DEVELOPMENT OF A PATH ANALYSIS PROGRAM COMPATIBLE WITH INTERACTIVE SPSS

John F. Hall and A. James Ring

(Principal Lecturer in Sociology
Department of Applied Social Studies,
Polytechnic of North London
Ladbroke House,
62/66, Highbury Grove,
London, N5 2AD)

(Research Officer:
National Institute of Social Work
Mary Ward House,
5, Tavistock Place,
London, WC1H 9SS)

Paper to the British Sociological Association Quantitative Sociology Group Colloquium on "Causality, Determination & Explanations"
Hiatt-Baker Hall, University of Bristol, April 5-6, 1977.

ABSTRACT

An interactive computer program developed at the SSRC Survey Unit to perform path analysis on survey data, is being completed at the Polytechnic of North London. User-interface is conversational with English language type instructions, and the prototype was successfully used by students at the Survey Unit Summer Schools in Survey Methods, as well as by Unit staff for research publications. A stand-alone program, it is written in standard Fortran and conforms to both the language and syntax specifications of the interactive version of SPSS being developed at the Universities of Chicago, Edinburgh and Strathclyde. The aim of the project described in this paper is to complete the development of the program and to prepare full user-and program-documentation. Path analysis is a powerful analytical tool for social scientists and a generally available, easy-to-use interactive program would be welcomed by teachers, researchers and students alike. The program is intended to operate in small amounts of core (less than 20K) and will be suitable for small machines. A second-generation proto-type is operating on the CDC 6400 at the University of London Computer Centre; this version is being implemented on the ICL 1905E at the Polytechnic of North London. The project is financed by the Social Science Research Council (Grant No. HR 4598/1).

Development of a Path Analysis Program Compatible with Interactive SPSS

John F. Hall and A. James Ring

1. Introduction

1.1 Path analysis was first considered as a method for analysing causal models in social systems by the extension of regression analysis to cover recursive systems of structural equations containing "measured" variables (Duncan, 1966; Land, 1969). More recent developments of path analysis have been fragmented and consequently lacking in co-ordination. On the one hand, there has developed an analysis of purely quantitative variables in the form of "linear flow graphs" (Goodman, 1972; Davis, 1974; Davis and Schooler, 1974); on the other hand, consideration has been given to the use of confirmatory factor analysis and the introduction of "unmeasured" variables. (Hauser & Goldberger, 1971; Jöreskog, 1973). Other work has pursued the integration of different methods for the analysis of single equations (Nelder & Wedderburn, 1972) and the extension of causal models to non-recursive systems and two-stage regression (Heise, 1975).

1.2 A number of computer programs already exist which have been developed in parallel with the theoretical aspects of path analysis. Simple interactive programs have been available for some time (Nygreen, 1970) and more specialized programs have also been written, such as CATFIT (Taylor, 1975) LISREL (Joreskog, 1972) and GLIM (Nelder, 1974). More recently, general packages have been implemented which can provide an interactive environment for a variety of path analysis techni ques and form a basis for their integration into a single system, in particular the interactive version of SPSS (Nie et al, 1974).

1.3 The next major development in this area clearly has to be the integration of past theoretical and computational work into a single interactive computer program for the analysis of causal models. Work on such a program "Interactive Path Analysis" started at the SSRC Survey Unit (Ring, 1975) and had reached an advanced stage by the time the Unit closed in September 1976. (Ring, forthcoming). The work is being continued at the Polytechnic of North London on a grant from the Social Science Research Council.

*SSRC grant no. HR 4598/1 to Mr. J.F. Hall, Department of Applied Social Studies, Polytechnic of North London; one year starting 1st May 1977.

2. Objectives

2.1 The objectives of this project are two-fold. The principal technical objectives are to complete the development of Interactive Path Analysis (IPA) to the design worked out at the Survey Unit, to implement it on several machines, and to provide full documentation for both users and computing installations. A secondary technical objective is to provide an additional procedure to enhance Interactive SPSS (SPSS-II) and consequently IPA has been written to the language and syntax specifications of SPSS-II (Franklin et al, 1974).

2.2 The academic objectives involve the provision of a powerful tool for social science teaching and research which will integrate hitherto disparate software work in path analysis, user-oriented interactive packages and general linear models with theoretical and causal modelling work as outlined in paras. 1.1 to 1.3. IPA would thus constitute a means of enabling researchers to develop and test social theories on quantitative data. Its value will lie in its generality, its practicality and its usability in providing instant feedback, conversational language, flexibility of problem definition and a variety of output presentations. It relates to the fundamental process of theory-building, not only through the dialectic between the researcher and his data, but also through the probability that its use will encourage more rigorous thought prior to research design, data collection or analysis.

