Statistics 222, Education 351A Autumn 2020

Statistical Methods for Longitudinal Research

Autumn 2020 Remote Asynchronous Instruction

David Rogosa Sequoia 224, rag{AT}stanford{DOT}edu
Course web page: http://rogosateaching.com/stat222//

To see full course materials from Autumn 2018 go here

Course Welcome and Logistics (first day stuff, to be posted in August, call it Week0)
Lecture slides, week 0 (pdf) **Audio companion, week 0**
For recreation of in-classroom experience, linked below are youtube versions of the music I play
before starting lecture and after lecture concludes. Some may wish to reverse that ordering.

Registrar's information
STATS 222 (Same as EDUC 351A): Statistical Methods for Longitudinal Research Units: 2
Grading Basis: Letter or Credit/No Credit

Course Description:
STATS 222: Statistical Methods for Longitudinal Research (EDUC 351A)
Research designs and statistical procedures for time-ordered (repeated-measures) data.
The analysis of longitudinal panel data is central to empirical research on learning, development, aging, and the effects of interventions.
Topics include: measurement of change, growth curve models, analysis of durations including survival analysis,
experimental and non-experimental group comparisons, reciprocal effects, stability.
See http://rogosateaching.com/stat222/. Prerequisite: intermediate statistical methods
Terms: Aut | Units: 2 | Grading: Letter or Credit/No Credit
Instructors: Rogosa, D. (PI)

Preliminary Course Outline
Week 1. Course Overview, Longitudinal Research; Analyses of Individual Histories and Growth Trajectories
Week 2. Introduction to Data Analysis Methods for assessing Individual Change for Collections of Growth Curves (mixed-effects models)
Week 3. Analysis of Collections of growth curves: linear, generalized linear and non-linear mixed-effects models
Week 4. Special case of time-1, time-2 data; Traditional measurement of change for individuals and group comparisons
Week 5. Assessing Group Growth and Comparing Treatments: Traditional Repeated Measures Analysis of Variance and Linear Mixed-effects Models
Week 6. Comparing group growth continued: Power calculations, Cohort Designs, Cross-over Designs, Methods for missing data, Observational studies.
Week 7. Analysis of Durations: Introduction to Survival Analysis and Event History Analysis
Weeks 8-9. Further topics in analysis of durations: Diagnostics and model modification; Interval censoring, Time-dependence, Recurrent Events, Frailty Models, Behavioral Observations and Series of Events (renewal processes)
Dead Week. Assorted Special Topics: Assessments of Stability (including Tracking), Reciprocal Effects, (mis)Applications of Structural Equation Models, Longitudinal Network Analysis

Texts and Resources for Course Content
1. Garrett M. Fitzmaurice Nan M. Laird James H. Ware Applied Longitudinal Analysis (Wiley Series in Probability and Statistics; 2nd ed 2011)
Text Website second edition website Text lecture slides
2. Judith D. Singer and John B. Willett . Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence New York: Oxford University Press, March, 2003.
Text web page Text data examples at UCLA IDRE Powerpoint presentations good gentle intro to modelling collections of growth curves (and survival analysis) is Willett and Singer (1998)
3. Douglas M. Bates. lme4: Mixed-effects modeling with R February 17, 2010 Springer (chapters). A merged version of Bates book: lme4: Mixed-effects modeling with R January 11, 2010 has been refound
Manual for R-package lme4 and mlmRev, Bates-Pinheiro book datasets.
Additional Doug Bates materials. Collection of all Doug Bates lme4 talks Mixed models in R using the lme4 package Part 2: Longitudinal data, modeling interactions Douglas Bates 8th International Amsterdam Conference on Multilevel Analysis 2011-03-16 another version
Original Bates-Pinheiro text (2000). Mixed-Effects Models in S and S-PLUS (Stanford access). Appendix C has non-linear regression models.
Fitting linear mixed-effects models using lme4, *Journal of Statistical Software* Douglas Bates Martin Machler Ben Bolker. Technical topics: Mixed models in R using the lme4 package Part 4: Theory of linear mixed models
4. A handbook of statistical analyses using R (second edition). Brian Everitt, Torsten Hothorn CRC Press, Index of book chapters Stanford access
Longitudinal chapters: Chap11 Chap12 Chap13. Data sets etc Package 'HSAUR2' August 2014, Title A Handbook of Statistical Analyses Using R (2nd Edition)
There is now a third edition of HSAUR, but full text not yet available in crcnetbase.com. CRAN HSAUR3 page with Vignettes (chapter pieces) and data in reference manual
5. Peter Diggle , Patrick Heagerty, Kung-Yee Liang , Scott Zeger. Analysis of Longitudinal Data 2nd Ed, 2002
Amazon page Peter Diggle home page Book data sets
A Short Course in Longitudinal Data Analysis Peter J Diggle, Nicola Reeve, Michelle Stanton (School of Health and Medicine, Lancaster University), June 2011 earlier version associated exercises: Lab 1 Lab2 Lab3
6. Longitudinal and Panel Data: Analysis and Applications for the Social Sciences by Edward W. Frees (2004). Full book available and book data and

programs (mostly SAS).
7. Growth Curve Analysis and Visualization Using R. Daniel Mirman Chapman and Hall/CRC 2014 Print ISBN: 978-1-4665-8432-7 Stanford Access Mirman web page (including data links).
8. *Longitudinal Data Analysis* Edited by Geert Verbeke , Marie Davidian , Garrett Fitzmaurice , and Geert Molenberghs Chapman and Hall/CRC 2008. online supplement for LDA book .
9. Verbeke, G. and Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. Springer Series in Statistics. New-York: Springer. Extended presentation: Introduction to Longitudinal Data Analysis A shorter exposition: Methods for Analyzing Continuous, Discrete, and Incomplete Longitudinal Data
10. Survival analysis Rupert G. Miller. Available as Stanford Tech Report
11. Event History Analysis with R (Stanford access). Goran Brostrom CRC Press 2012. R-package `eha`
12. John D. Kalbfleisch , Ross L. Prentice The Statistical Analysis of Failure Time Data 2nd Ed
Amazon page online from Wiley
13. Klein J, Moeschberger M (2003). Survival Analysis, 2nd edition. New York: Springer.
14. Therneau TM, Grambsch PM (2000). Modeling Survival Data: Extending the Cox Model. New York: Springer.
15. Advanced survival analysis topics.
Interval-Censored Time-to-Event Data Methods and Applications Chapman and Hall/CRC 2012 (esp Chap 14--glrt).
Recurrent Events: Chapter 9 of Kalbfleisch and Prentice (2nd edition), "Modeling and Analysis of Recurrent Event Data".
Cook, R. J. and Lawless, J. F. (2007). The Statistical Analysis of Recurrent Events. (Stanford access) Springer, New. York.
Joint Models for Longitudinal and Time-to-Event Data. With Applications in R. Dimitris Rizopoulos. Chapman and Hall/CRC 2012(Stanford access) Book website

Additional Specialized Resources
Harvey Goldstein. The Design and Analysis of Longitudinal Studies: Their Role in the Measurment of Change (1979). Elsevier
Amazon page Goldstein Chap 6 Repeated measures data Multilevel Statistical Models by Harvey Goldstein with data sets
David Roxbee Cox, Peter A. W. Lewis The statistical analysis of series of events. Chapman and Hall, 1966
Google books Poisson process computing program
David J Bartholomew. Stochastic Models for Social Processes, Chichester 3rd edition: John Wiley and Sons.
David J Bartholomew web page

Grading, Exams, and Credit Units
Stat222/Ed351A is listed as Letter or Credit/No Credit grading for 2-units
For Autumn 2020 grading for the 2-units will be based on a 'take home'(i.e. do at home) Problem Set.
Each week are posted a few exercises for that week's content--towards the end of the qtr I will identify a subset of those exercises to be turned in.
Those selected problems will constitute the graded Problem Set.
Also as you will see, for each week's content a number of Review Questions with Solutions are posted.

Course Problem Set 2020 partial posting Oct 30; complete posting Nov 10

Statistical computing
Class presentation will be in, and students are encouraged to use, R (occasionally, some references to SAS and Mathematica).
Current version of R is R version 4.0.2 (Taking Off Again) released 2020-06-22.
For references and software: The R Project for Statistical Computing Closest download mirrors in the past, UCLA and Berkeley, seem no longer avaliable, pick your fave anywhere in the world.
The CRAN Task View: Statistics for the Social Sciences provides an overview of some relevant R packages. Also the new CRAN Task View: Psychometric Models and Methods and CRAN Task View: Survival Analysis and CRAN Task View: Computational Econometrics.
A good R-primer on various applications (repeated measures and lots else). Notes on the use of R for psychology experiments and questionnaires Jonathan Baron, Yuelin Li. Another version
A Stat209 text, Data analysis and graphics using R (2007) J. Maindonald and J. Braun, Cambridge 2nd edition 2007. 3rd edition 2010 has available a short version in CRAN .
According to Peter Diggle: "The best resource for R that I have found is Karl Broman's Introduction to R page."

Course Content: Files, Readings, Examples

Week 1. First class: Longitudinal Research Overview, Analysis of Individual Trajectories.

Lecture slides, week 1 (pdf)
Audio companion, week 1
parta partb partc

Lecture Topics
A. Longitudinal research overview
B. Examples, illustrations for longitudinal research overview, taken from course resources above:
Laird,Ware (#1) slides 1-16; Diggle (#5) slides 4-14, 22-28 Verbeke (#9) slides from Ch 2 and Sec3.3
C. Data Analysis Examples of Model Fitting for Individual Trajectories and Histories.
Motto: Individual trajectories are the proper starting point for longitudinal data analysis
ascii version of class handout annotated version pdf version with plots datasets
Starting up R-addendum: installing packages and obtaining data (sleepstudy in lme4)

Additional materials for the trajectory examples
For Count Data (glm) example. Link functions for generalized linear mixed models (GLMMs), Bates slides (pdf pages 11-18)
AIDS in Belgium example, (from Simon Wood) single trajectory, count data using glm. Rogosa R session for aids data
aditional expositions of AIDS data, Poisson regression: Duke Kentucky

WEEK 9 Exercises

1. Dialysis Hemodialysis Data from Brazil.
Data on 6805 hemodialysis patients in all federally funded clinics (67 centers) in Rio de Janeiro State, Brazil.
`data(Dialysis)` package `RcmdrPlugin.survival`

```
Format
A data frame with 6805 observations on the following 7 variables.
center a numeric code indicating in which of 67 centers the patient was treated.
age of the patient.
begin The month in which treatment began, with 1 representing January 1998.
end The month in which observation terminated, either because of death or censoring. The study ended in month 44 (August, 2000).
event 1, death, or 0, censoring.
time the difference between end and begin.
disease a factor with levels congen, (congenital); diabetes; hypert (hypertension); other; and renal.
```

Use Cox regression with age and disease as predictors to study survival time of patients. Follow the in-class analysis of the eortc data to compare results ignoring centers and using a mixed-effects models with random factor `center`.

2. Time-dependent covariate, Stanford Heart Transplant data.
Data are available in the `survival` package in various formats: see `heart` listing. In the Stanford Heart Transplant study, all patients start on medical treatment. When a heart becomes available, the selected patient 'converts' to the transplant treatment arm. Thus the intervention variable `transplant` is time-dependent. Use these survival data (including the additional covariates if desirable) in a format of your choosing to evaluate the effect of transplant on survival. Additional discussion of these data is contained in the Thernau overview of advanced survival analysis topics, Multiple and Correlated Events, linked in week 9 materials (pdf page 26 onward) and in the Therneau et al primer Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model.

Week 10, Dead Week: Special Topics for Longitudinal Data

Lecture topics: (pick some)
1. Observational Studies, Survival analysis. Marriage and Cancer Survival example (from weeks 5 and 6, Stat266 2016)
2. Stability over time (Myth 8). Change and Sameness
3. Structural equation models for longitudinal data (don't do it; Myth 7) (handout in week 10 handout collection)
4. Reciprocal effects (Myth 9) Rogosa, Encyclopedia of Social Science
5. Longitudinal Network Data

Resources

Applications of Structural Equation Models (LISREL, path analysis, Myth 7)
David Rogosa. Casual Models Do Not Support Scientific Conclusions: A Comment in Support of Freedman. Journal of Educational Statistics, Vol. 12, No. 2. (Summer, 1987), pp. 185-195. Jstor link Theme Song *Ballad of the casual modeler* http://www.stanford.edu/class/ed260/ballad.mp3
Rogosa, D. R. (March 1994). Longitudinal reasons to avoid structural equation models, UC Berkeley.
Rogosa, D. R. (1993). Individual unit models versus structural equations: Growth curve examples. In Statistical modeling and latent variables, K. Haagen, D. Bartholomew, and M. Diestler, Eds. Amsterdam: Elsevier North Holland, 259-281.
original publication on the longitudinal path analysis: Some Models for Analysing Longitudinal Data on Educational Attainment. Harvey Goldstein *Journal of the Royal Statistical Society. Series A* (General), Vol. 142, No. 4. (1979), pp. 407-442. Jstor link
Rogosa, D. R., & Willett, J. B. (1985). Satisfying a simplex structure is simpler than it should be. Journal of Educational Statistics, 10, 99-107. Jstor link Follow-up paper: Two Aspects of the Simplex Model: Goodness of Fit to Linear Growth Curve Structures and the Analysis of Mean Trends. Frantisek Mandys; Conor V. Dolan; Peter C. M. Molenaar. Journal of Educational and Behavioral Statistics, Vol. 19, No. 3. (Autumn, 1994), pp. 201-215. Jstor link

Stability: Consistency, Change and Sameness (Myth 8)
J.H. Ware Tracking in S. Kotz, N.L. Johnson (Eds.), The Encyclopedia of Statistical Sciences (13th Edn.), Vol. 9 John Wiley, New York (1988)
Rogosa, D. R., Floden, R. E., & Willett, J. B. (1984). Assessing the stability of teacher behavior. Journal of Educational Psychology, 76, 1000-1027. APA link also available from John Willet's pub page
Rogosa, D. R., & Willett, J. B. (1983). Comparing two indices of tracking. Biometrics, 39, 795-6. JStor link
Rogosa, D. R. Stability section of Individual unit models versus structural equations (link above)
Rogosa, D. R. Stability of school scores from educational assessments: Confusions about Consistency in Improvement David Rogosa, June 2003 ; Education Writers Association April 2004
A non-parametric index of tracking James A. Koziol Pages 181-191 Published online: 02 Aug 2010 Journal of Applied Statistics Volume 25, 1998 - Issue 2
Personality research. Stability versus change, dependability versus error: Issues in the assessment of personality over time David Watson Journal of Research in Personality 38 (2004) 319-350.
Some applications: A Stochastic Model for Analysis of Longitudinal AIDS Data J.M.G. Taylor, W.G. Cumberland, Sy J.P.; Journal of the American Statistical Association, Vol. 89, 1994
Body mass index from childhood to middle age: a 50-y follow-up V A Casey J T Dwyer K A Coleman I Valadian The American Journal of Clinical Nutrition, Volume 56, Issue 1, 1 July 1992, Pages 14-18
Tracking and its applicability to Physical Education and Sport
Tracking of objectively measured physical activity from childhood to adolescence: The European youth heart study. Scandinavian Journal of Medicine & Science in SportsVolume 18, Issue 2, 2007.
Factors Associated With Tracking of BMI: A Meta-Regression Analysis on BMI Tracking. Obesity (2011) 19 5, 1069-1076. doi:10.1038/oby.2010.250
Long-term tracking of cardiovascular risk factors among men and women in a large population-based health system The Vorarlberg Health Monitoring and

Promotion Programme. European Heart Journal (2003) 24, 1004-1013.
Journal of Traumatic Stress. Reliability of Reports of Violent Victimization and Posttraumatic Stress Disorder Among Men and Women With Serious Mental Illness Volume 12 Issue4 587 - 599 1999-10-01 Lisa A. Goodman Kim M. Thompson Kevin Weinfurt Susan Corl Pat Acker Kim T. Mueser Stanley D. Rosenberg
Computing: Foulkes-Davis gamma (not in R). A GAUSS program for computing the Foulkes-Davis tracking index for polynomial growth curves
TRACK: A FORTRAN program for calculating the Foulkes-Davis tracking index Gerard E. Dallal Computers in Biology and Medicine Volume 19, Issue 5, 1989, Pages 367-371

Reciprocal effects (Myth 9).
Rogosa, D. R. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, 88, 245-258. APA site version
An old review of reciprocal effects. Rogosa, D. R. (1985). Analysis of reciprocal effects. In International Encyclopedia of Education, T. Husen and N. Postlethwaite, Eds. London: Pergamon Press, 4221-4225. (reprinted in Educational Research,Methodology & Measurement: An international handbook, J. P. Keeves Ed. Oxford: Pergamon Press, 1988.)
Granger Causality. Nobel 2003. Complete Granger
Relationships--and the Lack Thereof--Between Economic Time Series, with Special Reference to Money and Interest Rates. David A. Pierce *Journal of the American Statistical Association*, Vol. 72, No. 357. (Mar., 1977), pp. 11-26. Jstor

Longitudinal Networks
R-package `RSiena` manual, RSiena Resource page
Huisman, M. E. and Snijders, T. A. B. (2003). Statistical analysis of longitudinal network data with changing composition. Sociological Methods and Research, 32:253-287.
Application: Kids' friends influence physical activity levels Publication: The Distribution of Physical Activity in an After-school Friendship Network Sabina B. Gesell, Eric Tesdahl, Eileen Ruchman, Pediatrics; originally published online May 28, 2012.

WEEK 10 Review Questions
1. Stability of Individual Differences
For the Ramus data (week 2 exercise, 20 individuals, 4 time points), the Foulkes-Davis (gamma) index of tracking has point estimate .83 and (bootstrap) standard error .06 for the 18-month time interval 8yrs to 9.5 years. Compare that estimate of consistency of individual differences with time1-time2 correlations for the time intervals [8, 9.5] and [9, 9.5].
2. Path Analysis. Replicate the 3-wave path analysis example based on the Goldstein model for UK reading scores. Use the perfectly measured data Xi(t) in Week 1 Myths chapter data examples and description. Repeat with the fallible X(t) and compare results.
3. Try a path analysis on the 4-waves of the Ramus data. Compare with the growth curve analyses in week 2 exercises. (Solution: see pages 52-54 of 1994 Myths chap on class memory stick)

Week 10 solutions

By **DENNIS THOMPSON / HEALTHDAY /** *April 11, 2016, 12:11 PM*

Marriage may help fight cancer

A new study finds married people have higher survival rates after a cancer diagnosis than patients who are single. / **LEV DOLGACHOV**

A wedding band may be powerful medicine against cancer, a new study suggests.

Married people seem to have a marked survival advantage, the researchers found: Single men with cancer had a death rate 27 percent higher than it was for married male patients, while the death rate for single female patients was 19 percent higher than their married counterparts.

"The effects that we find were actually quite notable," said study author Scarlett Lin Gomez, a research scientist with the Cancer Prevention Institute of California. "They are comparable to some of the more clinical factors we often see that are associated with cancer prognosis, like stage of disease or certain types of treatment."

What's more, this advantage appears to rely solely on the emotional bonds of matrimony, and not the financial advantages that marriage offers, Gomez said.

"These patterns were very minimally explained by the married patients having greater economic resources," Gomez said. "Specifically, we looked at health insurance and we looked at living in a higher socioeconomic status neighborhood. Even though these played a small role, they really didn't explain the greater survival among the married."

It should be noted that the study only found an association between marital status and cancer prognosis; it did not prove a cause-and-effect link.

Previous studies over the past 10 to 15 years have shown a similar marriage benefit for cancer patients, Gomez said, but the benefit has always been chalked up to the love and support a person receives from their spouse.

But married people also tend to have greater combined incomes and better access to insurance, Gomez said. She and her colleagues decided to see whether money played a role in the survival advantage of married couples.

For the analysis, the researchers examined health records of nearly 800,000 adults in California who were diagnosed with invasive cancer between 2000 and 2009 and were followed through 2012.

The study, published April 11 in the journal *Cancer*, revealed that financial resources did not have much impact on a person's chance of beating cancer, Gomez said.

Instead, it appears that cancer patients benefit from the support that a spouse provides, such as driving them to doctor's appointments, providing nourishing meals and making sure they take their medicine, Gomez said.

"Treatments can drag on for months and months," she said. "It can be very difficult if you're single and you don't have any other means to get to the doctor."

Married cancer patients also benefit from emotional support, which can reduce stress and improve a person's chances, Gomez added.

"You have somebody who's there to listen to you, to counsel you through the stress of cancer treatment," she said. "Cancer is a very scary thing, and it's good to have someone by your side."

Dr. Gregory Masters, an oncologist with the Helen F. Graham Cancer Center in Newark, Del., said the study "suggests the value of a devoted caregiver in helping to improve survival in cancer patients.

"Social support provided by devoted caregivers such as a spouse may be a major component of the improved survival for these patients," said Masters, an expert on patient care for the American Society of Clinical Oncology. "It suggests that a concerted effort to evaluate a patient's psychosocial resources may be as important as other factors in helping to improve cancer survival."

Single people may not have someone in their lives who will automatically be by their side, but these results suggest that it is vitally important for them to reach out to family and friends for care and support when facing cancer treatment, Gomez and Masters said.

"Certainly we're not advocating for getting married as a means to improve your outcome," Gomez said. "But single people can help themselves by maintaining stronger social networks, and being able to rely on friends and family members for help."

Being in a long-term relationship has been shown to improve health and quality of life across the board, said Rachel Cannady, strategic director of cancer caregiver support for the American Cancer Society.

"These findings aren't unique to cancer," Cannady said. "We know that people live longer when they're couples, period, irrespective of disease. There's something to be said about having that sort of deep and meaningful relationship."

The beneficial effect of being married on survival differed across racial and ethnic groups, according to a companion study also published in the same issue of the journal.

Among men and women, whites seemed to benefit the most from being married, while Hispanics and Asian/Pacific Islanders benefited less. Also, Hispanic and Asian/Pacific Islander cancer patients who were born in the United States experienced a greater benefit than those born outside the country.

Those racial and ethnic groups with an advantage might benefit from a better

Original Article

Differences in Marital Status and Mortality by Race/Ethnicity and Nativity Among California Cancer Patients

María Elena Martínez, PhD[1,2]; Kristin Anderson, MD[1]; James D. Murphy, MD[1]; Susan Hurley, MPH[3]; Alison J. Canchola, MS[3]; Theresa H. M. Keegan, PhD[4]; Iona Cheng, PhD[3,6]; Christina A. Clarke, PhD[3,5,6]; Sally L. Glaser, PhD[3,5,6]; and Scarlett L. Gomez, PhD[3,5,6]

BACKGROUND: It has been observed that married cancer patients have lower mortality rates than unmarried patients, but data for different racial/ethnic groups are scarce. The authors examined the risk of overall mortality associated with marital status across racial/ethnic groups and sex in data from the California Cancer Registry. **METHODS:** California Cancer Registry data for all first primary invasive cancers diagnosed from 2000 through 2009 for the 10 most common sites of cancer-related death for non-Hispanic whites (NHWs), blacks, Asians/Pacific Islanders (APIs), and Hispanics were used to estimate multivariable hazard ratios (HRs) and 95% confidence intervals (CIs) for marital status in relation to overall mortality by race/ethnicity and sex. The study cohort included 393,470 male and 389,697 female cancer patients and 204,007 and 182,600 deaths from all causes, respectively, through December 31, 2012. **RESULTS:** All-cause mortality was higher in unmarried patients than in married patients, but there was significant variation by race/ethnicity. Adjusted HRs (95% CIs) ranged from 1.24 (95% CI, 1.23-1.26) in NHWs to 1.11 (95% CI, 1.07-1.15) in APIs among males and from 1.17 (95% CI, 1.15-1.18) in NHWs to 1.07 (95% CI, 1.04-1.11) in APIs among females. All-cause mortality associated with unmarried status compared with married status was higher in US-born API and Hispanic men and women relative to their foreign-born counterparts. **CONCLUSIONS:** For patients who have the cancers that contribute most to mortality, being unmarried is associated with worse overall survival compared with being married, with up to 24% higher mortality among NHW males but only 6% higher mortality among foreign-born Hispanic and API females. Future research should pursue the identification of factors underlying these associations to inform targeted interventions for unmarried cancer patients. ***Cancer*** **2016;000:000–000**. © *2016 American Cancer Society.*

KEYWORDS: marriage, mortality, nativity, neighborhood socioeconomic status, race/ethnicity.

INTRODUCTION

Results from several studies have demonstrated a lower risk of mortality among married cancer patients compared with their unmarried counterparts. A large meta-analysis and a recent population-based study of data from the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) program indicate that never-married, divorced/separated, and widowed cancer patients had higher mortality than married patients, with relative risks ranging from 1.10 to 1.23.[1,2] Being married is also associated with earlier cancer stage at diagnosis and receipt of definitive treatment.[1,3-7] Proposed reasons for the beneficial effects of being married include having stronger social support and social networks, resulting in higher psychological well being and help with navigating the health care system[2]; having medical insurance[8]; and economic well being[9] as well as improved behavioral and psychological function.[1-4,7]

Despite the considerable literature on marital status and cancer outcome and the well recognized racial/ethnic differences in cancer mortality and survival,[10,11] data on whether the impact of marriage varies across racial/ethnic groups are lacking. It is noteworthy that, in the United States, the proportion of adults who have never been married has risen from 10% in 1960 to 23% in 2012 among men and from 8% to 17% among women.[12] Furthermore, wide variations by

Corresponding author: María Elena Martínez, PhD, University of California, San Diego, Moores Cancer Center, 3855 Health Sciences Drive, No. 0901, La Jolla, CA 92093-0829; Fax: (858) 822-2399; e8martinez@ucsd.edu

[1]Moores Cancer Center, University of California, San Diego, La Jolla, California; [2]Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, California; [3]Cancer Prevention Institute of California, Fremont, California; [4]Department of Internal Medicine, Division of Hematology and Oncology, University of California, Davis, Sacramento, California; [5]Department of Health Research and Policy (Epidemiology), School of Medicine, Stanford, California; [6]Stanford Cancer Institute, Stanford, California.

The ideas and opinions expressed herein are those of the author(s) and endorsement by the State of California, Department of Public Health the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors is not intended nor should be inferred.
Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.29886, **Received:** October 7, 2015; **Revised:** December 1, 2015; **Accepted:** December 11, 2015, **Published online** Month 00, 2016 in Wiley Online Library (wileyonlinelibrary.com)

TABLE 2. *Continued*

Characteristic	Percentage									
	All, N = 389,697		NHW, N = 260,529		Black, N = 24,693		Hispanic, N = 60,132		API, N = 41,253	
	Unmarried, n = 191,059	Married, n = 198,638	Unmarried, n = 128,640	Married, n = 131,889	Unmarried, n = 16,789	Married, n = 7904	Unmarried, n = 28,413	Married, n = 31,719	Unmarried, n = 15,602	Married, n = 25,651
Yes	64.1	76.9	63.8	76.8	61.1	73.5	66.8	77.7	65.9	77.7
Unknown	3.1	2.7	3.1	2.6	2.4	2.5	3.6	3.3	2.6	2.1
Radiation										
No	72.9	64.8	72.5	63.4	72.4	68	73.5	66.9	75.5	67.8
Yes	27.1	35.2	27.5	36.6	27.6	32	26.5	33.1	24.5	32.2
Nativity										
US born							47.4	45.8	18.8	17.2
Foreign born							52.6	54.2	78.7	79.7
Unknown									2.4	3.2

Abbreviations: API, Asian/Pacific Islander; IBD, intrahepatic bile duct; NA, not applicable; NHL, non-Hodgkin lymphoma; Q, quintile; SES, socioeconomic status.

occupation, employment, household income, poverty, rent and house values from the Census 2000 Summary File (for cases diagnosed 2000-2005) and from American Community Survey data from 2007 to 2011 (for cases diagnosed 2006-2009).[17,18]

Follow-Up and Vital Status

Follow-up for overall mortality was computed as the number of days between the date of diagnosis and the first occurrence of the following: date of death, date of last known contact, or end date of follow-up (December 31, 2012). We also considered cancer-specific deaths based on the underlying causes of death (coded to the *International Classification of Diseases, 10th Revision*), and follow-up was censored at the date of death for those who died from an underlying cause other than the primary cancer.

