This paper discusses path analysis and the applicability of this methodology to the field of family studies. The statistical assumptions made in path analysis are presented along with a description of the two types of models within path analysis, i.e., recursive and non-recursive. Methods of calculating in the path model and the advantages of using the path model over multiple regression are presented in a series of tables which illustrate each step in the process. The second section of the paper discusses the application of path analysis to family studies in three situations: family crisis, divorce, and marital power. Characteristics of each of the three hypothetical situations are described in detail and the path model of research for each situation is discussed and graphically illustrated in accompanying figures. The conclusion emphasizes path analysis as the link providing the needed reciprocity between theory and research. (MCF)
PATH ANALYSIS: A LINK BETWEEN FAMILY THEORY AND RESEARCH*

Mark R. Rank
Department of Sociology
University of Wisconsin-Madison

Ronald M. Sabatelli
Department of Child and Family Studies
University of Wisconsin-Madison

*Paper presented at the NCFR Pre-Conference Workshop on Theory Construction and Research Methodology, Washington, D.C., October 12-13, 1983. Analysis was aided by a "Center for Population Research" grant, No. HD05876, to the Center for Demography and Ecology, University of Wisconsin-Madison, from the Center for Population Research of the National Institute of Child Health and Human Development.
Path analysis provides the family scholar with an excellent opportunity to link theory and research in a dynamic process. The purpose of this paper is therefore twofold: 1) To describe and illustrate the assumptions, methods, and advantages associated with path analysis; and 2) To discuss the applicability of path analysis to the field of family studies. It is argued that it may be to our benefit to view theory and research as distinct but permeable components in a process of understanding. Path analysis allows one to bridge these two components, hopefully resulting in a clear and insightful analysis of a specific phenomenon.
During the past 20 years, the area of family studies has seen a significant increase in both theoretical and methodological development. Increasing methodological sophistication (e.g. multiple regression, factor analysis, log linear techniques, etc.) has been matched by the construction of middle range theories and broader conceptual frameworks. Yet throughout this development, relatively little attention has been devoted to the relationship between research and theory. In *Social Theory and Social Structure* (1968), Robert K. Merton advocated a reciprocity between theory and research, that is, theory should guide research, and research in turn, should guide theory. While Merton's advice has often been ignored, path analysis provides the family scholar with an excellent opportunity to link theory and research. The purpose of this paper therefore is to discuss the applicability of path analysis to the field of family studies. In order to accomplish this, we first describe and illustrate the assumptions, methods, and advantages associated with path analysis. We then focus on how path analysis may in fact further both family theory and research by demonstrating its applicability to several substantive areas.

**PATH ANALYSIS AND CAUSAL MODELLING**

Although Sewall Wright (1921) laid the foundations for path analysis and causal modelling 60 years ago; it has not been until recently that path analysis has found its way into the social sciences. Likewise, it has only been in the last decade that family researchers have utilized path analysis as a technique for analyzing data (for a review, see Schumm et al., 1980). However, there is no doubt that path analysis can offer the family researcher an important methodological tool to be utilized.

Currently there are a number of excellent sources reviewing the rudiments and dynamics underlying path analysis (i.e. Duncan, 1975). It is therefore not our purpose to expound on what has already been covered in various statistical texts. Rather, by briefly reviewing and highlighting the assumptions, methods, and advantages of this statistical technique, it is hoped that the reader will be able to determine if path analysis is indeed an appropriate procedure with respect to analyzing his or her data.

**Assumptions**

We begin by emphasizing that path analysis is an extension of multiple regression. Whereas in a regression analysis we have one dependent variable, in path analysis we may have several dependent variables within our path model. However the procedure to estimate the path coefficients for a specific dependent variable, is identical to the procedure used to calculate the beta coefficients in a regression analysis—ordinary least squares. Consequently, many of the statistical assumptions made in path analysis (specifically recursive path models) are identical to those made in multiple regression.

