



## Path analysis

The technique of path analysis originated with Sewell Wright in the late 1930s. It can be used to test the plausibility of putative causal relationships between one variable and another in non-experimental conditions. In general, the procedure can be formulated as one of estimating the coefficients of a set of linear structural equations representing the cause and effect relationships hypothesized by the investigator.

The system of relationships involves variables of two kinds: independent or cause variables  $x_1, x_2, \dots, x_q$  and dependent or effect variables  $y_1, y_2, \dots, y_p$ .

The classical approach to path analysis consists of solving the structural equations for the dependent variables in terms of the independent and the random disturbance terms  $\zeta_1, \zeta_2, \dots, \zeta_p$ . The object is to obtain the reduced form equations, to estimate the regression of the dependent variables on the independent variables and then to solve for the structural parameters in terms of the regression coefficients. The last step is not always possible.

Estimating a path analysis model for directly observed variables with LISREL is, in contrast, entirely straightforward. Rather than estimating each equation separately, LISREL considers the model as a system of equations and estimates all the structural coefficients directly. The reduced form is obtained as a by-product. The following example (Kerchoff (1974), Kenny (1979)) illustrate the procedure.

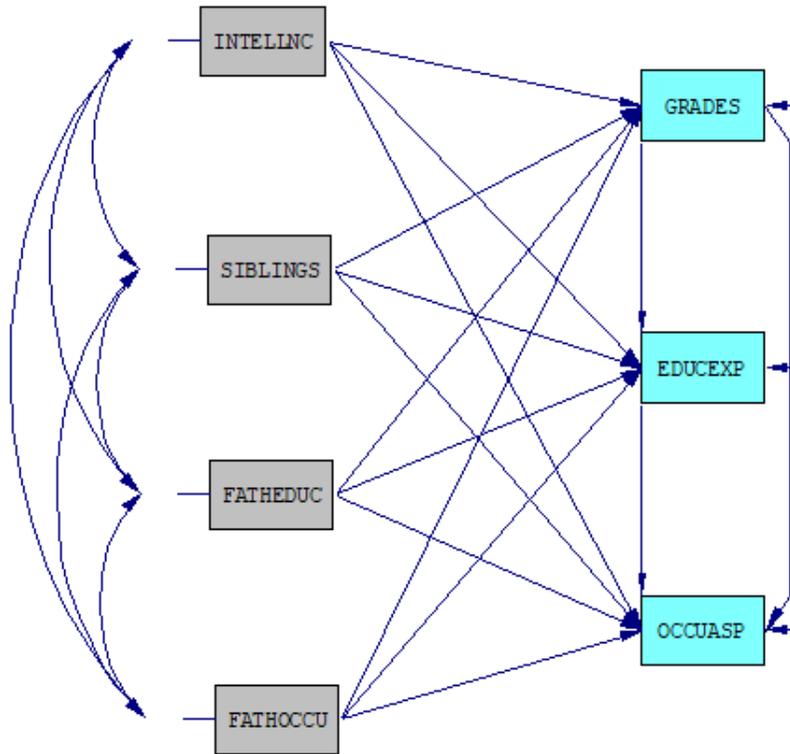
The table below gives the correlations between a number of attainment and background variables for 767 twelfth-grade males. The variables, in the order in which they appear in the table, are:

- $x_1$  = intelligence (INTELLNC)
- $x_2$  = number of siblings (SIBLINGS)
- $x_3$  = father's education (FATHEDUC)
- $x_4$  = father's occupation (FATHOCCU)
- $y_1$  = grades (GRADES)
- $y_2$  = educational expectation (EDUCEXP)
- $y_3$  = occupational aspiration (OCCUASP)

**Table: Correlations for background, ambition and attainment variables**

	$x_1$	$x_2$	$x_3$	$x_4$	$y_1$	$y_2$	$y_3$
INTELLNC	1.000						
SIBLINGS	-0.100	1.000					
FATHEDUC	0.277	-0.152	1.000				
FATHOCCU	0.250	-0.108	0.611	1.000			
GRADES	0.572	-0.105	0.294	0.348	1.000		
EDUCEXP	0.489	-0.213	0.446	0.410	0.597	1.000	
OCCUASP	0.335	-0.153	0.303	0.331	0.478	0.651	1.000

A possible path diagram to explain these correlations is given below.



