Longitudinal Reasons to **AVOID** Structural Equation Models

David Rogosa
Stanford University

rag@leland.stanford.edu

School of Education
Stanford CA 94305

Quantitative Social Science Seminar Series
U.C. Berkeley, March 14, 1994

1. WHERE DO DATA COME FROM?

Straight-line Growth Curve Formulation.

The within-unit model is a straight-line growth curve, attribute η, which exhibits systematic change over time. For individual p, growth curve in η is $\eta_p(t)$. A straight-line growth-curve is written as

$$\eta_p(t) = \eta_p(0) + \theta_p t$$

Note: Rewrite using the centering parameter τ . θ and η(τ) are uncorrelated over the population of individuals $\tau = -\sigma_{\eta(0)\theta} / \sigma_\theta^2$

$$\eta_p(t) = \eta_p(\tau) + \theta_p (t - \tau)$$

Main parameter of interest constant rate of change θ_p -- first two moments μ_θ σ_θ^2 . For systematic individual differences in growth (i.e. correlates of change) exogenous characteristic Z.

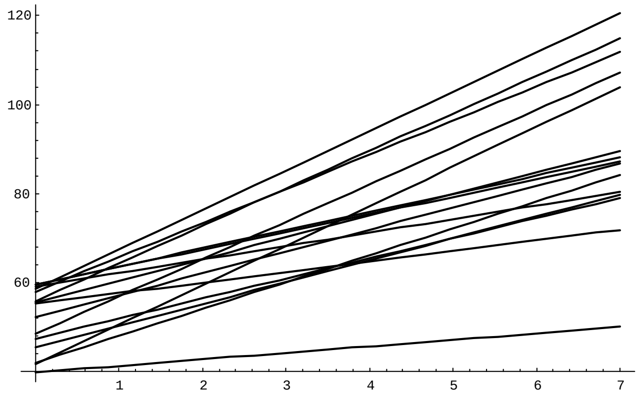
Conditional expectation E(θ|Z),

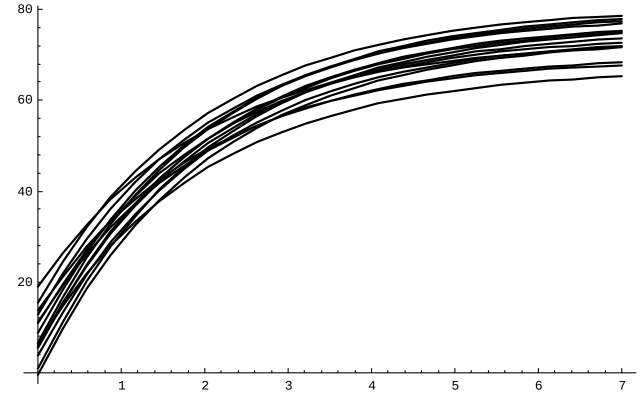
$$E(\theta \mid Z) = \mu_\theta + \gamma(Z - \mu_Z)$$

Where there is no measured exogenous variable, this between-unit model is $E(\theta|Z) = \mu_\theta$.

Observables. oversimplified version-- observable Y is an imperfectly measured η, relation between Y and η is through the simple classical test theory model. Times of observation are $t_1, \ldots, t_T$, observables for individual p are written as $Y_{p1}, \ldots, Y_{pT}$. $Y_p(t_i) = \eta_p(t_i) + \epsilon_i$

15 straight-line growth curves corresponding to population parameters $\tau = 2$; $\sigma_\theta^2 = 5.333$; $\sigma_{\eta(\tau)}^2 = 48$; $\theta \sim U[1, 9]$, $\eta(\tau) \sim U[38, 62]$. correlations among $\eta(t_i)$ for observation times $\rho_{\eta(1)\eta(4)} = .614$, $\rho_{\eta(1)\eta(6)} = .316$, $\rho_{\eta(4)\eta(6)} = .943$. For *Y,* var(ϵ) = 5, the population correlations are $\rho_{Y(1)Y(4)} = .567$, $\rho_{Y(1)Y(6)} = .297$, $\rho_{Y(4)Y(6)} = .894$.





