STRUCTURAL-EQUATION MODELS: A CASE STUDY

DAVID A. FREEDMAN

TECHNICAL REPORT NO. 22

REVISED MAY 1983

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY

RESEARCH PARTIALLY SUPPORTED BY

NATIONAL SCIENCE FOUNDATION

GRANT MCS-80-02535

May, 1983

Structural-equation models: a case study

by

D. A. Freedman[1]
Statistics Department
University of California, Berkeley

Abstract. In 1967, Blau and Duncan proposed a structural-equation model for occupational stratification. This is one of the most successful, and influential, applications of the techniques to social data. A point-by-point critique of the model is offered here, with a view to suggesting that the method of structural equations is not so useful in analyzing complex social phenomena.

Key words and phrases: Structural-equation models, causal models, path analysis, regression

[1]Research partially supported by National Science Foundation Grant MCS-80-02535. Persi Diaconis, David Hopelain, Allyn Romanow and Amos Tversky made a large number of helpful, if irritating, suggestions. William Navidi did the computer work.

Table of Contents

Section	Title	Page
1.	Introduction	1
2.	Background of the model	2
3.	Model specification	4
4.	Stochastic assumptions	7
5.	Some empirical tests	11
6.	How is the model used?	14
7.	Interpreting path coefficients	16
8.	Alternative stochastic assumptions	19
9.	What are the alternatives?	21
10.	Other literature	22
11.	Conclusion	23
References		24

1. Introduction

Structural-equation models are widely used in the social sciences, to explicate complex cause-and-effect relationships. Despite their popularity, I do not believe these models have in fact created much new understanding of the phenomena they are intended to illuminate; instead, the models seem to have created a lot of confusion. The main difficulty is that investigators tend to ignore the stochastic assumptions behind the models, perhaps because these assumptions would in practice be hard to validate. I will argue my view in detail for only one example: the model for social stratification developed by Blau and Duncan (1967). This was one of the first and most successful applications of the method, and the paradigm for much subsequent research, as any recent issue of The American Sociological Review will demonstrate. It was singled out for praise by e.g. Bielby-Hauser (1977, p. 137), or Adams-Smelser-Treiman (1982, p. 46), and even the skeptical Coser (1975, p. 694). However, the weaknesses in the method emerge clearly in the present example.

2. Background for the model

Blau and Duncan consider (p. vii) that "Men's careers occupy a dominant place in their lives today, and the occupational structure is the foundation of the stratification system of contemporary industrial society." They go on to say (pp. 19-20):

> The basic question is how the status individuals achieve in their careers is affected by the statuses ascribed in them earlier in life .... The basic model of [this] process ... is presented in Chapter 5. Occupational status in 1962, the survey date, is conceived as the outcome of a lifelong process in which ascribed positions at birth, intervening circumstances, and earlier attainments determine the level of ultimate achievement. A formalization in terms of a simple mathematical model permits an approximate assessment of the relative importance of the several measured determinants.

It is this "simple mathematical model" which will be analyzed in the present paper. Blau and Duncan do proceed to develop more complex models in later chapters, but the additional complexities seem irrelevant here.

Blau and Duncan fit their model to data collected during the March, 1962 Current Population Survey, on a nationwide probability sample of about 20,000 American men aged 20-64. This is a highly stratified area cluster sample of households, the ultimate sampling unit being 4 to 6 physically adjacent households. For an overview of the sample, see Freedman-Pisani-Purves (1978, Chapter 22); for more technical discussions, see U.S. Bureau of the Census (1963, 1978).

Modeling is often defended on the grounds that it is only "data analysis." That defense does not apply here. Blau and Duncan view their research (p. 1) as having "significant implications for social policy and action programs." They see the model as a tool for making inferences from the data, about complex social phenomena. The model is what ties the data to the phenomena, and the inferences depend for their validity on the assumptions in the model. Blau and Duncan are serious, and we in turn are obliged to be serious in our scrutiny of their assumptions.

3. Model specification

The model proposed by Blau and Duncan (pp. 165-170) involves five variables:

V	Father's educational level
X	Father's occupational status
U	Respondent's educational level
W	Status of respondent's first job
Y	Status of respondent's job in 1962

For convenience, I will standardize these variables to have mean 0 and variance 1.

To specify a model, it is necessary to write the equations down; what Blau and Duncan seem to have in mind is the following:

(1a) $$U_i = a\,V_i + b\,X_i + \delta_i$$

(1b) $$W_i = c\,U_i + d\,X_i + \varepsilon_i$$

(1c) $$Y_i = e\,U_i + f\,W_i + g\,X_i + \eta_i$$

Here, ζ_i denotes the value of variable ζ for subject i. The coefficients a, b, c, d, e, f, and g are parameters which characterize the stratification process. These are called <u>path coefficients</u>. The set of equations is called a <u>path model</u>. The δ_i, ε_i, η_i are disturbance terms, to be discussed in the next section.

