replaced by $\rho_{\xi\eta}^{-(y)}$ and $\rho_{\xi\eta}^{+(x)}$ be replaced by $\rho_{\xi\eta}^{+(y)}$ when the property of interchangeability is applied.

Therefore for β_y/β_η^* if $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(y)}$ or $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(y)}$ then $\beta_y/\beta_\eta^* = 1$, and expression (3.21a) for $\rho_{\xi\eta}^{-(x)}$ becomes

$$\rho_{\xi\eta}^{-(y)} = \frac{\dfrac{1-\rho_{yy}\rho_{xx}}{d} - \sqrt{\dfrac{(1-\rho_{yy}\rho_{xx})^2}{d^2} - 4\rho_{yy}(1-\rho_{xx})(1-\rho_{yy})}}{2\rho_{yy}(1-\rho_{xx})}$$

for $d \neq 0, \rho_{xx} \neq 1.$

Therefore

$$(3.21a') \quad \rho_{\xi\eta}^{-(y)} = \frac{(1-\rho_{yy}\rho_{xx}) - \sqrt{(1-\rho_{yy}\rho_{xx})^2 - 4d^2\rho_{yy}(1-\rho_{xx})(1-\rho_{yy})}}{2d\rho_{yy}(1-\rho_{xx})}$$

for $d \neq 0, \rho_{xx} \neq 1$

and (3.21b) for $\rho_{\xi\eta}^{+(x)}$ becomes

$$(3.21b') \quad \rho_{\xi\eta}^{+(y)} = \frac{(1-\rho_{yy}\rho_{xx}) + \sqrt{(1-\rho_{yy}\rho_{xx})^2 - 4d^2\rho_{yy}(1-\rho_{xx})(1-\rho_{yy})}}{2d\rho_{yy}(1-\rho_{xx})}$$

for $d \neq 0, \rho_{xx} \neq 1.$

Note here that the expressions for β_y/β_η^* produced directly from expressions for β_x/β_ξ^* using the property of interchangeability will not be assigned new expression numbers. The number assigned to the expression for β_y/β_η^* will be the same number as that of the original expression for β_x/β_ξ^*. To differentiate between the two expressions the expression number for β_y/β_η^* will always be primed (i.e., given a ' as a superscript).

Consider now the conditions for existence of $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ and hence for the existence of $\rho_{\xi\eta}^{-(y)} \in (-1, +1)$.

Result (3.23a) states: $\rho_{\xi\eta}^{-(x)}$ exists and $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ for $\rho_{yy} \neq 1$ and

$$\text{for} \quad 0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}} \quad \text{if} \quad d > 1.$$

This becomes for $\rho^{-(y)}_{\xi\eta}$: $\rho^{-(y)}_{\xi\eta}$ exists and $\rho^{-(y)}_{\xi\eta} \in (-1, +1)$ for $\rho_{xx} \neq 1$ and

$$\text{for} \quad 0 < \rho_{yy} \leq \frac{1}{2 - \rho_{xx}} \quad \text{if} \quad 1/d > 1.$$

Rearranging this result so that the interval of reliabilities is an interval of ρ_{xx} values rather than ρ_{yy} values and so that the condition is expressed in terms of d rather than $1/d$ produces:

$\rho^{-(y)}_{\xi\eta}$ exists and $\rho^{-(y)}_{\xi\eta} \in (-1, +1)$

(3.23a') $$\text{for} \quad \frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1 \qquad \text{if} \quad 0 < d < 1.$$

This rearrangement results from $0 < \rho_{yy} \leq \frac{1}{2 - \rho_{xx}} \Leftrightarrow \frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1$ by (B.3.3b) and (B.3.3a) from Appendix B.3 and from the fact that $d > 1$ for β_x/β^*_ξ also implies $d > 0$ and $1/d > 1 \Leftrightarrow d < 1$.

Result (3.23b) states: $\rho^{-(x)}_{\xi\eta}$ exists and $\rho^{-(x)}_{\xi\eta} \in (-1, +1)$ for $\rho_{yy} \neq +1$ and

$$\text{for} \quad \frac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1 \quad \text{if} \quad \sqrt{_x} \leq d \quad \text{or if} \quad -1 < d \leq -\sqrt{_x}\,.$$

This becomes for $\rho^{-(y)}_{\xi\eta}$: $\rho^{-(y)}_{\xi\eta}$ exists and $\rho^{-(y)}_{\xi\eta} \in (-1, +1)$

for $\rho_{xx} \neq +1$ and

$$\text{for} \quad \frac{1}{2 - \rho_{xx}} < \rho_{yy} \leq 1 \quad \text{if} \quad \sqrt{_{(y)}} \leq 1/d \quad \text{or if} \quad -1 < \frac{1}{d} \leq -\sqrt{_{(y)}}\,.$$

Rearranging this result as above produces: $\rho_{\xi\eta}^{-(y)}$ exists and $\rho_{\xi\eta}^{-(y)} \in (-1, +1)$

(3.23b') for $0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$ if $0 < d \leq \sqrt{y}$ or if

$$-\sqrt{y} \leq d < -1.$$

This rearrangement results from $\frac{1}{2\rho_{yy} - 1} < \rho_{yy} \leq 1 \Leftrightarrow 0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$ by (B.3.3c) from Appendix B.3 and from the fact that $\sqrt{x} \leq d$ for β_x/β_ξ^* also implies $d > 0$, that $-1 < 1/d \leq -\sqrt{(y)}$ for β_x/β_ξ^* implies $d < 0$, and that $\sqrt{(y)} = 1/\sqrt{y}$.

Result (3.24a) for $\rho_{\xi\eta}^{+(x)}$ states: $\rho_{\xi\eta}^{+(x)}$ exists and $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$ for $\rho_{yy} \neq 1$ and

for $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$ if $d < -1$.

This becomes for $\rho_{\xi\eta}^{+(y)}$: $\rho_{\xi\eta}^{+(y)}$ exists and $\rho_{\xi\eta}^{+(y)} \in (-1, +1)$ for $\rho_{xx} \neq 1$ and

for $0 < \rho_{yy} \leq 1/2 - \rho_{xx}$ if $1/d < -1$.

Rearranging this result produces: $\rho_{\xi\eta}^{+(y)}$ exists and $\rho_{\xi\eta}^{+(y)} \in (-1, +1)$

(3.24a') for $\frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1$ if $-1 < d < 0$.

This rearrangement results from the fact that $1/d < -1$ for β_x/β_ξ^* also implies that $d < 0$.

Result (3.24b) for $\rho_{\xi\eta}^{+(x)}$ states: $\rho_{\xi\eta}^{+(x)}$ exists and $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$ for $\rho_{yy} \neq 1$ and

$$\text{for} \quad 1/2 - \rho_{yy} < \rho_{xx} \leq 1 \quad \text{if} \quad \sqrt{x} \leq d < 1 \quad \text{or if}$$

$$d \leq - \sqrt{x} \; .$$

This becomes for $\rho_{\xi\eta}^{+(y)}$: $\rho_{\xi\eta}^{-(y)}$ exists and $\rho_{\xi\eta}^{+(y)} \in (-1, +1)$ for $\rho_{xx} \neq 1$ and

$$\text{for} \quad \frac{1}{2 - \rho_{xx}} < \rho_{yy} \leq 1 \quad \text{if} \quad \sqrt{(y)} \leq 1/d < 1 \quad \text{or if}$$

$$1/d \leq - \sqrt{(y)} \; .$$

Rearranging this result produces: $\rho_{\xi\eta}^{+(y)}$ exists and $\rho_{\xi\eta}^{+(y)} \in (-1, +1)$

$$\text{(3.24b')} \quad \text{for} \quad 0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}} \quad \text{if} \quad 1 < d \leq \sqrt{y} \quad \text{or if}$$

$$- \sqrt{y} \leq d < 0.$$

Consolidating the above results produces: $\rho_{\xi\eta}^{-(y)}$ exists and $\rho_{\xi\eta}^{-(y)} \in (-1, +1)$

$$\text{(3.23b')} \quad \text{for} \quad 0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}} \quad \text{if} \quad 0 < d < \sqrt{y} \quad \text{or if}$$

$$- \sqrt{y} \leq d < -1,$$

$$\text{(3.23a')} \quad \text{for} \quad \frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1 \quad \text{if} \quad 0 < d < 1,$$

and $\rho_{\xi\eta}^{+(y)}$ exists and $\rho_{\xi\eta}^{+(y)} \in (-1, +1)$

$$\text{(3.24b')} \quad \text{for} \quad 0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}} \quad \text{if} \quad 1 < d \leq \sqrt{y} \quad \text{or if}$$

$$- \sqrt{y} \leq d < 0,$$

$$\text{(3.24a')} \quad \text{for} \quad \frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1 \quad \text{if} \quad -1 < d < 0.$$

When $\rho_{yy} = 1$, $\beta_x/\beta_\xi^* = \dfrac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{(1 - \rho_{\xi\eta}^2\rho_{xx})}$ for $d\rho_{\xi\eta} \neq 1$

where $0 < \beta_x/\beta_\xi^* < 1$. (See Appendix B.2).

Therefore, when $\rho_{xx} = 1$, $\beta_y/\beta_\eta^* = \dfrac{(1 - \rho_{\xi\eta}^2)\rho_{yy}}{(1 - \rho_{\xi\eta}^2\rho_{yy})}$ for $\dfrac{\rho_{\xi\eta}}{d} \neq 1$

where $0 < \beta_y/\beta_\eta^* < 1$.

When $\rho_{xx} = 1$, $\beta_x/\beta_\xi^* = \dfrac{(1 - \rho_{\xi\eta}^2)(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{yy})(1 - d\rho_{\xi\eta})}$

and for $0 \leq d < 1$ then $\rho_{\xi\eta}^{-(x)} = 0$, $\rho_{\xi\eta}^{+(x)} = d$

for $d > 1$ then $\rho_{\xi\eta}^{-(x)} = 0$, $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$.

Therefore, when $\rho_{yy} = 1$, $\beta_y/\beta_\eta^* = \dfrac{(1 - \rho_{\xi\eta}^2)(1 - \frac{\rho_{\xi\eta}\rho_{xx}}{d})}{(1 - \rho_{\xi\eta}^2\rho_{xx})(1 - \frac{\rho_{\xi\eta}}{d})}$ for $d \neq 0$

which becomes $\beta_y/\beta_\eta^* = \dfrac{(1 - \rho_{\xi\eta}^2)(d - \rho_{\xi\eta}\rho_{xx})}{(1 - \rho_{\xi\eta}^2\rho_{xx})(d - \rho_{\xi\eta})}$ for $d \neq 0$,

and for $d > 1$ (i.e., $0 \leq 1/d < 1$) then $\rho_{\xi\eta}^{-(y)} = 0$,

$\rho_{\xi\eta}^{+(y)} = 1/d$

for $0 < d < 1$ (i.e., $1/d > 1$) then $\rho_{\xi\eta}^{-(y)} = 0$,

$\rho_{\xi\eta}^{+(y)} \notin (-1, +1)$.

When $d = 0$, expression (B.2.5) represents the appropriate expression for β_y/β_η^* i.e.,

$$\beta_y/\beta_\eta^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{yy}\rho_{xx}}{(1 - \rho_{\xi\eta}^2\rho_{yy}\rho_{xx})} \quad \text{for} \quad \rho_{\xi\eta} \neq 0.$$

Applying the property of interchangeability to the result presented in Appendix B.2 (for β_x/β_ξ^* when $a_\xi = 0$, i.e. d is undefined) yields: when $d = 0$, $0 < \beta_y/\beta_\eta^* < 1$. Also note when $d = 0$, β_y/β_η^* is symmetric about the line $\rho_{\xi\eta} = 0$. And as $\rho_{\xi\eta} \to 0$, $\beta_y/\beta_\eta^* \to \rho_{yy}\rho_{xx}$.

Three general categories of distributions of β_y/β_η^* as a function of $\rho_{\xi\eta}$ can be identified. Each of these categories is derived from one of the general categories of distributions of β_x/β_ξ^* as a function of $\rho_{\xi\eta}$ using the property of interchangeability.

Thus the first general category of distributions of β_y/β_η^* as a function of $\rho_{\xi\eta}$ (G.C. I (y)) is based on G.C. I for β_x/β_ξ^*. Recall G.C. I for β_x/β_ξ^* had two subcategories. Subcategory a) of G.C. I consisted of situations where $d > 1$ and $d\rho_{yy} > 1$ for any ρ_{xx}, ρ_{yy} $(\rho_{yy} \neq 1)$. Subcategory b) of G.C. I consisted of situations where $d > 1$ but $d\rho_{yy} < 1$ for any ρ_{xx}, ρ_{yy} $(\rho_{yy} \neq 1)$. Therefore subcategory a') for G.C. I (y) consists of situations where $0 < d < 1$ (i.e., $1/d > 1$) and $0 < d < \rho_{xx}$ (i.e., $\rho_{xx}/d > 1$) for any ρ_{yy}, ρ_{xx} $(\rho_{xx} \neq 1)$ and subcategory b') of G.C. I (y) consists of situations where $0 < d < 1$ (i.e., $1/d > 1$) and $\rho_{xx} < d < 1$ (i.e., $\rho_{xx}/d < 1$).

G.C. II for β_x/β_ξ^* consists of situations where $\sqrt{x} \leq d < 1$ for $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ with $\rho_{yy} \neq 1$. Therefore G.C. II (y) consists of situations where $1 < d \leq \sqrt{y}$ (i.e., $\sqrt{(y)} \leq 1/d < 1$) for $0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$ (i.e., $1/2 - \rho_{xx} < \rho_{yy} \leq 1$) with $\rho_{xx} \neq 1$.

G.C. III for β_x/β_ξ^* consists of two sets of situations; either $0 \leq d < \sqrt{x}$, for any ρ_{xx}, ρ_{yy} $(\rho_{yy} \neq 1)$ or $\sqrt{x} \leq d < 1$ for $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$ with $\rho_{yy} \neq 1$. Therefore G.C. III (y) consists of two sets of situations; either $d > \sqrt{y}$ (i.e. $0 < 1/d < \sqrt{(y)}$) for any ρ_{yy}, ρ_{xx} ($\rho_{xx} \neq 1$) or $1 < d \leq \sqrt{y}$ (i.e. $\sqrt{(y)} \leq 1/d < 1$) for $\frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1$.

Detailed information about the characteristics of each general category as well as information about the situation where $d = 1$ will be presented in the following tables (3.1a through 3.4b). In each table the characteristics of each general category for β_x/β_ξ^* will be presented on the left side of the table and the corresponding information for β_y/β_η^* will be presented on the right side of the table. Following each table the generic graph of β_y/β_η^* as a function of $\rho_{\xi\eta}$ will be presented for the general category displayed in the table.

Although the generic graphs for each category of β_y/β_η^* are identical to the generic graphs for corresponding categories of β_x/β_ξ^* they will still be presented. When the joint categories for β_x/β_ξ^* and β_y/β_η^* are constructed and generic graphs are presented there will be less chance of confusion if each category for β_x/β_ξ^* and β_y/β_η^* is clearly identified along with their generic graphs.

* * * * * * * * * * * * *

Insert Table 3.1a Here

* * * * * * * * * * * * *

Based on results (3.44a') through (3.44l') the generic graph of β_y/β_η^* as a function of $\rho_{\xi\eta}$ for G.C. I (y) subcategory a') has the following general shape:

Table 3.1a. General Categories G.C. I (x) and G.C. I (y). Subcategory a) and a').

GCI(x) for β_x/β_ξ^* subcategory a) $d > 1$, $d\rho_{yy} > 1$, any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$)			GCI(y) for β_y/β_η^* subcategory a') $0 < d < 1$, $0 < d < \rho_{xx}$, any ρ_{yy}, ρ_{xx} ($\rho_{xx} \neq 1$)		
(3.44a)	as $\rho_{\xi\eta} \to -1^+$	$\beta_x/\beta_\xi^* \to 0^+$	(3.44a')	as $\rho_{\xi\eta} \to -1^+$	$\beta_y/\beta_\eta^* \to 0^+$
(3.44b)	for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$	$0 < \beta_x/\beta_\xi^* < 1$	(3.44b')	for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(y)}$	$0 < \beta_y/\beta_\eta^* < 1$
(3.44c)	for $\rho_{\xi\eta} = 0$	$\beta_x/\beta_\xi^* = \rho_{xx}$	(3.44c')	for $\rho_{\xi\eta} = 0$	$\beta_y/\beta_\eta^* = \rho_{yy}$
(3.44d)	for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$	$\beta_x/\beta_\xi^* = 1$	(3.44d')	for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(y)}$	$\beta_y/\beta_\eta^* = 1$
(3.44e)	for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1/d$	$\beta_x/\beta_\xi^* > 1$	(3.44e')	for $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < d$	$\beta_y/\beta_\eta^* > 1$
(3.44f)	as $\rho_{\xi\eta} \to 1/d^-$	$\beta_x/\beta_\xi^* \to +\infty$	(3.44f')	as $\rho_{\xi\eta} \to d^-$	$\beta_y/\beta_\eta^* \to +\infty$
(3.44g)	for $\rho_{\xi\eta} = 1/d$	β_x/β_ξ^* is undefined	(3.44g')	for $\rho_{\xi\eta} = d$	β_y/β_η^* is undefined
(3.44h)	as $\rho_{\xi\eta} \to 1/d^+$	$\beta_x/\beta_\xi^* \to -\infty$	(3.44h')	as $\rho_{\xi\eta} \to d^+$	$\beta_y/\beta_\eta^* \to -\infty$
(3.44i)	for $1/d < \rho_{\xi\eta} < 1/d\rho_{yy}$	$\beta_x/\beta_\xi^* < 0$	(3.44i')	for $d < \rho_{\xi\eta} < d/\rho_{xx}$	$\beta_y/\beta_\eta^* < 0$
(3.44j)	for $\rho_{\xi\eta} = 1/d\rho_{yy}$	$\beta_x/\beta_\xi^* = 0$	(3.44j')	for $\rho_{\xi\eta} = d/\rho_{xx}$	$\beta_y/\beta_\eta^* = 0$
(3.44k)	for $1/d\rho_{yy} < \rho_{\xi\eta} < 1$	$0 < \beta_x/\beta_\xi^* < 1$	(3.44k')	for $d/\rho_{xx} < \rho_{\xi\eta} < 1$	$0 < \beta_y/\beta_\eta^* < 1$
(3.44l)	as $\rho_{\xi\eta} \to +1^-$	$\beta_x/\beta_\xi^* \to 0^+$	(3.44l')	as $\rho_{\xi\eta} \to +1^-$	$\beta_y/\beta_\eta^* \to 0^+$

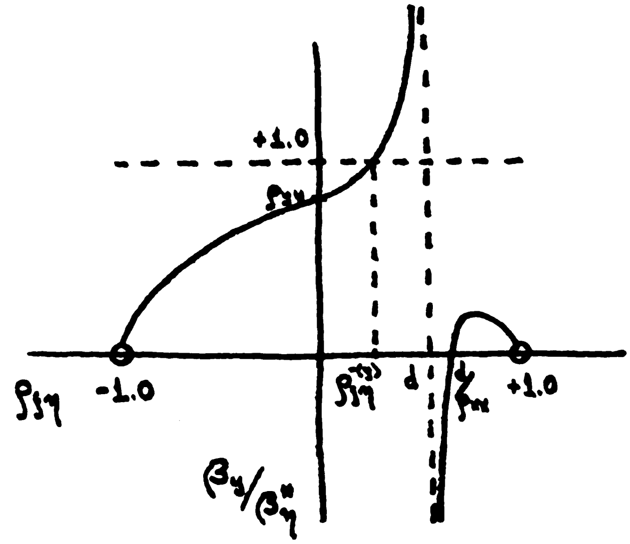


Figure 3.5a

* * * * * * * * * * * * *

Insert Table 3.1b Here

* * * * * * * * * * * * *

The generic graph of β_y/β^*_η as a function of $\rho_{\xi\eta}$ for G.C. I (y) subcategory b') has the following shape:

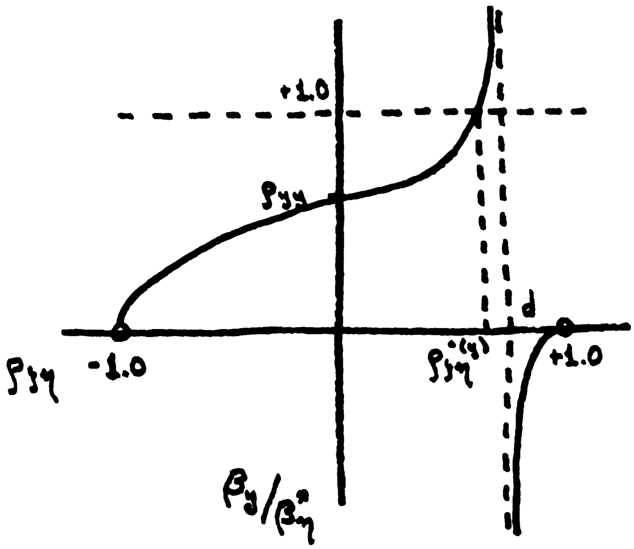


Figure 3.5b

Table 3.1b. General Categories G.C. I (x) and G.C. I (y). Subcategories b) and b').

GCI (x) for β_x/β^*_ξ - subcategory b) $d > 1$, $d\rho_{yy} < 1$, any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$)		GCI (y) for β_y/β^*_η - subcategory b') $0 < d < 1$, $\rho_{xx} < d < 1$, any ρ_{yy}, ρ_{xx} ($\rho_{xx} \neq 1$)	
(3.44a-h) Identical to subcategory a) (Table 3.1a)		(3.44a'-h') Identical to subcategory a') (Table 3.1a)	
(3.44m) for $1/d < \rho_{\xi\eta} < 1$	$\beta_x/\beta^*_\xi < 0$	(3.44m') for $d < \rho_{\xi\eta} < 1$	$\beta_y/\beta^*_\eta < 0$
3.44n) as $\rho_{\xi\eta} \to +1^-$	$\beta_x/\beta^*_\xi \to 0^-$	(3.44n') as $\rho_{\xi\eta} \to +1^-$	$\beta_y/\beta^*_\eta \to 0^-$

* * * * * * * * * * * * *

Insert Table 3.2 Here

* * * * * * * * * * * * *

The generic graph of β_y/β_η^* as a function of $\rho_{\xi\eta}$ for G.C. II (y) has the following shape:

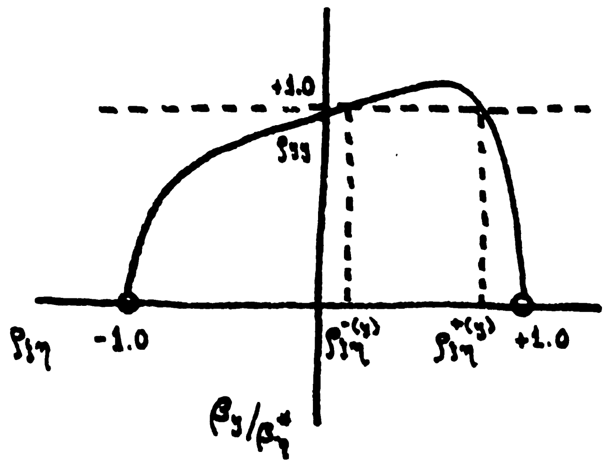


Figure 3.6

* * * * * * * * * * * * *

Insert Table 3.3 Here

* * * * * * * * * * * * *

The generic graph of β_y/β_η^* as a function of $\rho_{\xi\eta}$ for G.C. III (y) has the following shape:

Table 3.2. General Categories GC II (x) and GC II (y).

G.C. II (x) for β_x/β_ξ^* $\Gamma_x \le d < 1,\ 1/2 - \rho_{yy} < \rho_{xx} \le 1,\ \rho_{yy} \ne 1$		G.C. II (y) for β_y/β_η^* $1 < d \le \Gamma_y,\ 0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$	
(3.45a) as $\rho_{\xi\eta} \to -1^+$	$\beta_x/\beta_\xi^* \to 0^+$	(3.45a') as $\rho_{\xi\eta} \to -1^+$	$\beta_y/\beta_\eta^* \to 0^+$
(3.45b) for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$	$0 < \beta_x/\beta_\xi^* < 1$	(3.45b') for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(y)}$	$0 < \beta_y/\beta_\eta^* < 1$
(3.45c) for $\rho_{\xi\eta} = 0$	$\beta_x/\beta_\xi^* = \rho_{xx}$	(3.45c') for $\rho_{\xi\eta} = 0$	$\beta_y/\beta_\eta^* = \rho_{yy}$
(3.45d) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$	$\beta_x/\beta_\xi^* = 1$	(3.45d') for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(y)}$	$\beta_y/\beta_\eta^* = 1$
(3.45e) for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$	$\beta_x/\beta_\xi^* > 1$	(3.45e') for $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(y)}$	$\beta_y/\beta_\eta^* > 1$
(3.45f) for $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$	$\beta_x/\beta_\xi^* = 1$	(3.45f') for $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(y)}$	$\beta_y/\beta_\eta^* = 1$
(3.45g) for $\rho_{\xi\eta}^{+(x)} < \rho_{\xi\eta} < 1$	$0 < \beta_x/\beta_\xi^* < 1$	(3.45g') for $\rho_{\xi\eta}^{+(y)} < \rho_{\xi\eta} < 1$	$0 < \beta_y/\beta_\eta^* < 1$
(3.45h) as $\rho_{\xi\eta} \to +1^-$	$\beta_x/\beta_\xi^* \to 0^+$	(3.45h') as $\rho_{\xi\eta} \to +1^-$	$\beta_y/\beta_\eta^* \to 0^+$

Table 3.3. General Categories G.C. III (x) and G.C. III (y).

G.C. III (x) for β_x/β_ξ^* $0 < d < \sqrt{x}$, any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$) or $\sqrt{x} \leq d < 1$, $0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}}$, $\rho_{yy} \neq 1$	G.C. III (y) for β_y/β_η^* $\sqrt{y} < d$, any ρ_{yy}, ρ_{xx} ($\rho_{xx} \neq 1$) or $1 < d \leq \sqrt{y}$, $\frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1$, $\rho_{xx} \neq 1$
(3.46a) as $\rho_{\xi\eta} \to -1^+$ $\beta_x/\beta_\xi^* \to 0^+$	(3.46a') as $\rho_{\xi\eta} \to -1^+$ $\beta_y/\beta_\eta^* \to 0^+$
(3.46b) for $-1 < \rho_{\xi\eta} < +1$ $0 < \beta_x/\beta_\xi^* < 1$	(3.46n') for $-1 < \rho_{\xi\eta} < +1$ $0 < \beta_y/\beta_\eta^* < 1$
(3.46c) for $\rho_{\xi\eta} = 0$ $\beta_x/\beta_\xi^* = \rho_{xx}$	(3.46c') for $\rho_{\xi\eta} = 0$ $\beta_y/\beta_\eta^* = \rho_{yy}$
(3.46d) as $\rho_{\xi\eta} \to +1^-$ $\beta_x/\beta_\xi^* \to 0^+$	(3.46d') as $\rho_{\xi\eta} \to +1^-$ $\beta_y/\beta_\eta^* \to 0^+$

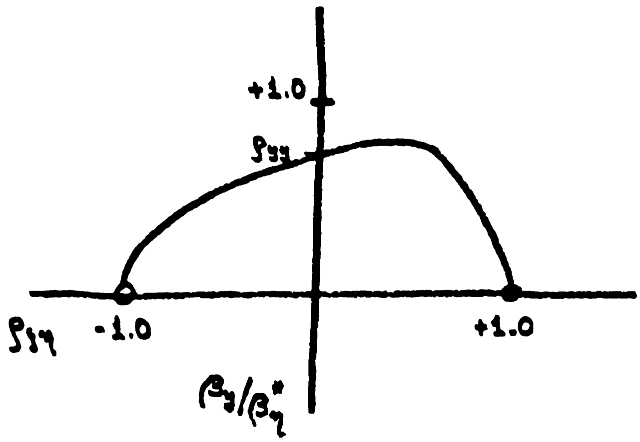


Figure 3.7

* * * * * * * * * * * * * * * * * * *

Insert Tables 3.4a and 3.4b Here

* * * * * * * * * * * * * * * * * * *

When $d = 1$ and $\frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} < 1$ the generic graph of β_y/β_η^* as a function of $\rho_{\xi\eta}$ has the following shape:

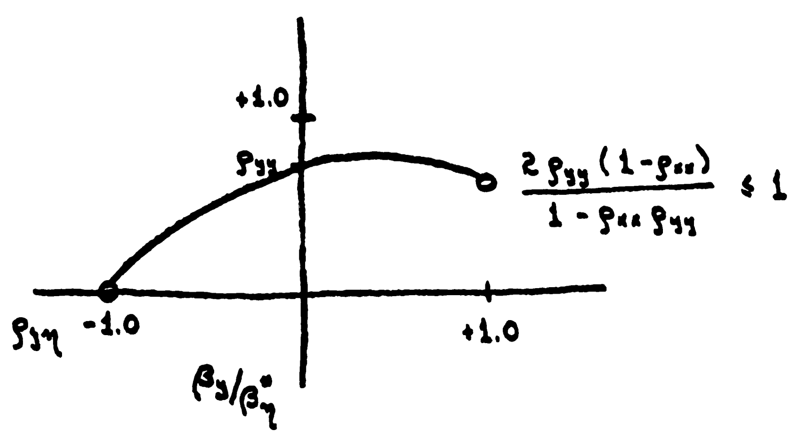


Figure 3.8a

When $d = 1$ and $0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$ the generic graph of β_y/β_η^* as a function of $\rho_{\xi\eta}$ has the following shape:

Table 3.4a. $d = 1$

Table 3.4a. $d = 1$

β_x/β_ξ^* $d = 1,\ 0 < \rho_{xx} \le \frac{1}{2 - \rho_{yy}},\ \rho_{yy} \ne 1$		β_y/β_η^* $d = 1,\ \frac{2\rho_{yy} - 1}{\rho_{yy}} \le \rho_{xx} < 1,\ \rho_{xx} \ne 1$	
(3.47a) as $\rho_{\xi\eta} \to -1^+$	$\beta_x/\beta_\xi^* \to 0^+$	(3.47a') as $\rho_{\xi\eta} \to -1^+$	$\beta_y/\beta_\eta^* \to 0^+$
(3.47b) for $-1 < \rho_{\xi\eta} < +1$	$0 < \beta_x/\beta_\xi^* < 1$	(3.46b') for $-1 < \rho_{\xi\eta} < +1$	$0 < \beta_y/\beta_\eta^* < 1$
(3.47c) for $\rho_{\xi\eta} = 0$	$\beta_x/\beta_\xi^* = \rho_{xx}$	(3.47c') for $\rho_{\xi\eta} = 0$	$\beta_y/\beta_\eta^* = \rho_{yy}$
(3.47d) as $\rho_{\xi\eta} \to +1^-$	$\beta_x/\beta_\xi^* \to$	(3.47d') as $\rho_{\xi\eta} \to +1^-$	$\beta_y/\beta_\eta^* \to$
	$\frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} \le 1$		$\frac{2\rho_{yy}(1 - \rho_{xx})}{1 - \rho_{xx}\rho_{yy}} \le 1$

Table 3.4

Table 3.4b. d = 1

β_x/β_ξ^* $d = 1,\ \frac{1}{2 - \rho_{yy}} < \rho_{xx} \le 1,\ \rho_{yy} \ne 1$			β_y/β_η^* $d = 1,\ 0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$		
(3.48a)	as $\rho_{\xi\eta} \to -1^+$	$\beta_x/\beta_\xi^* \to 0^+$	(3.48a')	as $\rho_{\xi\eta} \to -1^+$	$\beta_y/\beta_\eta^* \to 0^+$
(3.48b)	for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)}$	$0 < \beta_x/\beta_\xi^* < 1$	(3.48b')	for $-1 < \rho_{\xi\eta} < \rho_{\xi\eta}^{-(y)}$	$0 < \beta_y/\beta_\eta^* < 1$
(3.48c)	for $\rho_{\xi\eta} = 0$	$\beta_x/\beta_\xi^* = \rho_{xx}$	(3.48c')	for $\rho_{\xi\eta} = 0$	$\beta_y/\beta_\eta^* = \rho_{yy}$
(3.48d)	for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$	$\beta_x/\beta_\xi^* = 1$	(3.48d')	for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(y)}$	$\beta_y/\beta_\eta^* = 1$
(3.48e)	for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1$	$\beta_x/\beta_\xi^* > 1$	(3.48e')	for $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < 1$	$\beta_y/\beta_\eta^* > 1$
(3.48f)	as $\rho_{\xi\eta} \to +1^-$	$\beta_x/\beta_\xi^* \to$	(3.48f')	as $\rho_{\xi\eta} \to +1^-$	$\beta_y/\beta_\eta^* \to$
		$\frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} > 1$			$\frac{2\rho_{yy}(1 - \rho_{xx})}{1 - \rho_{xx}\rho_{yy}} > 1$

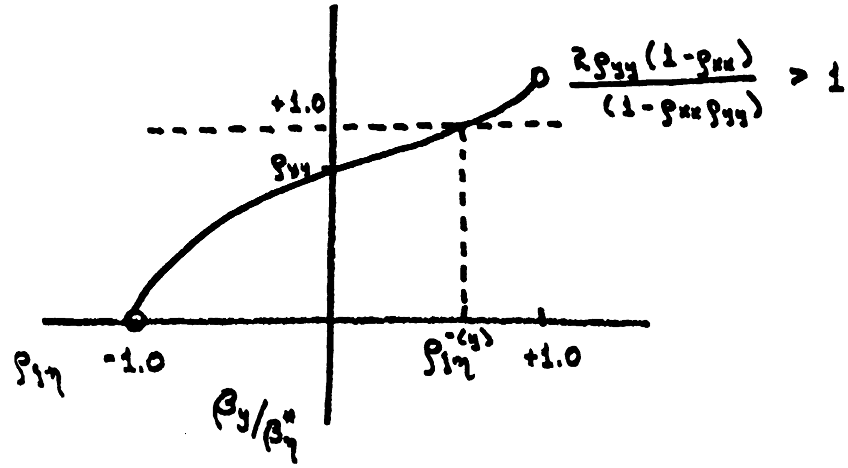


Figure 3.8b

Examples of graphs in each of the general categories of x and y for fixed values of d, ρ_{xx} and ρ_{yy} will be presented below in section VII.

VI. <u>Additional Algebraic Results Involving Both</u> β_x/β_ξ^* and β_y/β_η^*

Prior to combining the general categories of β_x/β_ξ^* with the general categories for β_y/β_η^* to form joint general categories some additional algebriac derivations involving both β_x/β_ξ^* and β_y/β_η^* are needed. In this section three algebraic derivations are to be presented. The first two derivations will be used in subsequent work but also represent interesting results by themselves. The third derivation provides an algebraic justification for a result which can be noted from the generic graphs for each general category of β_x/β_ξ^* and β_y/β_η^*.

<u>Derivation 1</u>: Show that when $\beta_x/\beta_\xi^* = 1$, then $\beta_y/\beta_\eta^* = \rho_{yy}$ and when $\beta_y/\beta_\eta^* = 1$, then $\beta_x/\beta_\xi^* = \rho_{xx}$.

Proof: From either (3.18b) or (3.19b) $\beta_x/\beta_\xi^* = 1 \Leftrightarrow Q = 0$ where $Q = \rho_{\xi\eta}^2 \rho_{xx}(1 - \rho_{yy}) - d\rho_{\xi\eta}(1 - \rho_{xx}\rho_{yy}) + (1 - \rho_{xx})$ from (3.17).

Therefore consider $\beta_y/\beta_\eta^* = \rho_{yy}$. Modifying expression (3.14) for β_y/β_η^* with $1/d = d_\eta$ yields:

$$\beta_y/\beta_\eta^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{yy}(d - \rho_{\xi\eta}\rho_{xx})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(d - \rho_{\xi\eta})}$$

and

$$\beta_y/\beta_\eta^* = \rho_{yy} \Leftrightarrow (1 - \rho_{\xi\eta}^2)\rho_{yy}(d - \rho_{\xi\eta}\rho_{xx})$$

$$= \rho_{yy}(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(d - \rho_{\xi\eta})$$

$$\Leftrightarrow \rho_{yy}(d - d\rho_{\xi\eta}^2 - \rho_{\xi\eta}\rho_{xx} + \rho_{\xi\eta}^3\rho_{xx})$$

$$= \rho_{yy}(d - d\rho_{\xi\eta}^2\rho_{xx}\rho_{yy} - \rho_{\xi\eta} + \rho_{\xi\eta}^3\rho_{xx}\rho_{yy})$$

$$\Leftrightarrow \rho_{yy}(\rho_{\xi\eta}^3\rho_{xx} - \rho_{\xi\eta}^3\rho_{xx}\rho_{yy} + d\rho_{\xi\eta}^2\rho_{xx}\rho_{yy} - d\rho_{\xi\eta}^2 + \rho_{\xi\eta} - \rho_{\xi\eta}\rho_{xx}) = 0$$

$$\Leftrightarrow \rho_{yy}\rho_{\xi\eta}(\rho_{\xi\eta}^2\rho_{xx}(1 - \rho_{yy}) - d\rho_{\xi\eta}(1 - \rho_{xx}\rho_{yy}) + (1-\rho_{xx})) = 0$$

$$\Leftrightarrow \rho_{yy}\rho_{\xi\eta}Q = 0 .$$

Therefore $\beta_y/\beta_\eta^* = \rho_{yy}$ if $\rho_{\xi\eta} = 0$ or $Q = 0$ since $\rho_{yy} > 0$. But if $\beta_x/\beta_\xi^* = 1$, then $Q = 0$ and thus $\beta_y/\beta_\eta^* = \rho_{yy}$.

(3.52) Therefore if $\beta_x/\beta_\xi^* = 1$, then $\beta_y/\beta_\eta^* = \rho_{yy}$.

And by the property of interchangeability of x and y:

(3.53) if $\beta_y/\beta_\eta^* = 1$, then $\beta_x/\beta_\xi^* = \rho_{xx}$.

Note that the converses of (3.52) and (3.53) are in general not true, since if $\rho_{\xi\eta} = 0$, $\beta_x/\beta_\xi^* = \rho_{xx}$ and $\beta_y/\beta_\eta^* = \rho_{yy}$. For

$\rho_{\xi\eta} \neq 0$, then $\beta_y/\beta_\eta^* = \rho_{yy} \Leftrightarrow \beta_x/\beta_\xi^* = 1$ and $\beta_x/\beta_\xi^* = \rho_{xx} \Leftrightarrow \beta_y/\beta_\eta^* = 1$.

Derivation 2: Show that β_x/β_ξ^* and β_y/β_η^* cannot both be greater than one for any situation, i.e., there exist no values of d, $\rho_{\xi\eta}$, ρ_{xx} and ρ_{yy} such that $\beta_x/\beta_\xi^* > 1$ and $\beta_y/\beta_\eta^* > 1$. For any given situation at most one of the observed weighting coefficients will be an overestimate of the corresponding latent weighting coefficient.

The procedure will first locate for β_x/β_ξ^* and β_y/β_η^* separately the general categories and then the regions within the general categories (from Tables 3.1a through 3.4b) where each ratio is greater than one. Then the regions will be compared to see if there are any regions where both are greater than one. If such regions exist then more detailed algebraic work will be performed to examine the situations in each region.

$\beta_x/\beta_\xi^* > 1$

(3.54a) 1. From Table 3.1a for G.C. I, when $d > 1$ for any ρ_{xx}, ρ_{yy} and $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1/d$ then $\beta_x/\beta_\xi^* > 1$.

(3.54b) 2. From Table 3.2 for G.C. II, when $\sqrt{x} \leq d < 1$, $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ and $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$ then $\beta_x/\beta_\xi^* > 1$.

(3.54c) 3. From Table 3.4b when $d = 1$, $1/2 - \rho_{yy} < \rho_{xx} \leq 1$ and $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1$, then $\beta_x/\beta_\xi^* > 1$.

$\beta_y/\beta_\eta^* > 1$

(3.55a) 1. From Table 3.1a for G.C. I (y) when $0 < d < 1$, any ρ_{xx}, ρ_{yy} and $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < d$ then $\beta_y/\beta_\eta^* > 1$.

(3.55b) 2. From Table 3.2 for G.C. II (y) when $1 < d \leq \sqrt{y}$,

$0 < \rho_{xx} < \dfrac{2\rho_{yy} - 1}{\rho_{yy}}$ and $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(y)}$ then

$\beta_y/\beta_\eta^* > 1$.

(3.55c) 3. From Table 3.4b when $d = 1$, $0 < \rho_{xx} < \dfrac{2\rho_{yy} - 1}{\rho_{yy}}$ and

and $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < 1$ then $\beta_y/\beta_\eta^* > 1$.

Note 1: $1/2 - \rho_{yy} \geq \dfrac{2\rho_{yy} - 1}{\rho_{yy}}$ for all ρ_{yy}, $0 < \rho_{yy} \leq 1$ with equality only when $\rho_{yy} = 1$. (See Appendix B.7 for proof.)

Note 2: $\sqrt{x} \leq 1$ for all ρ_{xx}, ρ_{yy} such that $\rho_{xx}\rho_{yy} \neq 1$ with equality for $\rho_{xx} = \dfrac{1}{2 - \rho_{yy}}$. (See Appendix B.3.)

Note 3: $\sqrt{y} \geq 1$ for all ρ_{xx}, ρ_{yy} such that $\rho_{xx}\rho_{yy} \neq 1$ with equality for $\rho_{xx} = \dfrac{2\rho_{yy} - 1}{\rho_{yy}}$. (By expression (3.51).)

Therefore the only regions where both $\beta_x/\beta_\xi^* > 1$ and $\beta_y/\beta_\eta^* > 1$ occur within the region are:

(3.56) when $\sqrt{x} \leq d < 1$ and $1/2 - \rho_{yy} < \rho_{xx} \leq 1$:

$\beta_x/\beta_\xi^* > 1$ if $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$ [from (3.54b)]

and $\beta_y/\beta_\eta^* > 1$ if $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < d$ [from (3.55a)].

(3.57) and, when $1 < d \le \sqrt{y}$ and $0 < \rho_{xx} < \dfrac{2\rho_{yy} - 1}{\rho_{yy}}$:

$$\beta_x/\beta_\xi^* > 1 \quad \text{if} \quad \rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < 1/d \qquad \text{[from (3.54a)]}$$

$$\text{and} \quad \beta_y/\beta_\eta^* > 1 \quad \text{if} \quad \rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(y)} \qquad \text{[from (3.55b)].}$$

Note however, that this second region can be found from the first region by use of the property of interchangeability. Therefore it is necessary to examine in detail only the situations in the first region, since all results for situations in the first region can be extended to corresponding situations in the second region.

Thus let $\sqrt{x} \le d < 1$ and $1/2 - \rho_{yy} < \rho_{xx} \le 1$. For any values of ρ_{yy}, ρ_{xx} and d in this region the question of interest is whether there exists a value $\rho'_{\xi\eta}$, $-1 < \rho'_{\xi\eta} < +1$ such that $\rho_{\xi\eta}^{-(x)} < \rho'_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$ and $\rho_{\xi\eta}^{-(y)} < \rho'_{\xi\eta} < d$.

In this region ($\sqrt{x} \le d < 1$ and $1/2 - \rho_{yy} < \rho_{xx} < 1$), $\rho_{\xi\eta}^{+(x)} < d$. (From expression (B.8.4b) in Appendix B.8). Therefore, if $\rho'_{\xi\eta}$ exists, that is if $\rho_{\xi\eta}^{-(x)} < \rho'_{\xi\eta} < \rho_{\xi\eta}^{+(x)} < d$ and $\rho_{\xi\eta}^{-(y)} < \rho'_{\xi\eta} < d$, then $\rho_{\xi\eta}^{-(y)} < \rho_{\xi\eta}^{+(x)}$ or else the two intervals $(\rho_{\xi\eta}^{-(x)}, \rho_{\xi\eta}^{+(x)})$ and $(\rho_{\xi\eta}^{-(y)}, d)$ will have no common values.

Note: for $\rho_{\xi\eta} < d$, $\beta_y/\beta_\eta^* \le 1$ occurs only when

$$\rho_{\xi\eta} \le \rho_{\xi\eta}^{-(y)} \text{ [from Table 3.1a for G.C. I (y)].}$$

Let $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$, then $\beta_x/\beta_\xi^* = 1$ (from (3.45f)) and thus $\beta_y/\beta_\eta^* = \rho_{yy} \le 1$ (from (3.52)). Therefore since $\beta_y/\beta_\eta^* \le 1$ for $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$, $\rho_{\xi\eta}^{+(x)} \le \rho_{\xi\eta}^{-(y)}$, hence $\rho_{\xi\eta}^{-(y)} \not< \rho_{\xi\eta}^{+(x)}$. Thus there does not exist a $\rho'_{\xi\eta}$ such that $\rho'_{\xi\eta} \in (\rho_{\xi\eta}^{-(x)}, \rho_{\xi\eta}^{+(x)})$ and $\rho'_{\xi\eta} \in (\rho_{\xi\eta}^{-(y)}, d)$.

Therefore there do not exist any combinations of values of $\rho_{\xi\eta}$, ρ_{yy}, ρ_{xx} and d, with $\sqrt{x} \leq d < 1$ and $1/2 - \rho_{yy} < \rho_{xx} < 1$ [region (3.56)] such that $\beta_x/\beta^*_\xi > 1$ and $\beta_y/\beta^*_\eta > 1$ simultaneously.

By the property of interchangeability of x and y this result for the region identified as (3.56) also holds for the region (3.57).

Therefore it is not possible to find a set of values of $\rho_{\xi\eta}$, ρ_{yy}, ρ_{xx} and d so that $\beta_x/\beta^*_\xi > 1$ and $\beta_y/\beta^*_\eta > 1$ simultaneously. That is in the two category, two predictor case it is not possible for both of the observed weighting coefficients to be overestimates of the latent weighting coefficients for the same combination of values of $\rho_{\xi\eta}$, ρ_{yy}, ρ_{xx} and d.

<u>Derivation 3</u>: Show that both $\beta_x/\beta^*_\xi < 1$ and $\beta_y/\beta^*_\eta < 1$ when $d\rho_{\xi\eta} < 0$. It is important to note here that this result represents a sufficient condition only.

Consider β_x/β^*_ξ from (3.12), i.e.,

$$\beta_x/\beta^*_\xi = \frac{(1 - \rho^2_{\xi\eta})\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})} .$$

Note that β_x/β^*_ξ can be considered as the product of two ratios R_1 and R_2 where $R_1 = \dfrac{(1 - \rho^2_{\xi\eta})\rho_{xx}}{(1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy})}$ and $R_2 = \dfrac{(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - d\rho_{\xi\eta})}$. Thus $\beta_x/\beta^*_\xi = R_1 \cdot R_2$. The examination of β_x/β^*_ξ for this derivation will proceed by examining R_1 and R_2 separately and then combining the results.

Work with R_1 first. There are three situations to examine. For what values of $\rho_{\xi\eta}$, ρ_{xx} and ρ_{yy} will R_1 (1) be greater than one $(R_1 > 1\ ?)$, (2) be equal to one $(R_1 = 1\ ?)$ and (3) be less than one $(R_1 < 1\ ?)$?

1.) $R_1 > 1$?

$$R_1 > 1 \Leftrightarrow \frac{(1 - \rho^2_{\xi\eta})\rho_{xx}}{1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy}} > 1$$

$$\Leftrightarrow (1 - \rho^2_{\xi\eta})\rho_{xx} > 1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy} \quad \text{since}$$

$$1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy} > 0$$

$$\Leftrightarrow \rho_{xx} - \rho^2_{\xi\eta}\rho_{xx} > 1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy}$$

$$\Leftrightarrow 0 > \rho^2_{\xi\eta}\rho_{xx} - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy} + (1 - \rho_{xx})$$

$$0 > \rho^2_{\xi\eta}\rho_{xx}(1 - \rho_{yy}) + (1 - \rho_{xx})$$

but note $\rho^2_{\xi\eta} \geq 0$, $\rho_{xx} > 0$, $1 - \rho_{yy} \geq 0$, $1 - \rho_{xx} \geq 0$.

Therefore $0 \not> \rho^2_{\xi\eta}\rho_{xx}(1 - \rho_{yy}) + (1 - \rho_{xx})$

(3.58a) and $R_1 \not> 1$

2.) $R_1 \leq 1$? $\quad R_1 \leq 1 \Leftrightarrow \dfrac{(1 - \rho^2_{\xi\eta})\rho_{xx}}{(1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy})} \leq 1$

using algebra from 1) above

(3.58b) $R_1 \leq 1 \Leftrightarrow 0 \leq \rho^2_{\xi\eta}\rho_{xx}(1 - \rho_{yy}) + (1 - \rho_{xx})$

true for all values of $\rho_{\xi\eta}$, ρ_{yy}, and ρ_{xx}.

Note: $R_1 = 1 \Leftrightarrow 0 = \rho^2_{\xi\eta}\rho_{xx}(1 - \rho_{yy}) + (1 - \rho_{xx})$.

(3.58c) Thus $R_1 = 1$ only if $\rho_{xx} = 1$ and either $\rho_{yy} = 1$ or

$\rho_{\xi\eta} = 0$.

Note: if $d = 0$,

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}} = R_1.$$

Work now with R_2. Consider only situations where $R_2 < 1$. Then $\beta_x/\beta_\xi^* = R_1 \cdot R_2 < 1$ by (3.58b). Note, for some situations where $R_2 > 1$, $\beta_x/\beta_\xi^* > 1$ and for other situations $\beta_x/\beta_\xi^* < 1$. The purpose of this derivation is to produce a sufficient condition for $\beta_x/\beta_\xi^* < 1$ and not reproduce the exhaustive study which has been done above in subsection IV for β_x/β_ξ^* and in subsection V for β_y/β_η^*.