Alternative: exponential growth to an asymptote
Exponential growth curve with asymptote λ_p and curvature γ

$$\eta_p(t) = \lambda_p - (\lambda_p - \eta_p(0))\exp(-\gamma t) \quad .$$

population parameters Figure 2 for 15 exponential growth curves

$$\tau = 2 ; \gamma = .5 ; \mu_\lambda = 75 ; \sigma^2_\lambda = 16 ; \mu_{\eta(\tau)} = 50 ; \sigma^2_{\eta(\tau)} = 9 .$$

correlations $\rho_{\eta(1)\eta(3)} = .657$, $\rho_{\eta(1)\eta(5)} = .435$, $\rho_{\eta(3)\eta(5)} = .965$.
exogenous variable *Z* could be linked with both λ_p and $\eta_p(\tau)$.

2. WHAT DO WE WANT TO KNOW?

Examples of longitudinal research questions

1. *Individual and Group Growth.* description of the form and amount of change. estimation of the individual (or group) growth curve, the heterogeneity (individual differences) in the individual growth curves, and the statistical and psychometric properties of these estimates. Parameters: f(θ; t), μ_θ σ^2_θ $\rho(\hat{\theta})$

2. *Correlates and Predictors of Change.* systematic individual differences in growth e.g., "What kind of persons learn (grow) fastest?". Parameters: $\rho_{\theta Z}$ $\beta_{\theta Z}$

3. *Stability over Time.* consistency over time of an individual and of consistency of individual differences over time.
Parameters: γ = Pr(two growth curves do not intersect)

4. *Comparing Experimental Groups.* standard, well-developed methods. More generally, individual differences in response to treatment.

5. *Comparing Nonexperimental Groups.* central topic in the methodology for the evaluation of social programs. Commonly employed statistical adjustment methods for pre-post data, often based on analysis of covariance, fail.

6. *Analysis of Reciprocal Effects.* empirical research has attempted to answer the oversimplified question, Does *X* cause *Y* or does *Y* cause *X*? from meager longitudinal data by casually comparing a couple of correlations (or structural regression coefficients) e.g., cross-lagged correlation approaches.

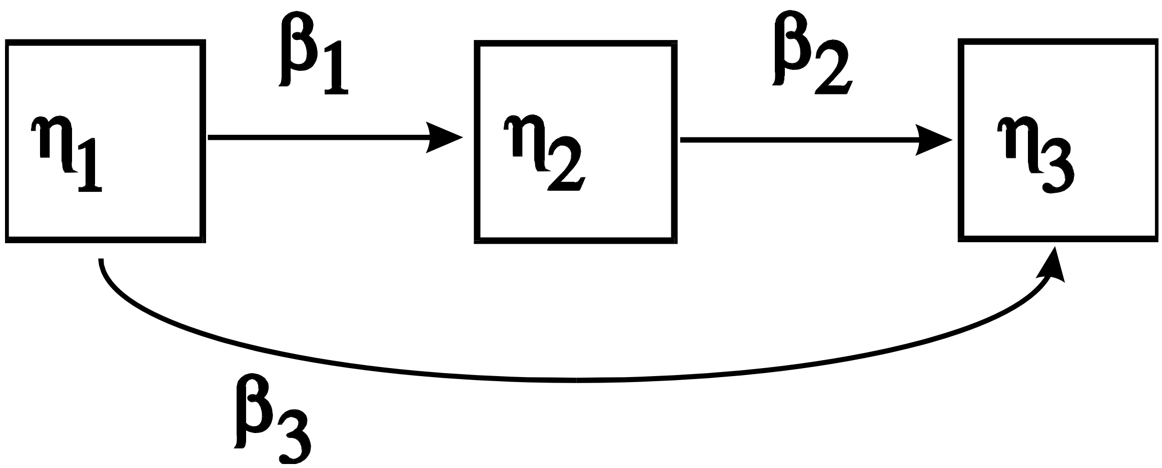
7. *Growth in Multiple Measures.* Natural questions include relative strengths and weaknesses in individual and group growth, or associations of rates of growth across attributes.