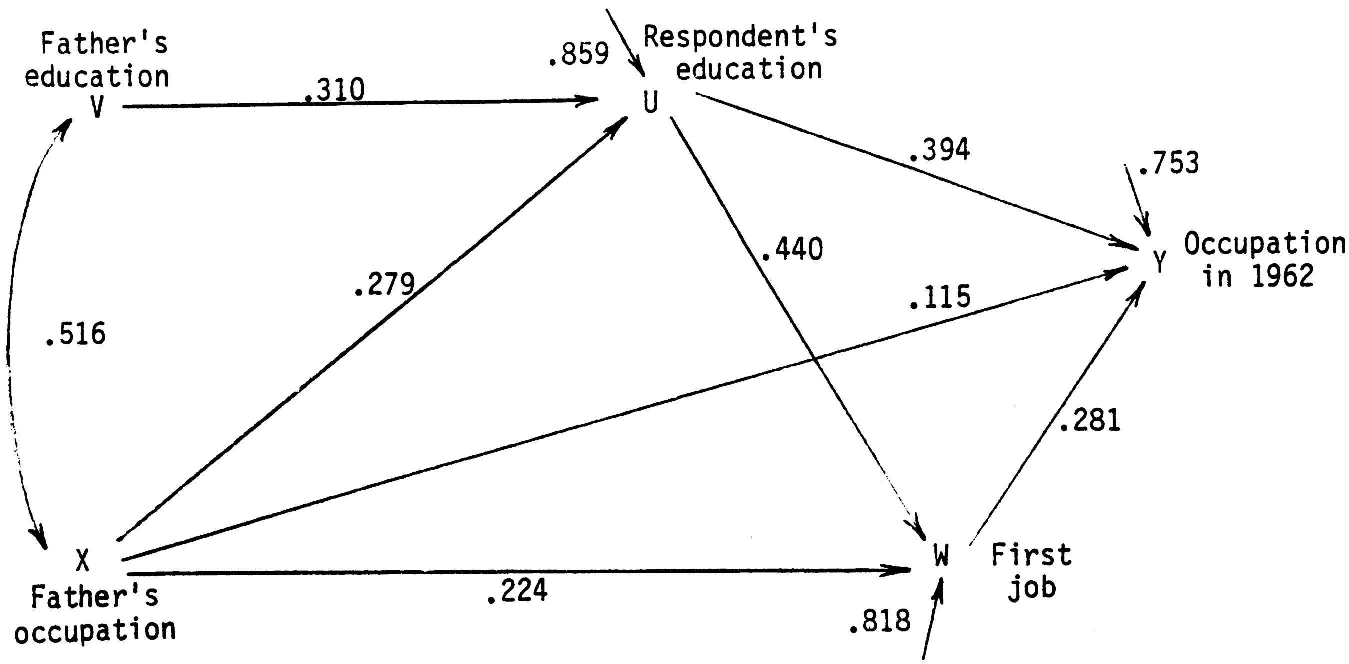


Figure 1. The Blau-Duncan path model for the process of stratification.

A <u>path diagram</u> is a graphical way of representing a path model. The Blau-Duncan path diagram is reproduced in Figure 1. The curved, two-headed arrow between V and X indicates that the two variables are correlated, but not related by any equation in the model; the empirical correlation of .516 is shown in the diagram. Next, take U for example. There are straight arrows leading to U from V and X; this is the graphical analog of (1a). The estimates for the coefficients are shown next to the arrows: $\hat{a} = .310$ and $\hat{b} = .279$. The free arrow pointing into U represents the disturbance term δ_i, whose standard deviation is estimated as .859. The rest of the diagram can be interpreted in a similar way.

In a good model, the equations have to be justified: the modelers have to make some showing that the structure of the equations reflects the structure of the phenomena. How are the equations in the present model derived? The only relevant considerations are presented on pp. 166-168. In effect, Blau and Duncan argue that V_i and X_i are determined prior in time to U_i; likewise, except "for an appreciable minority," U_i is determined prior to

W_i; and W_i is determined prior to Y_i. This is the whole of the argument.

How such an argument leads to the set of equations (1) is never made clear, but the primitive idea seems to be that since the equations are recursive, they can be solved in a way that parallels the assumed temporal ordering of the variables. Starting from V and X:

- Use (1a) to determine U.
- Then use (1b) to determine W.
- Finally, use (1c) to determine Y.