The situation to be examined then is for what values of $\rho_{\xi\eta}$, ρ_{yy} and d will R_2 be less than one $(R_2 < 1\ ?)$?

$$R_2 < 1? \quad R_2 < 1 \Leftrightarrow \frac{1 - d\rho_{\xi\eta}\rho_{yy}}{1 - d\rho_{\xi\eta}} < 1$$

1.) if $1 - d\rho_{\xi\eta} > 0 \quad d\rho_{\xi\eta} < 1$

then $R_2 < 1 \Leftrightarrow 1 - d\rho_{\xi\eta}\rho_{yy} < 1 - d\rho_{\xi\eta}$

$$\Leftrightarrow d\rho_{\xi\eta} - d\rho_{\xi\eta}\rho_{yy} < 0$$

$$\Leftrightarrow d\rho_{\xi\eta}(1 - \rho_{yy}) < 0$$

$$\Leftrightarrow d\rho_{\xi\eta} < 0 \quad \text{for} \quad \rho_{yy} \neq 1.$$

Since $R_2 < 1$ if $d\rho_{\xi\eta} < 1$ and $d\rho_{\xi\eta} < 0$. Then

(3.59a) $$R_2 < 1 \ \text{if} \ d\rho_{\xi\eta} < 0.$$

or 2.) if $1 - d\rho_{\xi\eta} < 0 \Leftrightarrow d\rho_{\xi\eta} > 1$

then $R_2 < 1 \Leftrightarrow 1 - d\rho_{\xi\eta}\rho_{yy} > 1 - d\rho_{\xi\eta}$

$$\Leftrightarrow d\rho_{\xi\eta}(1 - \rho_{yy}) > 0$$

$$\Leftrightarrow d\rho_{\xi\eta} > 0 \quad \text{for} \quad \rho_{yy} \neq 1 .$$

Since $R_2 < 1$ if $d\rho_{\xi\eta} > 1$ and $d\rho_{\xi\eta} > 0$. Then

(3.59b) $\quad R_2 < 1$ if $d\rho_{\xi\eta} > 1.$

Therefore, if $d\rho_{\xi\eta} < 0$ or $d\rho_{\xi\eta} > 1$ then $R_2 < 1$. Since by (3.58b) $R_1 \leq 1$ for all values of $\rho_{\xi\eta}$, ρ_{xx} and ρ_{yy} and since $\beta_x/\beta_\xi^* = R_1 \cdot R_2$ then $\beta_x/\beta_\xi^* < 1.$

(3.60) If $d\rho_{\xi\eta} < 0$ or $d\rho_{\xi\eta} > 1$ then $\beta_x/\beta_\xi^* < 1$. For $d\rho_{\xi\eta}$ to be less than zero, d and $\rho_{\xi\eta}$ must have opposite signs.

By the property of interchangeability of x and y (3.60) becomes for $d \neq 0$:

(3.60') If $\rho_{\xi\eta}/d < 0$ or $\rho_{\xi\eta}/d > 1$ then $\beta_y/\beta_\eta^* < 1.$

For $\rho_{\xi\eta}/d$ to be less than zero, d and $\rho_{\xi\eta}$ must have opposite signs, i.e., $d\rho_{\xi\eta} < 0$.

Therefore if d and $\rho_{\xi\eta}$ have opposite signs ($d\rho_{\xi\eta} < 0$) then both β_x/β_ξ^* and β_y/β_η^* will be less than one. Note that this is a sufficient condition only.

As noted above $d\rho_{\xi\eta}$ can also be interpreted as the ratio of the slope of the pooled within categories regression line of ξ on η over the slope of the between categories line joining the mid-points of the distributions in each category of ξ and η. This interpretation is presented in Appendix B.9 along with a presentation of results which corroborates the results in (3.60).

Using the ratio of slopes interpretation of $d\rho_{\xi\eta}$ in a situation where $d\rho_{\xi\eta} < 0$ indicates that the slope of the pooled within categories regression line has the opposite sign as the slope of the between categories line, i.e., the direction of the relationship between ξ and η as expressed by the pooled within categories regression line is the opposite of the direction of the relationship between category means as expressed by the slope of the between categories line.

Using this interpretation for $d\rho_{\xi\eta} < 0$ there may be some question about whether $d\rho_{\xi\eta}$ can be less than zero in practice. The concern at issue here is very similar to the concern involved in the study of ecological correlation where the interest is on using a correlation between group means (similar to the between categories situation here) to estimate either a total group or pooled within groups correlation (similar to the pooled within categories situation here). The study of ecological correlation produced some results which are applicable here as well.

If the groups (in this case two groups [categories]) are independent samples from the same population then the correlation in the population of individuals and the correlation in the population formed by the sampling distribution of the group means for a given sample size are identical. In this case the between groups correlation and the pooled within groups correlations are less likely to have different signs than the same sign. For the quantal response situation, if the individuals are arbitrarily assigned to each category of the criterion on the basis of a random sampling from a single population of individuals then the relationship between ξ and η as measured by the slope of the pooled within categories regression line will be more likely to have the same sign as the slope of the between categories line joining the category midpoints than to have a different sign. That is, $d\rho_{\xi\eta}$ is more likely to be greater than zero rather than less than zero.

For most quantal response situations it would seem unlikely and even contrary to the intents of quantal response procedures to arbitrarily define categories of a criterion as multiple random samples from some population. For most quantal response situations assignment of a subject to a category of the criterion is based on distinct and non-overlapping membership criteria, e.g., health status of an experimental animal (e.g., a rat) following an administration of an experimental drug (i.e., alive or dead) or group affiliation (Democrat, Republican, Independent, etc.).

For these types of situations it is not reasonable to assume generally that the between groups relationship will have the same sign as the within groups situation. That is, for a given situation there is no a priori basis to assume that $d\rho_{\xi\eta} > 0$ with any more likelihood than $d\rho_{\xi\eta} < 0$.

Although in many situations the ratio of slopes ($d\rho_{\xi\eta}$) will be positive there will exist situations where the ratio is negative. A hypothetical example can be constructed to illustrate thatthe ratio of slopes can be negative.

Consider two elementary schools. Each school represents a category. The mathematics curriculum of school 1 heavily emphasizes work on the basics of computation through rote memory and repeated drill under the assumption that students must have a sound basis of computational skills prior to tackling more advanced mathematics topics. The mathematics curriculum of school 2 emphasizes training in approaches to problem solving and the conceptualization of mathematical problems under the assumption that it is important to be able to identify approaches to the solution of problems and that specific computational skills can be more efficiently learned when the student is confronted with the need to compute as part of the solution of a problem.

The predictor variables in this situation are the mathematics computation subscale and the mathematics application subscale of some standardized test. It is reasonable to expect that there is a similar positive relationship between subscale scores on mathematics computation and application within each school, since factors such as general mathematics ability and motivation are likely to be underlying factors related to performance on both subscales within each school. Therefore, there is a positive within categories relationship between the predictor variables.

However, it is also reasonable to expect that students in school 1 will do better on the mathematics computation subscale than students in school 2. While, students from school 2 can be expected to do better on the mathematics application subscale than students in school 1. Therefore, the slope of the line which joins the midpoints of the distributions of the two schools on the computation and application subscale can be expected to be negative. That is, there is a negative between groups relationship. Therefore, the ratio of the pooled within categories slope to the between categories slope is negative.

If this or a similar situation were to be analyzed using a quantal response procedure then $d\rho_{\xi\eta} < 0$ and β_x and β_y will underestimate β_ξ^* and β_η^* respectively. Note if the ratio of slopes based solely on the observed predictor is negative then the ratio of slopes for the latent predictors ($d\rho_{\xi\eta}$) will also be negative. This follows since the errors of measurement will not affect the value of the slope of the line between the midpoints of the categories but will attenuate the value of the pooled within categories slope based on latent predictors (see Appendix B.9 for details). Thus the magnitude of the ratio based on observed predictors will be smaller than the magnitude of the ratio based on latent predictors but the signs of the two ratios will be identical.

VII. Joint General Categories of Distributions for β_x/β_ξ^* and β_y/β_η^* Together

In this section the results from section IV (for categories of β_x/β_ξ^* versus $\rho_{\xi\eta}$ for fixed values of d, ρ_{xx} and ρ_{yy}), section V (for categories of β_y/β_η^* versus $\rho_{\xi\eta}$ for fixed values of d, ρ_{xx} and ρ_{yy}) and section VI (Algebriac results for β_x/β_ξ^* and β_y/β_η^* together) are combined to derive joint general categories of distributions for β_x/β_ξ^* and β_y/β_η^* together for the same situations. For each joint general category, the generic graphs of β_x/β_ξ^* and β_y/β_η^* will be displayed to provide a visual indication of the generic shape of the distributions within the category. In addition actual graphs for specific situations (i.e., values of d, ρ_{xx} and ρ_{yy}) within each category will be referenced. For a more detailed indication of the shape of either β_x/β_ξ^* or β_y/β_η^* within any joint general category see the information from the tables in section IV or section V which applies.

An example of the notation for the joint general categories is: G.C. I (x,y). The x and the y in the parenthesis indicate that it is a joint general category involving both β_x/β_ξ^* and β_y/β_η^*.

G.C. I (x,y). When $0 < d < \sqrt{x}$ (Recall, $\sqrt{x}$ is defined by (3.22)), for any ρ_{xx}, ρ_{yy} ($\rho_{xx}, \rho_{yy} \neq 1$) or when $\sqrt{x} \leq d < 1$ for $0 < \rho_{xx} \leq 1/2 - \rho_{yy}$, G.C. III for β_x/β_ξ^* and G.C. I (y) for β_y/β_η^* apply. Ignoring the two subcategories of G.C. I (y), the generic graphs for this category of situations are:

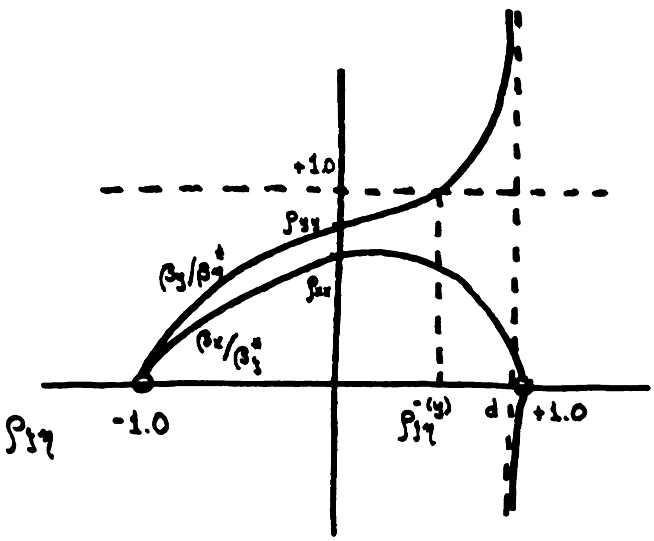


Figure 3.9

The exact shape of β_y/β_η^* when $d < \rho_{\xi\eta} < 1$ depends on whether $0 < d < \rho_{xx}$ or $\rho_{xx} < d < 1$. See section V for G.C. I (y) above for details.

Figure 3.11a through 3.11d provide examples for 4 specific situations in G.C. I (x,y).

G.C. II (x,y). When $\sqrt{x} \leq d < 1$ for $1/2 - \rho_{yy} < \rho_{xx} < 1$, G.C. II for β_x/β_ξ^*, G.C. I (y) for β_y/β_η^* and section VI results apply. Again ignoring the two subcategories of β_y/β_η^*, the generic graphs for this category of situations are:

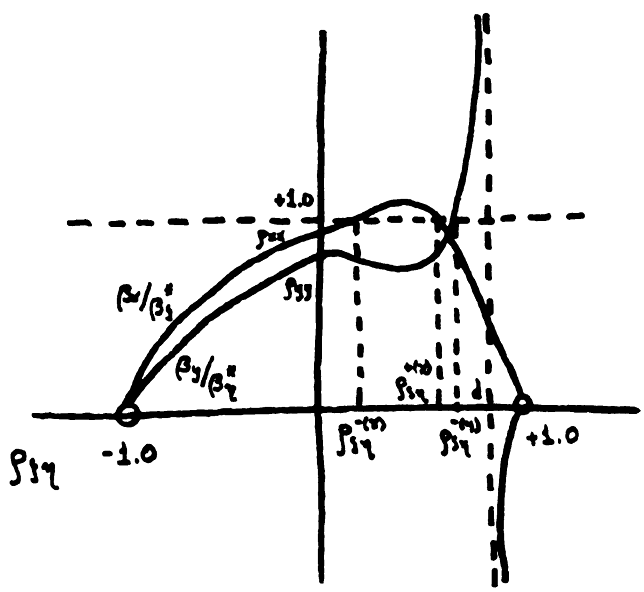


Figure 3.10

The exact shape of β_y/β_η^* when $d < \rho_{\xi\eta} < 1$ depends on whether $0 < d < \rho_{xx}$ or $\rho_{xx} < d < 1$. See section V for G.C. I (y) above the details.

Figures 3.12a through 3.12c provide examples for 3 specific situations in G.C. II (X,y).

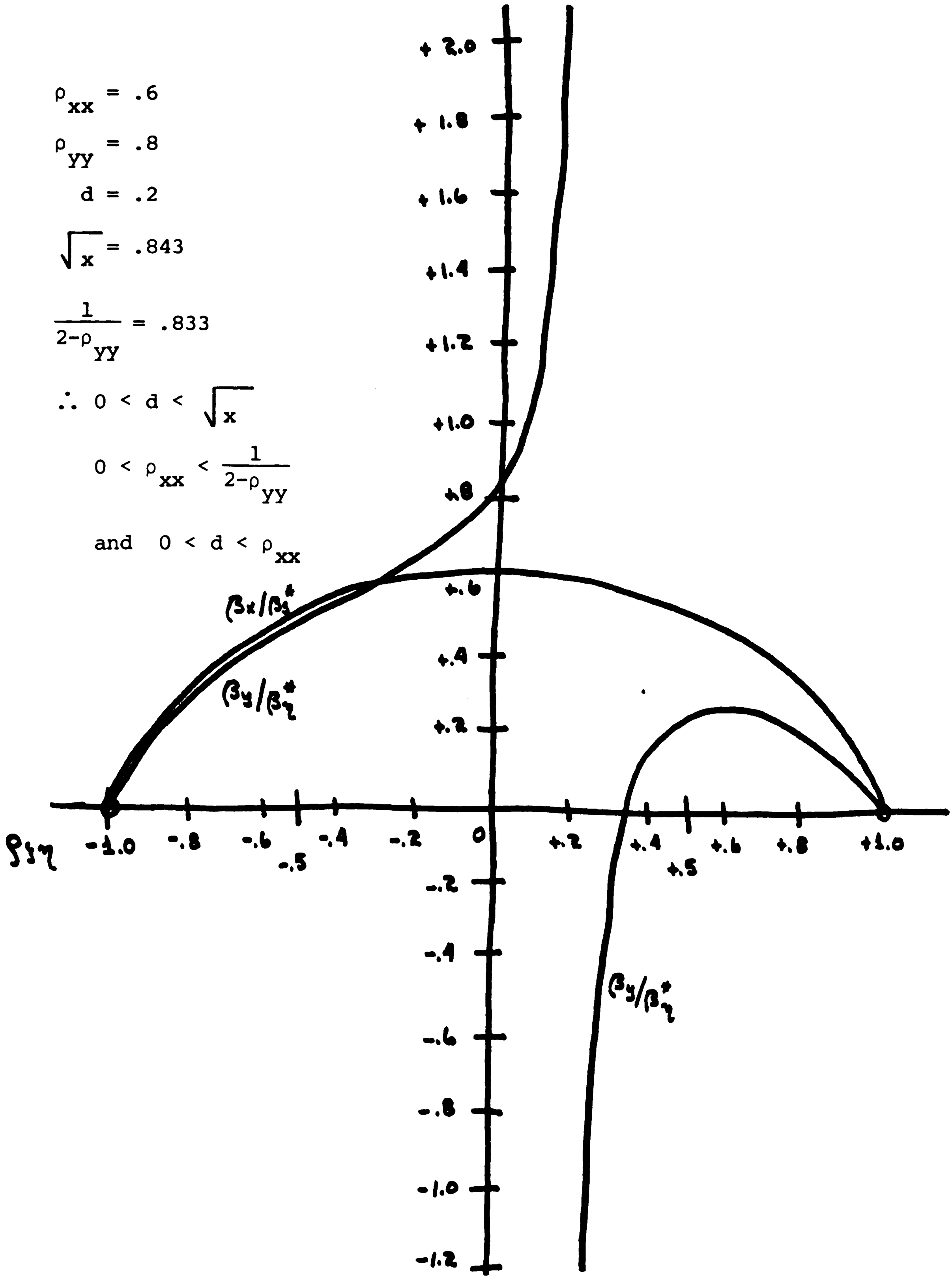


Figure 3.11a. G.C. I (x.y).

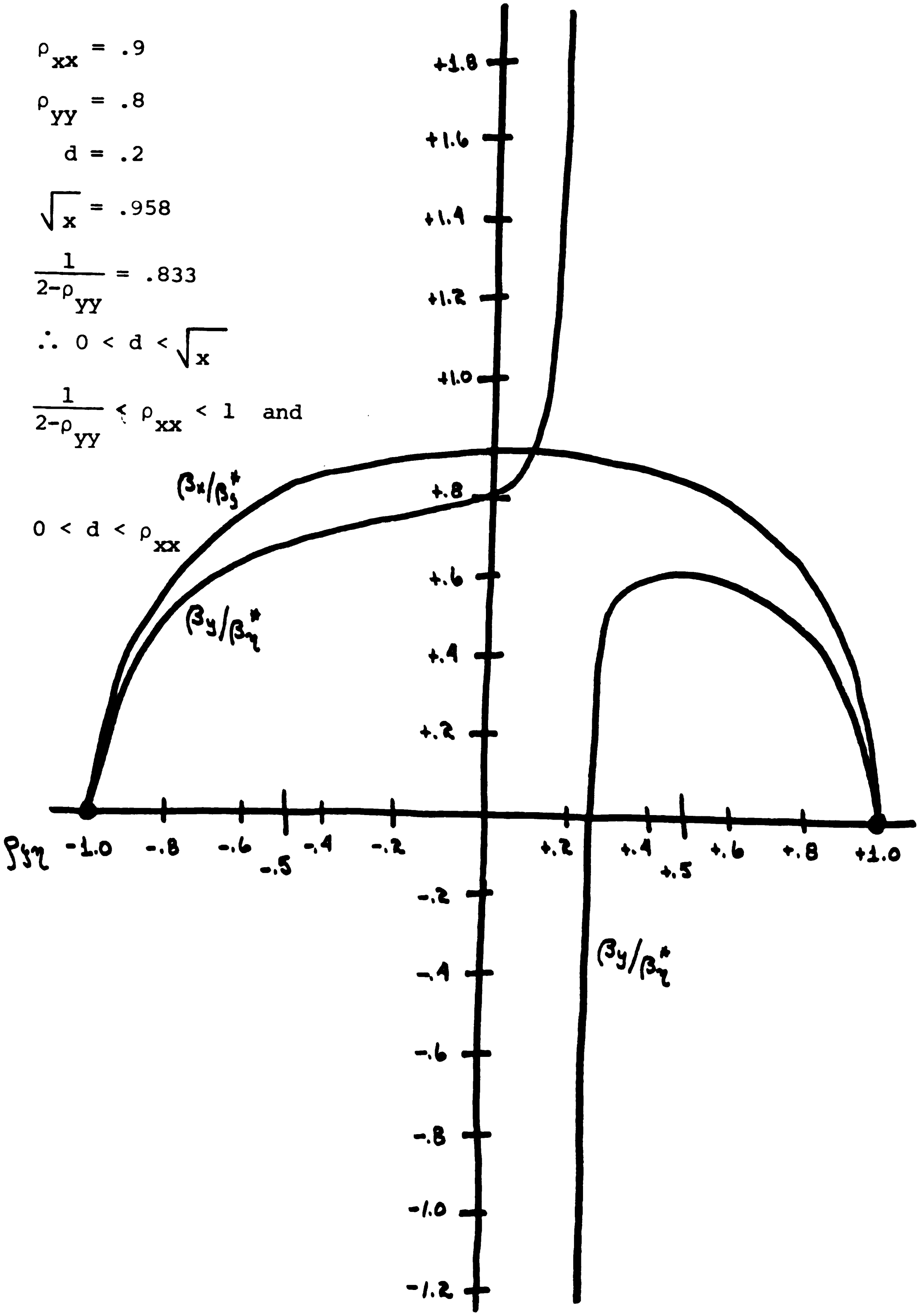


Figure 3.11b. G.C. I (x,y).

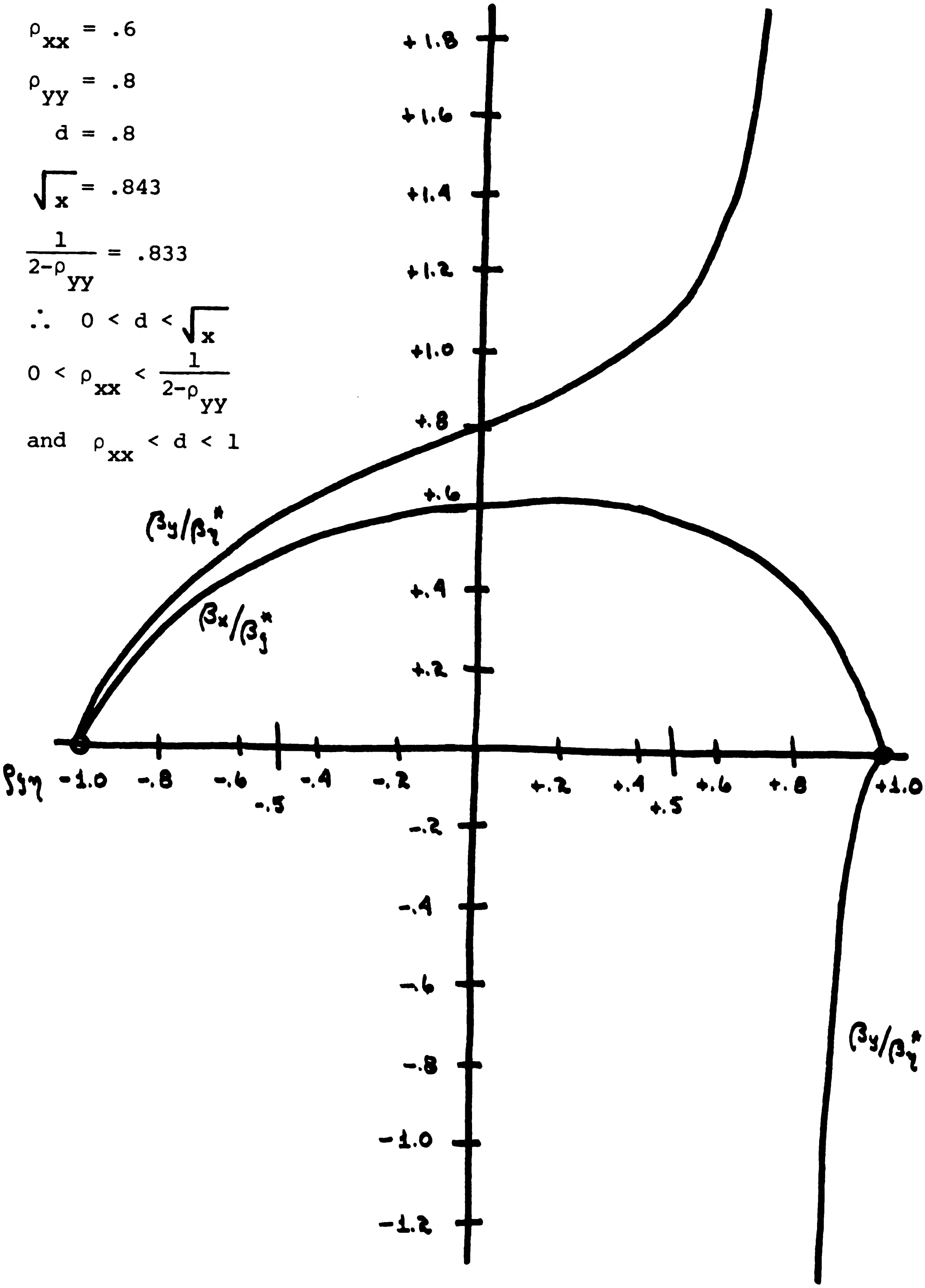


Figure 3.11c. G.C. I (x,y).

$\rho_{xx} = .6$

$\rho_{yy} = .8$

$d = .843$

$\sqrt{x} = .843$

$\frac{1}{2-\rho_{yy}} = .833$

$\therefore \sqrt{x} = d < 1,$

$0 < \rho_{xx} < \frac{1}{2-\rho_{yy}}$

and $\rho_{xx} < d < 1$

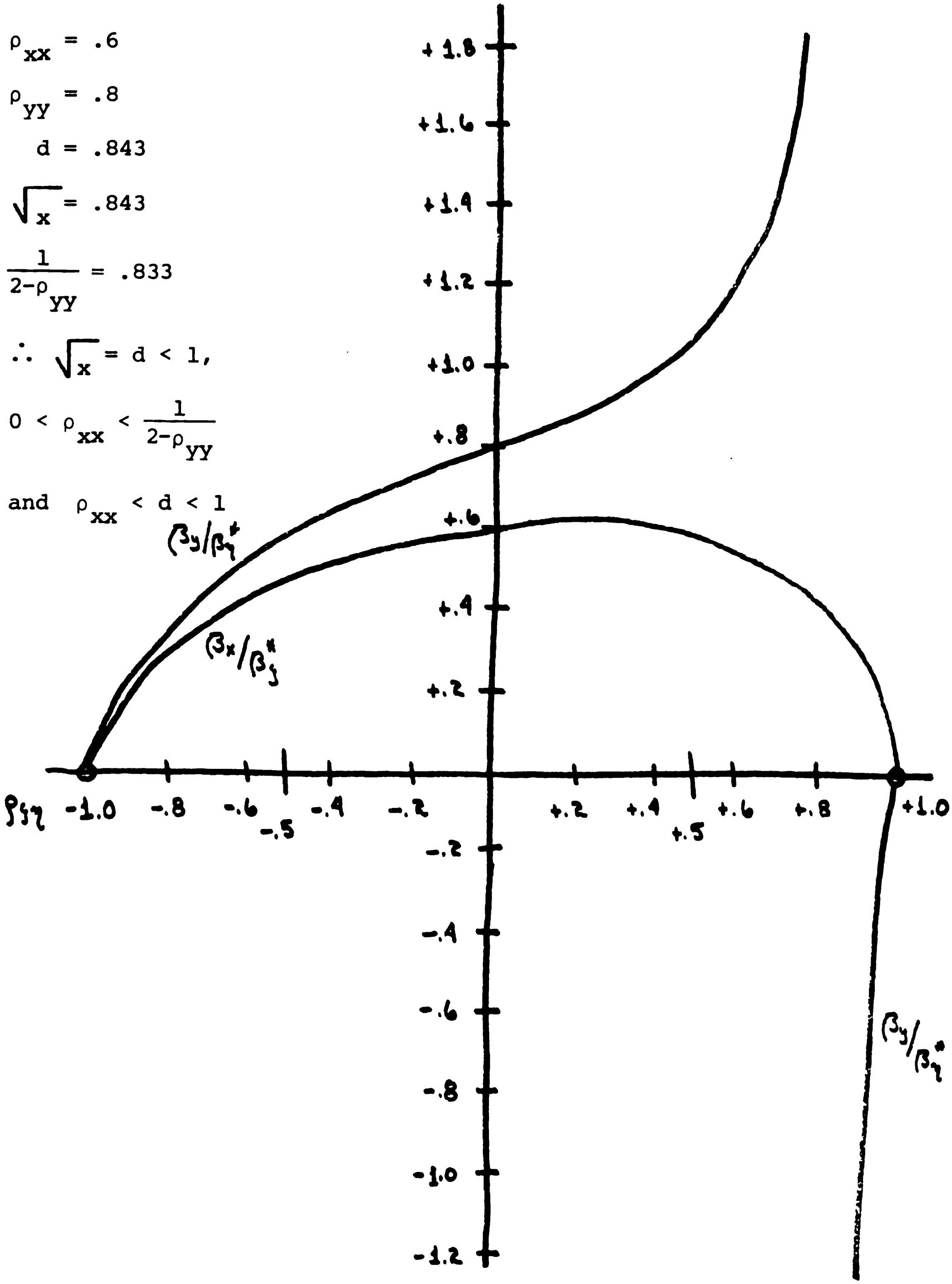


Figure 3.11d. G.C. I (x,y).

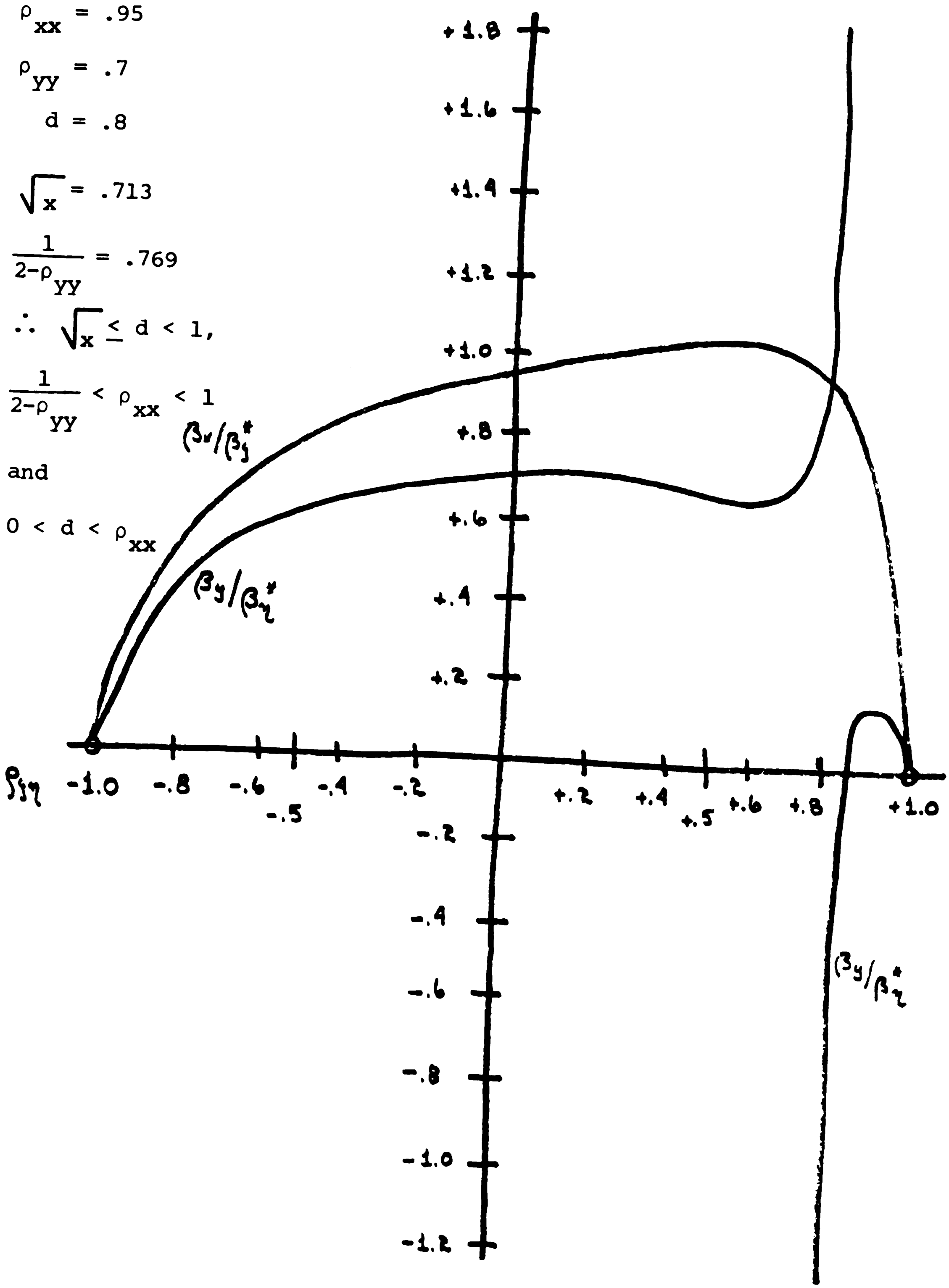


Figure 3.12a. G.C. II (x,y).

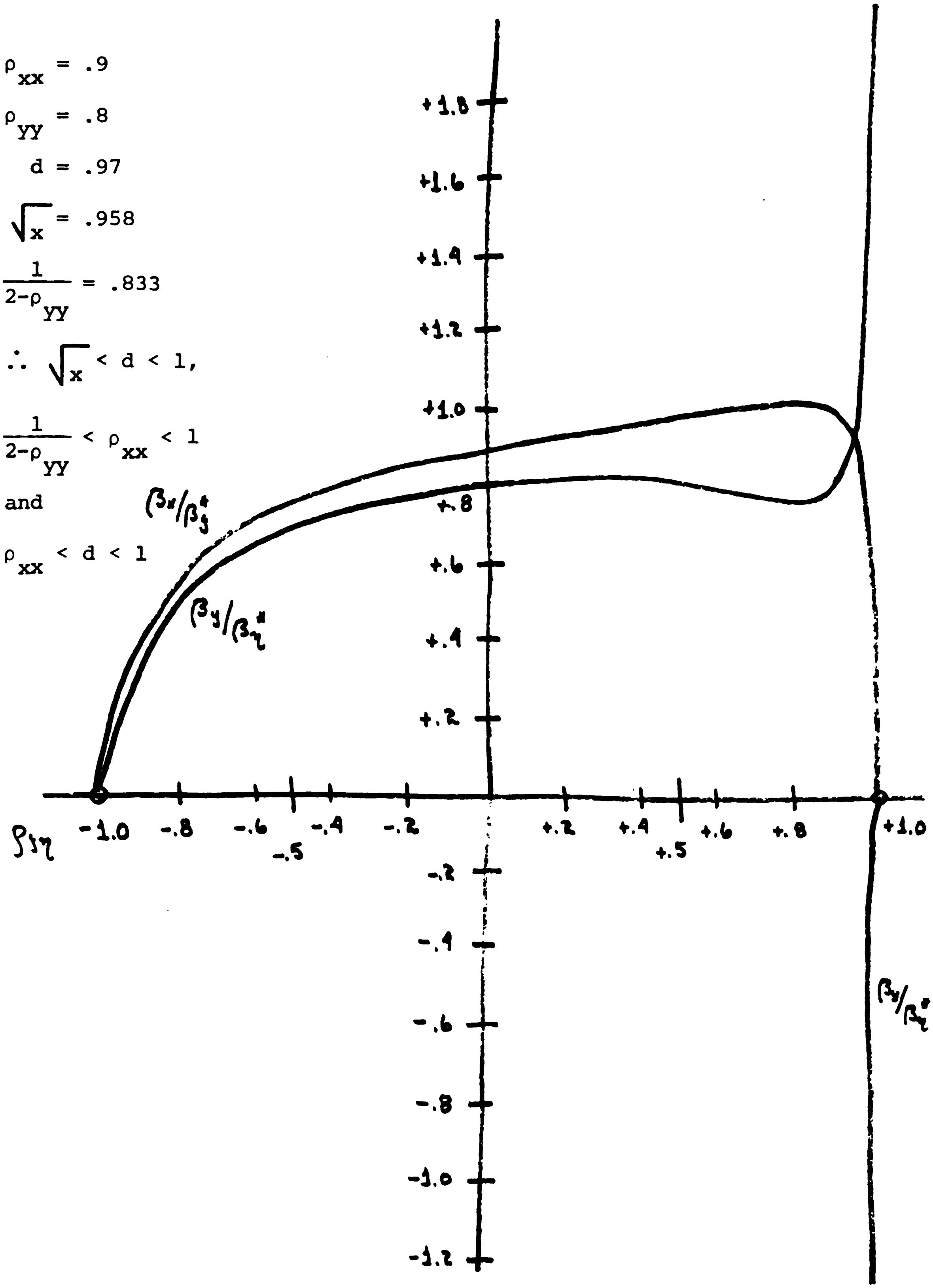


Figure 3.12b. G.C. II (x,y).

$\rho_{xx} = .9$

$\rho_{yy} = .7$

$d = .888$

$\sqrt{x} = .888$

$\frac{1}{2-\rho_{yy}} = .769$

$\therefore \sqrt{x} = d < 1,$

$\frac{1}{2-\rho_{yy}} < \rho_{xx} < 1$

and

$0 < d < \rho_{xx}$

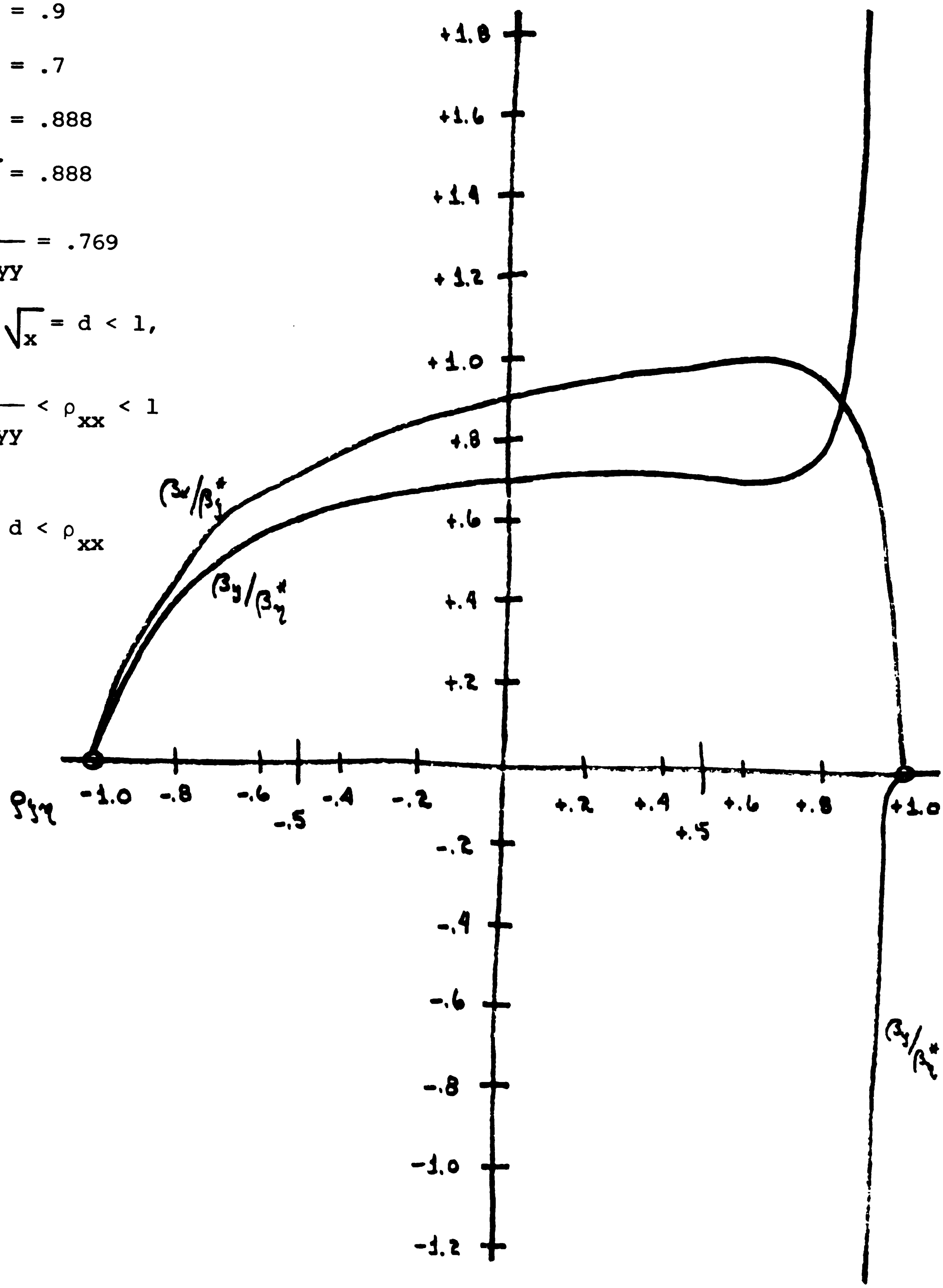


Figure 3.12c. G.C. II (x,y).

<u>G.C. III (x,y)</u>. When $1 < d < \sqrt{y}$ for $0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$ then G.C. I for β_x/β_ξ^*, G.C. II (y) for β_y/β_η^* and section VI results apply. Note, this general category is identical to G.C. II (x,y) but with the roles of x and y reversed. Therefore Figure 3.10 with the property of interchangeability of x and y applied provides the generic graphs for this category.

Figures 3.12a through 3.12c with the property of interchangeability of x and y applied provide examples for specific situations in G.C. III (x,y).

<u>G.C. IV (x,y)</u>. When $\sqrt{y} < d$ for any ρ_{xx}, ρ_{yy} $(\rho_{xx}, \rho_{yy} \neq 1)$ or when $1 < d < \sqrt{y}$ for $2\rho_{yy}-1/\rho_{yy} \leq \rho_{xx} < 1$, G.C. I for β_x/β_ξ^* and G.C. III (y) for β_y/β_η^* apply. Note, this general category is identical to G.C. I (x,y) but with the roles of x and y reversed. Therefore Figure 3.9 with the property of interchangeability of x and y applied provides the generic graphs for this category.

Figures 3.11a through 3.11d with the property of interchangeability of x and y applied provide examples for specific situations in G.C. IV (x,y).

When $d = 1$ for $0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$ with $\rho_{yy} \neq 1$ Figure 3.4a for β_x/β_ξ^* and Figure 3.8b for β_y/β_η^* apply, since $\frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \frac{1}{2 - \rho_{yy}}$ (see proof in Appendix B.7). Therefore the generic graphs for these situations are:

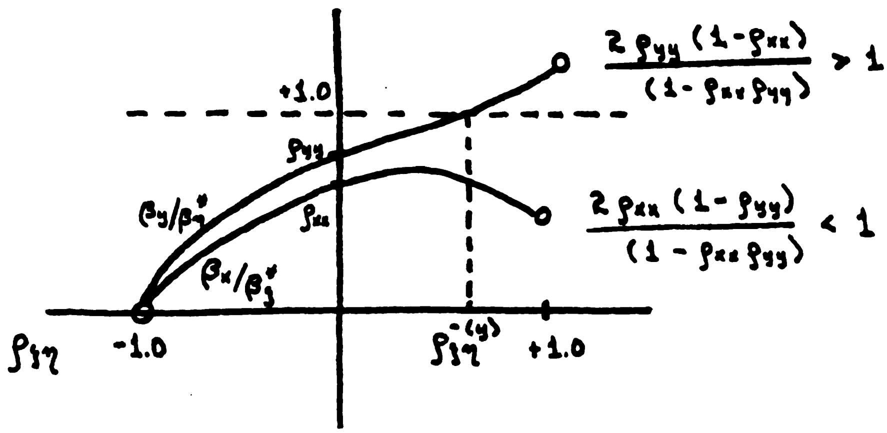


Figure 3.13a

When $d = 1$ for $\frac{2\rho_{yy} - 1}{\rho_{yy}} \leq \rho_{xx} \leq \frac{1}{2 - \rho_{yy}}$ with $\rho_{yy} \neq 1$ Figure 3.4a for β_x/β_ξ^* and Figure 3.8a for β_y/β_η^* apply. Therefore the generic graphs for these situations are:

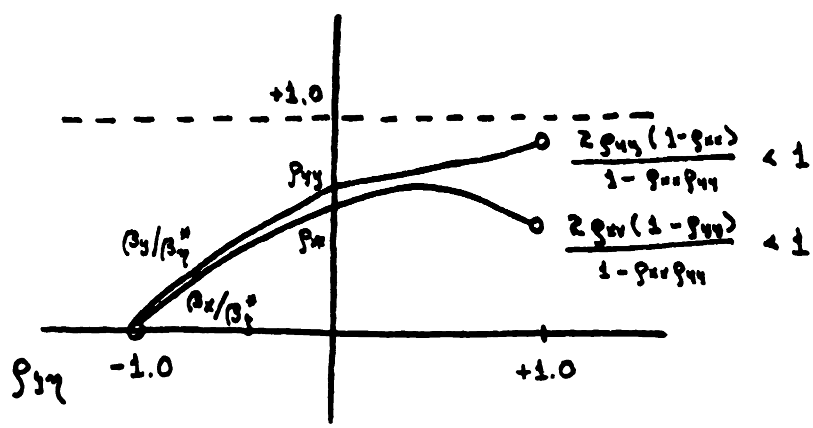


Figure 3.13b

Note: $\frac{2\rho_{yy}(1 - \rho_{xx})}{1 - \rho_{xx}\rho_{yy}} > \frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} \Leftrightarrow \rho_{yy} > \rho_{xx}$.

Proof:

$$\frac{2\rho_{yy}(1-\rho_{xx})}{1-\rho_{xx}\rho_{yy}} > \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} \Leftrightarrow \rho_{yy}(1-\rho_{xx}) > \rho_{xx}(1-\rho_{yy})$$

$$\Leftrightarrow \rho_{yy} - \rho_{yy}\rho_{xx} > \rho_{xx} - \rho_{xx}\rho_{yy}$$

(3.61a) $$\frac{2\rho_{yy}(1-\rho_{xx})}{1-\rho_{xx}\rho_{yy}} > \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} \Leftrightarrow \rho_{yy} > \rho_{xx}$$

In a similar manner it can be shown that

(3.61b) $$\frac{2\rho_{yy}(1-\rho_{xx})}{1-\rho_{xx}\rho_{yy}} \le \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} \Leftrightarrow \rho_{yy} \le \rho_{xx} .$$

When $d = 1$ for $\frac{1}{2-\rho_{yy}} < \rho_{xx} < 1$ with $\rho_{yy} \neq 1$ Figure 3.4b for β_x/β_ξ^* and Figure 3.8a for β_y/β_η^* apply. Applying the property of interchangeability of x and y to Figure 3.13a will provide the generic graphs for these situations.

Figures 3.14a and 3.14b provide examples for the first two sets of specific situations noted above when $d = 1$. Applying the property of interchangeability of x and y to Figure 3.14a provides an example for the third set of situations noted above.

J Category, Two Predictor Models ($J \geq 2$, $p = 2$)

The preceeding work has considered the case of 2 categories ($J = 2$) and 2 predictors ($p = 2$). To complete the examination of the two predictor special case, it is necessary to consider the most general two predictor model, that with J categories ($J \geq 2$). It is reasonably straightforward to show that results for the 2 category, 2 predictor model extend with only slight modification to the J category, 2 predictor model.

$\rho_{xx} = .6$

$\rho_{yy} = .8$

$d = 1$

$$\frac{2\rho_{yy} - 1}{\rho_{yy}} = .75$$

$$\frac{1}{2-\rho_{yy}} = .833$$

$$\therefore\ 0 < \rho_{xx} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$$

$$< \frac{1}{2-\rho_{yy}}$$

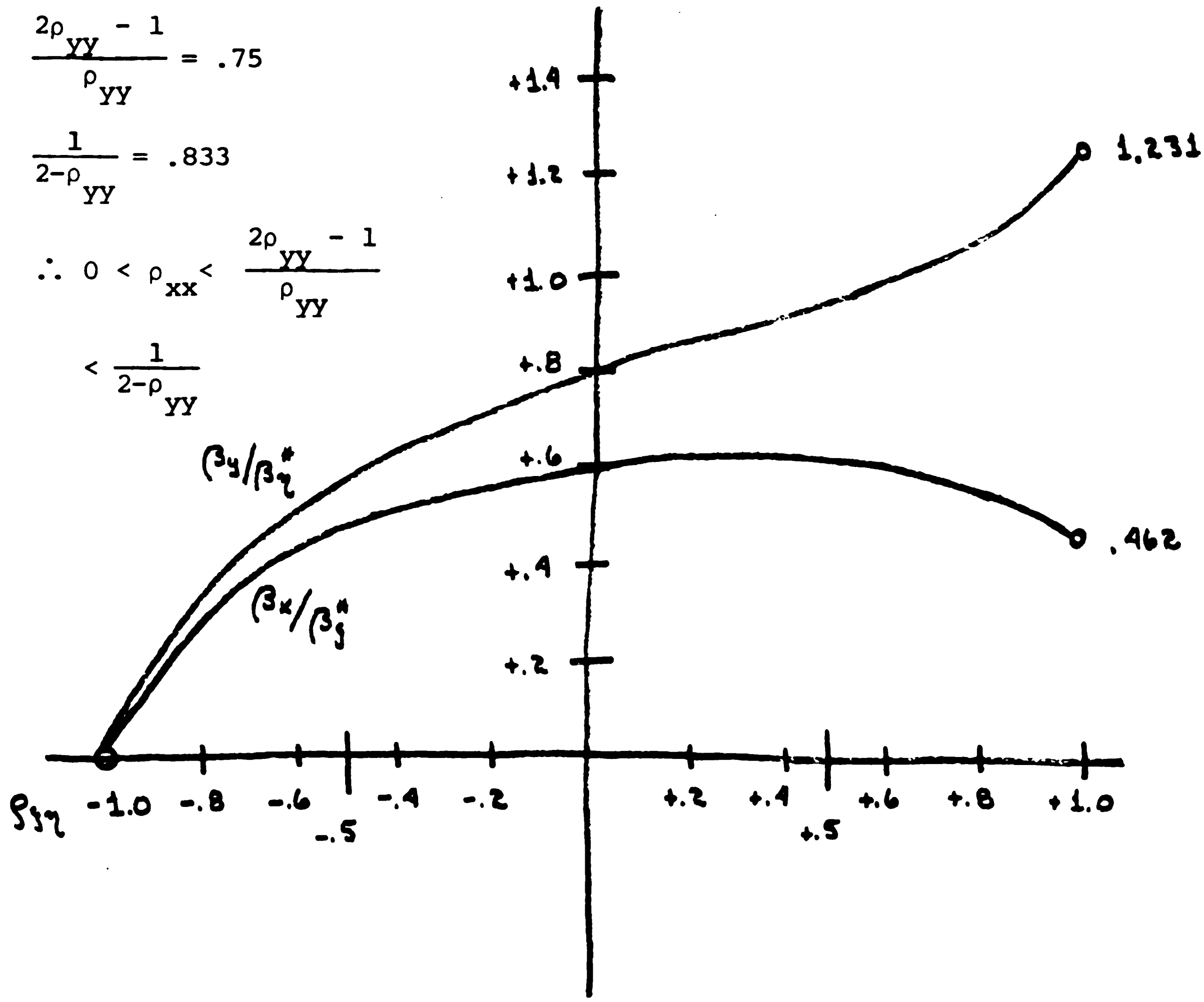


Figure 3.14a.

