Intensive Longitudinal Data: A Dynamic Structural Equation Modeling Perspective

Daniel J. Bauer & Patrick J. Curran

APA Science Training Sessions: The Collection and Analysis of Intensive Longitudinal Data

centerstat.org

Intensive Longitudinal Data

Part 3: A Dynamic Structural Equation Modeling Perspective

Daniel J. Bauer & Patrick J. Curran
University of North Carolina at Chapel Hill

AMERICAN PSYCHOLOGICAL ASSOCIATION

CenterStat

Objectives

- Overview the DSEM approach to ILD analysis
- Start with review of standard SEM
- Incorporate elements of time series to extend to $N = 1$ DSEM
- Incorporate elements of MLM to extend to $N > 1$ DSEM
- Conclude with example applications and directions for future work

3.2

Basics of DSEM

- DSEM is an approach to analyzing ILD in which one is primarily (but not solely) concerned with **stable within-person processes**
  - that is, processes that are not changing systematically with the passage of time
- Example: Negative Affect and Alcohol Use
  - Does my negative affect today predict my alcohol use this evening above and beyond my alcohol use yesterday?
  - Does my drinking tonight predict my negative affect tomorrow, above and beyond my negative affect today?
- Requires large number of measurements over a short time interval to obtain good estimates of "dynamic" processes

3.3

What's a *Stable* Process?

- DSEM typically assumes the parameters governing the process under study are identical over all time points
  - Neither alcohol use nor negative affect are systematically increasing or decreasing over time
  - How negative affect predicts alcohol use (and vice versa) is also not changing systematically over time
- This assumption known as *stationarity*
  - Implies that mean, variance, and correlations of repeated measures (at a given lag) do not change over time
- Often a reasonable assumption for ILD but not always
  - Can potentially pre-process data ("de-trend") or expand model to include measure of time to better meet assumption, but we won't get into this here

3.4

Architecture of DSEM

- DSEM combines elements of three modeling traditions
  - Structural equation modeling (SEM)
  - Time series analysis
  - Multilevel modeling (MLM)
- We begin with introduction to SEM then extend to DSEM by incorporating time series and MLM

3.5

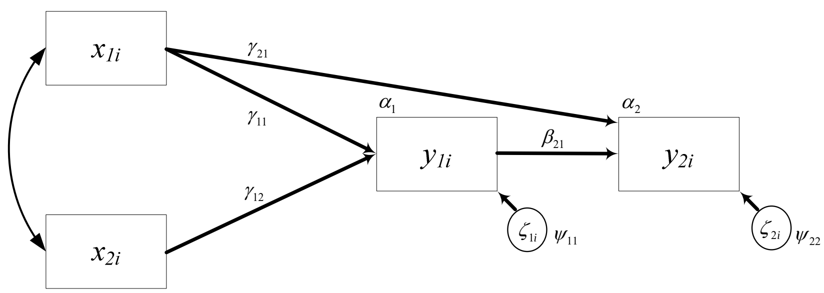
What is the Structural Equation Model?

- SEM is a general framework that subsumes a large number of models
  - *t*-test, ANOVA, MANOVA, regression, factor analysis can all be cast as an SEM
- Many extensions and advantages, but two will be focus for today:
  - modeling multiple dependent variables & complex chains of causal effects
  - estimating latent variables to account for measurement error
- SEM can be seen as a combination of path analysis with confirmatory factor analysis
  - also sub-models of the SEM
- Can see this visually through the depiction of SEM via path diagrams

3.6

Path Analysis

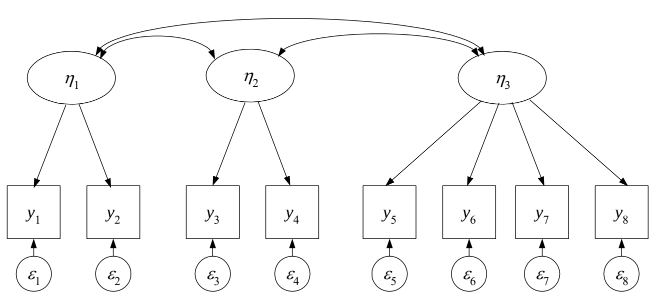
- Path analysis (a.k.a. *simultaneous equations model*) involves a multivariate model with structured relations among exclusively observed variables
  - Can have multiple x's and multiple y's
  - Can involve causal chains in which one y predicts another y