3. WHAT SHOULD WE DO?

A. Growth Curve Analyses for Longitudinal Panel Data
B. AABL (almost anything but LISREL)

4. CAUSAL INFLUENCES ON CHANGE

Three-waves, single variable.
For true scores structural regression model: $\eta_2 = \alpha_2 + \beta_1\eta_1 + e_2$; $\eta_3 = \alpha_3 + \beta_2\eta_2 + \beta_3\eta_1 + e_3$. Example: Goldstein (1979a, 1979b) for



reading test scores obtained for a nationwide British sample with measurements at ages 7, 11, and 16; estimates for the $\{\beta_k\}$: {.841, 1.11, −.147}. The negative estimate for β_3 causes considerable discomfort, summarized by Goldstein (1979a, p. 139): "This is difficult to interpret and may indicate that non-linear or interaction terms should be included in the model, or perhaps that the change in score between seven and 11 years is more important than the seven-year score itself. However, the addition of non-linear terms does not change this picture to any extent."

Main result:

$\beta_3 = (t_2 - t_3)/(t_2 - t_1) < 0$ and $\beta_2 = (t_3 - t_1)/(t_2 - t_1) > 0$.

Data example 40 cases, each observed at three time points {1, 3, 5} . true observations fall on a straight-line growth curve
Regression for $\eta(t_3)$ matches the theoretical results $\beta_3 = (3 - 5)/(3 - 1)$ and $\beta_2 = (5 - 1)/(3 - 1)$-- with squared multiple correlation of 1.0.

regression equation:

$$\eta(5) = -0.000003 - 1.00\ \eta(1) + 2.00\ \eta(3)$$

Predictor	Coef	Stdev	t-ratio	p
Constant	-0.00000309	0.00000000	*	*
η(1)	-1.00000	0.00000	*	*
η(3)	2.00000	0.00000	*	*

s = 0   R-sq = 100.0%   R-sq(adj) = 100.0%

Data Description.

	MEAN	MEDIAN	STDEV
η(1)	44.16	43.90	7.24
η(3)	54.21	53.63	7.24
η(5)	64.27	63.21	9.24
W	14.99	15.20	2.803

Correlations

	η(1)	η(3)	η(5)
η(3)	0.842		
η(5)	0.536	0.907	
W	0.766	0.765	0.598

Scatterplots between the η-values.

