



*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

*No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.*

# Calculating level-specific SEM fit indices for multilevel mediation analyses

W. Scott Comulada

Department of Psychiatry and Biobehavioral Sciences

Department of Health Policy and Management

University of California, Los Angeles

Los Angeles, CA

wcomulada@mednet.ucla.edu

**Abstract.** Stata’s `gsem` command provides the ability to fit multilevel structural equation models (SEM) and related multilevel models. A motivating example is provided by multilevel mediation analyses (MA) conducted on patient data from Methadone Maintenance Treatment clinics in China. Multilevel MA conducted through the `gsem` command examined the mediating effects of patients’ treatment progression and rapport with counselors on their treatment satisfaction. Multilevel models accounted for the clustering of patient observations within clinics. SEM fit indices, such as the comparative fit index and the root mean squared error of approximation, are commonly used in the SEM model selection process. Multilevel models present challenges in constructing fit indices because there are multiple levels of hierarchy to account for in establishing goodness of fit. Level-specific fit indices have been proposed in the literature but have not been incorporated into the `gsem` command. I created the `gsemgof` command to fill this role. Model results from the `gsem` command are used to calculate the level-specific comparative fit index and root mean squared error of approximation fit indices. I illustrate the `gsemgof` command through multilevel MA applied to two-level Methadone Maintenance Treatment data.

**Keywords:** `st0633`, `gsemgof`, `gsem`, `sem`, multilevel, structural equation model, mediation analysis, fit index

## 1 Introduction

Stata’s generalized structural equation model (SEM) command (`gsem`) extends the capabilities of the SEM command (`sem`) to fit multilevel SEM (Goldstein and McDonald 1988; Longford and Muthén 1992; Muthén 1994; Muthén and Satorra 1995) and related models to observations that are clustered across two or more levels of hierarchy. A motivating example is provided by Li et al. (2017) in their mediation analysis (MA) of baseline data from a randomized controlled trial. The trial evaluated an intervention designed to improve patient care provided by Methadone Maintenance Treatment (MMT) clinics in China. Patients at the first level were nested within clinics at the second level. Multilevel MA (MMA) examined the mediating effects of patients’ treatment progression and rapport with counselors on their treatment satisfaction. The `gsem` command was used to conduct MMA that accounted for patient observations at the first level clustered within clinics at the second level.

The `sem` command produces many popular fit indices to aid the model selection process, such as the comparative fit index (CFI; Bentler and Bonett 1980; Bentler 1990) and the root mean squared error of approximation (RMSEA; Browne and Cudeck 1993; Steiger 1990). Multilevel models present challenges in constructing fit indices because there are multiple levels of hierarchy to account for in establishing goodness of fit. A lack of fit, especially at higher levels, can be masked if the levels are combined to produce a fit index. As a result, there is a lack of consensus on suitable fit indices for multilevel models. Two main approaches have been proposed in the literature, including the segregating procedure by Yuan and Bentler (2007) and level-specific model fit evaluation by Ryu and West (2009). At present, fit indices are not produced by the `gsem` command. No fit indices are presented by Li et al. (2017).

In this article, I show how the `gsemgof` command, a new command that uses results from models fit through the `gsem` command, can be used to produce level-specific CFI and RMSEA fit indices. Both the level-specific and segregation procedures require additional computations to produce fit indices. The level-specific evaluation method is favored because it requires fewer programming steps and is less computationally intensive than the segregation procedure. We first review common fit indices in section 2. The level-specific evaluation method is discussed in section 3. Syntax for the `gsemgof` command is given in section 4. I then illustrate the level-specific evaluation method and the `gsem` command on MMA applied to the MMT data in section 5. Discussion follows in section 6.

