Overview and Implementation for Basic Longitudinal Data Analysis

David Rogosa
Stanford University

rag@leland.stanford.edu
http://www-leland.stanford.edu/~rag/
presentation materials at
http://tsg.stanford.edu/~rag/cresst/

C·R·E·S·S·T

PROJECT
3.3
PROGRESS

CRESST TFQ, September 6, 1997

MYTHS about longitudinal research

Rogosa, D. R.
Myths about longitudinal research. Presented at the Stanford-Berkeley Colloquium on Quantitative Methods in Behavioral Science, U.C. Berkeley, May 1983.

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

Rogosa, D. R. (1995). Myths and methods: "Myths about longitudinal research," plus supplemental questions. In The analysis of change, J. M. Gottman, Ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 3-65.

1. Two Observations a longitudinal study make.
2. The difference score is intrinsically unreliable and unfair
3. You can determine from the correlation matrix for the longitudinal data whether or not you are measuring the same thing over time
4. The correlation between change and initial status is:
(a) *negative;* (b) *zero;* (c) *positive;* (d) *all of the above.*
5. You can't avoid regression toward the mean
6. Residual change cures what ails the difference score
7. Analyses of covariance matrices inform about change
7.1 Path analysis informs about change
7.2 Structural regression models inform about change
7.3 Simplex models describe most longitudinal data
8. Stability coefficients estimate:
(a) *the consistency over time of an individual;*
(b) *the consistency over time of an average individual;*
(c) *the consistency over time of individual differences;*
(d) *none of the above;* (e) *some of the above.*
9. Casual analyses support causal inferences about reciprocal effects

OLD BUSINESS

Conditional versus Unconditional Analyses (Goldstein, Plewis...)

[*UK Reading example*]

Longitudinal Data Examples

Individual Change Analyses

output sheets and data listings for some of the following examples:

Textbook Examples

Dental From: lme and nlme: Mixed-effects Methods and Classes for S and S-plus Jose C. Pinheiro, Douglas M. Bates. From an orthodontic study presented in Potthoff and Roy (1964). Four measurements of the distance (in millimeters) from the center of the pituitary to the pteryomaxillary fissure made at ages 8, 10, 12, and 14 years on 16 boys and 11 girls (gender used as exogenous W).

Ramus 4 longitudinal observations on each of 20 cases. The measurement is the height of the mandibular ramus bone (in mm) for boys each measured at 8, 8.5, 9, 9.5 years of age. used by a number of authors, can be found in Table 4.1 of Goldstein (1979)

Rat Rat weight data, from 1989 HLM manual. The rat data consist of 10 individuals, with weight measurements (*Y*) at 5 occasions (weeks 0,1,2,3,4) and a background measure (*W*), the mother's weight

Education Data Examples

WISC 4 observations, Wechsler Intelligence Scale for Children,Performance Scale, 86 children (times: begin first, end first, third, fifth grades). Gender W (Osborne).

NC Fem North Carolina Achievement Data (see Williamson, Applebaum, Epanchin, 1991). These education data are eight yearly observations on achievement test scores in math (Y), for 277 females each followed from grade 1 to grade 8, with a verbal ability background measure (W)

Canonical Artificial Examples (known structure)

Structure and Parameter Values for Artificial Data on p.6

ArtN200 used in Rogosa Saner 1995, both full data and Y-missing versions: 5 waves on each of 200 individuals, with observation times {0,1,2,3,4}, and with an exogenous variable *W* for each individual.

Smearmiss. Artificial longitudinal data with known structure. Five observations (about 16% missing) on each of 100 individuals, with times of observation varying over individuals, and with an exogenous measure *W*

Data Structures.

First five examples (Textbook, Education) have the simplest structure, no missing data, and "synchronous"--i.e., all individual measures at same times. In practice, data are missing; different observation times across individuals. Timepath97 estimation procedures for the general case.