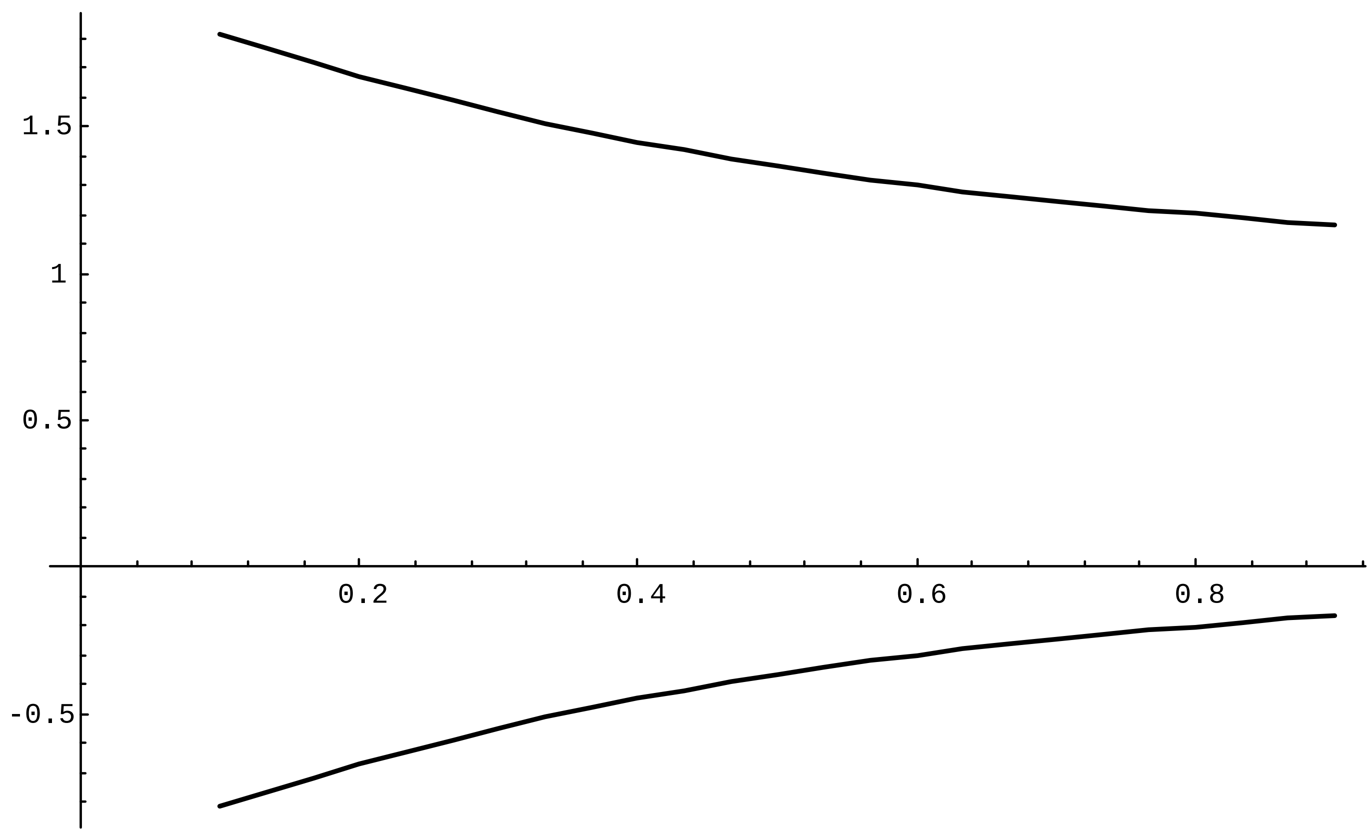
```
                               time 5 vs time 3
   -
   -
   -                                                    *
   -                                      2   *
 75+                                        *     2  *
   -                                    **   * **
   -
   -                           *       2        *
   -                *     ***     **     *
 60+                  * *2    *
   -             *     * *   2      *
   -                 *
   -            2
   -
 45+            *
   -     *
   -
     --------+---------+---------+---------+---------+-----
          42.0      48.0      54.0      60.0      66.0
                         time 5 vs time 1
   -
   -
   -                                            *
   -                        **    *
 75+                                  *         **    *
   -                            * *      *     **
   -
   -                       *          *   *               *
   -      *          ***           *  *              *
 60+              *   *2         *
   -         *           *  *     **           *
   -                   *
   -              * *
   -
 45+                       *
   -                  *
   -
    ----+---------+---------+---------+---------+---------+-
      30.0      36.0      42.0      48.0      54.0      60.0
                         time 3 vs time 1
   -                                            *
   -                                                  *
   -                                            **
   -                              *      *     **         *
 60+                        **        *              *
   -                            * *   *   *
   -                                  *        *
   -                                 *
   -                   *   *     ***
 50+                 *2     *
   -              *    2 *
   -      *            *
   -         *    * *      *
   -
 40+                  *
   -
   -
     ----+---------+---------+---------+---------+---------+
      30.0      36.0      42.0      48.0      54.0      60.0
```

Fallible measures give more complex relations. observed scores by adding measurement error variance 10; resulting reliabilities at times {1, 3, 5} are {.84, .84, .90}. Path analysis regression $Y(5) = 5.054 - .1212\ Y(1) + 1.19\ Y(3)$, with squared multiple correlation .552.