There are two types of models within path analysis—recursive and nonrecursive. By far the majority of applied research has focused upon recursive models.
models, and hence we shall limit the scope of this paper primarily to recursive path models. However, it is important to recognize the difference between these two types of models. A recursive path model is one in which the causal flow is unidimensional. Thus, within a recursive model, a variable cannot be both a cause and an effect of another variable (Kerlinger and Pedhazur, 1973). Figure 1 is an example of a recursive path model. Consequently, if there had been a path leading from variable 4 to variable 3, Figure 1 would then have represented a nonrecursive model, and would have violated several of the assumptions in this section. If our theoretical model assumes a reciprocal relationship among dependent variables, we violate the assumptions of ordinary least squares, and are forced to use alternative procedures such as two-stage least squares in order to calculate the appropriate paths. The presence of reciprocality among dependent variables implies a nonrecursive model.

X1 and X2 represent exogenous variables, that is, the causes of these variables are assumed to lie outside the model itself, while X3 and X4 are considered endogenous because the assumed causal direction is implied within the model. U and V represent error terms, or that amount of variance which is left unexplained in the endogenous variable. For example, V is the amount of unexplained variation in variable X4 that is not accounted for by variables X1, X2, and X3.

It should be clear from the path diagram in Figure 1, that an important assumption of path analysis (both recursive and nonrecursive models) is that the researcher possess a theoretical framework that guides the construction of a path diagram. As Kerlinger and Pedhazur note, "path analysis is useful in testing theory rather than in generating it. In fact, one of the virtues of the method is that in order to apply it the researcher is required to make explicit the theoretical framework within which he operates" (1973:305). Thus, path analysis enables the researcher to test the applicability of a theoretical design. Depending on how well the model fits the data, the researcher may feel satisfied with the path diagram, modify it, or abandon it in favor of an alternative design.

A second assumption, is that the relationships among the variables in the model are linear, additive and causal (both in recursive and nonrecursive path models). Therefore, nonlinear relationships such as curvilinearity are assumed to be absent.

Third, error terms or residuals in a recursive path model are assumed to be uncorrelated with variables preceding them in the model, as well as uncorrelated among themselves. The presence of reciprocal relationships among endogenous variables clearly violates this assumption. Subsequently, nonrecursive path models violate the assumption of uncorrelated error terms.

Fourth, reciprocal causation between variables is nonexistent in a recursive
path model. Again, reciprocality among variables can be accommodated within a nonrecursive model (see Duncan, 1975).

Fifth, it is assumed that one is aware of the time ordering of variables. The researcher should therefore be confident concerning which variables precede and follow others. This is true for both recursive and nonrecursive path models.

Finally, endogenous variables are assumed to be measured on an interval scale (recursive and nonrecursive models). There are, however, a number of discussions which address the implications of utilizing ordinal rather than interval measures in path analysis (e.g. Land, 1969; Bohnstedt and Carter, 1971). We would urge researchers interested in using path analysis, to examine whether their data meets the above assumptions. If the answer is yes, one is able to proceed by calculating the coefficients in a path model.

Methods

Let us assume that the researcher has computed the zero order correlations among three variables. These are found in Table 1. $X_1$ measures education, $X_2$ measures income, and $X_3$ measures support for equal job opportunity for women. Initially, the researcher (by looking at the correlation matrix), may feel that education has little or no effect on attitudes. However, working from a particular theory, he/she has hypothesized that the underlying arrangement of variables is the following:

![Figure 2](image)

The structural equation for $X_3$ is:

$$X_3 = P_{31}X_1 + P_{32}X_2 + P_{3u}U$$

while the normal equations are:

$$r_{13} = P_{31}r_{11} + P_{32}r_{21} + P_{3u}r_{u1} = P_{31} + P_{32}r_{21}$$

$$r_{23} = P_{31}r_{21} + P_{32}r_{22} + P_{3u}r_{u2} = P_{31}r_{12} + P_{32}.$$
In order to estimate the paths, algebraic substitution is used in order to arrive at:

\[ P_{31} = r_{13} - r_{32}r_{21} \]

\[ P_{32} = \frac{r_{23} - r_{13}r_{12}}{1 - r_{12}^2} \]

By inserting the correlations, we arrive at the path coefficients:

\[ P_{32} = \frac{.375 - (0)(-.5)}{1 - (-.5)} = .5 \]

\[ P_{31} = 0 - (.5)(.5) = .25 \]

This results in the following path diagram:

What the researcher has discovered is that there is indeed a direct effect of \( X_1 \) on \( X_3 \), however, in the correlation matrix, \( X_2 \) was acting as a suppressor variable.