For looking at Timepath97 and other estimation approaches, a simple 2x2 Table is useful:

Time Observations \ Outcome Observations	FULL	MISSING
SYNCHRONOUS	Rat, Ramus, WISC, Dental, NCFem	Rogosa-Saner Table 2
NOT SYNCHRONOUS	Rogosa-Saner Reply (smear)	Smearmiss

Estimation: Full-Synchronous– closed-form mle; lisrel; OLS exact
Missing-NonSynchronous– mixed model (SAS PROC MIXED, S-plus lme, HLM); OLS approximate.

Models for Collections of Growth Curves

Straight-line Growth Curve Formulation.

attribute η, which exhibits systematic change over time. For individual p, growth curve in η is $\eta_p(t)$.

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

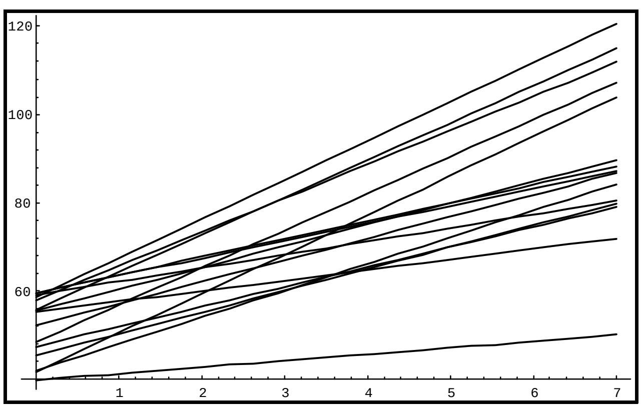
Note: Rewrite using the centering parameter t^o ; θ and $\eta(t^o)$ are uncorrelated over the population of individuals $t^o = -\sigma_{\eta(0)\theta}/\sigma^2_\theta$

$$\eta_p(t) = \eta_p(t^o) + \theta_p(t - t^o) \ .$$

Constant rate of change θ_p -- first two moments μ_θ σ^2_θ

Shown below 15 straight-line growth curves corresponding to pop. parameters $t^o = 2$; $\sigma^2_\theta = 5.333$; $\sigma^2_{\eta(t^o)} = 48$; $\theta \sim U[1, 9]$, $\eta(t^o) \sim U[38, 62]$. correlations among $\eta(t_i)$ for observation times $\rho_{\eta(1)\eta(4)} = .614$, $\rho_{\eta(1)\eta(6)} = .316$, $\rho_{\eta(4)\eta(6)} = .943$.

Straight-line Growth



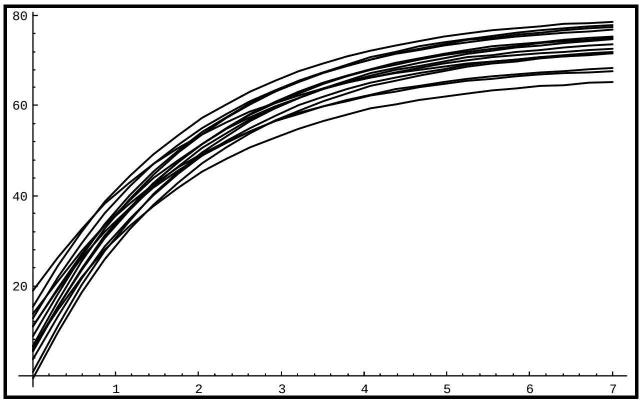
Observables. oversimplified version-- observable Y is an imperfectly measured η, relation between Y and η is simple classical test theory model: $Y_p(t_i) = \eta_p(t_i) + \epsilon_i$ For *Y*, with $var(\epsilon) = 5$, the pop. correlations are $\rho_{Y(1)Y(4)} = .567$, $\rho_{Y(1)Y(6)} = .297$, $\rho_{Y(4)Y(6)} = .894$.

For systematic individual differences in growth (i.e. correlates of change) exogenous characteristic W.

Conditional expectation $E(\theta|W) = \mu_\theta + \beta_{\theta W} (W - \mu_W)$. With no measured exogenous variable, this between-unit model is $E(\theta|W) = \mu_\theta$.

Alternative: exponential growth to an asymptote

Exponential growth curve with asymptote λ_p and curvature δ



$$\eta_p(t) = \lambda_p - (\lambda_p - \eta_p(0))\exp(-\delta t)$$

Figure shows a collection of 15 exponential growth curves

exogenous variable *W* could be linked with both λ_p and $\eta_p(t^o)$.

