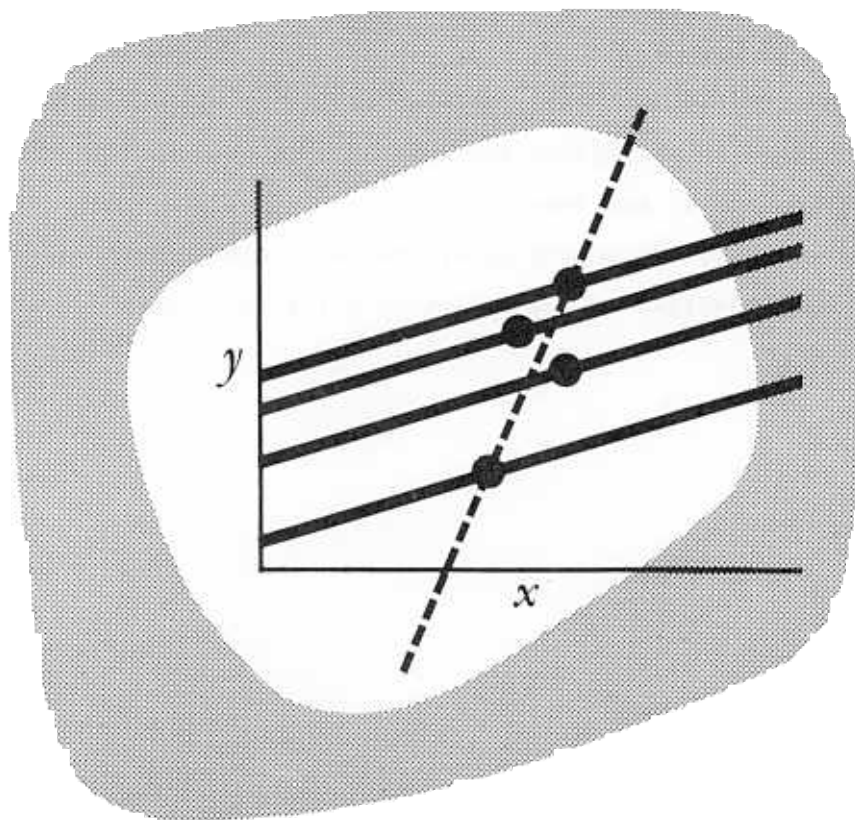


REX FORTRAN 4 SYSTEM

for combinatorial screening or
conventional analysis of multivariate regressions



NOTICE
=====

THIS COMPUTER-PRODUCED PUBLICATION IS AN EXPERIMENTAL EFFORT TO PUBLISH (MORE RAPIDLY AND MORE EFFICIENTLY) INFORMATION ON COMPUTER-ORIENTED THEORIES AND TECHNIQUES.

AT THE SAME TIME, WE ARE TRYING TO IMPROVE SUSCEPTIBILITY OF THE INFORMATION TO AUTOMATED SEARCH AND RETRIEVAL. THE INITIAL SUMMARY AND THE ENTIRE TEXT OF THE PAPER ARE IMMEDIATELY SUITABLE FOR COMPUTER SEARCH BY VIRTUE OF ALREADY BEING ON PUNCHED CARDS. AN IDENTIFIER AT THE TOP OF EACH PAGE SERVES TO MATCH IT WITH ITS PARENT DOCUMENT IN CASE OF SEPARATION OR MIXUPS.

FINALLY, COMPUTER-PROCESSED TEXT IS EASILY REVISED AND REPUBLISHED. THIS IS AN IMPORTANT CONSIDERATION IN FIELDS WHERE CHANGES AND NEW DEVELOPMENTS ARE OCCURRING SO RAPIDLY.

THE COMPUTER PROGRAM 'PRN', WRITTEN BY THE AUTHOR IN FORTRAN-4 LANGUAGE, WAS USED TO PRINT THIS RESEARCH PAPER AS WELL AS EARLIER U.S. FOREST SERVICE RESEARCH PAPERS PSW-13 AND PSW-21.

PROGRAM LISTING AND SOURCE DECKS FOR 'REX' CAN BE MADE AVAILABLE TO INTERESTED ORGANIZATIONS HAVING ACCESS TO A SUITABLE LARGE COMPUTER.

Grosenbaugh, L. R.

1967. *REX--Fortran-4 system for combinatorial screening or conventional analysis of multivariate regressions.* Berkeley, Calif., Pacific SW. Forest & Range Exp. Sta. 47 pp., illus (U.S. Forest Serv. Res. Paper PSW-44)

Describes an expansible computerized system that provides data needed in regression or covariance analysis of as many as 50 variables, 8 of which may be dependent. Alternatively, it can screen variously generated combinations of independent variables to find the regression with the smallest mean-squared-residual, which will be fitted if desired. The user can easily program additional processing.

OXFORD: U.519.1+U.519.27+U.681.3

RETRIEVAL TERMS: multivariate regression analysis; multivariate covariance analysis; combinatorial regression screening; digital computer program.

Grosenbaugh, L. R.

1967. *REX--Fortran-4 system for combinatorial screening or conventional analysis of multivariate regressions.* Berkeley, Calif., Pacific SW. Forest & Range Exp. Sta. 47 pp., illus (U.S. Forest Serv. Res. Paper PSW-44)

Describes an expansible computerized system that provides data needed in regression or covariance analysis of as many as 50 variables, 8 of which may be dependent. Alternatively, it can screen variously generated combinations of independent variables to find the regression with the smallest mean-squared-residual, which will be fitted if desired. The user can easily program additional processing.

OXFORD: U.519.1+U.519.27+U.681.3

RETRIEVAL TERMS: multivariate regression analysis; multivariate covariance analysis; combinatorial regression screening; digital computer program.

Grosenbaugh, L. R.

1967. *REX--Fortran-4 system for combinatorial screening or conventional analysis of multivariate regressions.* Berkeley, Calif., Pacific SW. Forest & Range Exp. Sta. 47 pp., illus (U.S. Forest Serv. Res. Paper PSW-44)

Describes an expansible computerized system that provides data needed in regression or covariance analysis of as many as 50 variables, 8 of which may be dependent. Alternatively, it can screen variously generated combinations of independent variables to find the regression with the smallest mean-squared-residual, which will be fitted if desired. The user can easily program additional processing.

OXFORD: U.519.1+U.519.27+U.681.3

RETRIEVAL TERMS: multivariate regression analysis; multivariate covariance analysis; combinatorial regression screening; digital computer program.

Grosenbaugh, L. R.

1967. *REX--Fortran-4 system for combinatorial screening or conventional analysis of multivariate regressions.* Berkeley, Calif., Pacific SW. Forest & Range Exp. Sta. 47 pp., illus (U.S. Forest Serv. Res. Paper PSW-44)

Describes an expansible computerized system that provides data needed in regression or covariance analysis of as many as 50 variables, 8 of which may be dependent. Alternatively, it can screen variously generated combinations of independent variables to find the regression with the smallest mean-squared-residual, which will be fitted if desired. The user can easily program additional processing.

OXFORD: U.519.1+U.519.27+U.681.3

RETRIEVAL TERMS: multivariate regression analysis; multivariate covariance analysis; combinatorial regression screening; digital computer program.

REX--FORTRAN-4 SYSTEM FOR COMBINATORIAL SCREENING OR
CONVENTIONAL ANALYSIS OF MULTIVARIATE REGRESSIONS

BY

L. R. GROSENBAUGH

CONTENTS

	PAGE
SUMMARY-----	1
REGRESSION FITTING, SCREENING, AND ANALYSIS-----	2
COVARIANCE ANALYSIS-----	9
BACKGROUND AND DEVELOPMENT OF SYSTEM-----	12
AIDS FOR USERS MODIFYING OR EXPANDING SYSTEM-----	15
LITERATURE CITED-----	16
APPENDIX A (DECK ARRANGEMENTS AND MODS)-----	17
APPENDIX B (EXAMPLES OF TEST INPUT DATA)-----	23
APPENDIX C (ILLUSTRATIVE OUTPUT)-----	39

+-----THE AUTHOR-----+

I
I L. R. GROSENBAUGH JOINED THE U. S. FOREST SERVICE I
I IN 1936 AFTER RECEIVING HIS MASTER OF FORESTRY I
I DEGREE FROM YALE UNIVERSITY. HE SPENT 25 YEARS I
I WITH THE SOUTHERN REGION AND THE SOUTHERN FOREST I
I EXPERIMENT STATION, AND IN 1961 STARTED THE FOREST I
I SERVICE'S FIRST PIONEERING RESEARCH UNIT (IN I
I FOREST MENSURATION) AT THE PACIFIC SOUTHWEST I
I FOREST AND RANGE EXPERIMENT STATION. I

+-----+

U.S. FOREST SERVICE RESEARCH PAPER PSW-44. (ORIGINAL VERSION DATED 5-01-67)
PACIFIC SOUTHWEST FOREST AND RANGE EXPERIMENT STATION, BERKELEY, CALIFORNIA
FOREST SERVICE, U.S. DEPARTMENT OF AGRICULTURE

REX -- FORTRAN-4 SYSTEM FOR COMBINATORIAL SCREENING OR CONVENTIONAL ANALYSIS
OF MULTIVARIATE REGRESSIONS.

L.R. GROSENBAUGH

===== SUMMARY =====

'REX' IS AN EXPANSIBLE COMPUTERIZED SYSTEM THAT PROVIDES DATA NEEDED IN REGRESSION OR COVARIANCE ANALYSIS OF AS MANY AS 50 VARIABLES, 8 OF WHICH MAY BE DEPENDENT. ALTERNATIVELY, IT CAN SCREEN VARIOUSLY GENERATED COMBINATIONS OF INDEPENDENT VARIABLES TO FIND THE REGRESSION WITH THE SMALLEST MEAN-SQUARED-RESIDUAL, WHICH WILL BE FITTED IF DESIRED. USERS CAN EASILY PROGRAM ADDITIONAL PROCESSING.

THE SYSTEM WILL ACCEPT INPUT IN THE FORM OF WEIGHTED OR UNWEIGHTED OBSERVATION VECTORS THAT CAN BE REARRANGED, TRANSFORMED, OR TRANSGENERATED. IT WILL ALSO ACCEPT MATRIX INPUT IN THE FORM OF THE COMPACT UPPER TRIANGLE OF A SYMMETRIC MOMENT MATRIX AND WILL CORRECT THE MATRIX FOR MEANS, IF THIS ACTION IS DESIRED. ALTERNATIVELY, IF REGRESSION THROUGH ORIGIN IS DESIRED DESPITE INPUT OF A CORRECTED MATRIX, THE SYSTEM WILL 'UNCORRECT' THE MATRIX. IN ANY CASE, THE COMPACT UPPER TRIANGLE OF THE FULL RESULTANT MATRIX CAN BE PUNCHED FOR LATER USE IF DESIRED.

REGRESSION ANALYSIS MAY INVOLVE THE FULL MATRIX OR ONLY SELECTED ELEMENTS. OUTPUT MAY INCLUDE MOMENT AND/OR CORRELATION MATRICES OF SELECTED ELEMENTS, WITH OR WITHOUT THE INVERSE MATRICES. STANDARDIZED REGRESSION COEFFICIENTS ARE AVAILABLE AS WELL AS CONVENTIONAL REGRESSION COEFFICIENTS, MEANS, VARIANCES, AND ANALYSIS OF SOURCES OF VARIATION. IF OBSERVATION VECTORS HAVE BEEN INPUT, COMPARISONS OF PREDICTIONS WITH OBSERVATIONS CAN BE PRINTED, ALONG WITH CORRESPONDING ERRORS.

WHERE COMBINATORIAL SCREENING HAS BEEN SPECIFIED AND THE NUMBER OF GENERATED COMBINATIONS OF VARIABLES EXCEEDS MAXIMUM MACHINE OR PROGRAM CAPABILITY, THE USER HAS SEVERAL METHODS AT HIS DISPOSAL FOR BRINGING THE GENERATED NUMBER DOWN TO SOME TOLERABLE NUMBER.

===== REGRESSION FITTING, SCREENING, AND ANALYSIS =====

PROBLEM SIZE--LIMITS FOR EACH APPLICATION APPEAR ON THE FIRST PAGE OF OUTPUT INVOLVING THAT APPLICATION. HOWEVER, MAXIMUM NUMBER OF INDEPENDENT VARIABLES THAT CAN BE FITTED TO ONE DEPENDENT VARIABLE IS CURRENTLY 49, WHILE MAXIMUM NUMBER OF VARIABLES THAT CAN BE COMBINED LINEARLY IN ALL POSSIBLE WAYS FOR PREDICTION OF ONE DEPENDENT VARIABLE IS CURRENTLY 13. BY FIXING OR FORCING VARIABLES, COMBINING THEM INTO SETS AND GROUPS, OR LIMITING THE MAXIMUM NUMBER OF OF SETS ALLOWED IN COMBINATION, THIS LATTER LIMIT CAN BE EXTENDED UPWARDS. BY EMPLOYING LESS THAN MAXIMUM NUMBER OF INDEPENDENT VARIABLES, AS MANY AS 8 DEPENDENT VARIABLES CAN BE PROCESSED AT ONCE. WHEN PARAMETERS EXCEED LIMITS HANDLED BY PROGRAM, A MESSAGE OR ASTERISK FLAGGING THE OFFENDER IS PRINTED OUT ON THE FIRST PAGE, AND NO FURTHER COMPUTATIONS ARE PERMITTED.

THE ABOVE LIMITS ASSUME AVAILABILITY OF 32K WORDS OF HIGH-SPEED STORAGE. IF 280K WORDS ARE AVAILABLE, CHANGING A FEW PARAMETERS AND DIMENSIONS WILL ALLOW EXPLORING ALL POSSIBLE LINEAR COMBINATIONS OF 30 INDEPENDENT VARIABLES TAKEN 1,2,3...,17 AT A TIME. THE PROGRAM CURRENTLY REQUIRES USE OF 'OVERLAY' (AN IMPROVED TYPE OF CHAIN EXECUTION) TO AVOID EXCEEDING AVAILABLE STORAGE. TWO BINARY SCRATCH TAPES 'JW' AND 'JX' (NOT NEEDED BY 'OVERLAY') MUST BE SPECIFIED IN BLOCK DATA SUBROUTINE 'BLRM', AND UNNEEDED BUFFERS MUST BE ELIMINATED BY SOME DEVICE SIMILAR TO SUBROUTINE 'BUFK'. APPENDIX A ILLUSTRATES DECK ARRANGEMENTS AND MODIFICATIONS NEEDED BY A NUMBER OF DIFFERENT COMPUTERS.

FIVE CONTROL CARDS MUST BE PROVIDED TO INITIATE PROCESSING OF EACH PROBLEM. THESE MUST BE FOLLOWED BY DATA VECTORS OF QUANTITIES OBSERVED IN ASSOCIATION OR BY A DATA MATRIX OF SUMMED SQUARES AND CROSSPRODUCTS UNLESS INPUT FROM EARLIER PROBLEM HAS BEEN SPECIFIED ON SECOND CONTROL CARD. IF OBSERVATION VECTORS NEED TO BE TRANSFORMED, TRANSGENERATED, OR REARRANGED, SUBROUTINE TRNX CAN BE MODIFIED APPROPRIATELY. THE SIMPLEST WAY OF DOING THIS IS TO USE THE '*ALTER' PROCEDURE AVAILABLE ON THE IBM 7090-7094 (ILLUSTRATED IN APPENDIX B), BUT ANOTHER METHOD AVAILABLE ON ANY COMPUTER IS TO REPLACE 'TRNX' CARDS NUMBERED 18, 19, 20, 21 WITH A GREATER OR LESS NUMBER OF CARDS SPECIFYING THE DESIRED FORTRAN MANIPULATION.