Of course, this parallelism does not establish (1) as the right model, since many other systems of equations have the same recursive structure. For example, the effects could be quadratic rather than linear, or multiplicative rather than additive. To sum up, the equations proposed by Blau and Duncan do not have any adequate theoretical foundation.

4. Stochastic assumptions

To complete the specification of a model, it is necessary to state the stochastic assumptions explicitly. This is particularly important here, since the disturbance terms account for more than half the variance. Blau and Duncan do not make their assumptions explicit, but seem to have the following in mind:

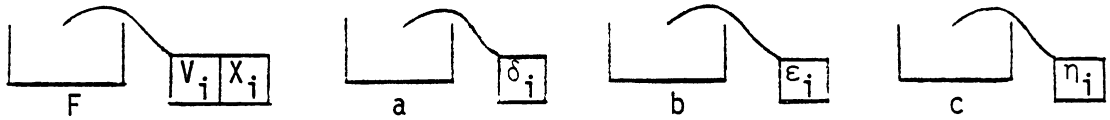
(2a) For the i^{th} respondent, V_i, X_i, δ_i, ε_i, η_i are random variables.

(2b) The 5-tuples $(V_i, X_i, \delta_i, \varepsilon_i, \eta_i)$ are independent and identically distributed across subjects.

(2c) (V_i, X_i), δ_i, ε_i, η_i are independent within each subject.

(2d) $E(\delta_i) = E(\varepsilon_i) = E(\eta_i) = 0$

To make the assumptions more vivid, introduce 4 boxes of tickets, labelled F, a, b, c. Each ticket in box F is marked with a pair of numbers; the average over the whole box of the first number in the pair is 0, with a standard deviation of 1; likewise for the second number. Each ticket in box a) is marked with a single number; the average over the whole box of these numbers is 0, with a standard deviation of 1. Likewise for boxes b) and c). For each subject i, we draw a ticket at random with replacement from each of the four boxes. The ticket for box F gives V_i and X_i; the other three tickets give the disturbance terms δ_i, ε_i and η_i. Then the system (1) is used to compute U_i, W_i and Y_i.



In a good model, the stochastic assumptions have to be justified. I will consider first the assumption in (2b) that (V_i, X_i) are independent and identically distributed (i.i.d) across subjects. This assumption is needed to justify the path calculus of correlation coefficients: see pp. 172ff of Blau and Duncan or section 4 below. The randomness in V_i and X_i seems to be viewed as due to sampling. If the sampling were done at random with replacement, the i.i.d. assumption would be fully justified. With a household cluster sample of the kind actually used to collect the data, the i.i.d. assumption is seriously in error. The clusters introduce correlations, which may be appreciable: people who live in the same household tend to be like each other. I think this is quite damaging to the logic of the model; however, section 8 below presents an alternative formulation, treating V_i and X_i as nonrandom data.

Next, I turn to the assumptions on the disturbance terms. Blau and Duncan view such terms as representing the effect of omitted variables. They argue (p. 175) as follows:

> The relevant question about [the disturbance term is] ... whether the unobserved factors it stands for are properly represented as being uncorrelated with the measured antecedent variables .... A delicate question in this regard is that of burden of proof. It is all too easy to make a formidable list of unmeasured variables that someone has alleged to be crucial to the process under study. But the mere existence of such variables is already acknowledged by the very presence of the residual. It would seem to be part of the task of the critic to *show*, if only hypothetically, but *specifically*, how the modification of the causal scheme to include a new variable would disrupt or alter the relationships in the original diagram. His argument to this effect could then be examined for plausibility and his evidence, if any, studied in terms of the empirical possibilities it suggests.

In other words, Blau and Duncan require their hypothetical critic to accept the causal-modeling framework, and to make statistical arguments about possible correlations in the disturbance terms within that framework. The end result would be a somewhat more complex model, with an explicit representation of some of the missing variables. I do not care to accept the burden, since my purpose is to question the whole framework.

With this declared, I turn to the assumptions about the disturbance terms. These can hardly be taken literally. For example, there are famous dynasties in banking, politics, and film; surely the corresponding δ's, ε's, and η's all tend to have the same sign. By virtue of its cluster design, the sample is bound to include brothers, relatives, neighbors; dynastic considerations are by no means irrelevant.

I want to demonstrate the correlations in the disturbance terms a bit more generally. Take for example the respondents whose fathers were simultaneously in the highest educational level and had the highest occupational status: $V_i = 8$ and $90 \leqq X_i \leqq 96$. (See pp. 122 and 166.) This will include fathers who got a degree in chiropractic medicine from Oral Roberts University and practice in Manhattan, Kansas; as well as fathers who got a law degree from Harvard, and practice in Manhattan, New York. The sons in the two groups may be expected to have different career trajectories, and correlated disturbances δ_i, ε_i, η_i.