Statistical Analysis

Prior analyses of these data[19] indicated significant differences in mortality by sex; therefore, all analyses presented here were conducted separately for males and females. Consistent with a recent review of mortality and marital status indicating that mortality risks did not vary across subcategories of unmarried status,[20] we conducted analyses using marital status coded as married and unmarried (never married, separated, divorced, and widowed). We used chi-square tests to compare demographic and clinical characteristics by marital status and multivariable Cox proportional hazards regression models to estimate hazard rate ratios[9] and 95% confidence intervals (CIs)[21] for overall mortality and cancer-specific mortality, by sex and race/ethnicity. The proportional hazards assumption was tested for marital status and for each covariate using correlation tests of time versus scaled Schoenfeld residuals. The assumption of proportional hazards was violated for SEER summary stage and age at diagnosis. Thus, we computed stage-stratified and age-stratified Cox regression models, which allowed the baseline hazards to vary by both disease stage and age at diagnosis. The models were adjusted for cancer site, surgery, radiation, systemic hormone agents, nSES, and sources of medical payment. We modeled the top 10 cancer sites combined for each sex as well as each cancer site separately. Models that included all of the cancer sites combined excluded cases with leukemia, because stage and surgery were not applicable. We also conducted separate analyses by nativity among Hispanics and APIs and for the 6 largest API ethnic groups: Chinese, Japanese, Filipino, Korean, South Asian, and Vietnamese. Greater than 40% of Hispanics had Spanish origin coded as "Hispanic, NOS," so we did not analyze

Cox Regression Cancer

NDVI dose quintiles this week

weeks 7-9 Stat 222

$h(t) = h_0(t) \exp(\beta \tilde{X})$ Cox regression PH

$\log h(t) = \beta_1 \text{Married} + \alpha(t)$

observational data

multiple regression

$+ \beta_2 \text{site} + \beta_3 \text{surgery}$

$+ \beta_4 \text{radiation} + \beta_5 \text{hormone}$

$+ \beta_6 \text{nSES} + \beta_7 \text{paym} + \varepsilon$

ancova in log(hazard), how nice

$\exp(\hat{\beta}_1)$ multiplicative effect on hazard

$h(t) = -d \log(S(t))/dt$ Instantaneous Failure rate

hazard ↑ survival ↓

integrated hazard $H(t) = \int_0^t h(t)\,dt$ $S(t) = \exp(-H(t))$

TABLE 3. Crude and Adjusted Hazard Ratios and 95% Confidence Intervals for All-Cause Mortality Associated With Marital Status (Unmarried vs Married) Stratified by Sex and Nativity

	Males: HR (95% CI)				
Variable	All Races, N = 377,932	NHW, N = 247,722	Black, N = 30,281	Hispanic, N = 56,754	API, N = 36,901
Overall, crude	1.57 (1.56–1.59)	1.63 (1.61–1.65)	1.64 (1.59–1.69)	1.56 (.52–1.59)	1.37 (1.32–1.41)
Overall, adjusted[a]	1.22 (1.21–1.24)	1.24 (1.23–1.26)	1.20 (1.16–1.24)	1.20 (1.17–1.23)	1.11 (1.07–1.15)
$P_{heterogeneity}$ for race[b]	6.2×10^{-9}				
By nativity					
US born, crude	—	—	—	1.59 (1.54–1.65)	1.49 (1.38–1.610
US born, adjusted[a]	—	—	—	1.21 (1.17–1.26)	1.21 (1.11–1.32)
Foreign born, crude	—	—	—	1.52 (1.46–1.57)	1.39 (1.34–1.44)
Foreign born, adjusted[a]	—	—	—	1.17 (1.13–1.21)	1.09 (1.05–1.13)
$P_{heterogeneity}$ for nativity[b]	—	—	—	.085	.0024

	Females: HR (95% CI)[c]				
Variable	All Races, N = 378,447	NHW, N = 253,102	Black, N = 24,098	Hispanic, N = 58,008	API, N = 40,309
Overall, crude	1.80 (1.78–1.82)	1.83 (1.81–1.85)	1.54 (1.48–1.60)	1.63 (1.59–1.68)	1.75 (1.69–1.81)
Overall, adjusted[a]	1.15 (1.14–1.16)	1.17 (1.15–1.18)	1.09 (1.05–1.13)	1.11 (1.08–1.14)	1.07 (1.04–1.11)
$P_{heterogeneity}$ for race[b]	$<1.0 \times 10^{-30}$				
By nativity					
US born, crude	—	—	—	1.64 (1.58–1.70)	1.88 (1.73–2.05)
US born, adjusted[a]	—	—	—	1.17 (1.13–1.22)	1.20 (1.09–1.32)
Foreign born, crude	—		—	1.63 (1.57–1.69)	1.74 (1.68–1.80)
Foreign born, adjusted[a]	—	—	—	1.06 (1.02–1.10)	1.06 (1.02–1.10)
$P_{heterogeneity}$ for nativity[b]	—	—	—	1.5×10^{-8}	4.6×10^{-6}

Abbreviations: API, Asian/Pacific Islander; CI, confidence interval; HR, hazard ratio; NHW, non-Hispanic white.
[a] Adjusted HRs were estimated from Cox proportional hazard models stratified by stage (localized, regional, distant, unknown) and age at diagnosis (in years) and were adjusted for: cancer site (prostate, lung, colon, non-Hodgkin lymphoma, bladder, liver and intrahepatic bile duct, pancreas, stomach, and esophagus), race/ethnicity (NHW, black, Hispanic, API, other/unknown for the all races model only), first course of treatment (yes or no for surgery, radiation, hormone therapy), insurance status (no insurance, private insurance only, Medicare only/Medicare and private insurance, any public/Medicaid/military insurance, unknown), and neighborhood socioeconomic status (quintiles).The analysis excluded 15,538 leukemia cases in males.
[b] The likelihood ratio test for interaction was computed based on cross-product terms.
[c] Models for females include sites in the breast, lung, colon, uterus, non-Hodgkin lymphoma, ovary, pancreas, brain, and liver and intrahepatic bile duct. The analysis excluded 11,250 leukemia cases in females.

TABLE 4. Crude and Adjusted Hazard Ratios and 95% Confidence Intervals for All-Cause Mortality Associated With Marital Status (Unmarried vs Married) by Sex and Asian and Pacific Islander Ethnicity

	Males			Females		
API Group	No.	Crude HR (95% CI)	Adjusted HR (95% CI)[a]	No.	Crude HR (95% CI)	Adjusted HR (95% CI)[a]
Chinese	10,752	1.35 (1.26–1.44)	1.11 (1.03–1.19)	10,656	1.87 (1.76–1.99)	1.08 (1.01–1.16)
Japanese	4147	1.30 (1.19–1.43)	1.17 (1.05–1.31)	5084	1.53 (1.41–1.66)	1.02 (0.93–1.13)
Filipino	9662	1.40 (1.31–1.50)	1.12 (1.04–1.21)	11,868	1.71 (1.61–1.81)	1.12 (1.05–1.20)
Korean	3225	1.46 (1.31–1.63)	1.16 (1.01–1.32)	2931	1.99 (1.77–2.23)	1.24 (1.06–1.44)
South Asian	1831	1.51 (1.25–1.82)	1.15 (0.88–1.50)	2039	1.73 (1.47–2.04)	0.98 (0.77–1.24)
Vietnamese	4255	1.33 (1.22–1.46)	1.09 (0.99–1.21)	3621	1.72 (1.55–1.90)	1.08 (0.96–1.22)

Abbreviations: API, Asian/Pacific Islander; HR, hazard ratio; CI, confidence interval.
[a] Adjusted HRs were estimated from Cox proportional hazard models stratified by stage (localized, regional, distant, unknown) and age at diagnosis (in years) and were adjusted for: cancer site (males: prostate, lung, colon, non-Hodgkin lymphoma, bladder, liver and intrahepatic bile duct, pancreas, stomach, esophagus; females: breast, lung, colon, uterus, non-Hodgkin lymphoma, ovary, pancreas, brain, and liver and intrahepatic bile duct), first course of treatment (yes or no for surgery, radiation, hormone therapy), insurance status (no insurance, private insurance only, Medicare only/Medicare and private insurance, any public/Medicaid/military insurance, unknown), neighborhood socioeconomic status (quintiles), and nativity (US born, foreign born). Leukemia cases were excluded.

data for specific Hispanic origin. All statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC). Statistical tests were 2-sided with an α value .05. We did not obtain informed consent from the patients, because we analyzed deidentified cancer registry data.

subclassification matching (say 5 bins); outcome analysis coxme analog to lmer or repeated KM 't-tests'

Package 'coxme'

June 15, 2015

Title Mixed Effects Cox Models

Priority optional

Version 2.2-5

Depends survival (>= 2.36.14), methods, bdsmatrix(>= 1.3)

Imports nlme, Matrix (>= 1.0)

Suggests mvtnorm, kinship2

LinkingTo bdsmatrix

LazyData Yes

LazyLoad Yes

Description Cox proportional hazards models containing Gaussian random effects, also known as frailty models.

License LGPL-2

URL http://r-forge.r-project.org

NeedsCompilation yes

Author Terry M. Therneau [aut, cre]

Maintainer Terry M. Therneau <therneau@mayo.edu>

Repository CRAN

Date/Publication 2015-06-15 17:08:23

R topics documented:

anova.coxme 2
coxme 3
coxme.control 5
coxme.object 6
coxmeFull 8
coxmeMlist 9
eortc 10
expand.nested 11
fixed.effects 12

`call` the calling sequence for the routine.

The computations require the original `x` matrix of the Cox model fit. Thus it saves time if the `x=TRUE` option is used in `coxph`. This function would usually be followed by both a plot and a print of the result. The plot gives an estimate of the time-dependent coefficient `beta(t)`. If the proportional hazards assumption is true, `beta(t)` will be a horizontal line. The printout gives a test for `slope=0`.

References

P. Grambsch and T. Therneau (1994), Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika,* **81**, 515-26.

See Also

`coxph`, `Surv`.

Alternatives: propensity w/out matching

ancova: propen covariate
IPTW: weights in coxph

Examples

```
fit <- coxph(Surv(futime, fustat) ~ age + ecog.ps,
             data=ovarian)
temp <- cox.zph(fit)
print(temp)                  # display the results
plot(temp)                   # plot curves
```

---

coxph *Fit Proportional Hazards Regression Model*

Description

Fits a Cox proportional hazards regression model. Time dependent variables, time dependent strata, multiple events per subject, and other extensions are incorporated using the counting process formulation of Andersen and Gill.

Usage

```
coxph(formula, data=, weights, subset,
      na.action, init, control,
      ties=c("efron","breslow","exact"),
      singular.ok=TRUE, robust=FALSE,
      model=FALSE, x=FALSE, y=TRUE, tt, method, ...)
```

Arguments

`formula` a formula object, with the response on the left of a ~ operator, and the terms on the right. The response must be a survival object as returned by the `Surv` function.

`data` a data.frame in which to interpret the variables named in the `formula`, or in the `subset` and the `weights` argument.

weights	vector of case weights. If weights is a vector of integers, then the estimated coefficients are equivalent to estimating the model from data with the individual cases replicated as many times as indicated by weights.
subset	expression indicating which subset of the rows of data should be used in the fit. All observations are included by default.
na.action	a missing-data filter function. This is applied to the model.frame after any subset argument has been used. Default is options()$na.action.
init	vector of initial values of the iteration. Default initial value is zero for all variables.
control	Object of class coxph.control specifying iteration limit and other control options. Default is coxph.control(...).
ties	a character string specifying the method for tie handling. If there are no tied death times all the methods are equivalent. Nearly all Cox regression programs use the Breslow method by default, but not this one. The Efron approximation is used as the default here, it is more accurate when dealing with tied death times, and is as efficient computationally. The "exact partial likelihood" is equivalent to a conditional logistic model, and is appropriate when the times are a small set of discrete values. See further below.
singular.ok	logical value indicating how to handle collinearity in the model matrix. If TRUE, the program will automatically skip over columns of the X matrix that are linear combinations of earlier columns. In this case the coefficients for such columns will be NA, and the variance matrix will contain zeros. For ancillary calculations, such as the linear predictor, the missing coefficients are treated as zeros.
robust	this argument has been deprecated, use a cluster term in the model instead. (The two options accomplish the same goal – creation of a robust variance – but the second is more flexible).
model	logical value: if TRUE, the model frame is returned in component model.
x	logical value: if TRUE, the x matrix is returned in component x.
y	logical value: if TRUE, the response vector is returned in component y.
tt	optional list of time-transform functions.
method	alternate name for the ties argument.
...	Other arguments will be passed to coxph.control

Details

The proportional hazards model is usually expressed in terms of a single survival time value for each person, with possible censoring. Andersen and Gill reformulated the same problem as a counting process; as time marches onward we observe the events for a subject, rather like watching a Geiger counter. The data for a subject is presented as multiple rows or "observations", each of which applies to an interval of observation (start, stop].

The routine internally scales and centers data to avoid overflow in the argument to the exponential function. These actions do not change the result, but lead to more numerical stability. However, arguments to offset are not scaled since there are situations where a large offset value is a purposefully used. Users should not use normally allow large numeric offset values.

IPTW outcomes; twang svyglm svycoxph

- The TSHRC implements a two stage procedure for comparing hazard functions.
- The Survgini package proposes to test the equality of two survival distributions based on the Gini index.
- The FHtest package offers several tests based on the Fleming-Harrington class for comparing surival curves with right- and interval-censored data.
- The LogrankA package provides a logrank test for which aggregated data can be used as input.
- The short term and long term hazard ratio model for two samples survival data can be found in the YPmodel package.

Regression Modelling

- ***Cox model:*** The `coxph` function in the survival package fits the Cox model. `cph` in the rms package and the eha package propose some extensions to the `coxph` function. The package coxphf implements the Firth's penalised maximum likelihood bias reduction method for the Cox model. An implementation of weighted estimation in Cox regression can be found in coxphw. The coxrobust package proposes a robust implementation of the Cox model. `timecox` in package timereg fits Cox models with possibly time-varying effects. The NestedCohort fits Cox models for covariates with missing data. A Cox model model can be fitted to data from complex survey design using the `svycoxph` function in survey. The multipleNCC package fits Cox models using a weighted partial likelihood for nested case-control studies. The intcox package implements the Cox model for interval-censored data using the ICM-algorithm. The MIICD package implements Pan's (2000) multiple imputation approach to Cox models for interval censored data. The ICsurv package fits Cox models for interval-censored data through an EM algorithm. The dynsurv package fits time-varying coefficient models for interval censored and right censored survival data using a Bayesian Cox model, a spline based Cox model or a transformation model. The CPHshape package computes the Cox proportional hazards model with shape constrained hazard functions. The OrdFacReg package implements the Cox model using an active set algorithm for dummy variables of ordered factors. The survivalMPL package fits Cox models using maximum penalised likelihood and provide a non parametric smooth estimate of the baseline hazard function. The `cumres` function in gof computes goodness-of-fit methods for the Cox proportional hazards model. The proportionality assumption can be checked using the `cox.zph` function in survival. The CPE package calculates concordance probability estimate for the Cox model, as does the `coxphCPE` function in clinfun. The `coxphQuantile` in the latter package draws a quantile curve of the survival distribution as a function of covariates. The multcomp package computes simultaneous tests and confidence intervals for the Cox model and other parametric survival models. The lsmeans package permits to obtain least-squares means (and contrasts thereof) from linear models. In particular, it provides support for the `coxph`, `survreg` and `coxme` functions. The multtest package on Bioconductor proposes a resampling based multiple hypothesis testing that can be applied to the Cox model. Testing coefficients of Cox regression models using a Wald test with a sandwich estimator of variance can be done using the saws package. The rankhazard package permits to plot visualisation of the relative importance of covariates in a proportional hazards model. The smoothHR package provides hazard ratio curves that allows for nonlinear relationship between predictor and survival. The paf package permits to compute the unadjusted/adjusted attributable fraction function from a Cox proportional hazards model. The PHeval package proposes tools to check the proportional hazards assumption using a standardised score process.
- ***Parametric Proportional Hazards Model:*** `survreg` (from survival) fits a parametric proportional hazards model. The eha and mixPHM packages implement a proportional hazards model with a parametric baseline hazard. The `pphsm` in rms translates an AFT model to a proportional hazards form. The polspline package includes the `hare` function that fits a hazard regression model, using splines to model the baseline hazard. Hazards can be, but not necessarily, proportional. The flexsurv package implements the model of Royston and Parmar (2002). The model uses natural cubic splines for the baseline survival function, and proportional hazards, proportional odds or probit functions for regression. The SurvRegCensCov package fits a Weibull regression model for right-censored data in which one covariate might be censored.
- ***Accelerated Failure Time (AFT) Models:*** The `survreg` function in package survival can fit an accelerated failure time model. A modified version of `survreg` is implemented in the rms package (`psm` function). It permits to use some of the rms functionalities. The eha package also proposes an implementation of the AFT model (function `aftreg`). An AFT model with an error distribution assumed to be a mixture of G-splines is implemented in the smoothSurv package. The NADA package proposes the front end of the `survreg` function for left-censored data. A least-square principled implementation of the AFT model can be found in the lss package. The simexaft package implements the Simulation-Extrapolation algorithm for the AFT model, that can be used when covariates are subject to measurement error. A robust version of the accelerated failure time model can be found in RobustAFT. The coarseDataTools package fits AFT models for interval censored data. The aftgee package implements both rank-based estimates and least square estimates (via generalised estimating equations) to the AFT model. An alternative weighting scheme for parameter estimation in the AFT model is proposed in the imputeYn package. The AdapEnetClass package implements elastic net regularisation for the AFT model.
- ***Additive Models:*** Both survival and timereg fit the additive hazards model of Aalen in functions `aareg` and `aalen`, respectively. timereg also proposes an implementation of the Cox-Aalen model (that can also be used to perform the Lin, Wei and Ying (1994) goodness-of-fit for Cox regression models) and the partly parametric additive risk model of McKeague and Sasieni. A version of the Cox-Aalen model for interval censored data is available in the coxinterval package.
- ***Buckley-James Models:*** The `bj` function in rms and `BJnoint` in emplik compute the Buckley-James model, though the latter does it without an intercept term. The bujar package fits the Buckley-James model with high-dimensional covariates (L2 boosting, regression trees and boosted MARS, elastic net).
- ***Other models:*** Functions like `survreg` can fit other types of models depending on the chosen distribution, *e.g.* , a tobit

The Washington Post

Dmdqfx!`mc!Dmuhqnml dms

Why living around nature could make you live longer

Ax!Bgdkrd`!G`qudx @odjk08`s0092 8@L

Living closer to nature is better for your health, new research suggests — and may even extend your life.

A study just published in the journal Environmental Health Perspectives found that people who live in "greener" areas, with more vegetation around, have a lower risk of mortality. The health benefits are likely thanks to factors such as improved mental health, social engagement and physical activity that come with living near green spaces.

The research relied on data from a vast long-term Harvard study funded by the National Institutes of Health called the Nurses' Health Study, which has collected health information biennially on more than 100,000 female registered nurses in the U.S. since 1976. The new paper analyzed participant data from between 2000 and 2008, taking note of any deaths that occurred and their causes. At the same time, the researchers used satellite data to assess the amount of green vegetation surrounding each participant's home during the study period.

The researchers found that people living in the greenest places — that is, people who had the most vegetation within 800 feet of their homes — had a 12 percent lower rate of mortality from any non-accidental cause than people living in the least green places. Specifically, they found that the relationship was strongest for deaths related to respiratory disease, cancer and kidney disease. These results were the same regardless of the participants' income, weight or smoking status and also did not significantly change between urban and suburban locations.

In statistical analyses, though, the researchers found that participants' mental health, social engagement, level of physical activity and exposure to air pollution likely explained how the green spaces were making a difference.

This is all in line with the ways previous research has suggested greenness can affect health. Places with more vegetation are generally thought to be less polluted, and the presence of vegetation, itself, can help keep air cleaner. And green spaces like parks can help encourage people to get outside, exercise and engage with other people — all factors that can improve overall health. The effects on mental health may be somewhat less straightforward, but nonetheless important, as this study suggested.

"We were really surprised to find that the mental health pathway explained about 30 percent of the relationship between greenness and mortality," said Peter James, the study's lead author and a research associate at the Harvard T.H. Chan School of Public Health.

There are several theories about how nature affects mental health, said Howard Frumkin, dean of the school of public health at the University of Washington, who was not involved with the new study. One of them is known as the "biophilia"

HEALTH APR 18 2016, 8:12 PM ET

Living Near Green Spaces Helps You Live Longer, New Study Shows

by **MAGGIE FOX**

SHARE

Go ahead, plant a tree. It might help you live longer.

A new study shows that living in or near green spaces can be a boon for longevity, and seeing more greenery may also boost mental health, according to a team at the Harvard School of Public Health.

A new study finds living in green spaces may help people live longer Maggie Fox/NBC News / NBC News

Women living in the greenest areas, as measured by satellite, were 34 percent less likely to die from a respiratory illness than women living in the most paved-over areas. And women living amidst greenery were 13 percent less likely to die of cancer.

"We were surprised to observe such strong associations between increased exposure to greenness and lower mortality rates," said Peter James, a research associate who worked on the study.

Codes E800-E999) were included as negative control outcomes to detect potential confounding bias (Lipsitch et al. 2010).

Statistical Analysis

Person-months of follow-up were accrued from the return date of the 2000 questionnaire until either death or the end of follow-up (May 31, 2008), whichever came first. We fit time-varying Cox proportional hazards models to compute hazard ratios (HRs) and 95% confidence intervals (CIs) for associations between each NDVI exposure measure and each mortality outcome. We examined the following covariates as potential confounders, effect modifiers, or mediators (all covariates time-varying unless otherwise indicated): fixed race/ethnicity (White non-Hispanic versus other), smoking status (current, past, never), pack-years of smoking, fixed individual-level socioeconomic status (SES), area-level SES, weight status (normal (body mass index (BMI) 18.5-24.9), overweight (BMI 25-29.9), obese (BMI > 30)), region, urbanicity, whether a participant had changed addresses over follow-up, physical activity, air pollution, social engagement, and mental health. Current smoking status and pack-years smoked was updated at each biennial questionnaire. To account for fixed individual SES, we included information on self-reported parental occupation for the participant's mother and father, whether the participant had a registered nursing degree, marital status, and husband's highest educational attainment (< high school, high school graduate, >high school, missing or not married) as reported in 1992. We examined area-level SES by including information on Census tract median home value and Census tract median income based on the Census tract containing the residential address at each questionnaire response and 2000 Census data. Urbanicity was determined by the participant's residence in a metropolitan (urban area ≥50,000 people), micropolitan (urban cluster of 10,000-

Environ Health Perspect DOI: 10.1289/ehp.1510363
Advance Publication: Not Copyedited

Table 2. Hazard ratios (HR) and 95% confidence intervals (CI) for greenness and non-accidental[a] all-cause mortality in the Nurses' Health Study (N=108,630, with 8,604 deaths from 2000-2008)

Exposure Metric	250m Buffer				1250m Buffer			
	Cumulative Average Greenness		Contemporaneous Greenness		Cumulative Average Greenness		Contemporaneous Greenness	
	Age-Adjusted HR (95% CI)	Fully Adjusted HR (95% CI)[b]	Age-Adjusted HR (95% CI)	Fully Adjusted HR (95% CI)[b]	Age-Adjusted HR (95% CI)	Fully Adjusted HR (95% CI)[b]	Age-Adjusted HR (95% CI)	Fully Adjusted HR (95% CI)[b]
Quintile 1	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Quintile 2	0.91 (0.85, 0.97)	0.92 (0.86, 0.98)	0.96 (0.90, 1.03)	0.96 (0.90, 1.02)	0.96 (0.90, 1.02)	0.95 (0.89, 1.01)	0.96 (0.90, 1.02)	0.95 (0.89, 1.01)
Quintile 3	0.88 (0.83, 0.94)	0.90 (0.84, 0.96)	0.93 (0.87, 0.99)	0.93 (0.87, 0.99)	0.95 (0.89, 1.01)	0.94 (0.88, 1.01)	0.95 (0.89, 1.01)	0.94 (0.88, 1.01)
Quintile 4	0.91 (0.85, 0.97)	0.94 (0.88, 1.00)	0.87 (0.82, 0.93)	0.89 (0.83, 0.95)	0.93 (0.87, 1.00)	0.94 (0.88, 1.01)	0.93 (0.87, 1.00)	0.94 (0.88, 1.01)
Quintile 5	0.83 (0.77, 0.88)	0.88 (0.82, 0.94)	0.89 (0.84, 0.96)	0.93 (0.87, 0.99)	0.86 (0.81, 0.92)	0.89 (0.83, 0.96)	0.86 (0.81, 0.92)	0.89 (0.83, 0.96)
P for Trend[c]	<0.0001	0.002	<0.0001	0.003	<0.0001	0.004	<0.0001	0.004
Continuous (per 0.1 unit)	0.83 (0.78, 0.89)	0.88 (0.82, 0.94)	0.97 (0.94, 1.00)	0.99 (0.95, 1.02)	0.84 (0.79, 0.90)	0.89 (0.83, 0.95)	0.96 (0.93, 1.00)	0.98 (0.94, 1.01)

[a] Excludes accidental ICD9 Codes E800-E999

[b] Hazard ratios are adjusted for age and calendar year, race/ethnicity, smoking status, pack-years smoked, parental occupation, RN degree, marital status, husband's highest education, Census tract median home value, and Census tract median income

[c] Based on linear test for trend using the ordinal rank for each quintile

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

2007 articles match your search: ""Journal of Research in personality" stability "

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Research in Personality 38 (2004) 319–350
JOURNAL OF RESEARCH IN PERSONALITY
www.elsevier.com/locate/jrp

Presidential Address

Stability versus change, dependability versus error: Issues in the assessment of personality over time

David Watson*

Department of Psychology, E11 Seashore Hall, University of Iowa, Iowa City, IA 52242-1407, USA

Abstract

Temporal instability can reflect either true psychological change or measurement error. I offer several recommendations to improve stability research and enhance the ability to detect error; these include the use of (a) theoretically meaningful retest intervals, (b) larger sample sizes, and (c) benchmark scales that permit comparative tests of stability. I illustrate this approach using retest data of obsessive–compulsive symptoms, dissociative tendencies, trait affectivity, and the Big Five. These data demonstrate that highly correlated measures of the same target constructs show significantly different levels of stability, even over 2-month retest intervals during which true change should be minimal. These discrepancies are not simply due to broad differences in content, but reflect more subtle differences in wording, instructions, and response formats.
© 2004 Elsevier Inc. All rights reserved.

1. Introduction

Trait psychologists have been slow to incorporate change into their models of personality development. Indeed, classical test theory was built on the assumption that basic dimensions of individual differences are extremely stable, even invariant, over time, such that any observed change is attributable to measurement error (Anastasi & Urbina, 1997; Kaplan & Saccuzzo, 2001; Murphy & Davidshofer, 2001). This assumption was incorporated into the concept of test–retest reliability, which is computed by correlating scores on the same test across two assessments separated by a specified time interval. As Murphy and Davidshofer (2001) have noted, "The

* Fax: 1-319-335-0191.
E-mail address: david-watson@uiowa.edu.

0092-6566/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jrp.2004.03.001

Questions About Stability

A. Is an individual consistent over time?

Examples: Personality characteristics/traits

Teacher behaviors

B. Are individual differences consistent over time?

Examples: Intelligence

Medical risk factors--
blood pressure, lipids
(Bogalusa, Chatham)

Assessing Consistency of Individual Differences

Time1 - Time2 Correlations

Intraclass Correlations/ G-Coefficients

Repeated Measures Anova

Path Analysis Regressions

Structural Regression With Exogenous Variables

Indices Of Tracking

Assessments of Stability

Correlation $\rho_{\xi(t_1)\xi(t_2)}$

Intraclass Correlation

"Path Regression" (Heise) $\beta_{\xi(t_2)\xi(t_1)}$

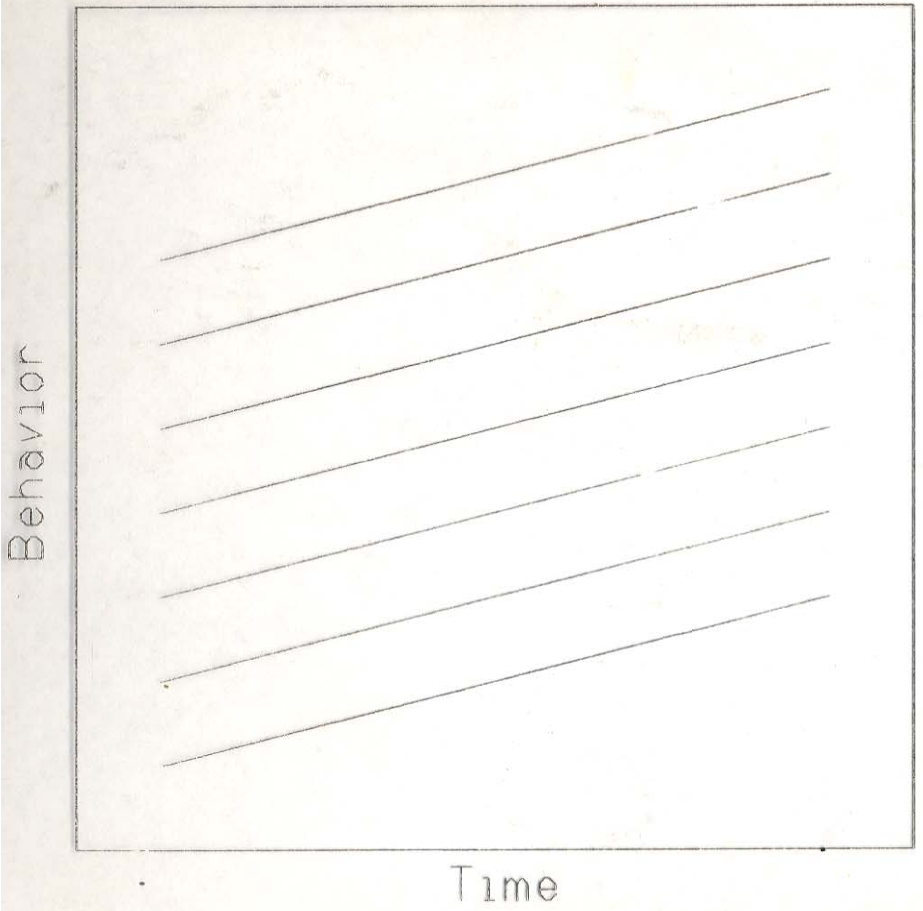
Structural Regression (Wheaton, Alwin, Summers, Muthen) $\beta_{\xi(t_2)\xi(t_1)\cdot W}$

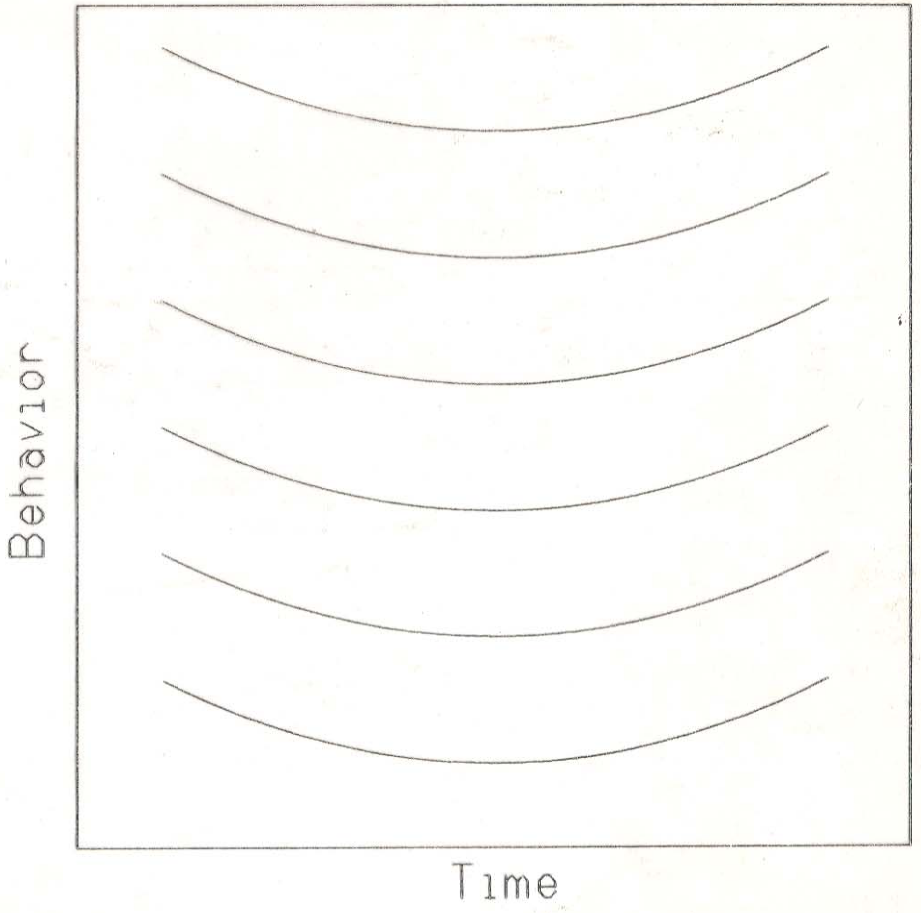
Indices of Tracking

Foulkes-Davis δ

McMahan γ

PERFECT CONSISTENCY OF INDIVIDUAL DIFFERENCES





Two illustrations of (perfect) consistency of individual differences over time

Foulkes-Davis γ (1981)

Measures consistency of rank order

$$\gamma(t_1, t_2) = \Pr\{\xi_p(t) \geq \xi_{p'}(t) \text{ or } \xi_p(t) \leq \xi_{p'}(t) \text{ for all } t \in [t_1, t_2]\}$$

Tracking exists for $\gamma > 1/2$

Estimate γ by counting crossings of fitted growth curves.