ARRANGEMENT OF VARIABLES IS CRITICAL ONLY IN SCREENING, WHEN IDENTIFICATION OF BEST-FITTING REGRESSIONS IS BEING ATTEMPTED. THEN 'TRNX' MUST ARRAY VARIABLES SO THAT FIXED VARIABLES OCCUR BEFORE ANY NONFIXED VARIABLES, AND ALL DEPENDENT VARIABLES MUST OCCUR CONSECUTIVELY IN A TERMINAL STRING. VARIABLES IN SAME SET AND SETS IN SAME GROUP MUST ALSO BE CONSECUTIVE, WHILE 'TRNX' SHOULD IGNORE UNNECESSARY VARIABLES.

A CARD PUNCHED 'DONE' IN COLUMNS 1-4 MUST FOLLOW LAST OBSERVATION VECTOR OR LAST CARD OF MATRIX ELEMENTS (OR LAST CONTROL CARD WHERE EARLIER DATA WILL BE USED) OF EACH PROBLEM.

A CARD PUNCHED 'DONE DONE DONE....' STARTING IN COLUMN 1 SHOULD FOLLOW THE 'DONE' CARD OF THE LAST PROBLEM TO RETURN CONTROL TO THE MONITOR BEFORE ENCOUNTERING END-OF-FILE.

AFTER READING THE 'DONE DONE DONE....' CARD FOLLOWING A 'DONE' CARD, 'REX' PRINTS NUMBER OF SUCCESSFUL COMPLETIONS AND TOTAL NUMBER OF PROBLEMS ATTEMPTED, THEN RETURNS CONTROL TO SYSTEM.

'REX' SHOULD RUN ON ANY COMPUTER WITH FULL FORTRAN-4 CAPABILITY, WORD-LENGTH OF AT LEAST 4 CHARACTERS, 'OVERLAY' OR 'CHAIN' CAPABILITY, AND AT LEAST 32K WORDS OF MEMORY. THIS INCLUDES IBM 7040-7044, 7090-7094, 360/65, ETC., AND CDC 6400-6600. WHERE MONITORS OCCUPY EXCESSIVE SPACE, DIMENSIONS OF 'A' IN 'SKRN' MUST BE SHRUNK BY CHANGING CARDS NUMBERED SKRN 11, PALM 63, PALM 284 (SEE CDC IN APPENDIX A).

SINCE MAJOR PROCESSING OPTIONS 0 AND 1 OFTEN INVOLVE NULL OR NEARLY NULL MATRICES, IT IS IMPERATIVE THAT USERS SET SYSTEM SUBROUTINES SO THAT UNDERFLOW IS RESET TO NORMAL ZERO WITHOUT ERROR MESSAGES OR ERROR TRACE.

FIRST CONTROL CARD -----

COL. 1- 4 --- ALWAYS BLANK
COL. 5-72 --- BCD PROBLEM IDENTIFICATION
COL. 73-76 --- LABEL FOR OPTIONAL INITIAL MATRIX PUNCHOUT

SECOND CONTROL CARD (FIELDS MUST BE RIGHT-JUSTIFIED) -----

COL. 1- 8 --- NUMBER OF OBSERVATION VECTORS = NOB
COL. 11-12 --- NUMBER OF RAW VARIABLES TO BE READ-IN BEFORE
TRANSFORMATION OR MATRIX FORMATION = NVR
(BLANK IMPLIES SAME AS COL. 15-16)
COL. 15-16 --- MAXIMUM SIZE OF MATRIX AVAILABLE = NVS
(EXCLUDES VECTOR OF TOTALS AND WEIGHT)
COL. 17 --- 'W' IF WEIGHTED REGRESSION IS DESIRED, ELSE
'BLANK' OR ANY CHARACTER BUT 'W'
COL. 19-20 --- LIMIT ON NUMBER OF SETS OF NONFIXED INDEPENDENT
VARIABLES TO BE INCLUDED IN LARGEST
COMBINATORIAL REGRESSION = LM (BLANK OR
ZERO IMPLIES NO LIMIT)

THE 5 PRECEDING FIELDS CAN BE OMITTED COMPLETELY WHEN DATA FROM EARLIER PROBLEM ARE BEING USED.

SIX GROUPS OF PROCESSING OPTIONS ARE CONTROLLED BY COLUMNS 67-72.

COL. 67 --- MAJOR PROCESSING ALTERNATIVES
0 = IDENTIFICATION OF BEST FITTING REGRESSIONS
1 = SAME AS 1 PLUS FITTING OF BEST REGRESSIONS
2 = FITTING OF SPECIFIED REGRESSIONS
3-9 ARE AVAILABLE FOR USER-PROGRAMMED OPTIONS

SECOND CONTROL CARD (CONTINUATION) -----

COL. 68 --- DATA DESCRIPTION, USE (**CORRECTED OR UNCORRECTED)

- 0 = OBSERVATION VECTOR INPUT,
TO BE **CORRECTED BEFORE USE
- 1 = OBSERVATION VECTOR INPUT,
TO BE USED UNCORRECTED
- 2 = UNCORRECTED MATRIX INPUT,
TO BE **CORRECTED BEFORE USE
- 3 = UNCORRECTED MATRIX INPUT,
TO BE USED UNCORRECTED
- 4 = **CORRECTED MATRIX INPUT,
TO BE USED **CORRECTED
- 5 = **CORRECTED MATRIX INPUT,
TO BE UNCORRECTED BEFORE USE
- 6 = UNCORRECTED MATRIX FROM EARLIER PROBLEM,
TO BE **CORRECTED BEFORE USE
- 7 = UNCORRECTED MATRIX FROM EARLIER PROBLEM,
TO BE USED UNCORRECTED
- 8 = **CORRECTED MATRIX FROM EARLIER PROBLEM,
TO BE USED **CORRECTED
- 9 = **CORRECTED MATRIX FROM EARLIER PROBLEM,
TO BE UNCORRECTED BEFORE USE

EVEN NUMBERS IMPLY REGRESSION THROUGH MEAN.
ODD NUMBERS IMPLY REGRESSION THROUGH ORIGIN.

WITH OBSERVATION VECTOR INPUT, NVR RAW ELEMENTS ARE AUTOMATICALLY READ INTO VECTOR D AND ARE AVAILABLE TO 'TRNX'. THESE MAY BE COMBINED OR TRANSFORMED, BUT IN ADDITION TO SUCH ACTIONS, THE USER MUST MODIFY 'TRNX' SO THAT RESULTANT QUANTITIES ARE STORED IN DESIRED ORDER AS THE NVS ELEMENTS OF VECTOR X. ALSO, 'TRNX' MUST SET W EQUAL TO AN APPROPRIATE WEIGHT IF WEIGHTING HAS BEEN SPECIFIED ON CONTROL CARDS.

'TRNX' IS BYPASSED WHEN MATRICES ARE INPUT, BUT DATA MUST CONFORM TO FORMAT (4E16.8), AND ORDER MUST CONFORM TO THAT DESCRIBED BELOW FOR MATRIX PUNCHOUT.

CODE NUMBER APPROPRIATE TO USE OF DATA ALREADY IN STORAGE FROM EARLIER PROBLEM IS 4 GREATER THAN THE CODE NUMBER REQUIRED TO INPUT AND SIMILARLY USE THE SAME MATRIX CURRENTLY.

EARLIER-STORED DATA (CODES 6-9) CANNOT BE USED WITH SCREENING (MAJOR ALTERNATIVE 0 OR 1), WITH FIRST PROBLEM OF ANY RUN, OR IMMEDIATELY AFTER ANY PROBLEM THAT HAS FAILED TO RUN.

SECOND CONTROL CARD (CONTINUATION) -----

COL. 69 --- PUNCHED-CARD OUTPUT

- 0 = NONE
- 1 = MAXIMUM MATRIX AVAILABLE
AFTER TRANSFORMATIONS AND /OR CORRECTION
BUT BEFORE SPECIFICATION OF ELEMENTS

CARDS WILL BE LABELLED AS SPECIFIED ON FIRST CONTROL CARD, WITH 0 OR 1 APPENDED TO SHOW WHETHER THE MOMENTS ARE ABOUT MEAN (CORRECTED) OR ABOUT ORIGIN (UNCORRECTED).

ROW ELEMENTS OF COMPACT UPPER TRIANGULAR MATRIX APPEAR SUCCESSIVELY, IMMEDIATELY FOLLOWED BY VECTOR OF SUMS OF WEIGHTED X'S OR Y'S. SUM OF WEIGHTS IS LAST. ALL MATRIX ELEMENTS HAVE BEEN MULTIPLIED BY FACTOR NEEDED TO MAKE SUM OF WEIGHTS EQUAL NUMBER OF OBSERVATION VECTORS.

COL. 70 --- MATRIX PRINTOUT AFTER SPECIFICATION OF ELEMENTS

- 0 = NONE
- 1 = MOMENT MATRIX
- 2 = CODED MATRIX (CORRELATION IF ABOUT MEAN)
- 3 = BOTH MATRICES

COL. 71 --- INVERSE MATRIX PRINTOUT AFTER SPECIFICATION

- 0 = NONE
- 1 = MOMENT MATRIX INVERSE
- 2 = CODED (CORRELATION) MATRIX INVERSE
WITH (STANDARDIZED) REGRESSION COEFF.
- 3 = BOTH MATRICES

COL. 72 --- PREDICTIONS, OBSERVATIONS, AND ERRORS

- 0 = NO COMPUTATION OR PRINTOUT
- 1 = COMPLETE COMPUTATION AND PRINTOUT, ALONG
WITH SUM OF WEIGHTED ERRORS AND
SUM OF WEIGHTED SQUARED ERRORS

THIS OPTION IS AVAILABLE ONLY WITH CURRENT/OR EARLIER INPUT OF DATA IN FORM OF OBSERVATION VECTORS, NOT WITH MATRIX INPUT. INTERVENING MATRIX INPUT OR FAILURE OF PROBLEM WIPES OUT AVAILABILITY OF EARLIER DATA.

SUM OF WEIGHTED ERRORS SHOULD APPROXIMATE ZERO WHEN REGRESSION IS THROUGH MEAN.

SUM OF WEIGHTED SQUARED ERRORS SHOULD ALWAYS EQUAL ERROR SUM OF SQUARES IN EARLIER TABLE, EXCEPT FOR ROUNDING ERRORS.

THIRD CONTROL CARD FOR MAJOR PROCESSING ALTERNATIVES 0 AND 1 -----
(IDENTIFICATION OF BEST FITTING REGRESSIONS) -----

COL. 1 --- 'BLANK' IMPLIES G'S PUNCHED IN THIS AND ALL
FOLLOWING COLUMNS UNTIL THE FIRST 'Y' OCCURS.

ANY COL. --- 'S' IN A COLUMN IMPLIES THAT THE ORDINAL OF THAT
COLUMN IS THE SUBSCRIPT OF THE FIRST
INDEPENDENT VARIABLE IN A SET. WHEN THE 'S'
OCCURS BEFORE ANY 'G', MEMBERS OF THE SET ARE
TREATED AS FIXED (I.E., THEY WILL ALWAYS BE
INCLUDED IN ANY REGRESSION TESTING COMBINATIONS
OF NONFIXED VARIABLES). MEMBERS OF SETS
OCCURRING AFTER A 'G' HAS APPEARED ARE TREATED
AS NONFIXED, AND ARE INCLUDED IN REGRESSIONS
ONLY WHEN CALLED FOR BY SPECIFIC COMBINATORIAL
RULES. ORDINALS OF BLANK COLUMNS FOLLOWING AN
'S' DENOTE SUBSCRIPTS OF OTHER INDEPENDENT
VARIABLES BELONGING TO THE SET. ALL MEMBERS
OF A GIVEN SET ARE SIMULTANEOUSLY PRESENT OR
ABSENT IN ANY PARTICULAR COMBINATION.

WHEN SEVERAL S'S OCCUR BEFORE ANY 'G', THEY WILL
BE FITTED CUMULATIVELY (NOT INDIVIDUALLY),
SO THAT A SPECIFIC PATH OF FIT CAN BE EXPLORED
AND OBSERVED AT EACH SUCCESSIVE STAGE.

'G' IN A COLUMN IMPLIES THAT THE ORDINAL OF THAT
COLUMN IS THE SUBSCRIPT OF THE FIRST
INDEPENDENT VARIABLE IN THE FIRST NONFIXED SET
OF A GROUP CONSISTING OF ONE OR MORE SETS.

NOT MORE THAN ONE SET FROM A GIVEN GROUP WILL
APPEAR IN THE SAME REGRESSION UNDER
COMBINATORIAL CONTROL (I.E., THEY ARE MUTUALLY
EXCLUSIVE). BLANK COLUMNS IMMEDIATELY AFTER
ANY 'G' DENOTE (BY COLUMN ORDINALS) SUBSCRIPTS
OF INDEPENDENT VARIABLES BELONGING TO THE FIRST
SET IN THAT GROUP, WHILE EACH SUBSEQUENT 'S'
BELONGS TO THE 'G' THAT STARTED THE SEQUENCE
OF SETS IN WHICH THE 'S' OCCURS.

'Y' IN A COLUMN IMPLIES THAT THE ORDINAL OF THAT
COLUMN IS THE SUBSCRIPT OF A DEPENDENT VARIABLE
THAT IS TO BE PREDICTED BY VARIOUS LINEAR
COMBINATIONS OF INDEPENDENT VARIABLES.
'Y' MUST APPEAR IN SUCCESSIVE COLUMNS AFTER ITS
FIRST APPEARANCE, TERMINATING WITH THE COLUMN
WHOSE ORDINAL CORRESPONDS TO THE NUMBER 'NVS'
PUNCHED IN COLUMNS 15-16 OF SECOND CONTROL
CARD. THE CHARACTER FOLLOWING THE LAST 'Y'
MUST BE SAME AS IN COLUMN 17 OF SECOND CONTROL
CARD ('W' REQUIRES WEIGHTING, ALL OTHER
CHARACTERS AND BLANKS IMPLY UNIT WEIGHTS).

THIRD CONTROL CARD FOR MAJOR PROCESSING ALTERNATIVE 2 -----
(FITTING OF SPECIFIED REGRESSIONS) -----

ANY COLUMN WHOSE ORDINAL DOES NOT EXCEED THE NUMBER 'NVS'
(PUNCHED OR IMPLIED IN COLUMNS 15-16 OF SECOND CONTROL CARD)
MAY CONTAIN AN 'X' OR 'Y' TO SHOW THAT THE VARIABLE
WITH SUBSCRIPT INDICATED IS TO BE CONSIDERED AS AN INDEPENDENT
OR DEPENDENT VARIABLE IN A PARTICULAR REGRESSION TO BE FITTED
BY LEAST SQUARES. THE SAME CHARACTER MUST APPEAR IN COLUMN
(NVS+1) OF 3RD CONTROL CARD AS IN COLUMN 17 OF 2ND CARD ('W'
REQUIRES WEIGHTING, ANYTHING ELSE IMPLIES UNIT WEIGHTS).

