Ivan Jacob Agaloos Pesigan

May 20, 2025

References

Andrews: Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space **Andrews-2000**

Donald W. K. Andrews. "Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space". In: *Econometrica* 68.2 (Mar. 2000), pp. 399–405. DOI: 10.1111/1468-0262.00114.

Arnett: The developmental context of substance use in emerging adulthood **Arnett-2005**

Jeffrey Jensen Arnett. "The developmental context of substance use in emerging adulthood". In: *Journal of Drug Issues* 35.2 (Apr. 2005), pp. 235–254. ISSN: 1945-1369. DOI: 10.1177/002204260503500202.

Abstract: The theory of emerging adulthood has been proposed as a way of conceptualizing the developmental characteristics of young people between the ages of 18 and 25. Here, the theory is applied to explaining the high rates of substance use in this age group. Specifically, five developmentally distinctive features of emerging adulthood are proposed: the age of identity explorations, the age of instability, the age of self-focus, the age of feeling in-between, and the age of possibilities. Then, each of these features is applied to an explanation of drug use in emerging adulthood.

Baker et al.: Addiction motivation reformulated: An affective processing model of negative reinforcement **Baker-Piper-McCarthy-etal-2004**

Timothy B. Baker et al. "Addiction motivation reformulated: An affective processing model of

negative reinforcement". In: *Psychological Review* 111.1 (2004), pp. 33–51. ISSN: 0033-295X. DOI: [10.1037/0033-295x.111.1.33](10.1037/0033-295x.111.1.33).

Abstract: This article offers a reformulation of the negative reinforcement model of drug addiction and proposes that the escape and avoidance of negative affect is the prepotent motive for addictive drug use. The authors posit that negative affect is the motivational core of the withdrawal syndrome and argue that, through repeated cycles of drug use and withdrawal, addicted organisms learn to detect interoceptive cues of negative affect preconsciously. Thus, the motivational basis of much drug use is opaque and tends not to reflect cognitive control. When either stressors or abstinence causes negative affect to grow and enter consciousness, increasing negative affect biases information processing in ways that promote renewed drug administration. After explicating their model, the authors address previous critiques of negative reinforcement models in light of their reformulation and review predictions generated by their model.

Bauer et al.: Probing interactions in fixed and multilevel regression: Inferential and graphical techniques   Bauer-Curran-2005

Daniel J. Bauer and Patrick J. Curran. "Probing interactions in fixed and multilevel regression: Inferential and graphical techniques". In: *Multivariate Behavioral Research* 40.3 (July 2005), pp. 373–400. ISSN: 1532-7906. DOI: [10.1207/s15327906mbr4003_5](10.1207/s15327906mbr4003_5).

Abstract: Many important research hypotheses concern conditional relations in which the effect of one predictor varies with the value of another. Such relations are commonly evaluated as multiplicative interactions and can be tested in both fixed- and random-effects regression. Often, these interactive effects must be further probed to fully explicate the nature of the conditional relation. The most common method for probing interactions is to test simple slopes at specific levels of the predictors. A more general method is the Johnson-Neyman (J-N) technique. This technique is not widely used, however, because it is currently limited to categorical by continuous interactions in fixed-effects regression and has yet to be extended to the broader class of random-effects regression

models. The goal of our article is to generalize the J-N technique to allow for tests of a variety of interactions that arise in both fixed- and random-effects regression. We review existing methods for probing interactions, explicate the analytic expressions needed to expand these tests to a wider set of conditions, and demonstrate the advantages of the J-N technique relative to simple slopes with three empirical examples.

Bauer et al.: Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations

Bauer-Preacher-Gil-2006

Daniel J. Bauer, Kristopher J. Preacher, and Karen M. Gil. "Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations". In: *Psychological Methods* 11.2 (2006), pp. 142–163. DOI: 10.1037/1082-989x.11.2.142.

Bentler: Can scientifically useful hypotheses be tested with correlations?

Bentler-2007

Peter M. Bentler. "Can scientifically useful hypotheses be tested with correlations?" In: *American Psychologist* 62.8 (2007), pp. 772–782. ISSN: 0003-066X. DOI: 10.1037/0003-066x.62.8.772.

Abstract: Historically, interesting psychological theories have been phrased in terms of correlation coefficients, which are standardized covariances, and various statistics derived from them. Methodological practice over the last 40 years, however, has suggested it is necessary to transform such theories into hypotheses on covariances and statistics derived from them. This complication turns out to be unnecessary, because the methodology now exists to test hypotheses on latent structures of correlations directly. Two examples are given. Limitations of correlation structures are also noted.

Beran: The impact of the bootstrap on statistical algorithms and theory   Beran-2003

Abstract: Bootstrap ideas yield remarkably effective algorithms for realizing certain programs in statistics. These include the construction of (possibly simultaneous) confidences sets and tests in classical models for which exact or asymptotic distribution theory is intractable. Success of the bootstrap, in the sense of doing what is expected under a probability model for data, is not universal. Modifications to Efron's definition of the bootstrap are needed to make the idea work for modern procedures that are not classically regular.

Boker: Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes   Boker-2002

Abstract: Notes that over 300 yrs ago Sir Isaac Newton wrote of a simple set of relations that could be used to predict the motions of objects relative to one another. The main advantage of this insight was that the relationship between the movements of the planets and stars could be predicted much more simply than with the accurate, but cumbersome Ptolemaic calculations. But perhaps the most important consequence of the acceptance of Newton's insight was that intrinsic properties such as mass could be distinguished from measurements such as weight. A similar revolution in thinking appears to be underway in the behavioral sciences. It is likely that intensive longitudinal measurement coupled with dynamical systems analyses will lead to simplified but powerful models of the evolution of psychological processes. In this case, it is reasonable to expect that a set of intrinsic psychological properties may be able to be extracted from the parameters of successful

dynamical systems models. The purpose of this article is to issue an invitation to the hunt, to provide a tentative map as to where the game might likely be found, and blow a call on the hunting horn.

Bolger et al.: Diary methods: Capturing life as it is lived   Bolger-Davis-Rafaeli-2003

Abstract: In diary studies, people provide frequent reports on the events and experiences of their daily lives. These reports capture the particulars of experience in a way that is not possible using traditional designs. We review the types of research questions that diary methods are best equipped to answer, the main designs that can be used, current technology for obtaining diary reports, and appropriate data analysis strategies. Major recent developments include the use of electronic forms of data collection and multilevel models in data analysis. We identify several areas of research opportunities: 1. in technology, combining electronic diary reports with collateral measures such as ambulatory heart rate; 2. in measurement, switching from measures based on between-person differences to those based on within-person changes; and 3. in research questions, using diaries to (a) explain why people differ in variability rather than mean level, (b) study change processes during major events and transitions, and (c) study interpersonal processes using dyadic and group diary methods.

Boos: Introduction to the bootstrap world   Boos-2003

Abstract: The bootstrap has made a fundamental impact on how we carry out statistical inference in problems without analytic solutions. This fact is illustrated with examples and comments that emphasize the parametric bootstrap and hypothesis testing.

Bühlmann: Bootstraps for time series   Buehlmann-2002

Peter Bühlmann. "Bootstraps for time series". In: *Statistical Science* 17.1 (May 2002). ISSN: 0883-4237. DOI: 10.1214/ss/1023798998.

Abstract: We review and compare block, sieve and local bootstraps for time series and thereby illuminate theoretical aspects of the procedures as well as their performance on finite-sample data. Our view is selective with the intention of providing a new and fair picture of some particular aspects of bootstrapping time series. The generality of the block bootstrap is contrasted with sieve bootstraps. We discuss implementational advantages and disadvantages. We argue that two types of sieve often outperform the block method, each of them in its own important niche, namely linear and categorical processes. Local bootstraps, designed for nonparametric smoothing problems, are easy to use and implement but exhibit in some cases low performance.

Casella: Introduction to the silver anniversary of the bootstrap   Casella-2003

George Casella. "Introduction to the silver anniversary of the bootstrap". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994967.

Cheong et al.: Investigation of mediational processes using parallel process latent growth curve modeling   Cheong-MacKinnon-Khoo-2003

JeeWon Cheong, David P. MacKinnon, and Siek Toon Khoo. "Investigation of mediational processes using parallel process latent growth curve modeling". In: *Structural Equation Modeling: A Multidisciplinary Journal* 10.2 (Apr. 2003), pp. 238–262. DOI: 10.1207/s15328007sem1002_5.

Abstract: This study investigated a method to evaluate mediational processes using latent growth curve modeling. The mediator and the outcome measured across multiple time points were viewed as 2 separate parallel processes. The mediational process was defined as the independent variable influencing the growth of the mediator, which, in turn, affected the growth of the outcome. To illustrate modeling procedures, empirical data from a longitudinal drug prevention program, Adolescents Training and Learning to Avoid Steroids, were used. The program effects on the growth of the mediator and the growth of the outcome were examined first in a 2-group structural equation model. The mediational process was then modeled and tested in a parallel process latent growth curve model by relating the prevention program condition, the growth rate factor of the mediator, and the growth rate factor of the outcome.