Exponential Growth The general result shows dependence only on the curvature paramater and the times of observation.
For $t_1 = 1$, $t_2 = 3$, $t_3 = 5$ structural parameters are $\beta_3 = -\exp[-2\gamma]$ and

$$\beta_3 = \frac{\exp[-\gamma t_3] - \exp[-\gamma t_2]}{\exp[-\gamma t_1] - \exp[-\gamma t_2]}$$

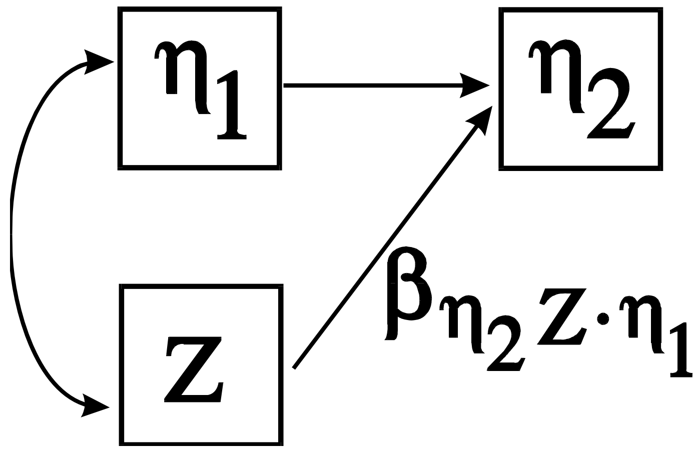
$$\beta_2 = \frac{\exp[-\gamma t_3] - \exp[-\gamma t_1]}{\exp[-\gamma t_2] - \exp[-\gamma t_1]}$$



$\beta_2 = 1 + \exp[-2\gamma]$. For the value of $\gamma = .5$ used in Figure $\beta_3 = -.3679$ and $\beta_2 = 1.3679$. Coefficients as a function of γ shown below.

5. EXOGENOUS CAUSAL INFLUENCES

DATA EXAMPLE Consequences of basing an analysis of the standard structural model shown in two waves with an exogenous variable.



In the population from which the example data are drawn there is no association between the background variable W and individual rate of change θ; $\rho_{Z\theta} = 0$. Structural/causal regression coefficients may be large positive or large negative even when $\rho_{Z\theta} = 0$. When η(3) is used as the initial value the structural coefficient for the influence of W on change is significant with a negative value and when η(1) is used as the initial value the structural coefficient for the influence of W on change is significant with a positive value.

1. η(3) as the initial value

The regression equation is η(5) = 0.68 - 0.757 W + 1.38 η(3)

Predictor	Coef	Stdev	t-ratio	p
Constant	0.683	4.555	0.15	0.882
W	-0.7570	0.3329	-2.27	0.029
η(3)	1.3822	0.1290	10.72	0.000