FOURTH AND FIFTH CONTROL CARDS (MUST BE PRESENT THO OFTEN BLANK) -----

THESE CARDS ARE FOR OBJECT-TIME SPECIFICATION OF VARIABLE
FORMAT NEEDED FOR DATA INPUT IN FORM OF OBSERVATION VECTORS.
NO FORMAT SPECIFICATION IS NEEDED FOR MATRIX INPUT, WHICH
MUST BE IN FORMAT (4E16.8). THUS, THESE TWO CONTROL CARDS
CAN BE LEFT BLANK EXCEPT WHERE COLUMN 68 OF SECOND CONTROL
IS PUNCHED 0 OR 1, BUT THE TWO CARDS MUST ALWAYS BE PRESENT.

WHEN USER ELECTS TO FIT AND ANALYZE A SPECIFIC REGRESSION,
MINIMUM SUCCESSFUL OUTPUT WILL CONSIST OF ONE PAGE LISTING IMPLICIT
OR EXPLICIT PARAMETERS OF PROBLEM, ONE PAGE OF REGRESSION COEFFICIENTS
WITH THEIR VARIANCES ON THE FOLLOWING PAGE, ONE PAGE FOR QUANTITIES
NEEDED IN REGRESSION-COVARIANCE ANALYSIS, AND ONE PAGE GIVING MEANS
AND VARIANCES OF ALL VARIABLES INVOLVED IN THE SPECIFICATION OF
VARIABLES ON THE THIRD CONTROL CARD.

WHEN USER ELECTS IDENTIFICATION OF BEST FITTING REGRESSIONS,
MINIMUM SUCCESSFUL OUTPUT WILL CONSIST OF ONE PAGE LISTING IMPLICIT
OR EXPLICIT PARAMETERS OF PROBLEM, AND ONE OR MORE PAGES OF RELATIVE
MEAN SQUARED RESIDUALS FOR EACH REGRESSION GENERATED BY THE SPECIFIC
COMBINATORIAL MODEL ESTABLISHED BY THE SECOND AND THIRD CONTROL CARDS.
THESE RELATIVE MEAN SQUARED RESIDUALS (RELATIVE TO SIMPLE VARIANCE
ABOUT MEAN Y) ARE IDENTIFIED AS TO COMPONENT COMBINATORIAL SETS
INVOLVED. ALL FIXED SETS ARE IMPLICITLY INCLUDED IN ANY COMBINATORIAL
REGRESSION. COMBINATORIAL REGRESSIONS HAVING MINIMUM VARIANCES ARE
IDENTIFIED IN TABLE ON LAST PAGE.

THE COEFFICIENT OF COLLINEARITY, A MEASURE OF RELATIVE NULLITY OF
THE MATRIX OF INDEPENDENT VARIABLES, IS ALSO LISTED AT THE LEFT OF
THE RELATIVE MEAN SQUARED RESIDUALS. A ZERO OR NEAR-ZERO COEFFICIENT
(WITH NEGATIVE EXPONENT MORE THAN DOUBLE THE NUMBER OF INDEPENDENT
VARIABLES INVOLVED) WARNS OF NULL OR NEARLY NULL MATRIX WHOSE INVERSE
MIGHT NOT BE COMPUTABLE BY SUBROUTINE CBXR, USUALLY BECAUSE OF HIGH
CORRELATION AMONG INDEPENDENT VARIABLES. COMPLETELY ORTHOGONAL
VARIABLES WOULD BE INDICATED BY A COEFFICIENT OF UNITY, SINCE IT IS
MERELY THE RATIO OF THE DETERMINANT OF THE UNCORRECTED MOMENT MATRIX
(ABOUT ORIGIN) OF INDEPENDENT VARIABLES TO THE CONTINUING PRODUCT
OF THE MATRIX DIAGONALS. REGRESSIONS THROUGH THE MEAN CONSIDER THE
VECTOR OF VARIABLE SUMS AND AGGREGATE WEIGHTS AS AUGMENTING THE
BASIC UNCORRECTED MATRIX, BUT OTHERWISE THE COMPUTATION IS THE SAME.

A COEFFICIENT FOR THE IMPLIED CONSTANT PSEUDO-VARIABLE 'UNITY' IS ALWAYS COMPUTED WHEN A PARTICULAR REGRESSION THROUGH MEAN IS SPECIFIED. THIS CONSTANT IS ALSO KNOWN AS THE INTERCEPT (ON THE Y AXIS) AND IS SUBSCRIPTED 'U' FOR UNITY. IT IMMEDIATELY FOLLOWS THE COEFFICIENT FOR THE LAST REAL VARIABLE, AS DOES ITS VARIANCE ON THE NEXT PAGE. SIMILARLY, A FINAL COLUMN OF ELEMENTS $C(I,U)$ FOR THE MOMENT MATRIX INVERSE IS ALWAYS COMPUTED WHEN REGRESSION THROUGH MEAN IS SPECIFIED. THIS IS THE IDENTICAL COLUMN OF ELEMENTS THAT WOULD BE COMPUTED LESS ACCURATELY IF REGRESSION THROUGH ORIGIN WERE SPECIFIED AND TERMINAL PSEUDO-VARIABLE OF UNITY WERE PROGRAMMED IN SUBROUTINE 'TRNX'. THE LATTER PROCEDURE RESULTS IN AUGMENTING THE UNCORRECTED MOMENT MATRIX BY A TERMINAL VECTOR OF SIMPLE TOTALS, AN UNDESIRABLE PROCEDURE EXCEPT FOR ITS CONVENIENCE IN COVARIANCE ANALYSIS.

ALTHOUGH SCREENING OF REGRESSIONS BY THE CONVENTIONAL STEPWISE PROCEDURE WILL ORDINARILY NOT LEAD TO AS GOOD REGRESSIONS AS THE COMBINATORIAL APPROACH DESCRIBED EARLIER, SOMETIMES THE USER MAY WISH TO VIEW THE BEHAVIOR OF THE MEAN SQUARED RESIDUAL AS VARIABLES ARE ADDED ALONG A SPECIFIC PATH OF FIT. THIS CAN BE DONE BY CHOOSING MAJOR PROCESSING ALTERNATIVES 0 OR 1 AND PUNCHING 'S' IN THE FIRST COLUMN OF THE SECOND CONTROL CARD, AND IN SUBSEQUENT COLUMNS WHERE A CUMULATIVE VALUE FOR THE MEAN SQUARED RESIDUAL IS DESIRED. ALTHOUGH NO 'BEST' REGRESSION ALONG THIS PATH WILL BE PROGRAM-SELECTED, THE USER CAN EASILY LOCATE IT BY INSPECTION. THE COMPLETE REGRESSION CONTAINING ALL VARIABLES ALONG THE PATH WILL NOT BE FITTED, EVEN WITH MAJOR PROCESSING ALTERNATIVE 1, UNLESS THE LAST SPECIFIED SET IS PUNCHED 'G'.

===== COVARIANCE ANALYSIS =====

THE SIMPLEST METHOD FOR ANALYZING COVARIANCE DOES NOT REQUIRE ORTHOGONAL DUMMY VARIABLES OR WEIGHTING INVERSELY ACCORDING TO GROUP SIZE, ALTHOUGH SUCH DEVICES IMPROVE ACCURACY WHEN MATRICES ARE ILL-CONDITIONED. IF (G) IS THE NUMBER OF GROUPS INTO WHICH OBSERVATIONS CAN MEANINGFULLY BE DIVIDED, AND IF (K) IS THE NUMBER OF INDEPENDENT VARIABLES, THEN 'REX' CAN EASILY ANALYZE COVARIANCE IN CASES WHERE $(G+1)*(K+1)$ DOES NOT EXCEED 49, AND CAN HANDLE SLIGHTLY LARGER PROBLEMS WITH A BIT MORE TROUBLE.

USER MUST MODIFY SUBROUTINE 'TRNX' SO THAT AN ENLARGED OBSERVATION VECTOR IS DEFINED WITH (G+1) SETS OF (K+1) INDEPENDENT VARIABLES AND WITH ONE OR MORE DEPENDENT VARIABLES. THE (K) ACTUAL INDEPENDENT VARIABLES ARE PLACED IN THE FIRST (K) ELEMENTS, WHILE UNITY IS PLACED IN THE (K+1)TH ELEMENT. THIS SAME SET OF (K+1) VARIABLES IS ALSO PLACED IN THE SUBSEQUENT SET OF ELEMENTS WHOSE ORDINAL IS ONE GREATER THAN THAT FOR THE GROUP. ALL OTHER INDEPENDENT VARIABLES MUST BE EQUATED TO ZERO.

THREE REGRESSIONS MUST NOW BE SPECIFIED TO OBTAIN ALL NEEDED DATA. ALTHOUGH THEY ARE NOMINALLY SPECIFIED TO PASS THROUGH THE ORIGIN, THE CONSTANT DUMMY FORCES THEM TO PASS THROUGH THE MEAN. THE SAME MAXIMUM MATRIX SERVES ALL 3. THE FIRST, CALLED A1, REPRESENTS UNGROUPED DATA. THE SECOND, CALLED B1, ALLOWS DIFFERENT INTERCEPTS FOR EACH GROUP BUT REQUIRES POOLED SLOPES. THE THIRD ALLOWS EACH GROUP TO HAVE ITS OWN REGRESSION.

THE THIRD CONTROL CARDS FOR THESE REGRESSIONS WOULD APPEAR AS FOLLOWS WHERE THERE WERE 6 GROUPS WITH 3 INDEPENDENT AND 2 DEPENDENT VARIABLES, AS IN THE LAST EXAMPLE OF INPUT IN APPENDIX B -----

```
A1'S CARD ==XXXX                                YY=
B1'S CARD ==XXX      X  X  X  X  X  X  XXX=
C1'S CARD ==          XXXXXXXXXXXXXXXXXXXXXXXYY=
```

FROM THE 3 OUTPUTS THUS SECURED, THE DEGREES OF FREEDOM (DFS) AND SUMS OF SQUARES (SS) ATTRIBUTABLE TO REGRESSION (R) AND TO ERROR (E) CAN BE OBTAINED, ALONG WITH THE MEAN SQUARED RESIDUALS (MSQR) OR ERROR VARIANCES. NEXT (A1E MINUS B1E) AND (B1E MINUS C1E) ARE COMPUTED BY HAND, BOTH FOR DFS AND FOR SS. LASTLY, CORRECTION TERMS MUST BE SUBTRACTED FROM B1R (BOTH DFS AND SS). THE APPROPRIATE SUBTRACTION IN THE CASE OF B1R DFS IS UNITY, WHILE IN THE CASE OF B1R SS, IT IS (SQUARED MEAN Y)*(NUMBER OF OBSERVATIONS).

AN ANALYSIS OF COVARIANCE USUALLY DETERMINES THE FOLLOWING
(DEGREES OF FREEDOM FOR 'F' CAN BE INFERRED FROM SS OR MSQR) ---

$F = (\text{CORRECTED B1R SS}) / ((\text{B1E MSQR}) * (\text{CORRECTED B1R DFS}))$
FOR SIGNIFICANCE OF POOLED REGRESSION COEFFICIENTS

$F = (\text{B1E SS MINUS C1E SS}) / ((\text{C1E MSQR}) * (\text{B1E DFS MINUS C1E DFS}))$
FOR SIGNIFICANCE OF DIFFERENCES AMONG SLOPES OF
INDIVIDUAL GROUP REGRESSIONS

$F = (\text{A1E SS MINUS B1E SS}) / ((\text{B1E MSQR}) * (\text{A1E DFS MINUS B1E DFS}))$
FOR SIGNIFICANCE OF DIFFERENCES AMONG GROUP INTERCEPTS
WHEN SLOPES ARE POOLED (TEST OF DIFFERENCES AMONG
ADJUSTED MEANS)

GENERALLY, THERE IS LITTLE USE IN CONDUCTING THE SECOND AND THIRD TESTS UNLESS THE FIRST SHOWS SIGNIFICANCE. SIMILARLY, THE THIRD TEST IS USUALLY OMITTED EXCEPT WHERE A SIGNIFICANT FIRST TEST AND A NONSIGNIFICANT SECOND TEST HAVE OCCURRED.

POOLED SLOPE COEFFICIENTS AND INDIVIDUAL GROUP INTERCEPTS MAY BE READ DIRECTLY FROM THE OUTPUT OF REGRESSION B1, WHILE INDIVIDUAL GROUP REGRESSION SLOPES AND INTERCEPTS OCCUR IN SETS OF (K+1) ENTITIES IN THE OUTPUT OF REGRESSION C1. THE VARIANCES OF THESE COEFFICIENTS ARE DERIVED FROM THE POOLED SUM OF SQUARED RESIDUALS FROM INDIVIDUAL REGRESSIONS RATHER THAN FROM SUM FOR JUST THE PARTICULAR GROUP, BUT OTHERWISE HAVE VALIDITY. THE OUTPUT OF MEANS AND VARIANCES FOR C1 REQUIRES SOME DECODING TO PLACE THEM ON AN INDIVIDUAL GROUP BASIS, BUT COMPUTATIONS ARE SIMPLE AND OBVIOUS.

AN ANALOGOUS PROCEDURE CAN BE FOLLOWED WITH ORTHOGONAL DUMMY VARIABLES SECURED FROM DELURY'S TABLES (*2) IF ALL 3 REGRESSIONS ARE SPECIFIED TO PASS THROUGH THE MEAN INSTEAD OF THE ORIGIN. HOWEVER, EACH GROUP MUST BE CHARACTERIZED BY (G-1) DUMMIES AND THEIR SUITABLY LOCATED CROSSPRODUCTS WITH THE (K) REAL VARIABLES, INSTEAD OF SIMPLY BY THE LOCATION OF A SECOND SET OF THE REAL VARIABLES AND UNITY. THESE DUMMIES MUST BE DIVIDED BY GROUP SIZE UNLESS ALL GROUPS CONTAIN THE SAME NUMBER OF OBSERVATIONS. ALTHOUGH NO CORRECTION IS NEEDED FOR THE SECOND REGRESSION (A0, B0, C0 WILL BE USED TO DISTINGUISH REGRESSIONS THROUGH THE MEAN HAVING ORTHOGONAL DUMMY VARIABLES), EACH GROUP INTERCEPT AND INDIVIDUAL GROUP COEFFICIENT MUST BE CALCULATED AS THE SEPARATE SUM OF (G) QUANTITIES, (G-1) OF WHICH ARE SECURED AS THE PRODUCT OF A DUMMY COEFFICIENT MULTIPLIED BY A DUMMY VARIABLE APPROPRIATE TO THAT PARTICULAR GROUP. PROBLEMS WHERE (G*(K+1))-1 DOES NOT EXCEED 49 CAN BE SIMPLY HANDLED BY THIS TECHNIQUE. LARGER PROBLEMS WILL REQUIRE GETTING THE C0 OR C1 REGRESSION COMPONENTS FROM SUMMING QUANTITIES SECURED BY SEPARATING DATA INTO GROUPS AND FITTING EACH GROUP REGRESSION INDIVIDUALLY. EXAMPLES AORX, BORX, CORX IN APPENDIX B ILLUSTRATE COVARIANCE INPUT USING ORTHOGONAL POLYNOMIALS.