$\hat{\gamma}$ asymptotically normal
S.e. easily obtained

Alternative - McMahan (1981) γ

Comparison γ, γ in Rogosa (1983) Biometrics

Scand J Med Sci Sports 2008: 18: 171–178
Printed in Singapore . All rights reserved
DOI: 10.1111/j.1600-0838.2006.00622.x

Copyright © 2007 The Authors
Journal compilation © 2007 Blackwell Munksgaard

SCANDINAVIAN JOURNAL OF
MEDICINE & SCIENCE
IN SPORTS

Tracking of objectively measured physical activity from childhood to adolescence: The European youth heart study

P. L. Kristensen[1], N. C. Møller[1], L. Korsholm[2], N. Wedderkopp[3], L. B. Andersen[4], K. Froberg[1]

[1]Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark, [2]Department of Statistics, University of Southern Denmark, Odense, Denmark, [3]Back Research centre; Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark, [4]Department of Sports Medicine, Norwegian school of Sport Sciences, Oslo, Norway
Corresponding author: Peter Lund Kristensen, Campusvej 55, 5230 Odense M. Tel: 65503639, E-mail: plkristensen@health.sdu.dk

Accepted for publication 8 November 2006

A number of studies have investigated tracking of physical activity from childhood to adolescence and, in general, these studies have been based on methods with some degree of subjectivity (e.g., questionnaires). The aim of the present study was to evaluate tracking of physical activity from childhood to adolescence using accelerometry, taking into account major sources of variation in physical activity. Both a crude and an adjusted model was fitted, and, in the adjusted model, analyses were corrected for seasonal variation, within-week variation, activity registration during night time sleep, within instrumental measurement error, and day-to-day variation in physical activity. In all, 208 subjects were included in the crude analyses. Stability coefficients estimated from the crude model were low (i.e., 0.18 and 0.19 for boys and girls, respectively) and only borderline significant. However, in the adjusted model highly significant stability coefficients of 0.53 and 0.48 for boys and girls, respectively, were observed. It was concluded that physical activity behavior tends to track moderately from childhood to adolescence.

There is a large amount of literature providing evidence of a positive health effect of being physically active in adulthood (Wannamethee & Shaper, 2001; Batty, 2002; Fransson et al., 2003; Barengo et al., 2004; Hu et al., 2004). Physical activity is associated with a lower risk of developing serious illnesses such as cardiovascular diseases (Lakka et al., 1994), some cancers (Steindorf et al., 2005; Kruk & boul Enein, 2006), and diabetes mellitus (Meisinger et al., 2005). Based on this knowledge, preventive efforts including health promotion campaigns have been designed to increase physical activity levels among adults. However, children and adolescents have also been targeted for physical activity promotion campaigns. The rationale underlying this strategy is based on several possible pathways through which physical activity in childhood could be related to present and future health status (Twisk et al., 2002). Physical activity in childhood could bring about immediate health benefits but it could also be directly associated with adult health. Moreover, habitual physical activity patterns in childhood could track into adulthood and thus be indirectly related to adult health.

Several studies have assessed tracking of physical activity, and, in general, results have shown low to moderate stability of physical activity through childhood and adolescent years (Raitakari et al., 1994; Sallis et al., 1995; Malina, 1996; Pate et al., 1996; Telama et al., 1996; Janz et al., 2000; Fortier et al., 2001; Malina, 2001a; McMurray et al., 2003; Anderssen et al., 2005; Janz et al., 2005). However, physical activity is a very complex behavior to quantify, and most studies have made use of methods with some degree of subjectivity (e.g., questionnaires). Although there is no single method of measuring habitual physical activity in large epidemiological studies, which is valid, precise, economical and practically possible accelerometers are now often considered to be the best choice available and preferred over self-report instruments. This is especially true when children younger than 10 years are examined, as validation studies of self-report instruments performed with this age group have revealed insignificant validation coefficients, indicating that most self-report instruments do not measure what they are intended to measure among young children (Kohl et al., 2000). Thus, results from earlier tracking studies of physical activity, which have generally used self-report methods and showed low to moderate stability coefficients from childhood to adolescence/adulthood may have underestimated the real tracking effect. To our knowledge, no study has yet

A Stochastic Model for Analysis of Longitudinal AIDS Data

Journal article by J.M.G. Taylor, W.G. Cumberland, Sy J.P.; Journal of the American Statistical Association, Vol. 89, 1994

Journal Article Excerpt See below...

A Stochastic Model for Analysis of Longitudinal AIDS Data

J. M. G. TAYLOR, W. G. CUMBERLAND, and J. P. SY*

In this paper we analyze serial CD4 T-cell measurements from the Los Angeles portion of the Multicenter AIDS Cohort Study. Our emphasis is on developing a plausible and parsimonious model to describe the stochastic process underlying the patterns of CD4 measurements. The stochastic process that we use enables us to investigate the concept of derivative tracking, for which it is assumed that the rank order of the individual's slopes is maintained over time. A general model for the analysis of longitudinal repeated measures data is

$$Y_i(t_{ij}) = X(t_{ij})\alpha + Z(t_{ij})b_i + W_i(t_{ij}) + \varepsilon_{ij}$$

where $Y_i(t_{ij})$ is the measurement of subject i at time t_{ij}, $X(t_{ij})\alpha$ represents fixed effect terms, $Z(t_{ij})b_i$ represents random effect terms, $W_i(t_{ij})$ is a stochastic process allowing correlation between measurements, and ε_{ij} is measurement error. In the simplest case, $X(t_{ij})$ and $Z(t_{ij})$ contain the times of measurements. For $W_i(t_{ij})$, we use a two-parameter integrated Ornstein–Uhlenbeck (OU) process. The OU process is the continuous mean zero Gaussian Markov process, which includes Brownian motion and white noise as special limiting cases. This model is a continuous-time version of an AR(1) process for the deviations of the derivative of y from the expected derivative of y with respect to t. This approach is flexible and tractable as the covariance structure has a closed-form expression. The model allows unequally spaced observations and can be generalized to multivariate responses. This model enables one to assess whether individuals maintain their trajectories; that is, whether their slope of Y tracks. We find no evidence in the data that the slopes of the CD4 values track.

KEY WORDS: AIDS; Ornstein–Uhlenbeck process; Repeated measures; Tracking.

1. INTRODUCTION

The concept of tracking for longitudinal data is popular with epidemiologists and has been discussed in the statistical literature. Foulkes and Davis (1981) and McMahan (1981) developed indices of tracking based on the concept of the maintenance over time of the relative or rank position of the response variable among the group of subjects. Ware and Wu (1981) regarded tracking as the ability to predict future observations for each individual, leading to a definition that a population tracks if for each individual, the expected values of serial measurements are given by a polynomial function of time. These two concepts are closely related, but not identical. In this article we adopt the first concept of tracking and apply it to the first derivative of the response variable. Thus we regard a population as having derivative tracking if the relative ranks of the rate of change of the outcome variable are maintained over time.

The typical structure of longitudinal data is numerous measurements of a possibly multivariate response variable on each subject. The measurements will usually be at unequally spaced time intervals, and the number of measurements per subject may differ between subjects. There could also be covariates, possibly time varying, that influence the response variable. The aim in the analysis of such data is to understand the changes in the mean structure of the response variable with time, to understand the effect of the covariates on the response variable, and to understand the within-subject correlation structure.

There are a number of approaches to the analysis of such data. One uses a mixed or random effects model (Laird and Ware 1982). Let $Y_i(t_{ij})$ be the response of person i at time t_{ij}; then in a random effects model, it is assumed that

$$Y_i(t_{ij}) = X(t_{ij})\alpha + X(t_{ij})b_i + \varepsilon_{ij},$$

where $X(t_{ij})\alpha$ represents fixed effects, $Z(t_{ij})b_i$ represents random effects, and ε_{ij} is an independent, identically distributed (iid) measurement error.

A second approach is to specify a structure for the within-subject covariance based possibly on a stochastic process. For example, it might be assumed that $Y_i(t_{ij}) = X(t_{ij})\alpha + W_i(t_{ij})$, where $W_i(t_{ij})$ is a realization of a stochastic process.

Other authors have suggested a combination of these two approaches (Chi and Reinsel 1989; Cullis and McGilchrist 1990; Diggle 1988; Jones and Bodi-Boateng 1991). The computational aspects of fitting these models to data are nontrivial and have been discussed by Jennrich and Schluchter (1986) and Lindstrom and Bates (1988).

In this article we focus on a model consisting of a combination of fixed effects, random effects, a specific stochastic process, and measurement error. In particular, we assume that

$$Y_i(t_{ij}) = X(t_{ij})\alpha + Z(t_{ij})b_i + W_i(t_{ij}) + \varepsilon_{ij}.$$

Our emphasis will be on interpreting the stochastic process with particular regard to the concept of derivative tracking. We propose using an integrated Ornstein–Uhlenbeck (IOU) process for the stochastic process. The motivation for this choice of stochastic process arose from an AIDS study of the natural history of CD4 T-cell numbers. CD4 numbers are a critical aspect of the immune system, with low values indicating more severe immune deficiency. An immunologic theory is that the rate of change of CD4 cell numbers "tracks"; that is, an immunologically weak individual who has an initial fast rate of decline of CD4 relative to other HIV-infected people will persist with a more rapid rate of

* J. M. G. Taylor is Associate Professor, W. G. Cumberland is Professor, and J. P. Sy is a graduate student, Department of Biostatistics, University of California, Los Angeles, CA 90024-1772. This work was partially supported by National Institutes of Health Grants AI29196 and AI72631.

© 1994 American Statistical Association

A Stochastic Model for Analysis of Longitudinal AIDS Data.

by J.M.G. Taylor , W.G. Cumberland , Sy J.P.

1. INTRODUCTION

The concept of tracking for longitudinal data is popular with epidemiologists and has been discussed in the statistical literature. Foulkes and Davis (1981) and McMahan (1981) developed indices of tracking based on the concept of the maintenance over time of the relative or rank position of the response variable among the group of subjects. Ware and Wu (1981) regarded tracking as the ability to predict future observations for each individual, leading to a definition that a population tracks if for each individual, the expected values of serial measurements are given by a polynomial function of time. These two concepts are closely related, but not identical. In this article we adopt the first concept of tracking and apply it to the first derivative of the response variable. Thus we regard a population as having derivative tracking if the relative ranks of the rate of change of the outcome variable are maintained over time.

The typical structure of longitudinal data is numerous measurements of a possibly multivariate response variable on each subject. The measurements will usually be at unequally spaced time intervals, and the number of measurements per subject may differ between subjects. There could also be covariates, possibly time varying, that influence the response variable. The aim in the analysis of such data is to understand the changes in the mean structure of the response variable with time, to understand the effect of the covariates on the response variable, and to understand the within-subject correlation structure.

There are a number of approaches to the analysis of such data. One uses a mixed or random effects model (Laird and Ware 1982). Let [Y.sub.i] ([t.sub.ij]) be the response of person i at time [t.sub.ij]; then in a random effects model, it is assumed that

[Y.sub.i]([t.sub.ij]) = X([t.sub.ij]][alpha] + X([t.sub.ij]}[b.sub.i] + [[epsilon].sub.ij] where X([t.sub.ij)[alpha] represents fixed effects, Z([t.sub.ij])[b.sub.i] represents random effects, and [[epsilon].sub.ij] is an independent, identically distributed (iid) measurement error.

A second approach is to specify a structure for the within-subject covariance based possibly on a stochastic process. For example, it might be assumed that [Y.sub.i] ([t.sub.ij]) = X([t.sub.ij])[alpha] [W.sub.i] ([t.sub.ij]), where [W.sub.i] ([t.sub.ij]) is a realization of a stochastic process.

Other authors have suggested a combination of these two approaches (Chi and Reinsel 1989; Cullis and McGilchrist 1990; Diggle 1988; Jones and Bodi-Boateng 1991). The computational aspects of fitting these models to data are nontrivial and have been discussed by Jennrich and Schluchter (1986) and Lindstrom and Bates (1988).

In this article we focus on a model consisting of a combination of fixed effects, random effects, a specific stochastic process, and measurement error. In particular, we assume that

[Y.sub.i]([t.sub.ij]) = X([t.sub.ij])[alpha] + Z([t.sub.ij])[b.sub.i] + [W.sub.i] ([t.sub.ij]) + [[epsilon].sub.ij].

Our emphasis will be on interpreting the stochastic process with particular regard to the concept of derivative tracking. We propose using an integrated Ornstein-Uhlenbeck (IOU) process for the stochastic process. The motivation for this choice of stochastic process arose from an AIDS study of the natural history of CD4 T-cell numbers. CD4 numbers are a critical aspect of the immune system, with low values indicating more

TRACKING

Investigators engaged in longitudinal studies of human populations have often found that individual patterns of change over time can be described in relatively simple ways that improve understanding of development and aging and facilitate the prediction of future values of biological variables (*see* LONGITUDINAL DATA ANALYSIS). Some investigators have emphasized the high correlation* between successive measurements of the same individuals [8], others have noted the tendency for individuals to maintain relative or percentile rank over time [5,8], and still others have noted that individual patterns of growth or aging can be described by simple functions of time [9]. These related phenomena have been described as *tracking*, often without benefit of a precise definition for the term. Recently, several papers in the statistical literature have provided explicit definitions and described statistical methods for investigating tracking in longitudinal data sets. We discuss three different approaches.

One definition is based on the notion that individuals remain at a constant percentile of the population distribution over time.

Definition 1. (McMahan). A population is said to track with respect to an observable characteristic if, for each individual, the expected value of the deviation from the population mean remains fixed over time [7].

To develop a quantitative analysis of tracking, McMahan assumes that the data of interest arise from a polynomial growth curve* model. Since other definitions of tracking are also related to growth curve analysis, we begin with a brief description of the growth curve model for a balanced design. For a more general formulation, see ref. 6.

Let $\mathbf{y}_i, i = 1, 2, \ldots, N$, be a $p \times 1$ vector of observations of the biological variable of interest at p different time points and $\mathbf{X}$ be a $p \times k$ within-individual design matrix*. The two-stage growth curve model is defined by
Stage I:

$$\mathbf{y}_i|\boldsymbol{\beta}_i \sim N(\mathbf{X}\boldsymbol{\beta}_i, \sigma^2\mathbf{I}),$$

where $N(\cdot,\cdot)$ denotes the multinormal distribution* and $\boldsymbol{\beta}_i$ is a $k \times 1$ vector of individual growth curve parameters;
Stage II:

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\beta}, \boldsymbol{\Sigma}).$$

The two-stage model implies that $\mathbf{y}_i$ has the marginal distribution

$$\mathbf{y}_i \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{X}\boldsymbol{\Sigma}\mathbf{X}' + \sigma^2\mathbf{I}).$$

The total variance of $\mathbf{y}$ consists of between-individual variation $\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}'$ and error variance $\sigma^2\mathbf{I}$.

Returning to the issue of tracking, Definition 1 implies that

$$E[y_{ij}|\boldsymbol{\beta}_i] = \mu_j + \kappa_i\delta_j, \quad i = 1, \ldots, N, j = 1, \ldots, p,$$

where μ_j and δ_j are the population mean and interindividual standard deviation respectively, at time j; κ_i is constant for the ith individual and $E(\kappa_i) = 0, \mathrm{var}(\kappa_i) = 1$ in the population. This can be expressed in matrix form as

$$\mathbf{X}\boldsymbol{\beta}_i = \mathbf{X}\boldsymbol{\beta} + \kappa_i\boldsymbol{\delta},$$

where $\boldsymbol{\delta}' = (\delta_1, \ldots, \delta_p)$ and $\delta_i^2 = (\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}')_{ii}$. Hence, tracking is equivalent to a restriction on the interindividual covariance matrix, namely,

$$\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}' = \boldsymbol{\delta}\boldsymbol{\delta}'.$$

This definition leads to a quantitative index of tracking for an arbitrary $\mathbf{X}$ and $\boldsymbol{\Sigma}$. One index,

$$\zeta = \boldsymbol{\delta}'\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}'\boldsymbol{\delta}/(\boldsymbol{\delta}'\boldsymbol{\delta})^2,$$

represents the fraction of the total interindividual variation, $\mathrm{tr}(\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}')$, attributable to tracking. Since ζ will be positive even when $\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}'$ is diagonal (implying mutually independent deviations and usually regarded as representing no tracking), McMahan defined a corrected index

$$\tau = \frac{\boldsymbol{\delta}'(\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}' - \mathbf{D}_\delta^2)\boldsymbol{\delta}}{\boldsymbol{\delta}'(\boldsymbol{\delta}\boldsymbol{\delta}' - \mathbf{D}_\delta^2)\boldsymbol{\delta}}$$

Encyclopedia of Statistical Sciences, Copyright © 2006 John Wiley & Sons, Inc.

where $\mathbf{D}_\delta$ is a diagonal matrix containing the elements of $\boldsymbol{\delta}$.

McMahan also defined tracking indices for deviations standardized by the interindividual component of population variance. These definitions clarify the interpretation of the two indices. If the tracking indices are defined in terms of $\mathbf{R} = \mathbf{D}_\delta^{-1}\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}'\mathbf{D}_\delta^{-1}$, the correlation matrix* of the expected deviations, and if $\mathbf{1}$ denotes a vector of 1's, then

$$\begin{aligned}\zeta_\rho &= \mathbf{1}'\mathbf{R}\mathbf{1}/(\mathbf{1}'\mathbf{1})^2 = \mathbf{1}'\mathbf{R}\mathbf{1}/p^2,\\ \tau_\rho &= [\mathbf{1}'(\mathbf{R}-\mathbf{I})\mathbf{1}]/[\mathbf{1}'(\mathbf{1}\mathbf{1}'-\mathbf{I})\mathbf{1}]\\ &= [\mathbf{1}'(\mathbf{R}-\mathbf{I})\mathbf{1}]/[p(p-1)].\end{aligned}$$

Then ζ_ρ is the average of the elements of $\mathbf{R}$ and τ_ρ is the average of the off-diagonal elements, the interoccasion correlation coefficients. McMahan estimated $\boldsymbol{\Sigma}$ by maximum likelihood* (Anderson [1]). Substitution of functions of $\hat{\boldsymbol{\Sigma}}$ into the various expressions for tracking coefficients gives their estimates. McMahan also proposed use of the Jackknife* to compute confidence intervals for each index. Although he considered only balanced data, the definitions extend directly to incomplete data sets and $\boldsymbol{\Sigma}$ can be estimated by either the maximum likelihood or the restricted maximum likelihood criterion, as described by Laird and Ware [6]. For unbalanced data sets, the concept remains clear but the indices are not well defined, since the design matrix $\mathbf{X}$ varies among individuals.

McMahan's definition of tracking is related to principal components* or factor analysis*. There is, however, a fundamental difference in that the vector $\boldsymbol{\delta}$ is fixed rather than computed from the spectral decomposition of $\mathbf{X}\boldsymbol{\Sigma}\mathbf{X}'$ as principal components analysis would require. For example, the one-factor model can be written as

$$\mathbf{y}_i = \boldsymbol{\mu} + \mathbf{a}f_i + \boldsymbol{\epsilon}_i,$$

where $\boldsymbol{\mu}$ is the vector of population means of the p occasions, $\mathbf{a}$ is a $p \times 1$ vector of unobservable factor loadings, f_i is a scalar factor with a standard normal distribution in the population, and $\boldsymbol{\epsilon}_i$ is the vector of errors or singularities. This formulation generalizes directly to several factors, but there is no natural definition of a tracking index corresponding to the factor analytic representation.

Often, individual patterns of change over time are regular but not as simple as Definition 1 would imply. For example, individuals may have different rates of development or aging. Hence, Ware and Wu [9] take the view that observations arising from a more complex growth curve model also exhibit a form of tracking. For example, if

$$\mathbf{X} = \begin{pmatrix} 1,\ldots,1 \\ t_1,\ldots,t_p \end{pmatrix},$$

the expected value of each individual's deviations from the population mean varies linearly with time.

This view of tracking suggests:

Definition 2. A population is said to track with respect to an observable characteristic if, for each individual, the expected values of the serial measurements are given by a polynomial function of time.

This definition suggests a more general index of tracking based on the proportion of total variation explained by a polynomial growth curve of lower order than that required to achieve good fit to the variance-covariance matrix. However, the model implied by Definition 1 is not a growth curve model unless interindividual variation is constant over time.

One motivation for interest in tracking is the desire to predict future values of known risk factors for disease. For example, since high blood pressure is a risk factor for cardiovascular disease in adult life, investigators have sought to determine whether children whose blood pressures are high for their age become adults with high blood pressure level. Ware and Wu [9] noted that, from standard multivariate normal* theory, if

$$\mathbf{y}_i = \begin{pmatrix} \mathbf{y}_{1i} \\ \mathbf{y}_{2i} \end{pmatrix}_1^{p_{i-1}},$$

$$\mathbf{X}_i = \begin{pmatrix} \mathbf{X}_{1i} \\ \mathbf{X}_{2i} \end{pmatrix}_1^{p_{i-1}},$$

and the observations arise from a two-stage growth curve model, then

$$E[y_{2i} - \mathbf{X}_{2i}\boldsymbol{\alpha}] = (\mathbf{X}_{2i}\boldsymbol{\Sigma}\mathbf{X}'_{1i})'(\mathbf{X}_{1i}\boldsymbol{\Sigma}\mathbf{X}'_{1i} + \sigma^2\mathbf{I})^{-1} \times(\mathbf{y}_{1i} - \mathbf{X}_{1i}\boldsymbol{\alpha})$$

with prediction variance

$$\sigma^2 + \mathbf{X}_{2i}\boldsymbol{\Sigma}\mathbf{X}'_{2i} - (\mathbf{X}_{2i}\boldsymbol{\Sigma}\mathbf{X}'_{1i}) \times (\mathbf{X}_{1i}\boldsymbol{\Sigma}\mathbf{X}'_{1i} + \sigma^2\mathbf{I})^{-1}(\mathbf{X}_{1i}\boldsymbol{\Sigma}\mathbf{X}'_{2i}).$$

For example, if, as in the intraclass correlation* model, $\mathbf{X}$ is $p_i \times 1$ and a column of 1's, the expected deviation at occasion p_i is the average of previous deviations corrected for regression to the mean*.

Foulkes and Davis [5] introduced a third definition, namely:

Definition 3. A population tracks with respect to an observable characteristic if, for each pair of individuals, the expected values of that characteristic remain in the same rank order over time.

For a time interval $[T_1, T_2]$, they defined an index of tracking

$$\gamma(T_1, T_2) = \Pr\{f(t, \boldsymbol{\beta}_1) \geqslant f(t, \boldsymbol{\beta}_2) \text{ or } f(t, \boldsymbol{\beta}_2) \geqslant f(t, \boldsymbol{\beta}_1) \text{ for all } t \in [T_1, T_2]\},$$

where $f(t, \boldsymbol{\beta}_i)[f(t, \boldsymbol{\beta}_j)]$ is the growth curve of expected values of the ith (jth) randomly chosen individual.

Foulkes and Davis recommend estimation of $\boldsymbol{\beta}_i$ by standard methods and estimate $\gamma(T_1, T_2)$ by

$$\hat{\gamma}(T_1, T_2) = \binom{n}{2}^{-1} \sum_{1<j} \phi(\mathbf{y}_i, \mathbf{y}_j, T_1, T_2),$$

$$\phi(y_i, y_j, T_1, T_2) = \begin{cases} 1, & \text{if } f(t, \hat{\boldsymbol{\beta}}_i) \geqslant f(t, \hat{\boldsymbol{\beta}}_j) \\ & \text{or } f(t, \hat{\boldsymbol{\beta}}_i) \leqslant f(t, \hat{\boldsymbol{\beta}}_j) \\ & \text{for all } t \in [T_1, T_2], \\ 0, & \text{otherwise.} \end{cases}$$

They appeal to the theory of U statistics* to establish asymptotic normality* and estimate the variance of $\hat{\gamma}(T_1, T_2)$. They also emphasize that $\gamma(T_1, T_2)$ depends heavily on the choice of the interval $[T_1, T_2]$.

These three definitions represent three closely related attempts to characterize patterns of growth and aging in populations. McMahan's definition represents a strong and intuitively appealing form of tracking. When it is satisfied, both description and prediction are straightforward. Even when it is not satisfied, however, relatively simple models for serial observations can be useful. For example, Fletcher et al. [4] have shown that individuals lose pulmonary capacity at different rates and that these rates remain constant over time in the absence of changes in smoking habits or health status. Although serial measurements of forced vital capacity do not track in the sense proposed by McMahan, the relatively simple patterns of temporal change facilitate prediction.

REFERENCES

1. Anderson, T. W. (1958). *Multivariate Statistical Analysis*. Wiley, New York.
2. Berenson, G. S., Foster, T. A., Frank, G. C., Freuchs, R. R., Srinivasan, S. R., Voors, A. W., and Webber, L. S. (1978). *Circulation*, **57**, 603–612.
3. Clarke, W. R., Schrott, H. G., Leaverton, P. E., Connor, W. E., and Lauer, R. M. (1978). *Circulation*, **58**, 626–634.
4. Fletcher, C., Peto, R., Tinker, C., and Speizer, F. E. (1976). *The Natural History of Chronic Bronchitis and Emphysema*. Oxford University Press, New York.
5. Foulkes, M. A. and Davis, C. E. (1981). *Biometrics*, **37**, 439–446.
6. Laird, N. M. and Ware, J. H. (1982). *Biometrics*, **38**, 963–974.
7. McMahan, C. A. (1981). *Biometrics*, **37**, 447–455.
8. Rosner, B., Hennekens, C. H., Kass, E. H., and Miall, W. E. (1977). *Amer. J. Epidemiol.*, **106**, 306–313.
9. Ware, J. H. and Wu, M. C. (1981). *Biometrics*, **37**, 427–438.

See also BIOSTATISTICS, CLASSICAL; EPIDEMIOLOGICAL STATISTICS—I; GROWTH CURVES; and LONGITUDINAL DATA ANALYSIS.

J. H. WARE

CORRESPONDENCE

Comparing Two Indices of Tracking

From: David R. Rogosa and John B. Willett
School of Education,
Stanford University,
Stanford, California 94305, *U.S.A.*

To the Editor of Biometrics:

In preparing computer programs (using SAS PROC MATRIX) to calculate values of the indices of tracking developed by Foulkes and Davis (*Biometrics* **37,** 439–446, September 1981) and McMahan (*Biometrics* **37,** 447–455, September 1981), we computed values of the Foulkes–Davis γ and the McMahan τ both for the data on serum cholesterol concentration in juvenile baboons used by McMahan and for the data on coronary uric potassium in dogs used by Foulkes and Davis. The results of our comparison of these two indices were somewhat surprising and are presented in the following table:

Index of tracking	Baboon data	Dog data			
		Group 1	Group 2	Group 3	Group 4
$\hat{\gamma}$	0.316	**0.444**	**0.711**	**0.500**	**0.444**
$\hat{\tau}$	**0.820**	0.672	0.996	0.750	0.735

Entries in bold face are the values calculated by Foulkes and Davis and by McMahan. Values of $\hat{\tau}$ are precisely the 'unjackknifed estimate after correction for within-subject error' (McMahan, p. 454). Standard errors for $\hat{\gamma}$ (computed using the formulae from Foulkes and Davis, p. 441) were found to be 0.041, 0.059, 0.060, 0.094 and 0.062, respectively, for the five columns of this table.

As no tracking is said to occur if $\gamma < 0.5$ (see Foulkes and Davis, p. 440), and the index τ is a slight modification of the proportion of variance accounted for by tracking (see McMahan, p. 450), the disagreement between the two indices for the baboon data is striking. That is, the Foulkes–Davis index does not indicate tracking, whereas the McMahan index does. Although the two indices are based on different (but not contradictory) definitions of tracking, and thus exact numerical agreement is not anticipated, it might be expected that they would agree as to whether tracking exists. Less disagreement between the two indices is seen for the four groups of dog data, although only for Group 2 does there appear to be strong agreement between the two indices.