Technical specifications for Artificial Data Structure.

The unobserved $\eta_p(t_i)$ follow the straight-line growth model, $\theta \sim N(5, 5)$; $\eta(0) \sim N(44, 52)$; $\eta(1) \sim N(49,47)$; $\eta(2) \sim N(54,52)$; $\eta(3) \sim N(59,67)$; $\eta(4) \sim N(64,92)$; an exogenous measure W : $W \sim N(10,4)$; with the observables Y including the measurement error ϵ : $\epsilon \sim N(0,12)$. Model parameter values for the artificial data are $t^o = 1$; $\sigma^2_\theta = 5.0$; $\sigma^2_{\eta(t^o)} = 47$ (yielding $\kappa = 3.066$); $\rho_{W\theta} = .60$; $\rho_{W\eta(1)} = .60$; for $\theta \sim N(5, 5)$, $\eta(t^o) \sim N(49, 47)$, $W \sim N(10, 4)$, $\epsilon \sim N(0,12)$

Values of parameters of interest.

For examining the performance of the data analysis procedures we have population values determined as:

a. rate $\mu_\theta = 5.0$ $\sigma^2_\theta = 5.0$ $\rho(\hat{\theta}) = .806$. .

b. Correlation between change and initial status. $\rho_{\eta(0)\theta}$ is $-.310$

c. Regression for exogenous variable, $\beta_{\theta W} = .671$).

Potential observables Y at the observation times $t_i = \{0,1,2,3,4\}$ have population reliabilities {.813, .797, .813, .848, .885}

The upper triangle of the population correlation matrix for the $\eta(t_i)$ at the observation times $t_i = \{0,1,2,3,4\}$ is

0.951	0.808	0.627	0.463
	0.951	0.838	0.715
		0.966	0.896
			0.981

longitudinal research questions and parameters of interest

for present purposes, analysis of individual change methods address the first three questions on my standard listing of research questions

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

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

3. *Stability over Time.* Questions about temporal stability fall into two broad headings--*Is an individual consistent over time? and Are individual differences consistent over time?* (Rogosa, Willett, Floden 1984) **.**
For consistency of individual differences over time:
Parameter: Foulkes-Davis γ = Pr(two growth curves do not intersect)
"tracking" if index > .50 (significantly)

Common Claims: (empirical and metaphysical)
*everyone changes at the same rate (people interchangable)
*change can't be measured reliably/accurately
*correlation of change and initial status is negative; regression toward mean pertains etc

Additional research questions

4. Comparing Experimental Groups.
5. *Comparing Nonexperimental Groups.*
(note: Dental, Wisc compare intact groups via W code)
6. *Analysis of Reciprocal Effects.*
7. *Growth in Multiple Measures.*

Theme for the day (and for the decades): Questions 1-3 addressed by growth-curve approach to the basic individual change problem.
This approach of modeling individual trajectories builds up to more complex settings such as-- group comparisons, hierarchical data structures, non-constant rate-of-change modelling etc.

Data Analysis and Parameter Estimation

Precursor: Descriptive Growth Curve Analyses

SFYS: fit Y on t regressions, describe resulting $\hat{\theta}_p$, fit $\hat{\theta}_p$ on W regr, Examples: WISC, frames 1-4; Ramus, frames 1-3; SmearMiss, frames 1-3. Even non-synchronous data, get variance comps and derived quants by approx method-of-moments (Rogosa-Saner 1995); works surprisingly well.