G. W. Cheung et al.: Testing mediation and suppression effects of latent variables
Cheung-Lau-2007

Abstract: Because of the importance of mediation studies, researchers have been continuously searching for the best statistical test for mediation effect. The approaches that have been most commonly employed include those that use zero-order and partial correlation, hierarchical regression models, and structural equation modeling (SEM). This study extends MacKinnon and colleagues (MacKinnon, Lockwood, Hoffmann, West, & Sheets, 2002; MacKinnon, Lockwood, & Williams, 2004, MacKinnon, Warsi, & Dwyer, 1995) works by conducting a simulation that examines the distribution of mediation and suppression effects of latent variables with SEM, and the properties of confidence intervals developed from eight different methods. Results show that SEM provides unbiased estimates of mediation and suppression effects, and that the bias-corrected bootstrap confidence intervals perform best in testing for mediation and suppression effects. Steps to implement the recommended procedures with Amos are presented.

M. W.-L. Cheung: Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models Cheung-2007

Abstract: Mediators are variables that explain the association between an independent variable and a dependent variable. Structural equation modeling (SEM) is widely used to test models with mediating effects. This article illustrates how to construct confidence intervals (CIs) of the mediating effects for a variety of models in SEM. Specifically, mediating models with 1 mediator, 2 intermediate mediators, 2 specific mediators, and 1 mediator in 2 independent groups are illustrated. By using phantom variables (Rindskopf, 1984), a Wald CI, percentile bootstrap CI, bias-corrected bootstrap CI, and a likelihood-based CI on the mediating effect are easily constructed with some existing SEM packages, such as LISREL, Mplus, and Mx. Monte Carlo simulation studies are used to compare the coverage probabilities of these CIs. The results show that the coverage probabilities of these CIs are comparable when the mediating effect is large or when the sample size is large. However, when the mediating effect and the sample size are both small, the bootstrap CI and likelihood-based CI are preferred over the Wald CI. Extensions of this SEM approach for future research are discussed.

M. W.-L. Cheung: A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling Cheung-2008

Abstract: Meta-analysis and structural equation modeling (SEM) are two important statistical methods in the behavioral, social, and medical sciences. They are generally treated as two unrelated topics in the literature. The present article proposes a model to integrate fixed-, random-, and mixed-

effects meta-analyses into the SEM framework. By applying an appropriate transformation on the data, studies in a meta-analysis can be analyzed as subjects in a structural equation model. This article also highlights some practical benefits of using the SEM approach to conduct a meta-analysis. Specifically, the SEM-based meta-analysis can be used to handle missing covariates, to quantify the heterogeneity of effect sizes, and to address the heterogeneity of effect sizes with mixture models. Examples are used to illustrate the equivalence between the conventional meta-analysis and the SEM-based meta-analysis. Future directions on and issues related to the SEM-based meta-analysis are discussed.

M. W.-L. Cheung: Comparison of methods for constructing confidence intervals of standardized indirect effects   Cheung-2009a

Abstract: Mediation models are often used as a means to explain the psychological mechanisms between an independent and a dependent variable in the behavioral and social sciences. A major limitation of the unstandardized indirect effect calculated from raw scores is that it cannot be interpreted as an effect-size measure. In contrast, the standardized indirect effect calculated from standardized scores can be a good candidate as a measure of effect size because it is scale invariant. In the present article, 11 methods for constructing the confidence intervals (CIs) of the standardized indirect effects were evaluated via a computer simulation. These included six Wald CIs, three bootstrap CIs, one likelihood-based CI, and the PRODCLIN CI. The results consistently showed that the percentile bootstrap, the bias-corrected bootstrap, and the likelihood-based approaches had the best coverage probability. Mplus, LISREL, and Mx syntax were included to facilitate the use of these preferred methods in applied settings. Future issues on the use of the standardized indirect effects are discussed.

Abstract: Confidence intervals (CIs) for parameters are usually constructed based on the estimated standard errors. These are known as Wald CIs. This article argues that likelihood-based CIs (CIs based on likelihood ratio statistics) are often preferred to Wald CIs. It shows how the likelihood-based CIs and the Wald CIs for many statistics and psychometric indexes can be constructed with the use of phantom variables (Rindskopf, 1984) in some of the current structural equation modeling (SEM) packages. The procedures to form CIs for the differences in correlation coefficients, squared multiple correlations, indirect effects, coefficient alphas, and reliability estimates are illustrated. A simulation study on the Pearson correlation is used to demonstrate the advantages of the likelihood-based CI over the Wald CI. Issues arising from this SEM approach and extensions of this approach are discussed.

Chow et al.: Age differences in dynamical emotion-cognition linkages

Chow-Hamagani-Nesselroade-2007

Abstract: The ability to maintain the separation between positive emotion and negative emotion in times of stress has been construed as a resilience mechanism. Emotional resiliency is particularly relevant in old age given concomitant declines in cognitive performance. In the present study, the authors examined the dynamical linkages among positive emotion, negative emotion, and cognition as individuals performed a complex cognitive task. Comparisons were made between younger (n

= 63) and older (n = 52) age groups. Older adults manifested significant unidirectional coupling from negative emotion to cognitive performance; younger adults manifested significant unidirectional coupling from negative emotion to positive emotion and from cognitive performance to both positive and negative emotions. Implications for age differences in emotion regulatory strategies are discussed.

Cole et al.: Empirical and conceptual problems with longitudinal trait-state models: Introducing a trait-state-occasion model   Cole-Martin-Steiger-2005

Abstract: The latent trait-state-error model (TSE) and the latent state-trait model with autoregression (LST-AR) represent creative structural equation methods for examining the longitudinal structure of psychological constructs. Application of these models has been somewhat limited by empirical or conceptual problems. In the present study, Monte Carlo analysis revealed that TSE models tend to generate improper solutions when N is too small, when waves are too few, and when occasion factor stability is either too large or too small. Mathematical analysis of the LST-AR model revealed its limitation to constructs that become more highly auto-correlated over time. The trait-state-occasion model has fewer empirical problems than does the TSE model and is more broadly applicable than is the LST-AR model.

Cole et al.: Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling.   Cole-Maxwell-2003

Abstract: R. M. Baron and D. A. Kenny (1986; see record 1987-13085-001) provided clarion conceptual and methodological guidelines for testing mediational models with cross-sectional data. Graduating from cross-sectional to longitudinal designs enables researchers to make more rigorous inferences about the causal relations implied by such models. In this transition, misconceptions and erroneous assumptions are the norm. First, we describe some of the questions that arise (and misconceptions that sometimes emerge) in longitudinal tests of mediational models. We also provide a collection of tips for structural equation modeling (SEM) of mediational processes. Finally, we suggest a series of 5 steps when using SEM to test mediational processes in longitudinal designs: testing the measurement model, testing for added components, testing for omitted paths, testing the stationarity assumption, and estimating the mediational effects.

Cribari-Neto: Asymptotic inference under heteroskedasticity of unknown form

CribariNeto-2004

Abstract: We focus on the finite-sample behavior of heteroskedasticity-consistent covariance matrix estimators and associated quasi-t tests. The estimator most commonly used is that proposed by Halbert White. Its finite-sample behavior under both homoskedasticity and heteroskedasticity is analyzed using Monte Carlo methods. We also consider two other consistent estimators, namely: the HC3 estimator, which is an approximation to the jackknife estimator, and the weighted bootstrap estimator. Additionally, we evaluate the finite-sample behavior of two bootstrap quasi-t tests: the test based on a single bootstrapping scheme and the test based on a double, nested bootstrapping scheme. The latter is very computer-intensive, but proves to work well in small samples. Finally, we propose a new estimator, which we call HC4; it is tailored to take into account the effect of leverage points in the design matrix on associated quasi-t tests.

Cribari-Neto et al.: A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model — CribariNeto-daSilva-2010

Abstract: The assumption that all random errors in the linear regression model share the same variance (homoskedasticity) is often violated in practice. The ordinary least squares estimator of the vector of regression parameters remains unbiased, consistent and asymptotically normal under unequal error variances. Many practitioners then choose to base their inferences on such an estimator. The usual practice is to couple it with an asymptotically valid estimation of its covariance matrix, and then carry out hypothesis tests that are valid under heteroskedasticity of unknown form. We use numerical integration methods to compute the exact null distributions of some quasi-t test statistics, and propose a new covariance matrix estimator. The numerical results favor testing inference based on the estimator we propose.

Cribari-Neto et al.: Inference under heteroskedasticity and leveraged data — CribariNeto-Souza-Vasconcellos-2007

Abstract: We evaluate the finite-sample behavior of different heteros-ke-das-ticity-consistent covariance matrix estimators, under both constant and unequal error variances. We consider the estimator proposed by Halbert White (HC0), and also its variants known as HC2, HC3, and HC4; the latter was recently proposed by Cribari-Neto (2004). We propose a new covariance matrix estimator: HC5. It is the first consistent estimator to explicitly take into account the effect that the maximal

leverage has on the associated inference. Our numerical results show that quasi-t inference based on HC5 is typically more reliable than inference based on other covariance matrix estimators.

Cribari-Neto et al.: Errata: Inference under heteroskedasticity and leveraged data, Communications in Statistics, Theory and Methods, 36, 1877–1888, 2007

CribariNeto-Souza-Vasconcellos-2008

Francisco Cribari-Neto, Tatiene C. Souza, and Klaus L. P. Vasconcellos. "Errata: Inference under heteroskedasticity and leveraged data, Communications in Statistics, Theory and Methods, 36, 1877–1888, 2007". In: *Communications in Statistics - Theory and Methods* 37.20 (Sept. 2008), pp. 3329–3330. DOI: 10.1080/03610920802109210.

Davison et al.: Recent developments in bootstrap methodology

Davison-Hinkley-Young-2003

Anthony Christopher Davison, David Victor Hinkley, and George Alastair Young. "Recent developments in bootstrap methodology". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994969.

