DISCUSSION

Robert P. Althauser, Princeton University

These three rather different papers run the gamut from "theoretical" to "empirical" methodology: from Land with a different ending for Heise's [2] theoretical piece on reliability and stability to Nestel on the practical realities of a "real" panel study. For me, at least, it is the mixture of theory and simulated, if not real data, that makes the third paper by Pelz and Faith the most interesting of the three.

As most of my attention will be given to this third paper, let me begin with some brief comments on the other two papers.

Nestel's paper inadvertently points up the increasing need for a joint consideration of theoretical and practical matters in the methodology of panel studies. He suggests at one point that, in effect, variation in measurement error should decrease over time in successive remeasures (i.e., respondents would make their answers more consistent with their "true" feelings). Yet, in other, more theoretical papers by Heise [2] and Wiley and Wiley [8], we find that a constant error variance is assumed in the solution of equations which lead to estimates of reliability and stability coefficients.

Consider a second example of the same thing. One of the three additional assumptions which Land proposes as a way of finessing an identification problem which thwarts Heise late in his paper [2] is one of a Markov process among errors (the other assumptions being equal stability coefficients and a Markov process among "true" scores). Yet Charles Werts (of E.T.S.) has found that in certain simulations, this first assumption is most unsound. While realism in theoretical assumptions need not be our only criteria for their use, these two examples may suggest that we ought to think more about the empirical realism of our theoretical assumptions.

One other comment -- about Land's test for the equality of stability coefficients. It would seem at first more straightforward to statistically compare the different stability coefficients as calculated by Heise's method [2]. In contrast, Land's approach which regresses later on earlier measures <u>must</u> assume perfect measurement. If that assumption is, in fact, violated, his test would be quite misleading. It confounds less than perfect epistemic path coefficients between measures and true scores with path coefficients of relationships between true scores.

Yet, in a way, Heise is in the same boat, similarly making other assumptions (e.g., constant reliability of successive measures) to permit an estimate of (possibly differing) stability coefficients. The trick, as Wiley and Wiley [8] illustrate, is to make the weakest possible assumptions. As between equal reliability coefficients with imperfect measures (Heise) and perfect measures (Land), the former appears to be the weaker, hence preferable assumption.

Pelz and Faith

Whatever the ultimate practicality of this line of continuing investigation into the inference structure of cross-lagged correlations by Pelz and his associates (and there are some critics: see Heise [3]; Rozelle and Campbell [7] and Duncan [1]), it has certainly resulted in some intriguing work. Though what follows is primarily a response to their paper above, scattered references will be made to Pelz and Lew [6] and their first Interim Report [5]. (Note that their paper above constitutes the second Interim Report of their "Causal Analysis Project.")

There are two fundamental question for a critic of this line of work to consider. Both require us to keep in mind the simple distinction between the problem of <u>recovering</u> a causal model, or at least an original $X - Y$ relationship built into simulated data, and the problem of making similar inferences from "real" data. About the second question -- whether the work so far which has been devoted to recovering already known relationships will have significant bearing on "real" inference -- more later. The prior question is how feasible does "recovery" of known relationships now appear to be, on balance?

I see three serious obstacles to the general applicability of this approach to recovery:

1) When Pelz and Lew [6] considered the effects of individual differences (which result in "long-term stability" -- autocorrelations persisting over time), they found that cross-correlations will lead to incorrect inferences about the direction of the causal relationship between X and Y. The question that remains is how large the relative variance of individual differences tends to be, for various types of data.

2) A second difficulty noted even in Pelz and Andrews [4] arises when the interval between measures (k) does not coincide with the causal interval (g). Working with X's and Y's with high to modest stability (.70 to .90), it is apparent in their paper above that when the difference between k and g is great, one may not be able to correctly recover the direction of causation. Thus, for p_{xx} and p_{yy} both equal to .70, we see that the cross-lagged differential may disappear

at $k \pm 15$. Practically speaking (and depending on the type of data), the problem does not look too serious so far.

But suppose that the stability coefficients were even more modest, approaching .5 or .4. What would the correlograms' curves look like then? Let us assume the same low value for the causal relationship $p_{yx} = .10$ as in the figures above.

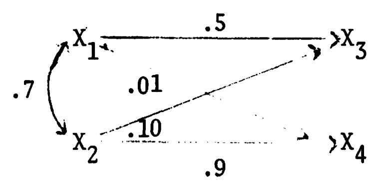
It is evident from equations 13 a - 13 d' in their technical appendix (which may not be reproduced above) that, first, the curves will retain the bell-shaped symmetry that the curves for $p_{xx} = p_{yy} = .70$ have. The cross-correlations will not persist over time, in other words. Second, the size of the maximum cross-correlations will drop, as they did previously when $p_{xx} = p_{yy}$ when down from .95 to .70.

But most importantly, the rate of increase in the cross-lagged correlations up to the interval g and the rate of decrease soon thereafter will greatly increase: the bell-shaped curve will now approach the shape of a needle.