## 2 Standard SEM fit indices

The likelihood-ratio (LR) test serves as a basic fit index. The log likelihood ( $\log \ell$ ) for a hypothesized model with a given set of parameters  $\theta$  is compared with the  $\log \ell$  for a saturated model with parameters  $\theta_S$ . The saturated model has the maximum number of parameters allowed by the data so that the model is not overidentified. Examples of hypothesized and saturated models are given in section 5. Under model assumptions, such as multivariate normality and a large sample size, the LR test follows a central  $\chi^2$  distribution and is expressed as

$$\chi^2_{\text{Hypothesized}} = -2 \times \left\{ \log \ell \left( \hat{\theta} \right) - \log \ell \left( \hat{\theta}_S \right) \right\} \quad (1)$$

When the hypothesized model provides an exact fit in line with the saturated model, the  $\chi^2_{\text{Hypothesized}}$  test statistic will take on a value of 0 in the population. Therefore, large  $\chi^2_{\text{Hypothesized}}$  values with a  $p$ -value  $\leq 0.05$  indicate poor model fit. A problem with the LR test is sample-size sensitivity. For example, the LR test may yield larger  $p$ -values in smaller samples, even though the hypothesized model does not provide adequate fit. Numerous fit indices have been developed to scale the LR test so that the fit indices are less sample-size dependent. This article discusses the CFI and RMSEA as two popular fit indices. Level-specific evaluation methods discussed in section 3 are applicable for the construction of other fit indices too.

The CFI incorporates both the  $\chi^2_{\text{Hypothesized}}$  test statistic and a  $\chi^2_{\text{Baseline}}$  statistic that compares the log  $\ell$  for a baseline model with the log  $\ell$  for a saturated model. The baseline model fits means and variances for all variables and covariances between exogenous (independent) variables. No additional parameters are estimated, such as paths from exogenous variables to an outcome. The degrees of freedom (d.f.) for the hypothesized (d.f.<sub>Hypothesized</sub>) and baseline models (d.f.<sub>Baseline</sub>) are also incorporated into the CFI calculation. The CFI is expressed as

$$\text{CFI} = 1 - \frac{\text{Max}\{(\chi^2_{\text{Hypothesized}} - \text{d.f.}_{\text{Hypothesized}}), 0\}}{\text{Max}\{(\chi^2_{\text{Baseline}} - \text{d.f.}_{\text{Baseline}}), 0\}} \quad (2)$$

Larger CFI values indicate better model fit. As a rule of thumb, CFI values  $\geq 0.90$  indicate good model fit.

The RMSEA is sometimes referred to as a “parsimony-adjusted” index because it adjusts the  $\chi^2_{\text{Hypothesized}}$  test statistic by both the d.f. and the total sample size ( $N$ ). The RMSEA is expressed as

$$\text{RMSEA} = \sqrt{\text{Max}\left\{\left(\frac{\chi^2_{\text{Hypothesized}} - \text{d.f.}_{\text{Hypothesized}}}{\text{d.f.}_{\text{Hypothesized}}(N - 1)}\right), 0\right\}} \quad (3)$$

Similar to the chi-squared test statistic, RMSEA values closer to 0 indicate better fit. RMSEA values  $< 0.08$  indicate good model fit.

### 3 Level-specific fit indices for multilevel models

For the sake of brevity, indices are discussed in terms of two-level models, but the method is applicable to multilevel models with additional levels. Chi-squared statistics for level-specific indices are formulated in a similar manner as the chi-squared test statistic shown in (1). Let  $\chi^2_{\text{Hypothesized},1}$  represent a chi-squared test statistic for a level-1 hypothesis, such as a hypothesized path between observed variables. Level-2 hypotheses correspond to random effects or latent variables, such as a hypothesized covariance between random effects.

Let  $\theta_1$  and  $\theta_2$  represent sets of parameters that correspond to level-1 and level-2 hypotheses, respectively. Let  $\theta_{1,S}$  and  $\theta_{2,S}$  represent sets of parameters for models that are saturated at levels 1 and 2, respectively. The  $\chi^2_{\text{Hypothesized},1}$  test statistic for a level-1 evaluation is

$$\chi^2_{\text{Hypothesized},1} = -2 \times \left\{ \log \ell \left( \hat{\theta}_1, \hat{\theta}_{2,S} \right) - \log \ell \left( \hat{\theta}_{1,S}, \hat{\theta}_{2,S} \right) \right\} \quad (4)$$