Abstract: Ever since its introduction, the bootstrap has provided both a powerful set of solutions for practical statisticians, and a rich source of theoretical and methodological problems for statistics. In this article, some recent developments in bootstrap methodology are reviewed and discussed. After a brief introduction to the bootstrap, we consider the following topics at varying levels of detail: the use of bootstrapping for highly accurate parametric inference; theoretical properties of nonparametric bootstrapping with unequal probabilities; subsampling and the m out of n bootstrap; bootstrap failures and remedies for superefficient estimators; recent topics in significance testing; bootstrap improvements of unstable classifiers and resampling for dependent data. The treatment is telegraphic rather than exhaustive.

Bradley Efron. "Second thoughts on the bootstrap". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994968.

Abstract: This brief review article is appearing in the issue of Statistical Science that marks the 25th anniversary of the bootstrap. It concerns some of the theoretical and methodological aspects of the bootstrap and how they might influence future work in statistics.

Ernst et al.: Utilizing a quantile function approach to obtain exact bootstrap solutions
Ernst-Hutson-2003

Michael D. Ernst and Alan D. Hutson. "Utilizing a quantile function approach to obtain exact bootstrap solutions". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994978.

Abstract: The popularity of the bootstrap is due in part to its wide applicability and the ease of implementing resampling procedures on modern computers. But careful reading of Efron (1979) will show that at its heart, the bootstrap is a "plug-in" procedure that involves calculating a functional $\theta\left(\hat{F}\right)$ from an estimate of the c.d.f. F. Resampling becomes invaluable when, as is often the case, $\theta\left(\hat{F}\right)$ cannot be calculated explicitly. We discuss some situations where working with the sample quantile function, $\hat{Q}$, rather than $\hat{F}$, can lead to explicit (exact) solutions to $\theta\left(\hat{F}\right)$.

Fairchild et al.: R^2 effect-size measures for mediation analysis
Fairchild-MacKinnon-Taborga-etal-2009

Amanda J. Fairchild et al. "R^2 effect-size measures for mediation analysis". In: *Behavior Research Methods* 41.2 (May 2009), pp. 486–498. ISSN: 1554-3528. DOI: 10.3758/brm.41.2.486.

Abstract: R^2 effect-size measures are presented to assess variance accounted for in mediation models. The measures offer a means to evaluate both component paths and the overall mediated effect in

mediation models. Statistical simulation results indicate acceptable bias across varying parameter and sample-size combinations. The measures are applied to a real-world example using data from a team-based health promotion program to improve the nutrition and exercise habits of firefighters. SAS and SPSS computer code are also provided for researchers to compute the measures in their own data.

Ferrer et al.: Alternative structural models for multivariate longitudinal data analysis Ferrer-McArdle-2003

Abstract: Structural equation models are presented as alternative models for examining longitudinal data. The models include (a) a cross-lagged regression model, (b) a factor model based on latent growth curves, and (c) a dynamic model based on latent difference scores. The illustrative data are on motivation and perceived competence of students during their first semester in high school. The 3 models yielded different results and such differences were discussed in terms of the conceptualization of change underlying each model. The last model was defended as the most reasonable for these data because it captured the dynamic interrelations between the examined constructs and, at the same time, identified potential growth in the variables.

Fritz et al.: Required sample size to detect the mediated effect   Fritz-MacKinnon-2007

Abstract: Mediation models are widely used, and there are many tests of the mediated effect. One of the most common questions that researchers have when planning mediation studies is, "How many subjects do I need to achieve adequate power when testing for mediation?" This article presents the

necessary sample sizes for six of the most common and the most recommended tests of mediation for various combinations of parameters, to provide a guide for researchers when designing studies or applying for grants.

Gatchel et al.: The biopsychosocial approach to chronic pain: Scientific advances and future directions.   Gatchel-Peng-Peters-etal-2007

Abstract: The prevalence and cost of chronic pain is a major physical and mental health care problem in the United States today. As a result, there has been a recent explosion of research on chronic pain, with significant advances in better understanding its etiology, assessment, and treatment. The purpose of the present article is to provide a review of the most noteworthy developments in the field. The biopsychosocial model is now widely accepted as the most heuristic approach to chronic pain. With this model in mind, a review of the basic neuroscience processes of pain (the bio part of biopsychosocial), as well as the psychosocial factors, is presented. This spans research on how psychological and social factors can interact with brain processes to influence health and illness as well as on the development of new technologies, such as brain imaging, that provide new insights into brain-pain mechanisms.

Graham et al.: How many imputations are really needed? Some practical clarifications of multiple imputation theory   Graham-Olchowski-Gilreath-2007

Abstract: Multiple imputation (MI) and full information maximum likelihood (FIML) are the two most common approaches to missing data analysis. In theory, MI and FIML are equivalent when identical models are tested using the same variables, and when m, the number of imputations performed with MI, approaches infinity. However, it is important to know how many imputations are necessary before MI and FIML are sufficiently equivalent in ways that are important to prevention scientists. MI theory suggests that small values of m, even on the order of three to five imputations, yield excellent results. Previous guidelines for sufficient m are based on relative efficiency, which involves the fraction of missing information (γ) for the parameter being estimated, and m. In the present study, we used a Monte Carlo simulation to test MI models across several scenarios in which γ and m were varied. Standard errors and p-values for the regression coefficient of interest varied as a function of m, but not at the same rate as relative efficiency. Most importantly, statistical power for small effect sizes diminished as m became smaller, and the rate of this power falloff was much greater than predicted by changes in relative efficiency. Based our findings, we recommend that researchers using MI should perform many more imputations than previously considered sufficient. These recommendations are based on γ, and take into consideration one's tolerance for a preventable power falloff (compared to FIML) due to using too few imputations.

Grundy et al.: Marital conflict and preadolescent behavioral competence: Maternal knowledge as a longitudinal mediator   Grundy-Gondoli-BlodgettSalafia-2007

Abstract: The present study considered whether maternal knowledge mediated the relation between overt marital conflict and preadolescent behavioral competence. Four years of self-report data were collected from 133 mothers and their preadolescents, beginning when the preadolescents were in 4th grade. Marital conflict, maternal knowledge, and preadolescent behavioral competence were

assessed at all 4 time points in order to apply a stringent methodology for assessing longitudinal mediating patterns. The results indicated that maternal knowledge mediated the relation between marital conflict and preadolescent behavioral competence. Thus, the present study identified one possible process through which marital conflict may affect preadolescent behavior.

Hall: A short prehistory of the bootstrap   Hall-2003

Abstract: The contemporary development of bootstrap methods, from the time of Efron's early articles to the present day, is well documented and widely appreciated. Likewise, the relationship of bootstrap techniques to certain early work on permutation testing, the jackknife and cross-validation is well understood. Less known, however, are the connections of the bootstrap to research on survey sampling for spatial data in the first half of the last century or to work from the 1940s to the 1970s on subsampling and resampling. In a selective way, some of these early linkages will be explored, giving emphasis to developments with which the statistics community tends to be less familiar. Particular attention will be paid to the work of P. C. Mahalanobis, whose development in the 1930s and 1940s of moving-block sampling methods for spatial data has a range of interesting features, and to contributions of other scientists who, during the next 40 years, developed half-sampling, subsampling and resampling methods.

Hamaker et al.: Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis   Hamaker-Dolan-Molenaar-2005

Abstract: Results obtained with interindividual techniques in a representative sample of a population are not necessarily generalizable to the individual members of this population. In this article the specific condition is presented that must be satisfied to generalize from the interindividual level to the intraindividual level. A way to investigate whether this condition is satisfied is by means of multivariate time series analysis. More generally, time series analysis can be used to investigate psychological processes situated within individuals. In this article we consider a well established class of multivariate stationary time series models that may be used to study the intraindividual covariance structure. We demonstrate the application of some of these models with an empirical example consisting of state measurements of behavior associated with the Five Factor Model of Personality. We illustrate how one can investigate whether individuals are similar with respect to their intraindividual structure of variation, and whether this structure is similar to the structure of interindividual variation.

Hatemi-J: A new method to choose optimal lag order in stable and unstable VAR models HatemiJ-2003

Abdulnasser Hatemi-J. "A new method to choose optimal lag order in stable and unstable VAR models". In: *Applied Economics Letters* 10.3 (Feb. 2003), pp. 135–137. DOI: 10.1080/1350485022000041050.

Abstract: A crucial aspect of empirical research based on the vector autoregressive (VAR) model is the choice of the lag order, since all inference in the VAR model is based on the chosen lag order. Here, a new information criterion is introduced for this purpose. The conducted Monte Carlo simulation experiments show that this new information criterion performs well in picking the true lag order in stable as well as unstable VAR models.

Hatemi-J: Multivariate tests for autocorrelation in the stable and unstable VAR models HatemiJ-2004

Abdulnasser Hatemi-J. "Multivariate tests for autocorrelation in the stable and unstable VAR

models". In: *Economic Modelling* 21.4 (July 2004), pp. 661–683. DOI: 10.1016/j.econmod.2003.09.005.