$\rho_{xx} = .76$

$\rho_{yy} = .8$

$d = 1$

$\frac{2\rho_{yy} - 1}{\rho_{yy}} = .750$

$\frac{1}{2-\rho_{yy}} = .833$

$\therefore \frac{2\rho_{yy} - 1}{\rho_{yy}} < \beta_y/\beta^*_\eta$

$\rho_{xx} < \frac{1}{2-\rho_{yy}}$

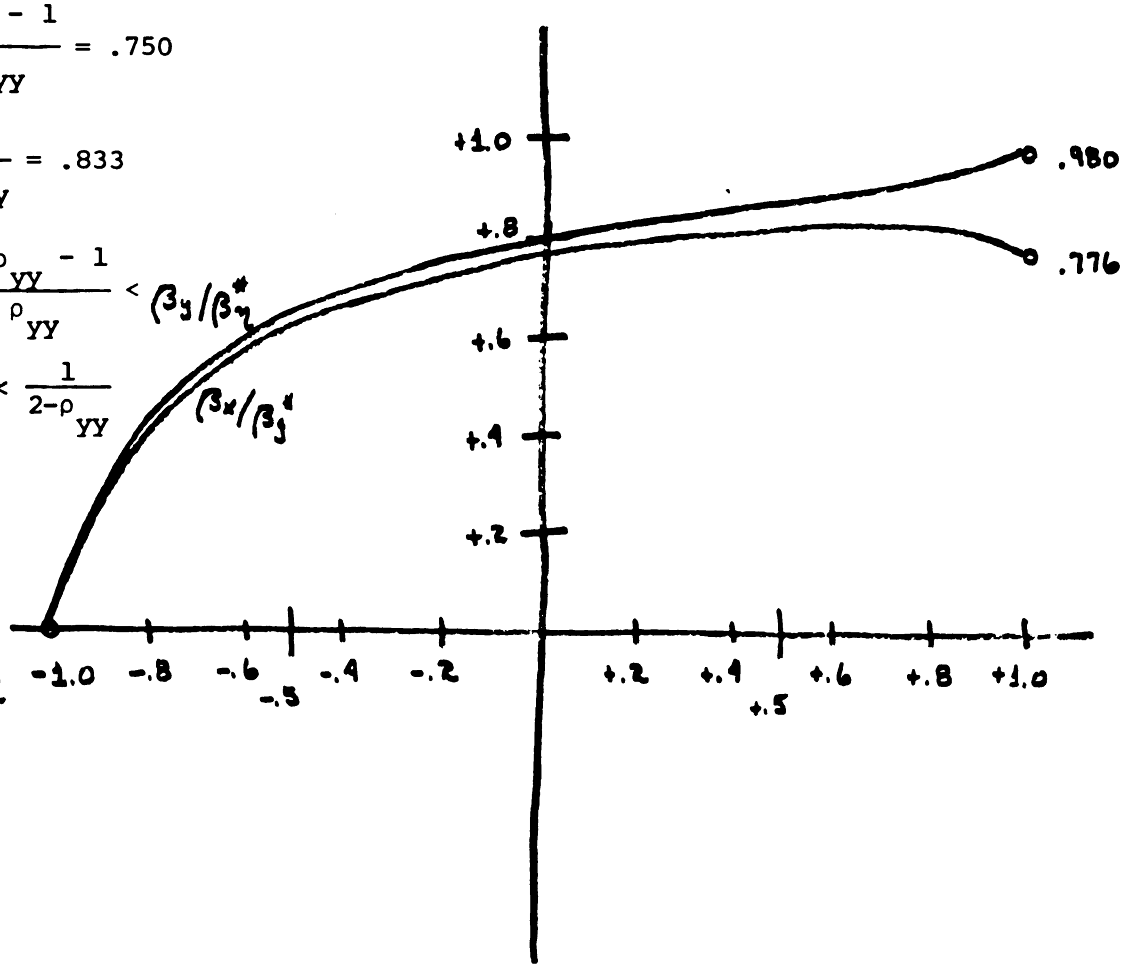


Figure 3.14b.

Consider any category k $(k = 1,2,\ldots,J)$, any vector of observed weighting coefficients from category k, $\underset{\sim}{\beta}_{k\cdot j}$ $(j \neq k,$ $j = 1,2,\ldots,J)$ and the corresponding vector of latent weighting coefficients $\underset{\sim}{\beta}^*_{k\cdot j}$. Since the model under consideration here is a 2 predictor model,

$$\underset{\sim}{\beta}_{k\cdot j} = \begin{bmatrix} \beta_{k\cdot j(x)} \\ \\ \beta_{k\cdot j(y)} \end{bmatrix} \quad \text{and} \quad \underset{\sim}{\beta}^*_{k\cdot j} = \begin{bmatrix} \beta_{k\cdot j(\xi)} \\ \\ \beta_{k\cdot j(\eta)} \end{bmatrix} .$$

Result (2.21) modified for the two predictor case (p = 2) becomes

$$(3.62a) \quad \beta^*_{k\cdot j(\xi)} = \frac{(\mu_\xi^{(k)} - \mu_\xi^{(j)}) - b_{\xi\cdot\eta}(\mu_\eta^{(k)} - \mu_\eta^{(j)})}{\sigma_\xi^2 - b_{\xi\cdot\eta}\sigma_{\xi\eta}} ,$$

where $b_{\xi\cdot\eta} = \rho_{\xi\eta}\dfrac{\sigma_\xi}{\sigma_\eta}$.

Let $a_\xi^{k\cdot j} = \mu_\xi^{(k)} - \mu_\xi^{(j)}$ and $a_\eta^{k\cdot j} = \mu_\eta^{(k)} - \mu_\eta^{(j)}$.

Note:

$$\sigma_\xi^2 - b_{\xi\cdot\eta}\sigma_{\xi\eta} = \sigma_\xi^2 - (\rho_{\xi\eta}\frac{\sigma_\xi}{\sigma_\eta})(\rho_{\xi\eta}\sigma_\xi\sigma_\eta) = \sigma_\xi^2(1 - \rho_{\xi\eta}^2).$$

Thus (3.62a) becomes

$$\beta^*_{k\cdot j(\xi)} = \frac{a_\xi^{k\cdot j} - a_\eta^{k\cdot j}\frac{\sigma_\xi}{\sigma_\eta}\rho_{\xi\eta}}{\sigma_\xi^2(1 - \rho_{\xi\eta}^2)} = \frac{a_\xi^{k\cdot j}\left(1 - \dfrac{a_\eta^{k\cdot j}/\sigma_\eta}{a_\xi^{k\cdot j}/\sigma_\xi}\rho_{\xi\eta}\right)}{\sigma_\xi^2(1 - \rho_{\xi\eta}^2)}$$

Let $d_\xi^{k\cdot j} = \dfrac{a_\eta^{k\cdot j}/\sigma_\eta}{a_\xi^{k\cdot j}/\sigma_\xi} = \dfrac{(\mu_\eta^{(k)} - \mu_\eta^{(j)})/\sigma_\eta}{(\mu_\xi^{(k)} - \mu_\xi^{(j)})/\sigma_\xi}$.

Therefore

(3.62b) $$\beta^{*}_{k\cdot j(\xi)} = \frac{a_{\xi}^{k\cdot j}(1 - d_{\xi}^{k\cdot j}\rho_{\xi\eta})}{\sigma_{\xi}^{2}(1 - \rho_{\xi\eta}^{2})} .$$

Result (2.17b) modified for the two predictor case (p = 2) becomes

(3.63a) $$\beta_{k\cdot j(x)} = \frac{(\mu_x^{(k)} - \mu_x^{(j)}) - b_{x\cdot y}(\mu_y^{(k)} - \mu_y^{(j)})}{\sigma_x^2 - b_{x\cdot y}\sigma_{xy}}$$

where $$b_{x\cdot y} = \rho_{xy}\frac{\sigma_x}{\sigma_y} .$$

Let $a_x^{k\cdot j} = \mu_x^{(k)} - \mu_x^{(j)}$ and $a_y^{k\cdot j} = \mu_y^{(k)} - \mu_y^{(j)}$, where $a_x^{k\cdot j} = a_{\xi}^{k\cdot j}$ and $a_y^{k\cdot j} = a_{\eta}^{k\cdot j}$ by a result comparable to (3.8).

Note: 1) $$b_{x\cdot y} = \rho_{xy}\sigma_x/\sigma_y = (\rho_{\xi\eta}\sqrt{\rho_{xx}\rho_{yy}})\left(\frac{\sigma_{\xi}}{\sqrt{\rho_{xx}}}\right)\left(\frac{\sqrt{\rho_{yy}}}{\sigma_{\eta}}\right)$$

$$= \frac{\sigma_{\xi}}{\sigma_{\eta}}\rho_{\xi\eta}\rho_{yy} .$$

By (3.10), (3.9a) and (3.9b).

2) $$\sigma_x^2 - b_{x\cdot y}\sigma_{xy} = \sigma_x^2 - \left(\rho_{xy}\frac{\sigma_x}{\sigma_y}\right)(\rho_{xy}\sigma_x\sigma_y) = \sigma_x^2(1 - \rho_{xy}^2)$$

$$= \frac{\sigma_{\xi}^2}{\rho_{xx}}(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}).$$

By (3.9a) and (3.10).

Therefore (3.63a) becomes

$$\beta_{k\cdot j(x)} = \frac{a_\xi^{k\cdot j} - a_\eta^{k\cdot j}\frac{\sigma_\xi}{\sigma_\eta}\rho_{\xi\eta}\rho_{yy}}{\frac{\sigma_\xi^2}{\rho_{xx}}(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})}$$

$$= \frac{\rho_{xx}\, a_\xi^{k\cdot j}(1 - \frac{a_\eta^{k\cdot j}/\sigma_\eta}{a_\xi^{k\cdot j}/\sigma_\xi}\rho_{\xi\eta}\rho_{yy})}{\sigma_\xi^2(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})}$$

(3.63b) $$\beta_{k\cdot j(x)} = \frac{\rho_{xx}\, a_\xi^{k\cdot j}(1 - d_\xi^{k\cdot j}\rho_{\xi\eta}\rho_{yy})}{\sigma_\xi^2(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})},$$

where $$d_\xi^{k\cdot j} = \frac{a_\eta^{k\cdot j}/\sigma_\eta}{a_\xi^{k\cdot j}/\sigma_\xi} .$$

The ratio of interest $\beta_{k\cdot j(x)}/\beta^*_{k\cdot j(\xi)}$ becomes

$$\beta_{k\cdot j(x)}/\beta^*_{k\cdot j(\xi)} = \frac{\rho_{xx}\, a_\xi^{k\cdot j}(1 - d_\xi^{k\cdot j}\rho_{\xi\eta}\rho_{yy})}{\sigma_\xi^2(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})} \Big/ \frac{a_\xi^{k\cdot j}(1 - d_\xi^{k\cdot j}\rho_{\xi\eta})}{\sigma_\xi^2(1 - \rho_{\xi\eta}^2)}$$

(3.64) $$\beta_{k\cdot j(x)}/\beta_{k\cdot j(\xi)} = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d_\xi^{k\cdot j}\rho_{\xi\eta})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d_\xi^{k\cdot j}\rho_{\xi\eta})}$$

where $$d_\xi^{k\cdot j} = \frac{a_\eta^{k\cdot j}/\sigma_\eta}{a_\xi^{k\cdot j}/\sigma_\xi} = \frac{(\mu_\eta^{(k)} - \mu_\eta^{(j)})/\sigma_\eta}{(\mu_\xi^{(k)} - \mu_\xi^{(j)})/\sigma_\xi} .$$

Note the close similarity between expression (3.64) for the J category, 2 predictor model and expression (3.12) for the 2 category, 2 predictor model. The only difference is in the use of $d_\xi^{k\cdot j}$ in (3.64) and d_ξ in (3.12). But note that d_ξ in (3.12) is based on a comparison between predictor means in the only two categories in the model whereas $d_{\xi}^{k\cdot j}$ in (3.64) is based on a comparison between predictor means from some two of the J categories.

Thus in the J category, 2 predictor model any ratio of the form $\beta_{k\cdot j(x)}/\beta^{*}_{k\cdot j(\xi)}$ for $j,k = 1,2,\ldots,J$ and $j \neq k$, will have a distribution which corresponds to one of the General Categories for x depending on the values of $d_{\xi}^{k\cdot j}$, ρ_{xx}, ρ_{yy} and $\rho_{\xi\eta}$, where the $d_{\xi}^{k\cdot j}$ value can be treated as a value of d_{ξ} from (3.12), since $d_{\xi}^{k\cdot j}$ will take on values $-\infty < d_{\xi}^{k\cdot j} < +\infty$ just as d_{ξ} does.

A generalized property of interchangeability applied to (3.64) produces the ratio $\beta_{k\cdot j(y)}/\beta^{*}_{k\cdot j(\eta)}$ for $j,k = 1,2,\ldots,J$ with $j \neq k$ which is an extension of (3.14).

$$(3.65)\quad \beta_{k\cdot j(y)}/\beta_{k\cdot j(\eta)} = \frac{(1-\rho_{\xi\eta}^{2})\rho_{yy}(1-d_{\eta}^{k\cdot j}\rho_{\xi\eta}\rho_{xx})}{(1-\rho_{\xi\eta}^{2}\rho_{xx}\rho_{yy})(1-d_{\eta}^{k\cdot j}\rho_{\xi\eta})}$$

$$\text{where}\quad d_{\eta}^{k\cdot j} = \frac{(\mu_{\xi}^{(k)} - \mu_{\eta}^{(j)})/\sigma_{\xi}}{(\mu_{\eta}^{(k)} - \mu_{\nu}^{(j)})/\sigma_{\eta}} .$$

Therefore all results which apply to (3.14) can be easily extended to apply to (3.65) with $d_{\eta}^{k\cdot j}$ in (3.65) taking the place of d_{η} in (3.14).

Therefore all results noted earlier for the 2 category, 2 predictor model extend simply to apply to corresponding cases in the J category $(J \geq 2)$, 2 predictor model.

Section D: Summary

The purpose of this chapter was to examine the effects of errors of measurement on the weighting coefficients of a Latent Random Predictor Quantal Response Model, given by (2.19) for the most general case. The approach to the problem involved selecting an arbitrary vector of weighting coefficients associated with some arbitrary category of the criterion variable and examining the individual weighting coefficients associated with each predictor.

From an arbitrary vector of latent weighting coefficients associated with some category of the criterion from the model given by (2.19), an individual latent weighting coefficient associated with some latent predictor T^q was selected, call it β_q^*. From the corresponding vector of observed weighting coefficients associated with the same category of the criterion, the individual observed weighting coefficient associated with observed predictor X^q was selected, call it β_q. Note that X^q and T^q are related through the measurement model (2.22) such that $X^q = T^q + E^q$ where E^q is the error portion of the observed predictor. Then the ratio β_q/β_q^* was examined. When $\beta_q/\beta_q^* > 1$, then the observed weighting coefficient (β_q) is an overestimate of the latent weighting coefficient (β_q^*). When $\beta_q/\beta_q^* = 1$, then β_q is an exact estimate of β_q^*. When $\beta_q/\beta_q^* < 1$, then β_q is an underestimate of β_q^*.

The research of this chapter included one and two predictor models only. No general results applicable to all models were discovered and the approach used in this chapter proved extremely difficulty for use with models involving more than two predictors.

For the one predictor models ($p = 1$) there is only one component in each vector of weighting coefficients. Result (2.28) indicated that for all one predictor models the value of the observed weighting coefficient will be an underestimate of the value of the latent weighting coefficient by a factor equal to the reliability of the single predictor variable. This result holds true for every pair of related observed and latent weighting coefficients associated with any category of the criterion. The only exception to this result occurs when the latent weighting coefficient is zero. In that case the observed weighting coefficient was also shown to be zero.

For the two predictor models ($p = 2$) no universally applicable result was found, such as that produced for one predictor models. The approach for the two predictor models involved a change to simplify the notation and make it consistent with the notation used by McSweeney and Schmidt (1974). Under this simplification the observed predictors are noted as x and y with the corresponding latent predictors being ξ and η, where $x = \xi + e_x$ and $y = \eta + e_y$ in an adaptation of the basic measurement model (2.22). The major work for the two predictor model was done for the two category ($J = 2$) case. All results for this simplest case of the two predictor model were then shown to extend easily to the general case ($J \geq 2$) of the two predictor model. In the 2 category, 2 predictor model the observed weighting coefficients were denoted β_x and β_y while the corresponding latent weighting coefficients were denoted β^*_ξ and β^*_η. Therefore, the ratios which were examined as they relate to one were β_x/β^*_ξ and β_y/β^*_η. Each of these ratios were shown to be functions of d, $\rho_{\xi\eta}$, ρ_{xx} and ρ_{yy}. d is a ratio of the differences between category means for the two predictors where the differences are in standardized latent units. See expression (3.6b) and the ensuing narrative for the definition and explanation of $d = d_\xi$. $\rho_{\xi\eta}$ is the correlation between the latent predictors. ρ_{xx} and ρ_{yy} are the predictor reliabilities for observed predictors x and y respectively, and indicate the presence of errors or measurement when either, or both, ρ_{xx} or ρ_{yy} are less than one.

In Appendix B.1, part B, it was shown that β_x/β_ξ^* and β_y/β_η^* need to be examined only for $d \geq 0$. Results for $d < 0$ can be derived simply for comparable results when $d > 0$ by the use of expressions (3.16a) and (3.16b).

Since there were no universally applicable results discovered, the work in this chapter identified four general categories of situations associated with β_x/β_ξ^* and β_y/β_η^*, where β_x/β_ξ^* and β_y/β_η^* were considered as functions of $\rho_{\xi\eta}$ for fixed values of d $(d > 0)$, ρ_{xx} and ρ_{yy}. Within each of these four joint general categories, defined in Section C under sub-heading VII above, the behavior of the ratios β_x/β_ξ^* and β_y/β_η^* follow the same general pattern for all situations (i.e., values of d, ρ_{xx} and ρ_{yy}) included in the category. In addition to the four joint general categories three categories related to the special case when $d = 1$ are also identified in Section C, sub-heading VII.

Three results do apply across all 4 joint general categories and the three special case categories. First, when there is no correlation between the latent predictors, i.e., $\rho_{\xi\eta} = 0$, then each ratio is equal to the reliability of the predictor. That is, when $\rho_{\xi\eta} = 0$, $\beta_x/\beta^*_\xi = \rho_{xx}$ and $\beta_y/\beta^*_\eta = \rho_{yy}$. This result applies to all situations when β^*_ξ and β^*_η are not equal to zero.

Second, when the correlation between the predictors, $\rho_{\xi\eta}$, and the ratio of standardized category mean differences, d, have opposite signs then the observed weighting coefficient will underestimate the latent weighting coefficient for both predictors, i.e., $\beta_x/\beta^*_\xi < 1$ and $\beta_y/\beta^*_\eta < 1$. In this case, i.e., $d\rho_{\xi\eta} < 0$, for fixed d, ρ_{xx} and ρ_{yy} the amount of the underestimate increases as the magnitude of the correlation increases. That is, if $d > 0$ then as $\rho_{\xi\eta}$ takes on values nearer to -1 the ratios β_x/β^*_ξ and β_y/β^*_η will become smaller, approaching zero as $\rho_{\xi\eta}$ approaches one.

In Appendix B.9 the interpretation of $d\rho_{\xi\eta}$ is given as a ratio of the slope of the pooled within category regression line of ξ on η over the slope of the line joining the midpoints of each category distribution of ξ and η. The potential for occurrence of a negative ratio of slopes is discussed under Derivation 3 in Section C, subheading VI above. Although in many situations the ratio of slopes will be positive (therefore $d\rho_{\xi\eta} > 0$) it is possible for the ratio of slopes to be negative (i.e., $d\rho_{\xi\eta} < 0$).

Third, Derivation 2 in Section C, sub-heading VI above demonstrates that it is impossible for both observed weighting coefficients to simultaneously overestimate the latent weighting coefficients for the same set of values for d, $\rho_{\xi\eta}$, ρ_{xx} and ρ_{yy}. At most one observed weighting coefficient will be an overestimate of the latent weighting coefficient in any given situation. In addition

Derivation 1 in Section C, sub-heading VI above indicates that for the two predictor model if the observed weighting coefficient for one predictor is equal to the latent weighting coefficient for that predictor then the observed weighting coefficient for the other predictor is an underestimate of its corresponding latent weighting coefficient by a factor equal to the reliability of this second predictor. That is, if $\beta_x/\beta_\xi^* = 1$ then $\beta_y/\beta_\eta^* = \rho_{yy}$ or if $\beta_y/\beta_\eta^* = 1$ then $\beta_x/\beta_\xi^* = \rho_{xx}$. The converses of these statements are true only when $\rho_{\xi\eta} \neq 0$.

An interesting result which occurs only for joint general categories one and two [G.C. I (x,y) and G.C. II (x,y)] occurs for values of $\rho_{\xi\eta}$ in an arbitrarily small neighborhood of d (in these categories d is positive and less than one). For values of $\rho_{\xi\eta}$ arbitrarily near d the magnitude of β_y/β_η^*, i.e., $|\beta_y/\beta_\eta^*|$ is unboundedly large. When $\rho_{\xi\eta} = d$ the ratio β_y/β_η^* is not defined since $\beta_\eta^* = 0$. A similar situation occurs for β_x/β_ξ^* in G.C. III (X,y) and G.C. IV (x,y) for values of $\rho_{\xi\eta}$ near $1/d$ (in these categories d exceeds one hence $1/d$ is less than one). In this case when $\rho_{\xi\eta} = 1/d$ the ratio β_x/β_ξ^* is not defined since $\beta_\xi^* = 0$.

The importance of this result occurs in interpreting the effects of errors of measurement using β_x/β_ξ^* when $\rho_{\xi\eta}$ is near $1/d$ or using β_y/β_η^* when $\rho_{\xi\eta}$ is near d. Consider some situation from G.C. I (x,y) where d, ρ_{xx} and ρ_{yy} are fixed. Here d will have some positive value which is less than one. For values of $\rho_{\xi\eta}$ which are arbitrarily close to d, $|\beta_y/\beta_\xi^*|$ will be arbitrarily large. However, depending on the specific value of the difference between the category means for predictor y, the magnitude of the latent weighting coefficient will be extremely small, i.e., $|\beta_\eta^*|$ will be near zero. In this situation the magnitude of the observed weighting coefficient, $|\beta_y|$, may also be quite small even though the ratio $|\beta_y/\beta_\eta^*|$ may be relatively large. For example, for some $\rho_{\xi\eta}$ near d, β_η^* might have a value of .005 while β_y might be .05. In this case $\beta_y/\beta_\eta^* = 10$ which represents a rather large factor. Even though β_y is an overestimate of β_η^* by a factor equal to 10, the magnitude of the overestimate, $\beta_y - \beta_\eta^* = .045$, is relatively small and in most interpretations a difference of this size for weighting coefficients of this magnitude is meaningfully insignificant. For values of $\rho_{\xi\eta}$ near d, β_η^* must be near zero. However, there is no necessary reason why β_y need be near zero also. In fact the difference $\beta_y - \beta_\eta^*$ may be significantly large for some situations.

Therefore, the value of β_y/β_η^* for values of $\rho_{\xi\eta}$ near d and the value of β_x/β_ξ^* for values of $\rho_{\xi\eta}$ near 1/d need to be interpreted with great care. A relatively large ratio may mask two rather small weighting coefficients which may have a neglible practical difference in magnitude. Or a large ratio may represent a significant discrepancy between β_y and β_η^*.

Note however that when $\rho_{\xi\eta} = 1/d$ then the value of β_x consists totally of effects of errors of measurement since in this case $\beta_\xi^* = 0$ indicating no relationship between the latent predictor and the probability of classification into category one of the criterion. See Appendix B.9 for an interpretation of this result in terms of the ratio of within group to between group slopes. The same conclusion applies for β_y as an estimate of β_η^* when $\rho_{\xi\eta} = d$.

Cochran (1968) reported results for the effect of errors of measurement on regression coefficients in linear regression models. Although the distributional assumptions included in linear regression models are different than the distributional assumptions included in random predictor quantal response models, the expression for a vector of regression coefficients has a structural similarity to the expression for a vector of quantal response weighting coefficients. Because of this structural similarity between vectors of regression coefficients and vectors of quantal response weighting coefficients it is not surprising that some of the results reported from this research for quantal response weighting coefficients have a similarity to results reported by Cochran (1968) for regression coefficients.

For the one predictor case Cochran (1968, p. 652) demonstrated that the observed regression coefficient is an underestimate of the latent regression coefficient by a factor equal to the reliability of the predictor. An identical result for the relationship between the observed quantal response weighting coefficient and the latent weighting coefficient is reported from this research.

In the two predictor situation where only one predictor is subject to error Cochran (1968, p. 656) provides an expression for the observed regression coefficient as a function of the latent regression coefficients and other parameters describing the two predictors. An identical expression also exists for the observed quantal response weighting coefficient as a function of the latent weighting coefficients and other parameters describing the two predictors.

Although the structures of the expressions for the vectors of regression coefficients and the vectors of quantal response weighting coefficients are similar, the derivations of the coefficients, which are based on the distributional assumptions of each model, are different. Thus it is also not surprising to discover that some results reported by Cochran (1968) for regression coefficients do not have exact counterparts among quantal response weighting coefficients. For example, in the two predictor situation where the reliabilities of the two predictors are equal, Cochran (1968, p. 656) reported that the ratio of the observed regression coefficient to the latent regression coefficient for either predictor will be somewhat greater than the reliability of the predictor when the correlation between predictors is positive. This conclusion is not true, in general, for ratios of quantal response coefficients. One quantal response counter-example occurs for $\rho_{xx} = \rho_{yy} = .8$, $d = d_\xi = .2$ and $\rho_{\xi\eta} = .3$. In this case $\beta_x/\beta_\xi^* = .782$ and $\beta_y/\beta_\eta^* = .310$ and both ratios are less than the common predictor reliability even though the correlation between predictor is positive.

Cochran (1968, p. 655ff) also reports that the observed regression coefficient, in a multiple linear regression, associated with some one predictor can be expressed as a linear function of the latent regression coefficient associated with that one predictor and the latent regression coefficients associated with any other predictors that are correlated with that one predictor. For the multiple predictor quantal response model no general result comparable to this result was found.

Thus, although there are similarities in the results reported by Cochran (1968) for the effects of errors of measurement on the regression coefficients of a linear regression model and the results reported from this research for quantal response models, the conclusions for the two models are not identical.

For the two predictor models (p = 2), reviewing the generic graphs and the tables of results which define each of the four joint general categories and the three categories of the special case $d = 1$, clearly indicates that for every situation where at least one predictor is less than perfectly reliable, i.e., either ρ_{xx} or ρ_{yy} is less than one, the observed weighting coefficient represents either an overestimate or an underestimate of the latent weighting coefficient for at least one of the predictors.

For the one predictor models (p = 1), the observed weighting coefficient is always an underestimate of the latent weighting coefficient by a factor of the reliability.

Therefore, for both one and two predictor models the presence of errors of measurement in the predictor variables does have an effect on the determination of the true relationship between a predictor and the probability of classification in a given category of a criterion. In all cases where errors of measurement are present in the predictors, use of the observed weighting coefficient as an estimate of the latent weighting coefficient will result in an incorrect estimate. This applies for at least one if not both predictors in a two predictor model and for the single predictor in a one predictor model. Determination of whether the discrepancy between the observed weighting coefficient and the latent weighting coefficient is large enough to be of practical significance for situations which typically occur in quantal response applications is beyond the scope of this research.

Since the use of observed weighting coefficients as estimates for latent weighting coefficients does not provide exact estimates, the work presented in chapter 4 will give a reformulation of the Observed Random Predictor Quantal Response Model (2.2) in terms of parameters from the Latent Random Predictor Quantal Response Model and parameters describing errors of measurement. The associated maximum likelihood estimation procedures which allow the estimation of the latent weighting coefficients from the observed data will also be presented.

CHAPTER 4

Section A: Introduction

The work in chapter 3 consisted of a theoretical, analytical comparison of the weighting coefficients from two quantal response models. In the Observed Random Predictor Quantal Response Model (2.2) the vectors of weighting coefficients are defined in terms of the variances of the observed predictors which include the error variances. In the Latent Random Predictor Quantal Response Model (2.19) the vectors of weighting coefficients are defined in terms of the variances of the latent predictors which include no error variance. The relationship between the two models is provided by the classical measurement model (2.22).

The relationships of interest between the predictors and the criterion are given by the vectors of latent weighting coefficients from (2.19). However, most variables encountered in practice which are reasonable candidates for use as predictor variables contain some errors of measurement. Thus the model (2.19) based on the availability of predictors with no errors of measurement will not typically be applicable. Hence the estimation of the latent weighting coefficients must come from the model (2.2) for observed predictors.

For the work to be presented below the Observed Random Predictor Quantal Response Model (2.2) will be reformulated in terms of parameters from the Latent Random Predictor Quantal Response Model (2.19), and parameters describing errors of measurement. Then the maximum likelihood procedures associated with the reformulated model for estimating the latent parameters will be described. The estimates of the parameters of the Latent Random Predictor Quantal Response Model can then be used to produce estimates of the vectors of latent weighting coefficients.

Section B: Reformulation of the Observed Random Predictor Quantal Response Model

The most general case of the Observed Random Predictor Quantal Response Model for J categories of the criterion ($J \geq 2$) and p observed predictors ($p \geq 1$) is given by (2.2) and repeated here for convenience.

For some category k $(k = 1,2,\ldots,J)$

$$\Pr\{Y = k|\underset{\sim}{X}\} = P_k = 1/[1 + \sum_{\substack{j=1 \\ j\neq k}}^{J} \exp\{-(\alpha_{k\cdot j} + \underset{\sim}{\beta}'_{k\cdot j}\underset{\sim}{X})\}]$$

where

$$\alpha_{k\cdot j} = -\ln(\frac{P_j}{P_k}) - \frac{1}{2}[\underset{\sim}{\mu}_x^{(k)'}\Sigma^{-1}\underset{\sim}{\mu}_x^{(k)} - \underset{\sim}{\mu}_x^{(j)'}\Sigma^{-1}\underset{\sim}{\mu}_x^{(j)}]$$

and

$$\underset{\sim}{\beta}_{k\cdot j} = \Sigma^{-1}(\underset{\sim}{\mu}_x^{(k)} - \underset{\sim}{\mu}_x^{(j)}) \quad \text{for} \quad j \neq k,\ j,k = 1,2,\ldots,J.$$

Applying the classical measurement model (2.22) together with some of the properties of the classical measurement model (2.24a) and (2.24b), to (2.2) produces a reformulation of the Observed Random Predictor Quantal Response Model in terms of parameters from the Latent Random Predictor Quantal Response Model (2.19) and parameters describing errors of measurement.

For some category k $(k = 1,2,\ldots,J)$

$$(4.1) \quad \Pr\{Y = k|\underset{\sim}{X}\} = P_k = 1/[1 + \sum_{\substack{j=1 \\ j\neq k}}^{J} \exp\{-(\alpha_{k\cdot j} + \underset{\sim}{\beta}'_{k\cdot j}\underset{\sim}{X})\}]$$

where

$$\alpha_{k\cdot j} = -\ln(\frac{P_j}{P_k}) - \frac{1}{2}[\underset{\sim}{\mu}_T^{(k)'}(\Phi + \Psi^2)^{-1}\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(j)'}(\Phi + \Psi^2)^{-1}\underset{\sim}{\mu}_T^{(j)}]$$

and

$$\beta_{k \cdot j} = (\Phi + \Psi^2)^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(j)}) \quad \text{for} \quad j \neq k,\ j,k = 1,2,\ldots,J.$$

Applying the expanded measurement model (2.26) together with some of the properties of this model, i.e., (2.27a) and (2.27b) to (2.2) produces another reformulation of the Observed Random Predictor Quantal Response Model in terms of parameters of the Latent Random Predictor Quantal Response Model (2.19), parameters describing errors of measurement and parameters allowing for different scales of measurement among the observed predictors. This reformulation incorporates the use of replicate observed measurements for each predictor.

For some category k $(k = 1,2,\ldots,J)$

$$(4.2) \quad \Pr\{Y = k|\underset{\sim}{X}\} = P_k = 1/[1 + \sum_{\substack{j=1 \\ j \neq k}}^{J} \exp\{-(\alpha_{k \cdot j} + \underset{\sim}{\beta}'_{k \cdot j}\underset{\sim}{X})\}]$$

where

$$\alpha_{k \cdot j} = -\ln(\frac{P_j}{P_k}) - \frac{1}{2}[(\Lambda\underset{\sim}{\mu}_T^{(k)})'(\Lambda\Phi\Lambda' + \Psi^2)^{-1}(\Lambda\underset{\sim}{\mu}_T^{(k)})$$

$$- (\Lambda\underset{\sim}{\mu}_T^{(j)})'(\Lambda\Phi\Lambda' + \Psi^2)^{-1}(\Lambda\underset{\sim}{\mu}_T^{(j)})]$$

and

$$\underset{\sim}{\beta}_{k \cdot j} = (\Lambda\Phi\Lambda' + \Psi^2)^{-1}(\Lambda\underset{\sim}{\mu}_T^{(k)} - \Lambda\underset{\sim}{\mu}_T^{(j)})$$

$$\text{for} \quad j \neq k,\ j,k = 1,2,\ldots,J.$$

The work presented below will describe the maximum likelihood procedures associated with the Observed Random Predictor Quantal Response Model for estimating the latent parameters. The term "latent parameters" as used here and in the work which follows includes the parameters from the Latent Random Predictor Quantal Response Model, the parameters describing the errors of measurement and for (4.2) the parameters which indicate a scale factor for each observed measurement. The term "observed parameters" includes the elements of Σ and $\underset{\sim}{\mu}_x^{(i)}$ (i = 1,...,J) as found in expression (2.2) of the Observed Random Predictor Quantal Response Model without application of any measurement model. That is, observed parameters, from Σ and $\underset{\sim}{\mu}_x^{(i)}$ (i = 1,2,...,J), represent population variances, covariances and means of the observed predictors with no consideration of latent predictors or errors of measurement.

The initial work will determine the conditions for the existence of estimates of the latent parameters and thus will demonstrate the need for (4.2) instead of (4.1) as the reformulation of the model (section C). Then the estimation procedure associated with reformulation (4.2) of the model will be described (section D).

Section C: Identifiability of the Models for the Covariance Matrix and Vectors of Category Means of the Reformulated Observed Random Predictor Quantal Response Model

Before estimation procedures associated with either (4.1) or (4.2) can be described, it is first necessary to determine the conditions under which estimates will exist.

Consider some model $\gamma = f(\theta)$ where γ and θ represent matrices of parameters and the elements of γ are known to be estimable.

<u>Definition</u>

(4.3) The parameters of θ are said to be <u>identifiable</u> if each parameter in θ can be uniquely defined as a function of parameters of γ.

When the parameters of some model are identifiable then the parameters can be estimated. Thus in order to describe estimation procedures for the parameters of the model for Σ and $\underset{\sim}{\mu}_x^{(i)}$ $(i = 1,2,\ldots,J)$ as given in (4.1), $\Sigma = \Phi + \Psi^2$ and $\underset{\sim}{\mu}_x^{(i)} = \underset{\sim}{\mu}_T^{(i)}$, or in (4.2), $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$ and $\underset{\sim}{\mu}_x^{(i)} = \Lambda\underset{\sim}{\mu}_T^{(i)}$, it is necessary to show that the latent parameters are identifiable. That is, it is necessary to show that each latent parameter can be expressed uniquely as a function of observed parameters.

If the latent parameters for models for Σ and $\underset{\sim}{\mu}_x^{(i)}$ $(i = 1,2,\ldots,J)$, whether given as in (4.1) or (4.2), are to be identifiable, definition (4.3) clearly implies that there must be at least as many observed parameters in Σ or $\underset{\sim}{\mu}_x^{(i)}$ as there are distinct latent parameters in the expression of the model.

Thus the approach to determining the identifiability of any model can begin by checking a simple counting condition. If there are at least as many observed parameters in Σ as there are latent parameters in the model for Σ then it is possible, but not guaranteed, that the model may be identified. However, if there are more distinct latent parameters in the expression of the model than there are observed parameters, the model is not identified and thus unique estimates of the latent parameters do not exist.

Before proceeding with the detailed examination of identifiability for the full models of (4.1) and (4.2) consider two examples.

Example 1. Let Σ be a covariance matrix where

$$\underset{2\times 2}{\Sigma} = \begin{bmatrix} \sigma^2_{X^1} & \sigma_{X^1X^2} \\ \sigma_{X^2X^1} & \sigma^2_{X^2} \end{bmatrix}$$

and let the structural model for Σ be $\underset{2\times 2}{\Sigma} = \underset{2\times 2}{\Phi} + \underset{2\times 2}{\Psi^2}$ where

$$\Phi = \begin{bmatrix} \sigma^2_{T^1} & \sigma_{T^1T^2} \\ \sigma_{T^1T^2} & \sigma^2_{T^2} \end{bmatrix} \quad \text{and} \quad \Psi^2 = \begin{bmatrix} \sigma^2_{E^1} & 0 \\ 0 & \sigma^2_{E^2} \end{bmatrix} .$$

There are 3 distinct observed parameters in Σ, i.e., $\sigma^2_{X^1}$, $\sigma^2_{X^2}$ and $\sigma_{X^1X^2}$ since $\sigma_{X^2X^1} = \sigma_{X^1X^2}$. There are 5 distinct latent parameters in the model for Σ, i.e., $\sigma^2_{T^1}$, $\sigma^2_{T^2}$, $\sigma_{T^1T^2}$, $\sigma^2_{E^1}$ and $\sigma^2_{E^2}$. Thus, since there are more distinct latent parameters than observed parameters, this model is not identified and unique estimates of the latent parameters do not exist.

Example 2. Let Σ be a covariance matrix where

$$\underset{2\times 2}{\Sigma} = \begin{bmatrix} \sigma^2_{X^1} & \sigma_{X^1X^2} \\ \sigma_{X^2X^1} & \sigma^2_{X^2} \end{bmatrix}$$

and let the structural model for Σ be $\underset{2\times 2}{\Sigma} = \underset{2\times 2}{\Phi} + \underset{2\times 2}{\Psi^2}$ where

$$\underset{2\times 2}{\Phi} = \begin{bmatrix} \sigma^2_{T^1} & \sigma_{T^1T^2} \\ \sigma_{T^1T^2} & \tfrac{1}{2}\sigma^2_{T^1} \end{bmatrix} \quad \text{and} \quad \underset{2\times 2}{\Psi^2} = \begin{bmatrix} \sigma^2_E & 0 \\ 0 & \sigma^2_E \end{bmatrix}.$$

There are 3 distinct observed parameters in Σ, i.e., $\sigma^2_{X^1}$, $\sigma^2_{X^2}$ and $\sigma_{X^1X^2}$. There are also 3 distinct latent parameters in the model for Σ, i.e., $\sigma^2_{T^1}$, $\sigma_{T^1T^2}$ and σ^2_E. Thus the preliminary counting requirement is satisfied. Now consider whether the latent parameters can be uniquely defined as functions of the observed parameters.

$$\Sigma = \begin{bmatrix} \sigma^2_{X^1} & \sigma_{X^1X^2} \\ \sigma_{X^2X^1} & \sigma^2_{X^2} \end{bmatrix} = \begin{bmatrix} \sigma^2_{T^1} + \sigma^2_{E^1} & \sigma_{T^1T^2} \\ \sigma_{T^1T^2} & \tfrac{1}{2}\sigma^2_{T^1} + \sigma^2_E \end{bmatrix} = \Phi + \Psi^2$$

i.e. $$\sigma^2_{X^1} = \sigma^2_{T^1} + \sigma^2_E$$

$$\sigma^2_{X^2} = \tfrac{1}{2}\sigma^2_{T^1} + \sigma^2_E$$

and $$\sigma_{X^1X^2} = \sigma_{X^2X^1} = \sigma_{T^1T^2} .$$

Therefore,

$$\sigma_{T^1T^2} = \sigma_{X^1X^2}$$

$$\sigma^2_{T^1} = 2(\sigma^2_{X^1} - \sigma^2_{X^2})$$

and
$$\sigma^2_E = \sigma^2_{X^1} - 2(\sigma^2_{X^1} - \sigma^2_{X^2}) = 2\sigma^2_{X^2} - \sigma^2_{X^1} .$$

The definition for identifiability (4.3) is satisfied for each latent parameter in the expression of the model for Σ. Since the model for Σ is identified, estimates of the latent parameters will exist. The two examples above pose two potential models for the same Σ, the second of which is identified while the first is not. A more detailed discussion of identifiability of the models for the general covariance matrix of this research will now be presented.

Consider a quantal response model with V observed predictors. Then Σ, the covariance matrix of observed predictors, assumed homogeneous across all categories, is a $V \times V$ matrix. Since Σ is symmetric, only the lower triangular portion of Σ (including the diagonal) will contain distinct observed parameters. There will be $1 + 2 + 3 + \ldots + V = \frac{V(V+1)}{2}$ distinct observed parameters in Σ. For some model for Σ, let r be the number of distinct latent parameters in the model. If $r > \frac{V(V+1)}{2}$ then there are more distinct latent parameters in the model for Σ than there are observed parameters in Σ and the model is not identified. If, however, $r \leq \frac{V(V+1)}{2}$ the counting condition is satisfied. That is, there are fewer distinct latent predictors in the model for Σ than there are observed parameters in Σ. Thus if each latent parameter can be expressed uniquely as a function of the observed parameters then the model is identified.

The question now arises about whether or not the model for Σ which results from applying the classical measurement model (2.24b), as in (4.1) above, is identified.

Recall, if there are p predictors, then Σ is a $p \times p$ symmetric matrix. The model for Σ (2.24b) is:

$$\underset{p\times p}{\Sigma} = \underset{p\times p}{\Phi} + \underset{p\times p}{\Psi^2}$$

where Φ is the covariance matrix of the latent predictors, and Ψ^2 is a diagonal matrix of error variances for the p predictors.

From example 1 above it was shown that when $p = 2$, this model for Σ is not identified. The general model for Σ with p predictors is also not identified for any value of p. There are $\frac{p(p+1)}{2}$ observed parameters in Σ. There are $\frac{p(p+1)}{2}$ distinct latent parameters in Φ and p distinct latent parameters in Ψ^2 for a total of $r = \frac{p(p+1)}{2} + p$ distinct latent parameters in the model for Σ. Thus there are more distinct latent parameters than observed parameters so the counting condition is not satisfied. Hence the model for Σ based on the classical measurement model (2.24b) and given in (4.1) is not identified.

Consider now the model for Σ based on the expanded measurement model (2.27b) and given in (4.2). This model for Σ, $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$, allows for the use of replicate measures, i.e., for multiple observed replications for a single latent predictor. This type of replicate measures is what Lord and Novick (1974) call nominally parallel measures.

In the model for Σ, $\underset{V\times V}{\Sigma} = \underset{V\times p}{\Lambda}\ \underset{p\times p}{\Phi}\ \underset{p\times V}{\Lambda'} + \underset{V\times V}{\Psi^2}$, there are p latent predictors and V observed replications associated with the p latent predictors by (2.26). Thus there are $\frac{V(V+1)}{2}$ observed parameters in Σ. There are $\frac{p(p+1)}{2}$ distinct latent parameters in Φ, $V - p$ distinct latent parameters in Λ, and V distinct latent parameters in Ψ^2 for a total of $r = V - p + \frac{p(p+1)}{2} + V$ distinct latent predictors in the model for Σ.

Note 1: There are $V - p$ latent predictors in Λ since each of the V observed replications is assigned a scale factor but at least one observed replication associated with each of the p latent predictors is assigned a scale factor of 1 which defines the metric of the true score.

Note 2: The counting condition will be satisfied when

$$\frac{V(V+1)}{2} \geq V - p + \frac{p(p+1)}{2} + V.$$

Consider the single predictor (p = 1) models. Here assume that there are V observed replications related to the single latent predictor by (2.26). Thus the model for Σ becomes $\underset{V\times V}{\Sigma} = \underset{V\times 1}{\Lambda}\ \sigma_T^2\ \underset{1\times V}{\Lambda'} + \underset{V\times V}{\Psi^2}$. There are $\frac{V(V+1)}{2}$ observed parameters in Σ. There are $V - 1$ distinct latent parameters in Λ, 1 latent parameter in Φ, i.e. σ_T^2, and V latent parameters in Ψ^2 for a total of $r = (V - 1) + 1 + V = 2V$ distinct latent parameters in the model for Σ. The counting condition will be satisfied if $\frac{V(V+1)}{2} \geq 2V$, that is if $V \geq 3$. Therefore there must be at least 3 observed replications of the single latent parameter to satisfy the counting condition.

Assume there are a total of K (where $K \geq 3$) observed replications for the single latent predictor, i.e., $\underset{\sim}{X}' = (X^1\ X^2\ \ldots\ X^K)$. The model for Σ is

$$(4.4) \qquad \underset{K\times K}{\Sigma} = \underset{K\times 1}{\tilde{\Lambda}}\ \underset{1\times 1}{\Phi}\ \underset{1\times K}{\tilde{\Lambda}'} + \underset{K\times K}{\Psi^2}$$

where

$$\underset{1\times K}{\tilde{\Lambda}'} = [1 \quad \lambda_2 \quad \ldots \quad \lambda_k]$$

$$\underset{1\times 1}{\Phi} = \sigma_T^2$$

and

$$\underset{K\times K}{\Psi^2} = \text{diag}[\sigma^2_{E^1} \quad \sigma^2_{E^2} \quad \ldots \quad \sigma^2_{E^K}] \ .$$

For this case $V = K \geq 3$. Expression (4.4) contains $\frac{K(K+1)}{2}$ observed parameters in Σ and $2K$ latent parameters in the model for Σ where $\frac{K(K+1)}{2} \geq 2K$ since $K \geq 3$. Expression (4.4) produces $\frac{K(K+1)}{2}$ simultaneous equations of the form

$$\Sigma_{ij} = f(\Lambda,\ \sigma_T^2,\ \Psi^2)$$

for $i,j = 1,2,\ldots,K$. Solving each of these expressions for latent parameters as functions of the observed parameters produces: (See Appendix C.1 for details)

$$\lambda_i = \frac{\sigma_{X^K X^i}}{\sigma_{X^K X^1}} \qquad \text{for} \quad i = 2,3,\ldots,K-1$$

$$\lambda_K = \frac{\sigma_{X^K X^2}}{\sigma_{X^2 X^1}}$$

$$\sigma_T^2 = \frac{\sigma_{X^2X^1}\,\sigma_{X^KX^1}}{\sigma_{X^KX^2}}$$

$$\sigma_{E^1}^2 = \sigma_{X^1}^2 - \frac{\sigma_{X^2X^1}\,\sigma_{X^KX^1}}{\sigma_{X^KX^2}}$$

$$\sigma_{E^i}^2 = \sigma_{X^i}^2 - \frac{\sigma_{X^KX^i}\,\sigma_{X^iX^1}}{\sigma_{X^KX^1}} \quad \text{for} \quad i = 2,3,\ldots,K-1$$

and

$$\sigma_{E^K}^2 = \sigma_{X^K}^2 - \frac{\sigma_{X^KX^2}\,\sigma_{X^KX^1}}{\sigma_{X^2X^1}} \;.$$

Thus a one predictor model is identified if there are at least three observed replications for the single latent predictor. Thus there will exist estimates of the latent parameters.

In this model there are $K(K + 1)/2$ observed parameters in Σ and $r = 2K$ distinct latent parameters in the model for Σ. When $K = 3$, there are 6 observed parameters and 6 latent parameters and since the model is identified it is said to be "just identified". When $K = 4$, there are 10 observed parameters and 8 latent parameters and the model is said to be "over-identified". When $K > 4$ the model will be over-identified. When $K = 2$, there are 3 observed parameters and 4 latent parameters and the model is not identified. The model with $K = 2$ is also said to be "under-identified".

Now consider the general model with p predictors $(p > 1)$. It will be shown that in order for the general model with p predictors to be identified <u>each</u> predictor must have at least two observed replications.

Suppose there are p predictors. Appendix C.2 demonstrates that there must be at least $p + 2$ observed replications, i.e., $V \geq p + 2$, in order for the counting condition to be satisfied. Therefore some one predictor must have at least three observed replications or at least two predictors must have two observed replications.