THE SIMPLER COVARIANCE PROCEDURE (USING UNITY AS A DUMMY VARIABLE WITH THE SAME RAW INPUT DATA) IS ILLUSTRATED BY EXAMPLES A1RX, B1RX, C1RX IN APPENDIX B. FORMAL REQUIREMENT THAT THE REGRESSIONS PASS THROUGH ORIGIN IS EFFECTIVELY CANCELLED BY USE OF UNITY AS A DUMMY VARIABLE.

IT CAN BE SEEN THAT THERE ARE 6 GROUPS, EACH CONTAINING 5 SETS OF OBSERVATIONS. ALTERATIONS TO 'TRNX' SHOW THAT THE GROUP CODE IS READ AS RAW VARIABLE D(2), AND THAT RAW VARIABLES D(3), D(6), D(10) SERVE AS FINAL VARIABLES X(1), X(2), X(3), WITH X(4) ALWAYS BEING UNITY. ADDITIONALLY, EXACTLY THE SAME SET OF VARIABLES IS STORED IN A SEQUENCE OF 4 X'S APPROPRIATE TO THE OBSERVED GROUP CODE. THE OTHER SUPERFLUOUS GROUPS OF 4 X'S ARE SET TO ZERO, WHILE THE 2 DEPENDENT VARIABLES ORIGINALLY READ AS D(11) AND D(12) ARE MOVED TO X(29) AND X(30) AT THE END OF THE 7 SEQUENCES OF 4 X'S RESERVED FOR THE POOLED REGRESSION AND 6 INDIVIDUAL GROUP REGRESSIONS. FROM THE SUBSEQUENT OUTPUT THAT REX PRODUCES, QUANTITIES APPEARING IN THE MULTIPLE COVARIANCE ANALYSIS BELOW ARE EASILY OBTAINED.

SOURCE	DFS	29Y SS	30Y SS	29Y MSQR	30Y MSQR
A1E	26	2187.50	1173.59		
B1E	21	28.80	731.73	1.371	34.84
C1E	6	21.94	71.64	3.657	11.94
B1R(+)	9	9426.20	4074.27		
(UNCORRECTED LINE ABOVE NEEDS CORRECTION TERMS BELOW)					
CORR.(-)	1	7207.50	3456.13		
=====					
B1R=B1R(+)-CORR.	8	2218.70	618.14		
AB = A1E-B1E	5	2158.70	441.87		
BC = B1E-C1E	15	6.86	660.09		
=====					

TO TEST SIGNIFICANCE OF DIFFERENCES OF POOLED REGRESSION COEFFICIENTS FROM ZERO --

$$F = (B1R \text{ SS}) / ((B1R \text{ DFS}) * (B1E \text{ MSQR})) = 202.2^{**} \text{ FOR } 29Y.$$

$$= 2.22 \text{ FOR } 30Y.$$

$$DFS = 8/21$$

TO TEST SIGNIFICANCE OF DIFFERENCES OF INDIVIDUAL GROUP REGRESSION COEFFICIENTS FROM POOLED REGRESSION COEFFICIENTS --

$$F = (BC \text{ SS}) / ((BC \text{ DFS}) * (C1E \text{ MSQR})) = .13 \text{ FOR } 29Y.$$

$$= 3.69 \text{ FOR } 30Y.$$

$$DFS = 15/6$$

TO TEST SIGNIFICANCE OF DIFFERENCES AMONG INDIVIDUAL GROUP MEANS ADJUSTED BY POOLED REGRESSION COEFFICIENTS (INTERCEPT DIFFERENCES) --

$$F = (AB \text{ SS}) / ((AB \text{ DFS}) * (B1E \text{ MSQR})) = 314.8^{**} \text{ FOR } 29Y.$$

$$= 2.54 \text{ FOR } 30Y.$$

$$DFS = 5/21$$

===== BACKGROUND AND DEVELOPMENT OF SYSTEM =====

PROGRAM LOGIC IS BRIEFLY AS FOLLOWS. 'REX' IS THE EXECUTIVE ROUTINE (OR ZERO LINK IN THE OVERLAY STRUCTURE) THAT CONTROLS ALL OTHER SUBROUTINES. 'PALM' INTERPRETS CONTROL CARDS 2 AND 3. THEN 'MATX' WITH 'TRNX' HANDLE DATA INPUT, TRANSFORMATIONS, CORRECTIONS TO CURRENT OR EARLIER DATA INPUT, AND WHATEVER PUNCHED-CARD OUTPUT THE USER MAY HAVE SPECIFIED. THEN, DEPENDING ON USER'S CHOICE OF MAJOR PROCESSING OPTIONS, CONTROL PASSES EITHER TO 'SKRN', TO 'SKRN' FOLLOWED BY 'CBXR', TO 'CBXR' ALONE, OR TO SOME USER-SUPPLIED SET OF SUBROUTINES WHICH MAY INCLUDE 'SKRN' AND/OR 'CBXR'. FINALLY, CONTROL RETURNS TO 'REX' FOR STARTING ON THE NEXT PROBLEM IN THE STACK.

THE BROAD OUTLINE FOR 'REX' WAS CONCEIVED BY THE AUTHOR AS A RESULT OF HIS EXPERIENCE IN DESIGNING AND USING A GROUP OF REGRESSION PROGRAMS FOR THE IBM 704 IN THE PERIOD 1957-60, THE FIRST OF WHICH WAS DESIGNATED G2-TV-REM (SHARE DISTRIBUTION AGENCY NUMBER 822). THIS WAS THE FIRST COMBINATORIAL APPROACH TO REGRESSION SCREENING, MADE POSSIBLE BY THE REVOLUTIONARY SPEEDS OF NEW ELECTRONIC COMPUTERS (*4). 'REM' WAS SOON SUPPLEMENTED BY 'REA' FOR MATRIX INPUT, 'XXR' FOR ACCUMULATION OF LARGER CORRELATION MATRICES, AND 'CBY' AND 'CBZ' FOR FITTING PARTICULAR REGRESSIONS BY THE MODIFIED FISHER-DOOLITTLE METHOD.

IN CONTRAST TO THE COMBINATORIAL APPROACH WHICH EXPLORES ALL POSSIBLE LINEAR COMBINATIONS OF VARIABLES WITHIN CERTAIN CONSTRAINTS IMPOSED TO KEEP THE DIMENSIONS OF THE PROBLEM WITHIN COMPUTABLE LIMITS, THE MORE WIDELY KNOWN STEPWISE APPROACH ADDS OR DELETES ONE VARIABLE AT A TIME TO SOME PREVIOUSLY SELECTED GROUP OF VARIABLES. UNFORTUNATELY, THIS PROCESS IS NOT CAPABLE OF FINDING THE COMBINATION OF INDEPENDENT VARIABLES THAT BEST EXPLAINS THE VARIATION OF THE DEPENDENT VARIABLE (EXCEPT IN THE TRIVIAL CASE INVOLVING NO MORE THAN 2 VARIABLES), AND MOST TESTS OF SIGNIFICANCE APPLIED TO REJECTION OR ACCEPTANCE OF VARIABLES FOR A GIVEN STEP ARE INVALID.

IT WAS THE AUTHOR'S DISSATISFACTION WITH THE MORE POPULAR STEPWISE PROCEDURE WHICH CAUSED HIM TO EXPLORE THE COMBINATORIAL APPROACH. HE EARLY DISCOVERED THAT SUMS OF SQUARES OR MULTIPLE CORRELATION COEFFICIENTS WERE NOT GOOD CRITERIA FOR DETERMINING THE 'BEST' REGRESSION FOR A PARTICULAR SET OF OBSERVATIONS, BECAUSE THEY INCREASED OR DECREASED MONOTONICALLY AS THE NUMBER OF VARIABLES INCREASED. HOWEVER, THE MEAN SQUARED RESIDUAL (OR VARIANCE OF THE REGRESSION ESTIMATE) IS NOT A MONOTONIC FUNCTION, BUT FLUCTUATES IN AN UNPREDICTABLE MANNER DEPENDING ON BOTH SUMS OF SQUARES AND DEGREES OF FREEDOM. IT IS A FIGURE OF MERIT WHOSE MINIMUM FLAGS A 'BEST BET'.

THE FIRST EXAMPLE IN APPENDIX B ('ABST') HAS OUTPUT FROM MAJOR OPTION 1 SHOWN IN APPENDIX C, AND ILLUSTRATES A SITUATION WHERE THE COMBINATORIAL APPROACH SUCCEEDS IN FINDING THE 'BEST' REGRESSION (INVOLVING VARIABLES 5, 6, 7).

STARTING WITH EXACTLY THE SAME DATA, NEITHER THE ASCENDING NOR THE DESCENDING STEPWISE REGRESSION ANALYSIS CAN REACH THE CORRECT ANSWER, SINCE THEIR SELECTION OF PATHS IS IRREVOCABLY LIMITED BY INITIAL INCLUSION OF VARIABLES 1 AND 2 OR BY INITIAL DELETION OF VARIABLES 5, 6, AND 7.

ALTHOUGH NO ONE CAN SAY WITH ASSURANCE THAT ANY VARIANCE IS A GLOBAL MINIMUM UNLESS ALL POSSIBLE COMBINATIONS HAVE BEEN TRIED, EXPERIENCE HAS SHOWN THAT THE 'BEST' COMBINATION FROM A LARGE NUMBER OF VARIABLES USUALLY DOES NOT INVOLVE MORE THAN 3 TO 7 VARIABLES. HENCE, THE INTRODUCTION OF SOME SUCH MODERATE LIMIT TO KEEP THE PROBLEM WITHIN CAPABILITIES OF THE PROGRAM USUALLY WILL NOT CHANGE THE OUTCOME.

THE PRESENT PROGRAM REFLECTS THE FOREGOING AS WELL AS A NUMBER OF ADDITIONAL IMPROVEMENTS STEMMING FROM THE AUTHOR'S EXPERIENCE. THERE ARE SEVERAL MAJOR IMPROVEMENTS FOR WHICH OTHERS SHOULD BE CREDITED, HOWEVER.

FURNIVAL (*3) REALIZED THAT RESIDUALS FROM REGRESSION COULD BE COMPUTED BY DETERMINANTS (METHOD OF SINGLE DIVISION) MUCH MORE EASILY THAN BY THE USUAL METHOD INVOLVING MATRIX INVERSION, IF OTHER REGRESSION STATISTICS WERE SACRIFICED. THIS MADE THE COEFFICIENT OF COLLINEARITY CHEAPLY AVAILABLE AS A BYPRODUCT. HE ALSO DEvised AN INGENIOUS COMPUTATIONAL SEQUENCE TO MINIMIZE DETERMINANTAL ARITHMETIC, AND RECOGNIZED THAT SET AND GROUP CONSTRAINTS MIGHT BE USEFUL TOOLS FOR KEEPING PROBLEM DIMENSIONS WITHIN REASONABLE BOUNDS. THESE IMPROVEMENTS HAVE BEEN EMBODIED IN SUBROUTINE 'SKRN'.

THE AUTHOR ORIGINALLY FELT THAT ACCUMULATION OF THE MATRIX OF SUMS OF SQUARES AND CROSSPRODUCTS FROM OBSERVATION VECTORS COULD BE MOST ACCURATELY AND EFFICIENTLY ACHIEVED BY INTEGER ARITHMETIC WITH MULTIPLE PRECISION, AS HE DID IN 'XXR'. RODDEN (*7) LATER SUPPORTED THIS VIEW, BUT THE AUTHOR DECIDED AGAINST IT IN 'REX' BECAUSE ASSEMBLY LANGUAGE WOULD BE REQUIRED IN AN OTHERWISE ALL-FORTRAN PROGRAM.

WELFORD'S TECHNIQUE (*9) INVOLVING USE OF SINGLE-PRECISION, FLOATING-POINT, PROGRESSIVE AVERAGES WAS EMPLOYED INITIALLY, BUT ITS NOTICEABLE INACCURACY ON SMALL TEST PROBLEMS WAS A DISAPPOINTMENT. NEELY (*6) LATER DOCUMENTED SIMILAR EXPERIENCE.

FINALLY, THE STRAIGHTFORWARD USE OF A SINGLE PASS WITH DOUBLE-PRECISION FLOATING-POINT ARITHMETIC SEEMED PREFERABLE TO THE USE OF TWO PASSES WITH SINGLE-PRECISION FLOATING-POINT ARITHMETIC (FIRST PASS TO COMPUTE MEANS, SECOND PASS TO CODE DEVIATIONS AND FORM CROSSPRODUCTS). HENCE, FORMATION OF MATRIX AND ITS SUBSEQUENT CORRECTION FOR MEANS (IF SPECIFIED) ARE PERFORMED WITH DOUBLE-PRECISION, BUT THE FINAL MATRIX IS STORED WITH SINGLE-PRECISION.

RECENT IMPROVEMENTS IN HIGH-SPEED COMPUTERS HAVE MADE THE SQUARE-ROOT METHOD OF MATRIX INVERSION MUCH MORE ATTRACTIVE THAN THE MODIFIED JORDAN ELIMINATION METHOD FOR SYMMETRIC MATRICES, ESPECIALLY IF THEY HAPPEN TO BE ILL-CONDITIONED. BODEWIG (*1) POINTS THIS OUT, ALTHOUGH DWYER AND OTHERS SURMISED THE SAME THING PRIOR TO THE AVAILABILITY OF MODERN COMPUTERS. CONSEQUENTLY, THE AUTHOR WROTE A DOUBLE-PRECISION FORTRAN-4 MATRIX INVERTER FOR 'REX' THAT USES THE SQUARE-ROOT METHOD.

THUS, REX EMPLOYS DOUBLE-PRECISION ARITHMETIC FOR ACCUMULATION AND CORRECTION OF MOMENT MATRIX, FOR MATRIX INVERSION BY THE SUPERIOR SQUARE-ROOT METHOD, AND FOR COMPUTATION OF MOST OF THE IMPORTANT REGRESSION STATISTICS. HENCE, 'REX' RETAINS MORE SIGNIFICANT DIGITS THAN DO LEAST-SQUARES PROGRAMS TESTED BY LONGLEY (*5), WHICH EMPLOYED SINGLE-PRECISION ARITHMETIC AND LESS ACCURATE INVERSION METHODS.

REGRESSION STATISTICS FOR 'REX' ARE COMPUTED BY THE MODIFIED FISHER-DOOLITTLE METHOD DESCRIBED IN (*8). HOWEVER, STANDARDIZED REGRESSION COEFFICIENTS ARE NOT CALCULATED UNLESS REQUESTED. CODING OF MATRIX ELEMENTS PRIOR TO MATRIX INVERSION IS ALWAYS BY DIVISION BY SQUARE-ROOT OF PRODUCT OF DIAGONAL ELEMENTS WITH SUBSCRIPTS INVOLVED, SO THE MATRIX DOES NOT BECOME A TRUE CORRELATION MATRIX EXCEPT WHERE REGRESSION THROUGH MEAN HAS BEEN SPECIFIED. RSQUARED THEREFORE, IS MERELY THE RATIO OF TWO SUMS OF SQUARES WHEN REGRESSION THROUGH ORIGIN HAS BEEN SPECIFIED. THE MAIN ARGUMENT IN FAVOR OF SUCH CODING IS THAT IT IMPROVES MATRIX CONDITION.