Maximum Likelihood estimation for parameters

Special, simple case; Complete, Synchronous Data.
ml estimation equations for full data in closed form (Blomqvist 1977)
example estimation of var(theta) σ^2_θ
MSR_p mean squared residual for OLS fit individual p; $\hat{\sigma}^2$ is Ave(MSR_p).
estimate for σ^2_θ : $\hat{\sigma}^2_\theta$ = SS($\hat{\theta}_p$)/"n" - $\hat{\sigma}^2/SSt$,
reliability estimate for $\hat{\theta}_p$: $\hat{\rho}(\hat{\theta})$ = $\hat{\sigma}^2_\theta$ /SS($\hat{\theta}_p$)/"n"

General strategy: get elements of 2x2 est. covariance matrix of θ and $\eta(0)$ for full or incomplete data. Common to **All programs** (**LISREL** HLM Tp) Tp: further substitute for derived quantities.
Also, fixed effects from separate run with W (when exists)–OLS equiv

properties of mle: bias, precision: Is reml best?
bias and mean-square-error : compare ML and REML
mle and reml simulation (50,000); complete synchronous data

Estimation of σ^2_θ var(theta) = 5.0

n	ML		REML	
10	4.37	[7.39]	4.99	[8.61]
15	4.58	[5.06]	4.99	[5.59]

MAJOR MESSAGES

1. OLS equivalences for fixed effects; Method-of-moments match for random effects

2. 2x2 covariance matrix ($\eta_p(0)$ θ_p)-- elements σ^2_θ $\sigma^2_{\eta(0)}$ $\sigma_{\eta(0)\theta}$ –starting point for growth statistics

3. uncertainty, via s.e. and CI, reporting essential–for small (or medium) n, BCa intervals vs standard

$\eta_p(t) = \eta_p(0) + \theta_p t$ Data $Y_{pj} = \eta_p(t_{pj}) + \varepsilon_{pj}$

ordered times $j = 1 \ldots T_p$ # obs for p (missing or not)

intercepts and slopes may differ across $p = 1, \ldots, N$

Now

$E_p(\theta_p) = \mu_\theta$

$E_p(\eta_p(0)) = \mu_{\eta(0)}$

Random: $\eta_p(0) - \mu_{\eta(0)}$, $\theta_p - \mu_\theta$

Fixed: $\mu_{\eta(0)}$, μ_θ

Mixed Effects Model

$$\underset{\sum_{p=1}^{N} T_p \times 1}{Y} = \underset{\sum T_p \times 2}{X}\ \underset{2\times 1}{\beta} + \underset{\sum T_p \times 2N}{Z}\ \underset{2N \times 1}{\gamma} + \varepsilon$$

fixed (ave. growth curve)

$$X = \begin{bmatrix} \begin{bmatrix} 1 & t_{p1} \\ \vdots & \\ 1 & t_{pT} \end{bmatrix} \\ \\ \end{bmatrix} \qquad \beta = \begin{bmatrix} \mu_{\eta(0)} \\ \mu_\theta \end{bmatrix}$$

random

$$\underset{\text{2 rows each } p}{\gamma} = \begin{bmatrix} \begin{bmatrix} \eta_p(0) - \mu_{\eta(0)} \\ \theta_p - \mu_\theta \end{bmatrix} \\ \\ \end{bmatrix}$$

blocks

$$Z = \begin{bmatrix} \begin{bmatrix} 1 & t_{p1} \\ \vdots & \\ 1 & t_{pT} \end{bmatrix} & & \\ & \ddots & \\ & & [\quad] \end{bmatrix}$$

$$Var(Y) = V = Z\underset{\text{also aka } D}{G}Z' + R$$

block elements of G $\begin{bmatrix} \sigma^2_{\eta(0)} & \sigma_{\eta(0)\theta} \\ & \sigma^2_\theta \end{bmatrix}$

Rogosa January 30 Page 7

With W

Fixed

$$E(\eta(0)|W) = \mu_{\eta(0)} + \beta_{\eta(0)W}(W - \mu_W)$$

$$E(\theta|W) = \mu_\theta + \beta_{\theta W}(W - \mu_W)$$

Random

$$\eta_p(0) - E(\eta(0)|W)$$

$$\theta_p - E(\theta|W)$$

$$Y = X\beta + Z\gamma + \varepsilon$$

fixed

$$\text{expanding } \underset{\Sigma T_p \times 4}{X} = \begin{bmatrix} \begin{bmatrix} 1 & t_{pj} & w_p & w_p * t_{pj} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_{pT_p} & w_p & w_p * t_{pT_p} \end{bmatrix} \\ \\ \end{bmatrix}$$