3.7

Confirmatory Factor Analysis

- In confirmatory factor analysis, we infer the presence of underlying, error-free latent variables (constructs) from correlated observed variables (items or indicators)

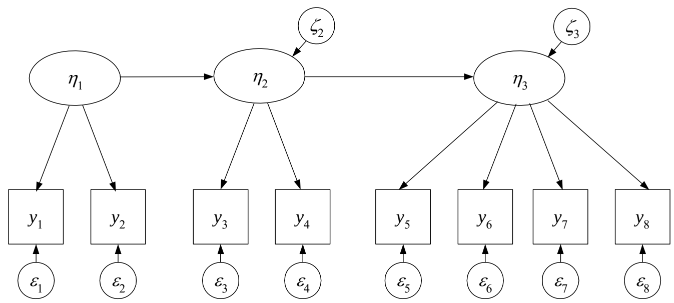


- Yet here we have not tested any structural model of interest

3.8

Full SEM

- In structural equation models, we combine the structural model of path analysis with the measurement model of CFA



- We are now estimating structural relations among latent variables that are unbiased by measurement error

3.9

Modeling Steps

1. Specification: What is the form of the model?
2. Identification: Possible to obtain unique parameter estimates?
3. Estimation: How obtain estimates of parameters of model?
4. Evaluation: How well does model fit data?
5. Re-specification: Should I modify my model?
6. Interpretation: Which effects significant? Size? Meaningful?

3.10

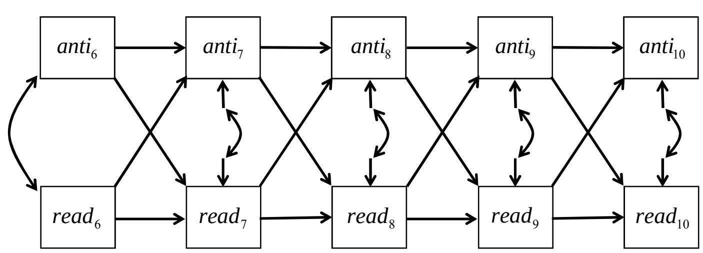
Traditional SEM With Repeated Measures

- Traditional SEM often fit to set of variables measured at one time point
  - many *cross-sectional* applications
- But there are well-developed SEMs for longitudinal "panel" data
  - e.g., consisting of say 3 to 6 assessments taken at 6 or 12 month intervals
- Two widely used approaches are the auto-regressive cross-lag panel model (ARCL) and the latent curve model (LCM)
- These are also easiest to see in path diagrams

3.11

Auto-Regressive Cross-Lagged Panel Model

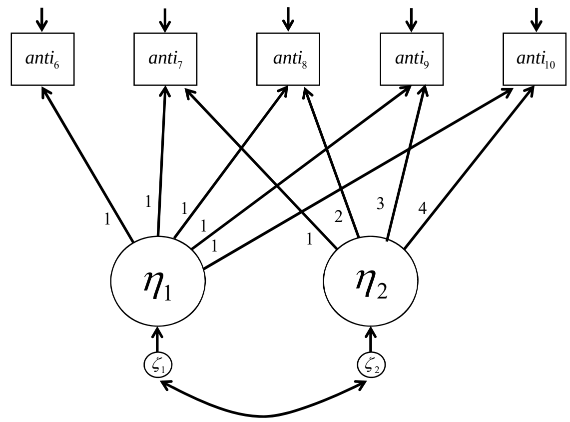
- Imagine we have measured *antisocial behavior* and *reading ability* in a sample of children between ages 6 and 10



- Motivating question is whether earlier antisocial behavior predicts later reading, and whether earlier reading predicts later antisocial behavior
  - because panel data, these relations are assessed at a spacing of one year

3.12

Latent Curve Model



- LCM uses precisely same data but estimates underlying *trajectory*
- Estimate means and variances of starting point and rate of change
- Can add predictors of latent factors or time-specific repeated measures
- Can also examine two or more constructs at same time
  - with or without lagged effects among the time-specific repeated measures

3.13

Moving From SEM to DSEM

- Both ARCL and LCM based on relatively small number of (typically) widely-spaced repeated measures
  - excellent for evaluating certain research hypotheses,
  - quite limited for assessing others, particularly those involving dynamic within-person processes that vary in magnitude over individuals
- Traditional longitudinal SEMs are not well suited to many features of ILD
  - high number of observations per person and complex patterns of serial dependence among repeated measures
- Indeed, conventional SEM often simply cannot be used with ILD
- Enter stage left: DSEM