The chi-squared test statistic for a level-2 evaluation is expressed as

$$\chi^2_{\text{Hypothesized},2} = -2 \times \left\{ \log \ell \left( \hat{\theta}_{1,S}, \hat{\theta}_2 \right) - \log \ell \left( \hat{\theta}_{1,S}, \hat{\theta}_{2,S} \right) \right\}$$

The chi-squared test statistic for the baseline model that is used to calculate the level-specific CFI also incorporates a partially saturated model. For example, the log likelihood for a level-1 baseline model,  $\log \ell(\hat{\theta}_1, \hat{\theta}_{2,S})$ , incorporates baseline parameters at

level 1 ( $\theta_1$ ) and parameters for a model that is saturated at level 2 ( $\theta_{2,S}$ ). Level-specific CFI and RMSEA formulas are similar to their standard formula counterparts in (2) and (3), respectively, except that level-specific chi-squared test statistics and d.f. replace the chi-squared statistics and d.f. for models that are not multilevel. The sample size is also adjusted in the level-specific RMSEA formula. For level-1 and level-2 evaluations,  $N - 1$  is replaced by  $N - J$  and  $J$ , respectively, where  $J$  is the number of level-2 clusters.

## 4 Syntax of gsemgof

The **gsemgof** command is run after an MMA is fit to two-level data by using the **gsem** command. The MMA should include an equation for a primary outcome and an equation for each mediator. Random effects need to be specified for the outcome in each equation. Importantly, level-1 fit indices require that level 2 is saturated, and vice versa for level-2 fit indices. The **gsemgof** command syntax is

```
gsemgof level [df]
```

where *level* is specified as 1 or 2 to request level-1 or level-2 fit indices, respectively. If level-2 fit indices are requested, *df* is specified to indicate the d.f. for the hypothesized model. Otherwise, *df* is left blank. The d.f. is automatically calculated by the **gsemgof** command for level-1 fit indices. Sample syntax for an MMA with a single outcome and two mediators is shown in section 5.

The **gsemgof** command works by modifying **gsem** command-line text for the hypothesized model to create command lines for additional models that are needed to calculate fit indices, such as a model that is saturated at levels 1 and 2. Additional models are run using the **gsem** and **sem** commands. Stored results in **e()** and **r()**, such as log likelihoods, are extracted, saved as scalar quantities, and plugged into fit index formulas. The following steps occur when level-1 fit indices are requested (that is, when *level* is set to 1).

1. Three numbers are extracted from **e()** and saved for later calculations: The total sample size ( $N$ ); the total number of dependent variables, including the primary outcome and mediators; and the log likelihood for the hypothesized model [ $\log \ell(\hat{\theta}_1, \hat{\theta}_{2,S})$ ].
2. The name of the cluster variable is identified in the **gsem** command-line text and is used to calculate the number of clusters in the dataset ( $J$ ).
3. The **gsem** command fits an MMA model that is saturated at levels 1 and 2. The log likelihood [ $\log \ell(\hat{\theta}_{1,S}, \hat{\theta}_{2,S})$ ] that is stored in **e()** is extracted and saved for later calculations.
4. The **gsem** command fits a level-1 baseline model that is saturated at level 2. The  $\log \ell$  is saved for later calculations.

5. The d.f. for chi-squared test statistics is not produced by the `gsem` command. Appropriate d.f. can be obtained by first fitting the hypothesized model without level-2 random effects through the `sem` command. Afterward, fit statistics are requested using `estat gof`, and the d.f. for the hypothesized and baseline model evaluations are extracted from `r()`. Level 2 can be ignored in the calculation of the d.f. for level-1 tests because level 2 is saturated.
6. Chi-squared statistics are calculated from log likelihoods by using (4).
7. Finally, fit indices are calculated using stored chi-squared statistics, d.f.,  $N$ , and  $J$ .