$$\underset{4\times 1}{\beta} = \begin{bmatrix} \mu_{\eta(0)} - \beta_{\eta(0)W}\mu_W \\ \mu_\theta - \beta_{\theta W}\mu_W \\ \beta_{\eta(0)W} \\ \beta_{\theta W} \end{bmatrix} \begin{matrix} \text{int} \\ \text{time} \\ W \\ \text{time} * W \end{matrix}$$

random

$$\underset{2N\times 1}{\gamma} = \begin{bmatrix} \begin{bmatrix} \eta_p(0) - E(\eta(0)|w_p) \\ \theta_p - E(\theta|w_p) \end{bmatrix} \\ \vdots \end{bmatrix}$$

G-matrix contains conditional variances

$Var(\theta|W)$ $\quad Var(\eta(0)|W)$

See NCFem, Frames 7, 8 etc

Rogosa January 30 Page 8

Timepath97–parameter estimation
* obtain estimates for growth curve quantities of interest from solutions (using Make, ODS facility for 6.11+) estimated covariance parameters give t^o, κ, variances and derived quantities (see page 12) etc;
* embed in jackboot.sas to obtain BCa confidence intervals for derived quantities (link to jackboot and docs on Timepath97 Web page)

Implementation of Estimation using SAS- PROC MIXED

(thanks to Neil Timm, Univ Pitt. & Russ Wolfinger, SAS Inc)
REML , ML etc available. (REML matches other E-M programs, e.g SmearMiss via HLM).

S-plus Alternative: `lme`-- Pinheiro & Bates, or further with `nlme`
http://netlib.bell-labs.com/cm/ms/departments/sia/project/nlme/index.html

put data in column form [ID, Y, t, W] Run PROC MIXED without and with W to obtain core quantities for parameter estimation
From no-W run obtain Covariance Parameter Matrix (G);

```
/* Proc mixed run */
 proc mixed data=yt;
   class case;
   model y = time / s;
   random int time / type=un sub=case gcorr;
   make 'CovParms' out=untot;
   make 'SolutionF' out=solfout;
   %bystmt;
 run;
```

fixed effects solution vector gives relations with W

```
 proc mixed data=yt;
   class case;
   model y = time W time*W / s;
   random int time / type=un sub=case gcorr;
   make 'SolutionF' out=solfout;
   %bystmt;
 run;
```

Raw SAS--- frames 7,8 NCFem; frame 7 Ramus; frames 7,8 Smearmiss;

What to do with covariance matrix ($\eta_p(0)$ θ_p) ?

Timepath97 Output, in each data example, constructed from SAS (reml or ml) core estimates

Extensions using properties of collections of growth curves

To estimate growth-curve quantities of interest, substitute core estimates into relations among moments: selected given below

variance

$$\sigma^2_{\eta(t)} = \sigma^2_{\eta(t^o)} + ((t - t^o)/\kappa)^2 \sigma^2_{\eta(t^o)}$$

covariance (also yields correlation, using above)

$$\sigma_{\eta(t_1)\eta(t_2)} = \sigma^2_{\eta(t^o)} + (t_1 - t^o)(t_2 - t^o)\sigma^2_\theta$$

correlation between change and status

$$\rho_{\eta(t)\theta} = \frac{(t - t^o)}{[\kappa^2 + (t - t^o)^2]^{1/2}}$$

correlation between exogenous variable, W and status

$$\rho_{W\eta(t)} = \frac{(t - t^o)\rho_{W\theta} + \kappa\rho_{W\eta(t^o)}}{[\kappa^2 + (t - t^o)^2]^{1/2}}$$

ASSESSMENTS OF STABILITY: index of tracking   Foulkes and Davis (1981): *“tracking”* if index > .50 (significantly)
Estimation. Fit individual trajectories (straight-line or polynomial etc).
For each individual compute the proportion of other trajectories not crossed.
Point estimate is the average over individuals of these proportions.
F-D p.441 use standard deviation of individual estimates divided by Sqrt[n] as the standard error and construct normal theory CI.