Abstract: This study investigates the size and power properties of three multivariate tests for autocorrelation, namely portmanteau test, Lagrange multiplier (LM) test and Rao F-test, in the stable and unstable vector autoregressive (VAR) models, with and without autoregressive conditional heteroscedasticity (ARCH) using Monte Carlo experiments. Many combinations of parameters are used in the simulations to cover a wide range of situations in order to make the results more representative. The results of conducted simulations show that all three tests perform relatively well in stable VAR models without ARCH. In unstable VAR models the portmanteau test exhibits serious size distortions. LM and Rao tests perform well in unstable VAR models without ARCH. These results are true, irrespective of sample size or order of autocorrelation. Another clear result that the simulations show is that none of the tests have the correct size when ARCH is present irrespective of VAR models being stable or unstable and regardless of the sample size or order of autocorrelation. The portmanteau test appears to have slightly better power properties than the LM test in almost all scenarios.

Hayes: Beyond Baron and Kenny: Statistical mediation analysis in the new millennium
Hayes-2009

Abstract: Understanding communication processes is the goal of most communication researchers. Rarely are we satisfied merely ascertaining whether messages have an effect on some outcome of focus in a specific context. Instead, we seek to understand how such effects come to be. What kinds of causal sequences does exposure to a message initiate? What are the causal pathways through which a message exerts its effect? And what role does communication play in the transmission of the effects of other variables over time and space? Numerous communication models attempt to describe the

mechanism through which messages or other communication-related variables transmit their effects or intervene between two other variables in a causal model. The communication literature is replete with tests of such models. Over the years, methods used to test such process models have grown in sophistication. An example includes the rise of structural equation modeling (SEM), which allows investigators to examine how well a process model that links some focal variable X to some outcome Y through one or more intervening pathways fits the observed data. Yet frequently, the analytical choices communication researchers make when testing intervening variables models are out of step with advances made in the statistical methods literature. My goal here is to update the field on some of these new advances. While at it, I challenge some conventional wisdom and nudge the field toward a more modern way of thinking about the analysis of intervening variable effects.

Hayes et al.: Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation   Hayes-Cai-2007

Andrew F. Hayes and Li Cai. "Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation". In: *Behavior Research Methods* 39.4 (Nov. 2007), pp. 709–722. DOI: 10.3758/bf03192961.

Higgins et al.: Quantifying heterogeneity in a meta-analysis   Higgins-Thompson-2002

Julian P. T. Higgins and Simon G. Thompson. "Quantifying heterogeneity in a meta-analysis". In: *Statistics in Medicine* 21.11 (May 2002), pp. 1539–1558. ISSN: 1097-0258. DOI: 10.1002/sim.1186.

Abstract: The extent of heterogeneity in a meta-analysis partly determines the difficulty in drawing overall conclusions. This extent may be measured by estimating a between-study variance, but interpretation is then specific to a particular treatment effect metric. A test for the existence of heterogeneity exists, but depends on the number of studies in the meta-analysis. We develop measures of the impact of heterogeneity on a meta-analysis, from mathematical criteria, that are independent of the number of studies and the treatment effect metric. We derive and propose three

suitable statistics: H is the square root of the χ^2 heterogeneity statistic divided by its degrees of freedom; R is the ratio of the standard error of the underlying mean from a random effects meta-analysis to the standard error of a fixed effect meta-analytic estimate, and I^2 is a transformation of H that describes the proportion of total variation in study estimates that is due to heterogeneity. We discuss interpretation, interval estimates and other properties of these measures and examine them in five example data sets showing different amounts of heterogeneity. We conclude that H and I^2, which can usually be calculated for published meta-analyses, are particularly useful summaries of the impact of heterogeneity. One or both should be presented in published meta-analyses in preference to the test for heterogeneity.

Holmes: Bootstrapping phylogenetic trees: Theory and methods   Holmes-2003a

Abstract: This is a survey of the use of the bootstrap in the area of systematic and evolutionary biology. I present the current usage by biologists of the bootstrap as a tool both for making inferences and for evaluating robustness, and propose a framework for thinking about these problems in terms of mathematical statistics.

Holmes: Bradley Efron: A conversation with good friends   Holmes-2003b

Abstract: Bradley Efron is Professor of Statistics and Biostatistics at Stanford University. He works on a combination of theoretical and applied topics, including empirical Bayes, survival analysis, exponential families, bootstrap and jackknife methods and confidence intervals. Most of his applied work has originated in biomedical consulting projects at the Stanford Medical School, mixed in with a few papers concerning astronomy and physics. Even his theoretical papers usually begin

with specific applied problems. All three of the interviewers here have been close scientific collaborators. Brad was born in St. Paul, Minnestora, May 1938, to Esther and Miles Efron, Jewish-Russian immigrants. A Merit Scholarship, in the program's inaugural year, brought him to Caltech, graduating in Mathematics in 1960. He arrived at Stanford that Fall, eventually gaining his Ph.D., under the direction of Rupert Miller and Herb Solomon, in the Statistics Department, whose faculty also included Charles Stein, Herman Chernoff, Manny Parzen, Lincoln Moses and Ingram Olkin. Brad has lived at Stanford since 1960, with sabbaticals at Harvard, Imperial College and Berkeley. He has held several administrative positions in the university: Chair of Statistics, Associate Dean of Science, Chairman of the University Advisory Board and Chair of the Faculty Senate. He is currently Chair of the Undergraduate Program in Applied Mathematics. Honors include doctorates from Chicago, Madrid and Oslo, a MacArthur Prize Fellowship, membership in the National Academy of Sciences and the American Academy of Arts and Sciences, fellowship in the IMS and ASA, the Wilks Medal, Parzen Prize, the newly inaugurated Rao Prize and the outstanding statistician award from the Chicago ASA chapter. He has been the Rietz, Wald, and Fisher lecturers and holds the Max H. Stein endowed chair as Professor of Humanities and Sciences at Stanford. Professional service includes Theory and Methods Editor of JASA and President of the IMS. Currently he is President-Elect of the American Statistical Association, becoming President in 2004.

Horowitz: The bootstrap in econometrics Horowitz-2003

Abstract: This paper presents examples of problems in estimation and hypothesis testing that demonstrate the use and performance of the bootstrap in econometric settings. The examples are illustrated with two empirical applications. The paper concludes with a discussion of topics on which further research is needed.

Kauermann et al.: A note on the efficiency of sandwich covariance matrix estimation

Kauermann-Carroll-2001

Abstract: The sandwich estimator, also known as robust covariance matrix estimator, heteroscedasticity-consistent covariance matrix estimate, or empirical covariance matrix estimator, has achieved increasing use in the econometric literature as well as with the growing popularity of generalized estimating equations. Its virtue is that it provides consistent estimates of the covariance matrix for parameter estimates even when the fitted parametric model fails to hold or is not even specified. Surprisingly though, there has been little discussion of properties of the sandwich method other than consistency. We investigate the sandwich estimator in quasi-likelihood models asymptotically, and in the linear case analytically. We show that under certain circumstances when the quasi-likelihood model is correct, the sandwich estimate is often far more variable than the usual parametric variance estimate. The increased variance is a fixed feature of the method and the price that one pays to obtain consistency even when the parametric model fails or when there is heteroscedasticity. We show that the additional variability directly affects the coverage probability of confidence intervals constructed from sandwich variance estimates. In fact, the use of sandwich variance estimates combined with t-distribution quantiles gives confidence intervals with coverage probability falling below the nominal value. We propose an adjustment to compensate for this fact.

Kenny et al.: Lower level mediation in multilevel models

Kenny-Korchmaros-Bolger-2003

Abstract: Multilevel models are increasingly used to estimate models for hierarchical and repeated measures data. The authors discuss a model in which there is mediation at the lower level and the mediational links vary randomly across upper level units. One repeated measures example is a case in which a person's daily stressors affect his or her coping efforts, which affect his or her mood, and both links vary randomly across persons. Where there is mediation at the lower level and the mediational links vary randomly across upper level units, the formulas for the indirect effect and its standard error must be modified to include the covariance between the random effects. Because no standard method can estimate such a model, the authors developed an ad hoc method that is illustrated with real and simulated data. Limitations of this method and characteristics of an ideal method are discussed.

Koob et al.: Addiction and the brain antireward system   Koob-LeMoal-2008

Abstract: A neurobiological model of the brain emotional systems has been proposed to explain the persistent changes in motivation that are associated with vulnerability to relapse in addiction, and this model may generalize to other psychopathology associated with dysregulated motivational systems. In this framework, addiction is conceptualized as a cycle of decreased function of brain reward systems and recruitment of antireward systems that progressively worsen, resulting in the compulsive use of drugs. Counteradaptive processes, such as opponent process, that are part of the normal homeostatic limitation of reward function fail to return within the normal homeostatic range and are hypothesized to repeatedly drive the allostatic state. Excessive drug taking thus results in not only the short-term amelioration of the reward deficit but also suppression of the antireward system. However, in the long term, there is worsening of the underlying neurochemical dysregulations that ultimately form an allostatic state (decreased dopamine and opioid peptide function, increased corticotropin-releasing factor activity). This allostatic state is hypothesized to

be reflected in a chronic deviation of reward set point that is fueled not only by dysregulation of reward circuits per se but also by recruitment of brain and hormonal stress responses. Vulnerability to addiction may involve genetic comorbidity and developmental factors at the molecular, cellular, or neurocircuitry levels that sensitize the brain antireward systems.