The authors replied as follows:

Rogosa and Willett find the lack of numerical agreement striking between the indices of tracking devleoped by ourselves and McMahan. Although both indices are based on a definition of tracking as the maintenance of relative rank over time, the indices actually estimate the probability of maintenance of relative rank over time, and the proportion of total variation accounted for by tracking less that proportion explained by a diagonal covariance matrix, respectively. These will not necessarily be in agreement, except in the case of perfect tracking, when $\gamma = \tau = 1$.

SpringerLink

Content Types Subject Collections

English Go

Journal Article

Institutional Login

Recognized as:

Stanford University Libraries (379-06-324)

NorthEast Research Libraries NERL 2008 (417-39-326)

Stanford Site License 2006 (678-49-492)

Welcome!

To use the personalized features of this site, please log in or register.

If you have forgotten your username or password, we can help.

My Menu

Marked Items

Alerts

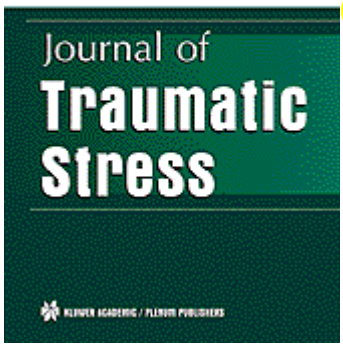
Order History

Saved Items

All

Favorites

Reliability of Reports of Violent Victimization and Posttraumatic Stress Disorder Among Men and Women With Serious Mental Illness



Journal	Journal of Traumatic Stress
Publisher	Springer Netherlands
ISSN	0894-9867 (Print) 1573-6598 (Online)
Issue	Volume 12, Number 4 / October, 1999
DOI	10.1023/A:1024708916143
Pages	587-599
Subject Collection	Behavioral Science
SpringerLink Date	Tuesday, November 30, 2004

find it @ Stanford

Add to marked items
Add to saved items
Permissions & Reprints
Recommend this article

PDF (6.3 KB)

Find more options

Go

Within all content
Within this journal
Within this issue

Export this article

Export this article as RIS | Text

Lisa A. Goodman[1], Kim M. Thompson[1], Kevin Weinfurt[2], Susan Corl[3], Pat Acker[3], Kim T. Mueser[4] and Stanley D. Rosenberg[4]

(1) Counseling Psychology Program, School of Education, Boston College, USA

(2) Department of Psychiatry, Georgetown University, USA

(3) Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD

(4) Dartmouth Medical School and New Hampshire-Dartmouth Psychiatric Research Center, Hanover, NH

Abstract *Although violent victimization is highly prevalent among men and women with serious mental illness (SMI; e.g., schizophrenia, bipolar disorder), future research in this area may be impeded by controversy concerning the ability of individuals with SMI to report traumatic events reliably. This article presents the results of a study exploring the temporal consistency of reports of childhood sexual abuse, adult sexual abuse, and adult physical abuse, as well as current symptoms of posttraumatic stress disorder (PTSD) among 50 people with SMI. Results show that trauma history and PTSD assessments can, for the most part, yield reliable information essential to further research in this area. The study also demonstrates the importance of using a variety of statistical methods to assess the reliability of self-reports of trauma history.*

reliability - measures - trauma exposure - serious mental illness - schizophrenia

---

Lisa A. Goodman
Email: goodmalc@bc.edu

Click here to visit our sponsor

Factors Associated With Tracking of BMI: A Meta-Regression Analysis on BMI Tracking*

Otmar Bayer[1], Helia Krüger[1], Rüdiger von Kries[1] and André M. Toschke[2]

Stable tracking of body composition is a prerequisite for the long-term effect of preventive measures against obesity and its harmful effects. As BMI tracking estimates reported by individual studies vary considerably, we performed a meta-regression analysis to provide a summary estimate and to assess determinants of BMI tracking. Using the Medline and EMBASE databases, a systematic review was conducted to identify publications reporting correlation coefficients as tracking estimates between BMI at baseline and follow-up measurements and the time interval between these measurements. Additional information recorded included age at baseline measurement, gender, and origin of the studied population. Based on the extracted data, a meta-regression analysis was performed using mixed effects models to account for multiple measurements of the same cohorts. Data on 55,072 individuals (797,094 person-years) extracted from 48 publications with follow-up times between 0.5 and 44 years entered the analysis. The overall estimates for the 1-year tracking correlation coefficient were strong (r = 0.78–0.86 depending on age at baseline measurement) and gradually decreasing over time (0.67–0.78 after 10 years, and 0.27–0.47 after 30 years). Study origin classified by continent was another significant predictor of BMI tracking whereas gender was not. In conclusion, this meta-regression analysis showed a high degree of BMI tracking across all age groups investigated and independent of BMI. Successful prevention in weight control is likely to have long-term effects at any age, thereby being beneficial with respect to the associated risks of over- and underweight.

Obesity (2011) **19,** 1069–1076. doi:10.1038/oby.2010.250

INTRODUCTION

The major burden of disease has shifted from communicable to noncommunicable diseases in high-income countries during the past century (1). Overweight is a major risk factor of major noncommunicable diseases such as cardiovascular disease (2,3). The increase of overweight and obesity in children and adolescents is believed to herald a further obesity epidemic in adulthood in coming years (4). The essential assumption underlying these projections is a high tracking of body composition. Tracking is important as allowing an early identification of individuals at high risk for later disease.

Despite being an indirect estimate of body composition, BMI offers a reliable and frequently used measurement in large cohort studies. However, the distribution of BMI is not comparable between children and adults, not even between the various childhood age groups (5), hampering measurement of BMI tracking from childhood to adulthood. Therefore, tracking estimates such as the correlation coefficient are preferable compared to difference estimates when assessing BMI tracking. Correlation coefficients are often used (6–8) and offer a quantitative surrogate of individual courses. In principle, tracking correlation coefficients can be calculated whenever two or more measurements of the same variable in the same individuals are available.

A number of studies on BMI tracking have been published, however, with tracking estimates ranging from no correlation (r = −0.18; ref. 9) to a strong correlation (r = 0.98; ref. 10). The consideration of heterogeneous study samples with respect to age at first measurement or observation time until follow-up measurements might explain these differences. An overall tracking estimate adjusted for the different baseline ages or follow-up times is lacking but is required for the estimation of potential long-term successes of prevention programs. Unlike conventional meta-analysis, a meta-regression takes care of heterogeneity in study characteristics (such as different follow-up times, ages, etc.) and additionally provides information on the impact of such covariates on the outcome of interest, i.e. BMI tracking.

The aim of this study was to summarize the findings of published studies on BMI tracking in a pooled and adjusted tracking estimate and to explore further determinants of BMI tracking.

[1]Ludwig-Maximilians University of Munich, Institute for Social Paediatrics and Adolescent Medicine, Department of Epidemiology, Munich, Germany; [2]Ludwig-Maximilians University of Munich, Department of Medical Informatics, Biometry and Epidemiology (IBE), Munich, Germany.
Correspondence: Otmar Bayer (Otmar.Bayer@lrz.uni-muenchen.de)
*This article was originally published as an Intervention and Prevention.

Received 10 May 2010; accepted 7 September 2010; published online 14 October 2010. doi:10.1038/oby.2010.250

RBCDH Revista Brasileira de CINEANTROPOMETRIA e Desempenho Humano

DOI: http://dx.doi.org/10.5007/1980-0037.2015v17n3p337

original article

Tracking and its applicability to Physical Education and Sport

A noção de tracking *e sua aplicação à Educação Física e ao Esporte*

Michele Caroline de Souza[1]
Cláudia Lúcia de Moraes Forjaz[2]
Joey Eisenmann[3]
José António Ribeiro Maia[1]

Abstract – Tracking refers to the idea of maintaining a relative position within a given group of individuals as they change in time. This paper presents several approaches to study and analyze tracking (i.e., stability and predictability) and its application in physical education and sport. We will use data from a mixed-longitudinal study conducted in the city of Porto, Portugal, comprising 486 girls that were divided into two age cohorts: 12-14 years and 14-16 years. Body mass index (BMI) was the chosen variable in all statistical analyses of tracking. Statistical techniques to describe tracking included: autocorrelations, Foulkes & Davis gamma and Goldstein constancy index. Regardless of statistical procedure used, tracking BMI was moderate to high in each cohort, which could be due to the short follow-up period. However, each tracking statistics showed different aspects of inter-individual differences in intra-individual changes of girls' BMI. The use of any of the suggested procedures to study aspects of stability and predictability (i.e., tracking) in longitudinal studies requires a careful scrutiny of main goals and hypotheses to be tested.

Key words: Body mass index; Growth and development; Longitudinal studies; Monitoring; Tracking.

Resumo – *O termo Tracking refere-se à noção de manutenção de posição relativa de valores de um dado grupo de sujeitos em função do tempo. O presente artigo apresenta diversas técnicas de estudo e análise do tracking (i.e., estabilidade e previsibilidade). Os dados utilizados provêm de um estudo longitudinal-misto da Região do Grande Porto, Portugal, compreendendo 486 meninas, divididas em duas coortes que abrangem as faixas etárias dos 12 aos 14 e dos 14 aos 16 anos. A variável eleita para as análises foi o índice de massa corporal (IMC). Os procedimentos estatísticos utilizados para descrever o tracking foram: autocorrelações, gama de Foulkes & Davis e índice de constância de Goldstein. Independentemente da estatística utilizada e face à curta duração do estudo, o tracking do IMC foi moderado a elevado em cada coorte. Contudo, cada procedimento de análise mostrou aspetos distintos das diferenças interindividuais nas mudanças intraindividuais do IMC das meninas. O uso parcimonioso de qualquer um dos procedimentos sugeridos para estudar aspetos da estabilidade e previsibilidade (i.e., do tracking) em estudos longitudinais exige o estabelecimento muito criterioso dos objetivos e hipóteses a serem testados.*

Palavras-chave: *Crescimento e desenvolvimento; Estudos longitudinais; Índice de massa corporal; Monitoramento.*

1. Universidade do Porto. Faculdade de Desporto. Laboratório de Cineantropometria e Estatística Aplicada, CIFI2D. Porto, Portugal.

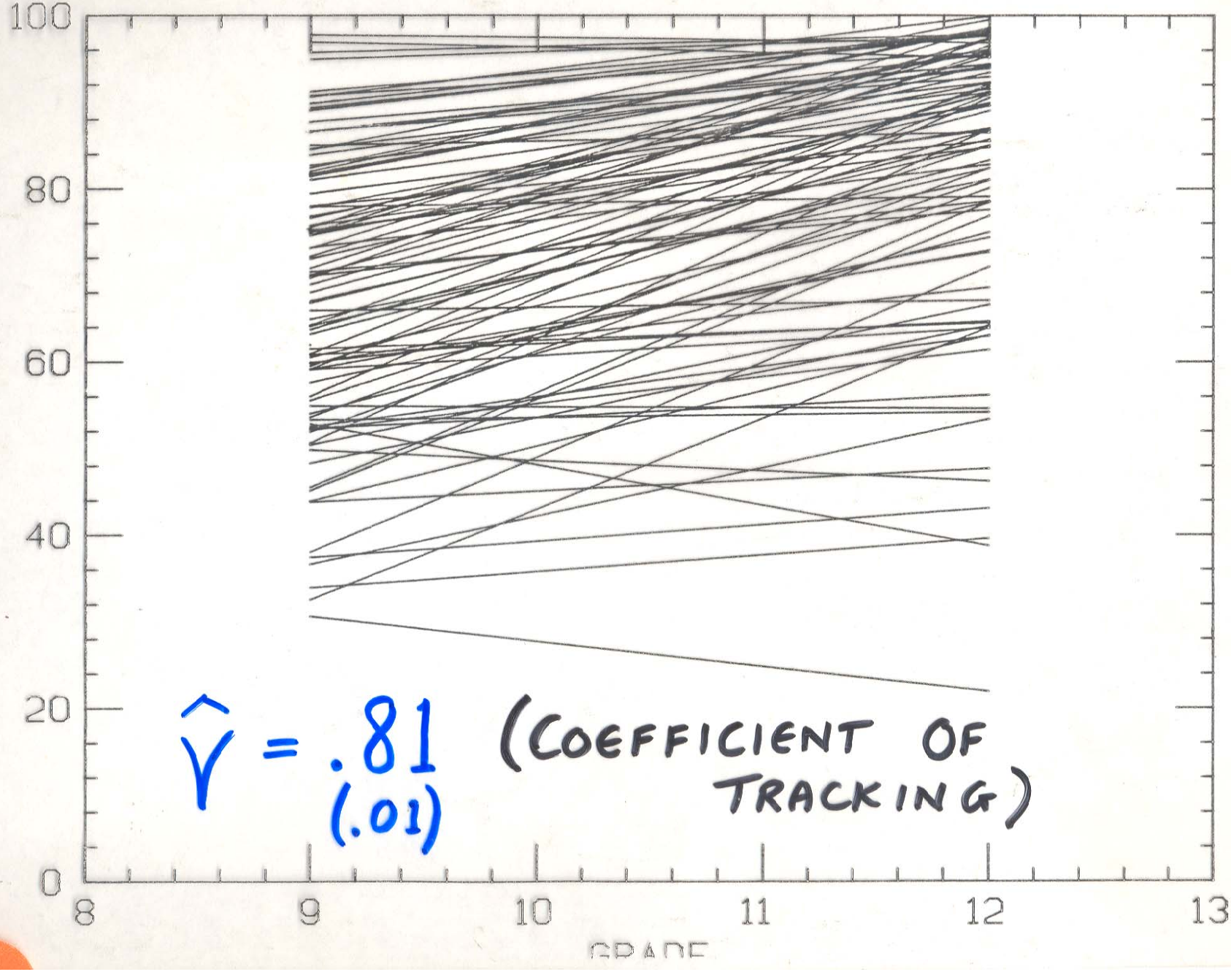
2. Universidade de São Paulo. Escola de Educação Física e Esporte. Laboratório de Hemodinâmica da Atividade Motora. São Paulo, SP, Brasil.

3. Michigan State University. Departamento de Radiologia. Divisão de Esporte e Nutrição Cardiovascular. East Lansing, Michigan. Estados Unidos.

Received: 17 March 2015
Accepted: 13 April 2015

Licence
Creative Commom

MATHEMATICS TØTAL RAW SCØRE--MALES



$\hat{\gamma} = .81$ (.01) (COEFFICIENT OF TRACKING)

ASSESSMENTS OF STABILITY.

Questions about temporal stability fall into two broad headings--
Is an individual consistent over time? and *Are individual differences consistent over time?* (Rogosa, Willett, and Floden 1984) .

index of tracking γ proposed by Foulkes and Davis (1981):
assess consistency of individual differences over a specified time interval.
the index estimates the probability that two randomly chosen individual's trajectories do not cross in the time interval specified.
"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.

odd?? Bootstrap s.e. typically is almost exactly twice as large as F-D.

RAMUS DATA

est.	mean.boot	se.boot	s.e. F-D
0.8263	0.7858	0.0655	.0320

$CI:	0.025	0.05	0.95	0.975
Standard	0.6978	0.7185	0.9341	0.9547
Percentile	0.6421	0.6684	0.8842	0.8947
BCa	0.7210	0.7473	0.9158	0.9263

WISC DATA

est.	mean.boot	se.boot	s.e. F-D
0.6700	0.6601	0.0367	.0185

NC Fem DATA

est.	mean.boot	se.boot	s.e. F-D
0.7212	0.7185	0.0174	.0087

RAT DATA (n=10)

est.	mean.boot	se.boot	s.e. F-D
0.6222	0.5248	0.100	.0529

ARTIFICIAL DATA (n=200)

est.	mean.boot	se.boot	s.e. F-D
0.6930	0.679	0.0445	.0223

Int J Biomed Comput, 32 (1993) 35–43
Elsevier Scientific Publishers Ireland Ltd.

A GAUSS PROGRAM FOR COMPUTING THE FOULKES-DAVIS TRACKING INDEX FOR POLYNOMIAL GROWTH CURVES

EMET D. SCHNEIDERMAN[a], STEPHEN M. WILLIS[a], CHARLES J. KOWALSKI[b] and THOMAS R. TEN HAVE[c]

[a]*Department of Oral and Maxillofacial Surgery, Baylor College of Dentistry, 3302 Gaston Avenue, Dallas, TX 75246,* [b]*Department of Oral Biology and* [c]*Department of Biostatistics, The University of Michigan, Ann Arbor, MI 48109 (USA)*

(Received June 1st, 1992)
(Accepted July 6th, 1992)

We have previously published a GAUSS program for computing the Foulkes-Davis tracking index, γ, from a one-sample longitudinal data set when no assumptions were made concerning the structure of the individual growth curves (Schneiderman et al., *Am J Hum Biol*, 4 (1992) 417–420). In this paper we consider the computation of the Foulkes-Davis index assuming that each individual growth curve may be adequately represented by a polynomial function in time and a GAUSS program performing these computations is made available. As with the two other tracking indices we have described, γ and κ (Schneiderman et al., *Am J Hum Biol*, 2 (1990) 475–490), this one can be used to evaluate regularity in patterns of growth or adaptation. An example is presented where statural growth in the same three groups considered in the earlier papers are analyzed. The small disparities between these and the earlier results are discussed in view of the different assumptions of the models and the differences in how they operationalize the concept of tracking.

Key words: Longitudinal studies; Growth curves; Polynomials; PC program; Tracking

Introduction

In a recent paper [1] we described a tracking index for longitudinal observations due to Foulkes and Davis [2] which did not assume any particular structure for the growth curves. Such an index has the obvious advantage of being widely applicable, but sharper results may be obtained when it is reasonable to assume that the growth curves have some specified functional form. The purpose of the present paper is to extend the Foulkes-Davis (FD) tracking index to the situation in which the growth curves may be assumed to be polynomials (tests of whether or not this assumption is tenable are included) and to provide a GAUSS program which will do the associated computations. An example illustrating the technique and the use of the program is given. The results are compared with those obtained using the 'non-parametric' form of the FD index and the index based on the kappa statistic [3].

Correspondence to: Emet D. Schneiderman, Department of Oral and Maxillofacial Surgery, Baylor College of Dentistry, 3302 Gaston Avenue, Dallas, TX 75246, USA.

0020-7101/93/$06.00 © 1993 Elsevier Scientific Publishers Ireland Ltd.
Printed and Published in Ireland

STABILITY

OF TEACHER BEHAVIOR AND TEACHER EFFECTIVENESS

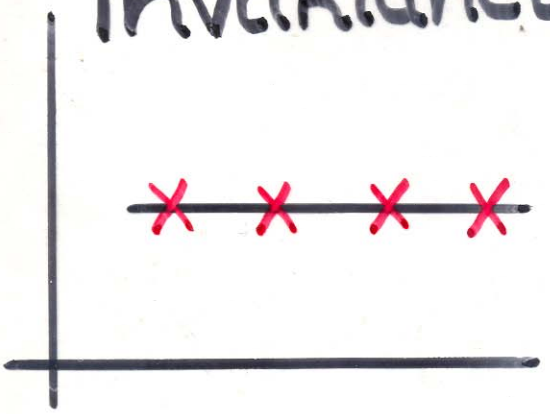
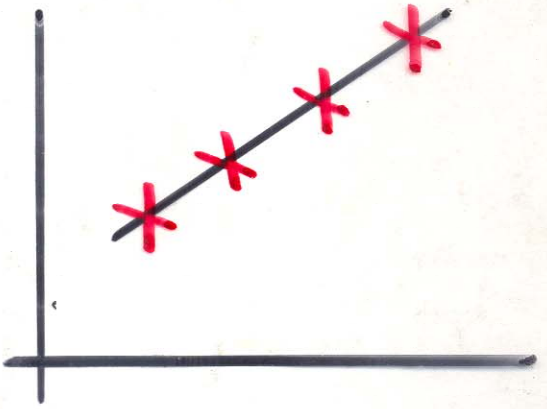
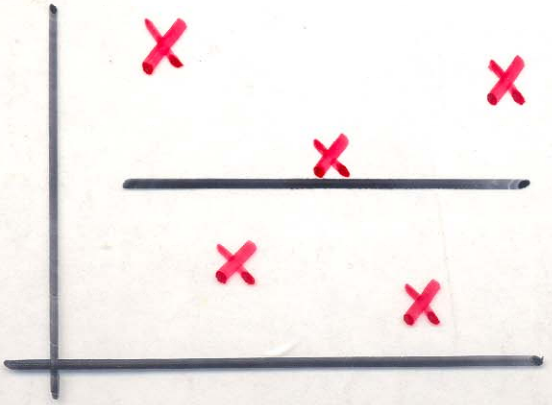
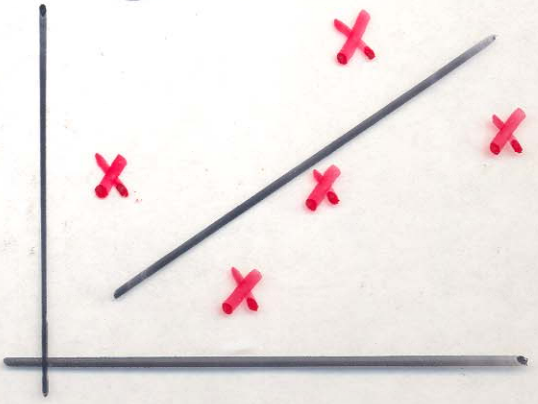
IS TEACHER BEHAVIOR STABLE?

Well-formulated questions about stability

A. Is the behavior of an individual teacher consistent over time?

B. Are between-teacher differences consistent over time?

PREDICTABILITY AND STATIONARITY

	STATIONARY	NONSTATIONARY
PREDICTABLE	Absolute Invariance 	Trend 
NONPREDICTABLE	Scatter 	Trend Plus Scatter 

ANSWERING QUESTION A

OBSERVE

target behavior(s) of an individual teacher on multiple occasions

---

THREE TYPES OF DATA ON TARGET BEHAVIORS

(i) Behavior counts

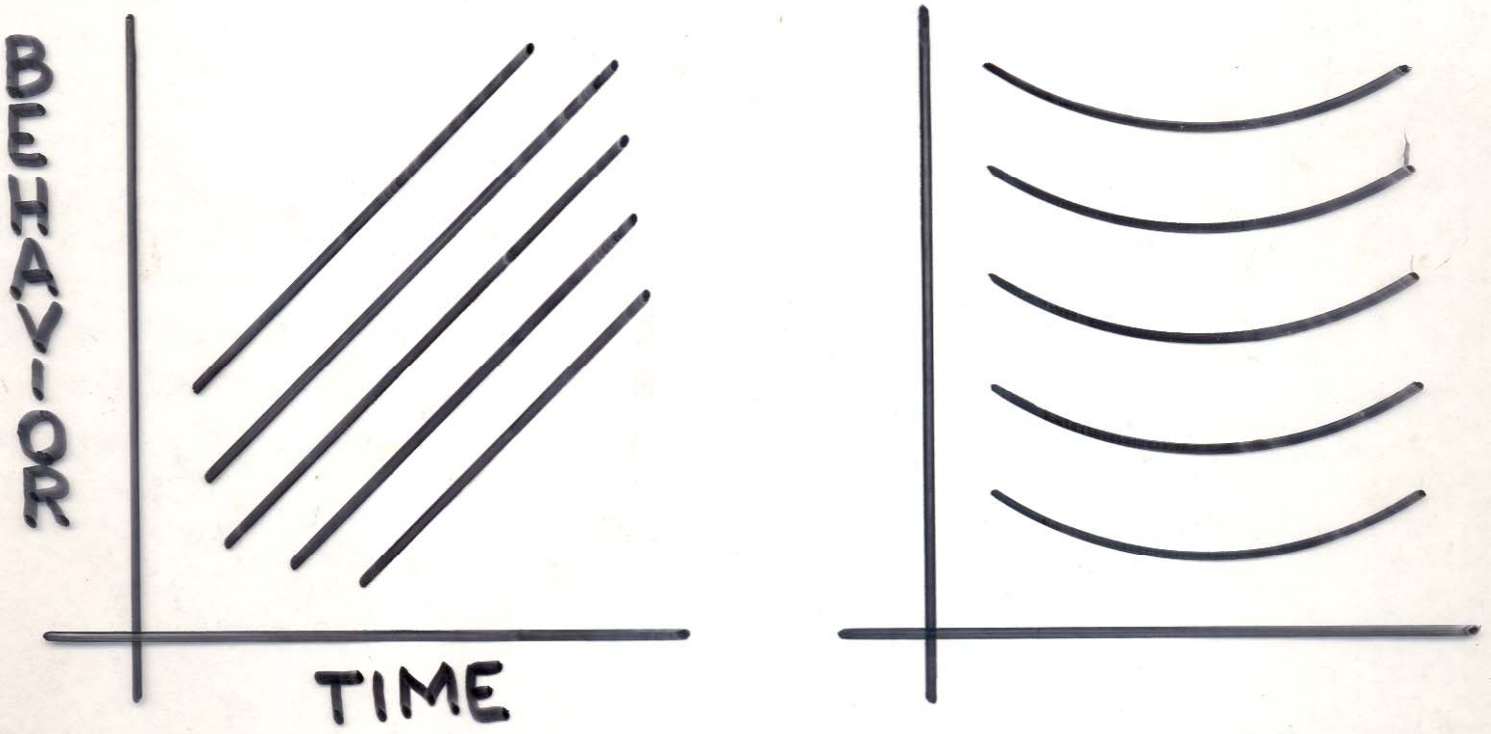
(ii) Bernoulli trials

(iii) Quantitative Measures

ANSWERING QUESTION B

DO TEACHERS TRACK?

PERFECT TRACKING



1. maintaining percentile rank
2. nonintersection of individual time paths

RECONSIDERING PREVIOUS APPROACHES AND METHODS

1. Correlations Among Observation Times
   a. Time 1 - Time 2
   b. Multiple times
2. Generalizability Theory
3. Mean teacher time trend

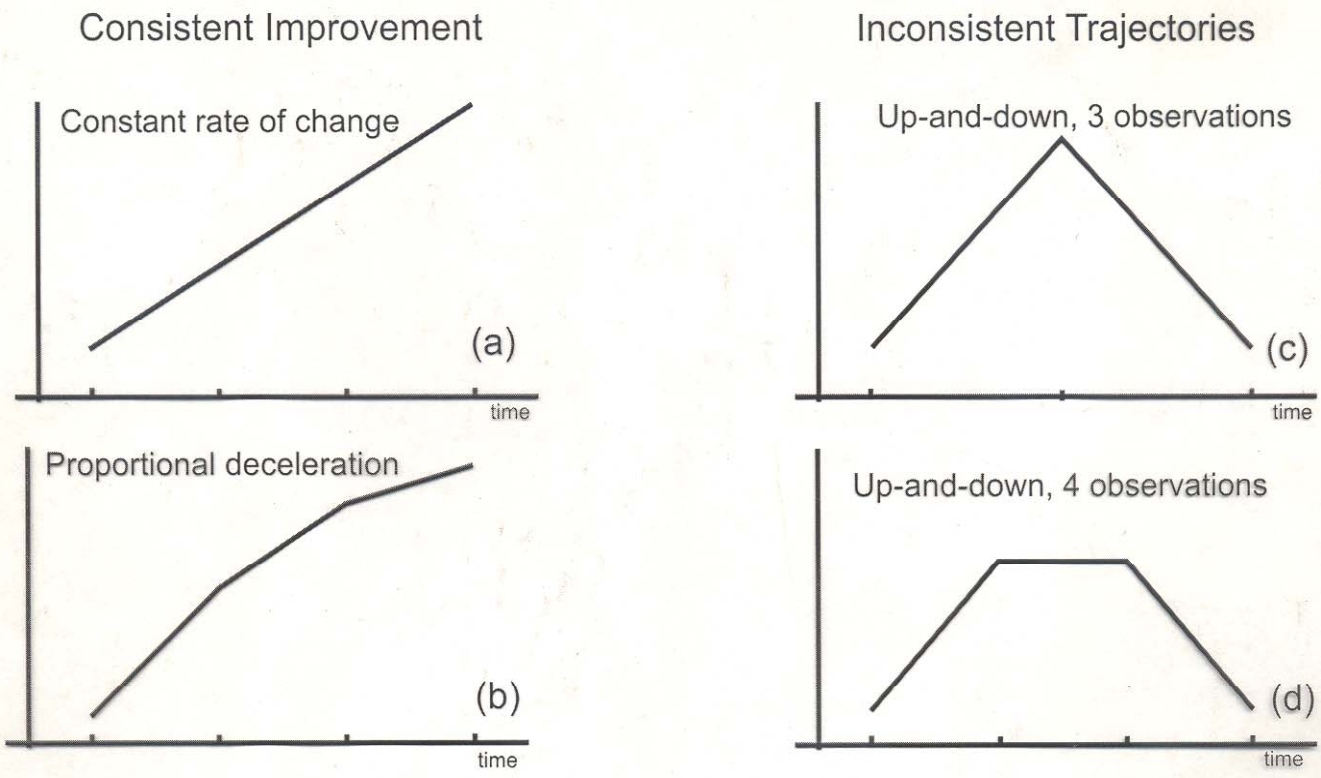


Figure 1.1. Depictions of Consistency of Individual School Improvement

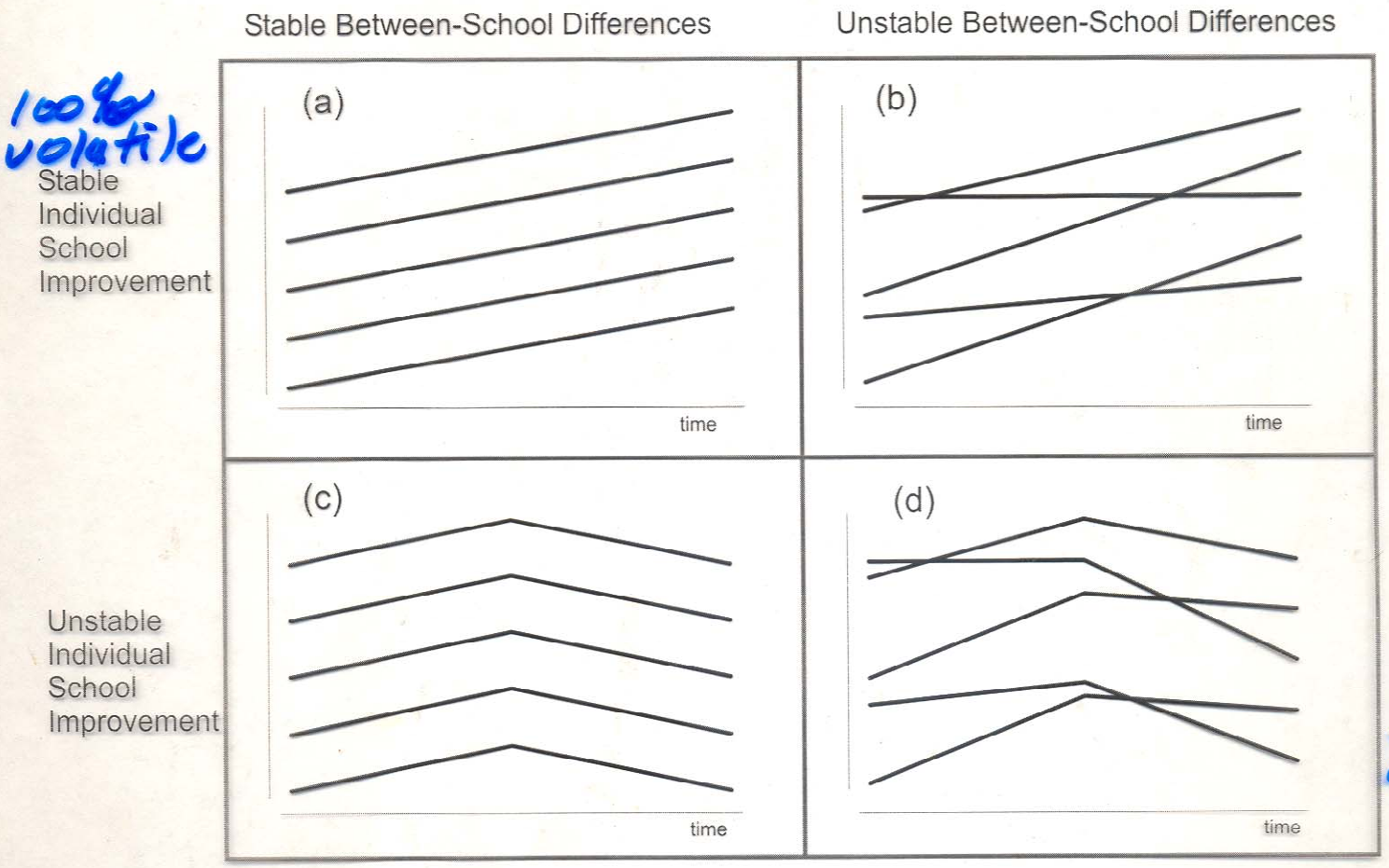


Figure 1.2. 2x2 Array of Stability of Improvement and Stability of Between-school Differences for Collection of Schools.