2.3 There will be direct contributions to methodology in the development of new routines for handling interactive programs, from the integration of existing algorithms and programs into a single program, and from the addition of a comprehensive procedure (IPA) to an existing package (SPSS). Indirect benefits will accrue through education in research and analytical approaches, encouragement of new research designs to assume a causal scientific approach. Secondary analysis of major surveys from the SSRC Survey Archive at Essex University is an obvious area of application for IPA.

3. Programme of action

3.1 The authors have been in close liaison with the SPSS-II team since the initial SSRC award was made. The major design conference on Interactive SPSS (held at LSE in April, 1974) was organized by John Hall, who also acted as rapporteur to the working group on user-interface and output options. Exchanges of documents have been frequent and regular. The National Opinion Research Center (home of SPSS) at

the University of Chicago has specifically refrained from developing a path analysis facility because they are aware of James Ring's work in London. Whilst IPA will itself be a stand-alone program, all documentation and coding will be made available to the SPSS-II team so that interfacing can be implemented as quickly as possible.

3.2 Whilst the major focus of liaison will be the SPSS-II team, the IPA team has a number of contacts in social science computing both in Britain and overseas, but particularly close contact will be maintained with Prof. James A. Davis of Dartmouth College. Information on progress will be reported directly to such contacts and indirectly through professional associations (British Sociological Association, Market Research Society, Royal Statistical Society) study groups (BSA Quantitative Sociology Group, BSA Survey Research Group, Study Group on Computers in Survey Analysis, SSRC Sponsored Research Seminars in Quantitative Social Science, Radical Statistics and publications (Quantitative Sociology Newsletter, SSRC Newsletter, SSRC Survey Archive Bulletin, Edinburgh Programm Library Unit SPSS Newsletter).

3.3 Most of the computing work will be done on the ICL1905E at the London Polytechnics' Computer Unit at Polytechnic of North London, but the CDC6400 version will be maintained at ULCC and, if time allows,a version will be implemented on the DEC10 at Polytechnic of the South Bank.

3.4 It is envisaged that the program will be in constant use, either by the applicants, or by researchers and students at PNL, or elsewhere as appropriate, so that feedback from users can be taken into account in the final versions of the program and its associated documentation. The authors would be grateful to hear from potential users who would like access to the program and who could provide the necessary feedback on bugs or user-interface problems. If all goes according to plan, the various stages of the work should be completed as follows.

Task		Completion by
Programming:	Programming for ICL 1905E	31.7.77
	Coding for SPSS subprogram	30.10.77
Manual:	Draft manual for PNL users	31. 7.77
	Draft manual for SPSS users	30.10.77
	Feed-back from users	31. 1.77
Completion of program & documentation		30. 4.77

3.5 For various reasons the project started seven months behind schedule, but with hard work and good will all round, we hope to provide, at relatively low cost and in a comparatively short time, a powerful, but easy-to-use facility for causal analysis and model-building in the Social Sciences.

4. The program

4.1 IPA is a program for performing path analysis interactively on data summaries of three types.

(i) Matrices of measures of association
(ii) Frequencies from n-way tables
(iii) Measures of central tendency and dispersion

User-interface is in conversational SPSS type language. The program supplies prompts and generates error messages, and saves all current information and models, thus enabling the user to re-enter at the point he left off in a previous session, even after a system failure. Full naming and labelling facilities are available for files, matrices, variables and models and it will be possible to change any part of the data or its labelling at will. The version to be released will operate in small amounts of core ($<$ 20K) thus widening the range of machines on which it could be implemented. This will benefit polytechnics and departments with small machines.

4.2 The program presents the user with a series of prompts, each of which defines a context with a specific free-formatted response or set of responses. For example, the program begins with the prompts GET FILE? to which the user must reply with either the file name of an existing IPA file on a local disk-file GTFILE, or with NULL indicating that there is no existing IPA file. The latter elicits the prompt NAME? to which user replies with a file name and an optional file lable. The program then moves into its variable definition context and prompts NAME OF VARIABLE? Variables are defined by name and by matrix row number with optional extended variable labelling. As soon as all variables have been defined the user types a NULL response to the variable name prompt, whereat the program moves into its data input context. Data can be input from some previously defined file INDATA or direct from the keyboard. Three types of data can be defined. These are means, standard deviations and coefficients of association. The prompt DATA? requests the next type of data to be read in, and can be answered by the responses MEANS , DEVIATIONS , or COEFFICIENTS . The response COEFFICIENTS must be used, and elicits the prompt LAYOUT? which can be answered by one of four layout types SQUARE , UPPER , LOWER or SPSS ; where UPPER and LOWER refer to triangular matrices of coefficients. When the data matrix has been defined,a NULL response to the prompt DATA?