Accordingly, it becomes much more crucial that our measurement interval k closely coincide with the true causal interval; otherwise, we cannot recover the (known) direction of causation from the cross-lagged correlations.

3) The third problem arises when we carry one step in the paper above a little further. The authors note that "with high stability... even if curves were moved leftward so that g = 0, the asymmetry would persist, and so would the cross-lagged differential." Consequently, "...the causal influence of X on Y remained apparent even when the causation was almost simultaneous." If the curves were moved still further leftward, so that g was less than zero, might not the cross-lagged differential still persist, leading us to the wrong causal inference?

The second question remains -- what sorts of connections are there going to be between recovery of a simulated X - Y relationship and causal inference from real data? That the authors are aware of this distinction is clear in their first Interim Report [5]. An example of the importance of the distinction has developed in a post-convention letter to me from Pelz, who wished to respond to my original discussion on this point. I borrowed Heise's argument [3] against cross-lagged correlations, using the following hypothetical example, with $r_{14} = p_{41} + p_{42}r_{12} = .01 + (.7)(.9) = .64$, and



$r_{23} = p_{32} + p_{31}r_{12} = .1 + (.7)(.5) = .45$. The effect of X_2 on X_3 is larger than that of X_1 on X_4 (.10 vs. .01), but the cross-lagged correlations would suggest the opposite. The result follows from the rather different stability coefficients shown (.5 vs. .9).

Pelz's response to this argument is quite interesting. He reports that the degree of stability in either X or Y does not reverse the cross-lagged differential in his simulations. To lead the way into his next point, further, he makes a trivial alteration in my example, assuming one-way causation by completely eliminating the relation between X_1 and X_4. This is necessary according to Faith for my example to correspond to a 'stable' time series.

Pelz then proposes a not-so-trivial alteration in the correlation between X_1 and X_2. This is required, following equation 14 in their technical appendix, so that stationary assumptions can be made (i.e., all autocorrelations and cross-correlations are independent of time, and the total correlations of the X's with themselves are unity). Adopting their equation 14 to my example, we have

$$r_{12} = \frac{p_{32}\ p_{42}}{1 - p_{31}\ p_{42}} = \frac{(.1)(.9)}{1 - (.5)(.9)} = .164$$

Thus, to have a stationary time series, given the path coefficients arbitrarily chosen in my example above, r_{12} must equal .164. If it does, the result is that "the cross-lagged differential is now consistent with the difference in causal coefficients." And so it is.

It would appear, then, that the nature of a stable time series precludes arbitrary examples like mine. Does their response answer Heise's [3] objections? If there was a stable time series hidden in some "real" data, his criticism would not apply. Cross-lagged differentials would be useful tools of causal inference.

But in real data, how do we know if there is a stable time series or not? Unless we know from past experience that certain types of data tend to embody stable time series, we are still faced with the problem Heise has raised. The estimates of paths or the interpretation of correlations still depend on the models postulated, as Duncan has vividly shown [1]. The problem is one of model testing, and cross-lagged differentials do not appear to be as trustworthy a tool of causal inference with "real" (as opposed to simulated) data, as is, for example, the use of instrumental variables.

In short, when Pelz and his associates later attack the basic problem of causal inference with real data, I suspect that their previous work may not be quite as useful as hoped. Yet there are enough surprises of methodological interest in their present work to give any skeptic pause!

REFERENCES

[1] Duncan, Otis Dudley, "Some Linear Models for Two-Wave, Two-Variable Panel Analysis," Psychological Bulletin, 72 (September 1969), 177-182.

[2] Heise, David, "Separating Reliability and Stability in Test-retest Correlation," American Sociological Review, 34 (February 1969), 93-101.

[3] Heise, David,"Causal Inference from Panel Data," in E.F. Borgatta and G.W. Bohrnstedt (eds), Sociological Methodology 1970, San Francisco: Jossey-Bass, 1970.

[4] Pelz, D.C., and Andrews, F.M., "Detecting Causal Priorities in Panel Study Data," American Sociological Review, 29 (December 1964), 836-848.

[5] Pelz, D.C., Magliveras, S., and Lew, R.A., "Correlational Properties of Simulated Panel Data with Causal Connections between Two Variances," Interim Report No. 1, Causal Analysis Project, Ann Arbor: Survey Research Center, University of Michigan, 1968.

[6] Pelz, D.C. and Lew, R.A.,"Heise's Causal Model Applied," in E.F. Borgatta and G.W. Bohrnstedt (eds), Sociological Methodology 1970, San Francisco: Jossey-Bass, 1970.

[7] Rozelle, Richard M. and Campbell, Donald T., "More Plausible Rival Hypotheses in the Cross-Lagged Panel Correlation Technique," Psychological Bulletin, 71 (January, 1969), 74-80.

[8] Wiley, David E. and Wiley, James A., "The Estimation of Measurement Error in Panel Data," American Sociological Review, 35 (February 1970), 112-117.