LH stability $Corr(Y_3 - Y_1, Y_4 - Y_2)$

THE VOLATILITY SCAM

(from Rogosa *Confusions about Consistency in Improvement*)

KS persistence $1 + 2\,Corr(Y_3 - Y_2, Y_2 - Y_1)$

Example A: Extreme Volatility?

	Improvement		
School	Yr1toYr2	Yr2toYr3	Yr3toYr4
A	40	50	50
B	40	40	50
C	50	40	40
D	59	41	49
E	45	45	45

$r_{LH} = 0$

$p_{KS} = .062$

If the above represents "volatile", then can consistent, stable improvement ever be identified?

Example D: Up-and-Down, LH stability. .

	Improvement		
School	Yr1toYr2	Yr2toYr3	Yr3toYr4
A	39	7	–47
B	39	–3	–54
C	49	–3	–41
D	49	–1	–36
E	44	2	–44

$r_{LH} = .977$

An up-and-down example for the KS procedure

	Improvement	
School	Yr1toYr2	Yr2toYr3
A	40	–40
B	40	–50
C	50	–50
D	50	–40
E	45	–45

$p_{KS} = 2$

Longitudinal Reasons to AVOID Structural Equation Models

David Rogosa

Stanford University

rag@leland.stanford.edu

School of Education
Stanford CA 94305

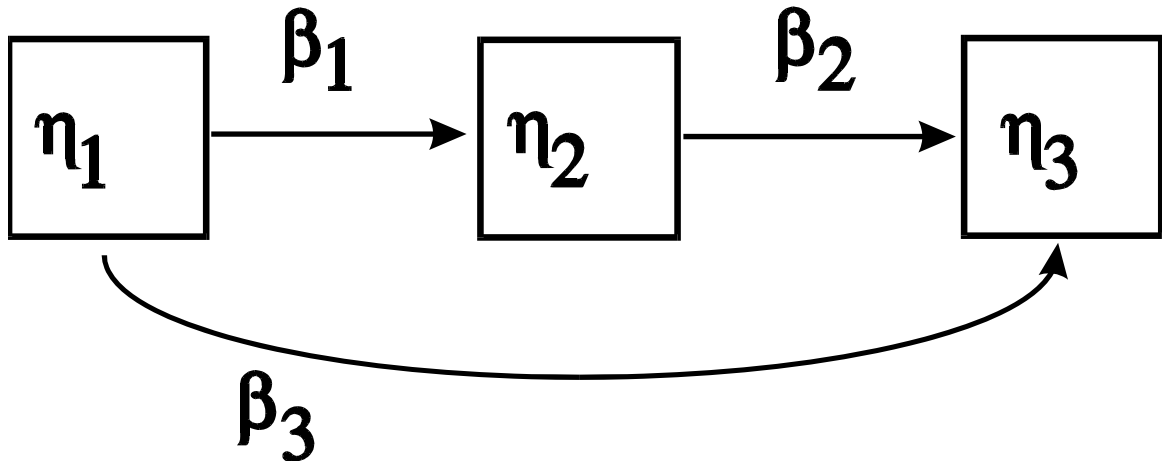
Quantitative Social Science Seminar Series
U.C. Berkeley, March 14, 1994

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

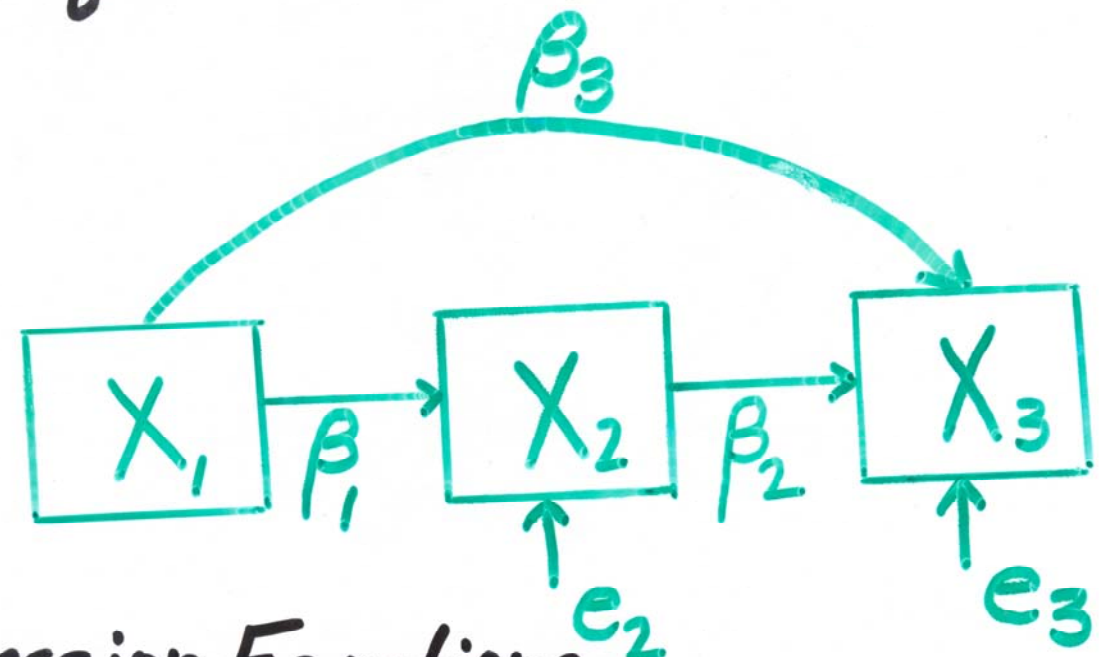
Data example 40 cases, each observed at three time points {1, 3, 5} . true observations fall on a straight-line growth curve
Regression for $\eta(t_3)$ matches the theoretical results $\beta_3 = (3 - 5)/(3 - 1)$ and $\beta_2 = (5 - 1)/(3 - 1)$-- with squared multiple correlation of 1.0.

Path Analysis

Example: Goldstein (JRSS-B, 1979)

X = reading test score

$t_1 = 7$ yr. $t_2 = 11$ yr. $t_3 = 16$ yr



Regression Equations

$$X_2 = \alpha_2 + \beta_1 X_1 + e_2$$

$$X_3 = \alpha_3 + \beta_2 X_2 + \beta_3 X_1 + e_3$$

Estimates (for transformed variables)

$\hat{\beta}_1 = .841$, $\hat{\beta}_2 = 1.11$, $\hat{\beta}_3 = -.147$

Interpretation??

"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 nonlinear terms does not change this picture to any extent. Nevertheless, the analysis is restricted to the measurements employed, and it might be possible to use other measuring instruments at 11 which would eliminate the seven-year-old contribution."

(Goldstein 1979)

Goldstein estimates

(for transformed, standardized variables)

$\hat{\beta}_1 = .841, \hat{\beta}_2 = 1.11, \hat{\beta}_3 = -.147$

Interpretation?

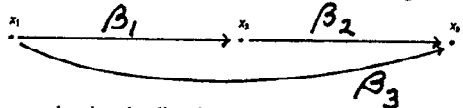
Path Analysis II
Longitudinal Data (Goldstein ea)

Rogosa 1988 Casual models.....

Does it work?

5. MODELS

The first and simplest model to be analysed is represented in Fig. 2 as a path diagram.



n=17000
ages 7 11 16

FIG. 2. Path diagram showing the direction of relationships between measurements at three occasions.

Rogosa (1988)
individual-level model straight-line growth in reading

specifies a constant rate of change denoted by θ. The straight-line growth curve for individual p is written:

$$\xi_p(t) = \xi_p(0) + \theta_p t. \qquad (1)$$

Path Regressions

Path analysis models for longitudinal data use the temporal ordering of the measurements to delimit the possible paths between the variables. Consider the example of a three-wave design with measures on X at times t_1, t_2, t_3. The path regressions for the unstandardized variables are:

$$X_2 = \alpha_2 + \beta_1 X_1 + e_2$$
$$X_3 = \alpha_3 + \beta_2 X_2 + \beta_3 X_1 + e_3. \qquad (2)$$

gression (path) coefficients from (2) are:

$$\beta_3 = \frac{t_2 - t_3}{t_2 - t_1} < 0$$

and $R^2 \geq 1$

$$\beta_2 = \frac{t_3 - t_1}{t_2 - t_1} > 0. \qquad (3)$$

data-free causal analysis

[Note: the results in (3) are easily verified by using (1) to substitute $X_2 = X_1 + \theta(t_2 - t_1)$ into the X_3 equation in (2). Substituting the values for β_2 and β_3 from (3) and collecting terms yields $X_3 = X_1 + \theta(t_3 - t_1)$.]

additional results in Trento (dec'93)

3.2. Exponential growth

For a population of individual growth curves given by (2) the population partial regression (structural) coefficients for the η_3 equation in (3) are:

$$\beta_3 = \frac{\exp[-\gamma t_3] - \exp[-\gamma t_2]}{\exp[-\gamma t_1] - \exp[-\gamma t_2]} > 0$$

$$\beta_2 = \frac{\exp[-\gamma t_3] - \exp[-\gamma t_1]}{\exp[-\gamma t_2] - \exp[-\gamma t_1]} < 0.$$

just depends on curvature

individ growth

$$\eta_p(t) = \lambda_p - (\lambda_p - \eta_p(0))e^{-\gamma t}$$

asymptote λ_p curvature γ
linear state dependence (7) $\frac{d\eta}{dt}$

For straight-line growth

$$\beta_3 = \beta_{X(t_3)X(t_1)\cdot X(t_2)} = \frac{t_2 - t_3}{t_2 - t_1} < 0$$

$$\beta_2 = \beta_{X(t_3)X(t_2)\cdot X(t_1)} = \frac{t_3 - t_1}{t_2 - t_1} > 0$$

Properties (Moments of Observables) of Collections of Growth Curves

STAT 222 Week 2 D Rogosa

for indv p $\quad \xi_p(t) = \xi_p(0) + \theta_p t$ $\quad t_i \ (i = 1, \ldots T)$ $\quad p \ (p = 1, \ldots n)$

"centering"

$$\xi_p(t) = \xi_p(t^0) + \theta_p (t - t^0)$$

centering, scale

$t^0 = -\sigma_{\xi(0)\theta} / \sigma^2_\theta$

$\rho_{\xi(t^0)\theta} = 0$, min Var($\xi$)

Scale $K = \sigma_{\xi(t^0)} / \sigma_\theta$

time metric

Moments

Covariance $\sigma_{\xi(t_1)\xi(t_2)} = (t_1 - t^0)(t_2 - t^0)\sigma^2_\theta + \sigma^2_{\xi(t^0)}$

Variance $\sigma^2_{\xi(t)} = \sigma^2_{\xi(t^0)} + ((t - t^0)/K)^2 \sigma^2_{\xi(t^0)}$

$$\sigma^2_{\xi(t)} / \sigma^2_{\xi(t^0)} = 1 + \left(\frac{t - t^0}{K}\right)^2$$

Correl change, initial status $\rho_{\theta\xi(t)} = \dfrac{t - t^0}{[K^2 + (t - t^0)^2]^{1/2}}$

exogenous var w

$$\rho_{w\xi(t)} = \frac{(t - t^0)\rho_{w\theta} + K\rho_{w\xi(t^0)}}{[K^2 + (t - t^0)^2]^{1/2}}$$

where $t^u = t^0 + K\left(\dfrac{\rho_{w\theta}}{\rho_{w\xi(t^0)}}\right)$ $\quad t^l = t^0 - K\left(\dfrac{\rho_{w\xi(t^0)}}{\rho_{w\theta}}\right)$

Week 1 example: (p. 64) $\theta \sim U[1, 9]$, $\xi(t^0) \sim U[38, 62]$

$t^0 = 2$ $\quad \sigma^2_\theta = 5.333$ $\quad \sigma^2_{\xi(t^0)} = 48$ $\quad \rho_{w\theta} = 0$ $\quad \rho_{w\xi(t^0)} =$

---

at time t_i $\quad X_{ip} = \xi_{ip} + \varepsilon$ $\quad \varepsilon \sim (0, \sigma^2_\varepsilon)$ errors in variables

weeks 1 ex $\sigma^2_\varepsilon = 10$

week 1 data

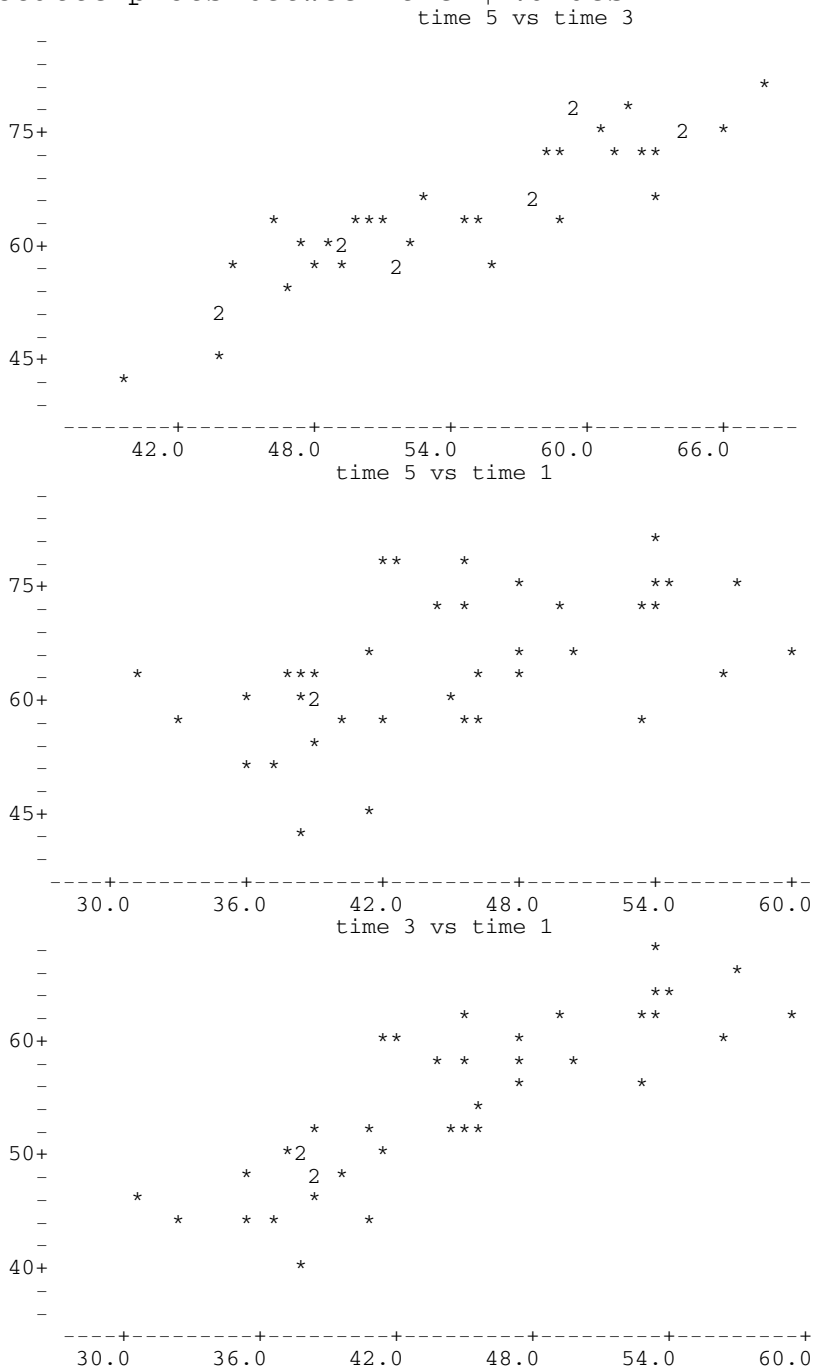
regression equation:

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

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%

Scatterplots between the η-values.



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

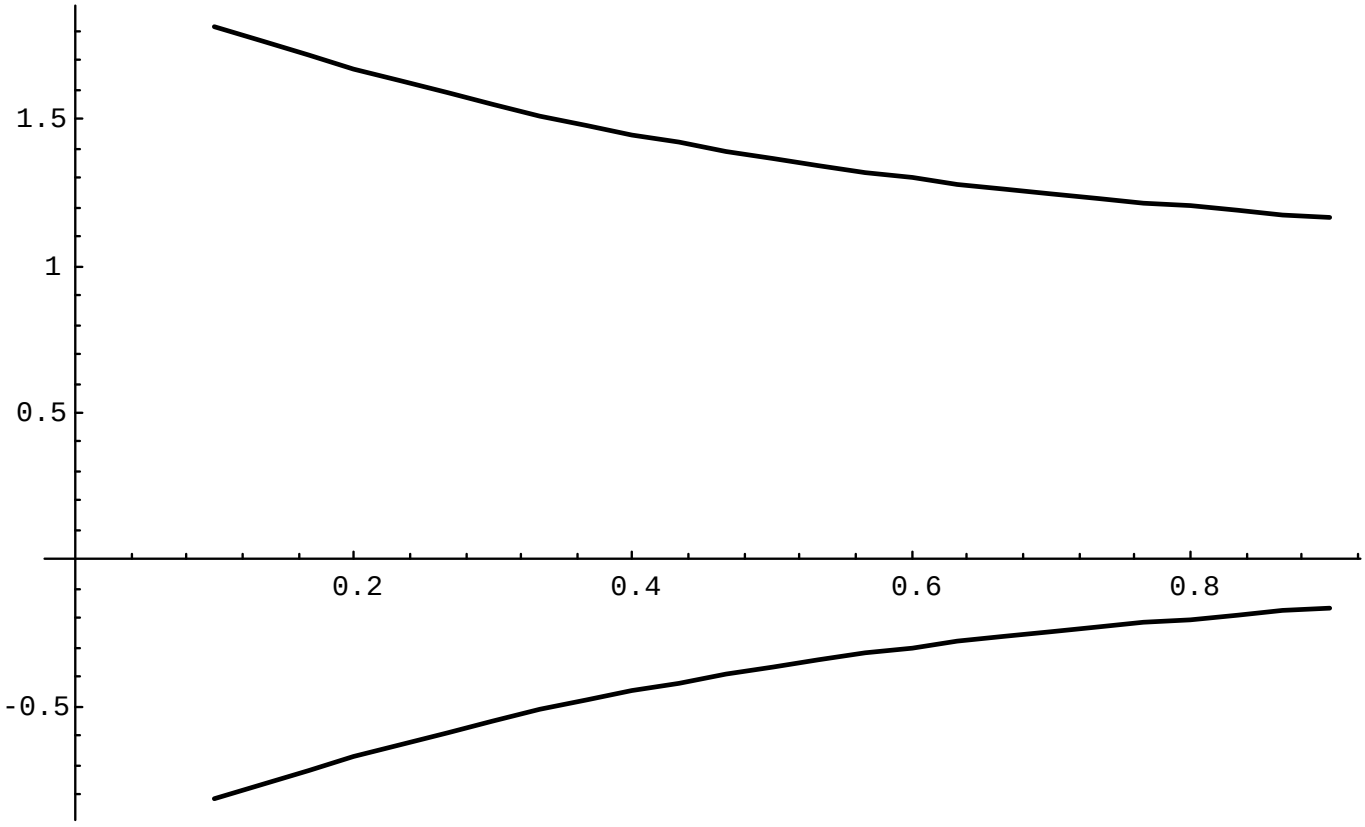
	η(1)	η(3)	η(5)
η(3)	0.842		
η(5)	0.536	0.907	
W	0.766	0.765	0.598

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.

Journal of Educational Statistics
Summer 1985, Volume 10, Number 2, pp. 99–107

SATISFYING A SIMPLEX STRUCTURE IS SIMPLER THAN IT SHOULD BE

DAVID ROGOSA and JOHN B. WILLETT
Stanford University

KEY WORDS. *Longitudinal data analysis, simplex model, growth curves.*

ABSTRACT. Markedly different types of growth (learning) curves may generate indistinguishable covariance structures. We illustrate with an example of a 5×5 covariance matrix representing longitudinal measurements at five occasions. This example appears to conform closely to a simplex correlation pattern, and a simplex covariance structure provides an excellent fit (using LISREL V) to this covariance matrix. However, the (known) structure of this example differs greatly from the simplex model. In addition to indicating that an excellent fit of a simplex structure can be misleading, this example provides an opportunity to question common uses of covariance structure models for the study of growth.

Since the introduction of the simplex model by Guttman (1954), simplex structures have been posited for many types of ordered variables (such as hierarchical tests or longitudinal measurements). In many analyses of longitudinal data, a simplex covariance or correlation structure serves as a basic model for statistical analysis (e.g., Humphreys, 1960; Jöreskog, 1970a, 1978; Jöreskog & Sörbom, 1977; Werts & Hilton, 1977; Werts, Linn, & Jöreskog, 1977, 1978). As Werts, Linn, and Jöreskog assert, "The simplex model appears to be particularly appropriate for studies of academic growth (Humphreys, 1960, 1968; Lunneborg & Lunneborg, 1970)" (p. 745).

Simplex Growth Model

Following Jöreskog (1970a), a discrete-time autoregressive model of lag 1, which gives rise to the simplex structure, has the form:

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

where η_{ip} indicates the value of η for the pth individual at time t_i, and where the $\delta_{i+1,p}$ are independent disturbances (increments to growth) with expectation zero and $\eta_{1p} = \delta_{1p}$. (See also the stochastic models considered in Anderson, 1960, Section 3.) An important property of Equation 1 is that the partial correlation is zero between η at two time points with η at an intervening time partialled out. This condition, which was used by Guttman (1954, Equation 15) to characterize the simplex structure, can be written

$$\rho_{\eta_i \eta_k \cdot \eta_j} = 0 \qquad (t_i < t_j < t_k). \tag{2}$$

"Unsimplex" Growth Models

An alternative simple model for growth specifies that each individual has a constant rate of change. This constant rate of change model yields a straight line growth curve for each individual:

$$\eta_{ip} = \eta_{1p} + \theta_p(t_i - t_1). \tag{3}$$

A striking consequence of the constant rate of change model is that it represents the most extreme violation of Guttman's condition for the simplex. That is, for straight-line growth,

$$\rho_{\eta_i \eta_k \cdot \eta_j} = -1 \qquad (t_i < t_j < t_k). \tag{4}$$

A related property of straight-line growth is that $\beta_{\eta_k \eta_i \cdot \eta_j} = (t_j - t_k)/(t_j - t_i)$ (e.g., with equally spaced time points $\beta_{\eta_3 \eta_1 \cdot \eta_2} = -1$). In contrast, the autoregressive model in Equation 1 specifies that $\beta_{\eta_k \eta_i \cdot \eta_j} = 0$ (for $t_i < t_j < t_k$). Also, for straight-line growth, $\beta_{\eta_k \eta_j \cdot \eta_i} = (t_k - t_i)/(t_j - t_i)$ (e.g., with equally spaced time points $\beta_{\eta_3 \eta_2 \cdot \eta_1} = 2$).

A second type of growth (learning) curve for which Equation 4 holds is the function that specifies that individuals approach an asymptote λ_p exponentially with rate parameter γ (where γ is identical for all individuals). This exponential growth curve can be written as

$$\eta_{ip} = \lambda_p - (\lambda_p - \eta_{1p}) e^{-\gamma(t_i - t_1)}. \tag{5}$$

Among the applications of the growth curve in Equation 5 to the study of learning are Hicklin (1976), Keats (1983), and Sörensen and Hallinan (1977).

The Numerical Example

A numerical example consisting of a covariance matrix of observed scores over five occasions of measurement is given in Table I. The relation between the observed score Y_{ip} and the true score η_{ip} is specified by the measurement model:

$$Y_{ip} = \eta_{ip} + \varepsilon_{ip}, \tag{6}$$

where Y_{ip} and η_{ip} indicate the values of observed and true scores for the pth individual ($p = 1, \ldots, 500$, as a sample size of 500 is used in this example) at the time t_i (for discrete times $t_1, t_2, \ldots, t_5$) and where ε_{ip} denotes the error of measurement in Y_{ip} at time t_i.

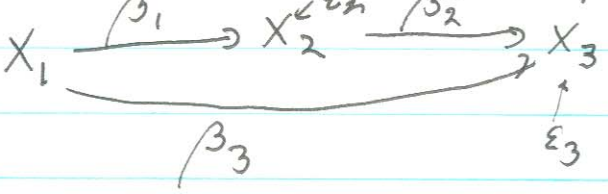
The corresponding correlation matrix for the Y_i in Table II displays the typical properties of a simplex correlation structure with the correlations decreasing for elements farther away from the main diagonal (and with nearly equal elements along the subdiagonals). Furthermore, Table II is similar to correlation matrices obtained in large-scale longitudinal studies of student

B Longitudinal Data (panel data)

observe $X_1\ X_2\ X_3$ measured w/out error at times $t_1\ t_2\ t_3$
following straight-line growth curve

$X_p(t) = X_p(0) + \theta_p t$, individual unit model with individual differences

Goldstein-style path analysis (canonical ex)



temporal = causal ordering

$$X_2 = \alpha_2 + \beta_1 X_1 + \varepsilon_2$$

$$X_3 = \alpha_3 + \beta_2 X_2 + \beta_3 X_1 + \varepsilon_3$$

Result (Rogosa 1989) also $R^2_{3\cdot 21} = 1$

$$\beta_3 = \frac{t_2 - t_3}{t_2 - t_1} \qquad \beta_2 = \frac{t_3 - t_1}{t_2 - t_1}$$

modified results if X measured w/ error (Trento paper)
also individual unit model exponential growth (handout)

Simplex Model Example (Rogosa, Willet 1994)

auto regressive lag-1

$$\eta_{i+1,p} = \beta_i \eta_p + \delta_{i+1,p} \Rightarrow \rho_{\eta_1 \eta_3 \cdot \eta_2} = 0$$

fit via LISREL (see slides) to data from maximally "unsimplex" straight-line growth

$$\eta_{ip} = \eta_{1p} + \theta_p (t_i - t_1) \Rightarrow \rho_{\eta_1 \eta_3 \cdot \eta_2} = -1$$

wonderful fit from LISREL
(cf Bracklow examples)

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)
$Corr(\eta_i, \eta_k \bullet \eta_j) = 0.$

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 $\rho_{\eta(t3)\eta(t1)\bullet\eta(t2)} = -1$.

$\tau = 3$; $\sigma^2_\theta = .0078$; $\sigma^2_{\eta(\tau)} = .4374$; $var(e) = .15$. 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, 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.