When using a standard statistical package such as SPSS, the calculation of paths is a relatively straightforward procedure. For example in Figure 3, in order to arrive at \( P_{21} \), one would regress \( X_1 \) on \( X_2 \). The beta coefficient for \( X_1 \) is synonymous with \( P_{21} \). One would then regress \( X_2 \) and \( X_1 \) on \( X_3 \) in order...
to estimate $P_{32}$ and $P_{31}$. Finally $X_3$, $X_2$, and $X_1$ would be regressed on $X_4$ to arrive at paths $P_{43}$, $P_{42}$, and $P_{41}$. Error or residual terms are calculated from each regression equation by simply taking the square root of $1 - R^2$.

Advantages

One of the clear advantages of path analysis over multiple regression is the ability to examine a correlation between two variables in terms of direct, indirect, and spurious effects, as well as associations between variables. Depending on the design of the path model, all of these relationships can be examined. For example, in Figure 4, $X_1$ has both a direct effect on $X_3$ ($P_{31}$) and an indirect effect ($P_{32}$($P_{21}$). In Figure 5 we have inserted hypothetical path coefficients into Figure 4. The direct effect of $X_1$ on $X_3$ is .4. The

![Figure 4](image)

indirect effect of $X_1$ on $X_3$ (operating though $X_2$) is .5 $\times$ .2 $=$ .1. Therefore the total effect of $X_1$ on $X_3$ is the direct effect (.4) plus the indirect effect (.1), or .5. A spurious relationship is indicated in Figure 6. Thus, while a change

![Figure 5](image)

in $X_2$ occurs with a change in $X_3$, the relationship is actually spurious, that is, $X_1$ is causing change in both $X_2$ and $X_3$. Finally, in Figure 7, variable $X_1$ not
only has a direct effect upon $X_3$ but there is also an association effect due to the correlation between $X_1$ and $X_2$. In short, these examples illustrate the complexity between variables that can be implied from path analysis, a complexity that multiple regression cannot attain. In path analytic terms, multiple regression is illustrated in Figure 8. Thus, path analysis could be thought of as an extension of multiple regression.

A second advantage of path analysis is the ability to test the adequacy of an entire theoretical model. Rather than having to focus on specific elements of a theory to the exclusion of others, path analysis allows one to examine various components of a model simultaneously.

A third advantage of path analysis is the ability to create a more parsimonious model through the deletion of paths. There are several rules of thumb that may be employed in deleting paths. One can compare differences between the original R matrix and one with deleted paths. If significant discrepancies are not apparent, the more parsimonious model may be preferable. Or, one might employ Heise's (1969) approach of "theory trimming." Thus one might judge the importance of a path by either a statistical or substantive criterion. If the path does not appear important, then it may be deleted from the model. By eliminating paths, one is able to "tighten up" a particular theory so that it may become more parsimonious with the data.

APPLICABILITY OF PATH ANALYSIS TO FAMILY STUDIES

Having reviewed the basic rudiments of path analysis, we now turn to several areas within family studies where the application of path analysis would seem particularly fruitful. These include: 1) family crisis; 2) divorce; and 3) marital power. Certainly there are a number of additional substantive areas in which path analysis might provide considerable insights, however we leave these investigations up to the reader.

Family Crisis.