moves the program on to its model-definition context. The prompt LINK? requires the initial path or paths to be specified by a statement of the form [< variable name> WITH < variable name > or < variable list>] in which the first variable specified is the dependent variable. Multiple paths can be specified until the response NULL indicates the completion of path definition. The program then moves into its display context with the prompt DISPLAY? and the user can request any or all of the options CORRELATIONS PATH or VARS . The latter request will display means, deviations and any current labels for variables in the model. CORRELATIONS will print the matrix of coefficients currently in use and PATH will print out for each model or section of model the standardised regression coefficients (beta-weights) and the unstandardised coefficients for the independent variables and the regression coefficient (multiple correlation coefficient :R^2) Examples of runs are given in Appendix B.

4.3 A conversational prompting system is ideal for the beginner or the diffident researcher, but experienced researchers with some knowledge of computing will find full prompting both tedious and repetitive. Consequently the system has been designed to allow such users shorthand responses and preemptive specifications, so that only the first three characters of system keywords need to be entered and the entry of a slash ('/') will preempt the next prompt to enable entire sequences to be entered without intervening prompts.

4.4 Even the most sophisticated users make keypunching errors and the program is equipped with a full system for error reporting.

The final version is designed for the user who has lost track of his analysis to enter the keyword /HELP at which point appropriate messages will appear.

4.5 A preemptive response of /STOP will end the run.

4.6 An additional response to the LINK? prompt has been provided to enable previously specified paths to be deleted. The user types /CUT ALL or <var. name> FROM <var. name> or <var. list> or ALL.

4.7 It needs to be stressed that IPA is not just another example of proliferation of programs for social scientists and survey analysis at a time when existing programs are more than adequate. On the contrary IPA has been designed from the start to avoid duplication of other work and this has been achieved through extensive consultation with others working in the field. As stated above, it has also been designed to the syntax and language specifications of SPSS-II. It is quite clear that, for social scientists, SPSS is now the de facto package in Britain for most of their requirements because of its availability, documentation and ease of use. This is the case not only in universities, but increasingly in polytechnics, local authorities, and government agencies. In addition to the advantages of SPSS-style English language and conversational operation, IPA offers instant feedback, a variety of output options and great flexibility in enabling the user to vary his assumptions quickly in accordance with the results obtained.

APPENDIX A: SYNTAX OF INTERACTIVE PATH ANALYSIS PROGRAM (IPA 1.2) 31.03.77

	PROMPT (FROM PROGRAM)	RESPONSE (FROM USER)	EXPLANATORY REMARKS
1.0	INTERACTIVE PATH ANALYSIS ...		Informative message
1.1	GET FILE?	<file name>	File GTFILE contains the IPA system-file
		NULL	No system-file has been defined: file definition sequence follows
2.0	FILE DEFINITIONS ...		Informative message
2.1	NAME?	<file name>	A new file must be given a name: label is optional, but must be in double quotes, otherwise NULL.
2.2	LABEL?	"<file label>"	
3.0	VARIABLE DEFINITIONS...		Informative message
3.1	NAME OF VARIABLE?	<var. name>	Each variable must be entered by name, followed by the row number of the matrix or table in which it occurs: label is optional. Program will detect and report errors in names, numbers, duplicates etc.
3.2	ROW NUMBER?	<matrix row no.>	
3.3	LABEL?	"<var. label>"	
	:		
3.1	NAME OF VARIABLE?	NULL	No more variables to be defined: matrix definitions sequence follows
4.0	MATRIX DEFINITIONS ...		Informative message
4.1	DATA?	MEANS	Optional. Program will expect next row of data to be means of variables.
		DEVIATIONS	Optional. Program will expect next row of data to be standard deviations of variables.
		COEFFICIENTS	Compulsory. Program expects matrix of coefficients of association, but will ask for format of matrix first.
4.2	LAYOUT?	UPPER	Program will expect upper triangle without diagonal:
		LOWER	Program will expect lower triangle without diagonal:
		SQUARE	Program will expect square matrix:

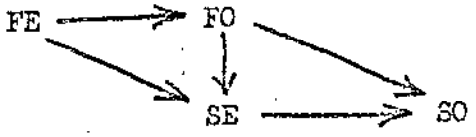
	PROMPT	RESPONSE	EXPLANATORY REMARKS
(4.2)	LAYOUT? (contd.)	SPSS	Program will expect square matrix of standard SPSS format. (eg. nF10.7) N.B. LAYOUT? prompt occurs only when COEFFICIENTS has been defined in response to the prompt DATA?
5.0	INPUT DEFINITIONS		Informative message
5.1	INPUT MEDIUM ("INPUT" OR "INDATA")?	INDATA	Matrix or table is stored on file INDATA. (N.B. This file name conforms to LSE conventions) Goes to 6.0
		INPUT	Default option. Program expects matrix or table to be entered direct via keyboard of terminal.
5.2	VALUES?	⟨free format list of real numbers⟩	Does not occur if data input is from INDATA
		NULL	No more data to be input. Program moves to model-building context.
6.0	PATH DEFINITIONS		Informative message
6.1	LINK?	⟨var. name⟩ WITH ⟨var. name⟩ or ⟨var. list⟩	Variable specified to left of WITH is dependent variable. ALL vars. to right of WITH are independent. A var. list may include keyword TO or commas as list separators.
		/CUT ALL or ⟨var. name⟩ FROM ⟨var. name⟩ or ⟨var. list⟩	Pre-emptive response to delete paths from existing model.
		NULL	No more paths to be defined: program enters display mode.
6.2	DISPLAY?	CORRELATIONS	Matrix is displayed in F5.4 format
		PATH	For each model or sub-model displays standardised and unstandardised regression coefficients and R^2
		VARIABLES	For each variable defined, displays mean, deviation and label (if any)
		NULL	Program returns to most recent context.
		/STOP	Pre-emptive entry allows user to finish at any time Current model and information automatically saved

APPENDIX B

This appendix gives examples of actual runs using IPA on existing data. Example 1 uses IPA on data given by Heise (Heise, 1975) to examine the relationship between education and occupational status of fathers and their sons. The father's occupational status (FO) is stated to be dependent on his education (FE) the son's education (SE) on both, and the son's occupational status (SO) on his own education and his father's occupational status. (See also Blau & Duncan, 1967, p. 169)

EXAMPLE 1

Diagrammatically:



To test this model using IPA results in the following runs:

(a) No pre-empting

```
INTERACTIVE PATH ANALYSIS...
GET FILE?NO
FILE DEFINITIONS...
NAME?MOBILITY
LABEL?"HEISE DATA"
VARIABLE DEFINITIONS...
NAME OF VARIABLE?FE,1,"FATHER'S EDUCATION"
NAME OF VARIABLE?FO,2,"FATHER'S OCCUPATIONAL STATUS"
NAME OF VARIABLE?SE,3,"SON'S EDUCATION"
NAME OF VARIABLE?SO,4,"SON'S OCCUPATIONAL STATUS"
NAME OF VARIABLE?NO
MATRIX DEFINITIONS...
DATA?COE
TYPE OF LAYOUT?UPPER
DATA?NO
INPUT DEFINITIONS...
MEDIUM ("INPUT" OR "INDATA")?INPUT
DATA = .516 .453 .322 .438 .405 .596
PATH DEFINITIONS...
LINK?SO WITH SE FO
LINK?SE WITH FO FE
LINK?FO WITH FE
LINK?NO
DISPLAY? CORR PATH
```

BOX 1

This will result in the following output:

```
HEISE DATA                                        CORRELATION MATRIX      PAGE  1

          FE         FO         SE         SO
FE        1.0000      .5160      .4530      .3220
FO         .5160     1.0000      .4380      .4050
SE         .4530      .4380     1.0000      .5960
SO         .3220      .4050      .5960     1.0000
TYPE "+" FOR NEXT PAGE..+
```

```
HEISE DATA                                        PATH ANALYSIS           PAGE  2

DEPENDENT VARIABLE FO          - FATHER'S OCCUPATIONAL STATUS
PATH FROM : BETA WT : B COEFF
FE           .5160     .5160
SQUARED MULTIPLE REGRESSION COEFFICIENT =     .266256

DEPENDENT VARIABLE SE          - SON'S EDUCATION
PATH FROM : BETA WT : B COEFF
FE           .3094     .3094
FO           .2784     .2784
SQUARED MULTIPLE REGRESSION COEFFICIENT =     .262067

DEPENDENT VARIABLE SO          - SON'S OCCUPATIONAL STATUS
PATH FROM : BETA WT : B COEFF
FO           .1781     .1781
SE           .5180     .5180
SQUARED MULTIPLE REGRESSION COEFFICIENT =     .380857
```