s = 3.752 R-sq = 84.4% R-sq(adj) = 83.5%

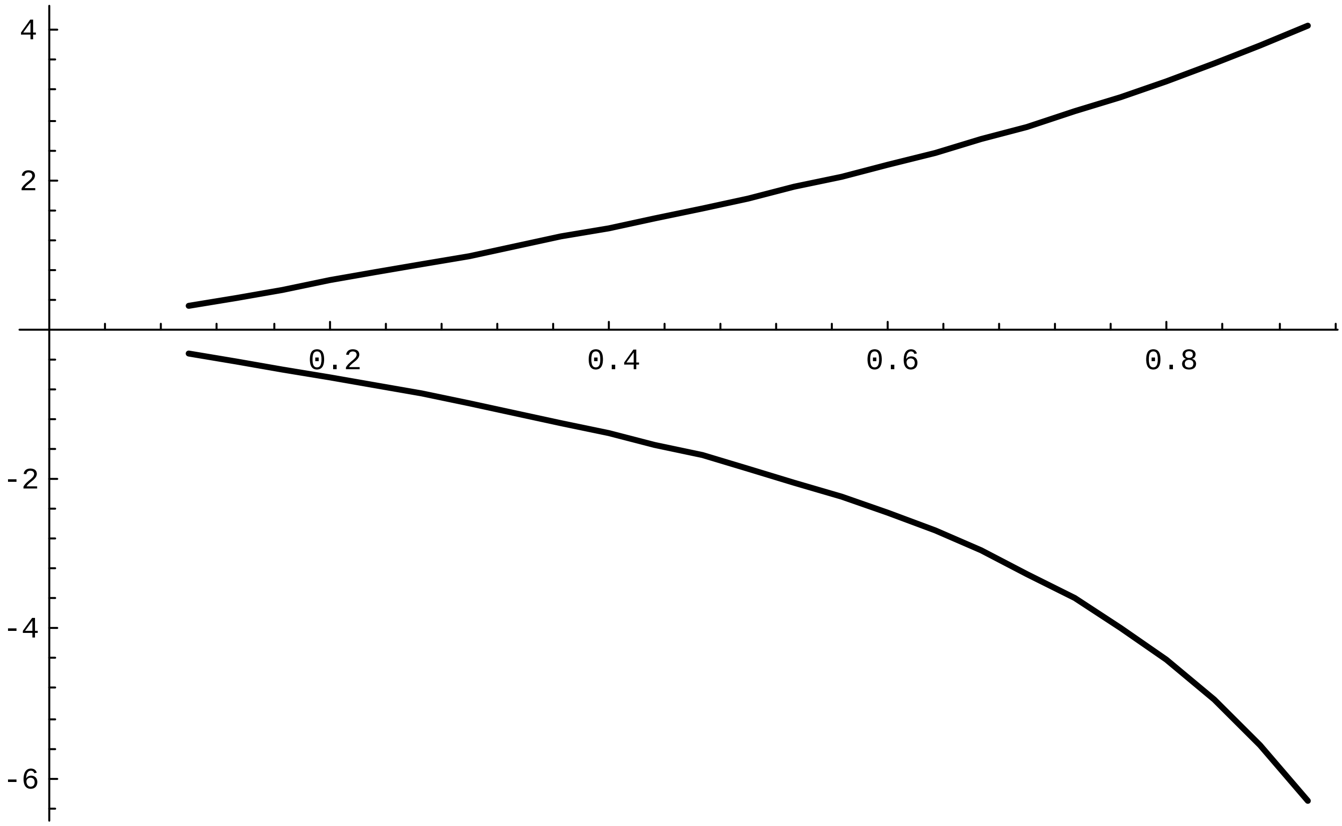
2. η(1) as the intial value

The regression equation is η(5) = 31.2 + 1.50 W + 0.239 η(1)

Predictor	Coef	Stdev	t-ratio	p
Constant	31.213	7.546	4.14	0.000
W	1.5004	0.6678	2.25	0.031
η(1)	0.2392	0.2587	0.92	0.361

Analytic results

Plot the value of structural parameter as a function of $\rho_{Z\eta(\tau)}$ for two choices of t_1 , 0 and 6, with $t_2 = t_1 + 5$. In figure are shown values of $\beta_{\eta(t2)Z\bullet\eta(t1)}$ with $\beta_{\eta(5)Z\bullet\eta(0)}$ having positive values and $\beta_{\eta(11)Z\bullet\eta(6)}$ having negative values.



structural parameter as a function of $\rho_{Z\eta(\tau)}$ for two choices of t_1 , 0 and 6, with $t_2 = t_1 + 5$.

6. SIMPLEX GROWTH MODELS

The claim
Joreskog (1979) states "For one measure administered repeatedly to the same group of people, an appropriate model is a simplex model.
Werts, Linn, and Joreskog (1977, p.745) assert "The simplex model appears to be particularly appropriate for studies of academic growth."

Autoregressive Lag-1 model

$$\eta_{i+1,p} = \beta_i \eta_{ip} + \delta_{i+1,p}$$

Partial Correlation Property (Guttman)
$\mathrm{Corr}(\eta_i, \eta_k \bullet \eta_j) = 0$.