As McCubbin et al. (1980) note, the area of family stress and crisis has been
one of the most productive areas in terms of theory construction within the field of family studies. Several well-developed theoretical models have been generated, most notably Reuben Hill's ABCX model (Hill, 1958). McCubbin et al. argue that Hill's ABCX model "has continued to serve as the foundation for the research and theory building efforts of the past decade of family stress investigation" (855). The model is composed of four elements. Hill's A component consists of the event or stressor itself. Obviously, a number of events or stressors occur to families which could lead to a crisis. These may include both normative events (e.g. the transition to parenthood, retirement, relocation, etc.) and non-normative events (e.g. occupational stress, chronic illness, long-term unemployment, etc.). The B component involves the family's crisis meeting resources and includes factors such as family adaptability and coherence, income, religious beliefs, education, good health, and so on. The third component of Hill's model (C), is the definition the family makes of the event. Thus, what is the perception and meaning attached to a particular event? Finally, these three elements (A, B, and C) produce X, or the crisis.

In path analytic terms, we could think of B, C, and X as interval variables. Thus, we might construct B to represent the extent of a family's crisis meeting resources ranging on a scale of low to high. Likewise, C could be constructed in a similar fashion, ranging from a lack of a serious perception of the event, to a very serious perception of the event. Finally, the extent of a crisis in the family (factor X) could also be operationalized in terms of an interval variable. A crisis scale might range from lack of crisis, to serious crisis. The event itself, or A, could be constructed as either categorical (the event exists or it does not) or interval (the degree of the event existing). Figure 9 diagrams Hill's ABCX model in terms of path analysis.

Figure 9.

A (the event) represents an exogenous variable, whereas B, C, and X are endogenous. Figure 9 represents a recursive model, as discussed earlier. There are several clear advantages in designing one's research in the above fashion. First, most studies of stress and crisis have focused on one of the elements in Hill's model, rather than testing the model as a whole. Path analysis allows us to test the adequacy of the entire model. We may find, for instance,
that for certain events the model fits quite well, while for other events it does not. The researcher would then be in a position to expand the theory by explaining why the model fits particular events and not others.

A second advantage of looking at Hill's ABCX model in terms of path analysis is that we are able to estimate the direct, indirect and spurious effects among the A, B, C, and X variables. Consequently we are able to examine, for example, what the direct and indirect effects of the event (A) are upon the crisis (X). The ability to examine such effects expands our understanding of the dynamics underlying Hill's model.

A final advantage, is that through the deletion of paths (as discussed earlier) we may be in a position to create a more parsimonious model. For example, we may find that crisis meeting resources (B) do not have a direct effect upon the crisis (X), rather they operate through perceptions as an indirect effect. In this case we may want to eliminate the path from B to X. This in turn would require revising or modifying our theory. Hence one is able to link theory and research (in Hill's ABCX model) via path analysis.

Divorce

Recently a number of theoretical models have been developed to explain the rising divorce rate in the United States (e.g. Levinger, 1979; Scanzoni, 1979). Concurrently, several models have emerged that deal with explanations of marital quality (see Spanier and Lewis, 1980). Many of these models have relied upon a social exchange perspective in guiding their theoretical formulations. For example George Levinger's (1979) "social psychological" model of marital dissolution is strongly influenced by social exchange theory. The model is based on attractions, barriers and alternative attractions to marriage. It is assumed that people stay in relationships because they are attracted to them and/or they are barred from leaving them, and that consciously or not, people compare their current relationships with alternative ones. If internal attraction and barrier forces become distinctly weaker than those from a viable alternative, the consequence is breakup (38).

This model is diagramed in Figure 10. Clearly Figure 10 lends itself to testing

Figure 10.

```
  Attractions  
/             
Barriers      
|             | 
|             |  Alternative Attractions  
|             |  Chance for Divorce  
```


via path analysis. As in our earlier example, Figure 10 is a recursive model, with attractions and barriers included as exogeneous variables. Again, the advantages of utilizing path analysis are clear. It allows us to: 1) test the applicability of the entire model; 2) examine the direct and indirect effects among variables; and 3) create a more parsimonious model.