BOX 2

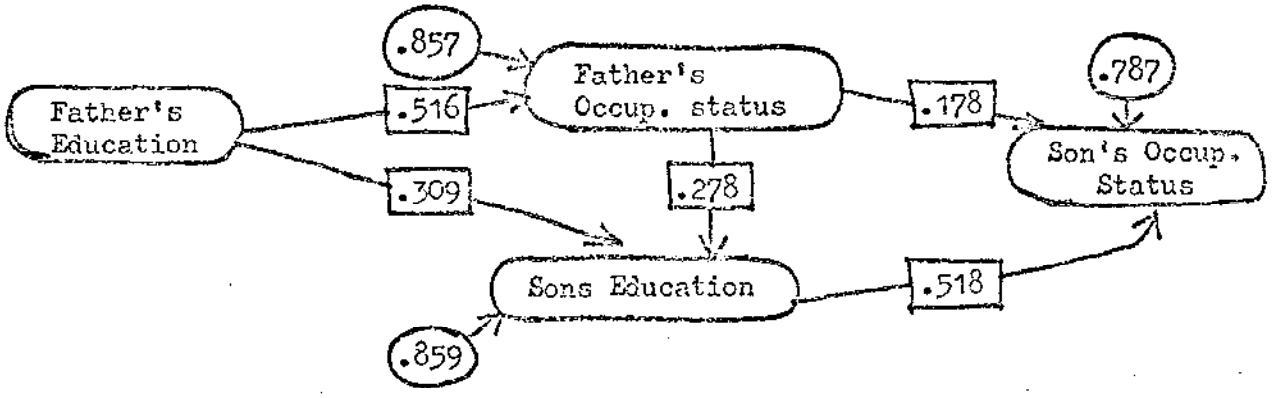
The contents of boxes 1 & 2 above are facsimiles from an actual run on a Data Dynamics teletype using a telephone link to the CDC 6400 at ULCC on 17th March 1977. The double quotes enclosing variable labels are not compatible with SPSS, but are necessitated temporarily by the need to use character handling routines from the program library of the London School of Hygiene and Tropical Medicine.

(b) Same run with pre-empting

```
GET FILE?        /VAR FE, 1,"FATHER'S EDUCATION"/FO, 2,"FATHER'S OCCUPATION"
                 /SE, 3,"SON'S EDUCATION"/SO, 4,"SON'S OCCUPATION"/INP MED INPUT
                 /DAT COE/LAY UPPER/DAT   .516 .453 .322  .438 .405 .596
                 /LIN SO WITH SE  FO/SE WITH FO  FE/FO WITH FE/DIS CORR  PATH

                 ****  DISPLAY OUTPUT HERE  ****
                       etc.,    etc.,
```

This run gives the same answers as Heise for his test exercise (Heise 1975 pp..403-406) the results of which can be displayed diagrammatically with path coefficients entered for explained variance, and residuals representing unexplained variance ($1 - R^2$) as follows



In crude terms, this model explains approx. 21% of the variation in occupational status of sons. Using IPA the researcher could proceed to add more variables into the model, and to define or delete paths, in an attempt to reduce the unexplained variance, or to simplify the model. Example 2 below is a demonstration run on a teletype using data from the Survey Unit's 1975 'Quality of Life' survey. An ampersand '&' at the end of a line indicates that the current response is continued on the next line.

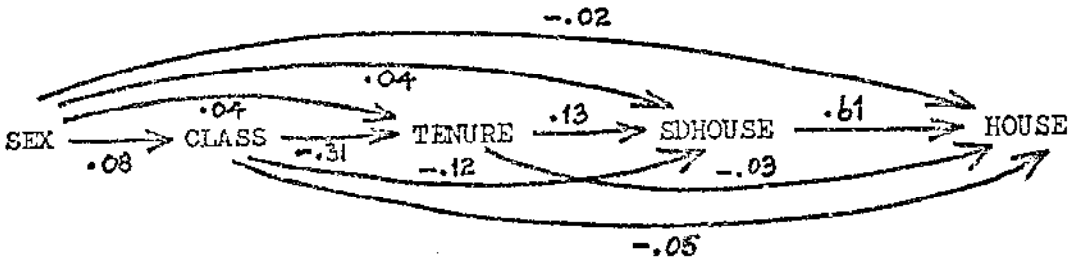
APPENDIX B contd.