Calculations for level-2 fit indices follow the same steps except that the d.f. in step 5 cannot be retrieved through the `sem` command. The d.f. for the level-2 baseline chi-squared statistic is calculated by the `gsemgof` command. The level-2 baseline d.f. is equal to the number of random-effects covariance parameters because they are not estimated and are assumed to be 0. Level-2 baseline model parameters that get estimated are the random-effects variances. Level 1 is saturated and does not factor into the d.f. calculations. If  $m$  is the number of random-effects variance parameters, then the level-2 baseline d.f. is  $m(m - 1)/2$ . The d.f. for the hypothesized model is specified in the `gsemgof` command by the user.

## 5 Level-specific fit index examples using the MMT data

The `gsemgof` command is illustrated through MMA applied to the MMT data. There are 2,427 patient observations and 68 clusters in the dataset. Analyses modeled the mediating effects of both treatment progression and counselor rapport on treatment satisfaction. This section focuses on the relationship between treatment progression and satisfaction at levels 1 and 2. Figure 1 presents diagrams for MMA that are evaluated. Diagrams were produced through Stata's SEM Builder visualization tool. Two level-1 evaluations are conducted. Figure 1a represents the model presented in Li et al. (2017). A path from treatment progression to satisfaction is modeled at level 1 through progression (`progress`) and `satisfaction` variables measured at the patient level. A path is also included from counselor rapport (`rapport`) to `satisfaction` at level 1.

The relationship between treatment progression and satisfaction is modeled at level 2 through the covariance between clinic-level random effects for treatment progression and satisfaction. Level 2 is saturated. The model includes random effects for both mediators and the outcome and three covariances between all possible pairs of random effects. There are six exogenous variables in the model: `age`, a yes-no indicator for `male` gender, a yes-no indicator for being `married`, the duration of MMT treatment (`mmt.duration`), a depressive symptoms score (`depression`), and a social and environmental support score (`env.support`). Figure 1b excludes the path from treatment progression to satisfaction.

A single level-2 evaluation is conducted. As shown in figure 1c, level 1 is saturated, and the covariance between random effects for treatment progression and satisfaction

is fixed at 0. Models represented by figures 1a, 1b, and 1c are compared with a model that is fully saturated at levels 1 and 2 (figure 1d). Models are fit using Stata/SE 16. Syntax for level-1 and level-2 fit indices is presented in sections 5.1 and 5.2, respectively. Results are presented in section 5.3.