Marital Power

As a final example of applying path analytic techniques to a substantive area, we turn to marital power. Data for this example were gathered in 1977 and 1978 with 349 married couples. (see Thomson, 1981; and Rank, 1982a, 1982b for a more detailed description). This example is included to illustrate the process of linking theory to research, and research in turn, to theory.

When dealing with panel data (as is the case here), a number of statistical techniques are available to the researcher, each with advantages and disadvantages. However, as Pendleton et al. (1979) note, path analysis provides an excellent means of analyzing data over time. In this example, the interest was in testing the applicability of a specific theoretical model. The effect of wife's and husband's income on wife's employment decision-making influence at times 1 (1977) and 2 (1978) were examined. It was hypothesized that the greater the wife's income, the greater the amount of influence she would experience over her employment decisions (based upon an economic self sufficiency argument). It was also hypothesized that the greater the husband's income, the greater the amount of influence his wife would experience over her employment decisions (based upon greater exposure to egalitarian norms). These hypotheses were supported in a cross-sectional analysis for the 1977 wave (Rank, 1982a).

The complete results of the following analysis are reported by Rank (1982b). The data were first examined to see if the assumptions of path analysis were met (which they were). Turning to Figure 11 we find that wife's and husband's incomes at time 1 positively affect the amount of influence a woman has at time 1 regarding whether she should work or not (as hypothesized). Moreover, the amount of influence a woman has at time 1 strongly affects the amount of influence she has at time 2. However, wife's and husband's incomes at time 2 (as well as time 1) fail to have a statistically significant direct effect on influence at time 2. There were several possible explanations for this. First, it could be that the high correlation between income at times 1 and 2 was washing out the effect of income at time 2 on influence. Or it was possible that wife's influence at time 1 was reducing the effect of income at time 2. Therefore, several modified path models were examined. When the path from influence 1 to influence 2 was eliminated, the direct effects of husband's and wife's incomes at time 1 on influence at time 2 became sizable, whereas the path coefficients of wife's and husband's incomes at time 2 on influence at time 2 decreased. Likewise, when the paths from wife's and husband's incomes at time 1 to wife's and husband's incomes at time 2 were eliminated, $P_{65}$ and $P_{64}$ remained slight. We therefore felt that our original model (and subsequently theory) were in need of revision.

Figure 12 reduced the original model by omitting wife's and husband's incomes at time 2. We can see that $X_1$ and $X_2$ have a positive effect on $X_3$. Wife's influence at time 1 ($X_3$) also has a significant direct effect on her influence at time 2 ($X_4$). Finally $X_1$ and $X_2$ are viewed as having both a direct and an indirect effect on $X_4$. The direct effects of $X_1$ and $X_2$ on $X_4$ are .04 and .09 respectively,
Wife's Income 1 \( X_1 \)

\[ .70^{**} (.799) \]

Wife's Influence 1 \( X_3 \)

\[ .22^{**} (.158) \]

Husband's Income 1 \( X_2 \)

\[ .75^{**} (.758) \]

Wife's Income 2 \( X_4 \)

\[ .06 (-.082) \]

Husband's Income 2 \( X_5 \)

\[ .05 (.033) \]

Wife's Influence 2 \( X_6 \)

\[ .03 (.023) \]

\[ .05 (.033) \]

\[ .70 \]

\[ .75 \]

\[ .06 \]

\[ .05 \]

\[ .85 \]

\[ .97 \]

\[ .71 \]

**significant at the .01 level
Unstandardized coefficients in parentheses
**Figure 12.**

Wife's Income 1
\(X_1\)

\(0.15^{**}(0.124)\)

Wife's Influence 1
\(X_3\)

\(0.49^{**}(0.455)\)

Wife's Influence 2
\(X_4\)

Husband's Income 1
\(X_2\)

\(0.22^{**}(0.158)\)

\(R_u\)

\(0.97\)

\(R_y\)

\(0.85\)

**significant at the .01 level**

Unstandardized coefficients in parentheses
while the indirect effects are .07 (.15 X .49) and .11 (.22 X .49), resulting in a total effect on $X_4$ of .11 ($X_1$) and .20 ($X_2$).