EXAMPLE 2

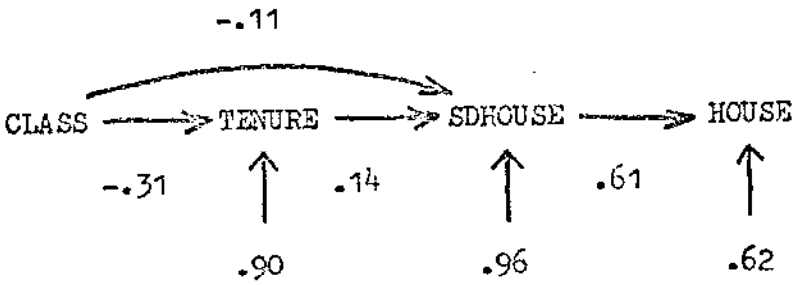
The model contains five variables causally ordered.

1.	Sex of respondent	(SEX)	Male, Female
2.	Class of head of household	(CLASS)	Middle class, Working class
3.	Tenure of dwelling	(TENURE)	Owner occupied, other
4.	Score on housing affect scale	(SDHOUSE)	0 - 48
5.	Satisfaction with house	(HOUSE)	0 - 10

The initial model is defined as follows:



By deleting paths with coefficients less than 0.1 the model is simplified as follows:



Far more complex models derived from many runs are described in Hall (1975) Marsh (1975) the analysis for which might have taken weeks or even months to complete, but was in fact done in two or three sessions of about half an hour each.

Appendix B Contd

EXAMPLE 2 TELETYPE RUN OF IPA 1.1

```
COMMAND - IPA
GET FILE?                          NULL
VARIABLE NAME?                     SEX, 1
VARIABLE NAME?                     CLASS, 2
VARIABLE NAME?                     TENURE, 3
VARIABLE NAME?                     HOUSE, 4
VARIABLE NAME?                     SDHOUSE, 5
VARIABLE NAME?                     NULL
INPUT MEDIUM ("INPUT" OR "INDATA")?  INPUT
MATRIX TYPE?                       LOWER
READ MATRIX                        .0827 .0169 -.3092 -.0013 -.1407 .0910  &
.0321 -.1565 .1718 .6139
LINK?                              HOUSE WITH SEX, CLASS, TENURE, SDHOUSE
LINK?                              SDHOUSE WITH SEX, CLASS, TENURE
LINK?                              TENURE WITH SEX, CLASS
LINK?                              CLASS WITH SEX
LINK?                              NULL
DISPLAY?                           CORR, PATH
```

CORRELATION MATRIX PAGE 1

	SEX	CLASS	TENURE	HOUSE	SDHOUSE
SEX	1.0000	.0827	.0169	-.0013	.0321
CLASS	.0827	1.0000	-.3092	-.1407	-.1565
TENURE	.0169	-.3092	1.0000	.0910	.1718
HOUSE	-.0013	-.1407	.0910	1.0000	.6139
SDHOUSE	.0321	-.1565	.1718	.6139	1.0000

TYPE "+" FOR NEXT PAGE... +

PATH ANALYSIS PAGE 2

```
DEPENDENT VARIABLE CLASS
PATH FROM:  BETA WT.
SEX          .0827
REGRESSION COEFFICIENT =          .006839

DEPENDENT VARIABLE TENURE
PATH FROM: BETA WT.
SEX          .0428
CLASS       -.3127
REGRESSION COEFFICIENT =          - .097421
```

Appendix B contd.

```
DEPENDENT VARIABLE HOUSE          -
PATH FROM: BETA WT.
SEX        -.0160
CLASS      -.0530
TENURE     -.0301
SDHOUSE     .6113
REGRESSION COEFFICIENT   =        .380012
TYPE "+"  FOR NEXT PAGE.... +

                                          PATH ANALYSIS     PAGE 3

DEPENDENT VARIABLE SDHOUSE        -
PATH FROM : BETA WT.
SEX         .0396
CLASS      -.1182
TENURE      .1346
REGRESSION COEFFICIENT =          .042886
LINK?     /CUT CLASS FROM SEX/TENURE FROM SEX/HOUSE FROM SEX TO  &
TENURE/SDHOUSE FROM SEX/DIS PATH

                                          PATH ANALYSIS     PAGE 1

DEPENDENT VARIABLE TENURE         -
PATH FROM: BETA WT.
CLASS      -.3092
REGRESSION COEFFICIENT =          .095605

DEPENDENT VARIABLE HOUSE          -
PATH FROM: BETA WT.
SDHOUSE     .6139
REGRESSION COEFFICIENT =          .376873

DEPENDENT VARIABLE SDHOUSE        -
PATH FROM: BETA WT.
CLASS      -.1143
TENURE      .1365
REGRESSION COEFFICIENT =          .041332
LINK?      NULL

DISPLAY?                  NULL
SAVE FILE NAME?           HOUSE
```