Krull et al.: Multilevel modeling of individual and group level mediated effects
Krull-MacKinnon-2001

Abstract: This article combines procedures for single-level mediational analysis with multilevel modeling techniques in order to appropriately test mediational effects in clustered data. A simulation study compared the performance of these multilevel mediational models with that of single-level mediational models in clustered data with individual- or group-level initial independent variables, individual- or group-level mediators, and individual level outcomes. The standard errors of mediated effects from the multilevel solution were generally accurate, while those from the single-level procedure were downwardly biased, often by 20% or more. The multilevel advantage was greatest in those situations involving group-level variables, larger group sizes, and higher intraclass correlations in mediator and outcome variables. Multilevel mediational modeling methods were also applied to data from a preventive intervention designed to reduce intentions to use steroids among players on high school football teams. This example illustrates differences between single-level and multilevel mediational modeling in real-world clustered data and shows how the multilevel technique may lead to more accurate results.

Lahiri: On the impact of bootstrap in survey sampling and small-area estimation

Abstract: Development of valid bootstrap procedures has been a challenging problem for survey samplers for the last two decades. This is due to the fact that in surveys we constantly face various complex issues such as complex correlation structure induced by the survey design, weighting, imputation, small-area estimation, among others. In this paper, we critically review various bootstrap methods developed to deal with these challenging issues. We discuss two applications where the bootstrap has been found to be effective.

Lele: Impact of bootstrap on the estimating functions

Abstract: Estimating functions form an attractive statistical methodology because of their dependence on only a few features of the underlying probabilistic structure. They also put a premium on developing methods that obtain model-robust confidence intervals. Bootstrap and jackknife ideas can be fruitfully used toward this purpose. Another important area in which bootstrap has proved its use is in the context of detecting the problem of multiple roots and searching for the consistent root of an estimating function. In this article, I review, compare and contrast various approaches for bootstrapping estimating functions.

Long et al.: Using heteroscedasticity consistent standard errors in the linear regression model   **Long-Ervin-2000**

J. Scott Long and Laurie H. Ervin. "Using heteroscedasticity consistent standard errors in the linear regression model". In: *The American Statistician* 54.3 (Aug. 2000), pp. 217–224. DOI: 10.1080/00031305.2000.10474549.

Lüdtke et al.: The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies   **Ludtke-Marsh-Robitzsch-etal-2008**

Oliver Lüdtke et al. "The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies". In: *Psychological Methods* 13.3 (Sept. 2008), pp. 203–229. ISSN: 1082-989X. DOI: 10.1037/a0012869.

Abstract: In multilevel modeling (MLM), group-level (L2) characteristics are often measured by aggregating individual-level (L1) characteristics within each group so as to assess contextual effects (e.g., group-average effects of socioeconomic status, achievement, climate). Most previous applications have used a multilevel manifest covariate (MMC) approach, in which the observed (manifest) group mean is assumed to be perfectly reliable. This article demonstrates mathematically and with simulation results that this MMC approach can result in substantially biased estimates of contextual effects and can substantially underestimate the associated standard errors, depending on the number of L1 individuals per group, the number of groups, the intraclass correlation, the sampling ratio (the percentage of cases within each group sampled), and the nature of the data. To address this pervasive problem, the authors introduce a new multilevel latent covariate (MLC) approach that corrects for unreliability at L2 and results in unbiased estimates of L2 constructs under appropriate conditions. However, under some circumstances when the sampling ratio approaches 100%, the MMC approach provides more accurate estimates. Based on 3 simulations and 2 real-data applications, the authors evaluate the MMC and MLC approaches and suggest when researchers should most appropriately use one, the other, or a combination of both approaches.

Abstract: This article describes a program, PRODCLIN (distribution of the PRODuct Confidence Limits for INdirect effects), written for SAS, SPSS, and R, that computes confidence limits for the product of two normal random variables. The program is important because it can be used to obtain more accurate confidence limits for the indirect effect, as demonstrated in several recent articles (MacKinnon, Lockwood, & Williams, 2004; Pituch, Whittaker, & Stapleton, 2005). Tests of the significance of and confidence limits for indirect effects based on the distribution of the product method have more accurate Type I error rates and more power than other, more commonly used tests. Values for the two paths involved in the indirect effect and their standard errors are entered in the PRODCLIN program, and distribution of the product confidence limits are computed. Several examples are used to illustrate the PRODCLIN program. The PRODCLIN programs in rich text format may be downloaded from www.psychonomic.org/archive.

Abstract: This paper describes the statistical similarities among mediation, confounding, and suppression. Each is quantified by measuring the change in the relationship between an independent and a dependent variable after adding a third variable to the analysis. Mediation and confounding are identical statistically and can be distinguished only on conceptual grounds. Methods to

determine the confidence intervals for confounding and suppression effects are proposed based on methods developed for mediated effects. Although the statistical estimation of effects and standard errors is the same, there are important conceptual differences among the three types of effects.

MacKinnon et al.: A comparison of methods to test mediation and other intervening variable effects   MacKinnon-Lockwood-Hoffman-etal-2002

David P. MacKinnon, Chondra M. Lockwood, Jeanne M. Hoffman, et al. "A comparison of methods to test mediation and other intervening variable effects". In: *Psychological Methods* 7.1 (2002), pp. 83–104. DOI: 10.1037/1082-989x.7.1.83.

Abstract: A Monte Carlo study compared 14 methods to test the statistical significance of the intervening variable effect. An intervening variable (mediator) transmits the effect of an independent variable to a dependent variable. The commonly used R. M. Baron and D. A. Kenny (1986) approach has low statistical power. Two methods based on the distribution of the product and 2 difference-in-coefficients methods have the most accurate Type I error rates and greatest statistical power except in 1 important case in which Type I error rates are too high. The best balance of Type I error and statistical power across all cases is the test of the joint significance of the two effects comprising the intervening variable effect.

MacKinnon et al.: Confidence limits for the indirect effect: Distribution of the product and resampling methods   MacKinnon-Lockwood-Williams-2004

David P. MacKinnon, Chondra M. Lockwood, and Jason Williams. "Confidence limits for the indirect effect: Distribution of the product and resampling methods". In: *Multivariate Behavioral Research* 39.1 (Jan. 2004), pp. 99–128. DOI: 10.1207/s15327906mbr3901_4.

Abstract: The most commonly used method to test an indirect effect is to divide the estimate of the indirect effect by its standard error and compare the resulting z statistic with a critical value from the standard normal distribution. Confidence limits for the indirect effect are also typically

based on critical values from the standard normal distribution. This article uses a simulation study to demonstrate that confidence limits are imbalanced because the distribution of the indirect effect is normal only in special cases. Two alternatives for improving the performance of confidence limits for the indirect effect are evaluated: (a) a method based on the distribution of the product of two normal random variables, and (b) resampling methods. In Study 1, confidence limits based on the distribution of the product are more accurate than methods based on an assumed normal distribution but confidence limits are still imbalanced. Study 2 demonstrates that more accurate confidence limits are obtained using resampling methods, with the bias-corrected bootstrap the best method overall.

Maxwell et al.: Bias in cross-sectional analyses of longitudinal mediation

Maxwell-Cole-2007

Abstract: Most empirical tests of mediation utilize cross-sectional data despite the fact that mediation consists of causal processes that unfold over time. The authors considered the possibility that longitudinal mediation might occur under either of two different models of change: (a) an autoregressive model or (b) a random effects model. For both models, the authors demonstrated that cross-sectional approaches to mediation typically generate substantially biased estimates of longitudinal parameters even under the ideal conditions when mediation is complete. In longitudinal models where variable M completely mediates the effect of X on Y, cross-sectional estimates of the direct effect of X on Y, the indirect effect of X on Y through M, and the proportion of the total effect mediated by M are often highly misleading.

J. J. McArdle: Latent variable modeling of differences and changes with longitudinal data

John J. McArdle. "Latent variable modeling of differences and changes with longitudinal data". In: *Annual Review of Psychology* 60.1 (Jan. 2009), pp. 577–605. DOI: 10.1146/annurev.psych.60.110707.163612.

Abstract: This review considers a common question in data analysis: What is the most useful way to analyze longitudinal repeated measures data? We discuss some contemporary forms of structural equation models (SEMs) based on the inclusion of latent variables. The specific goals of this review are to clarify basic SEM definitions, consider relations to classical models, focus on testable features of the new models, and provide recent references to more complete presentations. A broader goal is to illustrate why so many researchers are enthusiastic about the SEM approach to data analysis. We first outline some classic problems in longitudinal data analysis, consider definitions of differences and changes, and raise issues about measurement errors. We then present several classic SEMs based on the inclusion of invariant common factors and explain why these are so important. This leads to newer SEMs based on latent change scores, and we explain why these are useful.

Mehta et al.: People are variables too: Multilevel structural equations modeling.

Paras D. Mehta and Michael C. Neale. "People are variables too: Multilevel structural equations modeling." In: *Psychological Methods* 10.3 (Sept. 2005), pp. 259–284. ISSN: 1082-989X. DOI: 10.1037/1082-989x.10.3.259.

Abstract: The article uses confirmatory factor analysis (CFA) as a template to explain didactically multilevel structural equation models (ML-SEM) and to demonstrate the equivalence of general mixed-effects models and ML-SEM. An intuitively appealing graphical representation of complex ML-SEMs is introduced that succinctly describes the underlying model and its assumptions. The use of definition variables (i.e., observed variables used to fix model parameters to individual specific

data values) is extended to the case of ML-SEMs for clustered data with random slopes. Empirical examples of multilevel CFA and ML-SEM with random slopes are provided along with scripts for fitting such models in SAS Proc Mixed, Mplus, and Mx. Methodological issues regarding estimation of complex ML-SEMs and the evaluation of model fit are discussed. Further potential applications of ML-SEMs are explored.