It was argued that in order to longitudinally understand influence within wives' employment decision making, and perhaps in order to understand marital power in general, there is a need to develop a new model or theoretical framework, which was called a "carry over" model. It was argued that figure 12 is an approximate representation of the process of influence dynamics over time. In order to predict the extent of influence a wife has regarding the decision to work, one must be aware of previous levels of influence. However, at some point in time, probably very early in a couple's marital history (or perhaps premaritally), there is no level of previous influence (as in the case of our 1977 sample).

At this stage, the extent of the wife's economic self-sufficiency, and the extent of the husband's exposure to egalitarian norms come into play. These factors positively affect the amount of influence the wife (and conversely the husband) has. The greater the wife's ability to support herself independently of her husband, the greater the availability of marital alternatives. This in turn provides her (both consciously and unconsciously) with a viable bargaining chip in increasing her influence over employment decision making. Likewise, husband's socioeconomic status at this initial stage has a positive effect on wife's influence as well. The higher the husband's SES, the greater the likelihood that he has been exposed and adheres to a more egalitarian set of beliefs regarding male/female relations, which results in his spouse accruing greater influence. It was argued that this initial stage is paramount in understanding influence at a later date. There appears to be a carry over effect of influence at time 1 on influence at time 2. There is also a carry over effect of SES (as measured by income) at time 1 on influence at time 2. This effect is both indirect (through SES affecting influence at time 1 which in turn affects influence at time 2) and direct. Surprisingly, income at time 2 was shown to have little effect on influence at time 2, and was subsequently dropped from the model.

It was argued that the reason current levels of SES are perhaps not as important as previous levels of SES in predicting influence, is that the early stage of marriage, which often represents the period when guidelines and rules are established, is fundamental in determining how future patterns of interaction occur. Thus it is the early rather than the later levels of SES which are critical in predicting influence at a later time period. This idea is a radical departure from the existing theoretical frameworks dealing with marital power. However, previous theories have focused on the cross-section rather than the changing dynamics of power. It may be that such theories are in need of modification when dealing with panel data. It was suggested that to predict the current levels of influence in wives' employment decision making, one needs to focus on the levels of influence established early in the couple's marital career, and that these early levels of influence are affected by spousal levels of SES. Obviously the longer a couple is married, the less effect this early stage will have on current levels of influence. However, it may be that even for long-term couples, there is a carry over effect from the rules and guidelines that were established during the early years of marriage, to the current patterns of behavior, interaction, and influence.

This example illustrates that through the use of path analysis we are able to trim an original framework in order to derive a more parsimonious model. This in turn, caused us to revise our original theory, thereby linking theory...
research in a dynamic process. In short, path analysis may provide the analyst with a link between theory and research.

CONCLUSION

As noted in the introduction, there is a need for a reciprocity between theory and research. Theory should guide research, and research in turn should guide theory. Yet too often we are faced with the situation where theory has little empirical support, or research undertaken lacks any theoretical foundation. It is naive to assume that any one statistical technique can remedy this situation, however, it is not unrealistic to view path analysis as a first step towards linking theory and research.

There are a number of areas within family studies where theoretical models are fairly well developed. Path analysis provides an ideal opportunity for investigating how such models fit the data. This in turn, causes us to refine or perhaps discard particular theories in lieu of competing theoretical frameworks. Indeed, our example of marital power illustrated this process. It may be to our benefit to view theory and research as distinct but permeable components in a process of understanding. Path analysis allows one to bridge these two components, hopefully resulting in a clear and insightful analysis of a specific phenomenon.
Bohrnstedt, G. W. and T. M. Carter

Duncan, O. D.

Heise, D. R.

Hill, R.
1958 "Generic features of families under stress." Social Casework 49 (February-March):139-150.

Kerlinger, F. N. and E. J. Pedhazur

Land, K. C.

Levinger, G.


Merton, R. K.

Pendleton, B. F., R. D. Warren, and H. C. Chang

Rank, M. R.