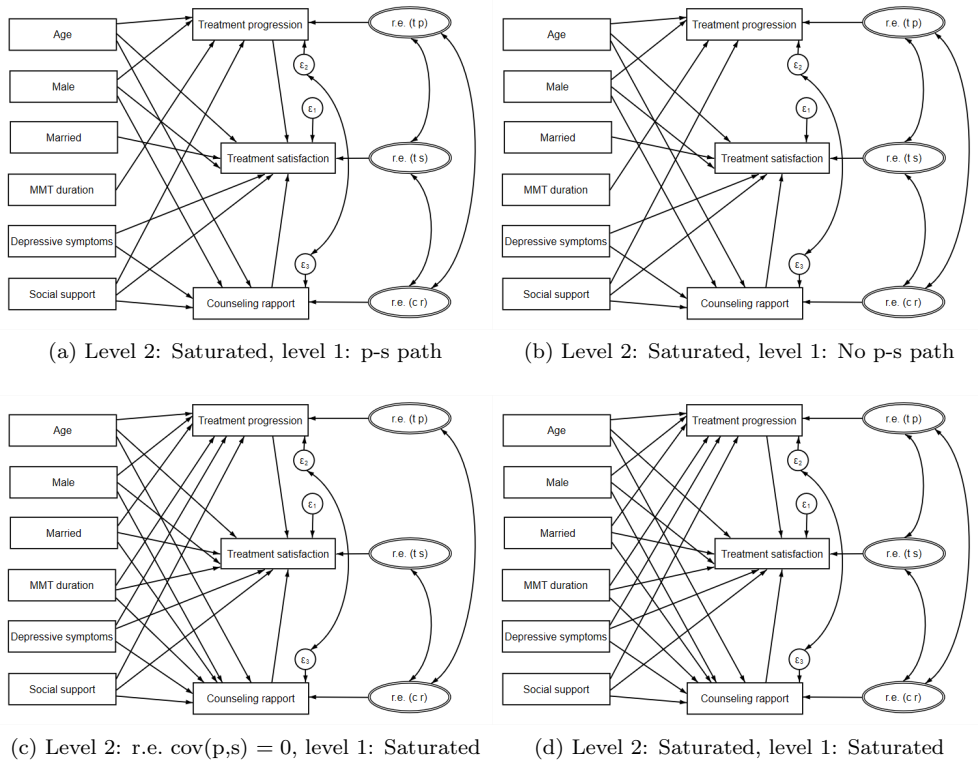


Figure 1. MMA diagrams for predictors of treatment satisfaction mediated by treatment progression and counseling rapport. Diagram 1a represents the hypothesized model in Li et al. (2017). Diagram 1b represents an MMA that excludes the path from treatment progression to satisfaction (p-s); diagram 1c represents an MMA that sets the random-effects covariance (r.e. cov) between treatment progression and satisfaction to 0; and diagram 1d represents an MMA that fits a saturated model where all possible r.e. covariances at level 2 and paths at level 1 are modeled. In line with standard SEM diagram representation, covariances between exogenous variables (for example, between age and male gender) are modeled but not represented in the diagrams.

## 5.1 Level-1 fit indices

In this section, we view the syntax for level-1 evaluations of models represented by diagrams 1a and 1b in figure 1. I first read in a dataset containing the MMT data.

```
. use mmt
```

I then fit the hypothesized model in diagram 1a by using the `gsem` command. A grouping variable (`clinic`) is used to index random effects by clinic. The `gsemgof` command follows each `gsem` command with *level* set to 1.

```
. // Model (a): Hypothesized model
. gsem (satisfaction <- progress rapport age male married depression
>   env_support Rs[clinic])
>   (progress <- age male mmt_duration env_support Rp[clinic])
>   (rapport <- age male depression env_support Rr[clinic]),
>   cov(e.progress*e.rapport)
(output omitted)
. gsemgof 1
(output omitted)
. // Model (b): Hypothesized model without path from progress to satisfaction
. gsem (satisfaction <- rapport age male married depression
>   env_support Rs[clinic])
>   (progress <- age male mmt_duration env_support Rp[clinic])
>   (rapport <- age male depression env_support Rr[clinic]),
>   cov(e.progress*e.rapport)
(output omitted)
. gsemgof 1
(output omitted)
```

## 5.2 Level-2 fit indices

In this section, we view syntax for a level-2 evaluation of the model represented in diagram 1c of figure 1. The random-effects covariance structure for the hypothesized model is specified through covariance matrix `Vre` and incorporated into the `gsem` model specification. The `gsemgof` command follows with *level* set to 2. In contrast to requests for level-1 tests, the user specifies the d.f. for the  $\chi^2_{\text{Hypothesized},2}$  test statistic. Referring to the level-2 hypothesized model in diagram 1c and the fully saturated model in diagram 1d, the hypothesized model has one fewer parameter. Therefore, the d.f. is set to 1. The `gsemgof` command automatically sets the d.f. for the  $\chi^2_{\text{Baseline},2}$  test statistic to 3, the number of random-effects covariance parameters [that is,  $m(m-1)/2 = 3(2)/2 = 3$ ].



```

. // Model (c): Level 2 hypothesis, no random-effects covariance between
> progress and satisfaction
. matrix Vre = (., 0, . \ 0, ., . \ ., ., .)
. gsem (satisfaction <- progress rapport age male married mmt_duration
>      depression env_support Rs[clinic])
>      (progress <- age male married mmt_duration depression env_support
>      Rp[clinic])
>      (rapport <- age male married mmt_duration depression env_support
>      Rr[clinic]), covstr(Rs[clinic] Rp[clinic] Rr[clinic],
>      fixed(Vre)) cov(e.progress*e.rapport)
(output omitted)
. gsemgof 2 1
(output omitted)

```

### 5.3 Results

Table 1 shows fit indices based on models fit to the MMT data. Indices in the left-hand column are based on MA that ignore clustering at level 2. Models are fit through the `sem` command. Indices in the right-hand column are based on MMA conducted through the `gsem` command. Fit indices are calculated by the `gsemgof` command.