The model is a *complete recursive system with exogenous variables*  $x_1$ ,  $x_2$ ,  $x_3$ , and  $x_4$ , and *jointly dependent variables*  $y_1$ ,  $y_2$ , and  $y_3$ . The structural equations are:

$$y_1 = \gamma_{11}x_1 + \gamma_{12}x_2 + \gamma_{13}x_3 + \gamma_{14}x_4 + \zeta_1$$

$$y_2 = \beta_{21}y_1 + \gamma_{21}x_1 + \gamma_{22}x_2 + \gamma_{23}x_3 + \gamma_{24}x_4 + \zeta_2$$

$$y_3 = \beta_{31}y_1 + \beta_{32}y_2 + \gamma_{31}x_1 + \gamma_{32}x_2 + \gamma_{33}x_3 + \gamma_{34}x_4 + \zeta_3$$

The LISREL command file **EX45A.LIS** (see the **LISREL Examples** folder) for fitting this model is as follows:

```
Ambition and attainment
DA NI=7 NO=767
LA
INTELLNC SIBLINGS FATHEDUC FATHOCCU GRADES EDUCEXP OCCUASP
KMATRIX
1 -.1 1 .277 -.152 1 .25 -.108 .611 1 .572 -.105 .294 .248 1
.489 -.213 .446 .41 .597 1 .335 -.153 .303 .331 .478 .651 1
SE
5 6 7 1 2 3 4
MO NY=3 NX=4 BE=SD PS=DI
OU SE TV EF
```

When there are no *a priori* fixed elements among the  $\beta$ 's or  $\gamma$ 's, the model is just identified and fits the data perfectly. The initial estimates are identical to the maximum likelihood estimates. The ML estimates and the *t*-values are shown below.

LISREL Estimates (Maximum Likelihood)

BETA				
	GRADES	EDUCEXP	OCCUASP	
	-----	-----	-----	
GRADES	- -	- -	- -	
EDUCEXP	0.405 (0.032) 12.562	- -	- -	
OCCUASP	0.158 (0.037) 4.294	0.550 (0.038) 14.626	- -	

GAMMA				
	INTELLNC	SIBLINGS	FATHEDUC	FATHOCCU
	-----	-----	-----	-----
GRADES	0.526 (0.031) 17.172	-0.030 (0.030) -1.011	0.119 (0.038) 3.166	0.041 (0.037) 1.094
EDUCEXP	0.160 (0.032) 4.989	-0.112 (0.026) -4.235	0.173 (0.034) 5.126	0.152 (0.033) 4.590
OCCUASP	-0.039 (0.034) -1.163	-0.019 (0.028) -0.679	-0.041 (0.036) -1.162	0.100 (0.035) 2.860

PSI

Note: This matrix is diagonal.

GRADES	EDUCEXP	OCCUASP
0.651	0.517	0.557
(0.033)	(0.026)	(0.028)
19.532	19.532	19.532

**Direct and indirect effects:**

The path diagram shows that there are both direct and indirect effects of the  $x$ -variables on  $y_2$  and  $y_3$ . These effects can be calculated from the estimated parameters of the model. For example, the direct effect of  $x_1$  on  $y_2$  is  $\gamma_{21} = 0.160$  and the indirect effect of  $x_1$  on  $y_2$  is  $\gamma_{11}\beta_{21} = 0.526 \times 0.405 = 0.213$ .

The sum of the direct effect and all indirect effects is called total effects. These total effects are given by LISREL when EF appears on the OU command. The total and indirect effects for the present example are shown below.