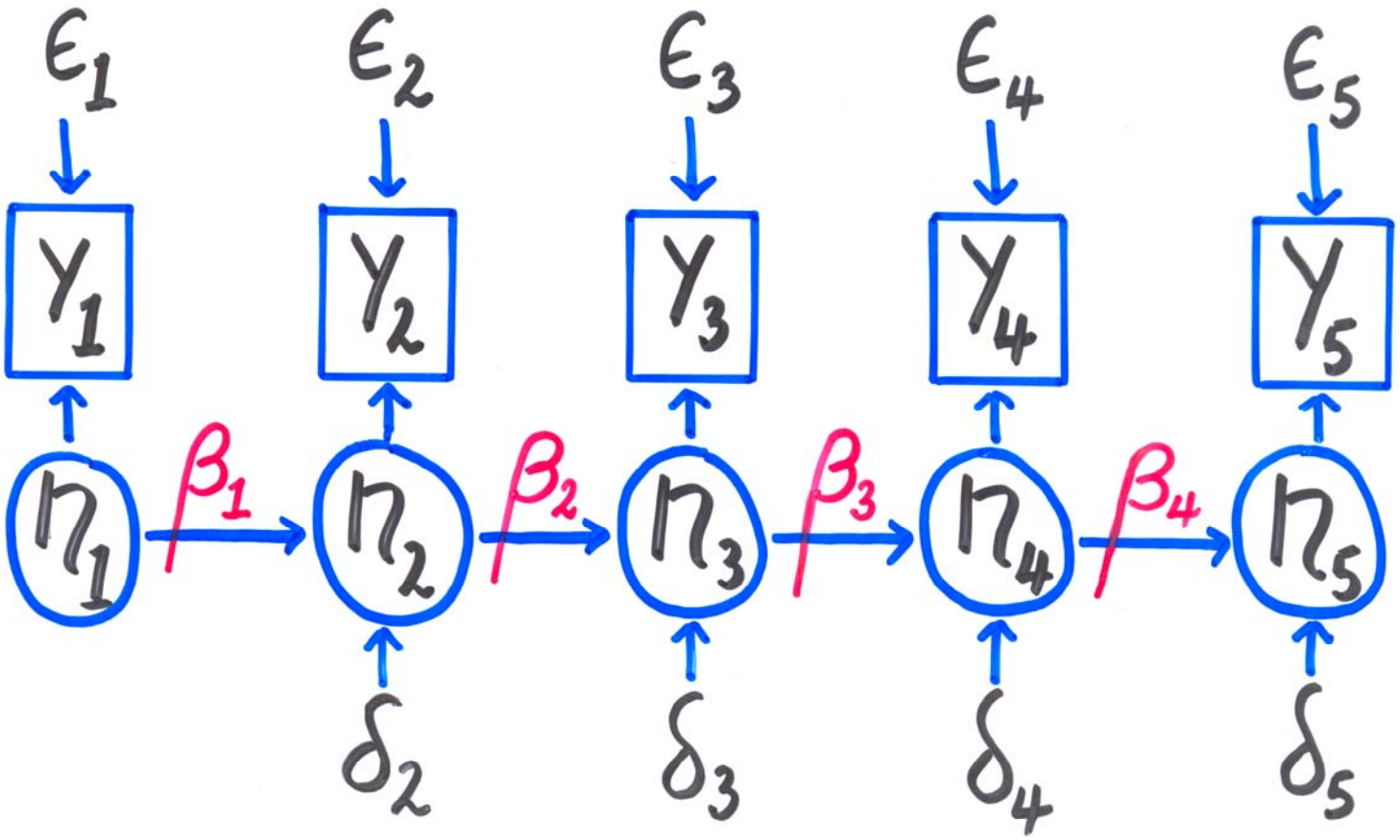
A dependable measure for assessing consistency of individual differences over a specified time interval is the index of tracking γ proposed by Foulkes and Davis (1981) ; this index estimates the probability that two randomly chosen individuals trajectories do not cross in the time interval specified. 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

THE ANALYSIS

FIT COVARIANCE STRUCTURE FOR Y USING LISREL:



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

$$Y_{ip} = \eta_{ip} + \epsilon_{ip}$$

Reproduced Correlation Matrix and Residuals

	Y_1	Y_2	Y_3	Y_4	Y_5
Y_1	1				
Y_2	.734	1			
	.002				
Y_3	.710	.751	1		
	.007	-.002			
Y_4	.694	.735	.771	1	
	-.001	.004	-.005		
Y_5	.680	.721	.755	.785	1
	-.014	.004	.006	-.003	

badness-of-fit chisq = 2.5, 5df

Root Mean Square Residual .006

The example (which is just one of many possible examples) provides one indication that very different types of individual growth curves may yield indistinguishable covariance or correlation structures. The example is not intended as a "contest" between the growth models, and hypothesis testing is not featured. The surprise in the example is the difficulty in detecting that the covariance matrix in Table I is not generated by Equations 1 and 6, given that the growth curves generated by the autoregressive model in Equation 1 are very distinct from straight-line growth. (Individual growth trajectories will differ greatly for the two models.) The purpose of this paper is less to attack covariance structure analysis than to appeal for explicit description and modeling of growth curves as a foundation for the analysis of longitudinal data.

APPENDIX A

Construction of the Example

For straight-line growth in Equation 3, the elements of the covariance matrix of the η_i can be expressed:

$$\sigma_{\eta_i\eta_j} = \sigma^2_{\eta_1} + (t_i + t_j - 2t_1)\sigma_{\eta_1\theta} + (t_i - t_1)(t_j - t_1)\sigma^2_{\theta}, \tag{A1}$$

where $t_k = k$ ($k = 1, \ldots, 5$). Our example used the values $\sigma^2_{\eta_1} = .4686$, $\sigma_{\eta_1\theta} = -.0156$, and $\sigma^2_{\theta} = .0078$. Finally, Table I was obtained by adding "measurement error variance" of .15 to the diagonal elements of the covariance matrix of the η_i. Also, this specification determines the values of the following quantities included in the text: $\rho_{\eta_i(\eta_{i+1}-\eta_i)}$, the variance of $\eta_{i+1} - \eta_i$, and the reliabilities of Y_i and $Y_k - Y_i$.

APPENDIX B

Specification of the Covariance Structure for the Quasi-Simplex

The standard covariance structure model with "no X" is written (see Jöreskog & Sörbom, 1981, Equation I.4 and Section I.3):

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}_Y(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B}')^{-1}\boldsymbol{\Lambda}_Y' + \boldsymbol{\Theta}_\epsilon. \tag{B1}$$

The covariance structure used in the LISREL analysis further specifies (B1) as follows:

$$\underset{5\times5}{\boldsymbol{\Lambda}_Y} = \mathbf{I},$$

$$\underset{5\times5}{\mathbf{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \beta_1 & 0 & 0 & 0 & 0 \\ 0 & \beta_2 & 0 & 0 & 0 \\ 0 & 0 & \beta_3 & 0 & 0 \\ 0 & 0 & 0 & \beta_4 & 0 \end{bmatrix}, \tag{B2}$$

$$\underset{5\times5}{\boldsymbol{\Psi}} = \operatorname{diag}(\sigma^2_{\delta_1}, \sigma^2_{\delta_2}, \sigma^2_{\delta_3}, \sigma^2_{\delta_4}, \sigma^2_{\delta_5}),$$

$$\underset{5\times5}{\boldsymbol{\Theta}_\epsilon} = \sigma^2_\epsilon \mathbf{I}.$$

Casual vs Serious Models

1. Representation of a story about an individual unit
   "Model processes that generate the data"
2. Representation of individual differences between units

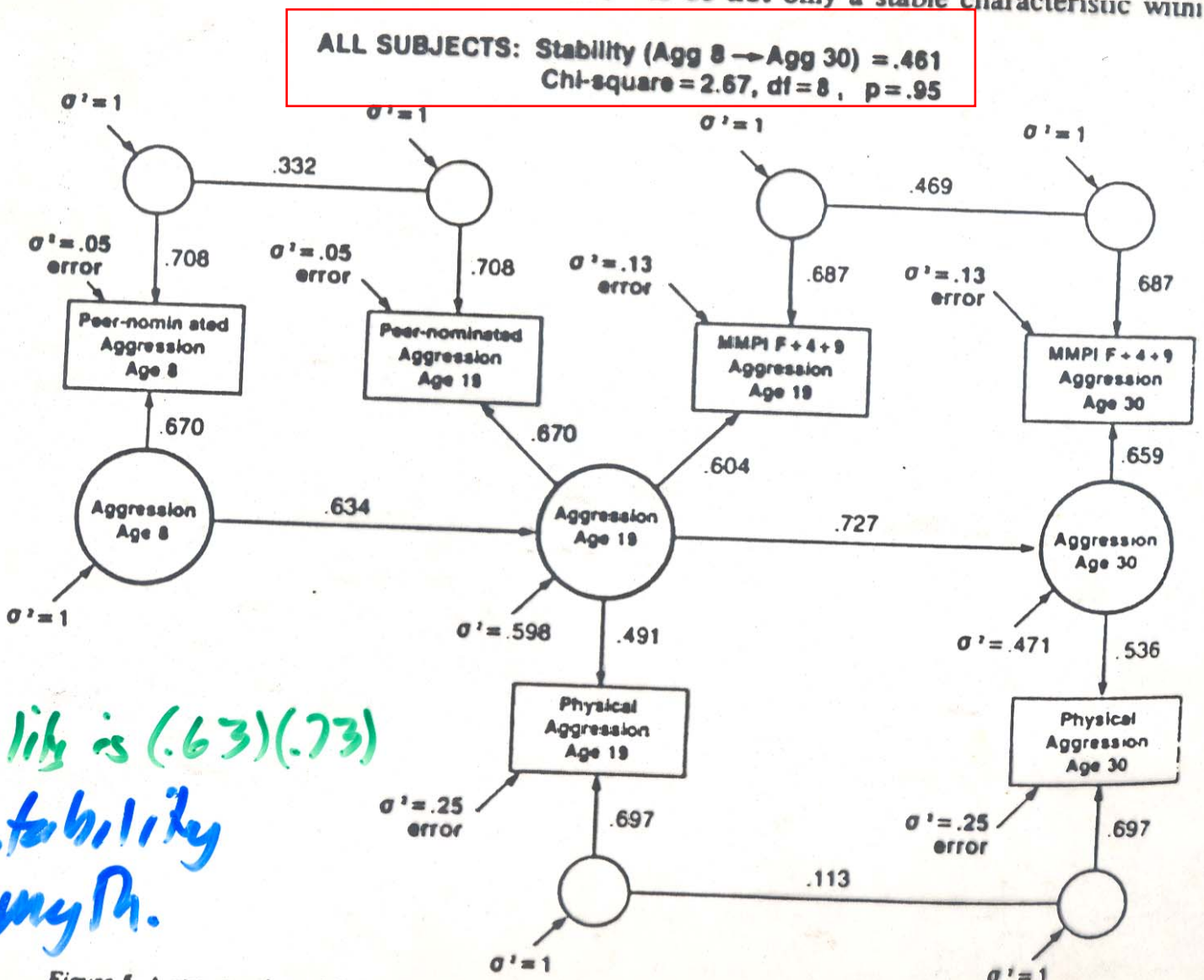
"Models" for relationships among

VARIABLES make little sense.

These are statistical models but not serious models.

Huesmann et al 1984 Dec 4 1984.
Stability of Aggression
Lisrel

to be not only a stable characteristic within



"stability is (.63)(.73)
See stability myPa.

Figure 5. A structural model showing the stability of an hypothesized latent trait of aggression within all subjects over 22 years.

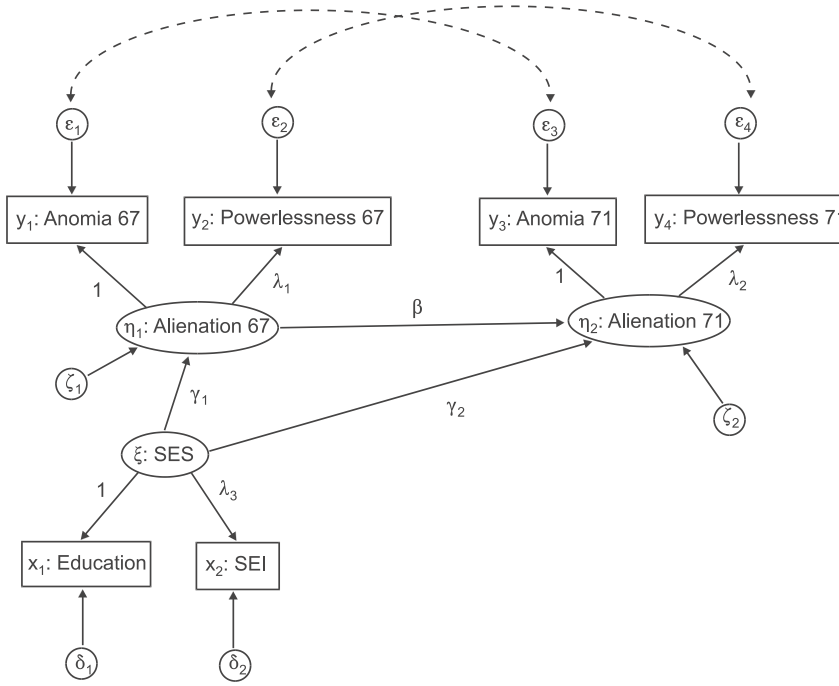


FIGURE 2 Conventional path diagram for Jöreskog and Sörbom's model for the Wheaton alienation data. Adapted with permission from Figure 6.5 in Jöreskog and Sörbom (1989, 171).

Using the common LISREL notation, this model consists of a structural submodel with two equations for the two latent endogenous variables (Alienation 67 and Alienation 71), and a measurement submodel with equations for the six indicators of the latent variables:

Structural submodel

$$\begin{aligned} \cdot_1 &= \cdot_1 \cdot + \cdot_1 \\ \cdot_2 &= \cdot_2 \cdot + \Downarrow \cdot_1 + \cdot_2 \end{aligned}$$

Measurement submodel

$$\begin{aligned} y_1 &= 1 \cdot_1 + \varepsilon_1 \\ y_2 &= \cdot_1 \cdot_1 + \varepsilon_2 \\ y_3 &= 1 \cdot_2 + \varepsilon_3 \\ y_4 &= \cdot_2 \cdot_2 + \varepsilon_4 \\ x_1 &= 1 \cdot + \delta_1 \\ x_2 &= \cdot_3 + \delta_2 \end{aligned}$$

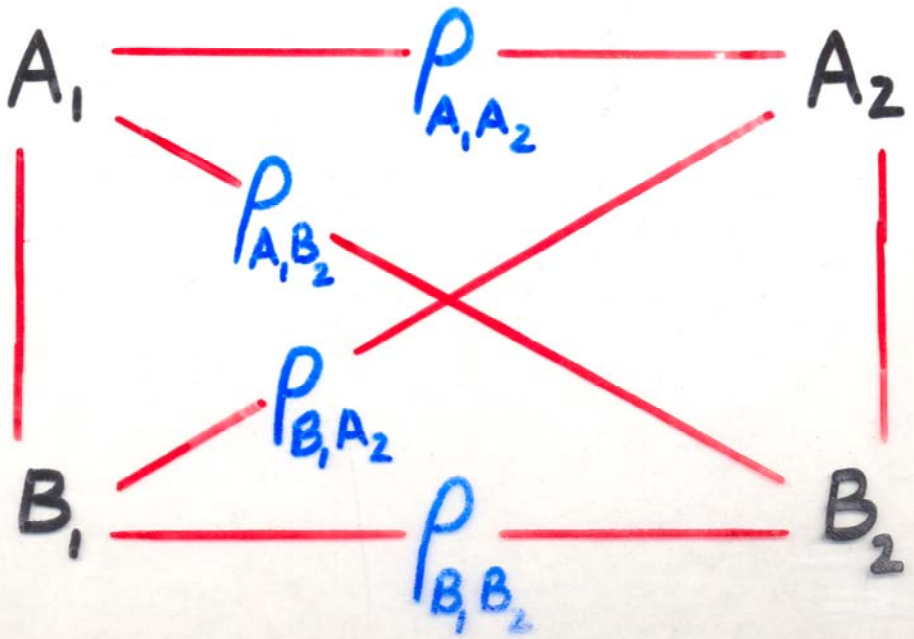
Reciprocal Effects

Does A cause B or does B cause A?

TV violence and aggression?

Personality characteristics and drug use?

Self esteem and achievement?



Weeks 6 [redacted]

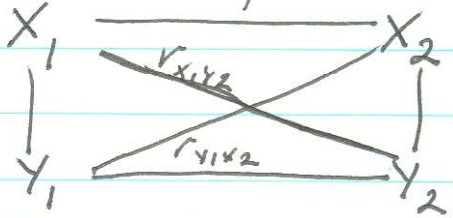
Stat 209 [redacted]

Reciprocal Effects, Simultaneous Equations

CLC (cross-lagged correlation)

compare $r_{X_1 Y_2}$ to $r_{Y_1 X_2}$

causal predominance to the larger (sig)

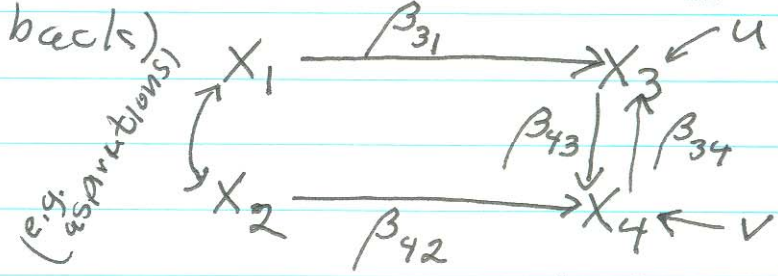


Non-Recursive (feedback)

$X_3 = \beta_{31} X_1 + \beta_{34} X_4 + u$

$X_4 = \beta_{42} X_2 + \beta_{43} X_3 + v$

$E(X_i) = 0 \quad E(u) = E(v) = 0$



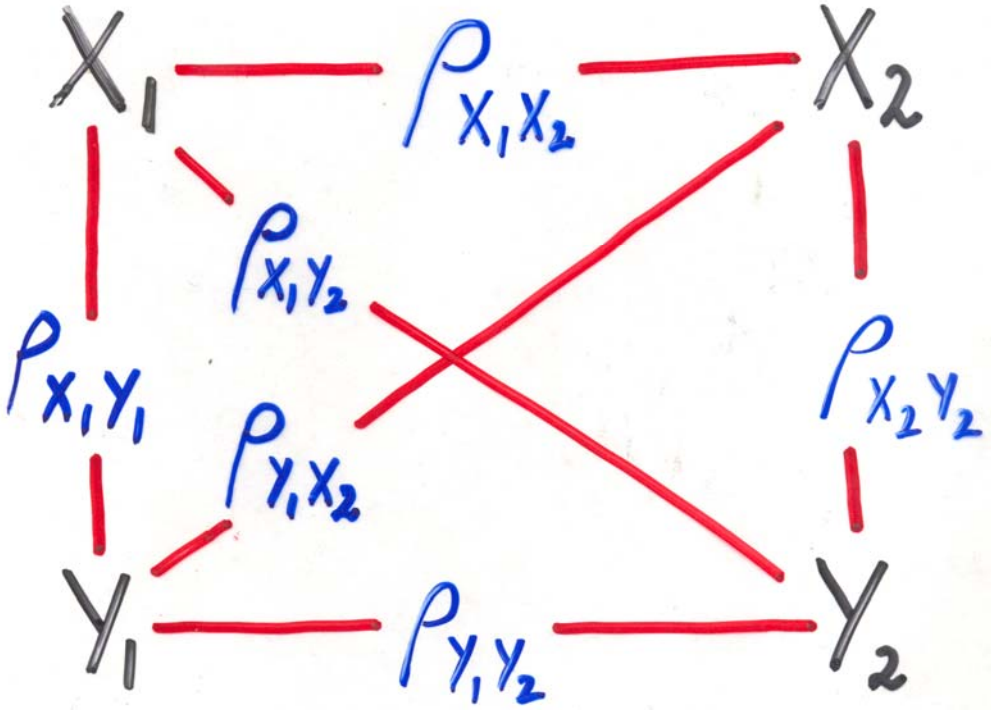
Cross-lagged credo

"With the introduction of the cross-lagged panel correlation method ..., causal inferences based on correlational data obtained in longitudinal studies can be made and enjoy the same logical status as those derived in the more standard experimental settings."

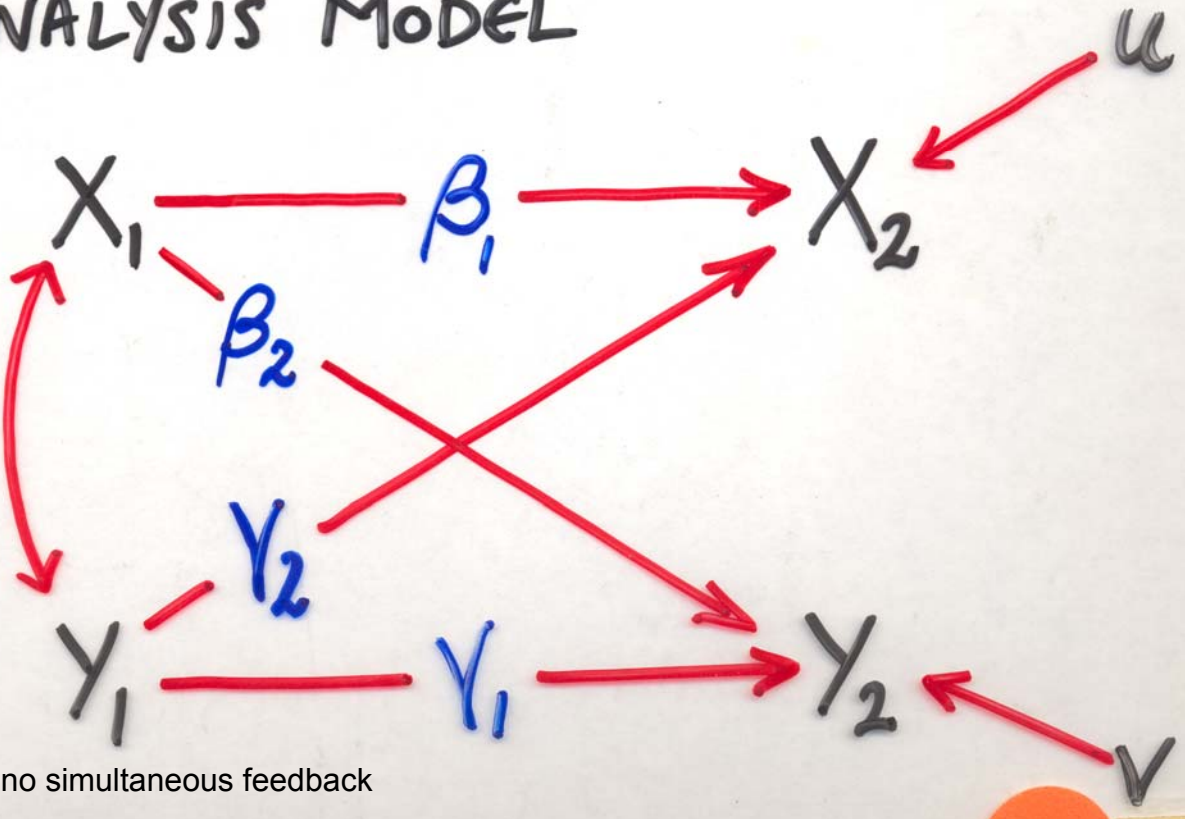
Crano and Mellon (1978) p. 41

RECIPROCAL EFFECTS

CROSS-LAGGED CORRELATION "MODEL"



PATH ANALYSIS "MODEL"



effects are lagged, no simultaneous feedback

[redacted]

Reciprocal Effects $X \rightleftarrows Y$

Longitudinal — TV violence, aggression

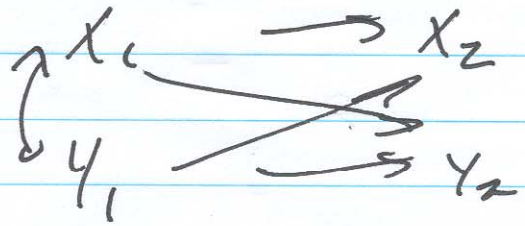
CLC $r_{x_1 y_2}$ vs $r_{y_1 x_2}$

Linked differential eq's

$$\frac{dY}{dt} = a + bX + cY$$

$$\frac{dX}{dt} = d + eX + fY$$

$\Rightarrow$ lagged regression discrete time



Time-series, Granger causality — Pierce + Haugh.

discrete data — Lazarsfeld 16-fold table (Lazarsfeld + Merton)

Simultaneous (reciprocal) effects

no lagged repeated measures

use IV exogenous vars to estimate relative effects

Duncan Haller Portes — aspiration

Rindflus (Juno) ed, fertility (ages)

the context of particular hypothesized models, and the assumptions contained in these models must be defended in a given situation. In particular, two-wave reciprocal effects models involve a set of possibly restrictive assumptions that can be relaxed only, as in cross-sectional research, by including outside variables. Three-wave and multiwave panels, however, can be estimated by imposing fewer constraints on the causal parameters.

Cross-Lagged Effects Models

The Two-Wave Model

The most basic model for estimating possible reciprocal effects is an extension to two dependent variables of the lagged effect static-score model considered in the previous chapter (Equation 2.14), with each variable at time 2 being predicted by its previous value as well as the time 1 value of the other variable of interest. This model is shown in diagram form in Figure 3.1. X_1 and X_2 represent one variable (e.g., group memberships) measured across the two panel waves, and Y_1 and Y_2 another variable (e.g., protest potential) measured at both time points. X_2 and Y_2 are hypothesized to be determined by their wave 1 values, the lagged value of the other variable, and an error term U. The correlation between the wave 1 variables is represented in the figure by ρ_1, and the correlation between the structural disturbances of the wave 2 equations is represented by $\rho_{U_1U_2}$. The two structural equations can be written as follows:

$$\begin{aligned} Y_2 &= \beta_1 X_1 + \beta_2 Y_1 + U_1 \\ X_2 &= \beta_3 Y_1 + \beta_4 X_1 + U_2 \end{aligned} \tag{3.1}$$

with all variables in the model expressed in mean deviation form to eliminate consideration of the intercept term.

The cross-lagged model has wide applicability in panel analysis. When change in the dependent variables is modeled according to a discrete time process, the cross-lagged model will be appropriate whenever the causal lags are approximately equal to the time period between measurements. When change in the dependent variables occur continuously, it can be shown that the cross-lagged model represents the integral solution to a system of differential equations where the instantaneous rates of change in X and Y are dependent on one another over time, as in

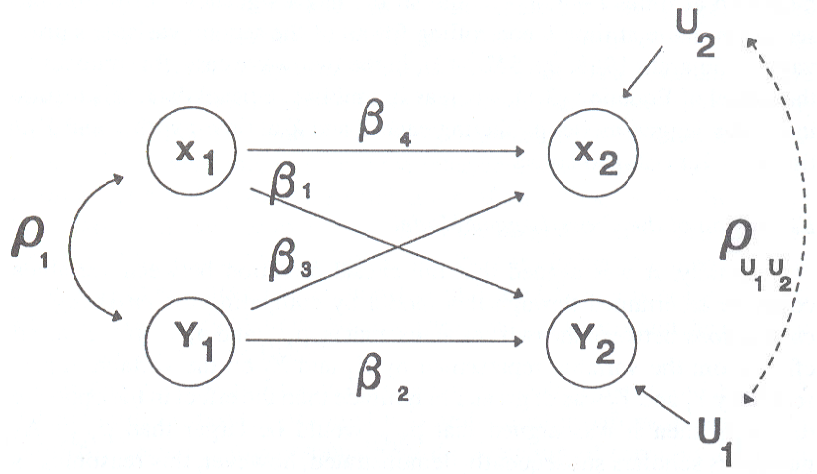


Figure 3.1. Two-Wave Model With Cross-Lagged Effects

$$\begin{aligned} \frac{dY_t}{dt} &= c_0 + c_1 X_t + c_2 Y_t \\ \frac{dX_t}{dt} &= c_3 + c_4 Y_t + c_5 X_t . \end{aligned} \tag{3.2}$$

The β_k in Equation (3.1) are, as in the single-equation continuous time model considered in the previous chapter, nonlinear functions of the c coefficients, the "fundamental parameters of change" in Equation (3.2), and the time between panel waves. The calculation of the c coefficients from the β_k involves complex mathematical manipulations beyond the scope of this monograph; interested readers are referred to Arminger (1986), Coleman (1968, pp. 448-452), and Tuma and Hannan (1984, Chapters 11-12) for details. For our purposes, it is sufficient to note that when the causal system is one of continuous reciprocal feedback from one variable to another, the cross-lagged model "tends not to be misleading about the *direction* of causal influence" (Dwyer, 1983, p. 352).

It may also be noted that the cross-lagged model corresponds to the so-called "Granger test" for causality in time series analysis, which posits that a variable "Granger-causes" the other if any value of the first variable

Psychological Bulletin
1980, Vol. 88, No. 2

A Critique of Cross-Lagged Correlation

David Rogosa
Department of Education, University of Chicago

Cross-lagged correlation is not a useful procedure for the analysis of longitudinal panel data. In particular, the difference between the cross-lagged correlations is not a sound basis for causal inference. Demonstrations of the failure of cross-lagged correlation are based mainly on results for the two-wave, two-variable longitudinal panel design. Extensions of these results to panels with multiple waves and multiple measures reveal additional problems.

The topic of this article is the analysis of reciprocal causal effects. Often, questions about reciprocal causal effects have been phrased, Does X cause Y or does Y cause X? More formally, many have spoken of a determination of causal predominance or of a preponderant causal effect. Examples of such research questions in developmental psychology include the reciprocal influences in mother–child interaction (Clarke-Stewart, 1973) and relationships between infant intelligence and infant behavior (Crano, 1977). Examples from educational research include the relationship between teacher expectation and student achievement (Crano & Mellon, 1978; Humphreys & Stubbs, 1977; West & Anderson, 1976) and the relationship between self-concept and achievement (Bachman & O'Malley, 1977; Calsyn & Kenny, 1977; Purkey, 1970). Empirical research on topics such as these has resulted in the collection and analysis of large amounts of longitudinal panel data.

Data from a longitudinal panel consist of observations on n cases at T ($t = 1, \ldots, T$) time points or waves. At each time point, observations on one or more variables are obtained. Much attention is given to the simplest relevant panel design: the two-wave, two-variable (2W2V) longitudinal panel. For 2W2V panels the two variables are labeled X and Y. These variables are subscripted to indicate the time of measurement. Thus for each individual case, measures X_1, Y_1, X_2, and Y_2 are available.

Cross-lagged correlation (CLC) is currently the most popular procedure in many areas of psychological and educational research for identifying causal effects from longitudinal panel data. Most often CLC is used to determine a predominant causal influence—the causal winner.