Bootstrap results array, in each data example, constructed by reformatting output from jackboot.sas. Choose quantities to bootstrap...
Examples: SmearMiss parameter est.; Ramus frame 6.

Timepath97 Output

Section 1– Descriptive analyses of growth rates.

Individual OLS Fits:
individual fits of a straight-line growth curve (the regression of *Y* on *t* for each *p*) by ordinary least-squares.

Cross-sectional Description: For *synchronous* data sets

OLS Fits: Descriptive Statistics:
comparisons of rates of change across individuals. Stem-and-leaf diagrams and accompanying boxplots for `RATE` (Empirical rate) `INIT_LVL` (Fitted Initial Level) `W` (Exogenous Variable).

OLS Fitted Values for Anchor Times

Rate and Fitted Initial Level: Scatter plot of RATE vs INIT_LVL

OLS Theta-hat on W Regression: When W is present, OLS regression and corresponding scatterplot is given–provides graphical diagnostic and shows correspondence (both point estimate and standard error) with fixed-effects estimates from mixed-model estimation (exact match for complete synchronous data)

OLS Fitted Initial Level on W Regression: When W is present, OLS regression is given–provides graphical diagnostic and shows correspondence (both point estimate and standard error) with fixed-effects estimates from mixed-model estimation (exact match for complete synchronous data)

Section 2– Parameter Estimates

Parameter Estimates collection of estimates based on the growth curve model
First parameters listed are "typical" rates of change μ_θ or median(θ), and a measure of heterogeneity σ^2_θ , the variance of the θ_p . Then estimate of the reliability $\rho(\hat{\theta})$. For each of the stated anchor times, the quantities estimated are: $\sigma^2_{\eta(t)}$ $\rho_{\eta(t)\theta}$ $\beta_{\theta\eta(t)}$ and Reliability Y(t) .
Using the chosen anchor time point, estimate of $\rho_{\eta(t_I)\theta}$;
The index of tracking γ (Foulkes & Davis 1981) ;
Systematic individual differences in growth $\rho_{\theta W}$ or by $\beta_{\theta W}$;

Section 3–Inference using Bootstrap Resampling

Bootstrap Confidence Intervals
Bootstrap estimation is provided for the following parameters

μ_θ `Mean(Rate);` Median(θ); σ^2_θ `var(Rate);` $\rho(\hat{\theta})$ `rel(Rate);`
$\rho_{\eta(t_I)\theta}$ `Corr(Rate, Initial status);` $\beta_{\theta W}$ `Beta(Rate, W)`
$\beta_{\eta(t_I)W}$ `Beta(Eta(T1), W)`

Notes on Data Analysis Examples

OLS equivs for ramus, artN200 (see auxiliary sheet) approx smear

```
Fixed Effects  Smearmiss output
         SAS Solution for Fixed Effects
Effect         Estimate     Std Error    t  Pr > |t|
INTERCEPT   30.47409042    3.83950806   7.94    0.0001
TIME        -0.92464420    1.18313096  -0.78    0.4364
W            1.29057304    0.37026530   3.49    0.0006
TIME*W       0.60101799    0.11408524   5.27    0.0001
============================================
OLS output in Timepath97
         Simple OLS Theta-hat on W Regression
             Parameter      Standard   T for H0:
Variable      Estimate         Error   Parameter=0  Prob > |T|
INTERCEP     -0.901105    1.16435047    -0.774        0.4408
W             0.594518    0.11248947     5.285        0.0001
         Simple OLS Fitted Initial Level on W Regression
             Parameter      Standard   T for H0:
Variable      Estimate         Error   Parameter=0  Prob > |T|
INTERCEP     30.601655    3.80705611    8.038        0.0001
W             1.286608    0.36780483    3.498        0.0007
```

Bootstrap output in Timepath97 (on auxiliary sheet)

Importance of Standard errors small n, estimation imprecise (most often).