THE AUTHOR CONSIDERED USING HIS COMBINATORIAL GENERATOR DEVELOPED FOR ANALYSIS OF VARIANCE PROGRAM G4-BC-ANV (SHARE DISTRIBUTION AGENCY NUMBER 3337), WHICH COULD HAVE HANDLED AN UNLIMITED NUMBER OF GROUPS COMBINATORIALLY WITHOUT ANY PRACTICAL STORAGE LIMITATION, BUT MATRIX OPERATIONS AFTER EACH GENERATION WOULD START ANEW, WHICH INVOLVED A CONSIDERABLE EXTRA EXPENDITURE OF TIME, SO THE IDEA WAS DISCARDED. PERSONS INTERESTED IN MASSIVE COMBINATORIAL ANALYSES MIGHT STILL WISH TO EMPLOY THIS ALTERNATIVE BY MODIFYING THE PRESENT VERSION OF 'SKRN'.

IN VIEW OF THE RELIABILITY OF THIRD-GENERATION SOLID-STATE COMPUTERS, THE ELABORATE SUM AND PRODUCT CHECKS OF 'REM', 'XXR', 'CBY', AND 'CBZ' HAVE BEEN DISCARDED. NO CHECKS AT ALL HAVE BEEN PROVIDED FOR THE SIMPLE SCREENING OPTION (MAJOR PROCESSING ALTERNATIVE 0), BUT MOST MACHINE ERRORS WOULD CAUSE OUTPUT FIELD OVERFLOW FLAGGED BY MONITOR. MAJOR PROCESSING ALTERNATIVE 1 HAS RMSQR OF BEST REGRESSIONS COMPUTED BY 2 INDEPENDENT PROCEDURES. MAJOR PROCESSING ALTERNATIVES 1 AND 2 (WITH PREDICTIONS) HAVE SUM OF SQUARED RESIDUALS COMPUTED BY 2 INDEPENDENT PROCEDURES, AND IF REGRESSION IS THROUGH MEAN, THE SUM OF RESIDUALS SHOULD APPROXIMATE ZERO FOR AN ADDITIONAL CHECK.

DYNAMIC MODIFICATION OF FORMATS STORED AS BCD ARRAYS SAVES SPACE AND HAS BEEN USED FREELY, ALTHOUGH THE AUTHOR ATTEMPTED TO MAINTAIN COMPATIBILITY BETWEEN MACHINES INCAPABLE OF HANDLING FORMATS BIGGER THAN A4 (IBM 360, CDC 3500) AND THOSE CAPABLE OF HANDLING A6 (SUCH AS THE IBM 7040-7044, 7090-7094, CDC 6400-6600).

===== AIDS FOR USERS MODIFYING OR EXPANDING SYSTEM =====

'REX' HAS BEEN DESIGNED AS AN OPEN-ENDED SYSTEM RATHER THAN AS A SINGLE PROGRAM. HENCE, USERS WISHING TO EXPAND IT CAN CALL ON OTHER MAJOR PROCESSING OPTIONS OF THEIR OWN. THIS WOULD MERELY INVOLVE PUNCHING AN APPROPRIATE INTEGER (FROM 3 THROUGH 9) IN COLUMN 67 OF OF THEIR SECOND CONTROL CARD AND REPLACING CORRESPONDING FORTRAN STATEMENT NUMBERED 300 THROUGH 900 IN 'REX' BY AN APPROPRIATE 'CALL'.

WITH A LITTLE ADDITIONAL PROGRAMMING, SUCH EXPANSIONS CAN EMPLOY 'PALM', 'MATX', AND 'TRNX' FOR INPUT PROCESSING AND ACCUMULATIONS, AND 'CBXR' FOR INTERMEDIATE CALCULATIONS. LOCATIONS OF RESULTS OBTAINED AND PRESERVED BY 'CBXR' ARE GIVEN BELOW.

ELEMENTS OF THE MOMENT MATRIX INVERSE (AUGMENTED BY C(I,U) WHERE APPROPRIATE) ARE STORED DOUBLE-PRECISION IN UPPER TRIANGLE OF A(I,J). B(I,K), THE MATRIX OF REGRESSION COEFFICIENTS, IS STORED SINGLE-PRECISION IN VB(I,K) AND DOUBLE-PRECISION IN B(I,K), WITH B(U,K) FOLLOWING THE LAST REAL COEFFICIENT WHERE APPROPRIATE. VB(51,K) CONTAINS RSQUARED, WHILE B(51,K) AND V(51,K) EACH CONTAIN THE VARIANCE OF THE REGRESSION ESTIMATE. VARIANCES OF THE REGRESSION COEFFICIENTS ARE STORED SINGLE-PRECISION IN V(I,K).

IT SHOULD BE NOTED THAT NEITHER OF THE DOUBLE-PRECISION ARRAYS 'A' AND 'B' IS IN COMMON STORAGE CURRENTLY. USERS WOULD HAVE TO DECLARE THEM AS LABELLED COMMON IF THEY ARE TO BE MADE AVAILABLE TO TO ANY OTHER SUBROUTINE THAN 'CBXR' IN THE SAME LINK AS 'CBXR'.

MNC(K) CONTAINS THE NUMBER OF INDEPENDENT VARIABLES INVOLVED, AND SUBSCRIPTS OF BOTH INDEPENDENT AND DEPENDENT VARIABLES ARE STORED IN KX(I). H(K,1) CONTAINS RSQUARED, H(K,2) CONTAINS (1-RSQUARED), AND H(K,3) CONTAINS TOTAL SUM OF SQUARED DEVIATIONS ABOUT MEAN Y OR ABOUT ORIGIN, ALL SINGLE-PRECISION. U(K) CONTAINS 'RMSQR', THE SINGLE-PRECISION VARIANCE OF REGRESSION ESTIMATE RELATIVE TO VARIANCE ABOUT MEAN.

===== LITERATURE CITED =====

- (*1) BODEWIG, E. 1959. MATRIX CALCULUS. INTERSCI. PUBL., INC., NEW YORK CITY. 2ND. ED., REV. 452 PP.
- (*2) DELURY, D.B. 1950. VALUES AND INTEGRALS OF THE ORTHOGONAL POLYNOMIALS UP TO $N = 26$. UNIV. OF TORONTO PRESS, TORONTO. 33 PP.
- (*3) FURNIVAL, G.M. 1965. MORE ON THE ELUSIVE FORMULA OF BEST FIT. PROCEEDINGS SOC. AMER. FORESTERS MEETING SEPT. 27-OCT. 1, 1964, DENVER, COL. (WASHINGTON, D.C.), PP. 201-207.
- (*4) GROSENBAUGH, L.R. 1958. THE ELUSIVE FORMULA OF BEST FIT -- A COMPREHENSIVE NEW MACHINE PROGRAM. U. S. FOREST SERV. SOUTH. FOREST EXPT. STA. OCCAS. PAPER 158, NEW ORLEANS, LA. 9 PP.
- (*5) LONGLEY, J. W. 1967. AN APPRAISAL OF LEAST SQUARES PROGRAMS FOR THE ELECTRONIC COMPUTER FROM THE POINT OF VIEW OF THE USER. JOUR. AMER. STATIS. ASSOC. 62, PP. 819-829.
- (*6) NEELY, P. M. 1966. COMPARISON OF SEVERAL ALGORITHMS FOR COMPUTATION OF MEANS, STANDARD DEVIATIONS AND CORRELATION COEFFICIENTS. COMMUNICATIONS OF THE ACM 9, PP. 496-499.
- (*7) RODDEN, B. E. 1967. ERROR-FREE METHODS FOR STATISTICAL COMPUTATIONS. COMMUNICATIONS OF THE ACM 10, PP. 179-180.
- (*8) WALKER, H. M., AND J. LEV. 1953. STATISTICAL INFERENCE. HENRY HOLT AND CO., NEW YORK CITY. 510 PP.
- (*9) WELFORD, B. P. 1962. NOTES ON A METHOD FOR CALCULATING CORRECTED SUMS OF SQUARES AND PRODUCTS. TECHNOMETRICS 4, PP. 419-420.

=====
APPENDIX A
=====

'REX' SOURCE DECK ARRANGEMENTS AND MODIFICATIONS NEEDED BY DIFFERENT COMPUTERS
=====

ARRANGEMENT OF PROGRAM DECKS FOR INPUT WITH APPROPRIATE OVERLAY CONTROL CARDS
FOR USE ON IBM 7090-7094 UNDER IBSYS

=====

\$JOB	XX	5,150,800	GROSENBAUGH-59.REX	
\$IBJOB	REX	MAP		
\$IBFTC	TRNXHH	DECK		TRNX 0
\$IBMAP	BUFK	DECK		BUFK 0
\$IBMAP	UN04	DECK		UN04 0
\$IBMAP	UN08	DECK		UN08 0
\$IBFTC	REX	DECK		REX 0
\$IBFTC	BLRM	DECK		BLRM 0
		DATA MRE/ 5/,MPR/ 6/,MPU/ 7/,JW/ 4/,JX/ 8/,MEOF/0/		BLRM 44
		(ASSIGNMENTS APPROPRIATE TO INSTALLATION I/O CONFIGURATION)		
\$ORIGIN		ABLE		
\$IBFTC	PALMH	DECK		PALM 0
\$IBFTC	MATXH	DECK		MATX 0
\$ORIGIN		ABLE,REW		
\$IBFTC	SKRNHH	DECK		SKRN 0
\$ORIGIN		ABLE,REW		
\$IBFTC	CBXRHH	DECK		CBXR 0
\$ENTRY		REX		
\$DATA				
		(FOLLOWED BY APPROPRIATE DATA DECK)		
\$EOF				

ARRANGEMENT OF PROGRAM DECKS FOR INPUT WITH APPROPRIATE CHAIN CONTROL CARDS
AND CHANGES FOR USE ON IBM 7040-7044 UNDER IBSYS

```

=====
$JOB      T      592026      GROSENBAUGH-8      18      20REXSKN      26
$*SCRATCH      L2 AND L3
$OPEN      S.SU02=I01,S.SU03=I02
$IBJOB REX      MAP
$CHAIN REX      U04
$NAME      SJXIT=S.JXIT
$IBFTC TRNXHH DECK                                TRNX      0
$IBMAP BUFK DECK                                BUFK      0
$IBFTC BLRM DECK                                BLRM      0
      DATA MRE/ 5/,MPR/ 6/,MPU/ 7/,JW/ 2/,JX/ 3/,MEOF/0/
      (ASSIGNMENTS APPROPRIATE TO INSTALLATION I/O CONFIGURATION)
$IBFTC REX DECK                                REX      0
      CALL CHAIN(1)                                REX      67
      CALL CHAIN(1)                                REX      70
      CALL CHAIN(2)                                REX      75
      200 CALL CHAIN(1)                            REX      83
      250 CALL CHAIN(3)                            REX      85
      1000 CALL SJXIT                                REX     102
$ENTRY      REX
$LINK LINK1
$IBFTC PALMHH DECK                                PALM      0
C      SUBROUTINE PALM                                PALM      1
      231(NONFIXED) REGRESSIONS (CURRENTLY MUST NOT EXCEED LPP= 8200/(NY+1)
      200 LPP= 8200/NYY                                PALM     63
      CALL CHNXIT                                PALM    284
      CALL CHNXIT                                PALM    375
$IBFTC MATXHH DECK                                MATX      0
$ENTRY
$LINK LINK2
$IBFTC SKRNHH DECK                                SKRN      0
C      SUBROUTINE SKRN                                SKRN      1
      1A( 8200),ORD(2),RMS(2),DFS(2),VOL(2),DEN(6)
      CALL CHNXIT                                SKRN     11
      CALL CHNXIT                                SKRN    337
$ENTRY
$LINK LINK3
$IBFTC CBXRHH DECK                                CBXR      0
C      SUBROUTINE CBXR                                CBXR      1
      261 CALL CHNXIT                                CBXR    308
$ENTRY
$ENDCH
      (FOLLOWED BY APPROPRIATE DATA DECK)
$IBSYS

```

ARRANGEMENT OF PROGRAM DECKS FOR INPUT WITH APPROPRIATE OVERLAY CONTROL CARDS
AND CHANGES FOR USE ON CDC 6400-6600 UNDER SCOPE 3.0

JXXXXAA,7,300,55000,XXXX,GROSENBAUGH-51.REX

RUN(P,55000,,,,,12000,1,1)

LGO.

EXIT.

DMP(101,3336)

DMP(3336,6062)

CDC RECORD SEPARATOR CARD, WITH 7,8,9 PUNCHED IN COLUMN ONE.

OVERLAY(REX,0,0)

OPROGRAM REX(INPUT,OUTPUT,PUNCH,TAPE5=INPUT,TAPE6=OUTPUT,
1TAPE7=PUNCH,TAPE4,TAPE8)

\$IBFTC REX DECK

CALL OVERLAY(4HREX0,1,0,6HRECALL)

CALL OVERLAY(4HREX0,1,0,6HRECALL)

CALL OVERLAY(4HREX0,2,0,6HRECALL)

200 CALL OVERLAY(4HREX0,1,0,6HRECALL)

250 CALL OVERLAY(4HREX0,3,0,6HRECALL)

\$IBFTC BLRM DECK

DATA MRE/ 5/,MPR/ 6/,MPU/ 7/,JW/ 4/,JX/ 8/,MEOF/0/

(ASSIGNMENTS APPROPRIATE TO PROGRAM CARD I/O EQUIVALENCES)

\$IBFTC TRNXHH DECK

OVERLAY(1,0)

\$IBFTC PALMHH DECK

C SUBROUTINE PALM

231(NONFIXED) REGRESSIONS (CURRENTLY MUST NOT EXCEED LPP= 8200/(NY+1))

200 LPP= 8200/NYY

\$IBFTC MATXHH DECK

REAL SS(1326),Z(51)

OVERLAY(2,0)

\$IBFTC SKRNHH DECK

C SUBROUTINE SKRN

1A(8200),ORD(2),RMS(2),OFS(2),VOL(2),DEN(6)

OVERLAY(3,0)

\$IBFTC CBXRHH DECK

C SUBROUTINE CBXR

OREAL A(51,51),B(51,8),DD(51),XX(51),YY(8),P,PP,PPP,QQ

4DZ/.0E 00/,D1/.1E 01/

DD(I)= SQRT (DD(I))

YY(JJ)= SQRT (DD(I))

A(J,J)= SQRT (A(J,J)-P)

CDC RECORD SEPARATOR CARD, WITH 7,8,9 PUNCHED IN COLUMN ONE.

(FOLLOWED BY APPROPRIATE DATA DECK)

CDC END-OF-FILE CARD, WITH 6,7,8,9 PUNCHED IN COLUMN ONE.