Users of CLC often make enthusiastic claims. For example, Crano and Mellon (1978) asserted,

> With the introduction of the cross-lagged panel correlational method. . ., causal inferences based on correla-

This research was supported by a seed grant from the Spencer Foundation. The assistance of David Brandt and Michele Zimowski in this research is gratefully acknowledged.

Requests for reprints should be sent to David Rogosa, who is now at the School of Education, Stanford University, Stanford, California 94305.

Copyright 1980 by the American Psychological Association, Inc. 0033-2909/80/8802-0245$00.75

be an improvement over current practice in CLC; the construction of the interval estimate can be found in Olkin (1966).

Also, significant differences between the cross-lagged correlations are interpreted without regard for the assumptions of CLC. Rarely is any assessment of stationarity made in applications of CLC. In many sets of longitudinal data, the difference between the synchronous correlations is about as large as the difference between the cross-lagged correlations (e.g., Humphreys & Stubbs, 1977, Table 4; Kenny, 1975, Tables 3 and 5). In a few applications a test of the null hypothesis—$H_0: \rho_{X_1Y_1} = \rho_{X_2Y_2}$—is performed. However, nonrejection does not prove the null hypothesis true, and in addition this null hypothesis is only a necessary condition for stationarity. The statistical test of equal cross-lagged correlations ignores the assumptions and the use of a preliminary test for stationarity. The complete null hypothesis should be that the cross-lagged correlations are equal, conditional on stationarity. That is, the conditional null hypothesis of interest is $H_0: \rho_{X_1Y_2} = \rho_{Y_1X_2} | \rho_{X_1Y_1} = \rho_{X_2Y_2}$. No exact statistical test is available for this conditional null hypothesis, although methods such as covariance structure analysis (Jöreskog & Sörbom, 1979) could be used to form a large-sample, normal-theory test. Of course, no improvement in the use of statistical inference procedures can offset the basic deficiencies of CLC.

Summary and Discussion

No justification was found for the use of CLC. In CLC both determinations of spuriousness and attributions of causal predominance are unsound. The results for 2W2V panels demonstrate that when reciprocal causal effects are absent, the difference between the cross-lagged correlations may be either small or large, and when reciprocal causal effects are present, the difference between the cross-lagged correlations may be either small or large. Also, the practice in CLC of reducing the analysis of data with multiple waves and multiple measures to a collection of 2W2V analyses produces additional problems. CLC is best forgotten.

It should be stressed that this article is not devoted to identifying perverse situations in which CLC might break down. Rather, a straightforward and explicit formulation of causal effects in panel data showed that CLC is not a sound basis for causal inference over a wide range of plausible situations. Possibly, CLC could be patched up, primarily through additional restrictive assumptions, in response to the deficiencies demonstrated in this article. This article is not intended to stimulate such activities. CLC should be set aside as a dead end.

In some sense this article is a flight of fantasy. The fantasy is the notion of a closed two-variable causal system on which the exposition and results are based. This simple formulation serves well for investigating the worth of CLC. Also, almost all applications and technical development of CLC have been limited to two-variable causal systems. CLC fails even in this idealized situation, and no grounds for optimism exist for better performance in more complex causal systems.

In his articles on longitudinal panel data, Duncan (1969, 1972, 1975) stressed that the analysis of panel data cannot be reduced to a mechanical procedure that yields trustworthy inferences about causal structures. A minimal requirement for success is the careful formulation of explicit (and often specialized) models for the substantive processes. An intent of this article is to emphasize this message. Trying to answer a causal question from a set of (longitudinal) data is asking a lot from those data. Minimal requirements are that the right variables be measured well. Often the state of theoretical and empirical knowledge in a substantive area is not sufficiently advanced that the relevant variables have been identified or that sufficient measurement techniques have been developed.

Alternative methods for analyzing panel data were not endorsed or discussed in detail. This omission reflects the fact that methods for detecting patterns of causal influence from panel data are far from fully developed. The contribution of this article is to demonstrate that CLC certainly is not the method to rely on for the analysis of panel data. It is hoped that this article will direct efforts away from further development and application of CLC and toward the development and evaluation of productive approaches for the analysis of panel data.

It seems appropriate to make the point that social scientists frequently have been attracted to methods for the analysis of nonexperimental data that are far more flawed and less justified than path analysis and relatives. A most vivid example is provided by the method of cross-lagged correlation, which remains a very popular procedure for the analysis of reciprocal effects from nonexperimental, longitudinal data. Cross-lagged correlation purports to answer the question—Does X cause Y or does Y cause X?—by a simple comparison of the lagged correlations between X and Y (i.e., the correlations between X_1 and Y_2 and Y_1 and X_2 for two time points). A remarkable attribution of as-if-by-experiment is provided by Crano and Mellon (1978): "With the introduction of the cross-lagged panel correlation method . . . , causal inferences based on correlational data obtained in longitudinal studies can be made and enjoy the same logical status as those

From: Casual Models.....

derived in the more standard experimental settings" (p. 41). In other words, the use of cross-lagged correlation dispenses with the need for experiments, statistical models, or careful data analysis; a quick comparison of a few correlation coefficients is all that's required to study reciprocal effects.

Rogosa (1980) was only one in a tradition of papers, starting with Duncan (1969), Goldberger (1971), and Heise (1970), sharply critical of cross-lagged correlation. Even Cook and Campbell (1979, Chap. 7) are unenthusiastic about the usefulness of cross-lagged correlation; yet most advocates and users of this procedure remain undaunted. Rogosa (1980) exposits a number of simple statistical models for reciprocal effects between two variables—path analysis models, continuous-time feedback models, and multiple time series models. The mathematical results in Rogosa (1980) demonstrate the inability of the method of cross-lagged correlation to recover the structure of the reciprocal effects specified by these models. Results and numerical examples are presented for two-wave and multiwave data. Rogosa (1985) provides a nontechnical overview and extensive references on approaches to the analysis of reciprocal effects.

The mathematical and numerical demonstrations of the failures of cross-lagged correlation in Rogosa (1980) had the following simple, limited structure. Start with a basic path-analysis regression model for two variables X and Y measured at times 1 and 2 (the popular two-wave, two-variable panel design)

$$X_2 = \beta_0 + \beta_1 X_1 + \gamma_2 Y_1 + u,$$
$$Y_2 = \gamma_0 + \beta_2 X_1 + \gamma_1 Y_1 + v. \qquad (4)$$

In the context of the statistical model in (4) the parameters β_1 and γ_1 represent the influence of a variable on itself over time. The parameters β_2 and γ_2 represent the lagged, reciprocal effects between X and Y; thus the relative magnitudes of β_2 and γ_2 are presumed to indicate the nature of the reciprocal causal effects. In Rogosa (1980) combinations of β_2 and γ_2 values are compared with the results of the method of cross-lagged correlation. The major (and perhaps only) virtue of the path analysis model (4) is the identification of specific parameters believed to represent the reciprocal effects. If this model of the reciprocal influences between X and Y were valid, then estimation of β_2 and γ_2 would inform about reciprocal effects. Perhaps the best way to think about (4) and the related structural regression models is that these comprise a simple statistical model for reciprocal effects that, however, may be a far-from-satisfactory scientific model of the psychological (etc.) process.

The real moral about the analysis of reciprocal effects is that you can't estimate something without first defining it, and statistical models at least allow definition of key parameters. Regrettably, the seductive simplicity of cross-lagged correlation has inhibited serious work on the complex question of reciprocal effects. Despite the complexity of research questions about

reciprocal effects, empirical research has attempted to answer the over-simplified question, Does X cause Y or does Y cause X? by casually comparing a couple of correlations.

(1981), explains how misbehaving children have benefitted from these techniques. Step eight (never give up) says, "Always look for a better way." Let the client or student know you care what happens, and try as many plans as you must to establish better behaviors.

4. Reality Therapy Programs in Schools

Glasser's book, *Schools Without Failure,* presents an understandable way to apply these ideas to schools. Published in 1969, it was the largest selling American book on education during the 1970s. Enthusiasm from educators who were using these ideas helped Glasser to found the Educator Training Center in 1968, to teach how to use the concepts in schools.

Starting with educational films and inservice training programs, the Educator Training Center has expanded until the staff now help universities. For example, the University of Wisconsin at LaCrosse offers master's degrees on these concepts. At least 250,000 teachers in the United States have been specifically trained by the Educator Training Center staff and countless others have taken courses by those who have had this training. In translating reality therapy to schools, educators have recognized that the most important concepts to be learned are: (a) to base instructional programs on success practices—increase the student's success within the curriculum; (b) use the classroom meeting as a device for involvement between teacher and student, and student and student. The class and teacher sit in a circle to communicate in a nonjudgmental way and discuss values, goals of the class, how to live together successfully in the classroom, and anything else that can increase involvement and help students to know someone cares. Use of these meetings shows significant improvement in behavior and increase in learning. This process is best described in the book *Schools Without Failure* (1969). In *Focus on Guidance*, Thompson and Cates (1976) quote a study in Tennessee showing specific data for frequency of inappropriate behavior and how it decreases as the classroom teacher uses these methods. Along with this comes increased learning.

The Educator Training Center has developed programs which help school personnel discipline students and involve them in their learning in a way that helps them to accept responsibility for all their behavior. A description of how this can be done can be studied in a pamphlet called "*Glasser's Approach to Discipline*" (Educator Training Center 1977). It is a 10-step program. The first 3 steps use the process of reality therapy to look at how one deals with disruptive students. Steps 4 to 7 are specific ways to deal with children having difficulty, and the last 3 steps bring in other resources within the school and/or community. Schools which have applied this program report decreases in suspensions by 50–80 percent in junior and senior high schools, and vandalism by 40–90 percent. Improvement in teacher morale and professional growth were also cited as being significant gains by schools using this program. Administrators gained time to be the educational leaders they were meant to be. The implementation of this program can be further understood by reading a book by a principal, Bill Borgers, *Return to Discipline* (1979), who applied it to the students, teachers, and counselors in his school. He found that by having everyone understand how to help students take responsibility for their own behaviors, schools could function in a way that produced more learning.

Another book, *Loneliness in the Schools* by Marc Roberts (1973) shows how to use the schools-without-failure principles to establish more effective interaction within schools. His plan to encourage humaneness and growth within the school environment was created out of his concern for the loneliness of people who spend many hours each day in school.

Although the Educator Training Center works throughout the United States, its main office is in Long Beach, California.

See also: Counseling Theories

Bibliography

Borgers W 1979 *Return to Discipline.* Privately printed, Austin, Texas

Educator Training Center 1977 *Glasser's Approach to Discipline.* Long Beach, California

Glasser N 1980 *What Are You Doing? How People are Helped Through Reality Therapy: Cases.* Harper and Row, New York

Glasser W 1965 *Reality Therapy: A New Approach to Psychiatry.* Harper and Row, New York

Glasser W 1969 *Schools Without Failure.* Harper and Row, New York

Glasser W 1981 *Stations of the Mind: New Directions for Reality Therapy.* Harper and Row, New York

Medick J M 1981 *Classroom Misbehavior: Turning It Around.* Fanning, East Lansing, Michigan

Roberts M 1973 *Loneliness in the Schools.* Argus Communications, Niles, Illinois

Thompson C L, Cates J T 1976 Teaching Discipline to Students: An Individualized Teaching–Counseling Approach. *Focus on Guid.* 9(4)

W. Glasser: N. Glasser

Reciprocal Effects, Analysis of

The analysis of reciprocal effects consists of the study of mutual influences between two (or more) variables. Often, researchers attribute causality to the influences among variables and speak of reciprocal causal effects. Examples from educational research include analyses of the reciprocal effects between: teacher expectations and student achievement, student self-concept and academic achieve-

ment, and vocabulary and comprehension skills. Also, among the many studies of reciprocal effects in psychology are analyses of mother–child interactions, infant intelligence and behavior, and aggressive behavior and viewing of violent television shows. Often, research questions about reciprocal effects have been simplified to, does X cause Y or does Y cause X? Consequently, many researchers have sought only determinations of causal predominance or of *the* causal ordering of the variables.

Longitudinal data are crucial to the analysis of reciprocal effects, because the temporal ordering is needed to unravel the influences linking the variables. Typically, observations on a large number of cases are obtained for each variable on a few (two or more) occasions. (Each case may be an individual or a unit such as a student–teacher or mother–child dyad.) Much of the empirical research and methodological discussion on reciprocal effects has been limited to the two-wave, two variable (2W2V) panel design, in which measures of X and Y are available on each of two occasions

Four different approaches and statistical methods for quantitative data have been used in analyses of reciprocal effects. These methods are cross-lagged correlation, structural regression models, continuous-time feedback models, and multiple time series. Separate methods for dichotomous and categorical data, which have not been much used in educational research, are Lazarsfeld's 16-fold table, log-linear models for contingency tables, and continuous-time Markov models (see *Structural Equation Models*; *Log-linear Models*; *Contingency Tables*).

1. Cross-lagged Correlation

Cross-lagged correlation has been the most popular procedure in educational and psychological research for the analysis of reciprocal effects. Most often, cross-lagged correlation is used to determine a predominant causal influence. Figure 1, which presents the population correlations among the variables in a 2W2V panel design, is the diagram that accompanies expositions of cross-lagged correlation. The population cross-lagged correlations are $\rho_{X_1 Y_2}$ and $\rho_{Y_1 X_2}$.

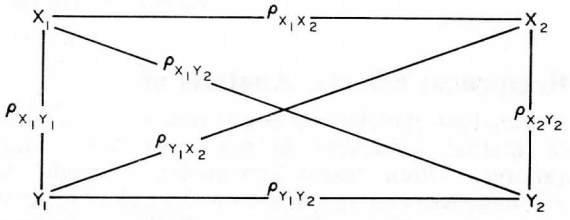


Figure 1
Population correlations for a 2W2V panel. The variables at time 1 are X_1 and Y_1 and the variables at time 2 are X_2 and Y_2

The difference between the population cross-lagged correlations, $\rho_{X_1 Y_2} - \rho_{Y_1 X_2}$, is the basis for attributions of a predominant causal influence. If the data indicate that $\rho_{X_1 Y_2} - \rho_{Y_1 X_2}$ is positive, the predominant causal influence is concluded to be in the direction of X causing Y. If the data indicate that $\rho_{X_1 Y_2} - \rho_{Y_1 X_2}$ is negative, the predominant causal influence is concluded to be in the direction of Y causing X. Usually, attributions of predominant causal influences are made only when the null hypothesis of equal population cross-lagged correlations is rejected. If this null hypothesis is not rejected, the usual interpretation is that no direct causal influences exist between X and Y; in particular that a common causal influence may be responsible for their observed associations. This interpretation has been adopted as a null hypothesis of spuriousness, which is represented through a model allowing no direct influences between X and Y but with an unmeasured third variable influencing both X and Y at each time (Kenny 1979 Fig. 12-2, Rogosa 1980 Fig. 3).

The extension of cross-lagged correlation to determinations of causal predominance when more than two waves of data are available is to compare the cross-lagged correlations from all possible two-wave combinations. Rogosa (1980) showed that this strategy of using the multiple waves for replication of differences between the cross-lagged correlations is more likely to generate confusion than corroboration.

A related statistical procedure for 2W2V panel data, which has seen a number of applications in educational research, is the frequency-in-change-in-product–moment procedure developed by Yee and Gage (1968). Although the Yee–Gage procedure differs from cross-lagged correlation in many important details, this procedure for the analysis of reciprocal effects suffers from the same basic deficiencies as cross-lagged correlation.

Cross-lagged correlation does not provide dependable information as to the causal structure underlying the data. Building upon earlier analyses by Duncan (1969) and Heise (1970), Rogosa (1980) demonstrated that, when there are no reciprocal causal effects, the difference between the cross-lagged correlations may be small or may be large; and when there are considerable reciprocal causal effects, the difference between the cross-lagged correlations may be small or may be large. Furthermore, a zero difference between the cross-lagged correlations (indicating spuriousness in cross-lagged correlation) is consistent with large reciprocal causal influences or with small or nonexistent causal influences between the variables. Moreover, cross-lagged correlation may indicate a causal predominance opposite to that of the actual causal structure of the data. Hence, neither determinations of spuriousness nor causal predominance can be trusted.

A basic deficiency in cross-lagged correlation is

the lack of an explicit definition of a causal effect. Without a clearly defined quantity to be estimated, it is not surprising that cross-lagged correlation fails to provide sound inferences. Also, the emphasis on causal predominance in cross-lagged correlation is unwise. The reciprocal nature of many social and educational processes makes determination of only causal predominance a serious oversimplification of the research problem. Measures of the strength and duration of the reciprocal relationship and of the specific causal effects are needed.

2. Structural Equation Models

Structural regression formulations of reciprocal effects in longitudinal panel data were originally introduced in the path analysis literature (Wright 1960, Duncan 1969, Heise 1970). The term causal model is popular for describing both the path analysis and the more general structural equation models. In these models, a causal effect is represented by the change in an outcome variable that results from an increment to an antecedent variable. For two variables, X and Y, the reciprocal influences are represented by the regression parameters of the path from a prior X to a later Y and from a prior Y to a later X.

Previous formulations of regression models for panel data with reciprocal causal effects have focused on models for 2W2V data. Figure 2 represents a specific regression model for 2W2V data, given by the regression equations:

$$X_2 = \beta_0 + \beta_1 X_1 + \gamma_2 Y_1 + u \quad (1)$$

$$Y_2 = \gamma_0 + \beta_2 X_1 + \gamma_1 Y_1 + v \quad (2)$$

The parameters β_2 and γ_2 represent the lagged, reciprocal causal effects between X and Y and thus are of central importance in the investigation of reciprocal causal effects.

Although the structural regressions do provide a model that defines reciprocal effects among the variables, the validity of inferences about the reciprocal effects depends crucially on the validity of the model. Foremost among the important assumptions built into Fig. 2 and Eqns. (1) and (2) is that X and Y constitute a closed system so that no important influences have been omitted from the regression model. Also important is the assumption that all causal influences are lagged; simultaneous causal influences between X and Y are not included. With only 2W2V data, frequently it is not possible to distinguish between different underlying models, which makes the determination of reciprocal effects very difficult. Additional observations can aid the formulation and testing of the regression models; one example of the use of regression models for three waves of data is the analysis of the influences between economic development and educational expansion in Hannan et al. (1974).

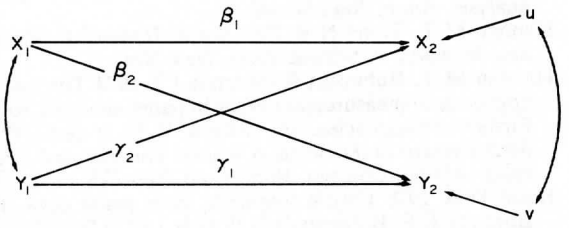


Figure 2
Representation of a regression (path analysis) model for 2W2V data

A generalization of these path analysis models is to specify X and Y to be latent variables having multiple indicators at each time point. The influences among the variables are represented by the parameters of the structural regression equations that relate the latent variables. (A measurement model connects the latent variables with their indicators.) For example, recasting Eqns. (1) and (2) in terms of latent variables, β_2 and γ_2 would then represent the reciprocal effects between the latent variables X and Y. Examples of the use of structural equation models for the analysis of reciprocal effects are: the analysis of attitudes and behaviors in Bentler and Speckart (1981), the analysis of intellectual flexibility and complexity of work in Kohn and Schooler (1978), and models for home environment and intellectual development in Rogosa (1979).

3. Continuous-time Feedback Models

An alternative formulation of reciprocal effects is to model the rates of change of the variables. A simple two variable continuous-time model that incorporates reciprocal influences between X and Y is:

$$\frac{dX(t)}{dt} = b_0 + b_1 X(t) + c_2 Y(t) \quad (3)$$

$$\frac{dY(t)}{dt} = c_0 + c_1 Y(t) + b_2 X(t) \quad (4)$$

Equations (3) and (4) are coupled differential equations which stipulate that the rates of change of X and Y at any time depend linearly on the levels of X and Y. The parameters b_2 and c_2 represent the cross effects or couplings between X and Y. Note that Eqns. (3) and (4) are deterministic. Similar models for change can be formulated which include stochastic components, exogenous variables, and other generalizations. Many applications of these models are presented in Hannan and Tuma (1983).

Rates of change are not directly observable. However, the solution of the system of differential equations in Eqns. (3) and (4) yields equations, in terms of the observable variables, of the same form as Eqns. (1) and (2). The parameters β_2 and γ_2 in

Eqns. (1) and (2) are nonlinear functions of the time between observations and the parameters of Eqns. (3) and (4). That the solutions of Eqns. (3) and (4) have the same form as the regression model for 2W2V data allows models for 2W2V data to be thought of as reflecting a more general process, that of causal influences and resulting adjustments that are continuous in time.

4. Multiple Time Series

The statistical analysis of reciprocal effects, or feedback, between two time series has been an active area in econometrics, with most applications investigating reciprocal effects between money supply and income or those between advertising and sales. The statistical methods are based on predictability criteria. Loosely speaking, one time series, $X(t)$, is said to cause another time series, $Y(t)$, if present Y values can be predicted better using past values of X than by not using past values of X, other relevant information (including past values of Y) being included in the prediction. This definition encompasses both lagged and instantaneous influences between X and Y. A comprehensive classification of the possible patterns of causal influences was provided by Pierce and Haugh (1977) who also presented data analysis procedures based on the correlations of residuals between the separately filtered time series for detecting these reciprocal effects. In addition, useful measures of linear dependence and feedback among multiple time series were developed by Geweke (1982).

In most economic research on reciprocal effects, the data consist of a single extensive time series of observations for each variable. The minimum number of observations over time needed for the application of time series statistical models is far beyond the usual design of longitudinal research in education. The longitudinal data for which methods commonly used in education (cross-lagged correlation, structural regression, and continuous-time feedback models) are applicable consist of a collection of many replications of very short time series (often with only two observations). Of course, such limited temporal data cannot support the complex time-series models used in the econometric analyses of reciprocal effects. One psychological application of models and analyses of reciprocal effects in time series data is the analysis of play behavior of individual mother–infant dyads in Gottman and Ringland (1981).

5. Methods for Categorical Data

Analyses of reciprocal effects using dichotomous or polychotomous variables have not been common in educational research. Lazarsfeld's 16-fold table for analyzing reciprocal effects among dichotomous variables is the best-known method for categorical data; Lazarsfeld (1978) provided a history of the development and application of this procedure. A natural extension of the 16-fold table analysis is the application of log-linear models for contingency tables (Goodman 1973). A third approach to the analysis of reciprocal effects is the use of continuous-time stochastic models, in particular, discrete-state, continuous-time Markov models. For additional references and for applications of these methods, see Coleman (1968), Hannan and Tuma (1979, 1983), and Markus (1979).

6. Conclusion

The investigation of reciprocal effects is an extremely difficult enterprise. Questions about reciprocal effects are some of the most complex in the design and analysis of longitudinal research. A humbling reality for research on reciprocal effects is that much simpler, preliminary research questions, namely those connected with the measurement of individual change and the assessment of correlates of change, remain controversial and unsolved. Clearly, reciprocal effects cannot be studied cheaply. Extensive, high-quality longitudinal data and theoretically based, explicit models of the reciprocal effects are absolutely necessary.

Bibliography

Bentler P M, Speckart G 1981 Attitudes "cause" behaviors: A structural equation analysis. *J. Pers. Soc. Psychol.* 40: 226–38

Coleman J S 1968 The mathematical study of change. In: Blalock H M, Blalock A B (eds.) 1968 *Methodology in Social Research.* McGraw-Hill, New York, pp. 428–78

Duncan O D 1969 Some linear models for two-wave, two-variable panel analysis. *Psychol. Bull.* 72: 177–82

Geweke J 1982 The measurement of linear dependence and feedback between multiple time series. *J. Am. Stat. Ass.* 77: 304–24

Goodman L A 1973 Causal analysis of data from panel studies and other kinds of surveys. *Am. J. Sociol.* 78: 1135–91

Gottman J M, Ringland J T 1981 The analysis of dominance and bidirectionality in social development. *Child Dev.* 52: 393–412

Hannan M T, Tuma N B 1979 Methods for temporal analysis. *Annu. Rev. Sociol.* 5

Hannan M T, Tuma N B 1983 *Social Dynamics: Models and Methods.* Academic Press, New York

Hannan M T, Rubinson R, Warren J T 1974 The causal approach to measurement error in panel analysis: Some further contingencies. In: Blalock H M Jr (ed.) 1974 *Measurement in the Social Sciences: Theories and Strategies.* Aldine, Chicago, Illinois, pp. 293–323

Heise D R 1970 Causal inference from panel data. In: Borgatta E F, Bohrnstedt G W (eds.) 1970 *Sociological Methodology.* Jossey-Bass, San Francisco, California

Kenny D A 1979 *Correlation and Causality.* Wiley, New York

Kessler R C, Greenberg D F 1981 *Linear Panel Analysis: Models of Quantitative Change.* Academic Press, New York

Kohn M L, Schooler C 1978 The reciprocal effects of the substantive complexity of work and intellectual flexibility: A longitudinal assessment. *Am. J. Sociol.* 84: 24–52

Lazarsfeld P F 1978 Some episodes in the history of panel analysis. In: Kandel D B (ed.) 1978 *Longitudinal Research on Drug Use.* Wiley, New York, pp. 249–65

Markus G B 1979 *Analyzing Panel Data.* Sage, Beverly Hills, California

Pierce D A, Haugh L D 1977 Causality in temporal systems: Characterizations and a survey. *J. Econometrics* 5: 265–93

Rogosa D R 1979 Causal models in longitudinal research: Rationale, formulation, and interpretation. In: Nesselroade J R, Baltes P B (eds.) 1979 *Longitudinal Methodology in the Study of Behavior and Development.* Academic Press, New York, pp. 263–302

Rogosa D R 1980 A critique of cross-lagged correlation. *Psychol. Bull.* 88: 245–58

Wright S 1960 The treatment of reciprocal interaction, with or without lag, in path analysis. *Biometrics* 16: 423–45

Yee A H, Gage N L 1968 Techniques for estimating the source and direction of causal influence in panel data. *Psychol. Bull.* 70: 115–26

D. R. Rogosa

Manual for RSiena

Ruth M. Ripley
Tom A.B. Snijders
Paulina Preciado

University of Oxford: Department of Statistics; Nuffield College

May 18, 2012

Abstract

SIENA (for Simulation Investigation for Empirical Network Analysis) is a computer program that carries out the statistical estimation of models for the evolution of social networks according to the dynamic actor-oriented model of Snijders (2001, 2005), Snijders et al. (2007), and Snijders et al. (2010a). This is the manual for RSiena, which also may be called SIENA version 4, a contributed package to the statistical system R. The manual is based on the earlier manual for SIENA version 3, and also contains contributions written for that manual by Mark Huisman, Michael Schweinberger, and Christian Steglich.

The Distribution of Physical Activity in an After-school Friendship Network

WHAT'S KNOWN ON THIS SUBJECT: New, effective approaches to obesity prevention are urgently needed. Social network interventions warrant our attention. Social networks play a significant role in adult and adolescent obesity. The role of social networks in pediatric obesity has not been examined.

WHAT THIS STUDY ADDS: Afterschool friendship ties play a critical role in setting physical activity patterns in children as young as 5 to 12 years. Children's activity levels can be changed by the activity level of their social network during a 12-week afterschool program.

abstract

OBJECTIVE: To examine whether a child's friendship network in an afterschool program influences his/her physical activity.

METHODS: Three waves of data were collected from school-aged children participating in aftercare (n = 81; mean [SD] age, 7.96 [1.74] years; 40% African American, 39% white, and 19% Latino) a name generator survey was used to map each child's social network, and accelerometers were used to measure physical activity. We applied stochastic actor-based modeling for social networks and behavior.

RESULTS: Children did not form or dissolve friendships based on physical activity levels, but existing friendships heavily influenced children's level of physical activity. The strongest influence on the amount of time children spent in moderate-to-vigorous activity in the afterschool hours was the activity level of their immediate friends. Children consistently made adjustments to their activity levels of 10% or more to emulate the activity levels of their peers (odds ratio [OR] = 6.89, $P < .01$). Age (OR = 0.92, $P < .10$) and obesity status (OR = 0.66, $P < .10$) had marginally significant and relatively small direct effects on the activity. Gender had no direct effect on activity.

CONCLUSIONS: These results suggest that friendship ties play a critical role in setting physical activity patterns in children as young as 5 to 12 years. Children's activity levels can be increased, decreased, or stabilized depending on the activity level of their immediate social network during a 12-week afterschool program. Network-based interventions hold the potential to produce clinically significant changes to children's physical activity. *Pediatrics* 2012;129:1064–1071

AUTHORS: Sabina B. Gesell, PhD,[a] Eric Tesdahl, MS,[b] and Eileen Ruchman, BA[a]

[a]*Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee; and* [b]*Department of Human and Organizational Development, Peabody College, Vanderbilt University, Nashville, Tennessee*

KEY WORDS

obesity, physical activity, attitude, causality, child, humans, longitudinal studies, social behavior, social support, sociology, medical, friendships, social networks, stochastic actor-based modeling, social influence, peer effects, selection

ABBREVIATIONS

MVPA—moderate to vigorous physical activity
OR—odds ratio

This trial has been registered at clinicaltrials.gov (identifier NCT01063413).

www.pediatrics.org/cgi/doi/10.1542/peds.2011-2567

doi:10.1542/peds.2011-2567

Accepted for publication Feb 15, 2012

Address correspondence to Sabina B. Gesell, PhD, Research Assistant Professor, Vanderbilt University School of Medicine, Department of Pediatrics, Pediatric Clinical and Translational Research, 2146 Belcourt Ave 2nd Fl, Nashville, TN 37212.E-mail: sabina.gesell@vanderbilt.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2012 by the American Academy of Pediatrics

FINANCIAL DISCLOSURES: *The authors have indicated they have no financial relationships relevant to this article to disclose.*

FUNDING: This work was supported by pilot funds awarded to Dr Gesell from the Vanderbilt Institute for Obesity and Metabolism and the Vanderbilt Clinical and Translational Science Award (NCRR/NIH 1 UL1 RR024975, Bernard PI). Dr Gesell was supported by the American Heart Association Clinical Research Grant Program (09CRP2230246) and by K23HD064700-01A1 (NIH/NICHD). Funded by the National Institutes of Health (NIH).