Appendix C Specifications for current and final versions of IPA

IPA is currently implemented as version 1.1 (ie. the first update of the first generation program). The version to be developed is the second generation program IPA 2.0

IPA 1.1:

1 Data and definitions are stored on disk in a system file identified by name and described by a label.

2 Variables must be defined separately, by name, together with the corresponding input correlation matrix row number and an optional label. Limit of no. of variables is 50.

3 All variables are assumed to be continuous or dummy variables.

4 Data must be read in the form of a correlation matrix and optionally means and standard deviations.

5 The data must be read in a specific order: means followed by deviations followed by coefficients. (This can be changed to the order in which data elements are defined in this version if required).

6 Data can be read either through the terminal or from a disk file.

7 The matrix may be defined in a number of formats: square, upper triangle, lower triangle or standard SPSS.

8 There are no facilities for reading totals and hence no significance tests.

9 Once the matrix is read, it is stored in the system file. It cannot be altered, neither can a new matrix be stored in its place, but additional rows can be defined and stored later.

10 Causal models are defined by links between each dependent variable and its predictors. Limit of no. of predictors is 25.

11 The model is stored in the system file, together with the resulting path coefficients.

12 The model can be changed at any time, either by introducing new links or by cutting old ones.

13 Three display options are available, each producing information on all variables defined:

- a) A variable list of all the variables in the file, including labels, row numbers, means and standard deviations.
- b) A correlation matrix of all variables.
- c) The path coefficients for each dependent variable specified, together with the unstandardised coefficients and the multiple regression.

14 Processing may be stopped at any time, but restarting begins from path definitions, irrespective of context when stopped.

Appendix C contd.

IPA 2.0 has following improvements over and above IPA 1.1

1. Variables can be defined in the form of a list, or series of lists
2. Variables can be of five types as in SPSS-II (alpha, zone, discrete, continuous or set)
3. Set variables can be used to define a list of variables
4. Values can be assigned to each category-scale variable (ie. 1st 3 types), together with missing definition and value label
5. Up to 100 variables may be defined, including set variables
6. Up to 20 different matrices may be defined, each with identifying name and a label
7. A variable list defines the rows of a matrix: category scale variables have one row for each category (or alternatively required categories are defined for each specified category scale variables in the list)
8. Missing treatment may be defined (ie. listwise, pairwise or included)
9. Totals for significance can be defined (either as a matrix of totals or as a single total)
10. The input matrix is in the form of correlations, frequencies & breakdown means, and requires a special format for data input.
11. Up to 20 models may be defined at the same time, identified by name and label
12. Display procedure defines variables to be presented, together with statistics and results for each variable.
13. Any information in the file may be modified or deleted at any time
14. Program restarts at next context after stopped.

Appendix D

Error Messages Generated by IPA 1.2

Error code	Message
0/1	TEXT SHOULD BE ON A NEW LINE OR PRECEEDED BY A PREEMPT
0/2	INSTRUCTION PROMPT WAS MISSING
1/1	AN ALPHANUMERIC FILE NAME WAS EXPECTED
1/2	FILE NAME DOES NOT MATCH SAVED FILE NAME
2/1	AN ALPHANUMERIC FILE NAME WAS EXPECTED
2/2	FILE LABEL SHOULD BE IN QUOTES
3/1	AN ALPHANUMERIC VARIABLE NAME WAS EXPECTED
3/2	VARIABLE NAME ALREADY EXISTS
3/3	AN INTEGER ROW NUMBER WAS EXPECTED
3/4	ROW NUMBER ALREADY ALLOCATED
3/5	A ROW NUMBER MUST BE DEFINED FOR EACH VARIABLE
4/1	A KEYWORD RESPONSE TO "DATA?" WAS EXPECTED
4/2	A KEYWORD RESPONSE TO "LAYOUT?" WAS EXPECTED
5/1	A KEYWORD RESPONSE TO "MEDIUM?" WAS EXPECTED
5/2	DEFAULT MATRIX LAYOUT "SQUARE" WAS
5/3	INPUT MEDIUM SET TO "INPUT" FOR NON-SPSS LAYOUT
5/4	DEFAULT INPUT MEDIUM "INPUT" WAS SUPPLIED
5/6	LAST ROW IN THIS LINE WAS INCOMPLETE - REPEAT THIS ROW
5/7	MORE ROWS REQUIRED - TYPE THE NEXT ROW
6/1	AN ALPHANUMERIC VARIABLE NAME WAS EXPECTED
6/2	VARIABLE NAME NOT RECOGNISED
6/3	A KEYWORD "WITH" OR "FROM" WAS EXPECTED
6/4	INCORRECT LINK KEYWORD WAS ACCEPTED
6/5	A KEYWORD RESPONSE TO "DISPLAY?" WAS EXPECTED