Artificial Data Example (5x5) Rogosa and Willett (1985)

	Y_1	Y_2	Y_3	Y_4	Y_5
Y_1	1				
Y_2	.75	1			
Y_3	.73	.74	1		
Y_4	.69	.72	.74	1	
Y_5	.66	.69	.73	.75	1

Correlation matrix looks like a simplex. Furthermore, a LISREL analysis based on the standard Quasi-simplex structure produced a marvelous fit by any of the standard indices: overall $\chi^2 = 2.77$; near perfect reproduction of covariance and correlation matrices.
PUNCH LINE the example covariance matrix was generated from a growth curve structure that maximally violates the assumptions of the simplex model. Straight-line growth curves can be thought of as maximally "unsimplex"; one common characterization of the simplex model is that $\rho_{\eta(t3)\eta(t1)\bullet\eta(t2)} = 0$, whereas for straight-line growth $\rho_{\eta(t3)\eta(t1)\bullet\eta(t2)} = -1$.

$\tau = 3$; $\sigma^2_\theta = .0078$; $\sigma^2_{\eta(\tau)} = .4374$; $\mathrm{var}(\epsilon) = .15$. were used for

this example. Exponential growth curves, for which also $\rho_{\eta(t3)\eta(t1)\bullet\eta(t2)} = -1$, can be used to generate a similar example.

WHY SHOULD WE CARE? Even when the simplex model fits wonderfully, the results of the structural equations can badly mislead. The covariance structure analyses usually go on to compute *growth statistics* and reliability estimates based on the fitted simplex model (Werts & Hilton, 1977; Werts et al. 1977). For example, the variance of true change over a time interval of one unit, $\mathrm{var}(\eta(t + 1) - \eta(t))$, is .0078 for all t. The LISREL analysis yields estimates nearly five times larger than the actual value; the LISREL estimates are {.038, .033, .033, .038} for $t = \{1, 2, 3, 4\}$. Similar discrepancies are found for estimates of the reliability of observed change--values estimated from the LISREL analysis differ markedly from the actual values.

7. ASSESSMENTS OF STABILITY.

Rogosa, Willett, and Floden (1984) organized questions about temporal stability into two broad headings--

Is an individual consistent over time? and

Are individual differences consistent over time?

Procedures for adressing both questions rely on modeling and analysis of the individual trajectories. Questions about consistency of individual differences have dominated attention in behavioral science. And a number of procedures based on stuctural equation models have been proposed and applied to the assessment of stability (e.g., Wheaton, Muthen, Alwin, & Summers, 1977; Huesman, Eron, Lefkowitz & Walder, 1984). Rogosa (1988, under Myth 8) gives examples of some of the wayward properties of structural parameters used to assess stability: e.g. $\beta_{\eta(t2)\eta(t1)}$ or $\beta_{\eta(t2)\eta(t1)\bullet Z}$ --the latter parameter for the path between $\eta(t_1)$ and $\eta(t_2)$ in the exogenous picture has gotten the most attention in expositions of structural equations.

A dependable measure for assessing consistency of individual differences over a specified time interval is the index of tracking γ proposed by Foulkes and Davis (1981) ; this index estimates the probability that two randomly chosen individuals trajectories do not cross in the time interval specified. Tables 5-13 through 5-15 in Rogosa (1988) contrast the index of tracking and the indices based on structural equation models; for example, two time intervals both having large tracking indices of .88 have values of the structural parameter $\beta_{\eta(t2)\eta(t1)\bullet Z}$ equal to 1.71 and .13. Stability large or not?

8. WHAT ABOUT INTERVENTIONS?

Rogosa (1991) : Longitudinal interventions, ATI research

Holland (1988): Encouragement Design

Encouragement Designs: Effects of Interventions

Exemplar Study: Random assignment of students to treatment-control conditions for intervention on improving study habits. *Measures:* Treatment/control assignment (G), amount of study (R), and outcome measure, achievement test score (Y).

Questions: **1.** Increase in study time from intervention? **2.** Increase in achievement from studying an hour longer (dose response)? **3.** Increase in achievement if no increase in study (placebo effect)? **4.** Total impact on achievement?

Counterfactual Data Formulation for Individual u. **1.** $R_t(u) - R_c(u) = \rho(u)$, treatment/control difference in amount of study. **2.** $Y_{Gr}(u) - Y_{Gr'}(u) = \beta(u)*(r - r')$, increment to outcome from study amount r′ vs r. **3.** $Y_{tr}(u) - Y_{cr}(u) = \tau(u)$, treatment/control difference in outcome with same amount of study r.