Consider a model with p predictors. Let some predictor i $(i = 1,2,\ldots,p)$ have exactly one observed measurement, i.e., $K_i = 1$. Let each of the other $p - 1$ predictors have K_j observed replications where $K_j \geq 1$ for $j = 1,2,\ldots,p$ with $j \neq i$, such that $V = \sum_{m=1}^{p} K_m \geq p + 2.$

If $\underset{1\times V}{\underset{\sim}{X}'} = [X_1^1 \ldots X_{K_1}^1 \mid \ldots \mid X_1^i \mid \ldots \mid X_1^p \ldots X_{K_p}^p]$ represents the observed replications of the p predictors then the model for Σ is

$$(4.5) \qquad \underset{V\times V}{\Sigma} = \underset{V\times p}{\Lambda} \; \underset{p\times p}{\Phi} \; \underset{p\times V}{\Lambda'} + \underset{V\times V}{\Psi^2} \; .$$

For this model there are $V(V + 1)/2$ observed parameters in Σ. There are $r = (V - p) + \frac{p(p + 1)}{2} + V = 2V + \frac{p(p - 1)}{2}$ distinct latent parameters in the model. Since $V \geq p + 2$ then $V(V + 1)/2 \geq 2V + \frac{p(p - 1)}{2}$ and the counting condition for identifiability is satisfied.

Expression (4.5) produces $V(V + 1)/2$ simultaneous equations of the form $\Sigma_{ij} = f(\Lambda, \Phi, \Psi^2)$ for $i,j = 1,2,\ldots,V$. For the predictor with only one observed replication, X_1^i, the equation for $\Sigma_{ii} = \sigma^2_{X_1^i}$ is:

$$(4.6) \qquad \sigma^2_{X_1^i} = \sigma^2_{T^i} + \sigma^2_{E_1^i} \; .$$

The parameters $\sigma^2_{T^i}$ and $\sigma^2_{E^i_1}$ occur together and only in the equation for $\sigma^2_{X^i_1}$ (4.6). Therefore a solution for $\sigma^2_{T^i}$ apart from $\sigma^2_{E^i_1}$ as a function of observed parameters will not exist.

Thus the definition of identifiability is not satisfied for the model for Σ as given by (4.5) if even one predictor has only one observed measurement. A more detailed algebraic exploration of this situation is contained in Appendix C.3.

Consider now a model with p predictors where each predictor has at least two observed replications, i.e., $K_i \geq 2$ for $i = 1,2,\ldots,p$. For $p > 1$, $V = \sum_{i=1}^{p} K_i \geq p + 2$ and thus the counting condition for identifiability will be satisfied. The model for Σ here has the same appearance as (4.5) only the internal structure differs.

$$\underset{V\times V}{\Sigma} = \underset{V\times p}{\Lambda} \; \underset{p\times p}{\Phi} \; \underset{p\times V}{\Lambda'} + \underset{V\times V}{\Psi^2} . \tag{4.7}$$

Expression (4.7) produces $V(V + 1)/2$ simultaneous equations of the form $\Sigma_{ij} = f(\Lambda, \Phi, \Psi^2)$ for $i,j = 1,2,\ldots,V$. There are $r = V - p + \frac{p(p + 1)}{2} + V = 2V + \frac{p(p - 1)}{2}$ latent parameters in the model for Σ. Each of these latent parameters can be expressed as a function of the observed parameters of Σ. See Appendix C.3 for the details. The results are presented below.

The $V - p$ parameters of Λ are:

$$\lambda^1_j = \frac{\sigma_{X^2_1 X^1_j}}{\sigma_{X^2_1 X^1_1}} \qquad \text{for} \quad j = 2,\ldots,K_1$$

($K_1 - 1$ number of parameters)

$$\lambda^i_j = \frac{\sigma_{X^i_j X^1_1}}{\sigma_{X^i_1 X^1_1}} \qquad \text{for } i = 2,\ldots,p,\ j = 2,\ldots,K_i$$

$$\Big(\sum_{i=2}^{p} (K_i - 1) = \sum_{i=2}^{p} K_i - (p-1) \quad \text{number of parameters}\Big)$$

where $K_1 - 1 + \sum_{i=2}^{p} K_i - (p-1) = \sum_{i=1}^{p} K_i - p = V - p.$

The $\frac{p(p + 1)}{2}$ parameters of Φ are:

$$\sigma_{T^i T^j} = \sigma_{X^i_1 X^j_1} \qquad \text{for } i = 1,2,\ldots,p \text{ with } i \neq j$$
$$j = 1,2,\ldots,p$$

$\Big(\frac{p(p - 1)}{2}$ number of distinct off-diagonal elements of $\Phi\Big)$

$$\sigma^2_{T^1} = \frac{\sigma_{X^1_2 X^1_1}\ \sigma_{X^2_1 X^1_1}}{\sigma_{X^2_1 X^1_2}}$$

(1 diagonal element of Φ)

$$\sigma^2_{T^i} = \frac{\sigma_{X^i_2 X^i_1}\ \sigma_{X^i_1 X^1_1}}{\sigma_{X^i_2 X^1_1}} \qquad \text{for } i = 2,\ldots,p$$

($p - 1$ number of diagonal elements of Φ)

where $\frac{p(p - 1)}{2} + 1 + p - 1 = \frac{p(p + 1)}{2}$.

The V parameters of Ψ^2 are:

$$\sigma^2_{E^1_1} = \sigma^2_{X^1_1} - \frac{\sigma_{X^1_2 X^1_1}\ \sigma_{X^2_1 X^1_1}}{\sigma_{X^2_1 X^1_1}} \qquad (1 \text{ element of } \psi^2)$$

for $i = 2,\ldots,p$

$$\sigma^2_{E^i_1} = \sigma^2_{X^i_1} - \frac{\sigma_{X^i_2 X^i_1}\,\sigma_{X^i_1 X^1_1}}{\sigma_{X^i_2 X^1_1}}$$

($p - 1$ number of elements of Ψ^2)

$$\sigma^2_{E^1_j} = \sigma^2_{X^1_j} - \left(\frac{\sigma_{X^2_1 X^1_j}}{\sigma_{X^2_1 X^1_1}}\right)^2 \cdot \left(\frac{\sigma_{X^1_2 X^1_1}\,\sigma_{X^2_1 X^1_1}}{\sigma_{X^2_1 X^1_2}}\right) \quad \text{for } j = 2,\ldots,K_1$$

($K_1 - 1$ number of elements of Ψ^2)

$$\sigma^2_{E^i_j} = \sigma^2_{X^i_j} - \left(\frac{\sigma_{X^i_j X^1_1}}{\sigma_{X^i_1 X^1_1}}\right)^2 \left(\frac{\sigma_{X^i_2 X^i_1}\,\sigma_{X^i_1 X^1_1}}{\sigma_{X^i_2 X^1_1}}\right) \quad \text{for } i = 2,\ldots,p \quad j = 2,\ldots,K_i$$

($\sum_{i=2}^{p} (K_i - 1)$ number of elements of Ψ^2)

where $1 + (p - 1) + (K_1 - 1) + \sum_{i=2}^{p} (K_i - 1) = p + \sum_{i=1}^{p} K_i - p = \sum_{i=1}^{p} K_i = V.$

Thus if <u>each</u> of the p predictors has at least two observed replications then each latent parameter in the model for Σ (4.7) can be expressed as a function of observed parameters of Σ. Thus the model (4.7) for Σ is identified, therefore estimates of the latent parameters will exist.

Consider now the model for the mean vector of observed replications for some category i ($i = 1,2,\ldots,J$), i.e., $\underset{\sim}{\mu}^{(i)}_X$. Applying the expanded measurement model (2.26) produces the following model for $\underset{\sim}{\mu}^{(i)}_X$:

(4.8) $$\underset{V\times 1}{\underline{\mu}_X^{(i)}} = \underset{V\times p}{\Lambda}\ \underset{p\times 1}{\underline{\mu}_T^{(i)}} \qquad \text{for some } i = 1,2,\ldots,J.$$

There are V observed parameters in $\underline{\mu}_X^{(i)}$. There are $V - p$ latent parameters in Λ and p latent parameters in $\underline{\mu}_T^{(i)}$ for a total of $r = V - p + p = V$ distinct latent parameters in the model. Thus the counting condition for identifiability is satisfied.

Because of the special nature of Λ (2.26), it is clear that $\Lambda'\Lambda$ will be a diagonal matrix of full rank i.e.

(4.9) $$\underset{p\times V}{\Lambda'}\ \underset{V\times p}{\Lambda} = \begin{bmatrix} 1 + \sum_{i=2}^{K_1} (\lambda_i^1)^2 & 0 & \cdots & 0 \\ 0 & 1 + \sum_{i=2}^{K_2} (\lambda_i^2)^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 + \sum_{i=2}^{K_p} (\lambda_i^p)^2 \end{bmatrix}.$$

Note: $\underset{p\times V}{\Lambda'}\ \underset{V\times p}{\Lambda}$ will be less than full rank if and only if $\sum_{i=1}^{K_i} (\lambda_i^j)^2 = -1$, for some predictor T^j $(j = 1,2,\ldots,p)$. This is impossible. Therefore $\Lambda'\Lambda$ will be of full rank p and thus will possess an inverse, $(\Lambda'\Lambda)^{-1}$.

Thus it is possible to express $\underline{\mu}_T^{(i)}$ as a function of $\underline{\mu}_X^{(i)}$ $(i = 1,2,\ldots,J)$ and Λ, i.e.,

(4.10) $$\underline{\mu}_T^{(i)} = (\Lambda'\Lambda)^{-1}\Lambda'\underline{\mu}_X^{(i)} \qquad \text{for } i = 1,2,\ldots,J.$$

Since the latent parameters of Λ can be expressed as a function of observed parameters in Σ based on the work for Σ above, result (4.10) indicates that the latent parameters in the mean vector for any category can be expressed as functions of observed parameters from that same category involving parameters from the covariance matrix and the mean vector for the observed replications. Thus the definition for identifiability is satisfied for $\underset{\sim}{\mu}_T^{(i)}$ $(i = 1,2,\ldots,J)$.

Therefore since the models for Σ and $\underset{\sim}{\mu}_X^{(i)}$ $(i = 1,2,\ldots,J)$ as used in reformulation (4.2) are both identified, it will then be possible to produce estimates for all the latent parameters in the reformulation (4.2). And using appropriate estimates it will be possible to construct an estimate for the latent weighting coefficients. This will be discussed in greater detail in Section D below as part of the description of the maximum likelihood estimation procedures associated with the reformulated model (4.2).

Before beginning the discussion on estimation procedures one additional topic relative to identifiability needs to be discussed briefly. Consider some non-identified model which expresses the covariance matrix Σ in terms of latent predictors, e.g., the model of example 1 above where $\Sigma = \Phi + \Psi^2$ with

$$\Sigma = \begin{bmatrix} \sigma^2_{X^1} & \sigma_{X^1X^2} \\ \sigma_{X^2X^1} & \sigma^2_{X^2} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \sigma^2_{T^1} & \sigma_{T^1T^2} \\ \sigma_{T^1T^2} & \sigma^2_{T^2} \end{bmatrix} \quad \text{and} \quad \Psi^2 = \begin{bmatrix} \sigma^2_{E^1} & 0 \\ 0 & \sigma_{E^2}^2 \end{bmatrix}.$$

It was shown above that this model for Σ is not identified.

A question which arises relative to non-identified models such as those in example 1 is whether it is possible to modify or extend the model in some fashion so that the modified or extended model is identified.

Two general approaches to the modification of non-identified models to produce an identified model are possible. For convenience, these two approaches will be presented in reference to models for the covariance matrix Σ as considered in this research.

The first approach attacks the problem of the non-identified model for Σ by attempting to increase the number of observed parameters in Σ without producing an equivalent increase in the number of latent parameters in the model. This is done by the use of multiple observed replications for each latent predictor.

The model for Σ based on the expanded measurement model (2.27b) is an example of the use of this approach. The model for Σ based on the classical measurement model (2.24b) was not identified. By the appropriate inclusion of replicate measures an identified model (2.27b) for Σ was produced. As noted above, it is not sufficient to indiscriminately include enough replicate measures to satisfy only the counting condition for identifiability. The pattern of replicate measures to be included in order to achieve identifiability of the model is crucial.

Since this approach was discussed in detail above for models for Σ it will not be pursued further here.

The second approach attacks the problem of the non-identified model for Σ by attempting to reduce the number of distinct latent parameters in the model for Σ. This is done by introducing constraints on the latent parameters. The process for introducing constraints on the latent parameters is to require that one or more of the latent parameters be given as unique functions of other latent parameters, thus reducing the number of distinct parameters in the model. Typically the constraints involve requiring two or more latent parameters to have the same value.

Example 2 above is an example of the use of constraints on the latent parameters to achieve identifiability. The model for example 2 can be produced from the model for example 1 by introducing the following constraints on some of the latent parameters of the model for example 1:

$$\text{let} \quad \sigma^2_{T^2} = 1/2\,\sigma^2_{T^1} \qquad \text{and} \qquad \sigma^2_E = \sigma^2_{E^1} = \sigma^2_{E^2} \; .$$

i.e.,

$$\Phi = \begin{bmatrix} \sigma^2_{T^2} & \sigma_{T^1T^2} \\ \sigma_{T^1T^2} & \tfrac{1}{2}\sigma^2_{T^1} \end{bmatrix} \qquad \text{and} \qquad \Psi^2 = \begin{bmatrix} \sigma^2_E & 0 \\ 0 & \sigma^2_E \end{bmatrix} \; .$$

By introducing these constraints on the latent parameters a non-identified model is modified into a model which is identified.

A word of caution is necessary here. The constraints to be imposed on the latent parameters of a model for Σ should be reasonable in terms of the situation to be analyzed. To introduce constraints which have no support in the situation merely to produce an algebraically identified model will provoke problems in the interpretation of results.

Since the number of possible combinations and types of constraints can be myriad even in a relatively simple model for Σ, further discussion for this approach will center on a few specific forms of constraints which may be reasonable in some situations under analysis.

Consider the single predictor situation (p = 1). With only one observed measurement of the single predictor, the only type of constraint which will produce an identified model is if the error variance can be considered to be a known function of the true score variance. This is rather unlikely for most situations and will not be pursued further.

Consider the multiple predictor situation (p > 1). Sometimes an identified model can be produced from a non-identified model by introducing constraints among the parameters of Ψ^2, i.e., among the parameters describing the errors of measurement. The simplest of this type of constraint assumes that some error variance is equal to some other error variance.

An example of a non-identified model for Σ where this simplest type of constraint among the error variances produces an identified model is a model for Σ similar to that given by (4.5). Recall, for this model there is one predictor i $(i = 1,2,\ldots,p)$ which has exactly one observed measurement $(K_i = 1)$. If each of the other predictors has at least two observed replications $(K_j \geq 2$ for $j = 1,2,\ldots,0$ with $j \neq i)$ then identifiability can be achieved by imposing a constraint of the form $\sigma^2_{E^i_1} = \sigma^2_{E^m_\ell}$ for some $m \neq i$, $m = 1,2,\ldots,p$ and some $1 = 1,2,\ldots,K_m$. This constraint requires that the error variance associated with the single predictor i $(i = 1,2,\ldots,p)$ is equal to the error variance for some replication 1 $(1 = 1,2,\ldots,K_m)$ of some other predictor variable m $(m = 1,2,\ldots,p$ with $m \neq 1)$.

It is possible to express all latent parameters of this model, except $\sigma^2_{T^i}$ and $\sigma^2_{E^i_1}$ (where $\sigma^2_{X^i_1} = \sigma^2_{T^i} + \sigma^2_{E^i_1}$) as functions of observed parameters using techniques similar to those used to show that the model (4.7) for Σ was identifiable. This is possible without the use of the constraint as long as $K_j \geq 2$ for $j = 1,2,...,p$ and $j \neq i$. Under the imposition of the constraint the expression for $\sigma^2_{E^m_\ell}$ as a function of observed parameters will also provide the expression for $\sigma^2_{E^i_1}$ as a function of observed parameters of Σ since $\sigma^2_{E^i_1} = \sigma^2_{E^m_\ell} = f(\Lambda, \Phi, \Psi^2)$. Thus $\sigma^2_{T^i} = \sigma^2_{X^i_1} - \sigma^2_{E^i_1}$ where $\sigma^2_{E^i_1}$ can be expressed as a function of observed parameters of Σ. Therefore with this one simple constraint imposed upon the error variances a non-identified model has been modified into an identified model.

A more extreme extension of the imposition of constraints on parameters of Ψ^2 occurs when all error variances are constrained to be equal across all V observed replications of the p predictors. This constraint can be expressed as $\Psi^2 = \sigma^2_E I$ where σ^2_E is the common value of all the error variances and I is the identity matrix of rank V.

An example of a non-identified model which can be modified into an identified model through the use of the constraint $\Psi^2 = \sigma^2_E I$ will be presented.

Consider a model with p predictors ($p > 1$). Let some one of the predictors have two observed replications, i.e., $K_i = 2$ for some $i = 1,2,...,p$, while each of the other predictors has precisely one observed measurement, i.e., $K_j = 1$ for $j = 1,2,...,p$ with

$j \neq i$. In this case $V = \sum_{m=1}^{p} K_m = p + 1$. As given without any constraints, the model is clearly not identified since it does not satisfy the counting condition that $V \geq p + 2$ (see Appendix C.2 for details). To reduce the number of distinct latent parameters in the model for Σ let $\Psi^2 = \sigma_E^2 I$ where I is the identity matrix of rank V. The two observed replications for predictor i are noted as X_1^i and X_2^i. The single observed measurement for the other j predictors is noted as X_1^j $(j \neq i)$.

Thus the model for Σ is:

$$\text{(4.11)} \qquad \underset{V\times V}{\Sigma} = \underset{V\times p}{\Lambda} \; \underset{p\times p}{\Phi} \; \underset{p\times V}{\Lambda'} + \underset{V\times V}{\Psi^2}$$

where

$$\underset{V\times V}{\Sigma} = \left[\begin{array}{cccccc}
\sigma^2_{X_1^1} & & & & & \\
\sigma_{X_1^2 X_1^1} & \sigma^2_{X_1^2} & & \text{symmetric} & & \\
\vdots & \vdots & \ddots & & & \\
\hline
\sigma_{X_1^i X_1^1} & \sigma_{X_1^i X_1^2} & \cdots & \sigma^2_{X_1^i} & & \\
\sigma_{X_2^i X_1^1} & \sigma_{X_2^i X_1^2} & \cdots & \sigma_{X_2^i X_1^i} & \sigma^2_{X_2^i} & \\
\hline
\vdots & \vdots & & \vdots & \vdots & \ddots \\
\sigma_{X_1^p X_1^1} & \sigma_{X_1^p X_1^2} & \cdots & \sigma_{X_1^p X_1^i} & \sigma_{X_1^p X_2^i} & \cdots \; \sigma^2_{X_1^p}
\end{array}\right]$$

$$\underset{V\times p}{\Lambda} = \left[\begin{array}{cccccccc}
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\hdashline
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & \lambda_2^1 & 0 & 0 & \cdots & 0 \\
\hdashline
0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1
\end{array}\right]$$

$$\underset{p\times p}{\Phi} = \left[\begin{array}{cccccc}
\sigma^2_{T^1} & & & \text{symmetric} & & \\
\sigma_{T^2T^1} & \sigma^2_{T^2} & & & & \\
\vdots & \vdots & \ddots & & & \\
\hdashline
\sigma_{T^iT^1} & \sigma_{T^iT^2} & \cdots & \sigma^2_{T^i} & & \\
\hdashline
\vdots & \vdots & & \vdots & \ddots & \\
\sigma_{T^pT^1} & \sigma_{T^pT^2} & \cdots & \sigma_{T^pT^i} & \cdots & \sigma^2_{T^p}
\end{array}\right]$$

$$\underset{V\times V}{\Psi^2} = \left[\begin{array}{cccccc}
\sigma^2_E & & & & & \\
0 & \sigma^2_E & & \text{symmetric} & & \\
\vdots & \vdots & \ddots & & & \\
\hdashline
0 & 0 & \cdots & \sigma^2_E & & \\
\hdashline
\vdots & \vdots & & \vdots & \ddots & \\
0 & 0 & \cdots & 0 & \cdots & \sigma^2_E
\end{array}\right]$$

In this model for Φ, (4.11), there are $V(V + 1)/2 = (p + 1)(p + 2)/2 = p(p + 1)/2 + p + 1$ observed parameters in Σ since $V = p + 1$. There are $p(p + 1)/2$ latent parameters in Φ, one latent parameter in Λ (i.e., λ_2^i) and one latent parameter in Ψ^2 (i.e., σ_E^2). Thus there is a total of $r = p(p + 1)/2 + 2$ latent parameters in the model for Σ. Since $p > 1$, then $p(p + 1)/2 + p + 1 > p(p + 1)/2 + 2$ and the counting condition for identifiability is satisfied since there are more observed parameters in Σ than there are distinct latent parameters in the model for Σ.

Expression (4.11) produces $p(p + 1)/2 + p + 1$ simultaneous equations of the form $\Sigma_{ij} = f(\Lambda, \Phi, \Psi^2)$. Each of the $p(p + 1)/2 + 2$ distinct latent parameters in the model for Σ can be expressed as functions of observed parameters. See Appendix C.4 for additional details. The results are presented below.

The one latent parameter of Λ is:

$$\lambda_2^i = \frac{\sigma_{X_2^i X_1^1}}{\sigma_{X_1^i X_1^1}} \qquad \text{for some specified } i \quad (i = 1,2,\ldots,p).$$

The $p(p + 1)/2$ latent parameters of Φ are:

$$\sigma_{T^k T^j} = \sigma_{X_1^k X_1^j} \qquad \text{for } k \neq j \text{ with } k,j = 1,2,\ldots,p$$

($p(p - 1)/2$ number of off-diagonal elements of Φ).

$$\sigma^2_{T^i} = \frac{\sigma_{X_2^i X_1^i}\ \sigma_{X_1^i X_1^1}}{\sigma_{X_2^i X_1^1}} \quad \text{for some specified } i \quad (i = 1,2,\ldots,p)$$

(one diagonal element of Φ)

$$\sigma^2_{T^j} = \sigma^2_{X^j_1} - \sigma^2_{X^i_1} + \frac{\sigma_{X^i_2 X^i_1}\sigma_{X^i_1 X^1_1}}{\sigma_{X^i_2 X^1_1}} \quad \text{for } j = 1,2,\ldots,p \text{ with } j \neq i$$

($p - 1$ number of diagonal elements of Φ).

The single latent parameter of Ψ^2 is:

$$\sigma^2_E = \sigma^2_{X^i_1} - \frac{\sigma_{X^i_2 X^i_1}\sigma_{X^i_1 X^1_1}}{\sigma_{X^i_2 X^1_1}} \quad \text{for some specified } i \quad (i = 1,2,\ldots,p).$$

Where $r = 1 + p(p - 1)/2 + 1 + p - 1 + 1 = p(p + 1)/2 + 2$ number of latent parameters in the model (4.11) for Σ. Thus the model (4.11) for Σ is identified, since each latent parameter in the model for Σ can be expressed as a function of observed parameters in Σ.

In the work above the only constraints which were considered involved the parameters of Ψ^2, that is, the error variances. These are not the only constraints which are possible for use. It is possible to impose constraints on elements of Λ or Φ as well as on elements of Ψ^2. It is even possible to impose constraints which involve elements of any of the three latent parameter matrices in the model for Σ simultaneously, e.g., $\lambda^1_3 = \sigma^2_{T^5} = \sigma^2_{E^4_2}$. The major question to be answered though, concerns not what constraints are possible but what constraints are reasonable for the given situation. This criterion of reasonableness should be the first priority in any consideration of constraints for a proposed model.

The brief work above does not even begin to exhaust the possibilities for the use of constraints to modify models to achieve identifiability. The few examples given were merely to illustrate some of the potential of this approach.

Summary for Section C

This section has included an examination of the identifiability of models for Σ and $\underset{\sim}{\mu}_{\mathbf{x}}^{(i)}$ $(i = 1,2,\ldots,J)$. The model for Σ based on the classical measurement model (2.24b) as included in (4.1) was shown to be not identified. Thus unique estimates for the latent parameters of the model will not exist. However, by the inclusion of multiple observed replications for each predictor (with at least two observed replications for each predictor) the model for Σ based on the expanded measurement model (2.27b) and the model for $\underset{\sim}{\mu}_{\mathbf{x}}^{(i)}$ $(i = 1,2,\ldots,J)$ based on (2.27a) were shown to be identified.

Two approaches to the modification of non-identified models in an attempt to produce identified models were presented. One approach involved the inclusion of replicate observed measurements for the predictors. The other approach involved imposing constraints on the latent parameters of the model. In many situations the most appropriate procedure to modify a non-identified model to produce an identified model will involve a combination of both approaches. That is, include observed replicate measurements and impose constraints on latent parameters of the model.

Any model for Σ in terms of latent parameters which is to be used in an estimation procedure should first be examined carefully to ensure that the model is identified. This examination for identifiability should be conducted whether or not observed replications of the predictors are included or whether or not constraints are imposed on the latent parameters.

For the remainder of this research, unless otherwise indicated, the assumption will be made that all models which involve a structure for Σ have been checked and found to be identified.

Section D: Maximum Likelihood Estimation Procedures Associated with the Reformulated Observed Random Predictor Quantal Response Model

For this section maximum likelihood estimation procedures associated with the Observed Random Predictor Quantal Response Model (2.2) will be described. In this model the vectors of category means, $\underset{\sim}{\mu}_X^{(i)}$ $(i = 1,2,\ldots,J)$, and the covariance matrix, Σ, have structures given by (2.27a) and (2.27b) based on the application of the expanded measurement model (2.26). Expression (4.2) results from (2.2) when the structures of the parameter matrices are displayed. The models of interest here will be assumed to be identified and thus estimates of the latent parameters in (4.2) will exist.

The structure imposed on the parameter matrices by the application of the expanded measurement model (2.26) is not apparent in the expression of the model given by (2.2). Thus the model (2.2) has the same appearance as the general case model examined by McSweeney and Schmidt (1974). Therefore the derivation of the likelihood function and the logarithm of the likelihood function produced by McSweeney and Schmidt (1974) is appropriate for presentation here.

Recall first that in (2.2) $\underset{\sim}{X}$ is the $V \times 1$ vector of observed replications for the p predictors which has the structure $\underset{\sim}{X} = \Lambda\underset{\sim}{T} + \underset{\sim}{E}$, from (2.26). For each category of the criterion $\underset{\sim}{X}$ is normally distributed with $V \times 1$ mean vector $\underset{\sim}{\mu}_X^{(i)}$, where $\underset{\sim}{\mu}_X^{(i)} = \underset{\sim}{\mu}_T^{(i)}$ $(i = 1,2,\ldots,J)$ from (2.27a), and $V \times V$ covariance matrix Σ which is assumed homogeneous across all categories, where $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$ from (2.27b).

In order to apply the maximum likelihood estimation procedures associated with reformulation (4.2) of model (2.2) it is necessary to have a random sample of subjects from each category of the criterion with n_j subjects from category j $(j = 1,2,\ldots,J)$ of the criterion. Thus there is a total of n subjects from all categories, i.e., $n = \sum_{j=1}^{J} n_j$. Let $\bar{\underset{\sim}{X}}^{(j)}$ represent the $V \times 1$ vector of sample means for the observed replications in category j and S_j^+ represent the $V \times V$ matrix of sums of squares and cross-product deviations about the respective means for the observed replications in categoty j $(j = 1,2,\ldots,J)$.

Based on the presentation in McSweeney and Schmidt (1974, p. 13) the effective part of the logarithm of the likelihood function can be written as:

$$(4.12)\quad \ln L' = \sum_{j=1}^{J-1} n_j \ln p_j + (n - \sum_{j=1}^{J-1} n_j)\ln(1 - \sum_{j=1}^{J-1} p_j)$$
$$- \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{j=1}^{J} \operatorname{tr}(\Sigma^{-1}S_j^+)$$
$$- \frac{1}{2}\sum_{j=1}^{J} n_j(\bar{\underset{\sim}{X}}^{(j)} - \underset{\sim}{\mu}_X^{(j)})'\Sigma^{-1}(\bar{\underset{\sim}{X}}^{(j)} - \underset{\sim}{\mu}_X^{(j)}) .$$

The maximum likelihood estimators for p_j and $\underset{\sim}{\mu}_X^{(j)}$ are then indicated:

$$\hat{p}_j = \frac{n_j}{n} \quad \text{for } j = 1,2,\ldots,J$$

$$(4.13)\qquad \text{and}\quad \hat{\underset{\sim}{\mu}}_X^{(j)} = \bar{\underset{\sim}{X}}^{(j)} \qquad \text{for } j = 1,2,\ldots,J.$$

The procedures presented by McSweeney and Schmidt (1974) for the estimation of Σ will be of little help in determining estimates of the latent parameter matrices $(\Lambda, \Phi$ and $\Psi^2)$ in the structure for Σ.

Thus consider the effective part of the logarithm of the likelihood (4.12) for estimating components of Σ, i.e.,

$$(4.14a) \quad \ln L'' = -\frac{n}{2} \ln|\Sigma| - \frac{1}{2} \sum_{j=1}^{J} tr(\Sigma^{-1} S_j^{+})$$

$$- \frac{1}{2} \sum_{j=1}^{J} n_j (\underset{\sim}{\bar{X}}^{(j)} - \underset{\sim}{\mu}_X^{(j)})' \Sigma^{-1} (\underset{\sim}{\bar{X}}^{(j)} - \underset{\sim}{\mu}_X^{(j)}) .$$

Consider the last term in (4.14a). Let $C = -\frac{1}{2} \sum_{j=1}^{J} n_j (\underset{\sim}{\bar{X}}^{(j)} - \underset{\sim}{\mu}_X^{(j)})' \Sigma^{-1} (\underset{\sim}{\bar{X}}^{(j)} - \underset{\sim}{\mu}_X^{(j)})$. McSweeney and Schmidt (1974) have shown that $\hat{\underset{\sim}{\mu}}_X^{(j)} = \underset{\sim}{\bar{X}}^{(j)}$ will maximize $\ln L''$ and L, the likelihood function. Thus $\hat{\underset{\sim}{\mu}}_X^{(j)} = \underset{\sim}{\bar{X}}^{(j)}$ is the maximum likelihood estimate for $\underset{\sim}{\mu}_X^{(j)}$. Therefore there is no need to continue to include C in the expression for $\ln L''$, since its contribution to maximizing $\ln L''$ occurs for $\hat{\underset{\sim}{\mu}}_X^{(j)} = \underset{\sim}{\bar{X}}^{(j)}$, i.e., $C = 0$.

Note: $S_j^{+} = n_j S_j$ where S_j is the sample covariance matrix of the V observed replications in category j.

Note also: $$\sum_{j=1}^{J} tr(\Sigma^{-1} S_j^{+}) = \sum_{j=1}^{J} tr(\Sigma^{-1} n_j S_j)$$

$$= tr\{ \sum_{j=1}^{J} (\Sigma^{-1} n_j S_j)\}$$

$$= tr\{ \Sigma^{-1} (\sum_{j=1}^{J} n_j S_j)\}$$

$$= tr\{\Sigma^{-1} n \; S_p\}$$

$$= n \; tr\{\Sigma^{-1} S_p\}$$

where $$n = \sum_{j=1}^{J} n_j$$

and $S_p = \frac{\sum_{j=1}^{J} n_j S_j}{n}$ i.e., S_p is the pooled sample covariance matrix of the observed replications.

Therefore (4.14a) can be rewritten as:

(4.14b) $\ln L'' = -\frac{n}{2}\ln|\Sigma| - \frac{n}{2}\operatorname{tr}\{\Sigma^{-1}S_p\}.$

Or when the structure for Σ is indicated (4.14b) can be reformulated as:

(4.14c) $\ln L'' = -\frac{n}{2}\ln|\Lambda\Phi\Lambda' + \Psi^2| - \frac{n}{2}\operatorname{tr}\{(\Lambda\Phi\Lambda' + \Psi^2)^{-1}S_p\}.$

The problem now is to find values of Λ, Φ and Ψ^2 which will maximize $\ln L''$. Let $F = -\ln L''$, thus maximizing $\ln L''$ is equivalent to minimizing F where F can be written as:

(4.15) $F = \frac{n}{2}\ln|\Lambda\Phi\Lambda' + \Psi^2| + \frac{n}{2}\operatorname{tr}\{(\Lambda\Phi\Lambda' + \Psi^2)^{-1}S_p\}.$

The values of the elements of Λ, Φ and Ψ^2 which minimize F and thus maximize $\ln L''$, for the given pooled sample covariance matrix S_p, will be the maximum likelihood estimates of the latent parameter elements of Λ, Φ and Ψ^2.

The problem of minimizing an expression F such as (4.15), which is a function of a covariance matrix Σ with a given structure, is a common problem encountered in the set of procedures termed Analysis of Covariance Structures (ANCOVST). Wiley, Schmidt, and Bramble (1973) indicate that "Covariance structure analysis is a term used to describe s recently developed series of procedures and models which are used for the structural analysis of covariance matrices" (p. 317). Both Jöreskog (1970) and Wiley, Schmidt and Bramble (1973) indicate that the minimization of F as a function of the elements of Λ, Φ and Ψ^2 in the structure for Σ can be carried out by an application of the numerical method of Fletcher and Powell (1963).

The application of this numerical method requires expressions for the derivatives of F with respect to the elements of each of the latent parameter matrices, Λ, Φ and Ψ^2. These derivatives are presented by Jöreskog (1970) for a more general model of the structure of Σ than that employed in this research. The results presented by Jöreskog (1970) for the derivatives of F have been verified by derivations contained in Appendix C.5 and are:

(4.16a) $$\frac{\partial F}{\partial \Lambda_{ij}} = \begin{cases} 0 & \text{if } \Lambda_{ij} = \text{constant} \\ 2(\Sigma^{-1}[\Sigma - S_p]\Sigma^{-1}\Lambda\Phi)_{ij} & \text{if } \Lambda_{ij} = \text{parameter} \end{cases}$$

for $i = 1,2,\ldots,V$

$j = 1,2,\ldots,p$

(4.16b) $$\frac{\partial F}{\partial \Phi_{ij}} = \begin{cases} 2(\Lambda'\Sigma^{-1}[\Sigma - S_p]\Sigma^{-1}\Lambda)_{ij} & \text{for } i \neq j \\ (\Lambda'\Sigma^{-1}[\Sigma - S_p]\Sigma^{-1}\Lambda)_{ij} & \text{for } i = j \end{cases}$$

for $i,j = 1,2,\ldots,p$

(4.16c) $$\frac{\partial F}{\partial \Psi_{ii}} = 2(\Sigma^{-1}[\Sigma - S_p]\Sigma^{-1}\Psi)_{ii} \qquad \text{for } i = 1,2,\ldots,V$$

where $\Psi^2 = \Psi \cdot \Psi$ and Ψ_{ii} is the ith diagonal element of Ψ.

A numerical approximation procedure is typically needed to produce values of the estimates of the latent parameters in Σ when ANCOVST procedures are being employed. When a structure is hypothesized for Σ such as (2.27b) the standard maximum likelihood estimation procedures will typically not be applicable, since the set of simultaneous equations gained by setting equal to zero the derivatives of F with respect to the elements of the parameter matrices in the structure for Σ will not, in general, be explicitly solvable.

Since the structure being hypothesized for Σ for this area of this research, that is, $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$ (2.27b) is completely consistent with a special case of the general model presented by Jöreskog (1970) and with model (8) presented by Wiley, Schmidt and Bramble (1973), the estimation procedures described in either reference (which differ only in minor details) will apply for the model for Σ for this research.

Thus numerical values for the estimates of each latent parameter can be produced. That is, the maximum likelihood estimates $\hat{\Lambda}$, $\hat{\Phi}$ and $\hat{\Psi}^2$ will exist. As noted above the values of these estimates will be the values which minimize F.

The original interest of this chapter was to develop estimates for the latent weighting coefficients, $\underset{\sim}{\beta}^*_{k\cdot j}$ $(j \neq k,\ j,k = 1,2,\ldots,J)$ of the Latent Random Predictor Quantal Response Model (2.19) using estimates of latent parameters from the reformulated Observed Random Predictor Quantal Response Model (4.2). Recall that by a result derived in Appendix A.2 only a base set of $J - 1$ vectors of weighting coefficients associated with some arbitrarily selected category need be derived. All other vectors of weighting coefficients can then be produced from linear combinations of vectors of weighting coefficients in the base set. Since any category can be selected to provide the reference for the base set of vectors, select the first category for convenience, that is, the category associated with $Y = 1$. Therefore, the $J - 1$ vectors of weighting coefficients in the base set will have the form:

$$\underset{\sim}{\beta}^{*}_{1\cdot j} = \Phi^{-1}(\underset{\sim}{\mu}_{T}^{(1)} - \underset{\sim}{\mu}_{T}^{(j)}) \quad \text{for } j = 2,3,\ldots,J. \tag{4.17}$$

In order to estimate the elements of $\underset{\sim}{\beta}^{*}_{1\cdot j}$, estimates of Φ (hence Φ^{-1}) and $\underset{\sim}{\mu}_{T}^{(i)}$ for $i = 1,2,\ldots,J$ are needed. The ANCOVST estimation procedures, applied to Φ, described by Jöreskog (1970) or Wiley, Schmidt and Bramble (1973) will produce an estimate for Φ, call it $\hat{\Phi}$. In order to estimate the vectors of latent predictor means, $\underset{\sim}{\mu}_{T}^{(i)}$ for $i = 1,2,\ldots,J$, recall that (4.10) provides a formulation for $\underset{\sim}{\mu}_{T}^{(i)}$ as a function of Λ and $\underset{\sim}{\mu}_{X}^{(i)}$, i.e., $\underset{\sim}{\mu}_{T}^{(i)} = (\Lambda'\Lambda)^{-1}\Lambda'\underset{\sim}{\mu}_{X}^{(j)}$ $(i = 1,\ldots,J)$. An estimate of Λ, call it $\hat{\Lambda}$, will be available from the ANCOVST estimation procedures applied to Σ. An estimate of $\underset{\sim}{\mu}_{X}^{(i)}$ $(i = 1,\ldots,J)$ was derived by McSweeney and Schmidt (1974), that is, $\hat{\underset{\sim}{\mu}}_{X}^{(i)} = \overline{X}^{(i)}$ (4.13) where $\overline{X}^{(i)}$ is the sample mean of the observed replications in category i $(i = 1,\ldots,J)$. Therefore an estimate of $\underset{\sim}{\mu}_{T}^{(i)}$, call it $\hat{\underset{\sim}{\mu}}_{T}^{(i)}$, can be written as:

$$\hat{\underset{\sim}{\mu}}_{T}^{(i)} = (\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}\hat{\underset{\sim}{\mu}}_{X}^{(i)} \tag{4.18}$$

or

$$\hat{\underset{\sim}{\mu}}_{T}^{(i)} = (\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}\ \underset{\sim}{\overline{X}}^{(i)} \qquad (i = 1,2,\ldots,J).$$

Thus the estimates of the vectors of weighting coefficients for the base set will have the following formulation:

$$\hat{\underset{\sim}{\beta}}^{*}_{1\cdot j} = \hat{\Phi}^{-1} (\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}\ \overline{\underset{\approx}{X}}^{(1)} - (\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}\ \overline{\underset{\approx}{X}}^{(j)}$$

or

$$\hat{\underset{\sim}{\beta}}^{*}_{1\cdot j} = \hat{\Phi}^{-1}(\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}(\overline{\underset{\approx}{X}}^{(1)} - \overline{\underset{\approx}{X}}^{(j)}) \qquad (j = 2,\ldots,J) \tag{4.19}$$

where the estimates $\hat{\Phi}$ and $\hat{\Lambda}$ will be produced from the application of ANCOVST numerical approximation procedures to the structure for Σ.

Section E: Summary

The purpose of this chapter was to describe models with their associated estimation procedures which would produce estimates of the latent weighting coefficients from the Latent Random Predictor Quantal Response Model (2.19). Since the variables which are available for use as predictors typically contain errors of measurement, direct application of the Latent Random Predictor Quantal Response Model is not appropriate.

In section B two major reformulations of the Observed Random Predictor Quantal Response Model were provided. The reformulation (4.1) is based on the application of the classical measurement model (2.22) while the reformulation (4.2) is based on the expanded measurement model (2.26) which allows for multiple observed replications of the predictors.

To determine whether or not estimates will exist, the identifiability of various models for Σ and $\underset{\sim}{\mu}_X^{(i)}$ $(i = 1,2,...,J)$, as contained in the two formulations, was examined in section C. Since the model for Σ contained in reformulation (4.1) was not identified, no unique estimates of the latent parameters in the model for Σ can be found. However, the models for Σ and $\underset{\sim}{\mu}_X^{(i)}$ $(i = 1,2,...,J)$ contained in reformulation (4.2) were shown to be identified under several combinations of inclusion of replicate measures and imposition of constraints. Thus the estimation procedures presented in section D were those associated with reformulation (4.2) of the Observed Random Predictor Quantal Response Model.

In section D, the estimation procedures described by McSweeney and Schmidt (1974) were shown to provide estimates of the unconditional probability of occurrence of each category, i.e., $\hat{p}_j$ $(j = 1,2,\ldots,J)$, and the vectors of means for the observed replications, i.e., $\hat{\mu}_X^{(j)}$ $(j = 1,2,\ldots,J)$. In order to provide estimates for the elements of the latent parameter matrices, Λ, Φ and Ψ^2, in the structure for Σ, the ANCOVST procedures described by Jöreskog (1970) and Wiley, Schmidt and Bramble (1973) are needed. The approach involved in these procedures was outlined in section D. Since production of the desired estimates of the vectors of weighting coefficients requires the use of components estimated through the application of ANCOVST procedures and since ANCOVST procedures typically require the use of numerical iteration in the calculation of the maximum likelihood estimates, the use of a computer program is a necessity if values of the estimates are to be produced.

Chapter 5 will briefly describe a computer program applying ANCOVST procedures described in chapter 4, which can provide the estimates of a base set of vectors of latent weighting coefficients in the form (4.19). The base set of latent weighting coefficients will be the set associated with category $Y = 1$. Also, included in chapter 5 will be an illustration of the application of the program.

CHAPTER 5

Section A: Introduction

The conclusion from chapter 3 indicated that when errors of measurement are present in the predictors, the observed weighting coefficient will not provide an exact estimate of the latent weighting coefficient for a great variety of situations. The relationship of interest for the quantal response analysis being considered in this research is given by the latent weighting coefficient, but the only observable data typically available for use as predictors contains errors of measurement. Therefore chapter 4 described a reformulated model (4.2) and its associated maximum likelihood estimation procedures which would allow for the estimation of the latent weighting coefficients based on observed data. However these maximum likelihood estimation procedures for elements of the model for Σ belong to a set of procedures (ANCOVST = Analysis of Covariance Structures) which typically require the use of a computer program to provide the numerical iteration procedures needed to produce the estimates of the elements of the model.

The purpose of this chapter is to describe, briefly, a computer program which can provide estimates of the elements of the structural model for Σ and using these estimates then provide estimates for the latent weighting coefficients.

In addition to the program description two examples will be presented to illustrate various estimates of the latent weighting coefficients, in particular, the maximum likelihood estimates produced by the computer program.

Section B: The Computer Program (TQUANER)

The major task in producing estimates of the latent weighting coefficients from an identified model is to produce estimates of the covariance matrix of the latent predictors, Φ, and of the scaling parameters, Λ. The production of estimates for elements of Φ and Λ involves the structural analysis of the model for Σ, that is ANCOVST. As there are several computer programs already available which allow for a structural analysis of a covariance matrix there was little utility in developing a totally new program. Thus the approach here was to produce a program (TQUANER) for estimating latent weighting coefficients by extensively modifying, specializing and extending one of the existing programs.

The program modified to produce TQUANER is ACOVSM: A General Computer Program for Analysis of Covariance Structures Including Generalized Manova by Karl G. Jöreskog, Marielle van Thillo and Gunnar T. Gruvaeus from the Education Testing Service, Princeton, New Jersey. The version of ACOVSM which was modified was the CDC 6500 conversion by Judy Pfaff dated January 1975. Both this version of ACOVSM and TQUANER are Fortran IV programs suitable for use on the CDC 6500 computer at Michigan State University.

ACOVSM provides estimation procedures associated with a general model for Σ described by Jöreskog (1970). It is a more general program with a more complex model for the covariance matrix, Σ, than is needed for the quantal response estimation task. And, of course, ACOVSM does not contain the appropriate algebraic manipulations needed to produce the estimated latent weighting coefficients from the appropriate estimated parameters from the structure for Σ.

Thus the task of producing TQUANER from ACOVSM was two-fold: first, to extensively modify ACOVSM to reduce the complexity of the model being considered by eliminating or bypassing unneeded components and to delete entirely the considerable portion of the program dealing with the generalized MANOVA; and second, to add the programming needed for the input of the sample category means for each observed replication, for the estimation of the latent predictor means and for the production of the estimated weighting coefficients from the appropriate estimates of elements of the models for Σ and $\underset{\sim}{\mu}_{x}^{(i)}$ $(i = 1,2,\ldots,J)$. These adjustments not only specialize the program for quantal response analysis but also realize an economy in both operational cost and amount of space occupied by the program in the computer.

Since TQUANER is based on ACOVSM, many of the input and output characteristics of ACOVSM were carried over to TQUANER. The only changes made were to facilitate special requirements related to quantal response analysis. Therefore a user who is familiar with ACOVSM should have little difficulty using TQUANER.

The information that TQUANER will accept as input consists of descriptions of the component matrices (Λ, Φ and Ψ^2) of the model for Σ, the vectors of sample means for each observed replication in each category of the criterion and the sample covariance matrix. TQUANER provides four options for the input of the sample covariance matrix. The sample covariance matrix for each category can be entered separately or a pooled sample covariance matrix can be entered once.

For either of these choices the individual matrix can be entered in rectangular form by rows or in packed form (i.e., as a vector consisting only of elements, taken by rows, in the lower triangular portion, including the diagonal).

TQUANER will produce and print out the estimates of the parameters in the model for Σ, the estimated vector of latent predictor means for each category and the base set of vectors of weighting coefficients associated with the category identified as the first category by the user.

Various options, most of them also common to ACOVSM, allow the user to request additional printed or punched output. Among these options are included the technical output which describes the behavior of the iterative procedure in the covariance structures analysis, the matrix of residuals for $\hat{\Sigma}$ and an option which allows punch card output of the final solution for the estimates of the elements of the parameter matrices in the model for Σ, i.e., $\hat{\Lambda}$, $\hat{\Phi}$, and $\hat{\Psi}$.

Section C: Two Examples

The purpose of this section is to illustrate the use of the computer program, TQUANER, in the production of estimates of the parameters in the model for Σ and of the latent weighting coefficients. Estimates of the latent weighting coefficients from two other sources will also be identified and derived.

It is important to note at the outset the limitations of the interpretations which can be drawn from this presentation. Only two specific and similar situations were selected. For each of these two situations two random samples were generated for each category. One sample included fifty (50) subjects per category and the second sample included three hundred (300) subjects per category. Thus the range of situations and samples is far too narrow to generalize the results presented below beyond the situations involved in the examples. These examples are provided solely as illustrations of the use of TQUANER and two other procedures for producing estimates for the latent parameters in the model for Σ and for the latent weighting coefficients. It is well beyond the scope of this research to provide a definitive study of the properties of these estimates across even a representative sample of situations. The results from these examples may, however, suggest directions for further study.

Before presenting the two examples a general description of the procedure used to develop each of the examples will be presented. Each example is a special case of the simplest multiple predictor quantal response model, i.e., the criterion has two categories and there are two latent predictors each with two observed replications for a total of four observed variables. The latent predictors are denoted as T^1 and T^2 where X_1^1 and X_2^1 are the observed replications associated with T^1 and, X_1^2 and X_2^2 are associated with T^2. This is a special case of model (4.7) from chapter 4 and thus is identified.

The situation for each example was selected first. This involved selecting values for $b^*_{T^1} = b^*_\xi$ (defined by (3.6b)), $d = d_\xi$ (3.6b), the value of correlation between latent predictors, $\rho_{\xi\eta}$, and the reliability coefficients for each predictor. The reliability coefficients are noted as ρ_{xx} and ρ_{yy} where the selected values are used solely to identify the ratio between the true variance and observed variance for the first observed replication associated with each latent predictor, i.e., X_1^1 and X_1^2. The ratios of true variance to observed variance for the second observed replications, X_2^1 and X_2^2 are selected to be nearly the same, but not identical, to the ratios for the corresponding first observed replication. It is also necessary to select values for the latent predictor variances, the vector of latent predictor means for one of the categories and the scaling factors (elements of Λ). All other parameter values can then be calculated to provide values of the population latent parameters for each element in the models for Σ and $\mu_x^{(i)}$ $(i = 1,2)$ as well as population values for the parameters of Σ and $\mu_x^{(i)}$ $(i = 1,2)$.

Using the population values of the parameters of Σ and $\mu_x^{(i)}$ $(i = 1,2)$ two random samples of size 50 and 300 were generated for each category from a multivariate normal distribution with mean vector $\mu_x^{(i)}$ $(i = 1,2)$ and covariance matrix Σ. A data generation program, GENDATA, developed and checked by Verda Scheifley for use with her doctoral research was used to generate the samples. The sample vectors of category means $\bar{\underset{\sim}{X}}^{(i)}$ (i = 1,2) and covariance matrices S_i (i = 1,2) were then entered into TQUANER as data. TQUANER produced estimates for the elements of Λ, Φ and Ψ (where $\Psi \cdot \Psi = \Psi^2$) as well as estimates for the latent weighting coefficients.