Still more generally, the model omits families, neighborhoods, geographical regions. It does not consider the quality of education, or when education was obtained, or when respondents entered the labor force: for a discussion of the relevance of such variables, see the Council of Economic Advisers (1974) or Polachek (1976). There is nothing of history in the model, and nothing of the economy. But the respondents were between 20 and 64 years old in the survey year. They or their parents lived through two world wars and the great depression, to name only three salient events. Do these have no effect on career trajectories? Or are their effects independent of the measured variables? Blau and Duncan seem to be making rather bold conjectures.

It may seem that my objections to the model are rather picky: that I am another purist trying to stop honest applied work: that, in short, my position is too academic. After all, what difference can it make if the disturbance terms were correlated? The answer is, a very large difference. With the independence assumptions (2), the path coefficients are appropriately estimated by ordinary least squares, the technique used by Blau and Duncan. If the disturbance terms are correlated, however, ordinary least squares may be badly biased, due to the recursive structure. For example, suppose the disturbance terms are independent of V and X, but δ is correlated .5 with ε --a very modest departure from (2). Assuming the rest of the Blau-Duncan numbers, the estimate of c is about 60% too big, while the estimate for d is about 15% too small. Assumptions matter, and the Blau-Duncan assumptions are rather fanciful.

5. Some empirical tests

So far, I have focused on the assumptions of the model; in detail, these seem quite unreasonable. Perhaps, however, the model works despite bad assumptions--for example, the model may make interesting and verifiable predictions. It is a fact, however, that Blau and Duncan do not make any predictions at all from their model. Their empirical argument is only that the model fits their data. Therefore, let us consider the fit.

Blau and Duncan do some data analysis in Chapter 4. A fair summary is that the data clearly but narrowly violate the assumptions of the model: the regression curves are nonlinear (pp. 137, 144); the residuals are heteroscedastic (pp. 139, 144); the slopes vary across subgroups (p. 148). The path coefficients in (1) therefore have no real existence. What are Blau and Duncan talking about?

In effect, their model ignores the data analysis. And the only empirical test they propose is whether the model "satisfactorily accounts for the observed correlations." The argument depends on a "path calculus," or formulas relating correlations to path coefficients. If (ξ_i, ζ_i) are i.i.d. across subjects i, let $r_{\xi\zeta}$ denote the theoretical correlation coefficient:

$$r_{\xi\zeta} = \frac{E(\xi_i \zeta_i) - E(\xi_i)\, E(\zeta_i)}{\sqrt{\operatorname{var} \xi_i} \cdot \sqrt{\operatorname{var} \zeta_i}}$$

where

$$\operatorname{var} \xi_i = E(\xi_i^2) - E(\xi_i)^2$$

and E stands for expectation.

Take r_{VW}, for example. I have standardized all 5 variables to have theoretical mean 0 and variance 1. Thus

$$r_{VW} = E(V_iW_i)$$

$$= cE(U_iV_i) + dE(X_iV_i) \qquad \text{by (1b) and (2c, d)}$$

That is,

(3) $$r_{VW} = cr_{UV} + dr_{XV}$$

This is an example of "path calculus."

On the data side, let $\hat{r}_{\xi\zeta}$ denote the sample correlation coefficient:

$$\hat{r}_{\xi\zeta} = \frac{\frac{1}{n}\Sigma_{i=1}^{n}\xi_i\zeta_i - (\frac{1}{n}\Sigma_{i=1}^{n}\xi_i)(\frac{1}{n}\Sigma_{i=1}^{n}\zeta_i)}{s_\xi\ s_\zeta}$$

where

$$s_\xi^2 = \frac{1}{n}\Sigma_{i=1}^{n}\xi_i^2 - (\frac{1}{n}\Sigma_{i=1}^{n}\xi_i)^2$$

Denote the ordinary least squares estimates by hats.

To test the goodness of fit of their model, Blau and Duncan (p. 173) look at the statistic suggested by (3):

$$\hat{T} = \hat{r}_{VW} - [\hat{c}\,\hat{r}_{UV} + \hat{d}\,\hat{r}_{XV}]$$

$$= .332 - [.440 * .453 + .224 * .516]$$

$$= .332 - .315$$

$$= .017$$

Since $\hat{T}$ is small, they conclude that the model fits the data.