REX	0
REX	67
REX	70
REX	75
REX	83
REX	85
BLRM	0
BLRM	44
TRNX	0
PALM	0
PALM	1
PALM	63
PALM	284
MATX	0
MATX	6
SKRN	0
SKRN	1
SKRN	11
CBXR	0
CBXR	1
CBXR	10
CBXR	54
CBXR	152
CBXR	160
CBXR	277

ARRANGEMENT OF PROGRAM DECKS FOR PART SOURCE AND PART BINARY INPUTS TO
CDC 6400-6600 UNDER SCOPE 3.0

=====

JXXXXAA,7,300,55000.XXXX,GROSENBAUGH-51.REX
 RUN(S,55000,,,,,12000,1,1)
 LOAD(INPUT)
 LOAD(LGO)
 LOAD(INPUT)
 EXECUTE,REX.
 CDC RECORD SEPARATOR CARD, WITH 7,8,9 PUNCHED IN COLUMN ONE.
 \$IBFTC TRNXHH DECK (TRANSGENERATION SOURCE PROGRAM SUPPLIED BY USER) TRNX 0
 CDC RECORD SEPARATOR CARD, WITH 7,8,9 PUNCHED IN COLUMN ONE.
 (FOLLOWED BY BINARY DECKS FOR ALL OF 'REX' LINK (0,0) EXCEPT 'TRNX')
 CDC RECORD SEPARATOR CARD, WITH 7,8,9 PUNCHED IN COLUMN ONE.
 (FOLLOWED BY BINARY DECKS FOR ALL OF 'REX' LINKS (1,0),(2,0),(3,0))
 CDC RECORD SEPARATOR CARD, WITH 7,8,9 PUNCHED IN COLUMN ONE.
 (FOLLOWED BY APPROPRIATE DATA DECK)
 CDC END-OF-FILE CARD, WITH 6,7,8,9 PUNCHED IN COLUMN ONE.

ARRANGEMENT OF PROGRAM DECKS FOR INPUT WITH APPROPRIATE CONTROL CARDS
AND CHANGES FOR USE WITH IBM 360 UNDER OPERATING SYSTEM 360

```

//JYYYYRX JOB (YYYY,50,50,1000),GROSENBAUGH,MSGLEVEL=1
//CLG EXEC FORTGCLG,PARM.FORT='DECK,LOAD,SOURCE,BCD,NAME=REX',      X1
//          PARM.LKED='XREF,LET,LIST,OVLY',COND.GO=(16,LT)          2
//FORT.SYSIN DD *
  (USE DECKS DENOTED BY FOLLOWING LABELS, BUT DELETE ALL $IBMAP AND $IBFTC CARDS)
C REX--REGRESSION EXECUTIVE PROGRAM (GROSENBAUGH 05-01-67)          REX 1
  BLOCK DATA                                                     BLRM 1
  DATA MRE/ 5/,MPR/ 6/,MPU/ 7/,JW/ 4/,JX/ 8/,MEDF/0/             BLRM 44
  (ASSIGNMENTS APPROPRIATE TO INSTALLATION I/O CONFIGURATION AND DD CARDS)
  SUBROUTINE TRNX                                                  TRNX 1
  SUBROUTINE PALM                                                  PALM 1
  231(NONFIXED) REGRESSIONS (CURRENTLY MUST NOT EXCEED LPP= 8200/(NY+1)) PALM 63
  200 LPP= 8200/NYY                                               PALM 284
  SUBROUTINE MATX                                                  MATX 1
  SUBROUTINE SKRN                                                  SKRN 1
  1A( 8200),ORD(2),RMS(2),OFS(2),VOL(2),DEN(6)                   SKRN 11
  SUBROUTINE CBXR                                                  CBXR 1
/*
//LKED.SYSIN DD *
  OVERLAY ABL
  INSERT PALM#,MATX#
  OVERLAY ABL
  INSERT SKRN#
  OVERLAY ABL
  INSERT CBXR#
  ENTRY REX
/*
//GO.FT04F001 DD DSNAME=&TAPE4,UNIT=SYSSQ,SPACE=(CYL,(20,20)),    X1
//          DISP=(NEW,DELETE)                                     2
//GO.FT08F001 DD DSNAME=&TAPE8,UNIT=SYSSQ,SPACE=(CYL,(20,20)),    X1
//          DISP=(NEW,DELETE)                                     2
//GO.SYSIN DD *
  (FOLLOWED BY INPUT DATA DECK PUNCHED ACCORDING TO EBCDIC CODE)
/*
//

```

COND.GO=(16,LT) ON EXEC CARD IS NEEDED ONLY BECAUSE OF BUG IN IBM LINKEDITOR.
IN THE EVENT THAT 131K-BYTE COMPILER IS UNABLE TO COMPILE SUBROUTINE CBXR,
IT CAN BE COMPILED ON A LARGER MACHINE AND OBJECT DECK OBTAINED. OBJECT DECKS
FOR THE REDUCED VERSION OF REX SHOWN ABOVE WILL RUN ON A 131K-BYTE COMPUTER.
PARENS AND PLUS SIGN MAY NOT PRINT PROPERLY UNLESS EBCDIC SOURCE DECKS ARE USED.

=====

APPENDIX B

=====

EXAMPLES OF INPUT DATA THAT TEST 'REX' PROCESSING AND OUTPUT OPTIONS

=====

MODIFICATION TO SUBROUTINE TRNX NEEDED BY TEST PROBLEMS 'ABST' THROUGH 'HCBW'
FITTING SPECIFIC REGRESSIONS AND COMPARING RELATIVE GOODNESS OF FIT
(TEST PROBLEMS 'ABST' THRO 'HCBW' CAN USE SAME VERSION OF 'TRNX')

*ALTER 18,21	TRNX	
C KEEPS VARIABLES AND SEQUENCE UNCHANGED WITH WEIGHT(IF ANY) FOLLOWING	TRNX	18
DO 1 J=1,NVS	TRNX	19
1 X(J)=D(J)	TRNX	20
W=D(NVR)	TRNX	21
*ENDAL	TRNX	

EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION

30 8=
Y=
(8F4.0)

-2	0	0	2	0	0	1	-2
-2	0	0	2	0	0	1	2
6	-2	0	0	0	0	4	2
6	-2	0	0	0	0	4	6
8	0	0	0	0	0	8	6
8	0	0	0	0	0	8	10
10	2	50	-50	0	0	12	10
10	2	-50	50	0	0	12	14
16	0	0	0	0	0	16	14
16	0	0	0	0	0	16	18
7	2	0	0	0	9	0	7
7	2	0	0	0	9	0	11
19	0	-3	0	0	16	0	14
19	0	-3	0	0	16	0	18
23	2	0	0	0	25	0	23
23	2	0	0	0	25	0	27
46	-10	0	0	0	36	0	34
46	-10	0	0	0	36	0	38
34	2	0	0	0	36	0	34
34	2	0	0	0	36	0	38
29	0	0	-4	25	0	0	23
29	0	0	-4	25	0	0	27
14	2	-50	50	16	0	0	14
14	2	50	-50	16	0	0	18
27	-2	0	0	25	0	0	23
27	-2	0	0	25	0	0	27
34	2	0	0	36	0	0	34
34	2	0	0	36	0	0	38
42	2	5	0	49	0	0	47
42	2	5	0	49	0	0	51

DONE

ABST 1
100111ABST 2
ABST 3
ABST 4
ABST 5

01	ABST
02	ABST
03	ABST
04	ABST
05	ABST
06	ABST
07	ABST
08	ABST
09	ABST
10	ABST
11	ABST
12	ABST
13	ABST
14	ABST
15	ABST
16	ABST
17	ABST
18	ABST
19	ABST
20	ABST
21	ABST
22	ABST
23	ABST
24	ABST
25	ABST
26	ABST
27	ABST
28	ABST
29	ABST
30	ABST
31	ABST

REX 5-01-67

[illegible]

REX 5-01-67

PAGE 27

UNCORRECTED MATRIX INPUT USED CORRECTED (THRO MEAN)

30 12
GSG GG GYY

0.94550000E 04 0.16000000E 04 0.99000000E 03 0.33500000E 04
0.64500000E 04 0.30300000E 04 0.13450000E 05 0.10200000E 05
0.40900000E 05 0.10224000E 05 0.94400000E 04 0.55990000E 04
0.27500000E 03 0.15000000E 03 0.55000000E 03 0.11250000E 04
0.45000000E 03 0.22500000E 04 0.16500000E 04 0.67500000E 04
0.15000000E 04 0.15970000E 04 0.91000000E 03 0.18000000E 03
0.45000000E 03 0.55000000E 03 0.60000000E 03 0.16500000E 04
0.15000000E 04 0.55000000E 04 0.21240000E 04 0.99000000E 03
0.72700000E 03 0.16500000E 04 0.22500000E 04 0.15000000E 04
0.67500000E 04 0.55000000E 04 0.22500000E 05 0.53100000E 04
0.33440000E 04 0.22570000E 04 0.48950000E 04 0.16500000E 04
0.97900000E 04 0.67500000E 04 0.29370000E 05 0.55000000E 04
0.64350000E 04 0.37080000E 04 0.21240000E 04 0.55000000E 04
0.53100000E 04 0.19470000E 05 0.78000000E 04 0.30300000E 04
0.22810000E 04 0.29370000E 05 0.22500000E 05 0.97900000E 05
0.19470000E 05 0.13420000E 05 0.93130000E 04 0.19470000E 05
0.79650000E 05 0.19500000E 05 0.10170000E 05 0.74030000E 04
0.34656600E 06 0.71500000E 05 0.40750000E 05 0.30619000E 05
0.29340000E 05 0.10224000E 05 0.78910000E 04 0.94550000E 04
0.56190000E 04 0.48060000E 04 0.46500000E 03 0.75000000E 02
0.60000000E 02 0.15000000E 03 0.27500000E 03 0.18000000E 03
0.55000000E 03 0.45000000E 03 0.16500000E 04 0.60000000E 03
0.46500000E 03 0.32200000E 03 0.30000000E 02

DONE

BREX 1
120000BREX 2
BREX 3
BREX 4
BREX 5
BREX1 1
BREX1 2
BREX1 3
BREX1 4
BREX1 5
BREX1 6
BREX1 7
BREX1 8
BREX1 9
BREX1 10
BREX1 11
BREX1 12
BREX1 13
BREX1 14
BREX1 15
BREX1 16
BREX1 17
BREX1 18
BREX1 19
BREX1 20
BREX1 21
BREX1 22
BREX1 23
BREX1 24

REX 5-01-67

PAGE 27

UNCORRECTED MATRIX INPUT USED CORRECTED (THRO MEAN)

30 12
GSG GG GYY

BREX 1
120000BREX 2
BREX 3
BREX 4
BREX 5
BREX1 1
BREX1 2
BREX1 3
BREX1 4
BREX1 5
BREX1 6
BREX1 7
BREX1 8
BREX1 9
BREX1 10
BREX1 11
BREX1 12
BREX1 13
BREX1 14
BREX1 15
BREX1 16
BREX1 17
BREX1 18
BREX1 19
BREX1 20
BREX1 21
BREX1 22
BREX1 23
BREX1 24

0.94550000E 04	0.16000000E 04	0.99000000E 03	0.33500000E 04
0.64500000E 04	0.30300000E 04	0.13450000E 05	0.10200000E 05
0.40900000E 05	0.10224000E 05	0.94400000E 04	0.55990000E 04
0.27500000E 03	0.15000000E 03	0.55000000E 03	0.11250000E 04
0.45000000E 03	0.22500000E 04	0.16500000E 04	0.67500000E 04
0.15000000E 04	0.15970000E 04	0.91000000E 03	0.18000000E 03
0.45000000E 03	0.55000000E 03	0.60000000E 03	0.16500000E 04
0.15000000E 04	0.55000000E 04	0.21240000E 04	0.99000000E 03
0.72700000E 03	0.16500000E 04	0.22500000E 04	0.15000000E 04
0.67500000E 04	0.55000000E 04	0.22500000E 05	0.53100000E 04
0.33440000E 04	0.22570000E 04	0.48950000E 04	0.16500000E 04
0.97900000E 04	0.67500000E 04	0.29370000E 05	0.55000000E 04
0.64350000E 04	0.37080000E 04	0.21240000E 04	0.55000000E 04
0.53100000E 04	0.19470000E 05	0.78000000E 04	0.30300000E 04
0.22810000E 04	0.29370000E 05	0.22500000E 05	0.97900000E 05
0.19470000E 05	0.13420000E 05	0.93130000E 04	0.19470000E 05
0.79650000E 05	0.19500000E 05	0.10170000E 05	0.74030000E 04
0.34656600E 06	0.71500000E 05	0.40750000E 05	0.30619000E 05
0.29340000E 05	0.10224000E 05	0.78910000E 04	0.94550000E 04
0.56190000E 04	0.48060000E 04	0.46500000E 03	0.75000000E 02
0.60000000E 02	0.15000000E 03	0.27500000E 03	0.18000000E 03
0.55000000E 03	0.45000000E 03	0.16500000E 04	0.60000000E 03
0.46500000E 03	0.32200000E 03	0.30000000E 02	

DONE

S)
1
1
44
1
1
63
84
1
1
11
1

0.