The population values of the elements of Λ, Φ and Ψ^2 will be displayed along with the values of two estimates for each element. The values of the maximum likelihood estimates from TQUANER for each sample will be displayed as well as the values of the heuristic estimates. The values of the heuristic estimates are derived directly from the expressions for the latent parameters as a function of observed parameters produced to show that the model (4.7) was identified. To produce the estimate of the latent parameter the observed parameters in the function will be estimated by their sampled counterparts.

For instance, the expression for $\sigma^2_{T^1}$ as a function of observed parameters is:

$$\sigma^2_{T^1} = \frac{\sigma_{X_2^1 X_1^1}\,\sigma_{X_1^2 X_1^1}}{\sigma_{X_1^2 X_2^1}} \,.$$

Therefore, the heuristic estimate of $\sigma^2_{T^1}$, $\hat{\sigma}^2_{T^1}$, is:

$$\hat{\sigma}^2_{T^1} = \frac{\hat{\sigma}_{X_2^1 X_1^1}\,\hat{\sigma}_{X_1^2 X_1^1}}{\hat{\sigma}_{X_1^2 X_2^1}} = \frac{S_{X_2^1 X_1^1}\,S_{X_1^2 X_1^1}}{S_{X_1^2 X_2^1}}$$

where $S_{X_i^j X_k^m}$ is some element of the sample covariance matrix S_p which corresponds to the element $\sigma_{X_i^j X_k^m}$ of the population covariance matrix Σ.

These heuristic estimates will, in general, not be the values which maximize the likelihood function (or minimize F (4.15)), that is, the heuristic estimates will typically not be equivalent to the maximum likelihood estimates. If the model is just identified, i.e., there are an equal number of latent parameters in the model to be estimated as there are observed parameters, then the maximum likelihood estimates will be equivalent to the heuristic estimates. But when the model is over-identified, i.e., more latent parameters than observed parameters, the maximum likelihood estimates will not equal the heuristic estimates. In this case, the maximum likelihood estimates will generally be more accurate since they incorporate all the observed data simultaneously where the heuristic estimates do not.

The advantage of the heuristic estimates is their relative ease of computation. The basic algebraic manipulation to express each latent parameter as a function of observed parameters should be done as part of the determination of identifiability and thus should be available for use in producing the heuristic estimates. Therefore computation of heuristic estimates can be done without the use of a computer program.

A question which needs to be pursued is how well the heuristic estimates approximate the parameters they estimate over the total range of situations which typically occur in quantal response analysis. In fact, the more important question may be to determine whether there are recognizable situations where the heuristic estimates will perform nearly as well as the maximum likelihood estimates and thus because of their relative ease of calculation be a reasonable alternative to the maximum likelihood procedure. Determining the responses to these questions involves work beyond the scope of this research and thus these questions will not be pursued further here.

For the weighting coefficients, the population latent weighting coefficients for each predictor will be displayed along with the population observed weighting coefficient based on the first observed replication of each predictor only.

From each sample the values of three estimates of each latent weighting coefficient will also be displayed. The estimates of the latent weighting coefficients are produced (1) from the maximum likelihood estimates from the computer program, TQUANER, (2) from the heuristic estimates of the latent parameters described above and (3) from the estimated observed weighting coefficients based on the first observed replication of each predictor as derived by McSweeney and Schmidt (1974).

As argued above the estimates of the latent weighting coefficients based on the maximum likelihood estimates of the latent parameters can be expected to be the most accurate estimates over the population of all possible samples. Since the estimates of the latent weighting coefficients based on the heuristic estimates of the latent parameters do involve attempts to include errors of measurement it is reasonable to expect that these estimates would be generally more accurate than estimates using the observed weighting coefficients to estimate the latent weighting coefficients especially in situations where the observed predictor variance can be expected to contain a relatively large proportion of error variance (i.e., when one or more predictors have a relatively low reliability).

For each of the two examples the situation (i.e., values of $b^*_{T^1}$, d, ρ_{xx}, ρ_{yy}, $\rho_{\xi\eta}$, $\sigma^2_{T^1}$, $\sigma^2_{T^2}$, $\mu^{(1)}_{T^1}$, $\mu^{(1)}_{T^2}$, λ^1_2 and λ^2_2) will be stated. The population parameter values for Λ, Φ, Ψ^2, Σ, $\underset{\sim}{\mu}^{(i)}_T$ and $\underset{\sim}{\mu}^{(i)}_x$ (i = 1,2) will also be given. The values of the vectors of sample means, $\bar{\underset{\sim}{X}}^{(i)}$ (i = 1,2), for each category and the pooled sample covariance matrix, S_p, will be given for each sample, as generated. And finally, the various population parameter values and estimates of the latent parameters and latent weighting coefficients will be tabled.

Example 1

For this example the situation is:

$$b^*_{T^1} = 1, \quad d = 2, \quad \rho_{xx} = .8, \quad \rho_{yy} = .7, \quad \rho_{\xi\eta} = +.3$$

$$\sigma^2_{T^1} = 16, \ \sigma^2_{T^2} = 9, \ \mu^{(i)}_{T^1} = 82, \ \mu^{(i)}_{T^2} = 54, \ \lambda^1_2 = .7 \ \text{ and } \ \lambda^2_2 = 1.4.$$

Using these values the remaining population parameter values can be calculated. Thus the values of the population latent parameters are:

$$\underset{4\times 2}{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0.7 & 0 \\ 0 & 1 \\ 0 & 1.4 \end{bmatrix}, \qquad \underset{2\times 2}{\Phi} = \begin{bmatrix} 16 & 3.6 \\ 3.6 & 9 \end{bmatrix},$$

$$\underset{4\times 4}{\Psi^2} = \text{diag}\{4.0 \quad 2.16 \quad 3.857 \quad 5.36\},$$

$$\underset{2\times 1}{\mu_T^{(1)}} = \begin{bmatrix} 82 \\ 54 \end{bmatrix} \quad \text{and} \quad \underset{2\times 1}{\mu_T^{(2)}} = \begin{bmatrix} 97 \\ 78 \end{bmatrix}.$$

Under the models $\underset{\sim}{\mu}_x^{(i)} = \Lambda\, \underset{\sim}{\mu}_T^{(i)}$ $(i = 1,2)$ and $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$ from (2.27a) and (2.27b), the values of the population observed parameters are:

$$\Sigma = \begin{bmatrix} 20.0 & & & \\ 11.2 & 10.0 & & \text{symmetric} \\ 3.6 & 2.52 & 12.857 & \\ 5.04 & 3.528 & 12.60 & 23.0 \end{bmatrix},$$

$$\underset{\sim}{\mu}_x^{(1)} = \begin{bmatrix} 82.0 \\ 57.4 \\ 54.0 \\ 75.6 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{\mu}_x^{(2)} = \begin{bmatrix} 97.0 \\ 67.9 \\ 78.0 \\ 109.2 \end{bmatrix}.$$

From each category of this population one random sample of size 50 and one random sample of size 300 were generated with the following pooled covariance matrix and vectors of observed means:

$N = 50$/category

$$S_p = \begin{bmatrix} 19.15 & & & \\ 9.715 & 8.495 & & \text{symmetric} \\ 3.435 & 1.69 & 10.44 & \\ 6.105 & 3.355 & 9.77 & 20.905 \end{bmatrix}$$

$$\underset{\sim}{\bar{x}}^{(1)} = \begin{bmatrix} 82.0 \\ 57.45 \\ 53.71 \\ 74.51 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{\bar{x}}^{(2)} = \begin{bmatrix} 97.33 \\ 68.07 \\ 77.60 \\ 108.85 \end{bmatrix} .$$

N = 300/category

$$S_p = \begin{bmatrix} 20.635 & & \text{symmetric} & \\ 11.485 & 10.225 & & \\ 4.47 & 3.29 & 12.97 & \\ 5.305 & 4.06 & 12.325 & 22.98 \end{bmatrix}$$

$$\underset{\sim}{\bar{x}}^{(1)} = \begin{bmatrix} 81.64 \\ 57.56 \\ 54.16 \\ 75.68 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{\bar{x}}^{(2)} = \begin{bmatrix} 97.58 \\ 68.10 \\ 77.97 \\ 108.89 \end{bmatrix} .$$

Table 5.1 presents the population parameter values and for each sample, the heuristic and the maximum likelihood (from TQUANER) estimates.

* * * * * * * * * * * *

Insert Table 5.1 here

* * * * * * * * * * * *

Table 5.2 presents, for each latent predictor, the population latent weighting coefficient value and the population observed weighting coefficient value based solely on the first observed replication of each predictor. These results for the observed weighting coefficient are those that would be derived if the classical measurement model

Table 5.1. Estimates of the Latent Parameters of Λ, Φ, Ψ^2 and $\mu_T^{(i)}$ (i = 1,2) for Example 1.

Parameter Type	Element Subscript	Population Parameter Value	Sample 1 (N = 50) Heuristic Estimate Value	Sample 1 (N = 50) Maximum Likelihood Estimate Value	Sample 2 (N = 300) Heuristic Estimate Value	Sample 2 (N = 300) Maximum Likelihood Estimate Value
Λ	21	.7	.49	.54	.74	.75
	42	1.4	1.78	1.80	1.19	1.22
Φ	11	16.0	19.75	18.01	15.60	15.39
	21	3.6	3.44	3.41	4.47	4.43
	22	9.0	5.50	5.44	10.39	10.13
Ψ^2	11	4.0	0.00*	1.14	5.03	5.24
	22	2.16	3.72	3.25	1.77	1.66
	33	3.857	4.94	5.00	2.59	2.84
	44	5.36	3.55	3.37	8.35	7.99
$\mu_T^{(1)}$	1	82.0	88.78	87.52	80.43	80.04
	2	54.0	44.76	44.40	59.77	58.96
$\mu_T^{(2)}$	1	97.0	105.32	103.83	95.80	95.33
	2	78.0	65.19	64.66	86.02	84.85

*Estimate was negative, so it was set to zero.

(2.22) were applied to relate each latent predictor to the first observed replication associated with that latent predictor. Under that interpretation the situation for this example falls into joint general category four (G.C. IV (x,y)) from chapter 3. The information from the generic graphs and tables in chapter 3 for G.C. IV (x,y) suggest that the observed weighting coefficient for the observed predictor associated with T^1 (observed predictor x in chapter 3 notation) may be a reasonably good estimator of the latent weighting coefficient for T^1 for moderately sized positive values of the correlation between predictors, $\rho_{\xi\eta}$, (since $\rho_{\xi\eta} < 1/d$ but $\rho_{\xi\eta}$ is not "too close" to $1/d$ here). The same information suggests that the observed weighting coefficient for the observed predictor associated with T^2 (observed predictor y in chapter 3 notation) will be an underestimate of the latent weighting coefficient for T^2.

Table 5.2 also contains, for each sample, the three estimates of the latent weighting coefficient described above. To help judge the accuracy of each estimate for the given sample the value in the parenthesis below each estimate is the ratio of the estimate over the population latent weighting coefficient value. As in chapter 3 ratios greater than one represent overestimates, ratios less than one represent underestimates and ratios equal to one represent exact estimates.

* * * * * * * * * * * * *

Insert Table 5.2 here

* * * * * * * * * * * * *

Table 5.2. Parameter Values and Estimates of Weighting Coefficients for Example 1.

Latent Predictor	Population Latent Weighting Coefficient Value	Population Observed Weighting Coefficient Value[a]	Estimated Observed Coefficient Value	Sample 1 (N = 50) Heuristic Estimate of Latent Coefficient Value	Sample 1 (N = 50) Maximum Likelihood Estimate of Latent Coefficient Value	Estimated Observed Coefficient Value	Sample 2 (N = 300) Heuristic Estimated Latent Coefficient Value	Sample 2 (N = 300) Maximum Likelihood Latent Coefficient Value
T^1	.371	.436 (1.175)[b]	.415 (1.118)[c]	.215 (.580)	.229 (.617)	.405 (1.09)	.298 (.803)	.295 (.795)
T^2	2.518	1.745 (.693)	2.152 (.855)	3.581 (1.422)	3.579 (1.421)	1.696 (.674)	2.399 (.953)	2.427 (.964)

[a]Based solely on the first observed replication associated with each latent predictor.

[b]ratio in parenthesis = $\frac{\text{population observed value}}{\text{population latent value}}$.

[c]ratio in parenthesis = $\frac{\text{estimated value}}{\text{population latent value}}$.

For this example and the given sample of 50 subjects per category the estimated observed weighting coefficient based solely on the first observed replication provides the most accurate estimate of both population latent weighting coefficients, with considerably less accuracy shown by both the heuristic and maximum likelihood estimates of the latent weighting coefficients.

For the sample of 300 per category, the estimated observed coefficients only slightly overestimate the corresponding latent coefficient for T^1 but provide a rather substantial underestimate of the corresponding latent coefficient for T^2. For each latent coefficient, the heuristic estimate is nearly identical to the maximum likelihood estimate. For T^1 they both provide slightly less accurate estimates than the estimated observed coefficient. However for T^2 they both provide considerably more accurate estimates than the estimated observed coefficient. Considering both latent coefficients the maximum likelihood and heuristic estimates seem to perform equally well and better than the estimated observed coefficient.

One reason for the relatively strong performance of the estimated observed coefficient as an estimate of the latent coefficient here is due to the relationship between the population observed coefficient and the population latent weighting coefficient. Derivation 1 from chapter 3, section C, subsection VI, indicates that when the population observed coefficient is equal to the population latent coefficient for one predictor the population observed coefficient for the other predictor will underestimate the population latent coefficient for that predictor by a factor equal to the reliability of the observed predictor. In this case, the population observed coefficient is nearly identical to the population latent coefficient for T^1 and the estimated observed coefficient is also nearly identical to the population latent coefficient. Since the reliability of the second observed predictor X_1^2 (associated with T^2) is not particularly low ($\rho_{yy} = .7$) the estimated observed coefficient would not be expected to be too inaccurate in this case. Example 2 involves a similar application of Derivation 1, but since the predictor reliabilities are much lower than in this example, the estimated observed coefficient proves to be a poor estimator of the latent coefficient for predictor T^2.

Example 2

Example 2 involves a situation somewhat similar to the situation examined in example 1 but with lower predictor reliabilities. Thus since a greater proportion of observed variance consists of error variance for this situation than for the situation in example 1 it does not seem likely that the estimates of observed weighting coefficients should provide as accurate estimates of both latent weighting coefficients as in example 1.

For this example the situation is:

$$b^*_{T^1} = 4, \quad d = 2, \quad \rho_{xx} = .4, \quad \rho_{yy} = .5, \quad \rho_{\xi\eta} = +.4$$

$$\sigma^2_{T^1} = 100, \quad \sigma^2_{T^2} = 64, \quad \mu^{(1)}_{T^1} = 80, \quad \mu^{(1)}_{T^2} = 60, \quad \lambda^1_2 = .8 \quad \text{and} \quad \lambda^2_2 = 1.3.$$

Using these values the remaining population parameter values can be calculated. Thus the values of the population latent parameters are:

$$\underset{4\times 2}{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0.8 & 0 \\ 0 & 1 \\ 0 & 1.3 \end{bmatrix}, \qquad \Phi = \begin{bmatrix} 100 & 32 \\ 32 & 64 \end{bmatrix},$$

$$\underset{4\times 4}{\Psi^2} = \text{diag}\{150.0 \quad 101.0 \quad 64.0 \quad 111.84\},$$

$$\underset{2\times 1}{\mu_T^{(1)}} = \begin{bmatrix} 80 \\ 60 \end{bmatrix} \quad \text{and} \quad \underset{2\times 1}{\mu_T^{(2)}} = \begin{bmatrix} 480 \\ 700 \end{bmatrix}.$$

Under the models $\mu_x^{(i)} = \Lambda\, \mu_T^{(i)}$ $(i = 1,2)$ and $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$ (from (2.27a) and (2.27b)), the values of the population observed parameters are:

$$\Sigma = \begin{bmatrix} 250.0 & & \text{symmetric} & \\ 80.0 & 165.0 & & \\ 32.0 & 25.6 & 128.0 & \\ 46.8 & 37.4 & 83.2 & 220.0 \end{bmatrix},$$

$$\mu_x^{(i)} = \begin{bmatrix} 80 \\ 64 \\ 60 \\ 78 \end{bmatrix} \quad \text{and} \quad \mu_x^{(2)} = \begin{bmatrix} 480 \\ 384 \\ 700 \\ 910 \end{bmatrix}.$$

From each category of this population one random sample of size 50 and one random sample of size 300 were generated with the following pooled covariance matrix and vectors of observed means:

N = 50/category

$$S_p = \begin{bmatrix} 267.455 & & & \text{symmetric} \\ 92.495 & 149.47 & & \\ 49.61 & 50.25 & 153.425 & \\ 59.805 & 57.98 & 96.945 & 210.075 \end{bmatrix}$$

$$\bar{\underset{\sim}{X}}^{(1)} = \begin{bmatrix} 81.43 \\ 64.50 \\ 60.42 \\ 78.13 \end{bmatrix} \quad \text{and} \quad \bar{\underset{\sim}{X}}^{(2)} = \begin{bmatrix} 479.72 \\ 382.48 \\ 700.63 \\ 908.95 \end{bmatrix} .$$

N = 300/category

$$S_p = \begin{bmatrix} 257.595 & & & \text{symmetric} \\ 77.275 & 153.64 & & \\ 31.27 & 36.785 & 133.995 & \\ 48.32 & 51.65 & 91.455 & 231.585 \end{bmatrix} ,$$

$$\bar{\underset{\sim}{X}}^{(1)} = \begin{bmatrix} 79.05 \\ 63.08 \\ 59.40 \\ 77.99 \end{bmatrix} \quad \text{and} \quad \bar{\underset{\sim}{X}}^{(2)} = \begin{bmatrix} 480.32 \\ 385.18 \\ 700.36 \\ 911.64 \end{bmatrix} .$$

Table 5.3 presents the population parameter values and, for each sample, the heuristic and the maximum likelihood (from TQUANER) estimates.

* * * * * * * * * * * *

Insert Table 5.3 here

* * * * * * * * * * * *

Table 5.3. Estimates of the Latent Parameters of Λ, Φ, Ψ^2 and $\mu_T^{(i)}$ (i = 1,2) for Example 2.

Parameter Type	Element Subscript	Population Parameter Value	Sample 1 (N = 50) Heuristic Estimate Value	Sample 1 (N = 50) Maximum Likelihood Estimate Value	Sample 2 (N = 300) Heuristic Estimate Value	Sample 2 (N = 300) Maximum Likelihood Estimate Value
Λ	21	.8	1.01	.98	1.18	1.11
	42	1.3	1.21	1.17	1.55	1.44
Φ	11	100.0	91.32	94.57	65.69	69.55
	21	32.0	49.61	50.92	31.27	32.69
	22	64.0	80.42	83.26	59.19	63.75
Ψ^2	11	150.0	176.14	173.50	191.91	188.05
	22	101.0	55.76	58.28	62.79	67.78
	33	64.0	73.01	70.28	74.81	70.24
	44	111.84	93.11	97.04	90.31	100.38
$\mu_T^{(1)}$	1	80.0	72.44	73.67	64.30	66.75
	2	60.0	63.00	64.23	53.11	56.01
$\mu_T^{(2)}$	1	480.0	427.98	435.24	391.65	406.49
	2	700.0	732.08	746.22	622.62	656.68

Table 5.4 presents information for example 2 which is comparable to the information presented in Table 5.2 for example 1.

* * * * * * * * * * * *

Insert Table 5.4 here

* * * * * * * * * * * *

It is interesting to note that the population observed weighting coefficient for the single observed replication associated with latent predictor T^1 is nearly identical to the population latent weighting coefficient for T^1 and that the estimated observed weighting coefficient is the most accurate estimate of the latent weighting coefficient for T^1 across both samples. But also note that the population weighting coefficient for the single observed replication associated with latent predictor T^2 is about half the size of the latent weighting coefficient for T^2. The estimated observed weighting coefficients are also less than half of the latent weighting coefficients across both samples. This illustrates result (3.52) from chapter 3, that if $\beta_x/\beta^*_\xi = 1$ then $\beta_y/\beta^*_\eta = \rho_{yy}$. Here $\beta_x/\beta^*_\xi \approx 1$ thus $\beta_y/\beta^*_\eta \approx \rho_{yy} = .5$.

For the sample of size 50/category, both the heuristic and maximum likelihood estimates associated with T^1 have the wrong sign. The heuristic estimate for T^2 is the most accurate but only slightly more accurate than the maximum likelihood estimate.

For the sample of size 300/category, the maximum likelihood estimate for T^1 is slightly more accurate than the heuristic estimate but for T^2 the maximum likelihood estimate is slightly less accurate than the heuristic estimate.

Table 5.4. Parameter Values and Estimates of Weighting Coefficients for Example 2.

				Sample 1 (N = 50)			Sample 2 (N = 300)	
Latent Predictor	Population Latent Weighting Coefficient Value	Population Observed Weighting Coefficient Value[a]	Estimated Observed Coefficient Value	Heuristic Estimate of Latent Coefficient Value	Maximum Likelihood Estimate of Latent Coefficient Value	Estimated Observed Coefficient Value	Heuristic Estimated Latent Coefficient Value	Maximum Likelihood Latent Coefficient Value
T^1	.953	.992 (1.04)[b]	.761 (.798)[c]	-.942 (-.988)	-.875 (-.918)	1.005 (1.055)	.538 (.565)	.601 (.631)
T^2	9.524	4.752 (.499)	3.927 (.412)	8.901 (.935)	8.727 (.916)	4.549 (.478)	9.338 (.980)	9.115 (.957)

[a] Based solely on the first observed replication associated with each latent predictor.

[b] ratio in parenthesis = $\frac{\text{population observed value}}{\text{population latent value}}$.

[c] ratio in parenthesis = $\frac{\text{estimated value}}{\text{population latent value}}$.

The result which is reinforced by this example is that the use of estimated observed weighting coefficients as estimates of the latent weighting coefficients in a two predictor model can provide precise estimates for at most one latent weighting coefficient. The other latent weighting coefficient will be underestimated by the estimated observed weighting coefficient by a factor approximately equal to the reliability of the observed predictor, which in this example is rather low ($\rho_{yy} = .5$). In the first example, where the reliability of the second predictor was higher, ($\rho_{yy} = .7$), the observed weighting coefficient provided a slightly more accurate estimate across both samples.

Section D: Summary

This chapter has presented a brief description of a computer program, TQUANER, which can provide estimates of the latent parameters in identified models for Σ and $\mu_x^{(i)}$ $(i = 1,2,\ldots,J)$ and estimates of the vectors of latent weighting coefficients using the procedures described in chapter 4.

Two similar examples were presented in section C of this chapter to illustrate the use of the computer program, TQUANER, to produce estimates from actual data examples. The two examples are insufficient to allow generalization of results and conclusions across a broad category of situations to which quantal response analysis may be applied.

Further research is needed to determine and describe the distribution of the maximum likelihood estimates of the latent weighting coefficients for samples of various sizes. Part of this research could include an examination of the relative accuracy and utility of the heuristic estimates since they are relatively easy to calculate, not requiring the use of a computer. Another part of this examination might also focus on the relative accuracy and utility of the estimated observed weighting coefficient, for the most reliable observed replication of each predictor, as an estimate of the latent weighting coefficient. This examination might be restricted to those situations where each of the predictors has a reasonably high reliability since it is clear that problems can exist when even one predictor has a low reliability. But this restriction may not be a major drawback since many situations where quantal response analysis is to be employed may use predictors with reasonably high reliabilities.

CHAPTER 6

Section A: Summary and Conclusions

In chapter 1, a general discussion of research on various quantitative and qualitative data analysis models indicated that the presence of errors of measurement in the variables under analysis can cause problems in interpretation of data analysis results. The purpose of the research reported here was then identified as expanding this previous research to include investigation of the effects of errors of measurement on a quantal response analysis technique. The particular quantal response technique to be examined involves a qualitative criterion with two or more categories and one or more quantitative predictor variables, where the predictor variables are assumed to be random variables possessing a normal distribution (a multivariate normal distribution for two or more predictors). For this particular quantal response technique, the interest is focused on the weighting coefficients associated with each predictor. The weighting coefficient associated with each predictor provides a measure of the relationship between that predictor and the probability of classification into one category of the criterion versus classification into some other category of the criterion. Classification into categories of the criterion is assumed to be without error. However, the predictors are assumed to be measured with error, where the presence of errors of measurement in any predictor is indicated by a reliability coefficient less than one for that predictor.

The weighting coefficients of interest for this approach to quantal response analysis are the weighting coefficients derived from the use of the latent (error-free) predictors and given by the Latent Random Predictor Quantal Response Model (2.19) of chapter 2. However, the variables which are available for use as predictors typically contain errors of measurement. Therefore direct application of a model (2.19) based solely on error-free predictors is not generally appropriate. The Observed Random Predictor Quantal Response Model (2.2) based on predictors which contain errors of measurement, called observed predictors, is also presented in chapter 2. It is this model which is generally appropriate for use. To relate the model (2.2) based on observed predictors (i.e., with errors of measurement) to the model (2.19) based on latent predictors (i.e., with no errors of measurement), two measurement models, (2.2) and (2.26), are presented in chapter 2. Properties of each of the measurement models which are useful in relating the two quantal response models are also presented.

In chapter 3, the classical measurement model (2.2) is used to relate the two quantal response models (2.2) and (2.19). The classical measurement model associates one observed predictor with each latent predictor. Thus the weighting coefficient associated with each latent predictor called the latent weighting coefficient for a given category of the criterion has related to it a single weighting coefficient associated with the corresponding observed predictor (called the observed weighting coefficient). The relationship between corresponding observed and latent weighting coefficients was examined in chapter 3.

Since no generally applicable results of the effects of errors of measurement were found for the most general case of the quantal response model, i.e., for a polychotomous criterion and multiple predictors, the research in chapter 3 focused on one and two predictor models with a polychotomous criterion.

For one predictor model, the observed weighting coefficient was shown (3.3) to be an underestimate of the latent weighting coefficient by a factor equal to the reliability of the single predictor. Thus, the use of a single predictor with low reliability in a quantal response analysis can lead to misinterpretations if the observed weighting coefficient (from (2.2)) is used as an estimate of the latent weighting coefficient (from (2.19)).

For two predictor models, no universally applicable result as found for one predictor models was discovered. Since no universally applicable results was found for two predictor models, the approach used in chapter 3 involved the search for general categories of situations where the relationship between the observed and latent weighting coefficients would have a similar pattern for all situations within a general category. As used in chapter 3 a "situation" is completely defined when values sufficient to specify the relationship between the observed and latent weighting coefficient are given. For the approach used in chapter 3 then a situation is completely defined for values of the relationship between standardized category mean differences on each of the predictors, and for values of the reliabilities of each predictor. For given situations, values of the relationship between the observed and latent weighting coefficients for each predictor are considered as functions of the correlation between predictors.

In chapter 3 four general categories of situations were described. In addition to the four general categories three special case sets of situations were also described. Although there are some similarities among results for the various categories only a few results are found which apply across all four general categories and the three special case situations.

First, when there is no correlation between the two latent predictors then each of the observed weighting coefficients is an underestimate of the corresponding latent weighting coefficient by a factor equal to the reliability of the given observed predictor. This result applies in all cases where the latent weighting coefficient has a non-zero value. This result is comparable to the result for one predictor models, which is not surprising since if there is no relationship between the two latent predictors the value of one latent predictor cannot be expected to influence the value of the latent weighting coefficient of the other latent predictor. The lack of correlation between latent predictors then results in the value of the latent weighting coefficient for each predictor being produced as if that predictor were the only predictor in the model, i.e., almost as if each predictor was included in a one predictor model. See Appendix B.2, section C, for details.

Second, when the ratio of the slope of the pooled within categories regression line of one predictor on the second predictor over the slope of the between categories line joining the midpoints of the joint distributions of predictors within each category is negative, then both of the observed weighting coefficients are underestimates of the latent weighting coefficients. See Appendix B.9 for further details on this interpretation and Derivation 3 under subheading VI of section C in chapter 3 for the proof and an illustrative example.

Third, it is not possible for both of the observed weighting coefficients to be overestimates of the corresponding latent weighting coefficients for the same situation. That is, at most one observed weighting coefficient, in a two predictor model, can overestimate the corresponding latent weighting coefficient. In addition, Derivation 1 under subheading VI of section C in chapter 3 proves that if the observed weighting coefficient is equal in value to the corresponding latent weighting coefficient for some one predictor then for the other predictor the observed weighting coefficient will be an underestimate of the latent weighting coefficient by a factor equal to the reliability of that observed predictor.

The only nearly universal conclusion for the two-predictor model is that for only a few special case situations will the value of either observed predictor precisely equal the value of the corresponding latent predictor.

Since the observed weighting coefficient typically does not precisely estimate the corresponding latent weighting coefficient (at least for one and two predictors models), there is some utility in describing procedures for estimating the latent weighting coefficients from observed data.

Chapter 4 presents two reformulations of the Observed Random Predictor Quantal Response Model (2.2) in terms of latent parameters, i.e., parameters from the Latent Random Predictor Quantal Response Model (2.19), parameters describing errors of measurement, parameters indicating a relative scale of measure for the observed predictors, and the vectors of latent predictor means for each category of the criterion. The first reformulation (4.1) is based on the application of the classical measurement model (2.22) while the second reformulation (4.2) is based on the expanded measurement model (2.26) which includes the use of multiple observed replications associated with each latent predictor.

Much of the work in chapter 4 (section C) defines and examines a sufficient condition for the existence of estimates of the latent parameters, primarily from the model for the population covariance matrix from (2.2). The sufficient condition for existence of estimates of the latent parameters is that the elements of a model be identifiable (see expression (4.3) for the definition). If a model is not identified, two approaches were discussed for the modification of the model to produce an identified model, (1) use of replicate observed measurements for each predictor or (2) imposition of constraints upon the latent parameters of the model. By a careful use of one or both approaches a non-identified model can usually be modified into an identified model.

The model for the covariance matrix based on the classical measurement model (2.22) was shown to be non-identified. However, under a variety of conditions involving either imposition of constraints or the use of replicate observations, or both, the models for the covariance matrix and the vectors of observed predictor means for each category, based on the expanded measurement model (2.26), were shown to be identified. Thus for reformulation (4.2) unique estimates of the latent parameters exist.

Maximum likelihood procedures associated with reformulation (4.2) were described in section D of chapter 4. These procedures involved the use of covariance structures analysis (ANCOVST). As with most situations where ANCOVST procedures are employed the estimates of the latent parameters in the model for the covariance matrix cannot be expressed as specific functions of the observed data. Therefore the derivation of the values of the maximum likelihood estimates must be accomplished by application of a numerical iteration process. Since ANCOVST procedures typically require the use of a computer program to perform the necessary numerical iterations, chapter 5 describes a computer program (TQUANER) which was programmed to provide the maximum likelihood estimates for the latent parameters and the latent weighting coefficients. The computer program (TQUANER) is a modification of ACOVSM: A General Computer Program for Analysis of Covariance Structures Including Generalized Manova by Karl G. Joreskog, Marielle van Thillo and Gunnar Gruvaeus.

In addition to describing the computer program (TQUANER), chapter 5 provides two simulated data examples to illustrate the use of the program. However, no general conclusions about the distribution or accuracy of the maximum likelihood estimates of the latent weighting coefficients can be drawn from the limited application of TQUANER to the two simulated data examples.

This research has shown that errors of measurement in the random predictor variables of a quantal response analysis technique can cause problems in using the observed weighting coefficients as estimates of the latent weighting coefficients which represent the relationships of interest between the error-free predictors and the criterion. A quantal response model based on observed data was presented with its associated estimation procedures with provide for the estimation of the latent weighting coefficients from observed data. And a computer program (TQUANER) which can produce estimates of the latent weighting coefficients was described and its use was illustrated on two simulated data examples.

Section B: Recommendations for Further Study

One obvious possibility for further study is to attempt to extend the results of chapter 3 to models with more than two predictors. Such an extension is easy to identify but will be difficult to do. Approaches to this extension which were tried with little success involved a general matrix manipulation approach and an algebraic derivation of each individual weighting coefficient in the base set of weighting coefficients. No generally useful results were discovered from the matrix approach. The algebriac derivation approach proved an extremely tedious way to extend the results since there is no guarantee that the patterns of results for three predictor models will extend to four or more predictor models, although at the present, this seems to be the most promising approach to extending the results beyond two predictor models.

A second possibility for further study also based on chapter 3 results, is to restrict the examination of the effects of errors of measurement to situations which are likely to occur in practical applications of quantal response analysis. The question of interest here is whether, within the set of typically occurring situations for the application of quantal response analysis, the effects of errors of measurement are sufficiently severe to prevent the use of the observed weighting coefficient as an estimate of the latent weighting coefficient.

A third possibility for further study is to examine the value of the heuristic estimates of the latent parameters described in chapter 5. The question here is whether the heuristic estimates are reasonable competitors of the maximum likelihood estimates especially considering that the maximum likelihood estimates require the use of a computer program on a reasonably sophisticated computer system while the heuristic estimates can typically be produced using no more than a simple calculation.

A fourth possibility to further study is to examine the properties of the sampling distribution of the maximum likelihood estimates of the latent parameters and the latent weighting coefficients. This will probably entail a Monte Carlo simulation study.

Four possibilities for further study have been given above. No priority is implied by the given ordering. In fact, some combination of the second and fourth possibilities may be the most fruitful approach for further study.

APPENDICES

APPENDIX A.1

IDENTIFICATION AND JUSTIFICATION OF THE EXISTENCE OF A BASE SET OF WEIGHTING COEFFICIENTS IN THE MULTIPLE OBSERVED PREDICTOR, POLYCHOTOMOUS CRITERION MODEL

Consider any category of the criterion k $(k = 1,2,\ldots,J)$

where

$$P_k = \text{Prob}\{Y = k|\underset{\sim}{X}\} = \frac{1}{1 + \sum_{\substack{j=1 \\ j \neq k}}^{J} \exp\{-(\alpha_{k \cdot j} + \underset{\sim}{\beta}'_{k \cdot j} X)\}}$$

$$\text{where} \quad \alpha_{k \cdot j} = -\ln\left(\frac{P_j}{P_k}\right) - \frac{1}{2}\left[\underset{\sim}{\mu}_X^{(k)'} \Sigma^{-1} \underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)'} \Sigma^{-1} \underset{\sim}{\mu}_X^{(j)}\right]$$

$$\text{and} \quad \underset{\sim}{\beta}_{k \cdot j} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)}).$$

Associated with category k are $J-1$ vectors of weighting coefficients $\underset{\sim}{\beta}_{k \cdot j} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(j)})$ with $j \neq k$, i.e.

$\underset{\sim}{\beta}_{k \cdot 1}, \underset{\sim}{\beta}_{k \cdot 2}, \ldots, \underset{\sim}{\beta}_{k \cdot (k-1)}, \underset{\sim}{\beta}_{k \cdot (k+1)}, \ldots, \underset{\sim}{\beta}_{k \cdot J}$.

Consider any category k' where $k' \neq k$, and consider any vector of weighting coefficients associated with category k', $\underset{\sim}{\beta}_{k' \cdot \ell}$ $(\ell \neq k', \ \ell, k' = 1,2,\ldots,J)$.

Therefore $\underset{\sim}{\beta}_{k' \cdot \ell} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k')} - \underset{\sim}{\mu}_X^{(\ell)})$.

If $\ell \neq k$, $\underset{\sim}{\beta}_{k'\cdot\ell} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k')} - \underset{\sim}{\mu}_X^{(\ell)})$

$$= \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k')} - \underset{\sim}{\mu}_X^{(k)} + \underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(\ell)})$$

$$= \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k')} - \underset{\sim}{\mu}_X^{(k)}) + \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(\ell)})$$

$$= -\Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(k')}) + \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} + \underset{\sim}{\mu}_X^{(\ell)})$$

$$= -\underset{\sim}{\beta}_{k\cdot k'} + \underset{\sim}{\beta}_{k\cdot\ell} .$$

Therefore $\underset{\sim}{\beta}_{k'\cdot\ell} = \underset{\sim}{\beta}_{k\cdot\ell} - \underset{\sim}{\beta}_{k\cdot k'}$ where $\underset{\sim}{\beta}_{k\cdot\ell}$ and $\underset{\sim}{\beta}_{k\cdot k'}$ are vectors of weighting coefficients associated with category k.

If $\ell = k$, $\underset{\sim}{\beta}_{k'\cdot\ell} = \underset{\sim}{\beta}_{k'\cdot k} = \Sigma^{-1}(\underset{\sim}{\mu}_X^{(k')} - \underset{\sim}{\mu}_X^{(k)})$

$$= -\Sigma^{-1}(\underset{\sim}{\mu}_X^{(k)} - \underset{\sim}{\mu}_X^{(k')}) .$$

Therefore $\underset{\sim}{\beta}_{k'\cdot k} = -\underset{\sim}{\beta}_{k\cdot k'}$ where $\underset{\sim}{\beta}_{k\cdot k'}$ is a vector of weighting coefficients associated with category k.

The above proof shows that given the J-1 vectors of observed weighting coefficients for some one category of the criterion k (k = 1,2,...,J), then all vectors of weighting coefficients associated with each of the remaining J-1 categories of the criterion can be expressed as combinations of the vectors of weighting coefficients associated with category k.

APPENDIX A.2

IDENTIFICATION AND JUSTIFICATION OF THE EXISTENCE OF A BASE SET OF WEIGHTING COEFFICIENTS IN THE MULTIPLE LATENT PREDICTOR, POLYCHOTOMOUS CRITERION MODEL.

Consider any category k $(k = 1,2,\ldots,J)$ of the criterion, where

$$P_k^* = \text{Prob}\{Y = k|\underset{\sim}{T}\} = \frac{1}{1 + \sum_{\substack{j=1 \\ j\neq k}}^{J} \exp\{-(\alpha_{k\cdot j}^* + \underset{\sim}{\beta}_{k\cdot j}^{*\prime}\, \underset{\sim}{T})\}}$$

where $\alpha_{k\cdot j}^* = -\ln(\frac{p_j}{p_k}) - \frac{1}{2}[\underset{\sim}{\mu}_T^{(k)\prime}\,\Phi^{-1}\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(j)\prime}\,\Phi^{-1}\underset{\sim}{\mu}_T^{(j)}]$

and $\underset{\sim}{\beta}_{k\cdot j}^* = \Phi^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(j)})$.

Associated with category k are $J-1$ vectors of weighting coefficients $\underset{\sim}{\beta}_{k\cdot j}^* = \Phi^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(j)})$ with $j \neq k$, i.e. $\underset{\sim}{\beta}_{k\cdot 1}^*, \underset{\sim}{\beta}_{k\cdot 2}^*, \ldots, \underset{\sim}{\beta}_{k\cdot (k-1)}^*, \underset{\sim}{\beta}_{k\cdot (k+1)}^*, \ldots, \underset{\sim}{\beta}_{k\cdot J}^*$.

Consider any category k' where $k' \neq k$, and consider any vector of weighting coefficients associated with category k', $\underset{\sim}{\beta}_{k'\cdot \ell}^*$ $(\ell \neq k',\ \ell, k' = 1,2,\ldots,J)$.

Therefore $\underset{\sim}{\beta}_{k'\cdot \ell} = \Phi^{-1}(\underset{\sim}{\mu}_T^{(k')} - \underset{\sim}{\mu}_T^{(\ell)})$.

If $\ell \neq k$,

$$\begin{aligned}
\underset{\sim}{\beta}_{k'\cdot \ell}^* &= \Phi^{-1}(\underset{\sim}{\mu}_T^{(k')} - \underset{\sim}{\mu}_T^{(\ell)}) \\
&= \Phi^{-1}(\underset{\sim}{\mu}_T^{(k')} - \underset{\sim}{\mu}_T^{(k)} + \underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(\ell)}) \\
&= \Phi^{-1}(\underset{\sim}{\mu}_T^{(k')} - \underset{\sim}{\mu}_T^{(k)}) + \Phi^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(\ell)}) \\
&= -\Phi^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(k')}) + \Phi^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(\ell)}) \\
&= -\underset{\sim}{\beta}_{k\cdot k'}^* + \underset{\sim}{\beta}_{k\cdot \ell}^* .
\end{aligned}$$

Therefore $\underset{\sim}{\beta}^{*}_{k\cdot\ell} = \underset{\sim}{\beta}^{*}_{k\cdot\ell} - \underset{\sim}{\beta}^{*}_{k\cdot k'}$ where $\underset{\sim}{\beta}^{*}_{k\cdot\ell}$ and $\underset{\sim}{\beta}^{*}_{k\cdot k'}$ are vectors of weighting coefficients associated with category k.

If $\ell = k$, $\underset{\sim}{\beta}^{*}_{k'\cdot\ell} = \underset{\sim}{\beta}^{*}_{k'\cdot k} = \Phi^{-1}(\underset{\sim}{\mu}_T^{(k')} - \underset{\sim}{\mu}_T^{(k)})$

$$= -\Phi^{-1}(\underset{\sim}{\mu}_T^{(k)} - \underset{\sim}{\mu}_T^{(k')}).$$

Therefore $\underset{\sim}{\beta}^{*}_{k'\cdot k} = -\underset{\sim}{\beta}^{*}_{k\cdot k'}$ where $\underset{\sim}{\beta}^{*}_{k\cdot k'}$ is a vector of weighting coefficients associated with category k.

The above proof shows that given the J-1 vectors of latent weighting coefficients associated with some one category of the criterion k (k = 1,2,...,J), then all vectors of weighting coefficients associated with each of the remaining J-1 categories of the criterion can be expressed as combinations of the vectors of weighting coefficients associated with category k.

APPENDIX B.1

A. Development of the Property of Interchangeability of x and y

Consider expression (3.12) for β_x/β_ξ^* and expression (3.14) for β_y/β_η^*. A close examination of these expressions shows an identical structure for each expression. To demonstrate the relationship between these expressions, consider some given situation, i.e. values of $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_ξ such that $\beta_\xi^* \neq 0$. For this situation $\beta_x/\beta_\xi^* = \dfrac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d_\xi\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})}$ takes on some value call it R. Consider now some second situation $\rho'_{\xi\eta}$, ρ'_{xx}, ρ'_{yy} and d'_ξ where:

(B.1.1) $\rho'_{\xi\eta} = \rho_{\xi\eta}$

(B.1.2) $\rho'_{xx} = \rho_{yy}$

(B.1.3) $\rho'_{yy} = \rho_{xx}$

(B.1.4) $d'_\eta = d_\xi$.

The two situations are not, in general, identical.

Consider now the value of β_y/β_η^* for the second situation.

$$\beta_y/\beta_\eta^* = \frac{(1 - \rho'^2_{\xi\eta})\rho'_{yy}(1 - d'_\eta\rho'_{\xi\eta}\rho'_{xx})}{(1 - \rho'^2_{\xi\eta}\rho'_{xx}\rho'_{yy})(1 - d'_\eta\rho'_{\xi\eta})} \qquad \text{[from (3.14)]}$$

$$= \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d_\xi\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{yy}\rho_{xx})(1 - d_\xi\rho_{\xi\eta})} = R \quad \text{[from (B.1.1)-(B.1.4)]}.$$

But $\beta_x/\beta_\xi^* = R$ for $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_ξ, i.e. the first situation.

Therefore, for any given situation $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_ξ which produces a given value for β_x/β_ξ^* there exists a second, typically different, situation related to the first situation by expressions (B.1.1)-(B.1.4) which produces the same value for β_y/β_η^* as the first situation produces for β_x/β_ξ^*.

A careful consideration of expressions (B.1.1)-(B.1.4) indicates that the right side of each expression can be produced by replacing each x with a y, each y with an x, each ξ with an η and each η with a ξ in the left side of each expression.

In fact if this same procedure of interchanging the x's and y's as well as the ξ's and η's were done on expression (3.12) for β_x/β_ξ^*, the result would be expression (3.14) for β_y/β_η^*.

Thus it is necessary to examine only β_x/β_ξ^*. All results for β_x/β_ξ^* can be translated into results for β_y/β_η^* by the use of this property of interchangeability of x and y discussed above.

B. Justification of the Need to Examine β_x/β_ξ^* Only for Values of $d \geq 0$.

Consider expression (3.12) for β_x/β_ξ^*. Let d_ξ be some negative number ($d_\xi < 0$). Therefore the situation $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_ξ ($d_\xi < 0$) produces some value for

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d_\xi\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d_\xi\rho_{\xi\eta})} \quad \text{call it } R.$$

Consider now some second situation $\rho''_{\xi\eta}$, ρ''_{xx}, ρ''_{yy}, and d''_{ξ} where:

(B.1.5) $\rho''_{\xi\eta} = -\rho_{\xi\eta}$

(B.1.6) $\rho''_{xx} = \rho_{xx}$

(B.1.7) $\rho''_{yy} = \rho_{yy}$

(B.1.8) $d''_{\xi} = -\, d_{\xi}$ (i.e. $d''_{\xi} > 0$, since $d_{\xi} < 0$).

The value of β_x/β^*_{ξ} for this second situation is:

$$\beta_x/\beta^*_{\xi} = \frac{(1 - \rho''^2_{\xi\eta})\rho''_{xx}(1 - d''_{\xi}\,\rho''_{\xi\eta}\rho''_{yy})}{(1 - \rho''^2_{\xi\eta}\,\rho''_{xx}\rho''_{yy})(1 - d''_{\xi}\,\rho''_{\xi\eta})}$$

$$= \frac{(1 - \rho^2_{\xi\eta})\rho_{xx}(1 - d_{\xi}\rho_{\xi\eta}\rho_{yy})}{(1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy})(1 - d_{\xi}\rho_{\xi\eta})} \quad \text{since} \quad \rho''^2_{\xi\eta} = (-\rho_{\xi\eta})^2 = \rho^2_{\xi\eta}$$

$$\text{and} \quad d''_{\xi}\rho''_{\xi\eta} = (-d)(-\rho_{\xi\eta}) = d\rho_{\xi\eta}$$

$$= R.$$

But $\beta_x/\beta^*_{\xi} = R$ for $\rho_{\xi\eta}$, ρ_{xx}, ρ_{yy} and d_{ξ} ($d_{\xi} < 0$). Thus for any situation where $d_{\xi} < 0$ there exists a comparable situation related by (B.1.5)-(B.1.8) with $d_{\xi} < 0$ which produces the same value for β_x/β^*_{ξ}.

Note that expressions (B.1.5) and (B.1.8) represent the only differences between the two situations. These changes can be expressed as: the expression for β_x/β^*_{ξ} with $d_{\xi} < 0$ is the reflection through the line $\rho_{\xi\eta} = 0$ of the expression for β_x/β^*_{ξ} with $|d_{\xi}| > 0$.

C. Summary

Therefore, part A indicates that only β_x/β_ξ^* need be examined and part B indicates β_x/β_ξ^* need be examined only for values of $d = d_\xi \geq 0$. Values of β_y/β_η^* or values of β_x/β_ξ^* with $d < 0$ can be simply stated from results for β_x/β_ξ^* with $d \geq 0$ by the use of expressions (B.1.1)-(B.1.4) or (B.1.5)-(B.1.8).

APPENDIX B.2

Examination of β_x/β_ξ^* for special case situations

A) $d_\eta = 0$ (i.e. $d = d_\xi = 1/d_\eta$ is undefined), B) $d = d_\xi = 0$, C) $\rho_{\xi\eta} = 0$, D) $\rho_{xx} = \rho_{yy} = 1$, and E) $\rho_{xx} = 1$, $\rho_{yy} < 1$ or $\rho_{xx} < 1$, $\rho_{yy} = 1$.

A.) $d_\eta = 0$ (i.e. $d = d_\xi = 1/d_\eta$ is undefined)

Recall $d_\eta = \dfrac{a_\xi/\sigma_\xi}{a_\eta/\sigma_\eta}$. Therefore $d_\eta = 0$ requires that $a_\xi = \mu_\xi^{(1)} - \mu_\xi^{(0)} = 0$, that is $\mu_\xi^{(1)} = \mu_\xi^{(0)}$. But if $a_\xi = 0$ then $d = d_\xi = \dfrac{a_\eta/\sigma_\eta}{a_\xi/\sigma_\xi}$ is undefined and expression (3.12) for β_x/β_ξ^* is not appropriate. In fact the expressions for β_x (3.11b) and β_ξ^* (3.6b) which were used to produce (3.12) are not appropriate since their derivations involve a division of a_ξ. Thus it is necessary to derive new expressions for β_x and β_ξ^* from (3.11a) and (3.6a), respectively, for $a_\xi = 0$.

Consider (3.11a)

$$\beta_x = \frac{1}{1 - \rho_{xy}^2}\left[\frac{a_x}{\sigma_x^2} - \frac{a_y\rho_{xy}}{\sigma_x\sigma_y}\right].$$

With $a_\xi = a_x = 0$,

$$\beta_x = \frac{1}{1 - \rho_{xy}^2}\left[- \frac{a_y\rho_{xy}}{\sigma_x\sigma_y}\right].$$

Using expressions (3.8) - (3.10) and expressing β_x in terms of latent parameters:

$$\text{(B.2.1)} \qquad \beta_x = \frac{1}{(1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy})}\left[- \frac{a_\xi \rho_{\xi\eta}\rho_{xx}\rho_{yy}}{\sigma_\xi \sigma_\eta}\right].$$

Consider (3.6a)

$$\beta^*_\xi = \frac{1}{1 - \rho^2_{\xi\eta}}\left[\frac{a_\xi}{\sigma^2_\xi} - \frac{a_\eta \rho_{\xi\eta}}{\sigma_\xi\sigma_\eta}\right].$$

With $a_\xi = 0$

$$\text{(B.2.2)} \qquad \beta^*_\xi = \frac{1}{1 - \rho^2_{\xi\eta}}\left[- \frac{a_\eta \rho_{\xi\eta}}{\sigma_\xi\sigma_\eta}\right].$$

Consider now β_x/β^*_ξ formed from (B.2.1) and (B.2.2):

$$\text{(B.2.3)} \qquad \beta_x/\beta^*_\xi = \frac{(1 - \rho^2_{\xi\eta})\rho_{xx}\rho_{yy}}{(1 - \rho^2_{\xi\eta}\rho_{xx}\rho_{yy})} .$$

Expression (B.2.3) for β_x/β^*_ξ when $a_\xi = 0$ will exist if:

1. $\rho_{\xi\eta} \neq \pm 1$ (Needed for Φ^{-1} to exist.)