Their $\hat{T}$ is indeed small. But is it statistically significant or not? To answer this question, a "bootstrap experiment" can be performed to attach a standard error to $\hat{T}$: see Efron (1979) or Bickel and Freedman (1981) or

Freedman and Peters (1983). The idea is simple. Assume the model (1-2); take as correct the Blau-Duncan estimates for the path coefficients, for the correlation between U and V, and for the variances of δ_i, ε_i, and η_i. For simplicity, assume all the variables are gaussian.

This now completely specifies a simulation model, which can be used to generate "data" on 20,000 subjects--the size of the Blau-Duncan sample. From these simulated data, correlations and path coefficients can be estimated: denote such estimates by stars. Then compute

$$\hat{T}^* = \hat{r}^*_{VW} - [\hat{c}^* \hat{r}^*_{UV} + \hat{d}^* \hat{r}^*_{XV}]$$

Thus, $\hat{T}^*$ is the Blau-Duncan statistic, computed from simulated data, generated exactly according to their stochastic model. Repeating this process 100 times, the mean of $\hat{T}^*$ is nearly 0, with a standard deviation of 5×10^{-3}. Thus, $\hat{T}$ is quite unusual, being more than 3.3 SD's away from average; indeed, none of 100 replications had $|\hat{T}^*| \geq |\hat{T}|$. The conclusion: the model narrowly but definitely fails the first statistical test proposed by Blau and Duncan.

Setting this objection aside, the model has 7 coefficients, as well as the correlation between V and X, for a total of 8 parameters. The 5 variables admit 10 observed correlations. Thus Blau and Duncan are using 8 parameters to reproduce 10 observations. Good agreement therefore is not surprising.[1]

---

[1]On their behalf, I do have to say that the Blau-Duncan model reproduces the observed correlations better than alternative 8-parameter path models which I fitted. For a more general critique of the correlation test, see Rogosa (1983).

6. How is the model used?

What sort of conclusions do Blau and Duncan draw from their model? A main one is stated clearly on pp. 402-403:

> Thus the entire influence of father's education on son's occupational status is mediated by father's occupation and son's education.

It seems difficult to interpret this statement except in terms of a path diagram, with no arrows from V to W or Y. Thus, the main assertion drawn from the model relates only to the model. This is perplexing for one who does not accept the model. Still, some analysis is possible.

The Blau-Duncan path diagram in Figure 1 shows no arrow from V to W, and no arrow from V to Y. This is so by assumption. Blau and Duncan just do not draw those arrows. Indeed, they are quite frank on p. 177.

> The technique of path analysis is not a method for discovering causal laws but a procedure for giving a quantitative interpretation of the manifestation of a *known or assumed* causal system .... [emphasis supplied]

The conclusion at issue is that father's education V has no direct influence on son's occupational status (W, Y). Is this totally without empirical proof? Not quite. Blau and Duncan have 3 pieces of evidence weakly supporting their conclusion:

i) The correlation coefficients computed from their path model are close to the observed ones.

ii) In a regression of Y on V, X, U and W, the coefficient of V is only -.014.

iii) In a regression of W on V, X and U, the coefficient of V is only .026: see p. 174.

I have already discussed i), and now take up ii) and iii). It is easy to compute standard errors for regression coefficients. With the Blau-Duncan sample size of 20,000, in a regression of Y on V, X, U, and W, the standard error for the coefficient of V is .0065. The observed coefficient of -.014 is 2.2 times its standard error, and is significant. Likewise, in a regression of W on V, X, and U, the standard error for the coefficient of V is .0071; the observed coefficient of .026 is 3.7 times its standard error, and is highly significant. This would seem to dispose of arguments ii) and iii).

The conclusion drawn by Blau and Duncan is unwarranted. Remember, they say they prove that

> ... the entire influence of father's education on son's occupational status is mediated by father's occupation and son's education.

But a fair statement of their results is only as follows: Roughly, the data conform to the equations (1) and (2), as depicted in the path diagram, although the differences are highly significant. In the path diagram, by assumption, father's education has no direct influence on son's occupation; no arrows were drawn from V to W or Y. Blau and Duncan seem to have been misled by their methodology into confusing assumptions with conclusions.

7. Interpreting path coefficients

Even given the model, I find "path coefficients" like a and b in (1) quite hard to interpret. Blau and Duncan see no problem. For them, the path coefficient measures the "direct effect" of one variable on the other, with no further qualifications. They are quite literal about it: see e.g. p. 176 or p. 201. Indeed, the charm of structural-equation models for sociologists appears to stem from the "direct effect" interpretation. For example, Coser (1975, p. 694) writes:

> Stratification studies [have] benefited a great deal from modern path analytical methods whose power is perhaps shown at its best in Blau and Duncan's *The American Occupational Structure* (1967). Path analysis allows these authors systematically to trace the impact of such factors as father's occupation, father's educational attainments, and son's education and first job on the son's placement in the occupational hierarchy. It allows for the first time the assessment in precise detail of the ways in which occupational status in a modern industrial society is based on a combination of achieved and ascribed characteristics.