CORRECTED MATRIX INPUT USED CORRECTED WITH MAX. NO. SETS PUT AT 5				CREX	
30	12	5		140000	CREX
YY					
0.22475000E 04	0.43750000E 03	0.60000000E 02	0.10250000E 04		CREX 1
0.21875000E 04	0.24000000E 03	0.49250000E 04	0.32250000E 04		CREX 2
0.15325000E 05	0.92400000E 03	0.22325000E 04	0.60800000E 03		CREX 3
0.87500000E 02	-0.00000000E-38	0.17500000E 03	0.43750000E 03		CREX 4
-0.00000000E-38	0.87500000E 03	0.52500000E 03	0.26250000E 04		CREX 5
-0.00000000E-38	0.43450000E 03	0.10500000E 03	0.60000000E 02		CREX 6
0.15000000E 03	-0.00000000E-38	0.24000000E 03	0.55000000E 03		CREX 7
0.60000000E 03	0.22000000E 04	0.92400000E 03	0.60000000E 02		CREX 8
0.83000000E 02	0.90000000E 03	0.87500000E 03	0.60000000E 03		CREX 9
0.40000000E 04	0.32500000E 04	0.14250000E 05	0.23100000E 04		CREX 10
0.10190000E 04	0.64700000E 03	0.23741667E 04	-0.00000000E-38		CREX 11
0.47483333E 04	0.26250000E 04	0.14245000E 05	-0.00000000E-38		CREX 12
0.21725000E 04	0.75633333E 03	0.10440000E 04	0.22000000E 04		CREX 13
0.26100000E 04	0.95700000E 04	0.42000000E 04	0.24000000E 03		CREX 14
0.34900000E 03	0.19286667E 05	0.14250000E 05	0.67650000E 05		CREX 15
0.84700000E 04	0.48950000E 04	0.34096667E 04	0.12720000E 05		CREX 16
0.54900000E 05	0.10500000E 05	0.31950000E 04	0.25730000E 04		CREX 17
0.25581600E 06	0.38500000E 05	0.15175000E 05	0.12909000E 05		CREX 18
0.17340000E 05	0.92400000E 03	0.14510000E 04	0.22475000E 04		CREX 19
0.62800000E 03	0.13498667E 04	0.46500000E 03	0.75000000E 02		CREX 20
0.60000000E 02	0.15000000E 03	0.27500000E 03	0.18000000E 03		CREX 21
0.55000000E 03	0.45000000E 03	0.16500000E 04	0.60000000E 03		CREX 22
0.46500000E 03	0.32200000E 03	0.30000000E 02			CREX 23
					CREX 24

DONE

CALCULATING PARTICULAR REGRESSION FROM LARGER MATRIX (MAX. NO. Y'S) DREX

$$30 \quad 18 \quad 18 =$$

200000DREX	2
------------	---

XXXXXXXXXXYYYYYYYYY=

DREX

(18F4.0)

DREX

DREX 5

																		DREX	
1	0	0	0	0	0	0	0	0	9	2	2	2	2	2	2	2	2	01DREX	
2	0	1	0	0	1	0	0	0	1	23	1	1	1	1	1	1	1	02DREX	
3	0	2	0	0	4	0	0	0	8	17	4	4	4	4	4	4	4	03DREX	
4	0	3	0	0	9	0	0	0	27	5	3	3	3	3	3	3	3	04DREX	
5	0	4	0	0	16	0	0	0	64	8	6	6	6	6	6	6	6	05DREX	
6	1	0	0	1	0	0	0	0	0	9	5	5	5	5	5	5	5	06DREX	
7	1	1	1	1	1	1	1	1	8	8	8	8	8	8	8	8	8	07DREX	
8	1	2	2	1	4	2	4	4	8	5	7	7	7	7	7	7	7	08DREX	
9	1	3	3	1	9	3	9	9	27	15	10	10	10	10	10	10	10	09DREX	
10	1	4	4	1	16	4	16	16	64	2	9	9	9	9	9	9	9	10DREX	
11	2	0	0	4	0	0	0	0	0	3	12	12	12	12	12	12	12	11DREX	
12	2	1	2	4	1	4	2	4	1	3	11	11	11	11	11	11	11	12DREX	
13	2	2	4	4	4	8	8	16	8	7	14	14	14	14	14	14	14	13DREX	
14	2	3	6	4	9	12	18	36	27	4	13	13	13	13	13	13	13	14DREX	
15	2	4	8	4	16	16	32	64	64	9	16	16	16	16	16	16	16	15DREX	
16	3	0	0	9	0	0	0	0	0	7	15	15	15	15	15	15	15	16DREX	
17	3	1	3	9	1	9	3	9	1	11	18	18	18	18	18	18	18	17DREX	
18	3	2	6	9	4	18	12	36	8	7	17	17	17	17	17	17	17	18DREX	
19	3	3	9	9	9	27	27	81	27	4	20	20	20	20	20	20	20	19DREX	
20	3	4	12	9	16	36	48	144	64	18	19	19	19	19	19	19	19	20DREX	
21	4	0	0	16	0	0	0	0	0	9	22	22	22	22	22	22	22	21DREX	
22	4	1	4	16	1	16	4	16	1	5	21	21	21	21	21	21	21	22DREX	
23	4	2	8	16	4	32	16	64	8	11	24	24	24	24	24	24	24	23DREX	
24	4	3	12	16	9	48	36	144	27	14	23	23	23	23	23	23	23	24DREX	
25	4	4	16	16	16	64	64	256	64	23	26	26	26	26	26	26	26	25DREX	
26	5	0	0	25	0	0	0	0	0	7	25	25	25	25	25	25	25	26DREX	
27	5	1	5	25	1	25	5	25	1	18	28	28	28	28	28	28	28	27DREX	
28	5	2	10	25	4	50	20	100	8	13	27	27	27	27	27	27	27	28DREX	
29	5	3	15	25	9	75	45	225	27	19	30	30	30	30	30	30	30	29DREX	
30	5	4	20	25	16	100	80	400	64	29									

PAGE 30

REX 5-01-67

SPECIFIC REGRESSION WITH NONTERMINAL Y, USING DATA INPUT EARLIER		EREX	1
		280000EREX	2
YXXXXXXXXX	=	EREX	3
		EREX	4
		EREX	5
DONE		EREXLAST	

REX 5-01-67

PAGE 31

MINIMUM WEIGHTED REGRESSION TEST FOR REX (OLD XXR-CBY, THRO ORIGIN)					FCBW	
10	5	4W			110000	FCBW
YYW						FCBW
(5F4.0)						FCBW
82	-17	82	5	2		01 FCBW
85	-17	75	5	2		02 FCBW
89	-13	62	4	2		03 FCBW
98	-15	62	4	2		04 FCBW
77	-15	50	3	2		05 FCBW
62	-13	49	3	2		06 FCBW
133	-17	94	6	2		07 FCBW
114	-15	89	6	2		08 FCBW
110	-14	87	6	2		09 FCBW
80	-14	79	5	2		10 FCBW
DONE						11 FCBW

0
0
0
0
0
0
44

0
0

0
0

PAGE 32

REX 5-01-67

WEIGHTED UNCORRECTED MATRIX INPUT USED CORRECTED BY REX (THRO				MEAN)	GCBW	
10	4W			120000	GCBW	1
YYW					GCBW	2
					GCBW	3
					GCBW	4
					GCBW	5
0.90432000E 05	-0.14058000E 05	0.70119000E 05	0.45420000E 04		GCBW1	1
0.22720000E 04	-0.11049000E 05	-0.71200000E 03	0.55505000E 05		GCBW1	2
0.35930000E 04	0.23300000E 03	0.93000000E 03	-0.15000000E 03		GCBW1	3
0.72900000E 03	0.47000000E 02	0.10000000E 02			GCBW1	4
DONE					GCBW1	5

REX 5-01-67

PAGE 33

WEIGHTED CORRECTED MATRIX INPUT USED UNCORRECTED (THRO ORIGIN) HCBW 1
10 4W 050000HCBW 2
SGYYW HCBW 3
(4E16.8) HCBW 4
HCBW 5
0.39420000E 04 -0.10800000E 03 0.23220000E 04 0.17100000E 03 HCBW0 1
0.22000000E 02 -0.11400000E 03 -0.70000000E 01 0.23609000E 04 HCBW0 2
0.16670000E 03 0.12100000E 02 0.93000000E 03 -0.15000000E 03 HCBW0 3
0.72900000E 03 0.47000000E 02 0.10000000E 02 HCBW0 4
DONE HCBW0 5
DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE

MODIFICATION TO SUBROUTINE TRNX NEEDED BY TEST PROBLEMS 'AORX' THROUGH 'CORX'
 COVARIANCE ANALYSIS USING ORTHOGONAL POLYNOMIALS
 (TEST PROBLEMS 'AORX' THRO 'CORX' CAN USE SAME VERSION OF 'TRNX')

```

      *ALTER 18,21
C  COVARIANCE (6 GROUPS, 3 X'S, 2 Y'S, CORRECTED MATRIX, ORTHO. POLY.) TRNX
      X(1)=D(3)
      X(2)=D(6)
      X(3)=D(10)
      X(24)=D(11)
      X(25)=D(12)
      IF (D(2) .NE. 0.0) GO TO 20
      X(4)=-5./5.
      X(5)=+5./5.
      X(6)=-5./5.
      X(7)=+1./5.
      X(8)=-1./5.
      GO TO 70
20  IF (D(2) .NE. 1.0) GO TO 30
      X(4)=-3./5.
      X(5)=-1./5.
      X(6)=+7./5.
      X(7)=-3./5.
      X(8)=+5./5.
      GO TO 70
30  IF (D(2) .NE. 2.0) GO TO 40
      X(4)=-1./5.
      X(5)=-4./5.
      X(6)=+4./5.
      X(7)=+2./5.
      X(8)=-10./5.
      GO TO 70
40  IF (D(2) .NE. 3.0) GO TO 50
      X(4)=+1./5.
      X(5)=-4./5.
      X(6)=-4./5.
      X(7)=+2./5.
      X(8)=+10./5.
      GO TO 70
50  IF (D(2) .NE. 4.0) GO TO 60
      X(4)=+3./5.
      X(5)=-1./5.
      X(6)=-7./5.
      X(7)=-3./5.
      X(8)=-5./5.
      GO TO 70

```

```

60 X(4)=+5./5.
   X(5)=+5./5.
   X(6)=+5./5.
   X(7)=+1./5.
   X(8)=+1./5.
70 K=8
   DO 75 J=1,3
   DO 75 I=1,5
   K=K+1
   L=I+3
75 X(K)=X(J)*X(L)
   *ENDAL

```

TRNX

```

COVARIANCE (6 GROUPS, 3 X'S, 2 Y'S) REAL
30 12 25=
XXX YY=
(12F4.0)

  1  0  0  0  0  0  0  0  0  0  2  9
  2  0  1  0  0  1  0  0  0  1  1 23
  3  0  2  0  0  4  0  0  0  8  4 17
  4  0  3  0  0  9  0  0  0 27  3  5
  5  0  4  0  0 16  0  0  0 64  6  8
  6  1  0  0  1  0  0  0  0  0  5  9
  7  1  1  1  1  1  1  1  1  1  8  8
  8  1  2  2  1  4  2  4  4  8  7  5
  9  1  3  3  1  9  3  9  9 27 10 15
 10  1  4  4  1 16  4 16 16 64  9  2
 11  2  0  0  4  0  0  0  0  0 12  3
 12  2  1  2  4  1  4  2  4  1 11  3
 13  2  2  4  4  4  8  8 16  8 14  7
 14  2  3  6  4  9 12 18 36 27 13  4
 15  2  4  8  4 16 16 32 64 64 16  9
 16  3  0  0  9  0  0  0  0  0 15  7
 17  3  1  3  9  1  9  3  9  1 18 11
 18  3  2  6  9  4 18 12 36  8 17  7
 19  3  3  9  9  9 27 27 81 27 20  4
 20  3  4 12  9 16 36 48 144 64 19 18
 21  4  0  0 16  0  0  0  0  0 22  9
 22  4  1  4 16  1 16  4 16  1 21  5
 23  4  2  8 16  4 32 16 64  8 24 11
 24  4  3 12 16  9 48 36 144 27 23 14
 25  4  4 16 16 16 64 64 256 64 26 23
 26  5  0  0 25  0  0  0  0  0 25  7
 27  5  1  5 25  1 25  5 25  1 28 18
 28  5  2 10 25  4 50 20 100  8 27 13
 29  5  3 15 25  9 75 45 225 27 30 19
 30  5  4 20 25 16 100 80 400 64 29 29
DONE
COVARIANCE (6 GROUPS, 3 X'S, 2 Y'S) REAL+DUMMY
XXXXXXXXX YY=

DONE
COVARIANCE (6 GROUPS, 3 X'S, 2 Y'S) REAL+DUMMY+INTERACTION
XXXXXXXXXXXXXXXXXXXXXXXXXXXXYY=

DONE
DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE DONE

```

```

AORX 1
200000AORX 2
AORX 3
AORX 4
AORX 5
01 AREX
02 AREX
03 AREX
04 AREX
05 AREX
06 AREX
07 AREX
08 AREX
09 AREX
10 AREX
11 AREX
12 AREX
13 AREX
14 AREX
15 AREX
16 AREX
17 AREX
18 AREX
19 AREX
20 AREX
21 AREX
22 AREX
23 AREX
24 AREX
25 AREX
26 AREX
27 AREX
28 AREX
29 AREX
30 AREX
31 AREX
BORX 1
280000BORX 2
BORX 3
BORX 4
BORX 5
BORXLAST
CORX 1
280000CORX 2
CORX 3
CORX 4
CORX 5
CORXLAST

```

MODIFICATION TO SUBROUTINE TRNX NEEDED BY TEST PROBLEMS 'AIRX' THROUGH 'C1RX'
 COVARIANCE ANALYSIS USING DUMMY UNITY VARIABLES
 (TEST PROBLEMS 'AIRX' THRO 'C1RX' CAN USE SAME VERSION OF 'TRNX')

```

        *ALTER 18,21
C  COVARIANCE (6 GROUPS, 3 X'S, 2 Y'S, UNCORRECTED MATRIX, UNITY VAR.)
    X(1)=D(3)
    X(2)=D(6)
    X(3)=D(10)
    X(4)=1.
    X(29)=D(11)
    X(30)=D(12)
    DO 5 I=5,28
5  X(I)=0.0
    K=4.0*(D(2)+1.)
    DO 6 I=1,4
    J=K+I
6  X(J)=X(I)
    *ENDAL
    
```

TRNX

TRNX

REX 5-01-67

[illegible]

=====
APPENDIX C
=====

ILLUSTRATIVE OUTPUT FROM INITIAL INPUT (PROBLEM LABELLED 'ABST')
USING 'REX' MAJOR PROCESSING OPTION 1
(PRINTOUT SLIGHTLY CONDENSED TO SAVE SPACE)

=====

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===  30 -0  8= -0  100111=====
===  Y=  =====

```

PARAMETER NUMBERS SPECIFIED IMPLICITLY OR EXPLICITLY BY CONTROL CARDS 2,3.
 PARENTHEZIZED ALLOWABLE LIMITS HAVE BEEN EXCEEDED WHEN PARAMETER NUMBERS BELOW ARE ASTERISKED.
 A BLANK COLUMN ONE ON THIRD CONTROL CARD IS INTERPRETED AS MEANING ALL X'S ARE PUNCHED WITH G'S.

```

(3,NOB, NOB) 30 = NUMBER OF OBSERVATIONS
(2,NVR, 51) 8 = NUMBER OF RAW VARIABLES (ALL TYPES INCLUDING WEIGHT IF NOT UNITY)
(2,NVS, 50) 8 = NUMBER OF FINAL VARIABLES (Y'S AND ALL X'S, BUT NOT INCLUDING WEIGHT)
(1,NY, 8) 1 = NUMBER OF DEPENDENT VARIABLES (Y'S)
(1,NXX, 49) 7 = TOTAL NUMBER OF INDEPENDENT VARIABLES (X'S IMPLIED BY ANY G, S, OR BLANK PRIOR TO FIRST Y)
(0,NXA, 49) 7 = NUMBER OF NONFIXED X'S (OCCURRENCE GOVERNED BY COMBINATORIAL RULES)
(0,NFA, 49) 0 = NUMBER OF FIXED X'S (ALWAYS INCLUDED WITH EACH COMBINATION OF NONFIXED X'S)
(1,NSET, 47) 7 = TOTAL NUMBER OF SETS OF X'S (G OR S STARTS SET, AND MEMBERS APPEAR OR DISAPPEAR TOGETHER)
(0,NXS, 30) 7 = NUMBER OF SETS OF NONFIXED X'S (ANY G, OR S AFTER OCCURRENCE OF FIRST G, STARTS SET)
(0,NF, 17) 0 = NUMBER OF SETS OF FIXED X'S (ANY S PRIOR TO OCCURRENCE OF FIRST G STARTS SET)
(0,NGX, 17) 7 = NUMBER OF GROUPS OF NONFIXED SETS (G STARTS GROUP, COMBINATIONS LIMITED TO ONE SET PER GROUP)
(0,LM, NGX) 7 = ARBITRARY MAXIMUM OR LIMITING NUMBER OF GROUPS OR SETS ALLOWED IN COMBINATION)
(0,LGK, LPP) 127 = TOTAL NUMBER OF GENERATED (NONFIXED) REGRESSIONS (CURRENTLY MUST NOT EXCEED LPP=16400/(NY+1))
(MX,NXX, MX) MACHINE ERROR--SORT OR SUM
(MS,NSET, MS) MACHINE ERROR--SORT OR SUM

```

```

GROUP LETTER  A  B  C  D  E  F  G
SET ORDINALS  1  2  3  4  5  6  7
X SUBSCRIPTS  1  2  3  4  5  6  7

```

DATA INPUT FORMAT
 (8F4.0)

LISTING OF FIRST TWO OBSERVATION VECTORS (AFTER TRANSFORMATIONS, IF ANY).