3.14

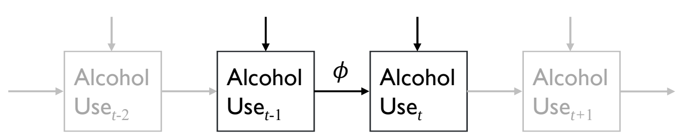
DSEM, $N = 1$

- Can begin with DSEM model for <u>one</u> unit measured intensively over time
- Leverages strengths of traditional time series analysis
  - One unit repeatedly measured high number of times to capture ***dynamics***
  - Focus usually on prediction of future state from past state of same process, strength of which is characterized as ***inertia*** or ***carry-over***
  - Can estimate prediction of future state of one process from past state of another to identify lead/lag effects, strength of which is characterized as ***spill-over***
  - Can model unexplained residual variance over time called ***innovations***
- But DSEM is fully multivariate, with greater flexibility in model specification and potential for incorporating latent variables
  - Will start with observed variable models and then note extensions

3.15

Why Dan Drinks

- Begin with simple univariate time series model to account for inertia in drinking behavior



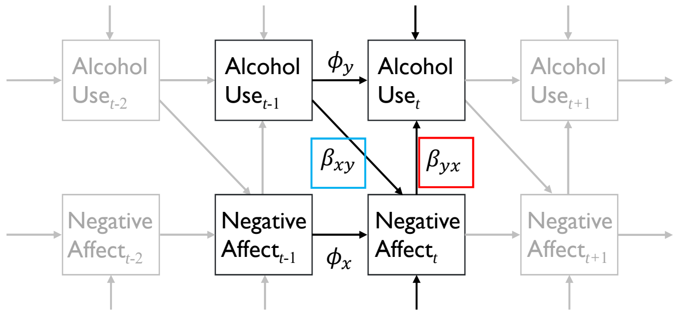
$$y_t = \nu_y + \underbrace{\phi y_{t-1}}_{\text{First-order autoregressive process}} + \zeta_t \qquad VAR(\zeta_t) = \psi$$

- Note similarity to traditional SEM-ARCL, but fit to highly dense data on just one person

3.16

Dan Drinks Because He is Sad ☹

- Expand to multivariate model to capture negative reinforcement process



Negative affect today may affect alcohol use tonight

Alcohol use tonight may lead to higher negative affect tomorrow

$$y_t = \nu_y + \phi_y y_{t-1} + \boxed{\beta_{yx} x_t} + \zeta_{yt} \qquad VAR(\zeta_{yt}) = \psi_y$$
$$x_t = \nu_x + \phi_x x_{t-1} + \boxed{\beta_{xy} y_{t-1}} + \zeta_{xt} \qquad VAR(\zeta_{xt}) = \psi_x$$

3.17

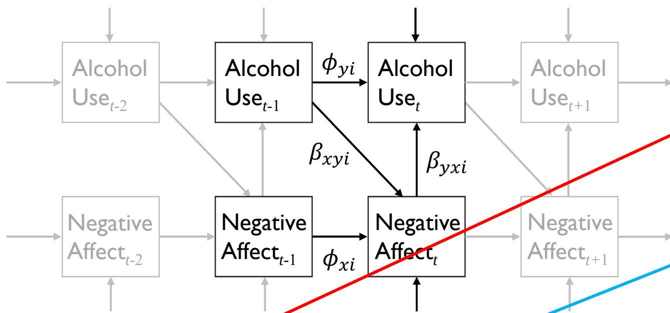
DSEM, N > 1

- As much as we may care about what motivates Dan to drink, we ultimately want to speak to a broader population of individuals
  - Move from ***idiographic*** to more ***nomothetic*** analysis to generalize inferences
- Can now incorporate aspects of MLM to allow for $N > 1$
  - Decompose variation in variables into within and between-person components
  - Look at between-person differences in parameters of within-person process
  - Generalize to population of individuals from which sample was drawn
- Like Level 1 / Level 2 representation in MLM, coefficients at Level 1 can be modeled as random effects that vary between persons
  - This is a big deal: carry-over and spill-over can vary randomly over individual
  - And if these vary randomly, might these be predictable?

3.18

Individual Differences in Process?