advantages of BCa (or ABC) intervals

RAT (HLM n=10)

```
           * var(theta) *
$Desc:  mle mean.boot  se.boot  se.SAS
   10.53901  11.86508 7.613102  9.59
$CI:             0.025      0.05     0.95    0.975
       BCa  1.2796840  1.751531 27.27126 31.87409
           * rel (thetahat) *
$Desc:  mle mean.boot   se.boot
   0.534604 0.5142543 0.1724859
$CI:            0.025      0.05      0.95     0.975
       BCa 0.1255408 0.1753536 0.7408227 0.7660244
           * corr trchange,initial status *
$Desc:   mle  mean.boot  se.boot
  -0.1227116 -0.2241844 0.418065
$CI:             0.025       0.05      0.95     0.975
       BCa -0.7646482 -0.6387207 0.8164504 0.8859661
```

Dental (n=27) rel(Rate) in {0,.8}, Corr(Rate, Initial Status) in {-.4,1.0}

Ramus (n=20) does better than Dental

Other Lessons: Corr(Rate, Initial Status) can be positive and large.

examples: NCFem, WISC Change can be assessed accurately, reliably

Further Investigations

LISREL equivalence adventure–Joreskog vs Gauss?
for Full, Synchronous data LISREL can be formulated (following roughly Willett-Sayer) to carry out a (reml) growth curve estimation. LISREL works almost as well as OLS. See attached model formulation and output (ver 8). In ArtN200 example results reasonable match reml estimates (except for s.e. of some fixed effects). Fit statistic?

Some extensions possible with LISREL that cannot be accomodated with other approaches.

AR(1) Excursion (with ArtN200)

1. AR(1) data construction
for $\eta_p(t_i)$ from ArtN200 construction add AR(1) ρ =.3 errors ϵ ~ N(0,12)

2. Effects of AR(1),
SAS

```
     AR(1) Errors                                 iid Errors

  Cov Parm        Estimate            Cov Parm      Estimate
  UN(1,1)      62.74926568            UN(1,1)    55.11544454
  UN(2,1)      -6.75750327            UN(2,1)    -3.80814129
  UN(2,2)       5.33756422            UN(2,2)     4.39372614
  Residual      8.72009620            Residual   11.90899971
```

LISREL (model mod) similar values to SAS AR(1)

```
 MAXIMUM MODIFICATION INDEX IS  9.86 FOR ELEMENT ( 5, 1) OF THETA-EPS
```

3. Amelioration–fitting AR(1) error structure
SAS line for PROC MIXED /* put in AR(1) */
  repeated /type=AR(1) sub=case;

```
 UN(1,1)         60.16344992
 UN(2,1)         -6.04826235
 UN(2,2)          4.92409041
 AR(1)            0.24049755
 Residual        11.05217550
```

LISREL more dicey to set up constraints–starting values problems, better way to do AR(1)? Band diag works pretty well

Assorted References

Blomqvist, N. (1977). On the relation between change and initial value. *Journal of the American Statistical Association*, 72, 746-749.
Bryk, A.S. & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, **101**, 147-58
Bryk, A.S. & Raudenbush, S. W.(1992). Hierarchical linear models: Applications and data analysis methods. Sage Publications:CA:Lnd.
Bryk, A.S, Raudenbush, S.W, Seltzer,M. Congdon,R.T (1989) An Introduction to HLM: Computer Program and User's guide.
Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
Fearn, T. (1975). A Bayesian approach to growth curves. *Biometrika, 62,* 89-100.
Foulkes, M.A. , & Davis,C.E. (1981) An index of tracking for longitudinal data. *Biometrics*,37, 439-446.
Hui, S. L., & Berger, J. O. (1983). Empirical Bayes estimation of rates in longitudinal studies. *Journal of the American Statistical Association , 78* , 753-760.
Laird, N.M., & Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics* , 38, 963-974.
Goldstein, H. (1979). *The design and analysis of longitudinal studies*. London: Academic Press.
Rogosa, D. R. (1995). Myths and methods: "Myths about longitudinal research," plus supplemental questions. In The analysis of change, J. M. Gottman, Ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 3-65.
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., and Saner, H. M. (1995). Longitudinal data analysis examples with random coefficient models. Journal of Educational and Behavioral Statistics, 20, 149-170. Also: Reply to Discussants, 234-238.
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. (1983). Comparing two indices of tracking. *Biometrics, 39,* 795-6.