4. $Y_{tR}(u) - Y_{cR}(u) = \tau(u) + \rho(u)\beta(u)$, overall treatment/control difference.

Individual Level Model. $R_G(u) = R_c(u) + \rho(u)G$; $Y_{Gr}(u) = Y_{c0}(u) + \tau(u)G + \beta(u)r$.

> ***Path Analysis Regressions***
> *Path Coefficients:* $\gamma_1, \gamma_2, \gamma_3$
> $R = \alpha_R + \gamma_1 G + \epsilon_R$

> **ALICE** specification: $\rho(u) = \rho; \beta(u) = \beta; \tau(u) = \tau$.
>
> **Path Analysis Results Under ALICE.** $\gamma_1 = \rho$; $\gamma_2 = \beta + \delta$; $\gamma_3 = \tau - \rho\delta$.

Alwin, D. F. (1988). Structural equation models in research on human development and aging. In K. W. Schaie, R. T. Campbell, W. M. Meredith, & S. M. Rawlings (Eds.), *Methodological issues in aging research* (pp. 71-170). New York: Springer Publishing Co.

Goldstein, H. (1979a). *The design and analysis of longitudinal studies*. London: Academic Press.

Goldstein, H. (1979b). Some models for analysing longitudinal data on educational attainment. *Journal of the Royal Statistical Society, A, 142,* 407-442.

Holland, P. W. (1988). Causal inference, path analysis and recursive structural equation models. In C. Clogg (Ed.), *Sociological Methodology 1988* (pp. 449-484). Washington, D.C.: ASA

Joreskog, K. & Sorbom, D. (1979). *Advances in Factor Analysis and Structural Equations Models*. Cambridge MA: ABT

Rogosa, D. R. (1987). Casual models do not support scientific conclusions: A comment in support of Freedman. *Journal of Educational Statistics, 12,* 185-195.

Rogosa, D. R. (1988). Myths about longitudinal research. In K. W. Schaie, R. T. Campbell, W. M. Meredith, & S. M. Rawlings (Eds.), *Methodological issues in aging research* (pp. 171-209). New York: Springer Publishing Co.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92,* 726-748.

Rogosa, D. R., Floden, R. E., & Willett, J. B. (1984). Assessing the stability of teacher behavior. *Journal of Educational Psychology, 76,* 1000-1027.

Rogosa, D. R. & Willett, J. B. (1985a). Understanding correlates of change by modeling individual differences in growth. *Psychometrika*, *50*, 203-228.

Rogosa, D. R. & Willett, J. B. (1985b). Satisfying a simplex structure is simpler than it should be. *Journal of Educational Statistics*, *10,* 99-107.

Rogosa, D. R., & Willett, J. B. (1983). Comparing two indices of tracking. *Biometrics, 39,* 795-6.

BALLAD OF THE CASUAL MODELER

Words and Music by David Rogosa 1988

When I was a student
in seventy-three
I heard of new ways
to do psychology

If you had you some data
and you knew a little math
you didn't need that much
thinkin'
you just draw in the path
CH
Now my model is busted
I can't make it fit
I drew in more arrows
but it still don't mean shit

Then they did it one better
way beyond you and me
with structures for variables
you never can see

Then came computer
programs
seminars too
by these fellows from
Sweeden
to teach me and you

And professors were travellin'
from miles around
just to see the sight of that
chi-square go down
CH
Now my model is busted
I can't make it fit
I drew in more arrows
but it still don't mean shit

Now you might come to
wonder
after all of this fuss
where is the science
the knowledge you trust

And I've been a askin'
anyone I see
why I should take this stuff
seriously
CH
Now my model is busted
I can't make it fit
I drew in more arrows
but it still don't mean shit

I read all their papers
and I looked for the facts
All I found's a lotta claims
they just oughta retract

For they talk about models
in a casual way
If you think that means causal
then I'm Doris Day
CH
Now my model is busted
I can't make it fit
I drew in more arrows
but it still don't mean shit