Molenaar: A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever   Molenaar-2004

Peter C. M. Molenaar. "A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever". In: *Measurement: Interdisciplinary Research & Perspective* 2.4 (Oct. 2004), pp. 201–218. ISSN: 1536-6359. DOI: 10.1207/s15366359mea0204_1.

Abstract: Psychology is focused on variation between cases (interindividual variation). Results thus obtained are considered to be generalizable to the understanding and explanation of variation within single cases (intraindividual variation). It is indicated, however, that the direct consequences of the classical ergodic theorems for psychology and psychometrics invalidate this conjectured generalizability: only under very strict conditions-which are hardly obtained in real psychological processes-can a generalization be made from a structure of interindividual variation to the analogous structure of intraindividual variation. Illustrations of the lack of this generalizability are given in the contexts of psychometrics, developmental psychology, and personality theory.

Molenaar et al.: The new person-specific paradigm in psychology

Molenaar-Campbell-2009

Peter C. M. Molenaar and Cynthia G. Campbell. "The new person-specific paradigm in psychology". In: *Current Directions in Psychological Science* 18.2 (Apr. 2009), pp. 112–117. ISSN: 1467-8721. DOI: 10.1111/j.1467-8721.2009.01619.x.

Abstract: Most research methodology in the behavioral sciences employs interindividual analyses, which provide information about the state of affairs of the population. However, as shown by classical mathematical-statistical theorems (the ergodic theorems), such analyses do not provide information for, and cannot be applied at, the level of the individual, except on rare occasions when the processes of interest meet certain stringent conditions. When psychological processes violate these conditions, the interindividual analyses that are now standardly applied have to be replaced by analysis of intraindividual variation in order to obtain valid results. Two illustrations involving analysis of intraindividual variation of personality and emotional processes are given.

Oud et al.: Continuous time state space modeling of panel data by means of SEM
Oud-Jansen-2000

Johan H. L. Oud and Robert A. R. G. Jansen. "Continuous time state space modeling of panel data by means of SEM". In: *Psychometrika* 65.2 (June 2000), pp. 199–215. DOI: 10.1007/bf02294374.

Abstract: Maximum likelihood parameter estimation of the continuous time linear stochastic state space model is considered on the basis of largeN discrete time data using a structural equation modeling (SEM) program. Random subject effects are allowed to be part of the model. The exact discrete model (EDM) is employed which links the discrete time model parameters to the underlying continuous time model parameters by means of nonlinear restrictions. The EDM is generalized to cover not only time-invariant parameters but also the cases of stepwise time-varying (piecewise time-invariant) parameters and parameters varying continuously over time according to a general polynomial scheme. The identification of the continuous time parameters is discussed and an educational example is presented.

Patrick et al.: Short-term changes in plans to drink and importance of positive and negative alcohol consequences
Patrick-Maggs-2007

Megan E. Patrick and Jennifer L. Maggs. "Short-term changes in plans to drink and importance of

positive and negative alcohol consequences". In: *Journal of Adolescence* 31.3 (July 2007), pp. 307–321. ISSN: 1095-9254. DOI: 10.1016/j.adolescence.2007.06.002.

Abstract: Experienced consequences predicted short-term changes in alcohol use plans and perceptions of the importance of alcohol-related consequences. Participants were 176 traditionally aged first-year university students who completed a 10-week telephone diary study (total weeks = 1735). In multi-level models, men and students who experienced more positive and negative consequences on average planned to drink more and rated avoiding negative consequences as less important. Students who experienced more positive consequences rated them as more important (between-person analyses). Following weeks of experiencing relatively more positive drinking consequences, students planned to drink more and rated experiencing positive consequences as more important for the subsequent week (within-person analyses). Challenges for intervening in the ongoing formation of anticipatory cognitions regarding alcohol use are discussed.

Patrick et al.: Profiles of motivations for alcohol use and sexual behavior among first-year university students   Patrick-Maggs-2009

Megan E. Patrick and Jennifer L. Maggs. "Profiles of motivations for alcohol use and sexual behavior among first-year university students". In: *Journal of Adolescence* 33.5 (Nov. 2009), pp. 755–765. ISSN: 1095-9254. DOI: 10.1016/j.adolescence.2009.10.003.

Abstract: The links between motivations for alcohol use and for sexual behaviors are not well understood. Latent profile analysis was used to identify drinking motivational profiles (based on motivations for: fun/social, relaxation/coping, image, sex; motivations against: physical, behavioral) and sex motivational profiles (motivations for: enhancement, intimacy, coping; motivations against: not ready, health, values) among college students (N = 227, 51% male). Latent profiles for drinking were: low for/high against drinking (34%), average drinking motives (53%), and high for/low against drinking (13%). Profiles for sex were: low for/high against sex (35%), high for/low against sex (42%), and high for with coping/moderate against sex (23%). Motivational profiles were related

across behaviors. Drinking motivational profiles were associated with alcohol use and psychosocial adjustment; sex motivational profiles were associated with sexual experiences. Distinct profiles of motivations support the need for differentiated intervention programs targeting individuals with different patterns of reasons for engaging in risk behaviors during late adolescence.

Peugh et al.: Missing data in educational research: A review of reporting practices and suggestions for improvement   Peugh-Enders-2004

Abstract: Missing data analyses have received considerable recent attention in the methodological literature, and two "modern" methods, multiple imputation and maximum likelihood estimation, are recommended. The goals of this article are to (a) provide an overview of missing-data theory, maximum likelihood estimation, and multiple imputation; (b) conduct a methodological review of missing-data reporting practices in 23 applied research journals; and (c) provide a demonstration of multiple imputation and maximum likelihood estimation using the Longitudinal Study of American Youth data. The results indicated that explicit discussions of missing data increased substantially between 1999 and 2003, but the use of maximum likelihood estimation or multiple imputation was rare; the studies relied almost exclusively on listwise and pairwise deletion.

Politis: The impact of bootstrap methods on time series analysis   Politis-2003

Abstract: Sparked by Efron's seminal paper, the decade of the 1980s was a period of active research on bootstrap methods for independent data–mainly i.i.d. or regression set-ups. By contrast, in the 1990s much research was directed towards resampling dependent data, for example, time series and

random fields. Consequently, the availability of valid nonparametric inference procedures based on resampling and/or subsampling has freed practitioners from the necessity of resorting to simplifying assumptions such as normality or linearity that may be misleading.

Preacher et al.: Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis   Preacher-Curran-Bauer-2006

Abstract: Simple slopes, regions of significance, and confidence bands are commonly used to evaluate interactions in multiple linear regression (MLR) models, and the use of these techniques has recently been extended to multilevel or hierarchical linear modeling (HLM) and latent curve analysis (LCA). However, conducting these tests and plotting the conditional relations is often a tedious and error-prone task. This article provides an overview of methods used to probe interaction effects and describes a unified collection of freely available online resources that researchers can use to obtain significance tests for simple slopes, compute regions of significance, and obtain confidence bands for simple slopes across the range of the moderator in the MLR, HLM, and LCA contexts. Plotting capabilities are also provided.

Preacher et al.: SPSS and SAS procedures for estimating indirect effects in simple mediation models   Preacher-Hayes-2004

Abstract: Researchers often conduct mediation analysis in order to indirectly assess the effect of a proposed cause on some outcome through a proposed mediator. The utility of mediation analysis stems from its ability to go beyond the merely descriptive to a more functional understanding of the relationships among variables. A necessary component of mediation is a statistically and practically significant indirect effect. Although mediation hypotheses are frequently explored in psychological research, formal significance tests of indirect effects are rarely conducted. After a brief overview of mediation, we argue the importance of directly testing the significance of indirect effects and provide SPSS and SAS macros that facilitate estimation of the indirect effect with a normal theory approach and a bootstrap approach to obtaining confidence intervals, as well as the traditional approach advocated by Baron and Kenny (1986). We hope that this discussion and the macros will enhance the frequency of formal mediation tests in the psychology literature. Electronic copies of these macros may be downloaded from the Psychonomic Society's Web archive at www.psychonomic.org/archive/.

Preacher et al.: Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models   Preacher-Hayes-2008

Abstract: Hypotheses involving mediation are common in the behavioral sciences. Mediation exists when a predictor affects a dependent variable indirectly through at least one intervening variable, or mediator. Methods to assess mediation involving multiple simultaneous mediators have received little attention in the methodological literature despite a clear need. We provide an overview of simple and multiple mediation and explore three approaches that can be used to investigate indirect processes, as well as methods for contrasting two or more mediators within a single model. We present an illustrative example, assessing and contrasting potential mediators of the relationship

between the helpfulness of socialization agents and job satisfaction. We also provide SAS and SPSS macros, as well as Mplus and LISREL syntax, to facilitate the use of these methods in applications.

Raghunathan et al.: A multivariate technique for multiply imputing missing values using a sequence of regression models

Raghunathan-Lepkowski-vanHoewyk-etal-2001

Trivellore E. Raghunathan et al. "A multivariate technique for multiply imputing missing values using a sequence of regression models". In: *Survey Methodology* 27.1 (2001), pp. 85–95.