   Note: $\rho_{xy} = \rho_{\xi\eta}\sqrt{\rho_{xx}\rho_{yy}}$, therefore $|\rho_{xy}| \leq |\rho_{\xi\eta}|$.
   Hence, if $\rho_{\xi\eta} \neq \pm 1$ then $\rho_{xy} \neq \pm 1$ and Σ^{-1} will exist.

2. $a_\eta \neq 0$

   If $a_\eta = 0$, as well as $a_\xi = 0$, then $\beta_x = \beta^*_\xi = 0$ and by (3.7a) and (3.13a), $\beta_y = \beta^*_\eta = 0$. That is, if there is no mean difference between categories on either of the predictor variables then there will be no information gained by the use of the predictor variables. In this case the unconditional probability of classification (assuming no knowledge of the predictors) will be equal to the conditional probability of classification (assuming knowledge of the predictors).

3. $\rho_{\xi\eta} \neq 0$

If $\rho_{\xi\eta} = 0$ and $a_{\xi} = 0$, then $\beta_x = \beta_{\xi}^* = 0$. That is, the observed predictor x contributes no weight to the probability of classification. Any value to the use of the conditional probability of classification over the unconditional probability (i.e. in the use of the predictors) must come solely from predictor y.

Note: When $a_{\xi} = 0$, β_x/β_{ξ}^* as given by (B.2.3) has the following property:

$$0 < \beta_x/\beta_{\xi}^* \leq 1 \text{ with equality only if } \rho_{xx} = \rho_{yy} = 1.$$

Proof: When $a_{\xi} = 0$, expression (B.2.3) for β_x/β_{ξ}^* is:

$$\beta_x/\beta_{\xi}^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}\rho_{yy}}{1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}} .$$

1.) Is $\beta_x/\beta_{\xi}^* > 1$? $\beta_x/\beta_{\xi}^* > 1 \Leftrightarrow \dfrac{(1 - \rho_{\xi\eta}^2)\rho_{xx}\rho_{yy}}{1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}} > 1$

$$\Leftrightarrow \rho_{xx}\rho_{yy} - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy} > 1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}$$

$$\Leftrightarrow \rho_{xx}\rho_{yy} > 1 \text{ which is impossible.}$$

Therefore $\beta_x/\beta_{\xi}^* \not> 1$.

2.) Is $\beta_x/\beta_{\xi}^* = 1$? $\beta_x/\beta_{\xi}^* = 1 \Leftrightarrow \dfrac{(1 - \rho_{\xi\eta}^2)\rho_{xx}\rho_{yy}}{1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}} = 1.$

Using algebra from part 1) above provides:

$$\beta_x/\beta_\xi^* = 1 \Leftrightarrow \rho_{xx}\rho_{yy} = 1 \Leftrightarrow \rho_{xx} = \rho_{yy} = 1 .$$

3.) Is $\beta_x/\beta_\xi^* < 1$? $\beta_x/\beta_\xi^* < 1 \Leftrightarrow \dfrac{(1 - \rho_{\xi\eta}^2)\rho_{xx}\rho_{yy}}{1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}} < 1.$

Using algebra from part 1) above provides:

$$\beta_x/\beta_\xi^* < 1 \Leftrightarrow \rho_{xx}\rho_{yy} < 1 \Leftrightarrow \rho_{xx} < 1 \text{ or } \rho_{yy} < 1.$$

Therefore $\beta_x/\beta_\xi^* \leq 1$ for any ρ_{xx}, ρ_{yy} when $a_\xi = 0$ with equality only for $\rho_{xx} = \rho_{yy} = 1$.

When $a_\xi = 0$, both the numerator and denominator of expression (B.2.3) for β_x/β_ξ^* are positive. Hence $\beta_x/\beta_\xi^* > 0$ for any ρ_{xx}, ρ_{yy} when $a_\xi = 0$.

Thus $0 < \beta_x/\beta_\xi^* \leq 1$ with equality only if $\rho_{xx} = \rho_{yy} = 1$.

To consider β_y/β_η^* where $d_\eta = 0$, the property of interchangeability of x and y can be applied to expression (3.12) for β_x/β_ξ^*. Using expressions (3.15a) - (3.15d) and interchanging the x's and y's and the ξ's and η's produces an expression for β_y/β_η^* with $d_\eta = 0$:

(B.2.4) $$\beta_y/\beta_\eta^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{yy}}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})} .$$

This same result can be obtained directly from expression (3.14) with $d_\eta = 0$. Therefore the conditions for existence of expression (3.14) for β_y/β_η^* also apply for expression (B.2.4).

Since expression (B.2.4) for β_y/β_η^* represents a special case of expression (3.14) no special considerations are needed for

β_y/β_η^* when $d_\eta = 0$ (i.e. $a_\xi = 0$). However, when $d_\eta = 0$ (i.e. $a_\xi = 0$) where $d = d_\xi$ is undefined, then expression (B.2.3) must be used for β_x/β_ξ^* instead of (3.12).

B.) $d = d_\xi = 0$

Recall $d = d_\xi = \dfrac{a_\eta/\sigma_\eta}{a_\xi/\sigma_\xi}$. Therefore $d_\xi = 0$ requires that $a_\eta = \mu_\eta^{(1)} - \mu_\eta^{(0)} = 0$ that is, $\mu_\eta^{(1)} = \mu_\eta^{(0)}$. Therefore, the general expression (3.12) for β_x/β_ξ^* is applicable, that is:

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy}} \quad \text{when} \quad d = d_\xi = 0.$$

To consider β_y/β_η^* when $d = d_\xi = 0$, the property of interchangeability of x and y cannot be applied to expression (3.12) for β_x/β_ξ^* to derive a comparable expression for β_y/β_η^*. When $d = d_\xi = 0$, then $a_\eta = 0$ thus $d_\eta = \dfrac{a_\xi/\sigma_\xi}{a_\eta/\sigma_\eta}$ is undefined. The derivation of expression (3.14) for β_y/β_η^*, which also results from applying the property of interchangeability of x and y to expression (3.12) for β_x/β_ξ^*, requires $a_\eta \neq 0$.

However, the property of interchangeability of x and y can be applied to expression (B.2.3), since this expression for β_x/β_ξ^* depends on $a_\xi = 0$ (i.e. $a_\eta = 0$ when the property of interchangeability is applied). Applying the property of interchangebability to expression (B.2.3) for β_x/β_ξ^* when $a_\xi = 0$ produces the following expression for β_y/β_η^* when $a_\eta = 0$:

$$\text{(B.2.5)} \qquad \beta_y/\beta_\eta^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{yy}\rho_{xx}}{(1 - \rho_{\xi\eta}^2\rho_{yy}\rho_{xx})}.$$

Existence conditions comparable to those for expression (B.2.3) for β_x/β_ξ^*, with the property of interchangeability applied, apply for (B.2.5).

Thus when $d = d_\xi = 0$ (i.e. $a_\eta = 0$), expression (3.12) is appropriate for β_x/β_ξ^*. But expression (B.2.5) must be used instead of (3.14) for β_y/β_η^* since d_η is undefined.

C.) $\rho_{\xi\eta} = 0$

When $\rho_{\xi\eta} = 0$, then expression (3.12) for β_x/β_ξ^* becomes

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})} = \rho_{xx}.$$

This result is identical to the result from the one predictor case. That is, in the two category, two predictor case if there is no correlation between the latent predictors (hence no correlation between the observed predictors either since $\rho_{xy} = \rho_{\xi\eta}\sqrt{\rho_{xx}\rho_{yy}}$) then the observed weighting coefficient will be attenuated by a factor equal to the reliability of the predictor, i.e. $\beta_x = \rho_{xx}\beta_\xi^*$.

Using the property of interchangeability a similar result and conclusion emerges for β_y/β_η^*:

$$\beta_y/\beta_\eta^* = \rho_{yy} \quad \text{or} \quad \beta_y = \rho_{yy}\beta_\eta^* .$$

These results are not surprising, since if there is no correlation between the predictors then the value and characteristics of one predictor cannot be expected to influence the conditional weighting coefficient of the other predictor.

D.) $\rho_{xx} = \rho_{yy} = 1$

For any $\rho_{\xi\eta} \in (-1, +1)$ and $\rho_{xx} = \rho_{yy} = 1$, that is both predictors are perfectly reliable, then (3.12) becomes

$$\beta_x/\beta_\xi^* = 1 \quad \text{that is} \quad \beta_x = \beta_\xi^* .$$

And (3.14) becomes

$$\beta_y/\beta_\eta^* = 1 \quad \text{that is} \quad \beta_y = \beta_\eta^* .$$

Again this result is not surprising. Logic suggests that if there are no errors of measurement in either predictor (i.e. $\rho_{xx} = \rho_{yy} = 1$ or $\sigma_\xi^2 = \sigma_x^2$, $\sigma_\eta^2 = \sigma_y^2$ or $x = \xi$, $y = \eta$) then the observed predictor model (3.4) and the latent predictor model (3.5) are, in fact, identical and thus the observed weighting coefficient for each predictor will be equal to the corresponding latent weighting coefficient.

E.) $\rho_{xx} = 1$, $\rho_{yy} < 1$ or $\rho_{xx} < 1$, $\rho_{yy} = 1$

Section D) above examined the case where both predictors are perfectly reliable (i.e. have no errors of measurement). This section examines the case when one predictor is perfectly reliable and the other predictor is fallible (i.e. has a reliability coefficient of less than one) which means that errors of measurement are present in only one of the two predictors.

If $\rho_{xx} = 1$, but $\rho_{yy} < 1$ and $\rho_{\xi\eta} \neq 0$ the expression (3.12) for β_x/β_ξ^* does not simplify. That is, (3.12) becomes

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})} .$$

This expression has no simple interpretation relative to one even though it represents the ratio of conditional weighting coefficients associated with an error-free predictor. This indicates that, based on the work so far, the observed weighting coefficient is neither a consistent underestimate or overestimate of the latent weighting coefficient. More work on this special case is included in Chapter 3 of this research.

If however, $\rho_{yy} = 1$ and $\rho_{xx} < 1$ the expression for β_x/β_ξ^* simplifies to become:

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{(1 - \rho_{\xi\eta}^2\rho_{xx})} \quad \text{for} \quad d\rho_{\xi\eta} \neq 1 .$$

In this case β_x/β_ξ^* can be shown to have a ratio between zero and one for all situations where $d\rho_{\xi\eta} \neq 1$ indicating that the observed weighting coefficient will be an underestimate of the latent weighting coefficient for the fallible predictor.

Proof: Let $\rho_{xx} < 1$, $\rho_{yy} = 1$, and $d\rho_{\xi\eta} \neq 1$.

1. Is $\beta_x/\beta_\xi^* > 1$?

$$\beta_x/\beta_\xi^* > 1 \Leftrightarrow \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{1 - \rho_{\xi\eta}^2\rho_{xx}} > 1$$

$$\Leftrightarrow (1 - \rho_{\xi\eta}^2)\rho_{xx} > 1 - \rho_{\xi\eta}^2\rho_{xx}$$

$$\Leftrightarrow \rho_{xx} - \rho_{\xi\eta}^2\rho_{xx} > 1 - \rho_{\xi\eta}^2\rho_{xx}$$

$$\Leftrightarrow \rho_{xx} > 1 \quad \text{which is impossible.}$$

Therefore $\beta_x/\beta_\xi^* \not> 1$.

2. Is $\beta_x/\beta_\xi^* = 1$?

$$\beta_x/\beta_\xi^* = 1 \Leftrightarrow \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{1 - \rho_{\xi\eta}^2\rho_{xx}} = 1 \ .$$

Using algebra from part 1 above produces:

$$\beta_x/\beta_\xi^* = 1 \Leftrightarrow \rho_{xx} = 1.$$

3. Is $\beta_x/\beta_\xi^* < 1$?

$$\beta_x/\beta_\xi^* < 1 \Leftrightarrow \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{1 - \rho_{\xi\eta}^2\rho_{xx}} < 1 \ .$$

Using algebra from part 1 above produces;

$$\beta_x/\beta_\xi^* < 1 \Leftrightarrow \rho_{xx} < 1.$$

(B.2.6) Therefore, when $\rho_{yy} = 1$, $\rho_{xx} < 1$, and $d\rho_{\xi\eta} \neq 1$, then $\beta_x/\beta_\xi^* < 1$.

Consider now the relationship of β_x/β_ξ^* to zero. When $\rho_{yy} = 1$, $\rho_{xx} < 1$, and $d\rho_{\xi\eta} \neq 1$, $\beta_x/\beta_\xi^* = \dfrac{(1 - \rho_{\xi\eta}^2)\rho_{xx}}{1 - \rho_{\xi\eta}^2\rho_{xx}}$. Since $-1 < \rho_{\xi\eta} < +1$, and $0 < \rho_{xx} < 1$, then $0 < \rho_{\xi\eta}^2 < 1$ and $0 < \rho_{\xi\eta}^2\rho_{xx} < 1$. Therefore both the numerator and denominator of β_x/β_ξ^* are positive.

(B.2.7) Hence, for any $\rho_{\xi\eta} \in (-1, +1)$, such that $d\rho_{\xi\eta} \neq 1$, $0 < \rho_{xx} < 1$ and $\rho_{yy} = 1$, $\beta_x/\beta_\xi^* > 0$ [Note: even for $\rho_{xx} = 1$, $\beta_x/\beta_\xi^* = \rho_{xx} > 0$.]

Combining results (B.2.6) and (B.2.7) produces:

(B.2.8) For $\rho_{yy} = 1$, $0 < \rho_{xx} < 1$ and $\rho_{\xi\eta} \in (-1, +1)$ such that $d\rho_{\xi\eta} \neq 1$, then $0 < \beta_x/\beta^*_\xi < 1$.

Comparable results for β_y/β^*_η could be stated using the property of interchangeability for x and y.

Therefore in a two category two predictor model with one error-free ($\rho_{ii} = 1$) and one error-ful ($\rho_{jj} < 1$) predictor the results above and the property of interchangeability of x and y indicate that the ratio of conditional weighting coefficients associated with the error-free predictor ($\rho_{ii} = 1$) has no simple interpretation while the ratio of conditional weighting coefficients associated with the error-ful predictor ($\rho_{jj} < 1$) will be less than one indicating that the observed weighting coefficient will always be an underestimate of the latent weighting coefficient for the error-ful predictor.

APPENDIX B.3

ALGEBRAIC EXAMINATIONS OF RELATIONSHIPS BETWEEN EXPRESSIONS NEEDED FOR WORK IN APPENDIX B.4

Relationships among three expressions will be of interest for the algebra to be produced in Appendix B.4. The three expressions are:

$$\sqrt{x} = \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} \; ; \quad f_x = \frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} \; ; \; 1 \; .$$

a) Consider $\sqrt{x} \; \mathcal{R} \; 1$ i.e. find the values of ρ_{xx} and ρ_{yy} such that:

1) $\sqrt{x} = 1$

$$\sqrt{x} = 1 \Leftrightarrow \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} = 1$$

$$\Leftrightarrow \frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2} = 1$$

$$\Leftrightarrow 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}) = (1 - \rho_{xx}\rho_{yy})^2$$

$$\Leftrightarrow 4\rho_{xx} - 4\rho_{xx}\rho_{yy} - 4\rho_{xx}^2 + 4\rho_{yy}\rho_{xx}^2$$

$$= 1 - 2\rho_{xx}\rho_{yy} + \rho_{xx}^2\rho_{yy}^2$$

$$\Leftrightarrow 0 = \rho_{xx}^2\rho_{yy}^2 + 4\rho_{xx}^2 - 4\rho_{yy}\rho_{xx}^2 + 2\rho_{xx}\rho_{yy} - 4\rho_{xx} + 1$$

$$\Leftrightarrow 0 = \rho_{xx}^2(\rho_{yy}^2 - 4\rho_{yy} + 4) + 2\rho_{xx}(\rho_{yy} - 2) + 1$$

$$\Leftrightarrow 0 = \rho_{xx}^2(\rho_{yy} - 2)^2 + 2\rho_{xx}(\rho_{yy} - 2) + 1$$

$$\Leftrightarrow 0 = (\rho_{xx}(\rho_{yy} - 2) + 1)^2$$

$$\Leftrightarrow 0 = \rho_{xx}(\rho_{yy} - 2) + 1$$

$$\Leftrightarrow (2 - \rho_{yy})\rho_{xx} = 1 \qquad \Leftrightarrow 2\rho_{xx} - \rho_{yy}\rho_{xx} = 1$$

$$\Leftrightarrow \rho_{xx} = \frac{1}{2 - \rho_{yy}} \qquad \Leftrightarrow 2\rho_{xx} - 1 = \rho_{yy}\rho_{xx}$$

$$\Leftrightarrow \frac{2\rho_{xx} - 1}{\rho_{xx}} = \rho_{yy}$$

(B.3.1) Therefore $\sqrt{x} = 1 \Leftrightarrow \rho_{xx} = \dfrac{1}{2 - \rho_{yy}} \Leftrightarrow \rho_{yy} = \dfrac{2\rho_{xx} - 1}{\rho_{xx}}$

2) $\sqrt{x} < 1$

$$\sqrt{x} < 1 \Leftrightarrow \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} < 1$$

$$\Leftrightarrow \frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2} < 1 \qquad \text{since } 0 \leq \sqrt{x}$$

$$\Leftrightarrow 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}) < (1 - \rho_{xx}\rho_{yy})^2$$

if $\rho_{xx}\rho_{yy} \neq 1$.

Using algebra from a) 1) above with appropriate attention to the inequality here produces:

$\sqrt{x} < 1 \Leftrightarrow 0 < [\rho_{xx}(\rho_{yy} - 2) + 1]^2$ which is true for all values of ρ_{xx}, ρ_{yy} except for $\rho_{xx} = \dfrac{1}{2 - \rho_{yy}}$ (or $\rho_{yy} = \dfrac{2\rho_{xx} - 1}{\rho_{xx}}$).

3) $\sqrt{x} > 1$

$$\sqrt{x} > 1 \Leftrightarrow \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} > 1$$

Using algebra from parts a) 1) and a) 2) above produces:

$$\sqrt{x} > 1 \Leftrightarrow 0 > [\rho_{xx}(\rho_{yy} - 2) + 1]^2 \quad \text{impossible.}$$

Therefore

$$0 \leq \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} \leq 1 \quad \text{for all values of } \rho_{xx}, \rho_{yy} \qquad \text{(B.3.2)}$$

with equality to one (1) when $\rho_{xx} = \dfrac{1}{2 - \rho_{yy}}$ (or $\rho_{yy} = \dfrac{2\rho_{xx} - 1}{\rho_{xx}}$) and with equality to zero (0) when either $\rho_{xx} = 1$ or $\rho_{yy} = 1$ but $\rho_{xx}\rho_{yy} \neq 1$.

b) Consider $\sqrt{x} \; \mathcal{R} \; f_x$ i.e. find ρ_{xx} and ρ_{yy} such that

1) $\sqrt{x} = f_x$

$$\sqrt{x} = f_x \Leftrightarrow \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} = \frac{2\rho_{xx}(1 - \rho_{yy})}{(1 - \rho_{xx}\rho_{yy})}$$

$$\Leftrightarrow \frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2} = \frac{4\rho_{xx}^2(1 - \rho_{yy})^2}{(1 - \rho_{xx}\rho_{yy})^2}$$

since $\sqrt{x} \geq 0$ and $f_x \geq 0$

$$\Leftrightarrow 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}) = 4\rho_{xx}^2(1 - \rho_{yy})^2$$

$$\Leftrightarrow 0 = 4\rho_{xx}^2(1 - \rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})[\rho_{xx}(1 - \rho_{yy}) - (1 - \rho_{xx})]$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})[\rho_{xx} - \rho_{xx}\rho_{yy} - 1 + \rho_{xx}]$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})(2\rho_{xx} - \rho_{xx}\rho_{yy} - 1)$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})(\rho_{xx}(2 - \rho_{yy}) - 1).$$

(B.3.3a) Therefore $\sqrt{\ }_x = f_x$ if $1 - \rho_{yy} = 0 \Leftrightarrow \rho_{yy} = 1$

or if $\rho_{xx}(2 - \rho_{yy}) - 1 = 0 \Leftrightarrow \rho_{xx} = \dfrac{1}{2 - \rho_{yy}} \Leftrightarrow \rho_{yy} = \dfrac{2\rho_{xx} - 1}{\rho_{xx}}$.

2) $\sqrt{\ }_x < f_x$

$$\sqrt{\ }_x < f_x \Leftrightarrow \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} < \frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}}$$

$$\Leftrightarrow \frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2} < \frac{4\rho_{xx}^2(1 - \rho_{yy})^2}{(1 - \rho_{xx}\rho_{yy})^2}$$

since $\sqrt{\ }_x \geq 0$ and $f_x \geq 0$

$$\Leftrightarrow 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}) < 4\rho_{xx}^2(1 - \rho_{yy})^2.$$

Using algebra from b) 1) above with appropriate attention to the inequality produces

$$\sqrt{\ }_x < f_x \Leftrightarrow 0 < 4\rho_{xx}(1 - \rho_{yy})(\rho_{xx}(2 - \rho_{yy}) - 1)$$

Therefore

(B.3.3b) $\sqrt{\ }_x < f_x$ if $\rho_{yy} \neq 1$ and

$$\rho_{xx}(2 - \rho_{yy}) - 1 > 0 \Leftrightarrow \rho_{xx} > \frac{1}{2 - \rho_{yy}} \Leftrightarrow \rho_{yy} < \frac{2\rho_{xx} - 1}{\rho_{xx}}.$$

3) $\sqrt{x} > f_x$

$$\sqrt{x} > f_x \Leftrightarrow \sqrt{\frac{4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}{(1-\rho_{xx}\rho_{yy})^2}} > \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} .$$

Using algebra from b) 1) and b) 2) above produces:

$$\sqrt{x} > f_x \Leftrightarrow 0 > 4\rho_{xx}(1-\rho_{yy})(\rho_{xx}(2-\rho_{yy}) - 1).$$

Therefore

(B.3.3c) $$\sqrt{x} > f_x \Leftrightarrow \rho_{yy} \neq 1 \text{ and}$$

$$\rho_{xx}(2-\rho_{yy}) - 1 < 0 \Leftrightarrow \rho_{xx} < \frac{1}{2-\rho_{yy}} \Leftrightarrow \rho_{yy} > \frac{2\rho_{xx} - 1}{\rho_{xx}} .$$

c) Consider $f_x \mathrel{R} 1$, i.e. find values of ρ_{xx} and ρ_{yy} such that:

1) $f_x = 1$

$$f_x = 1 \Leftrightarrow \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} = 1$$

$$\Leftrightarrow 2\rho_{xx}(1-\rho_{yy}) = 1 - \rho_{xx}\rho_{yy}$$

$$\Leftrightarrow 2\rho_{xx} - 2\rho_{xx}\rho_{yy} = 1 - \rho_{xx}\rho_{yy}$$

$$\Leftrightarrow 2\rho_{xx} - \rho_{xx}\rho_{yy} - 1 = 0$$

$$\Leftrightarrow \rho_{xx}(2-\rho_{yy}) - 1 = 0.$$

Therefore

(B.3.4a) $$f_x = 1 \Leftrightarrow \rho_{xx} = \frac{1}{2-\rho_{yy}} \Leftrightarrow \rho_{yy} = \frac{2\rho_{xx} - 1}{\rho_{xx}}$$

c) f_x 1 (cont'd.)

2) $f_x < 1$

$$f_x < 1 \Leftrightarrow \frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} < 1$$

$$\Leftrightarrow 2\rho_{xx}(1 - \rho_{yy}) < 1 - \rho_{xx}\rho_{yy} \quad \text{for} \quad \rho_{xx}\rho_{yy} \neq 1.$$

Using algebra from c) 1) with appropriate attention to the inequality produces:

$$f_x < 1 \Leftrightarrow \rho_{xx}(2 - \rho_{yy}) - 1 < 0.$$

Therefore

$$\text{(B.3.4b)} \quad f_x < 1 \Leftrightarrow \rho_{xx} < \frac{1}{2 - \rho_{yy}} \Leftrightarrow \rho_{yy} > \frac{2\rho_{xx} - 1}{\rho_{xx}} .$$

3) $f_x > 1$

$$f_x > 1 \Leftrightarrow \frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} > 1 .$$

Using algebra from c) 1) and c) 2) above produces:

$$f_x > 1 \Leftrightarrow \rho_{xx}(2 - \rho_{yy}) - 1 > 0.$$

Therefore

$$\text{(B.3.4c)} \quad f_x > 1 \Leftrightarrow \rho_{xx} > \frac{1}{2 - \rho_{yy}} \Leftrightarrow \rho_{yy} < \frac{2\rho_{xx} - 1}{\rho_{xx}} .$$

Combining results from parts a), b) and c) above produces:

for $0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}}$, $0 \leq f_x \leq \sqrt{x} \leq 1$

(B.3.5) with the rightmost two equalities occurring only if

$$\rho_{xx} = \frac{1}{2 - \rho_{yy}} .$$

[Hence $-1 \leq -\sqrt{x} \leq -f_x \leq 0$.]

for $\frac{1}{2 - \rho_{yy}} < \rho_{xx} < 1$, $0 \leq \sqrt{x} < 1 < f_x$.

[Hence $-f_x < -1 < -\sqrt{x} \leq 0$.]

APPENDIX B.4

Examination of $Q = \rho_{\xi\eta}^2 \rho_{xx}(1 - \rho_{yy}) - d\rho_{\xi\eta}(1 - \rho_{xx}\rho_{yy}) + (1 - \rho_{xx})$ as a function of $\rho_{\xi\eta}$. Identification and determination of existence conditions of the roots of Q as a function of $\rho_{\xi\eta}$.

Expression (3.17) defines Q as:

$$Q = \rho_{\xi\eta}^2 \rho_{xx}(1 - \rho_{yy}) - d\rho_{\xi\eta}(1 - \rho_{xx}\rho_{yy}) + (1 - \rho_{xx}) \ .$$

Expression (3.20) for Q as a function of $\rho_{\xi\eta}$ is:

$$Q = a_x \rho_{\xi\eta}^2 + b_x \rho_{\xi\eta} + c_x \quad \text{where} \quad a_x = \rho_{xx}(1 - \rho_{yy})$$

$$b_x = -d(1 - \rho_{xx}\rho_{yy})$$

$$\text{and} \quad c_x = (1 - \rho_{xx}).$$

This clearly indicates that Q is a quadratic function of $\rho_{\xi\eta}$.

Note: $\dfrac{\partial Q}{\partial \rho_{\xi\eta}} = 2a_x \rho_{\xi\eta} + b_x$, $\dfrac{\partial^2 Q}{\partial \rho_{\xi\eta}^2} = 2a_x$.

Since $a_x = \rho_{xx}(1 - \rho_{yy}) > 0$, $\dfrac{\partial^2 Q}{\partial \rho_{\xi\eta}^2} = 2a_x > 0$.

Therefore, Q as a function of $\rho_{\xi\eta}$ is concave upward and will possess a minimum value at the point where $\dfrac{\partial Q}{\partial \rho_{\xi\eta}} = 0$ i.e. where $\rho_{\xi\eta} = \dfrac{-b_x}{2a_x} = \dfrac{d(1 - \rho_{xx}\rho_{yy})}{2\rho_{xx}(1 - \rho_{yy})}$, provided that

$$-1 < \frac{d(1 - \rho_{xx}\rho_{yy})}{2\rho_{xx}(1 - \rho_{yy})} < +1.$$

However, since the concern here is only with the relationship of Q to zero (thus the relationship of β_x/β_ξ^* to one by expressions (3.18a) - (3.18c) or (3.19a) - (3.19c)) as $\rho_{\xi\eta}$ varies over its domain, the existence and precise location of a minimum value for Q is of little importance.

Since Q is a quadratic function of $\rho_{\xi\eta}$ for fixed values of ρ_{xx}, ρ_{yy} and d, Q will possess two roots identified in Chapter 3 as $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ defined by expressions (3.21a) and (3.21b) respectively:

$$\rho_{\xi\eta}^{-(x)} = \frac{d(1 - \rho_{xx}\rho_{yy}) - \sqrt{d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}}{2\rho_{xx}(1 - \rho_{yy})}$$

$$\rho_{\xi\eta}^{+(x)} = \frac{d(1 - \rho_{xx}\rho_{yy}) + \sqrt{d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}}{2\rho_{xx}(1 - \rho_{yy})} .$$

Using expression (3.20) for Q, $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ can be expressed as:

$$\text{(B.4.1a)} \quad \rho_{\xi\eta}^{-(x)} = \frac{-b_x - \sqrt{b_x^2 - 4a_xc_x}}{2a_x} \quad \text{where} \quad a_x = \rho_{xx}(1 - \rho_{yy})$$

$$\text{(B.4.1b)} \quad \rho_{\xi\eta}^{+(x)} = \frac{-b_x + \sqrt{b_x^2 - 4a_xc_x}}{2a_x} \qquad b_x = -d(1 - \rho_{xx}\rho_{yy})$$

$$c_x = (1 - \rho_{xx}) .$$

Therefore examining the relationship of Q to 0 will begin by determining the conditions for existence of $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ such that $Q = 0$, using expression (3.20) for Q.

$$Q = 0 \Leftrightarrow a_x\rho_{\xi\eta}^2 + b_x\rho_{\xi\eta} + c_x = 0 .$$

Therefore $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ will exist as real numbers but not necessarily in the interval (-1, +1), if $b_x^2 - 4a_x c_x \geq 0$.

$$b_x^2 - 4a_x c_x \geq 0 \Leftrightarrow d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}) \geq 0$$

$$\Leftrightarrow d^2 \geq \frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2} \quad \text{for} \quad 1 - \rho_{xx}\rho_{yy} > 0$$

$$\rho_{xx}\rho_{yy} < 1$$

$$\rho_{xx}\rho_{yy} \neq 1.$$

(B.4.2) $$b_x^2 - 4a_x c_x \geq 0 \Leftrightarrow |d| \geq \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}} \equiv \sqrt{x}$$

[definition by (3.22)]

When $b_x^2 - 4a_x c_x = 0$ then $\rho_{\xi\eta}^{-(x)} = \rho_{\xi\eta}^{+(x)}$.

To translate this into information about the relationship between Q and 0, ignore temporarily the requirement that values of correlations, $\rho_{\xi\eta}$, need to be in the interval (-1, +1). Thus if $|d| \geq \sqrt{x}$ (that is, if $b_x^2 - 4a_x c_x \geq 0$) then there exists $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ where $\rho_{\xi\eta}^{-(x)} \leq \rho_{\xi\eta}^{+(x)}$.

Since $Q = a_x\rho_{\xi\eta}^2 + b_x\rho_{\xi\eta} + c_x$ is concave upward then for $\rho_{\xi\eta}^{-(x)} < \rho_{\xi\eta} < \rho_{\xi\eta}^{+(x)}$, $Q < 0$

$$\text{for} \quad \left.\begin{array}{l} \rho_{\xi\eta} < \rho_{\xi\eta}^{-(x)} \\ \rho_{\xi\eta} > \rho_{\xi\eta}^{+(x)} \end{array}\right\} \quad Q > 0$$

$$\text{and for} \quad \left.\begin{array}{l} \rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)} \\ \rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)} \end{array}\right\} \quad Q = 0 .$$

Now considering the requirement that $\rho_{\xi\eta} \in (-1, +1)$, if $b_x^2 - 4a_xc_x \geq 0$, then $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ will exist and

(B.4.3a) $Q < 0$ for $\max(-1,\rho_{\xi\eta}^{-(x)}) < \rho_{\xi\eta} < \min(+1,\rho_{\xi\eta}^{+(x)})$

(B.4.3b) $Q = 0$ for $\rho_{\xi\eta} = \rho_{\xi\eta}^{-(x)}$ provided that $-1 < \rho_{\xi\eta}^{-(x)} < +1$

or $\rho_{\xi\eta} = \rho_{\xi\eta}^{+(x)}$ provided that $-1 < \rho_{\xi\eta}^{+(x)} < +1$

(B.4.3c) $Q > 0$ for $-1 < \rho_{\xi\eta} < \max(-1,\rho_{\xi\eta}^{-(x)})$

or $\min(+1,\rho_{\xi\eta}^{+(x)}) < \rho_{\xi\eta} < +1$.

If, however, $b_x^2 - 4a_xc_x < 0$ then neither $\rho_{\xi\eta}^{-(x)}$ nor $\rho_{\xi\eta}^{+(x)}$ exist as real numbers, that is, there does not exist any $\rho_{\xi\eta}$ such that $Q = 0$. But since Q is concave upward for all $\rho_{\xi\eta}$ then

(B.4.4) $Q > 0$ for all $\rho_{\xi\eta} \in (-1, +1)$ when $b_x^2 - 4a_xc_x < 0 \Leftrightarrow |d| < \sqrt{x}$.

Now consider the conditions under which $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ or that $\rho_{\xi\eta}^{+(x)} \in (-1,+1)$. The approach to this problem will be to first examine $\rho_{\xi\eta}^{-(x)}$ when $|d| \geq \sqrt{x}$. The first question to be determined is: for what values of d, ρ_{xx} and ρ_{yy} will $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ [i.e. $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$?]. The second question will then be: for what values of d, ρ_{xx} and ρ_{yy} will $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$ [i.e. $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$?].

Both $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ will exist with $\rho_{\xi\eta}^{-(x)} \leq \rho_{\xi\eta}^{+(x)}$ if and only if

$$|d| \geq \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}}\ .$$

Let $\sqrt{x} \equiv \sqrt{\dfrac{4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}{(1-\rho_{xx}\rho_{yy})^2}}$ and

$$f_x = \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} .$$

The task here is to find values of d, ρ_{xx} and ρ_{yy} such that 1) $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ and 2) $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$. Consider only situations where $\rho_{\xi\eta}^{-(x)}$ and $\rho_{\xi\eta}^{+(x)}$ exist, that is, when $|d| \geq \sqrt{x}$.

1) $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$?

a) Consider $-1 < \rho_{\xi\eta}^{-(x)}$.

$$-1 < \rho_{\xi\eta}^{-(x)} \Leftrightarrow -1 < \frac{d(1-\rho_{xx}\rho_{yy}) - \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})}$$

$$\Leftrightarrow -2\rho_{xx}(1-\rho_{yy}) < d(1-\rho_{xx}\rho_{yy}) - \sqrt{}$$

since $2\rho_{xx}(1-\rho_{yy}) > 0$ for $\rho_{yy} \neq 1$.

$$\Leftrightarrow 0 \leq \sqrt{} < d(1-\rho_{xx}\rho_{yy}) + 2\rho_{xx}(1-\rho_{yy}).$$

Therefore

(1) $0 < d(1-\rho_{xx}\rho_{yy}) + 2\rho_{xx}(1-\rho_{yy})$

$$\Leftrightarrow \boxed{d > \frac{-2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} = -f_x}$$

and (2) $\sqrt{} < d(1-\rho_{xx}\rho_{yy}) + 2\rho_{xx}(1-\rho_{yy})$

$$\Leftrightarrow d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}) < d^2(1-\rho_{xx}\rho_{yy})^2 + 4\rho_{xx}^2(1-\rho_{yy})^2 + 4d\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$\Leftrightarrow 0 < 4d\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy}) + 4\rho_{xx}^2(1-\rho_{yy})^2$$

$$+ 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})[d(1-\rho_{xx}\rho_{yy}) + \rho_{xx}(1-\rho_{yy}) + (1-\rho_{xx})]$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})[d(1-\rho_{xx}\rho_{yy}) + (1-\rho_{xx}\rho_{yy})]$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})(d + 1)$$

$$\Leftrightarrow 0 < d + 1 \quad \text{for} \quad \rho_{yy} \neq 1$$

$$\Leftrightarrow \boxed{d > -1} \, .$$

Therefore $\rho_{\xi\eta}^{-(x)}$ will exist and $-1 < \rho_{\xi\eta}^{-(x)}$ for $\rho_{yy} \neq 1$

$$\text{if} \quad |d| \geq \sqrt{\frac{4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}{(1-\rho_{xx}\rho_{yy})^2}} \equiv \sqrt{x} \Leftrightarrow d \geq \sqrt{x} \quad \text{or} \quad d \leq -\sqrt{x}$$

$$\text{and} \quad d \geq \frac{-2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} = -f_x$$

and $d > -1$.

Using the results (B.3.5) from Appendix B.3 produces:

(B.4.5a) $\rho_{\xi\eta}^{-(x)}$ exists and $-1 < \rho_{\xi\eta}^{-(x)}$ for $\rho_{yy} \neq 1$

$$\text{for} \quad 0 < \rho_{xx} \leq \frac{1}{2-\rho_{yy}} \quad \text{if} \quad d \geq \sqrt{x}$$

$$\text{for} \quad \frac{1}{2-\rho_{yy}} < \rho_{xx} \leq 1 \quad \text{if} \quad d \geq \sqrt{x}$$

$$\text{or if} \quad -1 < d < -\sqrt{x}$$

1) (cont'd.)

b) Consider $\rho_{\xi\eta}^{-(x)} < +1$

$$\rho_{\xi\eta}^{-(x)} < +1 \Leftrightarrow \frac{d(1-\rho_{xx}\rho_{yy}) - \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2-4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})} < +1$$

$$\Leftrightarrow d(1-\rho_{xx}\rho_{yy}) - \sqrt{} < 2\rho_{xx}(1-\rho_{yy})$$

$$\Leftrightarrow d(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy})$$

(I) $\rho_{\xi\eta}^{-(x)} < +1$ if $d(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy}) < 0$

[since $\sqrt{} \geq 0$]

$$\Leftrightarrow \boxed{d < \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} = f_x}$$

or (II) $\rho_{\xi\eta}^{-(x)} < +1$ if $0 \leq d(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy}) < \sqrt{}$

(1) $0 \leq d(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy})$

$$\Leftrightarrow \boxed{d \geq \frac{2\rho_{xx}(1-\rho_{yy})}{(1-\rho_{xx}\rho_{yy})} = f_x}$$

and (2) $d(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy}) < \sqrt{}$

$$\Leftrightarrow d^2(1-\rho_{xx}\rho_{yy})^2 + 4\rho_{xx}^2(1-\rho_{yy})^2 - 4d\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$< d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})$$

$$\Leftrightarrow 0 < -4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}) - 4\rho_{xx}^2(1-\rho_{yy})^2$$

$$+ 4d\rho_{xx}(1-\rho_{xx}\rho_{yy})(1-\rho_{yy})$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})[\rho_{xx}-1-\rho_{xx}(1-\rho_{yy}) + d(1-\rho_{xx}\rho_{yy})]$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})[-(1-\rho_{xx}\rho_{yy}) + d(1-\rho_{xx}\rho_{yy})]$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})(d - 1)$$

$$\Leftrightarrow 0 < d - 1 \quad \text{for} \quad \rho_{yy} \neq 1$$

$$\Leftrightarrow \boxed{d > 1} \quad \text{for} \quad \rho_{yy} \neq 1.$$

Therefore $\rho_{\xi\eta}^{-(x)}$ will exist and $\rho_{\xi\eta}^{-(x)} < +1$

if $|d| \geq \sqrt{x}$

and either

$d < f_x$

or $d > 1$ and $d \geq f_x$.

Using the results (B.3.5) from Appendix B.3

(B.4.5b) $\rho_{\xi\eta}^{-(x)}$ exists and $\rho_{\xi\eta}^{-(x)} < +1$, for $\rho_{yy} \neq 1$

for $0 < \rho_{xx} \leq \dfrac{1}{2 - \rho_{yy}}$ if $d > +1$

or if $d \leq -\sqrt{x}$

for $\dfrac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ if $d \leq -\sqrt{x}$

or if $d \geq +\sqrt{x}$

Combining results (B.4.5a) and (B.4.5b) produces:

(B.4.5c) $\rho_{\xi\eta}^{-(x)}$ exists and $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$, for $\rho_{yy} \neq 1$,

for $0 < \rho_{xx} \leq \dfrac{1}{2 - \rho_{yy}}$ if $d > +1$

for $\dfrac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ if $d \geq +\sqrt{x}$

or if $-1 < d < -\sqrt{x}$.

2) $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$?

a) Consider $-1 < \rho_{\xi\eta}^{+(x)}$.

$$-1 < \rho_{\xi\eta}^{+(x)} \Leftrightarrow -1 < \frac{d(1-\rho_{xx}\rho_{yy}) + \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})}$$

$$\Leftrightarrow -2\rho_{xx}(1-\rho_{yy}) < d(1-\rho_{xx}\rho_{yy}) + \sqrt{\quad}$$

$$\Leftrightarrow -2\rho_{xx}(1-\rho_{yy}) - d(1-\rho_{xx}\rho_{yy}) < \sqrt{\quad}$$

(I) $-1 < \rho_{\xi\eta}^{+(x)}$ if $-2\rho_{xx}(1-\rho_{yy}) - d(1-\rho_{xx}\rho_{yy}) < 0$

[Since $\sqrt{\quad} \geq 0$]

$$\Leftrightarrow \boxed{d > \frac{-2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} = -f_x}$$

or (II) $-1 < \rho_{\xi\eta}^{+(x)}$ if $0 \leq -2\rho_{xx}(1-\rho_{yy}) - d(1-\rho_{xx}\rho_{yy}) < \sqrt{\quad}$

(1) $0 \leq -2\rho_{xx}(1-\rho_{yy}) - d(1-\rho_{xx}\rho_{yy})$

$$\Leftrightarrow \boxed{d \leq \frac{-2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} = -f_x}$$

and (2) $-2\rho_{xx}(1-\rho_{yy}) - d(1-\rho_{xx}\rho_{yy}) < \sqrt{\quad}$

$$\Leftrightarrow 4\rho_{xx}^2(1-\rho_{yy})^2 + d^2(1-\rho_{xx}\rho_{yy})^2 + 4d\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$< d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})$$

$$\Leftrightarrow 4\rho_{xx}^2(1-\rho_{yy})^2 + 4d\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$+ 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}) < 0$$

$$\Leftrightarrow 4\rho_{xx}(1-\rho_{yy})[\rho_{xx}(1-\rho_{yy}) + d(1-\rho_{xx}\rho_{yy}) + 1-\rho_{xx}] < 0$$

$$\Leftrightarrow 4\rho_{xx}(1-\rho_{yy})\left[\, d(1-\rho_{xx}\rho_{yy}) + (1-\rho_{xx}\rho_{yy}) \right] < 0$$

$$\Leftrightarrow 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})(d+1) < 0$$

$$\Leftrightarrow d + 1 < 0 \quad \text{for} \quad \rho_{yy} \neq 1$$

$$\Leftrightarrow \boxed{d < -1} \quad \text{for} \quad \rho_{yy} \neq 1.$$

Therefore $\rho_{\xi\eta}^{+(x)}$ will exist and $-1 < \rho_{\xi\eta}^{+(x)}$ if $\rho_{yy} \neq 1$ and

if $|d| \geq \sqrt{x}$

and either $d > -f_x$

or $d < -1$ and $d \leq -f_x$.

Using results (B.35) from Appendix B.3 produces:

(B.4.6a) $\rho_{\xi\eta}^{+(x)}$ exists and $-1 < \rho_{\xi\eta}^{+(x)}$ for $\rho_{yy} \neq 1$

for $0 < \rho_{xx} \leq \dfrac{1}{2-\rho_{yy}}$ if $d \geq \sqrt{x}$

or if $d < -1$

for $\dfrac{1}{2-\rho_{yy}} < \rho_{xx} \leq 1$ if $d \geq \sqrt{x}$

or if $d \leq -\sqrt{x}$

b) Consider $\rho_{\xi\eta}^{+(x)} < +1$

$$\rho_{\xi\eta}^{+(x)} < +1 \Leftrightarrow \frac{d(1-\rho_{xx}\rho_{yy}) + \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})} < +1$$

$$\Leftrightarrow d(1-\rho_{xx}\rho_{yy}) + \sqrt{} < 2\rho_{xx}(1-\rho_{yy})$$

$$\Leftrightarrow \sqrt{} < 2\rho_{xx}(1-\rho_{yy}) - d(1-\rho_{xx}\rho_{yy}).$$

2) b) (cont'd.)

(1) $0 < 2\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy})$

$$\Leftrightarrow \boxed{d < \frac{2\rho_{xx}(1 - \rho_{yy})}{1 - \rho_{xx}\rho_{yy}} = f_x}$$

and (2) $\sqrt{} < 2\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy})$

$$\sqrt{} < 2\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy})$$

$$\Leftrightarrow d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}) < 4\rho_{xx}^2(1-\rho_{yy}^2)$$

$$+ d^2(1-\rho_{xx}\rho_{yy})^2 - 4d\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}) + 4\rho_{xx}^2(1-\rho_{yy})^2$$

$$- 4d\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})[1-\rho_{xx} + \rho_{xx}(1-\rho_{yy}) - d(1-\rho_{xx}\rho_{yy})]$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1-\rho_{yy})[1-\rho_{xx}\rho_{yy} - d(1-\rho_{xx}\rho_{yy})]$$

$$\Leftrightarrow 0 < 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}\rho_{yy})(1 - d)$$

$$\Leftrightarrow 0 < 1 - d \qquad \text{for} \quad \rho_{yy} \neq 1$$

$$\Leftrightarrow \boxed{d < 1} \qquad \text{for} \quad \rho_{yy} \neq 1.$$

Therefore $\rho_{\xi\eta}^{+(x)}$ will exist and $\rho_{\xi\eta}^{+(x)} < +1$ for $\rho_{yy} \neq 1$

if $|d| \geq \sqrt{x}$

and if $d < f_x$

and if $d < 1$.

Using the results (B.3.5) from Appendix B.3 produces:

(B.4.6b) $\rho_{\xi\eta}^{+(x)}$ exists and $\rho_{\xi\eta}^{+(x)} < +1$ for $\rho_{yy} \neq 1$

for $0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}}$ if $d \leq -\sqrt{x}$

for $\frac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ if $\sqrt{x} < d < 1$

or if $d \leq -\sqrt{x}$.

Combining results (B.4.6a) and (B.4.6b) produces:

(B.4.6c) $\rho_{\xi\eta}^{+(x)}$ exists and $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$ for $\rho_{yy} \neq 1$

for $0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}}$ if $d < -1$

for $\frac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ if $\sqrt{x} \leq d < 1$

or if $d \leq -\sqrt{x}$.

<u>Summary</u>

If $|d| < \sqrt{x} \equiv \sqrt{\frac{4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}{(1 - \rho_{xx}\rho_{yy})^2}}$ then neither $\rho_{\xi\eta}^{-(x)}$ nor $\rho_{\xi\eta}^{+(x)}$ exist but $Q > 0$ for all $\rho_{\xi\eta} \in (-1, +1)$. [By (B.4.4).]

If $|d| \geq \sqrt{x}$ then

$\rho_{\xi\eta}^{-(x)}$ will exist with $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$, for $\rho_{yy} \neq 1$

for $0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}}$ if $d > +1$

for $\frac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1$ if $d \geq \sqrt{x}$

or if $-1 < d < -\sqrt{x}$ [By (B.4.5c)],

and

$\rho_{\xi\eta}^{+(x)}$ will exist with $\rho_{\xi\eta}^{+(x)} \in (-1, +1)$, for $\rho_{yy} \neq 1$

$$\text{for} \quad 0 < \rho_{xx} \leq \frac{1}{2 - \rho_{yy}} \quad \text{if} \quad d < -1$$

$$\text{for} \quad \frac{1}{2 - \rho_{yy}} < \rho_{xx} \leq 1 \quad \text{if} \quad \sqrt{x} \leq d < 1$$

$$\text{or if} \quad d \leq -\sqrt{x}$$ [By (B.4.6c)].

APPENDIX B.5

RELATIONSHIP OF $\rho_{\xi\eta}^{-(x)}$ TO 1/d WHEN d > 1.

$$\rho_{\xi\eta}^{-(x)} = \frac{d(1-\rho_{xx}\rho_{yy}) - \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})}$$

1. $\rho_{\xi\eta}^{-(x)} \geq 1/d$ when $d > 1$? That is, do there exist values of ρ_{xx}, ρ_{yy} and $d > 1$ such that $\rho_{\xi\eta}^{-(x)} \geq 1/d$?