The interpretation of path coefficients as measures of direct effects is often attributed to Boudon (1965), who in turn quotes Wright (1934) as saying that a path coefficient measures

> ... the fraction of the standard deviation of the dependent variable (with appropriate sign) for which the designated factor is directly responsible, in the sense of the fraction which would be found if this factor varies to the same extent as in the observed data while all others (including residual factors) are constant.

Boudon (p. 370) goes on to say that a path coefficient (with standardized variables) is "really a measure of the *direct* influence of one variable on another in a causal scheme."

To be more definite, consider e.g. the regression of U on V and X, where all 3 are random variables standardized to have mean 0 and variance 1:

(4) $$U = aV + bX + \delta$$

Here

(5) $$E(U) = E(V) = E(X) = 0 \qquad \text{(mean 0)}$$

(6) $$E(U^2) = E(V^2) = E(X^2) = 1 \qquad \text{(variance 1)}$$

The error term δ is required to be orthogonal to V and X:

$$E(\delta V) = E(\delta X) = 0$$

It follows from (4) that $E(\delta) = 0$. As before, E is the expectation operator.

Consider now equation (4) from Boudon's point of view. It is the case that U has standard deviation 1. Furthermore, U is the sum of the three terms aV, bX, and δ; the first has standard deviation $|a|$, the second $|b|$. So far, so good. But the rest of Wright and Boudon's interpretation seems to me a sophistry; I will indicate my reasons by example.[1]

Suppose U, V, and X are all functions of a more primitive variable ζ, which is uniformly distributed over $[0, 1]$. More specifically, let V be ζ, let X be ζ^2, and U be ζ^3, but standardized to have mean 0 and variance 1:

(7) $$V = 2\sqrt{3}\,(\zeta - \tfrac{1}{2}) \qquad X = \tfrac{1}{2}\sqrt{45}\,(\zeta^2 - \tfrac{1}{3}) \qquad U = \tfrac{4}{3}\sqrt{7}\,(\zeta^3 - \tfrac{1}{4})$$

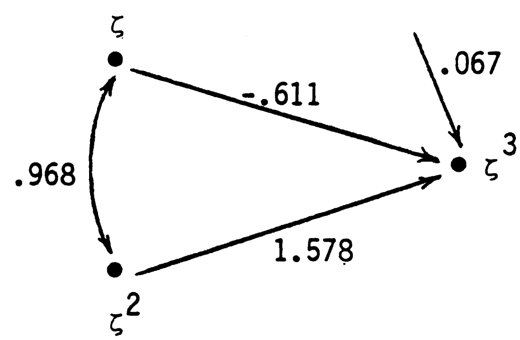
Multiply both sides of (4) by V and by X, and take expectations. This gives two linear equations in a and b. Solving,

[1] Compare Mosteller and Tukey (1977, Chapter 13). Also see Pratt and Schlaifer (1981).

$$a = -\frac{2}{15}\sqrt{21} \approx -.611 \qquad b = \frac{4}{15}\sqrt{35} \approx 1.578$$

The path diagram is shown in figure 2.

Figure 2. Path diagram for the polynomials. Rationale: ζ and ζ^2 occur before ζ^3 in the list of monomials.



Do we really want to say that the direct effect of ζ on ζ^3 is -.611, while the direct effect of ζ^2 on ζ^3 is 1.578? How can we vary ζ while keeping ζ^2 fixed? Or ζ^2 while fixing ζ?

The idea must be that structural equations are different from this artificial example. We need to have the difference spelled out. And then we need some showing that the Blau-Duncan equations really are "structural" in the relevant sense.

8. Alternative stochastic assumptions

The object of this section is to indicate two sets of stochastic assumptions alternative to (2), which might be thought to salvage the Blau-Duncan position, but do not. One idea is to condition on V_i and X_i, in effect treating them as data. The relevant assumptions become as follows: for the i^{th} respondent, V_i and X_i are non-random, but δ_i, ε_i, η_i are random variables. The 3-tuples $(\delta_i, \varepsilon_i, \eta_i)$ are independent and identically distributed across subjects. However, δ_i, ε_i, η_i are independent within each subject. Finally, $E(\delta_i) = E(\varepsilon_i) = E(\eta_i) = 0$. This set of assumptions is what statisticians now call "superpopulation theory."