**Total and Indirect Effects**

Total Effects of X on Y

	INTELLNC	SIBLINGS	FATHEDUC	FATHOCCU
GRADES	0.526 (0.031) 17.217	-0.030 (0.030) -1.014	0.119 (0.037) 3.174	0.041 (0.037) 1.097
EDUCEXP	0.373 (0.030) 12.489	-0.124 (0.029) -4.287	0.221 (0.037) 6.023	0.168 (0.036) 4.646
OCCUASP	0.249 (0.034) 7.317	-0.092 (0.033) -2.787	0.099 (0.042) 2.369	0.198 (0.041) 4.816

Indirect Effects of X on Y

	INTELLNC	SIBLINGS	FATHEDUC	FATHOCCU
GRADES	- -	- -	- -	- -
EDUCEXP	0.213 (0.021) 10.165	-0.012 (0.012) -1.011	0.048 (0.016) 3.078	0.016 (0.015) 1.093
OCCUASP	0.288 (0.027)	-0.073 (0.019)	0.140 (0.024)	0.099 (0.024)

10.869      -3.871      5.734      4.176

Total Effects of Y on Y

	GRADES	EDUCEXP	OCCUASP
	-----	-----	-----
GRADES	- -	- -	- -
EDUCEXP	0.405 (0.032) 12.595	- -	- -
OCCUASP	0.381 (0.038) 10.074	0.550 (0.037) 14.664	- -

Largest Eigenvalue of B\*B' (Stability Index) is 0.349

Indirect Effects of Y on Y

	GRADES	EDUCEXP	OCCUASP
	-----	-----	-----
GRADES	- -	- -	- -
EDUCEXP	- -	- -	- -
OCCUASP	0.223 (0.023) 9.555	- -	- -

The total effect of  $x_1$  on  $y_2$  is 0.373. This is the sum of the direct effect of 0.160 and the indirect effect 0.213 computed above.

The  $t$ -values for the solution reveal that the effects  $\gamma_{12}$ ,  $\gamma_{14}$ ,  $\gamma_{31}$ ,  $\gamma_{32}$ , and  $\gamma_{33}$  may not be significant. A formal test of the hypothesis that these five  $\gamma$ 's are zero should be based on an independent sample. This can be obtained by running the model again, inserting a command

FIX GA(1,2) GA(1,4) GA(3,1)-GA(3,3)

after the MO command. The last section of the output for the revised command file (**LISEX45B.LIS**) is given below.

Total and Indirect Effects

Total Effects of X on Y

	INTELLNC -----	SIBLINGS -----	FATHEDUC -----	FATHOCCU -----
GRADES	0.531 (0.030) 17.498	- -	0.147 (0.030) 4.835	- -
EDUCEXP	0.376 (0.030) 12.575	-0.112 (0.026) -4.249	0.232 (0.036) 6.526	0.152 (0.033) 4.605
OCCUASP	0.275 (0.024) 11.646	-0.060 (0.015) -4.088	0.145 (0.022) 6.582	0.158 (0.033) 4.761

Indirect Effects of X on Y

	INTELLNC -----	SIBLINGS -----	FATHEDUC -----	FATHOCCU -----
GRADES	- -	- -	- -	- -
EDUCEXP	0.215 (0.021) 10.232	- -	0.059 (0.013) 4.515	- -
OCCUASP	0.275 (0.024) 11.646	-0.060 (0.015) -4.088	0.145 (0.022) 6.582	0.082 (0.019) 4.403

Total Effects of Y on Y

	GRADES -----	EDUCEXP -----	OCCUASP -----
GRADES	- -	- -	- -
EDUCEXP	0.405 (0.032) 12.614	- -	- -
OCCUASP	0.356 (0.033) 10.817	0.537 (0.036) 15.024	- -

Largest Eigenvalue of B\*B' (Stability Index) is 0.327

Indirect Effects of Y on Y

	GRADES	EDUCEXP	OCCUASP
	-----	-----	-----
GRADES	- -	- -	- -
EDUCEXP	- -	- -	- -
OCCUASP	0.218 (0.023) 9.661	- -	- -