Other error and diagnostic messages

Context(s)	Message
1/1	FOUND FILE NAME @@@ - EXPECTED FILE NAME @@@
6/1,3	VARIABLE @@@ APPEARS ON BOTH SIDES - PATH IGNORED
6/1	PATH FROM @@@ TO @@@ ALREADY EXISTS - LINK REQUEST IGNORED
6/1	RECIPROCAL PATH FROM @@@ TO @@@ WAS CUT
6/3	NO PATH FROM @@@ TO @@@ - CUT REQUEST IGNORED
6/4	TOO MANY VARIABLES FOR MODEL - MAXIMUM IS 25
6/4	A SINGULARITY WAS ENCOUNTERED IN THE FOLLOWING SET.

REFERENCES

BLAU, P & DUNCAN, O.D. (1967) "The American Occupational Structure" Wiley, N.Y.

DAVIS, J A (1974) "Linear Flow Graphs as a Mathematical Model for Analysing Contingency Tables" NORC, Chicago,

DAVIS, J A and SCHOOLER S R (1974) "Nonparametric Path Analysis - The Multivariate Structure of Dichotomous Data When Using the Odds Ratio or Yule's Q. NORC, Chicago

DUNCAN, O D (1966) "Path Analysis, Sociological Examples," American Journal of Sociology, 75, 245-263

FRANKLIN M N et al (1974) "Interactive SPSS Project-Language Specifications 1-17" Program Library Unit, Edinburgh Reg. Comp. Centre.

GOODMAN L A (1972) "A General Model for the Analysis of Surveys" In American Journal of Sociology, 77, 1035-1086

HALL J F (1975) "The relationship between objective & Subjective indicators of individual well-being:a linear modelling approach", SSRC Survey Unit, UK/USA "Subjective Indicators of Quality of Life" Fitzwilliam College, Cambridge, Sept. 1975

HAUSER R L & GOLDBERGER A S (1971) "The Treatment of Unobservable Variables in Path Analysis" In Sociological Methodology (H L Costner, Ed.) Jossey-Bass, San Francisco.

HEISE D R (1975) "Causal Analysis" Wiley-Interscience, New York

JÖRESKOG K G (1972) "LISREL: A general computer program for estimating a linear structural equation system involvir multiple indicators of unmeasured variables" Princeton, New Jersey

JÖRESKOG K G (1973) "A General Model for Estimating a Linear Structural Equation System" In Structural Equation Models in the Social Sciences (A S Goldberger and O D Duncan, Eds) Seminar Press New York

LAND K C (9169) "Principles of Path Analysis" In Sociological Methodology (E F Borgatta, Ed.) Jossey-Bass, San Francisco

MARSH C (1975) "Job satisfaction - the search goes on" SSRC Survey Unit, UK/USA Seminar, "Subjective Indicators of Quality of Life" Fitzwilliam College, Cambridge, Sept. 1975.

NELDER J A (1974) "GLIM User's Manual" N.A.G. Oxford

NELDER, J A and WEDDERBURN R M W (1972) "Generalised Linear Models" In Journal of the Royal Statistical Society (series A), 135, 3, 370-384)

NIE N H., HULL C H and FRANKLIN M N (1974) "Interactive SPSS Project- General Introduction" Program Library Unit - Edinburgh

REFERENCES (Contd)

NYGREEN G T (1970)	"IPA-Interactive Path Analyser" Princeton, New Jersey
RING A J (1975)	"Interactive Path Analysis" BSA Quantitative Sociology Group, Essex University
RING A J (forthcoming)	"Causal Modelling & Path Analysis"
TAYLOR G (1975)	"User's Instructions for the DTSS Program *E85000: CATFITII" N.O.R.C., Chicago