Let $d > 1$ and $\rho_{yy} \neq 1$.

$$\rho_{\xi\eta}^{-(x)} \geq 1/d \Leftrightarrow \frac{d(1-\rho_{xx}\rho_{yy}) - \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})} \geq 1/d$$

$$\Leftrightarrow d^2(1-\rho_{xx}\rho_{yy}) - d\sqrt{} \geq 2\rho_{xx}(1-\rho_{yy})$$

since $2\rho_{xx}(1-\rho_{yy}) > 0$ for $\rho_{yy} \neq 1$ and $d > 0$ by definition.

$$\Leftrightarrow d^2(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy}) \geq d\sqrt{}$$

$\rho_{\xi\eta}^{-(x)} \geq 1/d$ for $\rho_{yy} \neq 1$ if

I.) $d^2(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy}) \geq 0$

$$\Leftrightarrow \boxed{d^2 \geq \frac{2\rho_{xx}(1-\rho_{yy})}{1-\rho_{xx}\rho_{yy}} = f_x}$$

and II.) $d^2(1-\rho_{xx}\rho_{yy}) - 2\rho_{xx}(1-\rho_{yy}) \geq d\sqrt{}$

$$\Leftrightarrow d^4(1-\rho_{xx}\rho_{yy})^2 + 4\rho_{xx}^2(1-\rho_{yy})^2 - 4d^2\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$\leq d^4(1-\rho_{xx}\rho_{yy})^2 - 4d^2\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})$$

1. II.) continued

$$\Leftrightarrow 4\rho_{xx}^2(1-\rho_{yy})^2 - 4d^2\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy}) + 4d^2\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}) \geq 0$$

$$\Leftrightarrow 4\rho_{xx}(1-\rho_{yy})[\rho_{xx}(1-\rho_{yy}) - d^2(1-\rho_{xx}\rho_{yy}) + d^2(1-\rho_{xx})] \geq 0$$

$$\Leftrightarrow 4\rho_{xx}(1-\rho_{yy})[\rho_{xx}(1-\rho_{yy}) - d^2 + d^2\rho_{xx}\rho_{yy} + d^2 - d^2\rho_{xx}] \geq 0$$

$$\Leftrightarrow 4\rho_{xx}(1-\rho_{yy})[\rho_{xx}(1-\rho_{yy}) - d^2\rho_{xx}(1-\rho_{yy})] \geq 0$$

$$\Leftrightarrow 4\rho_{xx}^2(1-\rho_{yy})^2[1-d^2] \geq 0$$

$$\Leftrightarrow 1-d^2 \geq 0 \quad \text{since} \quad 4\rho_{xx}^2(1-\rho_{yy})^2 > 0 \quad \text{for} \quad \rho_{yy} \neq 1$$

$$\Leftrightarrow 1 \geq d^2.$$

But $d > 1$ therefore $d^2 \not\leq 1$. Therefore $\rho_{\xi\eta}^{-(x)} \not\geq 1/d$.

2) Thus $\rho_{\xi\eta}^{-(x)} < 1/d$ for $d > 1$ for any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$). Since (B.4.5c) indicates that for $d > 1$, $\rho_{\xi\eta}^{-(x)}$ will exist and $\rho_{\xi\eta}^{-(x)} \in (-1, +1)$ for any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$), the algebra results from 1) $\rho_{\xi\eta}^{-(x)} \not\geq 1/d$ guarantee that $\rho_{\xi\eta}^{-(x)} < 1/d$ for any ρ_{xx}, ρ_{yy} ($\rho_{yy} \neq 1$) when $d > 1$.

APPENDIX B.6

Examination of the arithmetic sign of β_x/β_ξ^* and β_y/β_η^*.

Consider expression (3.12) for β_x/β_ξ^*

$$\beta_x/\beta_\xi^* = \frac{(1 - \rho_{\xi\eta}^2)\rho_{xx}(1 - d\rho_{\xi\eta}\rho_{yy})}{(1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy})(1 - d\rho_{\xi\eta})}$$

where $0 < 1 - \rho_{\xi\eta}^2 \leq 1$, since $-1 < \rho_{\xi\eta} < +1$

$0 < 1 - \rho_{\xi\eta}^2\rho_{xx}\rho_{yy} \leq 1$, since $-1 < \rho_{\xi\eta} < +1$,

$0 < \rho_{xx} \leq +1$

$0 < \rho_{yy} \leq +1$.

Thus the arithmetic sign of β_x/β_ξ^* depends solely on the arithmetic sign of

$$\text{(B.6.1)} \qquad \frac{1 - d\rho_{\xi\eta}\rho_{yy}}{1 - d\rho_{\xi\eta}} .$$

That is,

$$\text{(B.6.2a)} \qquad \frac{1 - d\rho_{\xi\eta}\rho_{yy}}{1 - d\rho_{\xi\eta}} > 0 \Leftrightarrow \beta_x/\beta_\xi^* > 0$$

$$\text{(B.6.2b)} \qquad \frac{1 - d\rho_{\xi\eta}\rho_{yy}}{1 - d\rho_{\xi\eta}} = 0 \Leftrightarrow \beta_x/\beta_\xi^* = 0$$

$$\text{(B.6.2c)} \qquad \frac{1 - d\rho_{\xi\eta}\rho_{yy}}{1 - d\rho_{\xi\eta}} < 0 \Leftrightarrow \beta_x/\beta_\xi^* < 0.$$

Consider the denominator of expression (B.6.1) i.e. $1 - d\rho_{\xi\eta}$. When $1 - d\rho_{\xi\eta} > 0$ $(d\rho_{\xi\eta} < 1)$, consider the numerator of (B.6.1).

If $1 - d\rho_{\xi\eta}\rho_{yy} > 0$, (B.6.1) is positive and by (B.6.2a), $\beta_x/\beta_\xi^* > 0$.

If $1 - d\rho_{\xi\eta}\rho_{yy} = 0$, (B.6.1) is zero and by (B.6.2b), $\beta_x/\beta_\xi^* = 0$.

If $1 - d\rho_{\xi\eta}\rho_{yy} \quad 0$, (B.6.1) is negative and by (B.6.2c), $\beta_x/\beta_\xi^* < 0$.

When $1 - d\rho_{\xi\eta} < 0$ $(d\rho_{\xi\eta} > 1)$, consider the numerator of (B.6.1).

If $1 - d\rho_{\xi\eta}\rho_{yy} > 0$, (B.6.1) is negative and by (B.6.2c), $\beta_x/\beta_\xi^* < 0$.

If $1 - d\rho_{\xi\eta}\rho_{yy} = 0$, (B.6.1) is zero and by (B.6.2b), $\beta_x/\beta_\xi^* = 0$.

If $1 - d\rho_{\xi\eta}\rho_{yy} < 0$, (B.6.1) is positive and by (B.6.2a), $\beta_x/\beta_\xi^* > 0$.

When $1 - d\rho_{\xi\eta} = 0$ $(d\rho_{\xi\eta} = 1)$, neither (B.6.1) nor β_x/β_ξ^* are defined.

<u>Note 1</u>:

(B.6.3a) $$1 - d\rho_{\xi\eta}\rho_{yy} > 0 \Leftrightarrow d\rho_{\xi\eta} < \frac{1}{\rho_{yy}}$$

(B.6.3b) $$1 - d\rho_{\xi\eta}\rho_{yy} = 0 \Leftrightarrow d\rho_{\xi\eta} = \frac{1}{\rho_{yy}}$$

(B.6.3c) $$1 - d\rho_{\xi\eta}\rho_{yy} < 0 \Leftrightarrow d\rho_{\xi\eta} > \frac{1}{yy}$$

where $\frac{1}{\rho_{yy}} \geq 1$ since $0 < \rho_{yy} \leq 1$.

<u>Note 2</u>: For an interpretation of expressions (B.6.3a) through (B.6.3c) in terms of ratios of slopes of lines see Appendix B.9.

<u>Note 3</u>: When the numerator of (B.6.1) is zero, then β_x is zero. That is, $1 - d\rho_{\xi\eta}\rho_{yy} = 0 \Rightarrow \beta_x = 0$.

Combining results from above produces:

(B.6.4a) $\beta_x/\beta_\xi^* > 0$ if $d\rho_{\xi\eta} < 1$ or $d\rho_{\xi\eta} > \frac{1}{\rho_{yy}}$

(B.6.4b) $\beta_x/\beta_\xi^* < 0$ if $1 < d\rho_{\xi\eta} < \frac{1}{\rho_{yy}}$

(B.6.4c) β_x/β_ξ^* is undefined if $d\rho_{\xi\eta} = 1$

(B.6.4d) $\beta_x/\beta_\xi^* = 0$ if $d\rho_{\xi\eta} = \frac{1}{\rho_{yy}}$ for $0 < \rho_{yy} < 1$.

In order to compare the distributions of β_x/β_ξ^* and β_y/β_η^* for common situations, it will be useful to have the properties of β_y/β_η^* expressed in the same parameters as the properties of β_x/β_ξ^*. Thus instead of using d_η in the expression of properties of β_y/β_η^*, $1/d$ (where $d = d_\xi$) will be used. Note that $d = d_\xi$ in expressions for β_x/β_ξ^* will be replaced by d_η when the property of interchangeability is applied. But since $d\eta = \frac{1}{d_\xi} = \frac{1}{d}$, d in expressions for β_x/β_ξ^* will be replaced by $1/d$ in comparable expressions for β_y/β_η^*.

Using the property of interchangeability expressions (B.6.4a) - (B.6.4d) for β_x/β_ξ^* become expressions for β_y/β_η^* as follows:

From (B.6.4a)

$$\beta_y/\beta_\eta^* > 0 \text{ if } \frac{\rho_{\xi\eta}}{d} < 1 \text{ or } \frac{\rho_{\xi\eta}}{d} > \frac{1}{\rho_{xx}} .$$

Therefore

(B.6.5a) $\beta_y/\beta_\eta^* > 0$ for $d > 0$ if $\rho_{\xi\eta} < d$ or $\rho_{\xi\eta} > \dfrac{d}{\rho_{xx}}$

[Note: $d \leq \dfrac{d}{\rho_{xx}}$, here.]

for $d < 0$ if $\rho_{\xi\eta} > d$ or $\rho_{\xi\eta} < \dfrac{d}{\rho_{xx}}$

[Note: $d \geq \dfrac{d}{\rho_{xx}}$, here.]

From (B.6.4b)

$$\beta_y/\beta_\eta^* < 0 \quad \text{if} \quad 1 < \frac{\rho_{\xi\eta}}{d} < 1/\rho_{xx} .$$

Therefore

(B.6.5b) $\beta_y/\beta_\eta^* < 0$ for $d > 0$ if $d < \rho_{\xi\eta} < d/\rho_{xx}$

for $d < 0$ if $d/\rho_{xx} < \rho_{\xi\eta} < d$.

From (B.6.4c)

$$\beta_y/\beta_\eta^* \text{ is undefined if } \rho_{\xi\eta}/d = 1.$$

Therefore

(B.6.5c) β_y/β_η^* is undefined if $\rho_{\xi\eta} = d$.

From (B.6.4d)

$$\beta_y/\beta_\eta^* = 0 \quad \text{if} \quad \rho_{\xi\eta}/d = 1/\rho_{xx} \quad \text{for} \quad 0 < \rho_{xx} < 1.$$

Therefore

(B.6.5d) $\beta_y/\beta_\eta^* = 0$ if $\rho_{\xi\eta} = d/\rho_{xx}$ for $0 < \rho_{xx} < 1$.

APPENDIX B.7

COMPARE $1/2 - \rho_{yy}$ TO $2\rho_{yy} - 1/\rho_{yy}$

1.) $\frac{1}{2 - \rho_{yy}} = \frac{2\rho_{yy} - 1}{\rho_{yy}}$?

$$\frac{1}{2 - \rho_{yy}} = \frac{2\rho_{yy} - 1}{\rho_{yy}} \Leftrightarrow \rho_{yy} = (2\rho_{yy} - 1)(2 - \rho_{yy}) \quad \text{Note: } 2-\rho_{yy} > 0 \text{ and } \rho_{yy} > 0.$$

$$\Leftrightarrow \rho_{yy} = 4\rho_{yy} - 2 - 2\rho_{yy}^2 + \rho_{yy}$$

$$\Leftrightarrow 2\rho_{yy}^2 - 4\rho_{yy} + 2 = 0$$

$$\Leftrightarrow 2(\rho_{yy}^2 - 2\rho_{yy} + 1) = 0$$

$$\Leftrightarrow 2(\rho_{yy} - 1)^2 = 0$$

$$\Leftrightarrow \rho_{yy} = 1.$$

Therefore:

(B.7.1) $$\frac{1}{2 - \rho_{yy}} = \frac{2\rho_{yy} - 1}{\rho_{yy}} \quad \text{when} \quad \rho_{yy} = 1 .$$

2.) $\frac{1}{2 - \rho_{yy}} > \frac{2\rho_{yy} - 1}{\rho_{yy}}$?

$$\frac{1}{2 - \rho_{yy}} > \frac{2\rho_{yy} - 1}{\rho_{yy}} \Leftrightarrow \rho_{yy} > (2\rho_{yy} - 1)(2 - \rho_{yy}) \quad \text{since } 2 - \rho_{yy} > 0 \text{ and } \rho_{yy} > 0.$$

Using algebra results from 1) above produces

$$\frac{1}{2 - \rho_{yy}} > \frac{2\rho_{yy} - 1}{\rho_{yy}} \Leftrightarrow 2(\rho_{yy} - 1)^2 > 0 \Leftrightarrow \rho_{yy} < 1.$$

(B.7.2) Therefore $\frac{1}{2 - \rho_{yy}} > \frac{2\rho_{yy} - 1}{\rho_{yy}}$ for $\rho_{yy} < 1$.

3.) $\frac{1}{2 - \rho_{yy}} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$?

$$\frac{1}{2 - \rho_{yy}} < \frac{2\rho_{yy} - 1}{\rho_{yy}} \Leftrightarrow \rho_{yy} < (2\rho_{yy} - 1)(2 - \rho_{yy}).$$

Using algebra results from 1) above produces:

$\frac{1}{2 - \rho_{yy}} < \frac{2\rho_{yy} - 1}{\rho_{yy}} \Leftrightarrow 2(\rho_{yy} - 1)^2 < 0$ which is never true for any values of ρ_{yy}.

(B.7.3) Therefore $\frac{1}{2 - \rho_{yy}} < \frac{2\rho_{yy} - 1}{\rho_{yy}}$ is never true for any ρ_{yy}.

Therefore

(B.7.4) $\frac{1}{2 - \rho_{yy}} \geq \frac{2\rho_{yy} - 1}{\rho_{yy}}$ for all ρ_{yy}, $0 < \rho_{yy} \leq 1$ with equality only if $\rho_{yy} = 1$.

APPENDIX B.8

Comparison of $\rho_{\xi\eta}^{+(x)}$ to d

$$\rho_{\xi\eta}^{+(x)} = \frac{d(1 - \rho_{xx}\rho_{yy}) + \sqrt{d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})}}{2\rho_{xx}(1 - \rho_{yy})} .$$

For the purposes of this appendix $\rho_{\xi\eta}^{+(x)}$ will not be restricted to the range (-1, +1) for the initial algebraic work.

Let $d \geq \sqrt{x}$, $\rho_{yy} \neq 1$.

1) $\rho_{\xi\eta}^{+(x)} = d$?

$$\rho_{\xi\eta}^{+(x)} = d$$

$$\Leftrightarrow \frac{d(1-\rho_{xx}\rho_{yy}) + \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1-\rho_{yy})} = d$$

$$\Leftrightarrow d(1 - \rho_{xx}\rho_{yy}) + \sqrt{} = 2d\rho_{xx}(1 - \rho_{yy})$$

for $\rho_{yy} \neq 1$

$$\Leftrightarrow \sqrt{} = 2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy})$$

$\rho_{\xi\eta}^{+(x)} = d$

if a) $2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}) \geq 0$

$$\Leftrightarrow d[2\rho_{xx}(1 - \rho_{yy}) - (1 - \rho_{xx}\rho_{yy})] \geq 0$$

$$\Leftrightarrow d[2\rho_{xx} - 2\rho_{xx}\rho_{yy} - 1 + \rho_{xx}\rho_{yy}] \geq 0$$

$$\Leftrightarrow d[2\rho_{xx} - \rho_{xx}\rho_{yy} - 1] \geq 0$$

$$\Leftrightarrow d[\rho_{xx}(2 - \rho_{yy}) - 1] \geq 0$$

1) a) (cont'd.)

$$\Leftrightarrow \boxed{\frac{1}{2 - \rho_{yy}} \leq \rho_{xx}} \qquad \text{since} \quad d \geq 0.$$

and b) $d^2(1 - \rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})$

$$= 4d^2\rho_{xx}^2(1 - \rho_{yy})^2$$

$$+ d^2(1 - \rho_{xx}\rho_{yy})^2$$

$$- 4d^2\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx}\rho_{yy})$$

$$\Leftrightarrow 0 = 4d^2\rho_{xx}^2(1-\rho_{yy})^2 - 4d^2\rho_{xx}(1-\rho_{yy})(1-\rho_{xx}\rho_{yy})$$

$$+ 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})[d^2\rho_{xx}(1 - \rho_{yy})$$

$$- d^2(1 - \rho_{xx}\rho_{yy}) + (1 - \rho_{xx})]$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})[d^2\rho_{xx} - d^2\rho_{xx}\rho_{yy}$$

$$- d^2 + d^2\rho_{xx}\rho_{yy} + (1 - \rho_{xx})]$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})[-d^2(1 - \rho_{xx}) + (1 - \rho_{xx})]$$

$$\Leftrightarrow 0 = 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})(1 - d^2)$$

$$\Leftrightarrow \rho_{xx} = 1 \quad \text{or} \quad |d| = 1.$$

Therefore $\rho_{\xi\eta}^{+(x)} = d$ (for $d \geq \sqrt{x}$)

(B.8.1a) if $d = 1$ and $\rho_{xx} \geq \frac{1}{2 - \rho_{yy}}$

(B.8.ab) or if $\rho_{xx} = 1$.

2) $\rho_{\xi\eta}^{+(x)} > d$?

$$\rho_{\xi\eta}^{+(x)} > d$$

$$\Leftrightarrow \frac{d(1-\rho_{xx}\rho_{yy}) + \sqrt{d^2(1-\rho_{xx}\rho_{yy})^2 - 4\rho_{xx}(1-\rho_{yy})(1-\rho_{xx})}}{2\rho_{xx}(1 - \rho_{yy})} > d.$$

Using algebra from 1) above produces:

$$\rho_{\xi\eta}^{+(x)} > d \Leftrightarrow \sqrt{} > 2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}).$$

$\rho_{\xi\eta}^{+(x)} > d$ if a) $2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}) < 0.$

Using algebra from 1) a) above produces:

$$\Leftrightarrow d[\rho_{xx}(2 - \rho_{yy}) - 1] < 0$$

$$\Leftrightarrow \boxed{\rho_{xx} < \frac{1}{2 - \rho_{yy}}} \quad \text{for} \quad d > 0$$

or if b) $\sqrt{} > 2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}) \geq 0$

I) $2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}) \geq 0$

$$\Leftrightarrow \boxed{\rho_{xx} \geq \frac{1}{2 - \rho_{yy}}} \quad \text{for} \quad d \geq 0$$

and II) $\sqrt{} > 2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}).$

Using algebra from 1) b) above produces:

$\Leftrightarrow 0 > 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})(1 - d^2)$

$\Leftrightarrow 0 > 1 - d^2$

$\Leftrightarrow |d| > 1$

$\Leftrightarrow d > 1$ since $d \geq 0$ by definition.

Therefore $\rho_{\xi\eta}^{+(x)} > d$ (for $d \geq \sqrt{x}$, $\rho_{yy} \neq 1$)

(B.8.2a) if $0 < \rho_{xx} < \frac{1}{2 - \rho_{yy}}$

(B.8.2b) or if $\frac{1}{2 - \rho_{yy}} \leq \rho_{xx} \leq 1$ and $d > 1$.

3) $\rho_{\xi\eta}^{+(x)} < d$?

Using algebra from 1) and 2) above produces:

$$\rho_{\xi\eta}^{+(x)} < d \Leftrightarrow \sqrt{\quad\quad} < 2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy})$$

$$\rho_{\xi\eta}^{+(x)} < d \text{ if a) } 2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}) \geq 0$$

$$\Leftrightarrow d[\rho_{xx}(2 - \rho_{yy}) - 1] > 0$$

$$\Leftrightarrow \boxed{\rho_{xx} \geq \frac{1}{2 - \rho_{yy}}} \quad \text{for } d > 0$$

$$\text{and b) } \sqrt{\quad\quad} < 2d\rho_{xx}(1 - \rho_{yy}) - d(1 - \rho_{xx}\rho_{yy}).$$

Using algebra from 1) b) above produces:

$\Leftrightarrow 0 < 4\rho_{xx}(1 - \rho_{yy})(1 - \rho_{xx})(1 - d^2)$

$\Leftrightarrow 0 < 1 - d^2$ for $\rho_{xx} \neq 1$

$\Leftrightarrow \boxed{0 < d < 1}$ for $\rho_{xx} \neq 1$.

Therefore $\rho_{\xi\eta}^{+(x)} < d$ (for $\rho_{xx}, \rho_{yy} \neq 1$, $d \geq \sqrt{x}$)

(B.8.3) if $\frac{1}{2 - \rho_{yy}} \leq \rho_{xx}$ and $0 < d < 1$.

Summary (for $d \geq \sqrt{x}$, $\rho_{yy} \neq 1$)

(B.8.4a) When $0 < \rho_{xx} < \frac{1}{2 - \rho_{yy}}$, then $\rho_{\xi\eta}^{+(x)} > d$ [from (B.8.2a)]

When $\frac{1}{2 - \rho_{yy}} \leq \rho_{xx} < 1$

(B.8.4b) for $0 < d < 1$, then $\rho_{\xi\eta}^{+(x)} < d$ [from (B.8.3)]

(B.8.4c) for $d = 1$, then $\rho_{\xi\eta}^{+(x)} = d$ [from (B.8.1a)]

(B.8.4d) for $d > 1$, then $\rho_{\xi\eta}^{+(x)} > d$ [from (B.8.2b)]

(B.8.4e) When $\rho_{xx} = 1$, then $\rho_{\xi\eta}^{+(x)} = d$ [from (B.8.1b)].

Restricting $\rho_{\xi\eta}^{+(x)}$ to the range (-1, +1), i.e. applying results (B.4.6c) from Appendix B.4, produces adjustments in results (B.8.4a) - (B.8.4d) as follows:

When $d \geq 0$ and $\rho_{yy} \neq 1$

(B.8.5a) for $0 < \rho_{xx} < \frac{1}{2 - \rho_{yy}}$ then $\rho_{\xi\eta}^{+(x)} \notin (-1, +1)$

(B.8.5b) for $\frac{1}{2 - \rho_{yy}} \leq \rho_{xx} < 1$ and $\begin{cases} 0 < d < \sqrt{x} & \text{then } \rho_{\xi\eta}^{+(x)} \notin (-1,+1), \\ \sqrt{x} \leq d < 1 & \text{then } \rho_{\xi\eta}^{+(x)} < d, \\ d \geq 1 & \text{then } \rho_{\xi\eta}^{+(x)} \notin (-1,+1) \end{cases}$

(B.8.5c) for $\rho_{xx} = 1$ and $\begin{cases} 0 < d < \sqrt{x} & \rho_{\xi\eta}^{+(x)} \notin (-1, +1) \\ \sqrt{x} \leq d < 1 & \text{then } \rho_{\xi\eta}^{+(x)} = d \\ d \geq 1 & \text{then } \rho_{\xi\eta}^{+(x)} \notin (-1, +1). \end{cases}$

APPENDIX B.9

An Interpretation of $d\rho_{\xi\eta}$ as a Ratio of Two Slopes

Recall (from 3.6b)

$$d = d_{\xi} = \frac{a_{\eta}/\sigma_{\eta}}{a_{\xi}/\sigma_{\xi}} = \frac{(\mu_{\eta}^{(1)} - \mu_{\eta}^{(0)})/\sigma_{\eta}}{(\mu_{\xi}^{(1)} - \mu_{\xi}^{(0)})/\sigma_{\xi}} .$$

Therefore

$$\text{(B.9.1)} \qquad d\rho_{\xi\eta} = \frac{a_{\eta}/\sigma_{\eta}}{a_{\xi}/\sigma_{\xi}} \cdot \rho_{\xi\eta} = \frac{a_{\eta}}{a_{\xi}} \cdot \rho_{\xi\eta} \frac{\sigma_{\xi}}{\sigma_{\eta}} = \frac{\rho_{\xi\eta}\sigma_{\xi}/\sigma_{\eta}}{a_{\xi}/a_{\eta}} .$$

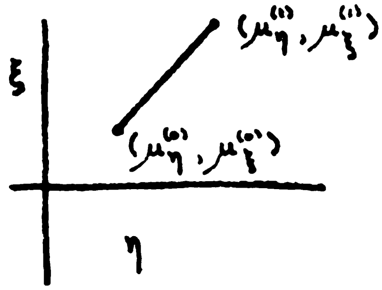
Consider the regression of ξ on η within each category

i.e. $\xi = b^{*(i)}_{\xi\cdot 0} + b^{*}_{\xi\cdot\eta}\eta + \varepsilon, \; i = 0,1$

(B.9.2) where $b^{*}_{\xi\cdot\eta} = \rho_{\xi\eta} \dfrac{\sigma_{\xi}}{\sigma_{\eta}}$ is the regression coefficient, assumed identical in each category,

and $b^{*(i)}_{\xi\cdot 0}$ = the constant in the regression for category i $(i = 0,1)$.

Consider also the midpoint of the bivariate distribution of values of ξ and η within each category i.e. $(\mu_{\eta}^{(i)}, \mu_{\xi}^{(i)})$, $i = 0,1$. For this two category case the line between the midpoints can be portrayed as:



where the slope of the line between the two midpoints can be denoted as m_ξ and is defined:

(B.9.3) $$m_\xi = \frac{\mu_\xi^{(1)} - \mu_\xi^{(0)}}{\mu_\eta^{(1)} - \mu_\eta^{(0)}} = \frac{a_\xi}{a_\eta} .$$

Therefore using (B.9.2) and (B.9.3) in (B.9.1) produces

(B.9.4) $$d\rho_{\xi\eta} = \frac{\rho_{\xi\eta}\sigma_\xi/\sigma_\eta}{a_\xi/a_\eta} = \frac{b^*_{\xi\cdot\eta}}{m_\xi} .$$

Thus $d\rho_{\xi\eta}$ has been expressed as the ratio of two slopes. $b^*_{\xi\cdot\eta}$ is the slope of the pooled <u>within</u> categories regression line $\xi = b^{*(i)}_{\xi\cdot 0} + b^*_{\xi\cdot\eta}\eta + \varepsilon$ $(i = 0,1)$ and m_ξ is the slope of the line <u>between</u> the midpoints of the distributions of η and ξ within the categories.

Note also expression (2.20) for β^*_ξ becomes

$$\beta^*_\xi = \frac{C^*_{11}}{|\Phi|}(b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0}) \quad \text{where} \quad \Phi = \begin{bmatrix} \sigma^2_\xi & \sigma_{\xi\eta} \\ \sigma_{\xi\eta} & \sigma^2_\eta \end{bmatrix}$$

$$|\Phi| = \sigma^2_\xi\sigma^2_\eta(1 - \rho^2_{\xi\eta})$$

$$\text{and} \quad C^*_{11} = \sigma^2_\eta .$$

(B.9.5) $$\beta^*_\xi = \frac{1}{\sigma^2(1 - \rho^2_{\xi\eta})}(b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0}).$$

Similarly, expression (2.15) for $\boldsymbol{\beta}_x$ becomes

(B.9.6) $$\beta_x = \frac{1}{\sigma^2_x(1 - \rho^2_{xy})}(b^{(1)}_{x\cdot 0} - b^{(1)}_{x\cdot 0}) \quad \text{where} \quad x = b^{(i)}_{x\cdot 0} + b_{x\cdot y}y + \varepsilon$$

is the regression of x on y for category i $(i = 0,1)$.

Note 1: $\sigma_\xi^2 = \rho_{xx}\sigma_x^2 \Rightarrow \sigma_\xi^2 < \sigma_x^2$

$$\rho_{\xi\eta} = \frac{\rho_{xy}}{\sqrt{\rho_{xx}\rho_{yy}}} \Rightarrow |\rho_{\xi\eta}| > |\rho_{xy}| \Rightarrow 1 - \rho_{\xi\eta}^2 < 1 - \rho_{xy}^2 .$$

(B.9.7a) Therefore $\sigma_\xi^2(1 - \rho_{\xi\eta}^2) < \sigma_x^2(1 - \rho_{xy}^2)$ i.e. the denominator of expression (B.9.5) for β_ξ^* is less than the denominator of expression (B.9.6) for β_x for all values of σ_ξ^2, $\rho_{\xi\eta}$, ρ_{xx} and ρ_{yy}.

Note 2: $b_{x\cdot y} = \rho_{yy} b_{\xi\cdot\eta}$

(B.9.7b) Therefore, $|b_{x\cdot y}| < |b_{\xi\cdot\eta}^*|$ i.e. the magnitude of the slope of the regression line for the observed predictors will be less than the magnitude of the slope of the regression line for the latent predictors.

Note 3:

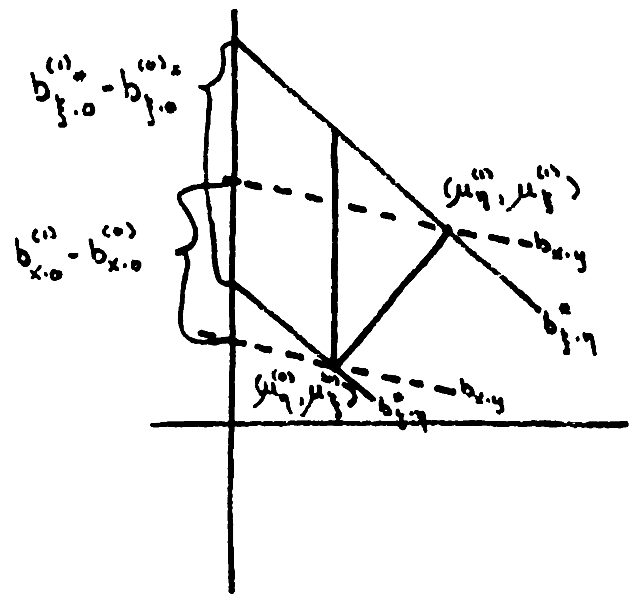
(B.9.7c) $m_x = \frac{a_x}{a_y} = \frac{a_\xi}{a_\eta} = m_\xi$ i.e. the slope of the line joining the midpoints of the distributions of the observed predictors between the two categories is equal to the slope of the line joining the midpoints of the distributions of the latent predictors between the two categories.

Now examine β_x and β_ξ for various combinations of situations of $d\rho_{\xi\eta} = b_{\xi\cdot\eta}^*/m_\xi$ using a pictorial approach.

Let $m_\xi > 0$. Comparable results for $m_\xi < 0$ can be found easily by considering $-b_{\xi\cdot\eta}^*$ in place of $b_{\xi\cdot\eta}^*$.

1) Let $b^*_{\xi\cdot\eta} < 0 \Rightarrow \frac{b^*_{\xi\cdot\eta}}{m_\xi} < 0 \Rightarrow d\rho_{\xi\eta} < 0.$

Here the within groups slope is negative while the between groups slope is positive.

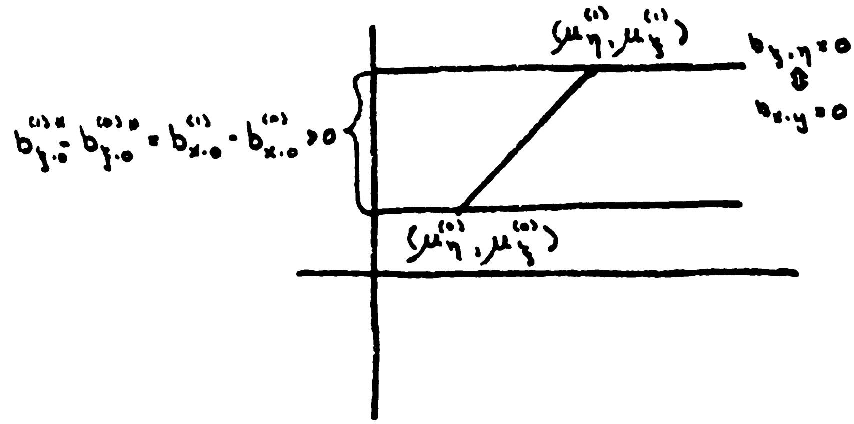


(B.9.8a) Here $0 < b^{(1)}_{x\cdot 0} - b^{(0)}_{x\cdot 0} < b^{(1)}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0}$, and thus $\beta_x > 0$ and $\beta^*_\xi > 0$. Thus the numerator of β^*_ξ is greater than the numerator of β_x and the denominator of β^*_ξ is less than the denominator of β_x.

Hence combining (B.9.8a) and (B.9.7a) produces $\beta^*_\xi > \beta_x > 0$:

(B.9.8b) $0 < \beta_x/\beta^*_\xi < 1$ when $\frac{b^*_{\xi\cdot\eta}}{m_\xi} = d\rho_{\xi\eta} < 0.$

2) Let $b^*_{\xi\cdot\eta} = 0 \Rightarrow \rho_{\xi\eta} = 0 \Rightarrow \frac{b^*_{\xi\cdot\eta}}{m_\xi} = 0 \Rightarrow d\rho_{\xi\eta} = 0$



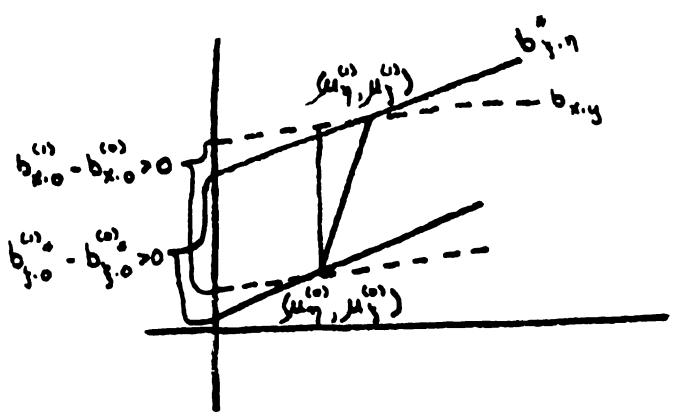
Since $\rho_{\xi\eta} = 0$ both $b^{*}_{\xi\cdot\eta} = 0$ and $b_{x\cdot y} = 0$.

(B.9.9a) Here $b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0} = b^{(1)}_{x\cdot 0} - b^{(0)}_{x\cdot 0} > 0.$

Since the numerators of β^{*}_{ξ} and β^{*}_{η} are equal and positive the ratio β_x/β^{*}_{ξ} formed from (B.9.5) and (B.9.6) when $\rho_{\xi\eta} = 0$ produces

(B.9.9b) $$\beta_x/\beta^{*}_{\xi} = \frac{\sigma^2_{\xi}}{\sigma^2_x} = \rho_{xx} \quad \text{for} \quad b^{*}_{\xi\cdot\eta} = 0 \Rightarrow d\rho_{\xi\eta} = 0 \ .$$

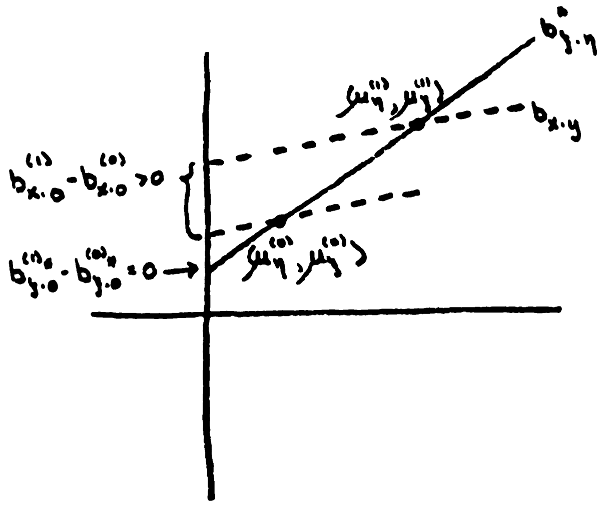
3) Let $0 < b^{*}_{\xi\cdot\eta} < m_{\xi} \Rightarrow 0 < \frac{b^{*}_{\xi\cdot\eta}}{m_{\xi}} < 1 \Leftrightarrow 0 < d\rho_{\xi\eta} < 1$



(B.9.10a) Here $0 < b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0} < b^{(1)}_{x\cdot 0} - b^{(0)}_{x\cdot 0}$, and thus $\beta^{*}_{\xi} > 0$ and $\beta_x > 0$. Thus both the numerator and denominator of β^{*}_{ξ} will be less than their counterparts in β_x.

(B.9.10b) Hence, the relationship of β_x/β^*_ξ to one is not clear when $0 < b^*_{\xi\cdot\eta} < m_\xi$ (i.e. $0 < d\rho_{\xi\eta} < 1$).

4) Let $b^*_{\xi\cdot\eta} = m_\xi \Rightarrow \dfrac{b^*_{\xi\cdot\eta}}{m_\xi} = 1 \Leftrightarrow d\rho_{\xi\eta} = 1$

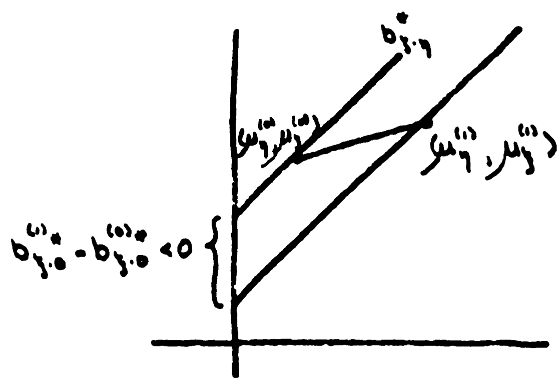


(B.9.11a) Here $b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0} = 0 \Rightarrow \beta^*_\xi = 0$, while $\beta_x > 0$.

(B.9.11b) Thus β_x/β^*_ξ is not defined for $b^*_{\xi\cdot\eta} = m_\xi$ (i.e. $d\rho_{\xi\eta} = 1$) since $\beta^*_\xi = 0$.

5) Let $b^*_{\xi\cdot\eta} > m_\xi \Rightarrow \dfrac{b^*_{\xi\cdot\eta}}{m_\xi} > 1 \Rightarrow d\rho_{\xi\eta} > 1$.

First map out the situation for the latent parameters $b^*_{\xi\cdot\eta}$ and m_ξ. Then examine three subcases for $b_{x\cdot y}$.

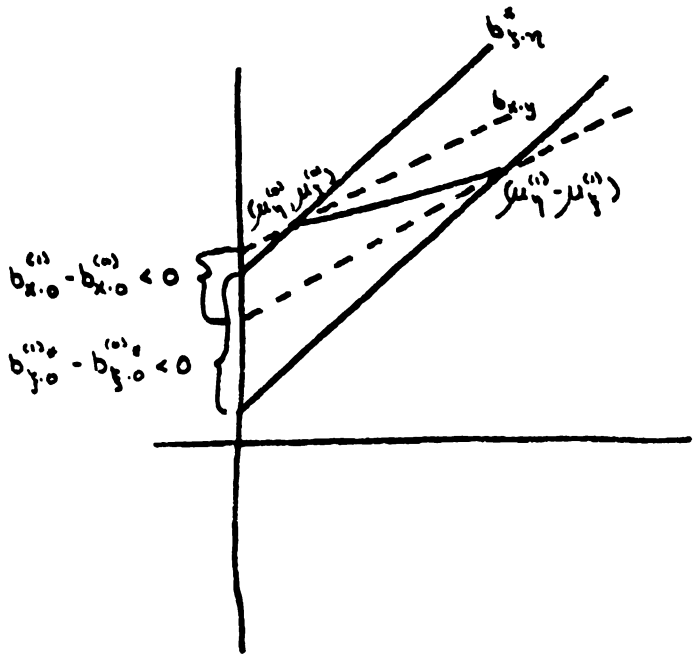


(B.9.12a) Here $b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0} < 0 \Rightarrow \beta^*_\xi < 0$.

<u>Subcase a</u> $b_{x\cdot y} > m_\xi$ as well as $b^*_{\xi\cdot\eta} > m_\xi$

$b_{x\cdot y} = \rho_{yy} b^*_{\xi\cdot\eta}$. Thus $b_{x\cdot y} > m_\xi \Rightarrow \rho_{yy} b^*_{\xi\cdot\eta} > m_\xi$

$\Rightarrow \frac{b^*_{\xi\cdot\eta}}{m_\xi} > \frac{1}{\rho_{yy}}$ (i.e. $d\rho_{\xi\eta} > 1/\rho_{yy}$). Here the slope of the line $x = b^{(i)}_{x\cdot 0} + b_{x\cdot y}\xi + \varepsilon$ for the observed predictors and the slope of the line $\xi = b^{(i)*}_{\xi\cdot 0} + b^*_{\xi\cdot\eta}\eta + \varepsilon$ (i = 0,1) for the latent predictors are both larger than the slope of the between categories line (m_ξ).



(B.9.12b) Here $b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0} < b^{(1)}_{x\cdot 0} - b^{(0)}_{x\cdot 0} < 0$ and $\beta^*_\xi < 0$, $\beta_x < 0$. But $|b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0}| > |b^{(1)}_{x\cdot 0} - b^{(0)}_{x\cdot 0}| > 0$. That is the numerator of β^*_ξ has a greater magnitude than the numerator of β_x.

Since the denominator of β_ξ^* has a smaller magnitude than the denominator of β_x, (B.9.7a) together with (B.9.12b) produce $\beta_\xi^* < \beta_x < 0$ (and $|\beta_\xi^*| > |\beta_x|$).

Therefore

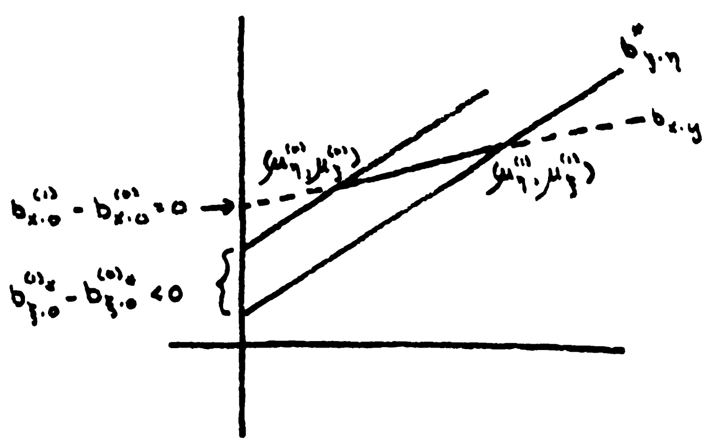
(B.9.12c) $0 < \beta_x/\beta_\xi^* < 1$ for $b_{\xi\cdot\eta}^* > b_{x\cdot y} > m_\xi$ (i.e. $d\rho_{\xi\eta} > 1/\rho_{yy}$).

<u>Subcase b</u> $b_{x\cdot y} = m$

$b_{x\cdot y} = \rho_{yy} b_{\xi\cdot\eta}^*$. Thus $b_{x\cdot y} = m_\xi \Rightarrow \rho_{yy} b_{\xi\cdot\eta}^* = m_\xi$

$\Rightarrow d\rho_{\xi\eta} = \dfrac{1}{\rho_{yy}} \Rightarrow \dfrac{b_{\xi\cdot\eta}^*}{m_\xi} = \dfrac{1}{\rho_{yy}}$.

That is, even though the slope of the regression line for the latent predictors exceeds the between categories slope, the slope of the regression line for the observed predictors equals the between groups slope.



(B.9.12d) Here $b_{\xi\cdot 0}^{(1)*} - b_{\xi\cdot 0}^{(0)*} < 0$, thus $\beta_\xi^* < 0$ and $b_{x\cdot 0}^{(1)} - b_{x\cdot 0}^{(0)} = 0$, thus $\beta_x = 0$.

(B.9.12e) Hence $\beta_x/\beta^*_\xi = 0$ for $b_{x\cdot y} = m_\xi$ (i.e. $d\rho_{\xi\eta} = 1/\rho_{yy}$).

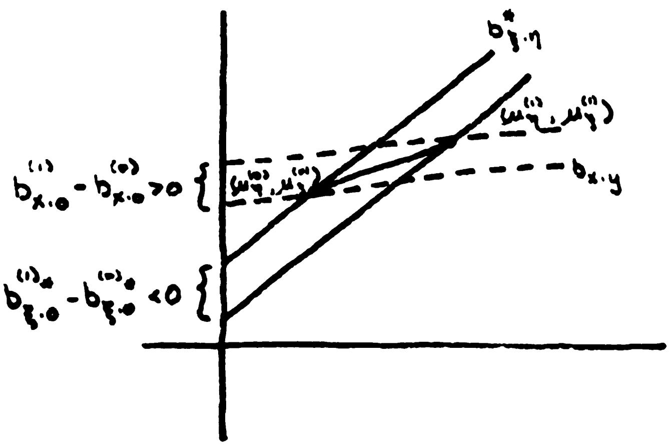
<u>Subcase c</u> $b_{x\cdot y} < m_\xi$ while $b^*_{\xi\cdot\eta} > m_\xi$

$b_{x\cdot y} = \rho_{yy} b^*_{\xi\cdot\eta}$. Thus $b_{x\cdot y} < m_\xi \Rightarrow \rho_{yy} b^*_{\xi\cdot\eta} < m_\xi \Rightarrow \dfrac{b^*_{\xi\cdot\eta}}{m_\xi} < \dfrac{1}{\rho_{yy}}$

but since $b^*_{\xi\cdot\eta} > m_\xi \Rightarrow \dfrac{b^*_{\xi\cdot\eta}}{m_\xi} > 1$, $b_{x\cdot y} < m_\xi \Rightarrow 1 < \dfrac{b^*_{\xi\cdot\eta}}{m_\xi} < \dfrac{1}{\rho_{yy}}$

(i.e. $1 < d\rho_{\xi\eta} < 1/\rho_{yy}$).

That is, the slope of the regression line for the observed predictors is less than the slope of the between categories line, while the slope of the regression line for the latent predictors is greater than the between categories slope.



(B.9.12f) Here $b^{(1)*}_{\xi\cdot 0} - b^{(0)*}_{\xi\cdot 0} < 0 < b^{(1)}_{x\cdot 0} - b^{(0)}_{x\cdot 0}$, thus

$$\beta^*_\xi < 0 \quad \text{and} \quad \beta_x > 0.$$

(B.9.12g) Hence $\beta_x/\beta^*_\xi < 0$ for $0 < b_{x\cdot y} < m_\xi < b^*_{\xi\cdot\eta}$

(i.e. $1 < d\rho_{\xi\eta} < 1/\rho_{yy}$).

APPENDIX C.1

The model with $P = 1$ predictor and $V = K \geq 3$ observed replications is

$$\underset{K\times K}{\Sigma} = \underset{K\times 1}{\Lambda}\ \underset{1\times 1}{\Phi}\ \underset{1\times K}{\Lambda'} + \underset{K\times K}{\Psi^2},$$

that is,

$$\begin{bmatrix} \sigma^2_{x_1} & & \text{Symmetric} & \\ \sigma_{x_2x_1} & \sigma^2_{x_2} & & \\ \vdots & \vdots & \ddots & \\ \sigma_{x_kx_1} & \sigma_{x_kx_2} & \cdots & \sigma^2_{x_k} \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix} \sigma^2_T \begin{bmatrix} 1 & \lambda_2 & \cdots & \lambda_k \end{bmatrix} + \begin{bmatrix} \sigma^2_{e_1} & & & \\ & \sigma^2_{e_2} & & \\ & & \ddots & \\ & & & \sigma^2_{e_k} \end{bmatrix}$$

or

$$\begin{bmatrix} \sigma^2_{x_1} & & \text{Symmetric} & & \\ \sigma_{x_2x_1} & \sigma^2_{x_2} & & & \\ \sigma_{x_3x_1} & \sigma_{x_3x_2} & \sigma^2_{x_3} & & \\ \vdots & \vdots & \vdots & \ddots & \\ \sigma_{x_kx_1} & \sigma_{x_kx_2} & \sigma_{x_kx_3} & \cdots & \sigma^2_{x_k} \end{bmatrix} =$$

$$\begin{bmatrix} \sigma_T^2 + \sigma_{e_1}^2 & & & & \\ \lambda_2 \sigma_T^2 & \lambda_2^2 \sigma_T^2 + \sigma_{e_2}^2 & & & \\ \lambda_3 \sigma_T^2 & \lambda_2 \lambda_3 \sigma_T^2 & \lambda_3^2 \sigma_T^2 + \sigma_{e_3}^2 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \lambda_k \sigma_T^2 & \lambda_2 \lambda_k \sigma_T^2 & \lambda_3\lambda_k\sigma_T^2 & \cdots & \lambda_k^2\sigma_T^2 + \sigma_{e_k} \end{bmatrix}$$

This model for Σ results in $K(K + 1)/2$ equations which relate the parameters in Σ to the $2K$ parameters in the model for Σ, e.g. $\sigma_{x_1}^2 = \sigma_T^2 + \sigma_{e_1}^2$, $\sigma_{x_2x_1} = \lambda_2\sigma_T^2$, etc. These equations can be found by equating the corresponding elements of the two matrices in the equation just above.

Therefore $\dfrac{\sigma_{x_kx_3}}{\sigma_{x_kx_1}} = \dfrac{\lambda_3\lambda_k\sigma_T^2}{\lambda_k\sigma_T^2} = \lambda_3$. In a similar fashion it is possible to see that

$$\frac{\sigma_{x_kx_i}}{\sigma_{x_kx_1}} = \frac{\lambda_i\lambda_k\sigma_T^2}{\lambda_k\sigma_T^2} = \lambda_i \quad \text{for} \quad i = 2,\ldots,k\text{-}1 \quad \text{and} \quad \frac{\sigma_{x_kx_2}}{\sigma_{x_2x_1}} = \frac{\lambda_2\lambda_k\sigma_T^2}{\lambda_2\sigma_T^2} = \lambda_k.$$

Thus expressions for the k-1 parameters of $\underset{\sim}{\Lambda}$, in terms of parameters from Σ, exist.

Consider:

$$\frac{\sigma_{x_2x_1}\ \sigma_{x_kx_1}}{\sigma_{x_kx_2}} = \frac{\lambda_2\sigma_T^2\lambda_k\sigma_T^2}{\lambda_2\lambda_k\sigma_T^2} = \sigma_T^2 .$$

This is one of several possible expressions for the single parameter in Φ.