Downloaded from pediatrics.aappublications.org at Stanford Univ Med Ctr on May 28, 2012

Siena in R: RSiena

The *SIENA* methods are available in *RSiena*, which is a package of the statistical system R. This replaces the older Windows-based *SIENA* version 3, which still is available **from the downloads page**, but no longer maintained. *RSiena* can be executed on all platforms for which R is available: Windows, Mac, Unix/Linux. R can be downloaded from **http://cran.r-project.org**, and *RSiena* can be loaded as one of the packages in R.

The development of *RSiena* is part of the project *Adolescent Peer Social Network Dynamics and Problem Behavior*, funded by NIH (Grant Number 1R01HD052887-01A2), Principal Investigator John M. Light (Oregon Research Institute). *RSiena* was programmed by Ruth Ripley and Krists Boitmanis, in collaboration with Tom Snijders. Several others contributed, among whom is Josh Lospinoso. Ruth, Josh, and Tom, together with others, are continuing with the extension of the package. The fact that R is open source implies that the source is available (most easily from R-Forge, see below) and the project is open to contributions by other researchers who know how to program in R and/or C/C++.

RSiena can be operated in several ways. The best (most complete) functionality is obtained by executing *RSiena* entirely within R, using the R functions supplied by this package.
Those who have no prior knowledge of R may find it easy to execute *RSiena* through the graphical user interface which allows the user to execute *RSiena* without prior knowledge of R, or to access *RSiena* from **visone**; a short guide is available **at the visone website**.

The installation and operation of *RSiena* are explained in the **users' manual**. It includes information on how to install *RSiena*; how to work with the *Siena* graphical user interface; and some help on getting started with operating *RSiena* within R. In addition, it gives extensive descriptions of the methods available in *RSiena*.

RSiena and its manual are being updated rather frequently. New versions of the manual are available from this page.

Downloads / links

- **RSienaTest_1.1-213.zip**, the 18-05-2012 version of RSienaTest (installer for Windows), which currently is not available on R-Forge. Compiled under R 2.15.0. A further update is expected soon.
  Windows users can use this in R to "install from local zip file".
  The source code is in the associated **tar ball** (which can be unzipped).
- Some websites and resources that can be very helpful for the beginning R user:
  1. ***Some Hints for the R Beginner* by Patrick Burns** with the memorable quote "I asked R users what their biggest stumbling blocks were in learning R. A common answer that I was quite surprised by was that the biggest stumbling block was thinking that R was hard".
  2. **Excellent simple intro to R from Princeton**.
  3. **Getting started with R** by J. Gardner.
  4. **The official R intro**

5. **A portal with pointers to other R resources** .
6. **R Starter kit from UCLA**.
7. **Quick R** website (intended especially for experienced users of other statistical software).

- Further general help on R:
  - **Course on R and R programming by Ruth Ripley** with **extra resources**.
  - The "Documentation" links on **http://cran.r-project.org**
  - **R reference card**.
  - **RStudio**, a new integrated development environment (IDE) for R, available for all major platforms (like g Windows, Mac OS X, and Linux). This is an alternative to the R gui available for Windows.
  - An **introductory R script** to learn R for Social Networks applications, using the **Lazega lawyers data** , which can also be directed downloaded from **this directory with data files**.
  - **Introduction to Social Network Analysis in R** by Michal Bojanowski (ICM, University of Warsaw).
- **The Siena 4 manual**.
- **The RSiena project at R-Forge**. To allow quick access to new versions of *RSiena* without bothering too much the maintainers of the R archive, current and beta versions will be made available at R-Forge.
  The R-Forge page of RSiena contains two packages: the regular package RSiena and the 'experimental' package RSienaTest, which may contain elements that are less extensively tested.
  Note that you can subscribe to RSS feeds from R-Forge.
  Also note that in order to successfully install the packages provided on R-Forge, you have to switch to the most recent version of R (2.15).
  At R-forge there is an **email discussion list for help concerning RSiena programming**.
- **The RSiena information at CRAN** (the *Comprehensive R Archive Network*).
- *RSiena* can also be accessed from **visone**; a short guide is available **at the visone website**.

Back to the main Siena page

Although obesity has stabilized in some US populations, it is still at epidemic proportions with >23 million children classified as overweight or obese.[1,2] The cascade of adverse health effects associated with childhood overweight and obesity is well established and includes type 2 diabetes, hypertension, hyperlipidemia, abnormal glucose tolerance, cardiovascular disease, and psychosocial problems such as weight prejudice, depression, social isolation, poor self-esteem, and poor academic performance.[3–5] Given the resistant nature of obesity once established, prevention efforts must start early in life. Antiobesity interventions have generally failed.[6–8] New, effective approaches to obesity prevention are urgently needed.

An innovative approach to public health interventions has been proposed by scientists studying social networks. Social networks, the "thick webs of social relations and interactions"[9] that connect individuals to one another, exert measurable influence on our health.[10–20] Several independent research teams have linked social networks to obesity in adults and adolescents.[12,14,20–22] In contrast, there are limited data on social networks and obesity in children.

The extent to which healthy lifestyle behaviors can be influenced by children's social networks is unknown. If healthy lifestyle behaviors can be facilitated through children's social ties, such that when one child engages in physical activity, those children with whom he or she is connected will be more likely to engage in physical activity, then we could develop novel intervention strategies that leverage the social influences of social networks to make a real impact on childhood obesity. Specifically, to deem physical activity a candidate for network-based intervention strategies, we would need evidence of not just individuals clustering (ie, choosing to form or dissolve social ties) based on physical activity levels, but evidence of changes in their behaviors (ie, being more or less active) because of influence from their social ties.

One promising model for research on this topic is derived from studies on the coevolution of adolescent social networks and the use of alcohol and tobacco.[23–26] Through the use of newly developed stochastic actor-based models, these studies have been able to disentangle the complicated interplay between friendship selection (along lines of visible health behaviors) and behavioral influence (due to social ties). In this study, we extended such actor-based models to the study of the coevolution of friendship networks and physical activity. Thus, the purpose of this article is to test two research questions:

Research Question 1: Do children form or dissolve friendships based on physical activity levels? (Selection Effect)

Research Question 2: Is there child-to-child influence on physical activity behaviors over time? (Influence Effect)

METHODS

Study Population and Design

Public school students in 2 structured aftercare programs were invited to participate in a study to evaluate the routine physical activity levels of children in aftercare programs. Inclusion criteria included the following: (1) child age ≥5 years; (2) enrolled in school; (3) parental permission to copy official school records. The aftercare programs followed similar formats, operated from 3:00 to 6:00 PM Monday through Friday, and included time for snacks, homework, and play (eg, playground, gymnasium, art, board games). One aftercare program was based in a school. The other was based in a community center and drew children from the same school that the school-based program was located in, and from other schools in the same district, as well. Baseline data were obtained from 83 children, of whom 81 provided useable social network data.* The study was approved by the Vanderbilt University Institutional Review Board (IRB#090986).

Data Collection

All data were collected by trained study staff during the normal operating hours of the afterschool programs. All data were collected at 3 points over a school semester (February through May 2010), with 6 weeks separating each wave of measurement. The measurement period was guided by the Cochrane Review that pediatric obesity prevention interventions should be at least 12 weeks in duration for behavior change to occur.[27]

Measures

Physical Activity

Physical activity was assessed by using the ActiGraphGT1M accelerometer (ActiGraph LLC, Pensacola, FL). The ActiGraph is a small, lightweight monitor that is worn on a belt around the waist and measures the intensity of physical activity associated with locomotion. Monitors were programmed to record in continuous 10-second epochs to capture the short, spurtlike activity characteristic of children. Accelerometry has been used successfully in studies with children,[28–32] including Latino and African American children,[28–32] with high reliability, $r = 0.93$.[33] For each of the 3 waves of measurement, children wore monitors for 5 consecutive days from the time they signed into the afterschool program until

*Because of the nature of the statistical methods used in this study (stochastic actor-based modeling) pre hoc power analysis in the traditional sense of the term was not directly applicable. Given 3 measurement waves of behavioral data on 81 participants and $N \times (N - 1) \times 3$ social network data points, we were confident at the outset of the study that sufficient power would be achieved. See ref 26.

Downloaded from pediatrics.aappublications.org at Stanford Univ Med Ctr on May 28, 2012

they were picked up. The monitors did not provide the wearer with any feedback. Every child wore an accelerometer for a minimum of 60 minutes each day of measurement. Validated threshold values were used to derive time spent in sedentary, light, moderate, and vigorous activity.[34] The analysis of raw accelerometer data was performed by using a procedure similar to that used to analyze the NHANES data.[35]

Social Network

To map each child's social network, an open-ended survey was administered in private 1:1 interviews. Students were asked, "Please tell me the names of the friends you hang around with and talk to and do things with the most here in this after-school program." Children were allowed to nominate as many friends as they liked. This sociometric question (known as a name generator) is comparable to that used in other youth social network studies.[20] At waves 2 and 3, children were not shown the names of friends they had previously generated, but rather were asked to report on their friendships entirely anew to capture the friendships that were most salient to children at each time of measurement. The size of each child's network was reflected in both the number of friend nominations made, as well as the number of nominations received. To maximize the statistical power available for these analyses, sociometric data from each afterschool setting ($n = 46$ and $n = 35$ respectively) were combined into 1 single network of 81 nodes, where ties between children not attending the same afterschool program were coded as structurally impossible. By combining these networks for analysis, we made the assumption that the dynamics of friendship selection and friendship influence on physical activity behavior were consistent across these 2 aftercare settings, an assumption we found reasonable.

Anthropometrics

Body weight was measured, while wearing light clothing without shoes after voiding, to the nearest 0.1 kg on a calibrated digital scale (Detecto, Webb City, MO, model 758C). Body height without shoes was measured to the nearest 0.1 cm with the attached stadiometer. BMI percentile was calculated by using the Centers for Disease Control and Prevention calculator, where children with BMI ≥95th percentile were classified as obese.[36]

Demographics

Parents completed a demographic survey on the child's date of birth, gender, race/ethnicity, and school.

Statistical Analysis

Physical Activity Data Analysis

Physical activity was obtained from ActigraphGT1M accelerometer data collected in 10-second epochs. Time spent in physical activity intensities was based on activity counts: thresholds <420 activity counts for sedentary, 420 to 1679 counts for light, 1680 to 3379 counts for moderate, and >3379 counts for vigorous intensity activity per minute.[37] Time spent in sedentary behavior or physical activity was determined by summing minutes in a day where the count met the criterion for that intensity. Daily averages for each level of intensity were computed across all days of measurement. Start and stop times for play were recorded at each site and used as precise cutoff points. This captured the timeframe during which the children had control over their activity level, as opposed to the scheduled snack and homework times, which imposed sedentary behavior. Children spent varying amounts of time in aftercare depending on their family needs. Thus, our outcome measure was the proportion of play time spent in moderate-to-vigorous physical activity (MVPA), rather than absolute minutes in MVPA. Although our outcome measure was continuous, the social network analysis package that was used (SIENA)[38] requires categorical outcome variables for the analysis of social influence on behavior. Therefore, for analysis, we collapsed the data into deciles of percentage of play time spent in MVPA: 0% to 9%, 10% to 19%, 20% to 29%, etc, to 100%.

Social Network Data Analysis

Longitudinal analysis of physical activity and social network data were performed by means of stochastic actor-based modeling as implemented in SIENA† version 4.0.[38] Actor-based modeling assumes that changes observed over time in social ties and individual behaviors are the result of actors' decisions to optimize their position in the network at a given point in time. Observed social networks and behavior, such as can be collected in panel data, are assumed to be outcomes of an underlying continuous-time Markov process. That is to say, between each of the successive observed states of network and behavior captured in panel data, a number of smaller changes occur that are unobserved.[39] The estimation algorithm of SIENA works to arrive at the series of microchanges between observed waves of data that are most likely to have occurred when actors change their social ties or behavior. The rules governing such microchanges take the form of model parameters. Complete descriptions of this method are given by Snijders, Steglich, and colleagues.[39–41]

This approach affords major advantages over those based in classic regression analysis. The first of these is that stochastic actor-based methods allow researchers to directly model the complex dependencies between actors (ie, individuals) in a network. As such, this approach allows for the simultaneous examination of network structural

†SIENA is an acronym that stands for Simulation Investigation for Empirical Network Analysis.

outcome: Prevalence, somewhat continuous time sampling

Downloaded from pediatrics.aappublications.org at Stanford Univ Med Ctr on May 28, 2012

Alternating Renewal Process Models for the Single Behavior Stream

Independent r.v.'s d^0 d^1

Event transition probability matrix $\begin{Bmatrix} 0 & 1 \\ 1 & 0 \end{Bmatrix}$

$d^0 \overset{iid}{\sim} g_0$ with mean μ_0

$d^1 \overset{iid}{\sim} g_1$ with mean μ_1

Corresponding parameters for behavioral measures

Incidence (Rate)
(rate of off-to-on transitions)

$$I = \frac{1}{\mu_0 + \mu_1}$$

Prevalence
(proportion time process "on")

$$\phi = \frac{\mu_1}{\mu_0 + \mu_1}$$

Duration
(Average length of "on bout")

$$\phi / I = \mu_1$$

Data Collection: Observation Schemes

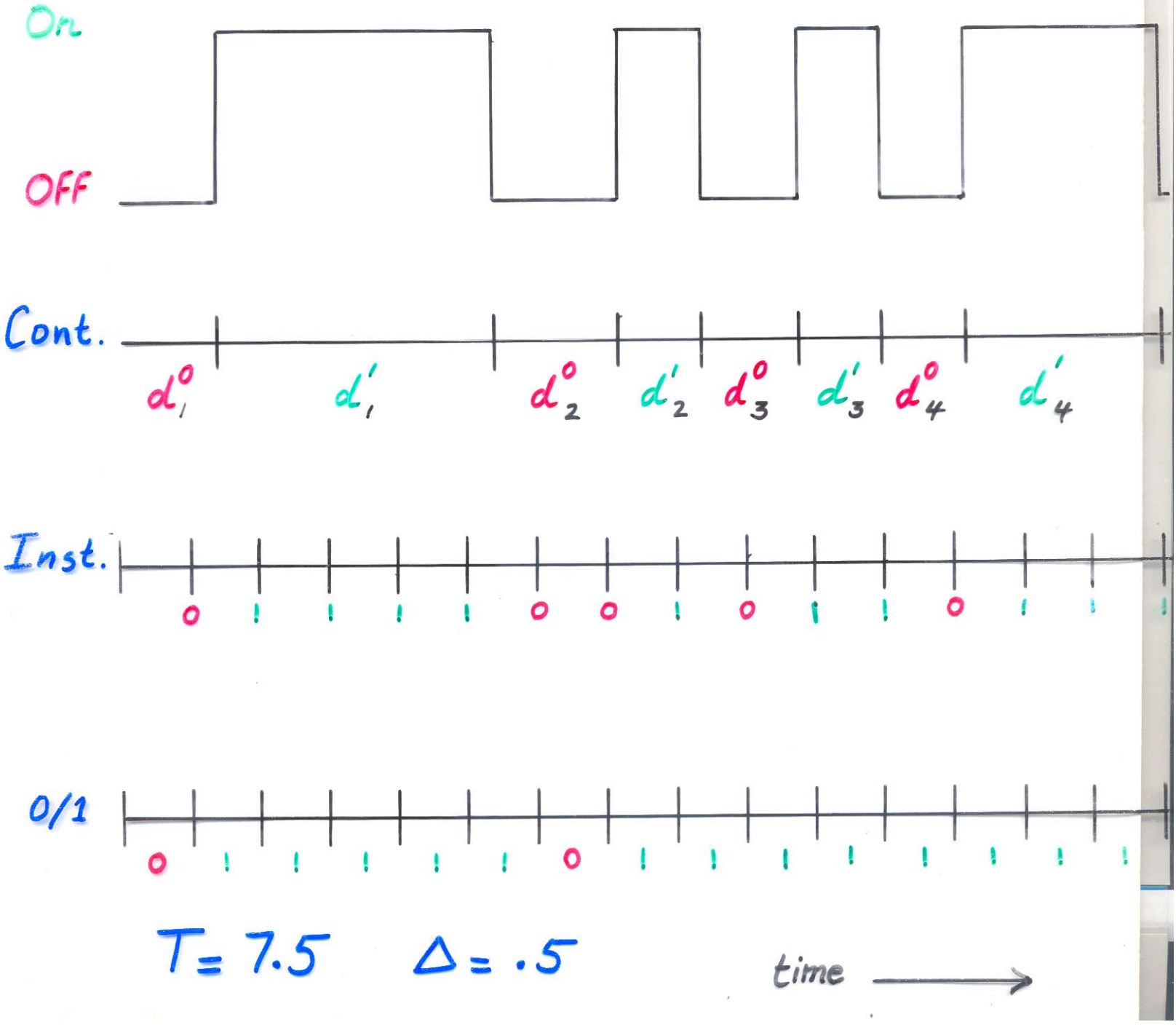
Continuous

Instantaneous (Point) sampling — snapshots Δ apart

Interval sampling / Hansen frequencies — continuous observation with discrete record: record no events or otherwise for each interval Δ.

SMBO determine statistical and psychometric properties of behavioral measures under alternative time-sampling.

Behavior Stream for a Specific Target Behavior



properties resulting from endogenous network effects such as transitive closure (the tendency for a friend of a friend to become a friend), dyadic effects (shared characteristics among members of a dyad), and exogenous effects (such as actor attributes).[40] Inclusion of such network structural properties in the modeling of socially mediated phenomena represents a notable advancement in methodological rigor in that such structural properties (ie, statistical nonindependence between cases), when neglected, have been shown to lead to biased results.[40,41] Of particular interest in this study is the ability of stochastic actor-based modeling procedures to disentangle the twin forces of selection and influence. Selection refers to factors observed in our data that may influence a child's friendship decision-making; these factors include structural properties such as those just named, but also may include actor attributes such as physical activity level, obesity status, gender, age similarity, etc. For each of these factors, the selection portion of our model will report on the likelihood of forming a tie with another child based on their particular attributes. Influence refers to factors observed in our data that may affect a child's behavior, in this case, physical activity level. In this study, we are primarily interested in understanding the effect of activity levels of friends upon the activity level of an individual child, but other direct effects on activity level are also included as controls. Changes in the activity level of a child's social network that precede changes to a child's activity level are the primary source of variability upon which this portion of the model is based. We describe our modeling in further detail in the Supplemental Information.

RESULTS

Sample Demographics

The sample (n = 81) was 40% African American, 39% white; 19% Latino; 65.4% female; 56% healthy weight, 23% overweight, 21% obese; and averaged 7.96 years of age (SD = 1.74).

Changes in Physical Activity

Table 1 describes changes in children's activity levels. The number of children who increased and decreased their physical activity was roughly equivalent; slightly more students maintained their activity levels. Table 2 reports the percentage of play time spent in MVPA.

Changes in Social Network Ties

Table 3 conveys changes in friendship status for all dyads. Children both formed and dissolved friendships. As is often observed in recently formed social networks, the amount of change over the first measurement period was greater than that of the subsequent period. This was reflected in the lower Jaccard coefficient (a measure of similarity in networks across time) for the first measurement period. Both periods satisfied the recommended guideline of Jaccard coefficient >0.30.[39] The overall network average degree (number of friendship ties) trended upward over the 3 waves of measurement, further illustrating that these friendship networks were evolving over the course of the study. On average, children nominated 3.7 (SD = 2.6) friends at wave 1, 5.25 (SD = 3.1) friends at wave 2, and 6.32 (SD = 3.5) friends at wave 3. Control variables are summarized in Table 4.

Results of SIENA Modeling

Results from the SIENA analysis are given in Table 5 and the Supplemental Information.

Selection Effect

The selection portion of the full model yielded large and highly significant parameters for rate of change in the friendship network, with changes in the first period (9.80, $P < .01$) greater than changes in the second period (4.95, $P < .01$). The outdegree parameter (odds ratio [OR] = 0.17, $P < .01$) indicated the general tendency away from having very large numbers of friendship ties in the network, which is a finding common to the vast majority of studies of this nature. The reciprocity parameter (OR = 1.99, $P < .01$) indicated the general tendency to reciprocate an incoming friendship tie. The transitive triplets term (OR = 1.32, $P < .01$), together with the term for 3 cycles (OR = 0.83, $P < .01$), indicated the general tendency for a friend of a friend to become a friend. Controls for similar age (OR = 2.40, $P < .01$), attending the same school (OR = 1.78, $P < .01$), being the same gender

TABLE 1 Changes in Physical Activity by Period

	Up	Down	Same	Missing
From t1 to t2	26	19	29	7
From t2 to t3	16	34	17	14

TABLE 2 Mean Percentage of Play Time Spent in Moderate and/or Vigorous Activity by Wave

	t1	t2	t3
Mean, %	30.35	31.12	29.40
SD	15.62	15.59	18.94

TABLE 3 Changes in Friendships for Each Dyad by Period

	Tie Change				Jaccard Coefficient
	0 → 0	0 → 1	1 → 0	1 → 1	
From t1 to t2	5927	253	128	172	0.311
From t2 to t3	5874	181	95	330	0.545

Jaccard Coefficient, measure of similarity in networks across time.

TABLE 4 Actor Attributes Used as Control Variables

	Mean/Proportion	SD
Female, %	65.40	—
Age	7.96	1.74
Weight category		
Normal or Overweight	79.00	—
Obese	21.00	—
Dyads same race, %	34.30	—
Dyads same school, %	28.50	—
Dyads same household, %	0.08	—

Downloaded from pediatrics.aappublications.org at Stanford Univ Med Ctr on May 28, 2012

19. Mednick SC, Christakis NA, Fowler JH. The spread of sleep loss influences drug use in adolescent social networks. *PLoS ONE*. 2010;5(3):e9775
20. de la Haye K, Robins G, Mohr P, Wilson C. Obesity-related behaviors in adolescent friendship networks. *Soc Networks*. 2010;32(3):161–167. Available at: www.sciencedirect.com/science/article/pii/S0378873309000495. Accessed March 26, 2012
21. Bahr DB, Browning RC, Wyatt HR, Hill JO. Exploiting social networks to mitigate the obesity epidemic. *Obesity (Silver Spring)*. 2009;17(4):723–728
22. de la Haye K, Robins G, Mohr P, Wilson C. Homophily and contagion as explanations for weight similarities among adolescent friends. *J Adolesc Health*. 2011;49(4):421–427
23. Mercken L, Snijders TA, Steglich C, de Vries H. Dynamics of adolescent friendship networks and smoking behavior: social network analyses in six European countries. *Soc Sci Med*. 2009;69(10):1506–1514
24. Mercken L, Snijders TA, Steglich C, Vertiainen E, de Vries H. Smoking-based selection and influence in gender-segregated friendship networks: a social network analysis of adolescent smoking. *Addiction*. 2010;105(7):1280–1289
25. Kiuru N, Burk WJ, Laursen B, Salmela-Aro K, Nurmi JE. Pressure to drink but not to smoke: disentangling selection and socialization in adolescent peer networks and peer groups. *J Adolesc*. 2010;33(6):801–812
26. Knecht AB, Burk WJ, Weesie J, Steglich C. Friendship and alcohol use in early adolescence: a multilevel social network approach. *J Res Adolesc*. 2011;21(2):475–487
27. Summerbell CD, Waters E, Edmunds LD, Kelly S, Brown T, Campbell KJ. Interventions for preventing obesity in children. *Cochrane Database Syst Rev*. 2005;3:CD001871
28. de Vries SI, Bakker I, Hopman-Rock M, Hirasing RA, van Mechelen W. Clinimetric review of motion sensors in children and adolescents. *J Clin Epidemiol*. 2006;59(7):670–680
29. Brage S, Wedderkopp N, Andersen LB, Froberg K. Influence of step frequency on movement intensity predictions with the CSA accelerometer: a field validation study in children. *Pediatr Exerc Sci*. 2003;15(3):277–287
30. Trost SG, Way R, Okely AD. Predictive validity of three ActiGraph energy expenditure equations for children. *Med Sci Sports Exerc*. 2006;38(2):380–387
31. Treuth MS, Schmitz K, Catellier DJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc*. 2004;36(7):1259–1266
32. Mattocks C, Leary S, Ness A, et al. Calibration of an accelerometer during free-living activities in children. *Int J Pediatr Obes*. 2007;2(4):218–226.
33. Puyau MR, Adolph AL, Vohra FA, Zakeri I, Butte NF. Prediction of activity energy expenditure using accelerometers in children. *Med Sci Sports Exerc*. 2004;36(9):1625–1631
34. Puyau MR, Adolph AL, Vohra FA, Butte NF. Validation and calibration of physical activity monitors in children. *Obes Res*. 2002;10(3):150–157
35. Tudor-Locke C, Johnson WD, Katzmarzyk PT. Accelerometer-determined steps per day in US adults. *Med Sci Sports Exerc*. 2009;41(7):1384–1391
36. Children's Hospital of Philadelphia. Body mass index and z-score calculation in children. 2010. Available at: http://stokes.chop.edu/web/zscore/. Accessed March 26, 2012
37. Pate RR, Almeida MJ, McIver KL, Pfeiffer KA, Dowda M. Validation and calibration of an accelerometer in preschool children. *Obesity (Silver Spring)*. 2006;14(11):2000–2006
38. Ripley R, Snijders T. Manual for SIENA version 4.0. 2011. Available at: www.stats.ox.ac.uk/~snijders/siena/s_man400.pdf. Accessed March 26, 2012
39. Snijders TAB, Van De Bunt GG, Steglich CEG. Introduction to stochastic actor-based models for network dynamics. *Soc Networks*. 2010;32(1):44–60
40. Snijders T, Steglich C, Schweinberger M. Modeling the co-evolution of networks and behavior. In: van Montfort K, Oud J, Satorra A, eds. *Longitudinal Models in the Behavioral and Related Sciences*. Mahwah, NJ: Lawrence Erlbaum; 2007
41. Steglich C, Snijders T, Pearson M. Dynamic networks and behavior: separating selection from influence. *Sociol Methodol*. 2010;40(1):329–393
42. de la Haye K, Robins G, Mohr P, Wilson C. How physical activity shapes, and is shaped by, adolescent friendships. *Soc Sci Med*. 2011;73(5):719–728
43. Chen KY, Bassett DR Jr. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc*. 2005;37(suppl 11):S490–S500
44. Beets MW, Rooney L, Tilley F, Beighle A, Webster C. Evaluation of policies to promote physical activity in afterschool programs: are we meeting current benchmarks? *Prev Med*. 2010;51(3-4):299–301
45. Beets MW, Wallner M, Beighle A. Defining standards and policies for promoting physical activity in afterschool programs. *J Sch Health*. 2010;80(8):411–417
46. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. *Med Sci Sports Exerc*. 2000;32(5):963–975
47. Van Der Horst K, Paw MJ, Twisk JW, Van Mechelen W. A brief review on correlates of physical activity and sedentariness in youth. *Med Sci Sports Exerc*. 2007;39(8):1241–1250

Downloaded from pediatrics.aappublications.org at Stanford Univ Med Ctr on May 28, 2012

Sociological Methods & Research
http://smr.sagepub.com/

Statistical Analysis of Longitudinal Network Data With Changing Composition

Mark Huisman and Tom A. B. Snijders
Sociological Methods & Research 2003 32: 253
DOI: 10.1177/0049124103256096

The online version of this article can be found at:
http://smr.sagepub.com/content/32/2/253

Published by:
SAGE
http://www.sagepublications.com

Additional services and information for *Sociological Methods & Research* can be found at:

Email Alerts: http://smr.sagepub.com/cgi/alerts

Subscriptions: http://smr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://smr.sagepub.com/content/32/2/253.refs.html

>> Version of Record - Nov 1, 2003

What is This?

Downloaded from smr.sagepub.com at Stanford University Libraries on May 31, 2012

3. The *gratification function*, indicating the instant gratification experienced by an actor i when changing the relation with an actor j, given the current state of the network $x(t) = x$:

$$g_i(\gamma, x, j), \quad i, j \in \mathcal{I}, \quad j \neq i, \quad x \in \mathcal{X}.$$

This function is motivated by the experience that a given effect may be stronger (or weaker) for initiating new relations than for breaking existing relations (Van de Bunt 1999 describes the gratification effect of initiating and breaking reciprocated relations). Like the objective function, the gratification function is a weighted sum of network and attribute effects:

$$g_i(\gamma, x, j) = \sum_{h=1}^{H} \gamma_h r_{ijh}(x), \tag{2}$$

with γ_h the statistical parameters indicating the strength of the effect $r_{ijh}(x)$ and H the number of included effects. The effects $r_{ijh}(x)$ always include a factor x_{ij} (reflecting the effect of breaking a relation) or $1 - x_{ij}$ (reflecting the effect of initiating a relation). See Snijders (2001) for examples of possible effects.

The choice made by an actor to perform an action is based on the function

$$r(\theta, i, j, x) = f_i(\beta, x(i \rightsquigarrow j)) + g_i(\gamma, x, j), \tag{3}$$

where $x(i \rightsquigarrow j)$ denotes the state of the network after actor i changed the outgoing relation to actor j: x_{ij} changed into $1 - x_{ij}$, with $i, j \in \mathcal{I}$ and $j \neq i$. The actor tries to optimize this function plus a random component; that is, actor i chooses to change his or her relation with that actor j such that

$$r(\theta, i, j, x) + U_i(t, x, j) \tag{4}$$

is maximized. The random variables $U_i(t, x, j)$ are assumed to be independent and identically distributed for all i, j, x, and t, having the Type I extreme value distribution with mean 0 and scale parameter 1. This assumption is commonly made in random utility modeling in econometrics (Maddala 1983). Given this distribution, the probability that a given actor i chooses actor j and changes the relation x_{ij} is

$$p_{ij}(\theta, x) = \frac{\exp(r(\theta, i, j, x))}{\sum_{l=1, l\neq i}^{n} \exp(r(\theta, i, l, x))}, \quad j \neq i. \tag{5}$$

Downloaded from smr.sagepub.com at Stanford University Libraries on May 31, 2012