With this set of assumptions, the variables V, X, U, W, Y should be standardized a bit differently; the path calculus goes through, if correlation is replaced by the appropriate "inner product;" details are omitted. These modifications leave most of the Blau-Duncan argument intact; nor do they affect my objections to that argument.

Another idea is to view each person in the population as having the 5 characteristics; define the coefficients in (1) by regression at the level of the population. The randomness in the observables is then due just to the sampling. There are several difficulties with this approach:

i) Ordinary least squares is not the appropriate estimation procedure: the sample design must be taken into account.

ii) The disturbance terms will be correlated across equations, so the path calculus does not go through.

iii) The model is not "structural" in the sense used by Duncan (1975 p. 151):

> The structural form of the model is that parameterization--among the various possible ones--in which the coefficients are (relatively) unmixed, invariant, and autonomous ... if the coefficients in the model are indeed relatively invariant across [sub] populations, somewhat autonomous, and not inseparable mixtures of the coefficients that "really" govern how the world works--then your model is actually in the "structural" form.

In the sampling model, the regression coefficients are "inseparable mixtures" of the numbers "that 'really' govern how the world works."

9. What are the options?

The strongest caveat I could find in Blau and Duncan is on p. 172:

> We are a long way from being able to make causal inferences with confidence, and schemes of the kind presented here had best be regarded as crude first approximations to *adequate causal models*. [emphasis supplied]

For Blau and Duncan, we go from simple models to more complicated ones.

My view is different. I think Blau and Duncan would have written a better book without the models. They have interesting questions, and interesting data. They succeed very well in documenting the great inter-generational mobility across occupations. (See Chapter 2.) The data analysis in Chapter 4 is revealing. Some of the regression equations are suggestive, if not pushed too hard: e.g. in a regression of Y on V, X, U, and W, the coefficient of V is small and negative. Why not stop there?

The apparent power and sophistication of path models have beguiled even Blau and Duncan. Later investigators have forgotten the assumptions and limitations of the method; they make very literal or even naive interpretations of model results; and they are distracted from other lines of inquiry. In sum, path models may be worse than the alternative of no path models.

10. Other literature

I have argued my position at some length in Freedman (1983). For other critical reviews of causal modeling, see Baumrind (1983) or Ling (1983). Also see Karlin (1979) on path models in genetics. For related issues in handling experimental data, see Zeisel (1982). Similar issues for econometric models are discussed in Freedman (1981) and Freedman-Rothenberg-Sutch (1983a, b); also see Christ (1975), Hausman and Wise (1982), Hendry (1979), Leamer (1983), Lucas and Sargent (1978), Sims (1980, 1982).

11. Conclusion

I invite a reply from proponents of structural-equation models. If I have the Blau-Duncan model wrong, let them state it correctly, but at the same level of clarity and detail as in my exposition. Let them argue the specification, and justify the stochastic assumptions. Let them draw some conclusions from the model, particularly conclusions testable by empirical observation. On the other hand, if the Blau-Duncan model is no longer considered such a success, let the proponents of the method point to a better example.

I will close by quoting Duncan himself (1975, pp. 150-151):

> Do not undertake the study of structural equation models ... in the hope of acquiring a technique that can be applied mechanically to a set of numerical data with the expectation that the result will automatically be 'research.' Over and over again, sociologists have seized upon the latest innovation in statistical method, rushed to their calculators or computers to apply it, and naively exhibited the results as if they were contributions to scientific knowledge. The lust for 'instant sociology,' the superstition that it is to be achieved merely by a complication if not perfection of formal or statistical methods, and the instinct to suppose that any old set of data, tortured according to the prescribed ritual, will yield up interesting scientific discoveries--all these pathological habits of thought are grounded ... in the fallacy of induction. It might appear that the literature of sociological investigations using structural equation models provides just another episode of seizing upon the latest methodological fad in the incessant quest for the formula for instant sociology. Your present author would not care to defend too vigorously the contrary view. But if any of these sociological examples *are* contributions to science (and not merely exercises in quantitative technique), it is because the models rest on creative, substantial, and sound sociological theory.

My point is that the equations of path models, let alone the stochastic assumptions, have no basis in "sound sociological theory."

REFERENCES

[1] R. McC. Adams, N. J. Smelser, and D. J. Treiman eds. (1982). _Behavioral and Social Science Research_: _A National Resource_. Vols. I and II. National Academy Press, Washington, D.C.

[1a] D. Baumrind (1983). Specious causal attributions in the social sciences: the reformulated stepping-stone theory of heroin use as exemplar. _J. of Personality and Social Psychology_. To appear.

[2] P. Bickel and D. Freedman (1981). Some asymptotic theory for the bootstrap. _Ann. Statist._, Vol. 9, pp. 1196-1217.