- Now allow parameters of model to vary over persons (i subscript)



Patrick drinks less than Dan in general (differences in intercepts)

Negative affect a stronger predictor of drinking for Dan than Patrick (differences in slopes)

These parameters had only a single value in prior model for just Dan's data

Level 1 (within-person):

$$y_{it} = \boxed{\nu_{yi}} + \phi_{yi} y_{it-1} + \boxed{\beta_{yxi}} x_{it} + \zeta_{yit} \qquad VAR(\zeta_{yt}) = \psi_{yi}$$
$$x_{it} = \nu_{xi} + \phi_{xi} x_{it-1} + \beta_{xyi} y_{it-1} + \zeta_{xit} \qquad VAR(\zeta_{xt}) = \psi_{xi}$$

3.19

Within- and Between-Person Variability

- Like in a standard MLM, each parameter of the within-person process model now has its own equation expressing individual differences:

Level 2 (between-person):

$$\nu_{yi} = \gamma_{y0} + u_{y0i}$$
$$\nu_{xi} = \gamma_{x0} + u_{x0i}$$
$$\vdots$$
$$\beta_{xyi} = \gamma_{xy} + u_{xyi}$$

Fixed effects: represent across-persons average values, we estimate their specific values

Random effects: represent between-persons differences, we estimate their variances and covariances

- Residual variances expressed via log-linear expressions (not shown)
- Can incorporate individual difference variables as predictors in these equations as well
  - e.g., does strength or direction of dynamics vary by sex or treatment condition

3.20

Addressing Measurement Error

- Unlike standard MLM, DSEM allows multiple DVs, auto-regressions, reciprocal effects, and the potential for random variability in these effects
  - Yet thus far we have only considered DSEMs with *observed variables*
- We have implicitly assumed these are variables are measured *without error*
  - Also standard assumption of regression, path analysis, ARCL, LCM, MLM, etc.
- However, if observed variables are measured with error, this can result in biased estimates of effects
  - stress, anxiety, gratitude, self-esteem, depression…measured error free? Nope.
  - measurement error in IVs can bias regression coefficients and in DVs standard errors
- Fortunately, DSEM also allows us to specify models with *latent factors*

3.21

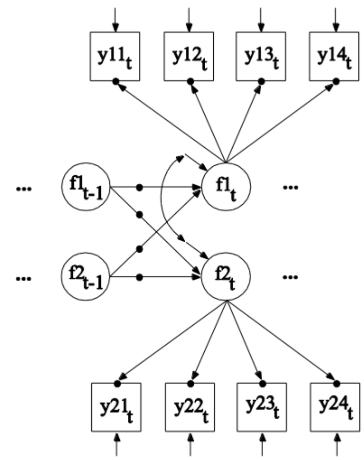
Addressing Measurement Error

- A key feature of the traditional SEM is the ability to estimate *latent factors* (also referred to as *latent variables*)
- Instead of computing a scale score using a set of items (e.g., a mean or a sum), the items themselves are used to infer an underlying latent factor
- Measurement error can then be explicitly estimated as part of the model, and thus "removed" at the level of the latent factors
  - regressions among latent factors are then unbiased (under assumptions)
- A version of this same approach can be incorporated into the DSEM

3.22

Multiple Indicator Latent Factors

- Within the DSEM, factors may reside at the within-person (Level 1) or between-person (Level 2) portions of the model, or both
- At level-1: estimate a latent factor for each *assessment point* instead of computing time-specific means of a set of items
  - e.g., latent factor for *depression* at each time point
- At level-2: estimate a latent factor for *person-level* characteristic that is used as predictors of dynamic processes
  - e.g., a client's *therapeutic alliance* at the start of an intervention
- Added complexities in both estimation and interpretation



3.23

Summary of DSEM Thus Far

- DSEM combines features of SEM, MLM, and time series that allows for:
  - large numbers of observations taken on modestly sized samples
  - multiple dependent variables
  - auto-regressive relations among repeated measures
  - reciprocal relations among repeated measures
  - estimation of individual variability in dynamic processes
  - prediction of individual variability in dynamic process
  - several other expansions not discussed here
- Helpful to see some recent applications