```

-0.2000000E 01 0.0000000E-38 0.0000000E-38 0.2000000E 01 0.0000000E-38 0.0000000E-38 0.1000000E 01-0.2000000E 01
0.1000000E 01

-0.2000000E 01 0.0000000E-38 0.0000000E-38 0.2000000E 01 0.0000000E-38 0.0000000E-38 0.1000000E 01 0.2000000E 01
0.1000000E 01

```

LISTING OF LAST OBSERVATION VECTOR--IF ABSENT, DATA CARDS ARE INCONSISTENT WITH NUMBER OF OBSERVATIONS
 ON SECOND CONTROL CARD (OR WRONG FORMAT HAS BEEN USED).

```

0.4200000E 02 0.2000000E 01 0.5000000E 01 0.0000000E-38 0.4900000E 02 0.0000000E-38 0.0000000E-38 0.5100000E 02
0.1000000E 01

```

=== EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION ABST
 === 30 8 8= 7 100111====
 ===GGGGGGGY= =====

COEFFICIENT OF COLLINEARITY AND RATIOS OF MEAN SQUARED RESIDUALS FROM VARIOUS
 'YI' REGRESSIONS TO MEAN SQUARED RESIDUALS FROM CORRESPONDING MEAN 'YI',
 WITH ORDINALS OF NONFIXED SETS INVOLVED (ALL FIXED ARE IMPLICIT)

NO.	SETS	C.OF C.	RMSQR. Y1	ORDINALS OF SETS INVOLVED
1	0.30E 00	0.0975	1	
1	0.10E 01	1.0283	2	
1	0.10E 01	1.0339	3	
1	0.10E 01	1.0355	4	
1	0.71E 00	0.6075	5	
1	0.72E 00	0.8871	6	
1	0.77E 00	0.8926	7	
2	0.27E 00	0.0436	1 2	
2	0.30E 00	0.1009	1 3	
2	0.30E 00	0.1011	1 4	
2	0.15E 00	0.0772	1 5	
2	0.17E 00	0.0938	1 6	
2	0.20E 00	0.1011	1 7	
2	0.10E 01	1.0644	2 3	
2	0.10E 01	1.0661	2 4	
2	0.68E 00	0.5741	2 5	
2	0.62E 00	0.9160	2 6	
2	0.76E 00	0.9220	2 7	
2	0.11E-01	1.0057	3 4	
2	0.71E 00	0.6299	3 5	
2	0.72E 00	0.9174	3 6	
2	0.77E 00	0.9239	3 7	
2	0.71E 00	0.6300	4 5	
2	0.72E 00	0.9196	4 6	
2	0.77E 00	0.9254	4 7	
2	0.43E 00	0.1131	5 6	
2	0.47E 00	0.6045	5 7	
2	0.49E 00	0.8488	6 7	
3	0.27E 00	0.0452	1 2 3	
3	0.27E 00	0.0453	1 2 4	
3	0.10E 00	0.0434	1 2 5	
3	0.14E 00	0.0448	1 2 6	
3	0.18E 00	0.0452	1 2 7	
3	0.32E-02	0.0723	1 3 4	
3	0.15E 00	0.0801	1 3 5	
3	0.17E 00	0.0973	1 3 6	
3	0.20E 00	0.1048	1 3 7	

PAGE 42

REX 5-01-67

3	0.15E 00	0.0801	1	4	5
3	0.17E 00	0.0974	1	4	6
3	0.20E 00	0.1050	1	4	7
3	0.17E-01	0.0665	1	5	6
3	0.95E-01	0.0796	1	5	7
3	0.11E 00	0.0969	1	6	7
3	0.11E-01	1.0297	2	3	4
3	0.67E 00	0.5961	2	3	5
3	0.62E 00	0.9486	2	3	6
3	0.76E 00	0.9556	2	3	7
3	0.67E 00	0.5961	2	4	5
3	0.62E 00	0.9509	2	4	6
3	0.76E 00	0.9572	2	4	7
3	0.37E 00	0.1174	2	5	6
3	0.44E 00	0.5804	2	5	7
3	0.42E 00	0.8792	2	6	7
3	0.68E-02	0.6520	3	4	5
3	0.76E-02	0.8524	3	4	6
3	0.82E-02	0.8874	3	4	7
3	0.43E 00	0.1174	3	5	6
3	0.47E 00	0.6276	3	5	7
3	0.49E 00	0.8791	3	6	7
3	0.43E 00	0.1174	4	5	6
3	0.47E 00	0.6277	4	5	7
3	0.49E 00	0.8811	4	6	7
3	0.20E 00	0.0245	5	6	7
4	0.28E-02	0.0252	1	2	3 4
4	0.10E 00	0.0451	1	2	3 5
4	0.14E 00	0.0465	1	2	3 6
4	0.18E 00	0.0470	1	2	3 7
4	0.10E 00	0.0451	1	2	4 5
4	0.14E 00	0.0465	1	2	4 6
4	0.18E 00	0.0470	1	2	4 7
4	0.97E-02	0.0436	1	2	5 6
4	0.65E-01	0.0448	1	2	5 7
4	0.87E-01	0.0466	1	2	6 7
4	0.14E-02	0.0641	1	3	4 5
4	0.17E-02	0.0722	1	3	4 6
4	0.21E-02	0.0750	1	3	4 7
4	0.17E-01	0.0691	1	3	5 6
4	0.95E-01	0.0827	1	3	5 7
4	0.11E 00	0.1006	1	3	6 7
4	0.17E-01	0.0692	1	4	5 6
4	0.95E-01	0.0827	1	4	5 7
4	0.11E 00	0.1007	1	4	6 7
4	0.37E-02	0.0254	1	5	6 7
4	0.64E-02	0.6163	2	3	4 5
4	0.65E-02	0.8848	2	3	4 6
4	0.81E-02	0.9137	2	3	4 7
4	0.37E 00	0.1220	2	3	5 6
4	0.44E 00	0.6035	2	3	5 7

4	0.42E 00	0.9120	2	3	6	7
4	0.37E 00	0.1221	2	4	5	6
4	0.44E 00	0.6036	2	4	5	7
4	0.42E 00	0.9141	2	4	6	7
4	0.17E 00	0.0255	2	5	6	7
4	0.41E-02	0.1199	3	4	5	6
4	0.44E-02	0.6475	3	4	5	7
4	0.51E-02	0.8154	3	4	6	7
4	0.20E 00	0.0255	3	5	6	7
4	0.20E 00	0.0255	4	5	6	7

5	0.97E-03	0.0262	1	2	3	4	5
5	0.14E-02	0.0262	1	2	3	4	6
5	0.18E-02	0.0262	1	2	3	4	7
5	0.97E-02	0.0454	1	2	3	5	6
5	0.65E-01	0.0467	1	2	3	5	7
5	0.87E-01	0.0484	1	2	3	6	7
5	0.97E-02	0.0454	1	2	4	5	6
5	0.65E-01	0.0467	1	2	4	5	7
5	0.87E-01	0.0484	1	2	4	6	7
5	0.95E-03	0.0264	1	2	5	6	7
5	0.15E-03	0.0588	1	3	4	5	6
5	0.88E-03	0.0668	1	3	4	5	7
5	0.11E-02	0.0745	1	3	4	6	7
5	0.37E-02	0.0265	1	3	5	6	7
5	0.37E-02	0.0265	1	4	5	6	7
5	0.35E-02	0.1248	2	3	4	5	6
5	0.41E-02	0.6222	2	3	4	5	7
5	0.44E-02	0.8487	2	3	4	6	7
5	0.17E 00	0.0265	2	3	5	6	7
5	0.17E 00	0.0265	2	4	5	6	7
5	0.18E-02	0.0264	3	4	5	6	7

6	0.84E-04	0.0274	1	2	3	4	5	6
6	0.60E-03	0.0274	1	2	3	4	5	7
6	0.88E-03	0.0274	1	2	3	4	6	7
6	0.94E-03	0.0276	1	2	3	5	6	7
6	0.95E-03	0.0276	1	2	4	5	6	7
6	0.22E-04	0.0276	1	3	4	5	6	7
6	0.15E-02	0.0276	2	3	4	5	6	7

7	0.88E-07	0.0286	1	2	3	4	5	6	7
---	----------	--------	---	---	---	---	---	---	---

ORDINALS OF SETS IN REGRESSION WITH SMALLEST RELATIVE MEAN SQUARED RESIDUAL

Y1 0.20E 00 0.0245 5 6 7

MSQR ABOUT MEAN Y1= 0.19074023E 03
 MEAN Y1 = 0.20866667E 02

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===  30  8  8= -0 100111=====
===  XXXY=
VARIABLE SUBSCRIPT IN  MOMENT  MATRIX  (REGRESSION THROUGH MEAN)
XX          5          6          7          XX
=====
5  0.73658666E 04 -0.24562667E 04 -0.82546666E 03  5
6          0.51234666E 04 -0.66693333E 03  6
7          0.73786666E 03  7

```

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===  30  8  8= -0 100111=====
===  XXXY=
VARIABLE SUBSCRIPT IN  MOMENT  MATRIX  (REGRESSION THROUGH MEAN)
XY          8          XY
=====
5  0.41042667E 04  5
6  0.20165333E 04  6
7 -0.75106667E 03  7

```

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===  30  8  8= -0 100111=====
===  XXXY=
VARIABLE SUBSCRIPT IN  MOMENT  MATRIX INVERSE  (REGRESSION THROUGH MEAN)
C          5          6          7          U          C
=====
5  0.23958586E-03  0.16972008E-03  0.42143403E-03 -0.49441407E-02  5
6          0.34143514E-03  0.49848131E-03 -0.58480369E-02  6
7          0.22772862E-02 -0.14521333E-01  7
U          0.17036003E 00  U

```

```
===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===    30    8    8= -0                                         100111===
===  XXXY=                                                     =====
VARIABLE SUBSCRIPT IN REG. COEFF. MATRIX          (REGRESSION THROUGH MEAN)
B              8              B
=====
5    0.10090454E 01    5
6    0.10106991E 01    6
7    0.10244881E 01    7
U   -0.31167755E 00    U
```

```
===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===    30    8    8= -0                                         100111===
===  XXXY=                                                     =====
VARIABLE SUBSCRIPT IN B-VARIANCE MATRIX          (REGRESSION THROUGH MEAN)
VB              8              VB
=====
5    0.11189170E-02    5
6    0.15945749E-02    6
7    0.10635411E-01    7
U    0.79561761E 00    U
```

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===    30    8    8= -0                                         100111=====
===    XXXY=                                                     =====
( REGRESSION THROUGH MEAN )

SOURCE OF I DEG. OF I    SUMS OF SQUARED Y DEVIATIONS
VARIATION I FREEDOM I      8Y
=====
REGRESSION           3    0.54100410E 04
ERROR                26    0.12142554E 03
=====
TOTAL                29    0.55314666E 04
=====

MEAN SQUARED RESIDUAL OF PREDICTION MINUS MEAN Y
=VAR. OF REG. ESTIMATE      8Y
=MSQR (ABBREVIATION)=      0.46702130E 01
=====
RMSQR=MSQR/(VAR. Y) =      0.0245
1-RSQUARED              0.0220
RSQUARED                 0.9780
=====

```

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===    30    8    8= -0                                         100111=====
===    XXXY=                                                     =====
( REGRESSION THROUGH MEAN )
I    MEAN X OR Y    VARIANCE X OR Y
=====
5    0.10066667E 02    0.25399540E 03
6    0.81333333E 01    0.17667126E 03
7    0.27333333E 01    0.25443678E 02
8    0.20866667E 02    0.19074023E 03
=====

```

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===  30  8  8= -0  100111=====
===  XXXY=
(REGRESSION THROUGH MEAN)
PT. NO.  PREDICTION MINUS OBSERVED  8Y  =  ERROR
=====
  1  0.71281054E 00 -0.20000000E 01  0.27128105E 01
  2  0.71281054E 00  0.20000000E 01 -0.12871895E 01
  3  0.37862748E 01  0.20000000E 01  0.17862748E 01
  4  0.37862748E 01  0.60000000E 01 -0.22137252E 01
  5  0.78842272E 01  0.60000000E 01  0.18842272E 01
  6  0.78842272E 01  0.10000000E 02 -0.21157728E 01
  7  0.11982180E 02  0.10000000E 02  0.19821795E 01
  8  0.11982180E 02  0.14000000E 02 -0.20178205E 01
  9  0.16080132E 02  0.14000000E 02  0.20801318E 01
 10  0.16080132E 02  0.18000000E 02 -0.19198682E 01
 11  0.87846143E 01  0.70000000E 01  0.17846143E 01
 12  0.87846143E 01  0.11000000E 02 -0.22153857E 01
 13  0.15859508E 02  0.14000000E 02  0.18595082E 01
 14  0.15859508E 02  0.18000000E 02 -0.21404918E 01
 15  0.24955800E 02  0.23000000E 02  0.19557998E 01
 16  0.24955800E 02  0.27000000E 02 -0.20442002E 01
 17  0.36073490E 02  0.34000000E 02  0.20734897E 01
 18  0.36073490E 02  0.38000000E 02 -0.19265103E 01
 19  0.36073490E 02  0.34000000E 02  0.20734897E 01
 20  0.36073490E 02  0.38000000E 02 -0.19265103E 01
 21  0.24914457E 02  0.23000000E 02  0.19144573E 01
 22  0.24914457E 02  0.27000000E 02 -0.20855427E 01
 23  0.15833049E 02  0.14000000E 02  0.18330489E 01
 24  0.15833049E 02  0.18000000E 02 -0.21669511E 01
 25  0.24914457E 02  0.23000000E 02  0.19144573E 01
 26  0.24914457E 02  0.27000000E 02 -0.20855427E 01
 27  0.36013957E 02  0.34000000E 02  0.20139565E 01
 28  0.36013957E 02  0.38000000E 02 -0.19860435E 01
 29  0.49131547E 02  0.47000000E 02  0.21315470E 01
 30  0.49131547E 02  0.51000000E 02 -0.18684530E 01

```

```

===  EXAMPLE OF FAILURE OF STEPWISE PROCEDURE TO FIND BEST REGRESSION  ABST
===  30  8  8= -0  100111=====
===  XXXY=
(REGRESSION THROUGH MEAN)
PT. NO.  PREDICTION MINUS OBSERVED  8Y  =  ERROR
=====
CHECK SUM OF WEIGHTED ERRORS EQUALS  -0.14990568E-04
CHECK SUM OF WEIGHTED SQUARED ERRORS EQUALS  0.12142562E 03

```