[3] W. T. Bielby and R. M. Hauser (1977). Structural equation models. _Ann. Rev. Sociol._, Vol. 3, pp. 137-161.

[4] P. M. Blau and O. D. Duncan (1967). _The American Occupational Structure_. Wiley, New York.

[5] R. Boudon (1965). A method of linear causal analysis: dependence analysis. _Am. Soc. Rev._, Vol. 30, pp. 365-374.

[6] Council of Economic Advisors (1974). Economic Report of the President. Washington, D.C.

[7] C. Christ (1975). Judging the performance of econometric models of the United States Economy. _International Economic Review_, Vol. 16, pp. 54-74.

[8] L. Coser (1975). Presidential address: two methods in search of a substance. _Am. Soc. Rev._, Vol. 40, pp. 691-700.

[9] O. D. Duncan (1975). Introduction to Structural Equation Models. Academic Press, New York.

[10] O. D. Duncan, D. L. Featherman and B. Duncan (1972). Socioeconomic Background and Achievement. Seminar Press, New York.

[11] B. Efron (1979). Bootstrap methods: another look at the jackknife. Ann. Statist., Vol. 7., pp. 1-26.

[12] D. L. Featherman and R. M. Hauser (1978). Opportunity and Change. Academic Press, New York.

[13] D. Freedman (1981). Some pitfalls in large econometric models: a case study. J. of Business, Vol. 54, pp. 479-500.

[14] D. Freedman (1983). Statistics and the scientific method. Technical report no. 19. Department of Statistics, University of California, Berkeley. To appear in Cohort Analysis in Social Research: Beyond the Identification Problem. Eds. W. Mason and S. Fienberg, Academic Press, New York.

[15] D. Freedman and S. Peters (1983). Bootstrapping a regression equation: some empirical results. Technical report no. 10. Department of Statistics, University of California, Berkeley. To appear in J. Amer. Statist. Assoc.

[16] D. Freedman, R. Pisani and R. Purves (1978). Statistics. Norton, New York.

[17] D. Freedman, T. Rothenberg and R. Sutch (1983a). On energy policy models. J. of Business and Economic Statist., Vol. 1, pp. 24-36.

[18] D. Freedman, T. Rothenberg and R. Sutch (1983b). A review of a residential energy end use model. Technical report no. 14, Department of Statistics, University of California, Berkeley.

[19] J. Hausman and D. Wise (1982). Technical problems in social experimentation: cost versus ease of analysis. N.B.E.R. Working Paper No. 1061.

[20] D. Hendry (1979). Econometrics - alchemy or science? Inaugural lecture, London School of Economics. Econometrica, November 1980, Vol. 7, pp. 387-406.

[21] S. Karlin (1979). Comments on statistical methodology in medical genetics. pp. 497-520 in Genetic Analysis of Common Diseases: Application to Predictive Factors in Coronary Disease. Alan R. Liss, Inc., New York.

[21a] E. Leamer (1983). Taking the 'con' out of econometrics. American Economic Review. March , pp. 31-43.

[21b] R. F. Ling (1983). Review of Correlation and Causation by Kenny. J. Amer. Statist. Assoc. Vol. 77, pp. 489-491.

[22] R. E. Lucas and T. J. Sargent (1978). After Keynesian macro-economics. In After the Phillips curve: Persistence of High Inflation and High Employment, pp. 49-72, Conference Series No. 19, Federal Reserve Bank of Boston.

[23] F. Mosteller and J. W. Tukey (1977). Data Analysis with Regression. Addison-Wesley, Reading.

[24] S. W. Polachek (1976). Occupational segregation: an alternative hypothesis. J. Contemporary Business.

[25] J. W. Pratt and R. Schlaifer (1981). On the nature and discovery of structure. Technical report. Harvard Business School.

[25a] D. Rogosa (1983).

[26] C. A. Sims (1980). Macro-economics and reality. Econometrica, Vol. 48, pp. 1-48.

[27] C. A. Sims (1982). Scientific standards in econometric modelling, paper presented at the symposium on the developments in econometrics and related fields. Netherlands Econometric Institute.

[28] U.S. Bureau of the Census (1963). The Current Population Survey - a report on methodology. Technical Paper No. 7.

[28] U.S. Bureau of the Census (1978). The Current Population Survey - a report on methodology. Technical Paper No. 40.

[30] Sewall Wright (1934). The method of path coefficients. Ann. Math. Statist., Vol. 5, pp. 161-215.

[31] H. Zeisel (1982). Disagreement over the evaluation of a controlled experiment. American Journal of Sociology, Vol. 88, No. 2, pp. 378-389.