3.24

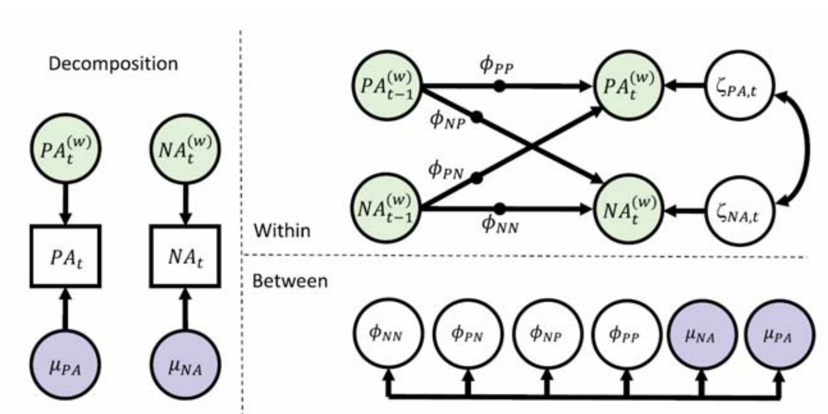
Example: Hamaker et al. (2018)

- Analyzed samples of n=101 younger and n=103 older subjects separately
  - each subject assessed daily for approximately 100 days
- Outcomes were composite scores of positive affect and negative affect
  - drawn from PANAS (Watson, Clark & Tellegen, 1988)
- Used Bayesian estimation to fit series of DSEMs in increasing complexity
- Motivating question was the dynamic and temporally-ordered relation between positive and negative affect
- We focus here on their Model 1, a vector autoregressive (VAR1) model
  - Best seen in diagram form

3.25

Example: Hamaker et al. (2018), con't

- Figure 2 from Hamaker et al. (2018)



3.26

Results

- A few of the key results:
  - Older adults had higher mean PA and lower mean NA
  - Older adults had stronger inertia for PA, weaker inertia for NA
  - On average, little cross-over from PA to NA in either group
  - Stronger positive cross-over from NA to PA for older adults
    - Higher NA today predicts greater PA tomorrow
- Basically, it's good to be old (yay for Patrick and Dan)
- Also found significant between-person differences in most effects
  - e.g., for a couple of old guys, Dan is grumpy (high NA mean and inertia) whereas Patrick is not (low NA mean and inertia)

3.27

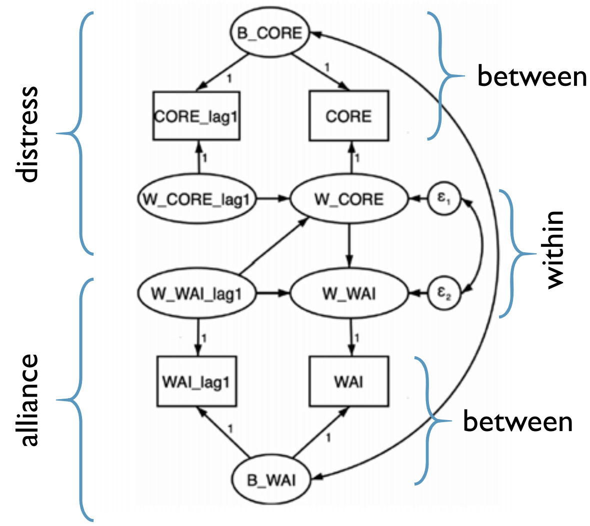
Example: Simons et al. (2020)

- Sample consisted of n=254 U.S. military veterans with repeated measures data collected in seven separate "bursts" spanning 1.5 years
  - assessment window varied over burst but spanned one to three weeks with random prompts at approximately two hour intervals
- Each subject provided a mean of 65 days of data across all bursts for a total of more than 90,000 person-by-time observations
- Series of DSEMs found that "*...veterans who experience greater sadness, anxiety, and anger may find these emotional states to be self-perpetuating and difficult alleviate. In addition, their experience of negative emotion may seem erratic and unpredictable.*" (p764).
- Notice the focus on dynamics – inertia and within-person variance -- that is not accessible in more traditional analytic approaches

3.28

Example: Gidhagen et al. (2021)

- Studied relation between *psychological distress* and *working alliance* in n=99 outpatients seeking treatment for substance use disorder (SUD)
- Repeated assessments taken in session ranging from 2 to 75 weeks
- Concluded "*...SUD patients' attachment orientation and type of abuse to a certain extent influence the associations between therapeutic alliance and outcome of psychological distress and substance use*" (p569)



3.29

Future Directions

- DSEM is a powerful and flexible methodology that allows us to test hypotheses in ways not previously possible
  - but this methodology is quite new and there are many issues to be resolved
- However, DSEM is also a target-rich environment for novel developments, rigorous evaluation, creative applications, and training and dissemination
- Indeed, as Hamaker et al. (2018, p 837) wonderfully concludes:

> "*We need psychometricians, applied statisticians, quantitative psychologists, and substantive researchers to explore this exciting new frontier, so that 10 years from now we can look back and smile at how little was known today.*"

3.30

Summary

- DSEM a hybrid of SEM, MLM, and time series models generally focused on:
  - estimation of intra-individual dynamic relations among measures over time
  - estimation of inter-individual variability in intra-individual dynamics
  - potential prediction of inter-individual variability by person-level covariates
  - expansion of all of above using latent factors to control for measurement error
- Despite great promise of DSEM, many significant issues yet to be resolved
- Limited software options, primarily M*plus*, although rapid development occurring throughout many methodological disciplines
- DSEM both a powerful new tool for studying dynamics and as a hot bed for future methodological research and dissemination

3.31

A Semi-Random Sampling of Resources

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 25*, 359-388.

Blanke, E. S., Neubauer, A. B., Houben, M., Erbas, Y., & Brose, A. (2021). Why do my thoughts feel so bad? Getting at the reciprocal effects of rumination and negative affect using dynamic structural equation modeling. *Emotion*. doi 10.1037/emo0000946

Gidhagen, Y., Holmqvist, R., Philips, B., & Falkenström, F. (2021). The role of the working alliance in psychological treatment of substance use disorder outpatients. *Psychotherapy Research, 31*, 557-572.

Grimm, K. J., & Ram, N. (2018). Latent growth and dynamic structural equation models. *Annual Review of Clinical Psychology, 14*, 55-89.

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research, 53*, 820-841.

Hamaker, E. L., Asparouhov, T., & Muthén, B. (2021). Dynamic structural equation modeling as a combination of time series modeling, multilevel modeling, and structural equation modeling. The handbook of structural equation modeling, 2nd Edition (R.H. Hoyle, Ed.) Guilford Press.

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods, 25*, 610.

Simons, J. S., Simons, R. M., Grimm, K. J., Keith, J. A., & Stoltenberg, S. F. (2021). Affective dynamics among veterans: Associations with distress tolerance and posttraumatic stress symptoms. *Emotion, 21*, 757.

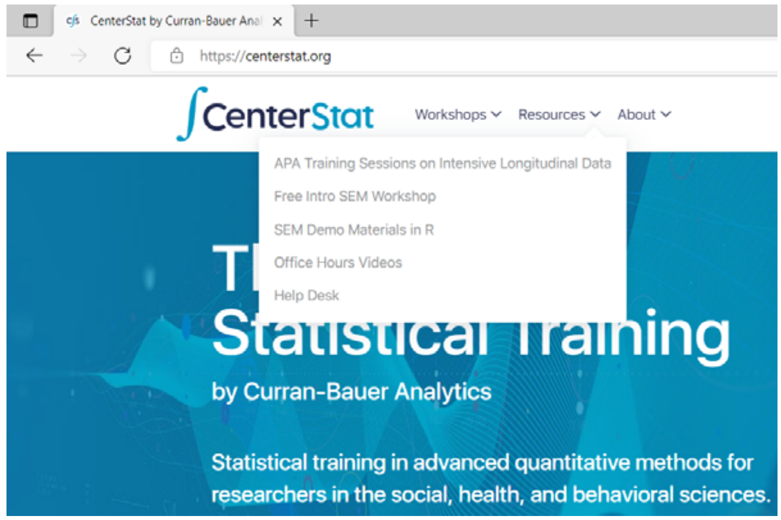
Weiss, N. H., Brick, L. A., Forkus, S. R., Goldstein, S. C., Thomas, E. D., Schick, M. R., ... & Sullivan, T. P. (2022). Modeling reciprocal relations between emotion dysregulation and alcohol use using dynamic structural equation modeling: A micro-longitudinal study. *Alcoholism: Clinical and Experimental Research, 46*, 1460-1471.

Zhou, L., Wang, M., & Zhang, Z. (2021). Intensive longitudinal data analyses with dynamic structural equation modeling. *Organizational Research Methods, 24*, 219-250.

3.32

Free Instructional Resources from CenterStat

- We offer a number of free instructional resources at ***centerstat.org***
  - **free three-day workshop** on structural equation modeling
  - tutorial lecture series on **YouTube**
  - written responses to submitted questions on **Help Desk**
  - informational posts on Twitter: **@curranbauer**
  - **informational emails** to which you can subscribe on the web page



3.33