Finally consider,

$$\sigma^2_{x_1} - \frac{\sigma_{x_2x_1}\,\sigma_{x_kx_1}}{\sigma_{x_kx_2}} = \sigma^2_T + \sigma^2_{e_1} - \sigma^2_T = \sigma^2_{e_1} \;;$$

$$\sigma^2_{x_i} - \frac{\sigma_{x_kx_i}\,\sigma_{x_ix_1}}{\sigma_{x_kx_1}} = \lambda_i^2\sigma_T^2 + \sigma^2_{e_i} - \frac{\lambda_i\lambda_k\sigma_T^2\lambda_i\sigma_T^2}{\lambda_k\sigma_T^2}$$

$$= \lambda_i^2\sigma_T^2 + \sigma^2_{e_i} - \lambda_i^2\sigma_T^2$$

$$= \sigma^2_{e_i} \quad \text{for} \quad i = 2,\ldots,k-1;$$

and

$$\sigma^2_{x_k} - \frac{\sigma_{x_kx_2}\sigma_{x_kx_1}}{\sigma_{x_2x_1}} = \lambda_k^2\sigma_T^2 + \sigma^2_{e_k} - \frac{\lambda_2\lambda_k\sigma_T^2 \cdot \lambda_k\sigma_T^2}{\lambda_2\sigma_T^2}$$

$$= \lambda_k^2\sigma_T^2 + \sigma^2_{e_k} - \lambda_k^2\sigma_T^2$$

$$= \sigma^2_{e_k} \; .$$

Thus expressions for the k parameters of Ψ^2, in terms of parameters from Σ, exist.

Since it is possible to express all $2k$ parameters of the model for Σ in terms of parameters in Σ, the model (4.4) is identified.

APPENDIX C.2

The necessary condition for identifiability will be satisfied when

$$\frac{V(V+1)}{2} \geq V - p + \frac{p(p+1)}{2} + V,$$

where V is the total number of observed replications and p is the number of predictors in the model.

$$\frac{V(V+1)}{2} \geq V - p + \frac{p(p+1)}{2} + V = 2V - p + \frac{p^2 + p}{2}$$

$$\Updownarrow \qquad\qquad = 2V + \frac{p^2}{2} - \frac{p}{2}$$

$$V^2 + V \geq 4V + p^2 - p$$

$$\Updownarrow$$

$$V^2 - 3V \geq p^2 - p \ .$$

Since each predictor must have at least one observed representative, then $V = p + A$ where A is the number of observed replications beyond the original representatives $(A \geq 0)$.

$$V^2 - 3V \geq p^2 - p$$

$$\Updownarrow$$

$$(p + A)^2 - 3(p + A) \geq p^2 - p$$

$$\Updownarrow$$

$$p^2 + 2pA + A^2 - 3p - 3A \geq p^2 - p$$

$$\Updownarrow$$

$$2pA + A^2 - 3p - 3A \geq -p$$

$$\Updownarrow$$

$$A^2 - 3A \geq 2p - 2pA = 2p(1 - A).$$

For $A < 1,\ 1 - A > 0$.

$$\therefore\ \frac{A(A - 3)}{2(1 - A)} \geq p\ .$$

But $A < 1 \Rightarrow A = 0 \Rightarrow \frac{A(A - 3)}{2(1 - A)} = -\frac{3}{2}$.

But $-\frac{3}{2} \geq p$ is impossible since by definition $p \geq 1$,

therefore $A \neq 0$.

For $A = 1,\ A(A - 3) = -2$ and $2p(1 - A) = 0$.

$$\therefore\ A(A - 3) \geq 2p(1 - A)$$

$$\Updownarrow$$

$$-2 \geq 0. \quad \text{Impossible.}$$

Therefore $A \neq 1$.

For $A > 1,\ 1 - A < 0$.

$$\therefore\ \frac{A(A - 3)}{2(1 - A)} \leq p.$$

For $A = 2,\ \frac{A(A - 3)}{2(1 - A)} = \frac{-2}{-2} = 1 \leq p$. This is possible.

For $A = 2,\ \frac{A(A - 3)}{2(1 - A)} = 0 \leq p$. This is possible.

For $A > 3,\ A - 3 > 0$ and $1 - A < 0$, therefore

$$\frac{A(A - 3)}{2(1 - A)} < 0 < p. \quad \text{This is possible.}$$

Therefore the counting condition for identifiability is satisfied when $A \geq 2 \Leftrightarrow V \geq p + 2$, since $V = p + A$. That is, the counting condition for identifiability is satisfied if there are at least two additional observed replications beyond the original set of p observed measurements.

APPENDIX C.3

EXAMINATION OF IDENTIFIABILITY FOR TWO MODELS FOR Σ: A) MODEL (4.5) FROM CHAPTER 4, AND B) MODEL (4.7) FROM CHAPTER 4.

Although each model, (4.5) and (4.7), has a similar external appearance their differences will become apparent upon close examination.

Consider the model with p latent predictors ($p > 1$) where each predictor, T^i, has K_i ($i = 1,2,\ldots,p$) observed replications, and where $V = \sum_{i=1}^{p} K_i$ with V being the total number of observed replications (including all predictors). The model for Σ is:

$$\underset{V\times V}{\Sigma} = \underset{V\times p}{\Lambda} \; \underset{p\times p}{\Phi} \; \underset{p\times V}{\Lambda'} + \underset{V\times V}{\Psi^2}$$

where

$$\underset{V\times p}{\Lambda} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \lambda^1_2 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \lambda^1_{K_1} & 0 & \cdots & 0 \\ \hline 0 & 1 & \cdots & 0 \\ 0 & \lambda^2_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \lambda^2_{K_2} & \cdots & 0 \\ \hline & \vdots & & \\ \hline 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & \lambda^p_2 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda^p_{K_p} \end{bmatrix},$$

$$\underset{p\times p}{\Phi} = \begin{bmatrix} \sigma^2_{T^1} & & \text{Symmetric} \\ \sigma_{T^2T^1} & \sigma^2_{T^2} & \\ \vdots & \vdots & \ddots \\ \sigma_{T^pT^1} & \sigma_{T^pT^2} & \cdots \ \sigma^2_{T^p} \end{bmatrix}$$

and

$$\underset{V\times V}{\Psi^2} = \text{DIAG}\left\{\sigma^2_{E^1_1}\sigma^2_{E^1_2} \cdots \sigma^2_{E^1_{K_1}} \,\Big|\, \sigma^2_{E^2_1}\ \sigma^2_{E^2_2} \cdots \sigma^2_{E^2_{K_2}} \,\Big|\cdots\Big|\, \sigma^2_{E^p_1}\ \sigma^2_{E^p_2} \cdots \sigma^2_{E^p_{K_p}}\right\}.$$

Thus Σ expressed in terms of parameters of the model becomes:

$\underset{V \times V}{\Sigma} =$

	x_2^1	x_2^1	$\cdots$	$x_{K_1}^1$	x_1^2	x_2^2	$\cdots$	$x_{K_2}^2$		x_1^P	x_2^P	$\cdots$	$x_{K_P}^P$
x_1^1	$\sigma^2_{T^1} + \sigma^2_{E_1^1}$							SYMMETRIC					
x_2^1	$\lambda_2^1 \sigma^2_{T^1}$	$(\lambda_2^1)^2 \sigma^2_{T^1} + \sigma^2_{E_2^1}$											
$\vdots$	$\vdots$	$\vdots$											
$x_{K_1}^1$	$\lambda^1_{K_1} \sigma^2_{T^1}$	$\lambda^1_{K_1} \lambda^1_2 \sigma^2_{T^1}$	$\cdots$	$(\lambda^1_{K_1})^2 \sigma^2_{T^1} + \sigma^2_{E^1_{K_1}}$									
x_1^2	$\sigma_{T^2 T^1}$	$\lambda^1_2 \sigma_{T^2 T^1}$	$\cdots$	$\lambda^1_{K_1} \sigma_{T^2 T^1}$	$\sigma^2_{T^2} + \sigma^2_{E^2_1}$								
x_1^2	$\lambda^2_2 \sigma_{T^2 T^1}$	$\lambda^2_2 \lambda^1_2 \sigma_{T^2 T^1}$	$\cdots$	$\lambda^2_2 \lambda^1_{K_1} \sigma_{T^2 T^1}$	$\lambda^2_2 \sigma^2_{T^2}$	$(\lambda^2_2)^2 \sigma^2_{T^2} + \sigma^2_{E^2_2}$	$\ddots$						
$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$							
$x_{K_2}^2$	$\lambda^2_{K_2} \sigma_{T^2 T^1}$	$\lambda^2_{K_2} \lambda^1_2 \sigma_{T^2 T^1}$	$\cdots$	$\lambda^2_{K_1} \lambda^1_{K_1} \sigma_{T^2 T^1}$	$\lambda^2_{K_2} \sigma^2_{T^2}$	$\lambda^2_{K_2} \lambda^2_2 \sigma^2_{T^2}$	$\cdots$	$(\lambda^2_{K_2})^2 \sigma^2_{T^2} + \sigma^2_{E^2_{K_2}}$					
$\vdots$		$\vdots$			$\vdots$				$\ddots$				
x_1^P	$\sigma_{T^P T^1}$	$\lambda^1_2 \sigma_{T^P T^1}$	$\cdots$	$\lambda^1_{K_1} \sigma_{T^P T^1}$	$\sigma_{T^P T^2}$	$\lambda^2_2 \sigma_{T^P T^2}$	$\cdots$	$\lambda^2_{K_2} \sigma_{T^P T^2}$		$\sigma^2_{T^P} + \sigma^2_{E^P_1}$			
x_2^P	$\lambda^P_2 \sigma_{T^P T^1}$	$\lambda^P_2 \lambda^1_2 \sigma_{T^P T^1}$	$\cdots$	$\lambda^P_2 \lambda^1_{K_1} \sigma_{T^P T^1}$	$\lambda^P_2 \sigma_{T^P T^2}$	$\lambda^P_2 \lambda^2_2 \sigma_{T^P T^2}$	$\cdots$	$\lambda^P_2 \lambda^2_{K_2} \sigma_{T^P T^1}$		$\lambda^P_2 \sigma^2_{T^P}$	$(\lambda^P_2)^2 \sigma^2_{T^P} + \sigma^2_{E^P_2}$		
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		
$x_{K_P}^P$	$\lambda^P_{K_P} c_{T^P T^1}$	$\lambda^P_{K_P} \lambda^1_2 \sigma_{T^P T^1}$	$\cdots$	$\lambda^P_{K_P} \lambda^1_{K_1} \sigma_{T^P T^1}$	$\lambda^P_{K_P} \sigma_{T^P T^2}$	$\lambda^P_{K_P} \lambda^2_2 \sigma_{T^P T^2}$	$\cdots$	$\lambda^P_{K_P} \lambda^2_{K_2} \sigma_{T^P T^2}$		$\lambda^P_{K_P} \sigma^2_{T^P}$	$\lambda^P_{K_P} \lambda^P_2 \sigma^2_{T^P}$	$\cdots$	$(\lambda^P_{K_P})^2 \sigma^2_{T^P} + \sigma^2_{E^P_{K_P}}$

In this model Σ contains $\frac{V(V + 1)}{2}$ observed parameters. There are $V - p$ parameters in Λ, $\frac{p(p + 1)}{2}$ parameters in Φ and V parameters in Ψ^2 for a total of $R = 2V + \frac{p(p - 1)}{2}$ parameters in the model for Σ.

PART A:

For model (4.5) it is assumed that $K_i = 1$ for some predictor $i = 1,\ldots,p$ and $K_j \geq 1$ for $j \neq i$, $j = 1,\ldots,J$ such that $V = \sum_{m=1}^{p} K_m \geq p + 2$.

Note: If $p = 2$ and $K_1 = 1$, then the counting condition for identifiability will not be satisfied unless $K_2 \geq 3$.

If $K_2 = 2$ there will be $V = 3$ total observed replications which gives 6 observed parameters in Σ. But there will be 1 parameter in Λ, 3 parameters in Φ and 3 parameters in Ψ^2 to be estimated. Since there are 7 latent parameters in the model for Σ with only 6 observed parameters in Σ, the model is not identified.

Since $V = \sum_{m=1}^{p} K_m \geq p + 2$, the counting condition for identifiability is satisfied. To show that the definition of identifiability is not satisfied, it is sufficient to show that there exists one latent parameter of the model which cannot be expressed as a function of the observed parameters in Σ.

By examination of the expression for Σ in terms of the latent parameters of the model it can be seen that if $K_i = 1$, $\sigma^2_{T^i}$ and $\sigma^2_{E^i_1}$ will each occur in only one location and they will occur together, i.e., $\sigma^2_{X^i_1} = \sigma^2_{T^i} + \sigma^2_{E^i_1}$. Since there is only one equation relating the two unknown latent parameters $\sigma^2_{T^i}$ and $\sigma^2_{E^i_1}$ to observed parameters in Σ there will exist no unique solution for either of the latent parameters separately and therefore the model (4.5) with $K_i = 1$ for some predictor ($i = 1,2,\ldots,p$) is not identified.

Models which have several predictors with only one observed measurement will have the same problem identified above with each expression for the predictor with a single observed measurement and thus will not be identified.

PART B:

For model (4.7) it it assumed that $K_i \geq 2$ for all predictors ($i = 1,2,\ldots,p$). Since $p > 1$, the counting condition for identifiability is satisfied. To show that the definition for identifiability is satisfied, it is necessary to show that each latent parameter can be expressed as a function of observed parameters in Σ.

By observation of expression (C.3.1) for Σ in terms of latent parameters of the model (4.7), the following result is easily obtained.

(C.3.2) $$\sigma_{T^iT^j} = \sigma_{X_1^iX_1^j} \quad \text{for} \quad i,j = 1,2,\ldots,p \quad \text{with} \quad i \neq j.$$

These expressions (C.3.2) solve for the $\frac{p(p-1)}{2}$ off-diagonal parameters of Φ.

(C.3.3) $$\lambda_j^1 = \frac{\sigma_{X_1^2X_j^1}}{\sigma_{X_1^2X_1^1}} \quad \text{for} \quad j = 2,\ldots,K_1$$

(C.3.4) $$\lambda_j^i = \frac{\sigma_{X_j^iX_1^1}}{\sigma_{X_1^iX_1^1}} \quad \text{for} \quad \begin{matrix} i = 2,\ldots,p \\ j = 2,\ldots,K_i \end{matrix} .$$

These expressions (C.3.3) and (C.3.4) solve for the $V - p$ parameters in Λ (since $K_1 - 1 + \sum_{i=2}^{p} (K_i - 1) = \sum_{i=1}^{p} (K_i - 1) = V - p$).

(C.3.5) $\sigma^2_{T^1} = \dfrac{\sigma_{X_2^1 X_1^1}}{\lambda_2^1}$ where λ_2^1 is given by (C.3.3).

(C.3.6) $\sigma^2_{T^i} = \dfrac{\sigma_{X_2^i X_1^i}}{\lambda_2^i}$ for $i = 2,\ldots,p$

where λ_2^i is given by (C.3.4).

These expressions (C.3.5) and (C.3.6) solve for the p diagonal elements of Φ.

(C.3.7) $\sigma^2_{E_1^i} = \sigma^2_{X_1^i} - \sigma^2_{T^i}$ for $i = 1,2,\ldots,p$

where $\sigma^2_{T^i}$ is given by either (C.3.5) or (C.3.6).

(C.3.8) $\sigma^2_{E_j^i} = \sigma^2_{X_j^i} - (\lambda_j^i)^2 \sigma^2_{T^i}$ for $i = 1,\ldots,p$
$j = 2,\ldots,K_i$

where λ_j^i is given either by (C.3.3) or (C.3.4) and $\sigma^2_{T^i}$ is given either by (C.3.5) or (C.3.6).

These expressions (C.3.7) and (c.3.8) solve for the V parameters of Ψ^2 (since $p + \sum_{i=1}^{p} (K_i - 1) = p + V - p = V$).

Thus all $R = \frac{p(p-1)}{2} + V - p + p + V = 2V + \frac{p(p-1)}{2}$ latent parameters in the model for Σ can be expressed as functions of the observed parameters in Σ.

Therefore, when there are at least two observed replications for each predictor when $p > 1$ the model (4.7) for Σ is identified.

CONCLUSION

For models with more than one predictor the model for Σ, given by $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$, will be identified if and only if there are at least two observed replications associated with each latent predictor.

APPENDIX C.4

Consider the model for Σ with p predictors where $K_i = 2$ for some $i = 1,2,\ldots,p$ and $K_j = 1$ for each $j = 1,2,\ldots,p$ where $j \neq i$. Here $V = p + 1$.

The model for Σ is

$$\underset{V\times V}{\Sigma} = \underset{V\times p}{\Lambda} \ \underset{p\times p}{\Phi} \ \underset{p\times V}{\Lambda'} + \underset{V\times V}{\Psi^2}$$

As given the model is not identified. To reduce the number of parameters to be estimated the constraint $\Psi^2 = \sigma_E^2 I$ is introduced, where I is the identity matrix of rank V. Notice that this constraint will not be reasonable for all situations.

Let $\underset{\sim}{X}$ be the $V \times 1$ vector of observed replications for the p predictors. Therefore, $\underset{\sim}{X}' = \left[X_1^1 X_1^2 \ldots X_1^{i-1} \,\vdots\, X_1^i X_2^i \,\vdots\, X_1^{i+1} \ldots X_1^p \right]$.

Thus

$$\Sigma = \Lambda\Phi\Lambda' + \sigma_e^2 I \quad \text{becomes}$$

$$
\Sigma = \begin{array}{c} x_1^1 \\ x_1^2 \\ \vdots \\ x_1^i \\ x_2^i \\ \vdots \\ x_1^p \end{array}
\begin{bmatrix}
\sigma^2_{x_1^1} & & & & & & \\
\sigma_{x_1^2 x_1^1} & \sigma^2_{x_1^2} & & \text{SYMMETRIC} & & & \\
\vdots & \vdots & & & & & \\
\sigma_{x_1^i x_1^1} & \sigma_{x_1^i x_1^2} & \cdots & \sigma^2_{x_1^i} & & & \\
\sigma_{x_2^i x_1^1} & \sigma_{x_2^i x_1^2} & \cdots & \sigma_{x_2^i x_1^i} & \sigma^2_{x_2^i} & & \\
\vdots & \vdots & & \vdots & & & \\
\sigma_{x_1^p x_1^1} & \sigma_{x_1^p x_1^2} & \cdots & \sigma_{x_1^p x_1^i} & \sigma_{x_1^p x_2^i} & \cdots & \sigma^2_{x_1^p}
\end{bmatrix} =
$$

$$
\begin{array}{c} x_1^1 \\ x_1^2 \\ \vdots \\ x_1^i \\ x_2^i \\ \vdots \\ x_1^p \end{array}
\begin{bmatrix}
\sigma^2_{T^1} + \sigma^2_E & & & & & & \\
\sigma_{T^2 T^1} & \sigma^2_{T^2} + \sigma^2_E & & & & & \\
\vdots & \vdots & & & & & \\
\sigma_{T^i T^1} & \sigma_{T^i T^2} & \cdots & \sigma^2_{T^i} + \sigma^2_E & & & \\
\lambda_2^i \sigma_{T^i T^1} & \lambda_2^i \sigma_{T^i T^2} & \cdots & \lambda_2^i \sigma^2_{T^i} & (\lambda_2^i)^2 \sigma^2_{T^i} + \sigma^2_E & & \\
\vdots & \vdots & & \vdots & \vdots & & \\
\sigma_{T^p T^1} & \sigma_{T^p T^2} & \cdots & \sigma_{T^p T^i} & \lambda_2^i \sigma_{T^p T^i} & \cdots & \sigma^2_{T^p} + \lambda^2_E
\end{bmatrix}
$$

First check about the counting condition for identifiability. There are $\frac{(p+1)(p+2)}{2} = \frac{p(p+1)}{2} + p + 1$ observed parameters in Σ. There is one parameter, λ_2^i, in Λ (all other elements are either zero or one), $\frac{p(p+1)}{2}$ parameters in Φ and one parameter in Ψ^2, σ_E^2, for a total of $\frac{p(p+1)}{2} + 2$ parameters in the model for Σ. Since $p > 1$, the counting condition for identifiability is satisfied.

Now check the definition fo ridentifiability.

(C.4.1) $\sigma_{T^kT^j} = \sigma_{X_1^kX_1^j}$ for $k \neq j$ with $k,j = 1,2,\ldots,p$.

This expression (C.4.1) solves for the $\frac{p(p-1)}{2}$ off-diagonal parameters of Φ.

(C.4.2) $\lambda_2^i = \dfrac{\sigma_{X_2^iX_1^1}}{\sigma_{X_1^iX_1^1}}$ for some given i $(i = 1,2,\ldots,p)$.

This expression (C.4.2) solves for the single parameter in Λ.

(C.4.3) $\lambda^2_{T^i} = \dfrac{\sigma_{X_2^iX_1^i}}{\lambda_2^i}$ for some given i $(i = 1,2,\ldots,p)$

where λ_2^i is given by (C.4.2).

This expression (C.4.3) solves for one of the diagonal parameters of Φ.

(C.4.4) $\sigma_E^2 = \sigma^2_{X_1^i} - \sigma^2_{T^i}$ for some given i $(i = 1,2,\ldots,p)$

where $\sigma^2_{T^i}$ is given by (C.4.3).

This expression (C.4.4) solves for the single parameter in Ψ^2.

(C.4.5) $\sigma^2_{T^j} = \sigma^2_{X_1^i} - \sigma_E^2$ for $j \neq i$ with $j = 1,2,\ldots,p$

where σ_E^2 is given by (C.4.4).

This expression (C.4.5) solves for the remaining $p-1$ diagonal elements of Φ.

Therefore all $\frac{p(p+1)}{2} + 2$ parameters in the model for Σ can be expressed as functions of the observed parameters in Σ.

Thus the model (4.11) with the constraint $\Psi^2 = \sigma_E^2 I$ is identified.

APPENDIX C.5

Derivatives of $F = \ln|\Sigma| + \mathrm{tr}\{\Sigma^{-1}S_p\}$, where $\Sigma = \Lambda\Phi\Lambda' + \Psi^2$.

The expressions for vector and matrix derivatives employed in this appendix are taken from Chapter 2, Section 5 of Multivariate Statistical Methods in Behavioral Research by R. Darrell Bock (1975), and are referenced by the chapter, section and statement number used by Bock.

A.) Derivatives of F with respect to elements of $\Lambda(V \times p)$

Recall: Λ is a $V \times p$ matrix of scale factors, where $V = \sum_{i=1}^{p} K_i$ with K_i being the number of observed replications for predictor i $(i = 1,2,\ldots,p)$ i.e.

$$\underset{V\times p}{\Lambda} = \left[\begin{array}{ccccc}
1 & 0 & 0 & \cdots & 0 \\
\lambda_2^1 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
\lambda_{K_1}^1 & 0 & 0 & \cdots & 0 \\
\hdashline
0 & 1 & 0 & \cdots & 0 \\
0 & \lambda_K^2 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
\hdashline
 & & \vdots & & \\
\hdashline
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & \lambda_2^p \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & \lambda_{K_p}^p
\end{array}\right]$$

The work on finding the derivative of F with respect to Λ will first assume that Λ is a general $V \times p$ matrix with elements λ_{ij}, $i = 1,2,\ldots,V$ and $j = 1,\ldots,p$. The derivative desired for the Λ of this research will then be a special case of the general derivative with the necessary adjustments in notation.

$$\frac{\partial F}{\partial \Lambda} = \frac{\partial \ell n|\Sigma|}{\partial \Lambda} + \frac{\partial \, tr\{\Sigma^{-1} S_p\}}{\partial \Lambda}.$$

Consider a) $\dfrac{\partial \; \ell n|\Sigma|}{\partial \Lambda} = \left[\dfrac{\partial \; \ell n|\Sigma|}{\partial \lambda_{ij}}\right]$.

For some element of Λ, λ_{ij},

$$\frac{\partial \; \ell n|\Sigma|}{\partial \Lambda_{ij}} = tr\left[\Sigma^{-1'} \frac{\partial \Sigma'}{\partial \lambda_{ij}}\right] \qquad \text{(Bock 2.5-32)}$$

$$= tr\left[\Sigma^{-1} \frac{\partial \Sigma}{\partial \lambda_{ij}}\right] \quad \text{since } \Sigma \text{ (hence } \Sigma^{-1}) \text{ is symmetric.}$$

$$\frac{\partial \Sigma}{\partial \lambda_{ij}} = \frac{\partial(\Lambda\Phi\Lambda' + \Psi^2)}{\partial \lambda_{ij}} = \frac{\partial(\Lambda\Phi\Lambda')}{\partial \lambda_{ij}} + \underbrace{\frac{\partial(\Psi^2)}{\partial \lambda_{ij}}}_{=0} = \frac{\partial(\Lambda\Phi\Lambda')}{\partial \lambda_{ij}}$$

$$= \Lambda \frac{\partial \Phi\Lambda'}{\partial \lambda_{ij}} + \frac{\partial \Lambda}{\partial \lambda_{ij}} \Phi\Lambda' \qquad \text{(Bock 2.5-3)}$$

$$= \Lambda\left[\Phi \frac{\partial \Lambda'}{\partial \lambda_{ij}} + \underbrace{\frac{\partial \Phi}{\partial \lambda_{ij}}}_{=0} \Lambda'\right] + \frac{\partial \Lambda}{\partial \lambda_{ij}} \Phi\Lambda'$$

$= \Lambda\Phi 1'_{ij} + 1_{ij}\Phi\Lambda'$ where 1_{ij} is the $V \times p$ matrix which has 1 as the ijth element and zero as all other elements.

$= \Lambda\Phi 1'_{ij} + (\Lambda\Phi 1'_{ij})'$ recall Φ is symmetric.

$$\therefore \frac{\partial \ln|\Sigma|}{\partial \lambda_{ij}} = tr\{\Sigma^{-1}[\Lambda\Phi 1'_{ij} + (\Lambda\Phi 1'_{ij})']\}$$

$$= tr\{\Sigma^{-1}\Lambda\Phi 1'_{ij}\} + tr\{\Sigma^{-1}(\Lambda\Phi 1'_{ij})'\}$$

$$= tr\{\Sigma^{-1}\Lambda\Phi 1'_{ij}\} + tr\{(\Lambda\Phi 1'_{ij}\Sigma^{-1})'\}$$

$$= tr\{\Sigma^{-1}\Lambda\Phi 1'_{ij}\} + tr\{\Lambda\Phi 1'_{ij}\Sigma^{-1}\} \quad \text{since} \quad tr\{A\} = tr\{A'\}$$

$$= tr\{\Sigma^{-1}\Lambda\Phi 1'_{ij}\} + tr\{\Sigma^{-1}\Lambda\Phi 1'_{ij}\}$$

$$= 2\ tr\{\Sigma^{-1}\Lambda\Phi 1'_{ij}\}$$

$$= 2(\Sigma^{-1}\Lambda\Phi)_{ij}$$

$$\therefore \frac{\partial \ln|\Sigma|}{\partial \Lambda} = 2\Sigma^{-1}\Lambda\Phi.$$

Now consider

b) $$\frac{\partial\, tr\{\Sigma^{-1}S_p\}}{\partial \Lambda} = -\frac{\partial\, tr\{(\overline{\Sigma^{-1}}S_p\overline{\Sigma^{-1}})\Sigma\}}{\partial \Lambda} \qquad \text{(Bock 2.5-22)}$$

where the bar over Σ^{-1} indicates that Σ^{-1} is "to be regarded as constant for the purposes of differentiation." (Bock, 1975, p. 69)

Let $\overline{\Sigma^{-1}}S_p\overline{\Sigma^{-1}} = C$ where C is considered constant with respect to the differentiation.

$$\therefore \frac{\partial\, tr\{\Sigma^{-1}S_p\}}{\partial \Lambda} = -\frac{\partial\, tr\{C\Sigma\}}{\partial \Lambda} = -\frac{\partial\, tr\{\Sigma C\}}{\partial \Lambda}$$

$$= -\frac{\partial\ tr\{(\Lambda\Phi\Lambda' + \Psi^2)C}{\partial \Lambda} = -\frac{\partial\ tr\{\Lambda\Phi\Lambda' C\}}{\partial \Lambda} - \underbrace{\frac{\partial\ tr\{\Psi^2 C\}}{\partial \Lambda}}_{=\,0}$$

$$= -\frac{\partial\ tr\{\Lambda\Phi\Lambda' C\}}{\partial \Lambda} = -\frac{\partial\ tr\{\overline{\Lambda\Phi}\Lambda' C\}}{\partial \Lambda} - \frac{\partial\ tr\{\Lambda\Phi\overline{\Lambda' C}\}}{\partial \Lambda}\ .$$

Let $\overline{\Lambda\Phi} = D_1$ and $\overline{\Lambda' C} = D_2$ with both D_1 and D_2 constant with respect to the differentiation.

$$= -\frac{\partial\ tr\{D_1\Lambda'C\}}{\partial\Lambda} - \frac{\partial\ tr\{\Lambda\Phi D_2\}}{\partial\Lambda}$$

$$= -\frac{\partial\ tr\{CD_1\Lambda'\}}{\partial\Lambda} - \frac{\partial\ tr\{\Lambda\Phi D_2\}}{\partial\Lambda}$$

$$\frac{\partial\ tr\{\Sigma^{-1}S_p\}}{\partial\Lambda} = -\frac{\partial\ tr\{\Lambda D_1'C'\}}{\partial\Lambda} - \frac{\partial\ tr\{\Lambda\Phi D_2\}}{\partial\Lambda}$$

$$= -(D_1'C')' - (\Phi D_2)' \qquad \text{(Bock 2.5-15)}$$

$$= -CD_1 - D_2'\Phi$$

$$= -\Sigma^{-1}S_p\Sigma^{-1}\Lambda\Phi - \Sigma^{-1}S_p\Sigma^{-1}\Lambda\Phi$$

$$\therefore \frac{\partial\ tr\{\Sigma^{-1}S_p\}}{\partial\Lambda} = -2\Sigma^{-1}S_p\Sigma^{-1}\Lambda\Phi$$

Combining results from part a) and part b):

$$\frac{\partial F}{\partial\Lambda} = \frac{\partial \ln|\Sigma|}{\partial\Lambda} + \frac{\partial\ tr\{\Sigma^{-1}S_p\}}{\partial\Lambda}$$

$$= 2\Sigma^{-1}\Lambda\Phi - 2\Sigma^{-1}S_p\Sigma^{-1}\Lambda\Phi$$

$$= 2\Sigma^{-1}(I - S_p\Sigma^{-1})\Lambda\Phi$$

$$\therefore \qquad \frac{\partial F}{\partial\Lambda} = 2\Sigma^{-1}(\Sigma - S_p)\Sigma^{-1}\Lambda\Phi$$

This result for $\frac{\partial F}{\partial\Lambda}$ is precisely appropriate only if all elements of Λ are latent parameters. Since many of the parameters in Λ of the model for Σ of this research are fixed values, A slight adjustment is needed on the above expression for $\frac{\partial F}{\partial\Lambda}$ to make it suitable to the Λ of this research. The adjustment is described in Chapter 4.

B.) Derivatives of F with respect to elements of $\Phi(p \times p)$

Typically Φ, as used in this research, will contain no fixed values and thus a general expression for $\frac{\partial F}{\partial \Phi}$ will be appropriate. If, however, in some applications Φ does contain one or more fixed values then modifications, suggested in Chapter 4, are needed for $\frac{\partial F}{\partial \Phi}$.

$$\frac{\partial F}{\partial \Phi} = \frac{\partial \ln|\Sigma|}{\partial \Phi} + \frac{\partial \operatorname{tr}\{\Sigma^{-1} S_p\}}{\partial \Phi}$$

Consider a) $\frac{\partial \ln|\Sigma|}{\partial \Phi} = \left[\frac{\partial \ln|\Sigma|}{\partial \Phi_{ij}}\right]$

for some element of Φ, Φ_{ij}

$$\frac{\partial \ln|\Sigma|}{\partial \Phi_{ij}} = \operatorname{tr}\{\Sigma^{-1} \frac{\partial \Sigma}{\partial \Phi_{ij}}\} \qquad \text{(Bock 2.5-32)}$$

$$\frac{\partial \Sigma}{\partial \Phi_{ij}} = \frac{\partial(\Lambda\Phi\Lambda' + \Psi^2)}{\partial \Phi_{ij}} = \frac{\partial(\Lambda\Phi\Lambda')}{\partial \Phi_{ij}} + \underbrace{\frac{\partial(\Psi^2)}{\partial \Phi_{ij}}}_{=0}$$

$$= \Lambda\Phi \underbrace{\frac{\partial \Lambda'}{\partial \Phi_{ij}}}_{=0} + \frac{\partial(\Lambda\Phi)}{\partial \Phi_{ij}} \Lambda' = \left[\Lambda \frac{\partial \Phi}{\partial \Phi_{ij}} + \underbrace{\frac{\partial \Lambda}{\partial \Phi_{ij}}}_{=0} \Phi\right]\Lambda'$$

$\therefore \frac{\partial \Sigma}{\partial \Phi_{ij}} = \Lambda 1_{ij}^{ji} \Lambda'$ where 1_{ij}^{ji} is a $p \times p$ matrix where elements (ij) and (ji) are equal to one and all other elements are equal to zero.

(Note: $\frac{\partial \Phi}{\partial \Phi_{ij}} = 1_{ij}^{ji}$ since Φ is symmetric.)

$$\therefore \frac{\partial \ln|\Sigma|}{\partial \Phi_{ij}} = \operatorname{tr}\{\Sigma^{-1}\Lambda 1_{ij}^{ji}\Lambda'\} = \operatorname{tr}\{\Lambda'\Sigma^{-1}\Lambda\ 1_{ij}^{ji'}\}$$

$$\therefore \frac{\partial \ln|\Sigma|}{\partial \Phi_{ij}} = \begin{cases} (\Lambda'\Sigma^{-1}\Lambda)_{ij} & \text{for } i = j,\ i,j = 1,2,\ldots,p \\ 2(\Lambda'\Sigma^{-1}\Lambda)_{ij} & \text{for } i \neq j,\ i,j = 1,2,\ldots,p \end{cases}$$

$\therefore \frac{\partial \ell n |\Sigma|}{\partial \Phi} = 2\Lambda'\Sigma^{-1}\Lambda - \text{DIAG}\{\Lambda'\Sigma^{-1}\Lambda\}$

Now consider b) $\frac{\partial \; tr\{\Sigma^{-1}S_p\}}{\partial \Phi}$

Before proceeding with this section it is necessary to establish a result which will be needed and is not given by Bock.

If X is an $r \times r$ symmetric matrix of variables and C is an $r \times r$ matrix of constants then

$$\frac{\partial \; tr\{XC\}}{\partial X} = C' + C - \text{DIAG}\{C\}$$

and if C is also symmetric

$$\frac{\partial \; tr\{XC\}}{\partial X} = 2C - \text{DIAG}\{C\}.$$

Proof:

By definition $tr\{XC\} = \sum_{k=1}^{r} [XC]_{kk}$ where $[XC]_{kk}$ is the k^{th} diagonal element of the matrix product XC.

$[XC]_{kk} = \sum_{\ell=1}^{r} X_{k\ell}C_{\ell k}$ by definition of matrix multiplication

$$\therefore tr\{SC\} = \sum_{k=1}^{r} \sum_{\ell=1}^{r} X_{k\ell}C_{\ell k}$$

For $i \neq j$, $i,j = 1,2,\ldots,r$

$$\frac{\partial \; tr\{XC\}}{\partial X_{ij}} = \frac{\partial}{\partial X_{ij}} \sum_{k=1}^{r} \sum_{\ell=1}^{r} X_{k\ell}C_{\ell k} = \frac{\partial}{\partial X_{ij}} (X_{ij}C_{ji} + X_{ji}C_{ij})$$

Note: $\frac{\partial}{\partial X_{ij}} (X_{k\ell}C_{\ell k}) = 0$ unless $k = i$ and $\ell = j$

or $k = j$ and $\ell = i$

$$= \frac{\partial (X_{ij}C_{ji})}{\partial X_{ij}} + \frac{\partial (X_{ji}C_{ij})}{\partial X_{ij}}.$$

But since X is symmetric $X_{ij} = X_{ji}$.

$$\therefore \frac{\partial \, tr\{XC\}}{\partial X_{ij}} = C_{ji} + C_{ij} \quad \text{for } i \neq j,\ i,j = 1,2,\ldots,r$$

and, for $i = j$, $i,j = 1,2,\ldots,r$

$$\frac{\partial \, tr\{XC\}}{\partial X_{ij}} = \frac{\partial}{\partial X_{ij}} \sum_{k=1}^{r} \sum_{\ell=1}^{r} X_{k\ell} C_{\ell k} = \frac{\partial}{X_{ij}} (X_{ii}C_{ii}) = \frac{\partial}{X_{ii}} (X_{ii}C_{ii})$$

Note: $\frac{\partial}{\partial X_{ii}} (X_{k\ell}C_{\ell k}) = 0$ unless $k = i$ <u>and</u> $\ell = i$.

$$\therefore \frac{\partial \, tr\{XC\}}{\partial X_{ij}} = C_{ii} \quad \text{for } i = j,\ i,j = 1,2,\ldots,r.$$

$$\therefore \frac{\partial \, tr\{XC\}}{\partial X} = \left[\frac{\partial \, tr\{XC\}}{\partial X_{ij}} \right] = C' + C - DIAG\{C\}.$$

If C is symmetric, then $C' = C$ and

$$\frac{\partial \, tr\{XC\}}{\partial X} = 2C - DIAG\{C\}.$$

Now $$\frac{\partial \, tr\{\Sigma^{-1} S_p\}}{\partial \Phi} = - \frac{\partial \, tr\{(\overline{\Sigma}^{-1} S_p \overline{\Sigma}^{-1})\Sigma\}}{\partial \Phi} \qquad \text{(Bock 2.5-32)}$$

Let $C = \overline{\Sigma}^{-1} S_p \overline{\Sigma}^{-1}$ as in Section A) part b.) above.

$$\frac{\partial \, tr\{\Sigma^{-1} S_p\}}{\partial \Phi} = - \frac{\partial tr\{C\Sigma\}}{\partial \Phi} = - \frac{\partial \, tr\{\Sigma C\}}{\partial \Phi}$$

$$= - \frac{\partial \, tr\{(\Lambda\Phi\Lambda' + \Psi^2)C\}}{\partial \Phi}$$

$$= - \frac{\partial \, tr\{\Lambda\Phi\Lambda' C\}}{\partial \Phi} - \underbrace{\frac{\partial \, tr\{\Psi^2 C\}}{\partial \Phi}}_{=\,0}$$

$$= - \frac{\partial \, tr\{\Phi\Lambda' C\Lambda\}}{\partial \Phi}.$$

Since Φ is a symmetric matrix of parameters and $\Lambda'C\Lambda$ is a symmetric matrix of constants (with respect to the differentiation) the result proved above applies.

$$\therefore \frac{\partial\ tr\{\Sigma^{-1}S_p\}}{\partial\Phi} = -[2(\Lambda'C\Lambda) - DIAG\{\Lambda'C\Lambda\}]$$

$$= -2\Lambda'\Sigma^{-1}S_p\Sigma^{-1}\Lambda + DIAG\{\Lambda'\Sigma^{-1}S_p\Sigma^{-1}\Lambda\} \quad .$$

Therefore combining results from part a) and part b)

$$\frac{\partial F}{\partial\Phi} = \frac{\partial \ell n|\Sigma|}{\partial\Phi} + \frac{\partial tr\{\Sigma^{-1}S_p\}}{\partial\Phi}$$

$$= 2\Lambda'\Sigma^{-1}\Lambda - DIAG\{\Lambda'\Sigma^{-1}\Lambda\} - 2\Lambda'\Sigma^{-1}S_p\Sigma^{-1}\Lambda + DIAG\{\Lambda'\Sigma^{-1}S_p\Sigma^{-1}\Lambda\}$$

$$= 2(\Lambda'\Sigma^{-1}\Lambda - \Lambda'\Sigma^{-1}S_p\Sigma^{-1}\Lambda) - DIAG\{\Lambda'\Sigma^{-1}\Lambda - \Lambda'\Sigma^{-1}S_p\Sigma^{-1}\Lambda\}$$

$$= 2(\Lambda'\Sigma^{-1}(I - S_p\Sigma^{-1})\Lambda) - DIAG\{\Lambda'\Sigma^{-1}(I - S_p\Sigma^{-1})\Lambda\}$$

$$\therefore \frac{\partial F}{\partial\Phi} = 2\Lambda'\Sigma^{-1}(\Sigma - S_p)\Sigma^{-1}\Lambda - DIAG\{\Lambda'\Sigma^{-1}(\Sigma - S_p)\Sigma^{-1}\Lambda\}.$$

C) Derivatives of F with respect to elements of $\Psi(V \times V)$.

Since Ψ^2, hence Ψ, will be a diagonal matrix for virtually all applications of the model of this research, the derivations, $\frac{\partial F}{\partial\Psi}$, will be found with respect to the diagonal elements alone.

$$\frac{\partial F}{\partial\Psi_{ii}} = \frac{\partial\ \ell n|\Sigma|}{\partial\Psi_{ii}} + \frac{\partial\ tr\{\Sigma^{-1}S_p\}}{\partial\Psi_{ii}} \quad \text{for} \quad i = 1,2,\ldots,V\ .$$

Note the use of Ψ_{ii} rather than Ψ^2_{ii} here.

Consider a) $\frac{\partial\ \ell n|\Sigma|}{\partial\Psi_{ii}} = tr\{\Sigma^{-1}\ \frac{\partial\Sigma}{\partial\Psi_{ii}}\}$

$$\frac{\partial\Sigma}{\partial\Psi_{ii}} = \frac{\partial(\Lambda\Phi\Lambda' + \Psi^2)}{\partial\Psi_{ii}} = \underbrace{\frac{\partial(\Lambda\Phi\Lambda')}{\partial\Psi_{ii}}}_{0} = \frac{\partial(\Psi^2)}{\partial\Psi_{ii}}$$

Note: $\Psi^2 = \Psi$ and since Ψ^2 is diagonal

$$(\Psi^2)_{ii} = [(\Psi)_{ii}]^2$$

$$\therefore \quad \frac{\partial \Sigma}{\partial \Psi_{ii}} = \frac{\partial (\Psi^2)}{\partial \Psi_{ii}} = 2\Psi_{ii} 1_{ii}$$

$$\therefore \quad \frac{\partial \, \ell n|\Sigma|}{\partial \Psi_{ii}} = tr\{\Sigma^{-1} 2\Psi_{ii} 1_{ii}\} = 2\Psi_{ii} tr\{\Sigma^{-1} 1_{ii}\} = 2\Psi_{ii} \Sigma^{-1}_{ii}$$

$= 2\Sigma^{-1}_{ii}\Psi_{ii}$ where Σ^{-1}_{ii} and Ψ_{ii} are the ith diagonal elements of Σ^{-1} and Ψ respectively.

And since Ψ is diagonal

$$\frac{\partial \, \ell n|\Sigma|}{\partial \Psi_{ii}} = 2[\Sigma^{-1}\Psi]_{ii} \qquad i = 1,2,\ldots,V$$

where $[\Sigma^{-1}\Psi]_{ii}$ is the ith diagonal elements of of the matric product $\Sigma^{-1}\Psi$.

Consider now b)

$$\frac{\partial \, tr\{\Sigma^{-1} S_p\}}{\partial \Psi_{ii}} = - \frac{\partial tr\{(\overline{\Sigma}^{-1} S_p \overline{\Sigma}^{-1})\Sigma\}}{\partial \Psi_{ii}}$$ (Bock 2.5-22 with respect to a single element of Ψ)

Again let $C = \overline{\Sigma}^{-1} S_p \overline{\Sigma}^{-1}$.

$$\frac{\partial \, tr\{\Sigma^{-1} S_p\}}{\partial \Psi_{ii}} = - \frac{\partial \, tr\{C\Sigma\}}{\partial \Psi_{ii}} = - \frac{\partial \, tr\{\Sigma C\}}{\partial \Psi_{ii}} = - \frac{\partial tr\{(\Lambda\Phi\Lambda' + \Psi^2)C\}}{\partial \Psi_{ii}}$$

$$= \underbrace{\frac{\partial \, tr\{\Lambda\Phi\Lambda' C\}}{\partial \Psi_{ii}}}_{=\,0} - \frac{\partial \, tr\{\Psi^2 C\}}{\partial \Psi_{ii}}$$

$$= - \frac{\partial\ tr\{\Psi\Psi C\}}{\partial\Psi_{ii}} = - \frac{\partial\ tr\{\overline{\Psi}\Psi C\}}{\partial\Psi_{ii}} - \frac{\partial\ tr\{\Psi\overline{\Psi}C\}}{\partial\Psi_{ii}}$$ (Bock 2.5-20 with respect to a single element of Ψ)

$$= - \frac{\partial\ tr\{\Psi C\overline{\Psi}\}}{\partial\Psi_{ii}} - \frac{\partial\ tr\{\Psi\overline{\Psi C}\}}{\partial\Psi_{ii}} .$$

By a result gained in the process of the proof in section B)

i.e. $\frac{\partial\ tr\{XC\}}{X_{ii}} = C_{ii}$,

$$\frac{\partial\ tr\{\Psi C\overline{\Psi}\}}{\partial\Psi_{ii}} = (C\Psi)_{ii}$$

and

$$\frac{\partial\ tr\{\Psi\overline{\Psi C}\}}{\partial\Psi_{ii}} = (\Psi C)_{ii}$$

$$\therefore \frac{\partial\ tr\{\Sigma^{-1}S_p\}}{\partial\Psi_{ii}} = -(C\Psi)_{ii} - (\Psi C)_{ii}$$

But since $C = \Sigma^{-1}S_p\Sigma^{-1}$ is symmetric and Ψ is diagonal, $C\Psi$ is symmetric and $C\Psi = \Psi C$.

$$\therefore \frac{\partial tr\{\Sigma^{-1}S_p\}}{\partial\Psi_{ii}} = -2(C\Psi)_{ii} = -2(\Sigma^{-1}S_p\Sigma^{-1}\Psi)_{ii}$$

Therefore combining part a) and b)

$$\frac{\partial F}{\partial\Psi_{ii}} = \frac{\partial\ \ell n|\Sigma|}{\partial\Psi_{ii}} + \frac{\partial\ tr\{\Sigma^{-1}S_p\}}{\partial\Psi_{ii}} \quad \text{for} \quad i = 1,2,\ldots,V$$

$$= 2(\Sigma^{-1}\Psi)_{ii} - 2(\Sigma^{-1}S_p\Sigma^{-1}\Psi)_{ii}$$

$$= 2(\Sigma^{-1}\Psi - \Sigma^{-1}S_p\Sigma^{-1}\Psi)_{ii}$$

$$\therefore \frac{\partial F}{\partial\Psi_{ii}} = 2[\Sigma^{-1}(\Sigma - S_p)\Sigma^{-1}\Psi]_{ii} \quad \text{for} \quad i = 1,2,\ldots,V.$$

LIST OF REFERENCES

LIST OF REFERENCES

Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis, John Wiley and Sons.

Assakul, K. and Proctor, C.H. (1967). "Testing Independence in Two-Way Contingency Tables with Data Subject to Misclassification," Psychometrika, 32, 67-76.

Bock, R.D. (1975). Multivariate Statistical Methods in Behavioral Research, McGraw Hill.

Bross, I.D.J. (1954). "Misclassification in 2 x 2 Tables," Biometrica, 10, 478-486.

Cochran, W.G. (1970). "Some Effects of Errors of Measurement on Multiple Correlation," Journal of the American Statistical Association, 65, No. 329, Applications Section, 22-34.

Cochran, W.G. (1968). "Errors of Measurement in Statistics," Technometrics, 10, No. 4, 637-666.

Cornfield, J., Gordon, T. and Smith, W.W. (1960). "Quantal Response Curves for Experimentally Uncontrolled Variables," International Statistical Institute - Bulletin, 38, Part 3, 97-115.

Davis, P.J. (1965). The Mathematics of Matrices, Ginn and Co.

Fletcher, R. and Powell, M.J.D. (1963). "A Rapidly Convergent Descent Method for Minimization," Computer Journal, 6, 163-168.

Jöreskog, K.G. (1970). "A General Method for Analysis of Covariance Structures," Biometrika, 57, 239-251.

Jöreskog, K.B., van Thillo, M. and Gruvaeus, G. (1975). ACOVSM: A General Computer Program for Analyais of Covariance Structures Including Generalized Manova - CDC 6500 Conversion by Judy Pfaff at Michigan State University.

Lord, F.M. and Novick, M.R. (1974). Statistical Theories of Mental Test Scores, Addison-Wesley.

McSweeney and Schmidt (1974). "Quantal Response Techniques for Random Predictor Variables," Paper presented at AERA annual convention.

Mote, V.L. and Anderson, R.L. (1965). "An Investigation of the Effect of Misclassification on the Properties of X^2-tests in the Analysis of Categorical Data," Biometrika, 52, 95-109.

Porter, A.C. (1971). "How Errors of Measurement Affect ANOVA, Regression Analyses, ANCOVA and Factor Analyses," Occasional Paper No. 14; Office of Research Consultation, School for Advanced Studies, College of Education, Michigan State University.

Scheifley, V. (1972). GENDATA.

Tatsuoka, M.M. (1974). Classification Procedures: Profile Similarity, Institute for Personality and Ability Testing, Champaign, Illinois.

Wiley, D.E. and Hornik, R. (1973). "Measurement Error and the Analysis of Panel Data," Mehr Licht! Studies of Educative Processes, Report No. 5.

Wiley, D.E., Schmidt, W.H. and Bramble, W.J. (1973). "Studies of a Class of Covariance Structure Models," Journal of the American Statistical Association, 68, No. 342, 317-323.