Abstract: This article describes and evaluates a procedure for imputing missing values for a relatively complex data structure when the data are missing at random. The imputations are obtained by fitting a sequence of regression models and drawing values from the corresponding predictive distributions. The types of regression models used are linear, logistic, Poisson, generalized logit or a mixture of these depending on the type of variable being imputed. Two additional common features in the imputation process are incorporated: restriction to a relevant subpopulation for some variables and logical bounds or constraints for the imputed values. The restrictions involve subsetting the sample individuals that satisfy certain criteria while fitting the regression models. The bounds involve drawing values from a truncated predictive distribution. The development of this method was partly motivated by the analysis of two data sets which are used as illustrations. The sequential regression procedure is applied to perform multiple imputation analysis for the two applied problems. The sampling properties of inferences from multiply imputed data sets created using the sequential regression method are evaluated through simulated data sets.

Raykov et al.: Using the delta method for approximate interval estimation of parameter functions in SEM

Raykov-Marcoulides-2004

Tenko Raykov and George A. Marcoulides. "Using the delta method for approximate interval es-

timation of parameter functions in SEM". In: *Structural Equation Modeling: A Multidisciplinary Journal* 11.4 (Oct. 2004), pp. 621–637. ISSN: 1532-8007. DOI: 10.1207/s15328007sem1104_7.

Abstract: In applications of structural equation modeling, it is often desirable to obtain measures of uncertainty for special functions of model parameters. This article provides a didactic discussion of how a method widely used in applied statistics can be employed for approximate standard error and confidence interval evaluation of such functions. The described approach is illustrated with data from a cognitive intervention study, in which it is used to estimate time-invariant reliability in multiwave, multiple indicator models.

Sakai et al.: Validity of transdermal alcohol monitoring: Fixed and self-regulated dosing
Sakai-MikulichGilbertson-Long-etal-2006

Abstract: Background: To study the validity of transdermal assessment of alcohol concentration measured by a lightweight, noninvasive device. Methods: Subjects wore a 227-g anklet that sensed transdermal alcohol concentrations (TACs) every 15 to 30 minutes, downloading results to a remote computer each day. Twenty-four subjects entered a laboratory and received a dose of 0, 0.28, or 0.56 g/kg of ethanol. Breath alcohol concentrations (BrAC) and TAC were measured every 15 to 30 minutes Twenty others [10 alcohol dependent (AD) and 10 not (NAD)] in the community who wore the anklet for 8 days kept a drinking log and provided a BrAC sample each day. Results: In the laboratory, no zero-dose subject, and every subject receiving alcohol, had alcohol-positive TACs. The device distinguished low- and high-alcohol–dosing groups using peak ($t_{14} = 3.37; p < 0.01$) and area under the curve ($t_{14} = 3.42; p < 0.01$) of TACs. Within dosing groups, average TAC curves were broader (right-shifted) and had lower peaks than average BrAC curves. For community participants, self-reported number of drinks ($t_{18} = -3.77; p < 0.01$), area under the TAC curve ($t_{9.5} = -3.56;$

$p < 0.01$), and mean TAC ($t_{9.9} = -3.35$; $p < 0.01$) all significantly distinguished the AD and NAD groups. However, individual transdermal readings were not reliably quantitatively equivalent to simultaneously obtained breath results. Conclusions: Within the limits of the laboratory study, the device consistently detected consumption of approximately 2 standard drinks. On average, the device shows discriminative validity as a semiquantitative measure of alcohol consumption but individual readings often are not equivalent to simultaneous BrACs.

Schafer et al.: Missing data: Our view of the state of the art   Schafer-Graham-2002

Joseph L. Schafer and John W. Graham. "Missing data: Our view of the state of the art". In: *Psychological Methods* 7.2 (2002), pp. 147–177. DOI: 10.1037/1082-989x.7.2.147.

Abstract: Statistical procedures for missing data have vastly improved, yet misconception and unsound practice still abound. The authors frame the missing-data problem, review methods, offer advice, and raise issues that remain unresolved. They clear up common misunderstandings regarding the missing at random (MAR) concept. They summarize the evidence against older procedures and, with few exceptions, discourage their use. They present, in both technical and practical language, 2 general approaches that come highly recommended: maximum likelihood (ML) and Bayesian multiple imputation (MI). Newer developments are discussed, including some for dealing with missing data that are not MAR. Although not yet in the mainstream, these procedures may eventually extend the ML and MI methods that currently represent the state of the art.

Selig et al.: Mediation models for longitudinal data in developmental research

Selig-Preacher-2009

James P. Selig and Kristopher J. Preacher. "Mediation models for longitudinal data in developmental research". In: *Research in Human Development* 6.2-3 (June 2009), pp. 144–164. DOI: 10.1080/15427600902911247.

Abstract: Mediation models are used to describe the mechanism(s) by which one variable influences another. These models can be useful in developmental research to explicate the relationship between variables, developmental processes, or combinations of variables and processes. In this article we describe aspects of mediation effects specific to developmental research. We focus on three central issues in longitudinal mediation models: the theory of change for variables in the model, the role of time in the model, and the types of indirect effects in the model. We use these themes as we describe three different models for examining mediation in longitudinal data.

Serlin: Testing for robustness in Monte Carlo studies   Serlin-2000

Ronald C. Serlin. "Testing for robustness in Monte Carlo studies". In: *Psychological Methods* 5.2 (2000), pp. 230–240. DOI: 10.1037/1082-989x.5.2.230.

Abstract: Monte Carlo studies provide the information needed to help researchers select appropriate analytical procedures under design conditions in which the underlying assumptions of the procedures are not met. In Monte Carlo studies, the 2 errors that one could commit involve (a) concluding that a statistical procedure is robust when it is not or (b) concluding that it is not robust when it is. In previous attempts to apply standard statistical design principles to Monte Carlo studies, the less severe of these errors has been wrongly designated the Type I error. In this article, a method is presented for controlling the appropriate Type I error rate; the determination of the number of iterations required in a Monte Carlo study to achieve desired power is described; and a confidence interval for a test's true Type I error rate is derived. A robustness criterion is also proposed that is a compromise between W. G. Cochran's (1952) and J. V. Bradley's (1978) criteria.

Shao: Impact of the bootstrap on sample surveys   Shao-2003

Jun Shao. "Impact of the bootstrap on sample surveys". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994974.

Abstract: This article discusses the impact of the bootstrap on sample surveys and introduces some of the main developments of the bootstrap methodology for sample surveys in the last twenty five years.

Shiffman: Ecological momentary assessment (EMA) in studies of substance use

Shiffman-2009

Abstract: Ecological momentary assessment (EMA) is particularly suitable for studying substance use, because use is episodic and thought to be related to mood and context. This article reviews EMA methods in substance use research, focusing on tobacco and alcohol use and relapse, where EMA has been most applied. Common EMA designs combine event-based reports of substance use with time-based assessments. Approaches to data organization and analysis have been very diverse, particularly regarding their treatment of time. Compliance with signaled assessments is often high. Compliance with recording of substance use appears good but is harder to validate. Treatment applications of EMA are emerging. EMA captures substance use patterns not measured by questionnaires or retrospective data and holds promise for substance use research.

Shiffman et al.: Ecological momentary assessment   Shiffman-Stone-Hufford-2008

Abstract: Assessment in clinical psychology typically relies on global retrospective self-reports collected at research or clinic visits, which are limited by recall bias and are not well suited to address how behavior changes over time and across contexts. Ecological momentary assessment (EMA) involves repeated sampling of subjects' current behaviors and experiences in real time, in subjects'

natural environments. EMA aims to minimize recall bias, maximize ecological validity, and allow study of microprocesses that influence behavior in real-world contexts. EMA studies assess particular events in subjects' lives or assess subjects at periodic intervals, often by random time sampling, using technologies ranging from written diaries and telephones to electronic diaries and physiological sensors. We discuss the rationale for EMA, EMA designs, methodological and practical issues, and comparisons of EMA and recall data. EMA holds unique promise to advance the science and practice of clinical psychology by shedding light on the dynamics of behavior in real-world settings.

Shrout et al.: Mediation in experimental and nonexperimental studies: New procedures and recommendations   Shrout-Bolger-2002

Patrick E. Shrout and Niall Bolger. "Mediation in experimental and nonexperimental studies: New procedures and recommendations". In: *Psychological Methods* 7.4 (2002), pp. 422–445. DOI: 10.1037/1082-989x.7.4.422.

Abstract: Mediation is said to occur when a causal effect of some variable X on an outcome Y is explained by some intervening variable M. The authors recommend that with small to moderate samples, bootstrap methods (B. Efron & R. Tibshirani, 1993) be used to assess mediation. Bootstrap tests are powerful because they detect that the sampling distribution of the mediated effect is skewed away from 0. They argue that R. M. Baron and D. A. Kenny's (1986) recommendation of first testing the $X \rightarrow Y$ association for statistical significance should not be a requirement when there is a priori belief that the effect size is small or suppression is a possibility. Empirical examples and computer setups for bootstrap analyses are provided.

P. S. Soltis et al.: Applying the Bootstrap in Phylogeny Reconstruction

Soltis-Soltis-2003

Pamela S. Soltis and Douglas E. Soltis. "Applying the Bootstrap in Phylogeny Reconstruction". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994980.

Abstract: With the increasing emphasis in biology on reconstruction of phylogenetic trees, questions have arisen as to how confident one should be in a given phylogenetic tree and how support for phylogenetic trees should be measured. Felsenstein suggested that bootstrapping be applied across characters of a taxon-by-character data matrix to produce replicate "bootstrap data sets," each of which is then analyzed phylogenetically, with a consensus tree constructed to summarize the results of all replicates. The proportion of trees/replicates in which a grouping is recovered is presented as a measure of support for that group. Bootstrapping has become a common feature of phylogenetic analysis. However, the interpretation of bootstrap values remains open to discussion, and phylogeneticists have used these values in multiple ways. The usefulness of phylogenetic bootstrapping is potentially limited by a number of features, such as the size of the data matrix and the underlying assumptions of the phylogeny reconstruction program. Recent studies have explored the application of bootstrapping to large data sets and the relative performance of bootstrapping and jackknifing.

Staudenmayer et al.: Measurement error in linear autoregressive models

Staudenmayer-Buonaccorsi-2005

Abstract: Time series data are often subject to measurement error, usually the result of needing to estimate the variable of interest. Although it is often reasonable to assume that the measurement error is additive (i.e., the estimator is conditionally unbiased for the missing true value), the measurement error variances often vary as a result of changes in the population/process over time and/or changes in sampling effort. In this article we address estimation of the parameters in linear autoregressive models in the presence of additive and uncorrelated measurement errors, allowing heteroscedasticity in the measurement error variances. We establish the asymptotic properties of naive estimators that ignore measurement error and propose an estimator based on correcting the Yule–Walker estimating equations. We also examine a pseudo-likelihood method based on normal-

ity assumptions and computed using the Kalman filter. We review other techniques that have been proposed, including two that require no information about the measurement error variances, and compare the various estimators both theoretically and via simulations. The estimator based on corrected estimating equations is easy to obtain and readily accommodates (and is robust to) unequal measurement error variances. Asymptotic calculations and finite-sample simulations show that it is often relatively efficient.

Swift: Transdermal alcohol measurement for estimation of blood alcohol concentration
Swift-2000

Robert Swift. "Transdermal alcohol measurement for estimation of blood alcohol concentration". In: *Alcoholism: Clinical and Experimental Research* 24.4 (Apr. 2000), pp. 422–423. ISSN: 1530-0277. DOI: 10.1111/j.1530-0277.2000.tb02006.x.

Taylor et al.: Tests of the three-path mediated effect   Taylor-MacKinnon-Tein-2007

Aaron B. Taylor, David P. MacKinnon, and Jenn-Yun Tein. "Tests of the three-path mediated effect". In: *Organizational Research Methods* 11.2 (July 2007), pp. 241–269. DOI: 10.1177/1094428107300344.

Abstract: In a three-path mediational model, two mediators intervene in a series between an independent and a dependent variable. Methods of testing for mediation in such a model are generalized from the more often used single-mediator model. Six such methods are introduced and compared in a Monte Carlo study in terms of their Type I error, power, and coverage. Based on its results, the joint significance test is preferred when only a hypothesis test is of interest. The percentile bootstrap and bias-corrected bootstrap are preferred when a confidence interval on the mediated effect is desired, with the latter having more power but also slightly inflated Type I error in some conditions.

van Buuren et al.: Fully conditional specification in multivariate imputation

vanBuuren-Brand-GroothuisOudshoorn-etal-2006

Stef van Buuren et al. "Fully conditional specification in multivariate imputation". In: *Journal of Statistical Computation and Simulation* 76.12 (Dec. 2006), pp. 1049–1064. DOI: 10.1080/10629360600810434.

Abstract: The use of the Gibbs sampler with fully conditionally specified models, where the distribution of each variable given the other variables is the starting point, has become a popular method to create imputations in incomplete multivariate data. The theoretical weakness of this approach is that the specified conditional densities can be incompatible, and therefore the stationary distribution to which the Gibbs sampler attempts to converge may not exist. This study investigates practical consequences of this problem by means of simulation. Missing data are created under four different missing data mechanisms. Attention is given to the statistical behavior under compatible and incompatible models. The results indicate that multiple imputation produces essentially unbiased estimates with appropriate coverage in the simple cases investigated, even for the incompatible models. Of particular interest is that these results were produced using only five Gibbs iterations starting from a simple draw from observed marginal distributions. It thus appears that, despite the theoretical weaknesses, the actual performance of conditional model specification for multivariate imputation can be quite good, and therefore deserves further study.

Wang et al.: Investigating the impact of the time interval selection on autoregressive mediation modeling: Result interpretations, effect reporting, and temporal designs

Wang-Zhang-2020

Lijuan Wang and Qian Zhang. "Investigating the impact of the time interval selection on autoregressive mediation modeling: Result interpretations, effect reporting, and temporal designs". In: *Psychological Methods* 25.3 (June 2020), pp. 271–291. ISSN: 1082-989X. DOI: 10.1037/met0000235.

Abstract: This study investigates the impact of the time interval (the time passed between 2 consecutive measurements) selection on autoregressive mediation modeling (AMM). For a widely used autoregressive mediation model, via analytical derivations, we explained why and how the conventionally reported time-specific coefficient estimates (e.g., $\hat{a}\hat{b}\hat{b}$ and $\hat{c}'$) and inference results in AMM provide limited information and can arrive in even misleading conclusions about direct and indirect effects over time. Furthermore, under the stationarity assumption, we proposed an approach to calculate the overall direct and indirect effect estimates over time and the time lag lengths at which they reach maxima, using AMM results. The derivation results revealed that the overall direct and indirect effect curves are asymptotically invariant to the time interval selection, under stationarity. With finite samples and thus sampling errors and potential computing problems, however, our simulation results revealed that the overall indirect effect curves were better recovered when the time interval is selected to be closer to half of the time lag length at which the overall indirect effect reaches its maximum. An R function and an R Shiny app were developed to obtain the overall direct and indirect effect curves over time and facilitate the time interval selection using AMM results. Our findings provide another look at the connections between AMM and continuous time mediation modeling and the connections are discussed.

K.-H. Yuan et al.: Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data   Yuan-Bentler-2000

Abstract: Survey and longitudinal studies in the social and behavioral sciences generally contain missing data. Mean and covariance structure models play an important role in analyzing such data. Two promising methods for dealing with missing data are a direct maximum-likelihood and a two-stage approach based on the unstructured mean and covariance estimates obtained by the EM-algorithm. Typical assumptions under these two methods are ignorable nonresponse and normality

of data. However, data sets in social and behavioral sciences are seldom normal, and experience with these procedures indicates that normal theory based methods for nonnormal data very often lead to incorrect model evaluations. By dropping the normal distribution assumption, we develop more accurate procedures for model inference. Based on the theory of generalized estimating equations, a way to obtain consistent standard errors of the two-stage estimates is given. The asymptotic efficiencies of different estimators are compared under various assumptions. We also propose a minimum chi-square approach and show that the estimator obtained by this approach is asymptotically at least as efficient as the two likelihood-based estimators for either normal or nonnormal data. The major contribution of this paper is that for each estimator, we give a test statistic whose asymptotic distribution is chisquare as long as the underlying sampling distribution enjoys finite fourth-order moments. We also give a characterization for each of the two likelihood ratio test statistics when the underlying distribution is nonnormal. Modifications to the likelihood ratio statistics are also given. Our working assumption is that the missing data mechanism is missing completely at random. Examples and Monte Carlo studies indicate that, for commonly encountered nonnormal distributions, the procedures developed in this paper are quite reliable even for samples with missing data that are missing at random.

Y. Yuan et al.: Bayesian mediation analysis. Yuan-MacKinnon-2009

Abstract: In this article, we propose Bayesian analysis of mediation effects. Compared with conventional frequentist mediation analysis, the Bayesian approach has several advantages. First, it allows researchers to incorporate prior information into the mediation analysis, thus potentially improving the efficiency of estimates. Second, under the Bayesian mediation analysis, inference is straightforward and exact, which makes it appealing for studies with small samples. Third, the Bayesian approach is conceptually simpler for multilevel mediation analysis. Simulation studies and analysis of 2 data sets are used to illustrate the proposed methods.

Achim Zeileis. "Econometric computing with HC and HAC covariance matrix estimators". In: *Journal of Statistical Software* 11.10 (2004). DOI: 10.18637/jss.v011.i10.

Abstract: Data described by econometric models typically contains autocorrelation and/or heteroskedasticity of unknown form and for inference in such models it is essential to use covariance matrix estimators that can consistently estimate the covariance of the model parameters. Hence, suitable heteroskedasticity consistent (HC) and heteroskedasticity and autocorrelation consistent (HAC) estimators have been receiving attention in the econometric literature over the last 20 years. To apply these estimators in practice, an implementation is needed that preferably translates the conceptual properties of the underlying theoretical frameworks into computational tools. In this paper, such an implementation in the package sandwich in the R system for statistical computing is described and it is shown how the suggested functions provide reusable components that build on readily existing functionality and how they can be integrated easily into new inferential procedures or applications. The toolbox contained in sandwich is extremely flexible and comprehensive, including specific functions for the most important HC and HAC estimators from the econometric literature. Several real-world data sets are used to illustrate how the functionality can be integrated into applications.

Achim Zeileis. "Object-oriented computation of sandwich estimators". In: *Journal of Statistical Software* 16.9 (Aug. 2006). DOI: 10.18637/jss.v016.i09.

Abstract: Sandwich covariance matrix estimators are a popular tool in applied regression modeling for performing inference that is robust to certain types of model misspecification. Suitable implementations are available in the R system for statistical computing for certain model fitting functions

only (in particular lm()), but not for other standard regression functions, such as glm(), nls(), or survreg(). Therefore, conceptual tools and their translation to computational tools in the package sandwich are discussed, enabling the computation of sandwich estimators in general parametric models. Object orientation can be achieved by providing a few extractor functions' most importantly for the empirical estimating functions' from which various types of sandwich estimators can be computed.