Table 1. Model fit indices obtained from MA and MMA conducted on MMT study data. Level-1 indices evaluate fit for models with and without paths from treatment progression to satisfaction (p-s). Level-2 indices evaluate fit for a multilevel model where the random-effects covariance between treatment progression and satisfaction is fixed at 0.

	MA	MMA
<b>Level 2: Saturated</b>		
Level 1: p-s path		
$\chi^2(5)$	8.524	10.308
CFI	0.999	0.999
RMSEA	0.017	0.021
Level 1: No p-s path		
$\chi^2(6)$	229.650**	152.144**
CFI	0.931	0.998
RMSEA	0.124	0.102
<b>Level 1: Saturated</b>		
Level 2: r.e. cov(p,s) = 0		
$\chi^2(1)$		4.334*
CFI		0.846
RMSEA		0.221

NOTE: \* $p < 0.05$ ; \*\* $p < 0.01$

Overall, standard and level-1 fit index values are similar. For level-1 evaluations, both the standard fit indices in the left-hand column and the level-1 fit indices in the right-hand column indicate that the model fitting a path from treatment progression to satisfaction provides a better fit than the model where the path is fixed at 0. The chi-squared test statistics and RMSEA are smaller for the model with the path. The CFI is larger for the model with the path, but neither the standard CFI nor the level-1 CFI values are small enough to indicate a lack of fit for the model without the path.

Table 1 highlights a key feature of the level-specific evaluation method in its ability to evaluate level-2 fit. Level-2 fit indices indicate that the covariance between random effects for treatment progression and satisfaction should be modeled. A model fixing the covariance at 0 results in a statistically significant chi-squared test at level 2. The CFI value of 0.846 and the RMSEA value of 0.221 are also outside the ranges of values that indicate good model fit.

## 6 Discussion

In this article, I showed how the `gsemgof` command can be used to calculate level-specific CFI and RMSEA fit indices. I illustrated calculations on MMA applied to the MMT data. Standard fit indices obtained from MA that did not include random effects and level-1 fit indices obtained from MMA produced similar values. This is not completely surprising. Simulation studies by Ryu and West (2009) showed that level 1 tends to dominate in determining fit index values. However, it should not be assumed that the standard and level-specific approaches will always lead to similar conclusions for level-1 evaluations. Moreover, level-2 evaluations cannot be conducted using standard fit indices. In regard to the MMT data, level-specific indices indicated that relationships between treatment progression and satisfaction should be captured at both the patient (level 1) and clinic (level 2) levels.

Given the benefits of the level-specific evaluation method, a few caveats are in order. The d.f. is an important component of the fit index calculations and may need to be calculated by hand for more-complex models than the MMA that were presented in this article. We refer to d.f. for level-1 tests shown in table 1 to illustrate hand calculations. The  $\chi^2_{\text{Hypothesized},1}$  test statistic based on diagram 1a of figure 1 has 5 d.f. for five possible paths from exogenous variables to mediators and the outcome that are excluded from figure 1a. The level-1 evaluation for diagram 1b excludes an additional path from treatment progression to satisfaction and has 6 d.f.

The d.f. for the  $\chi^2_{\text{Baseline},1}$  test (not shown in table 1) is the difference in the number of parameters between the partially saturated baseline model and the fully saturated model in diagram 1d. The partially saturated baseline model excludes 18 paths from exogenous variables to mediators and the outcome, two paths from mediators to the outcome, and a covariance between the two mediators. The d.f. is 21. An alternative method to count d.f. is outlined in Acock (2013). Between the six exogenous variables, two mediators, and an outcome, there are nine variables contributing to the variance-covariance matrix for the fully saturated model. The variance-covariance matrix has

$9(10)/2 = 45$  parameters. The baseline model has 24 variance and covariance parameters, including 21 parameters for the six exogenous variables, a variance for each mediator, and an outcome variance. The  $\chi^2_{\text{Baseline}}$  test has  $45 - 24 = 21$  d.f. General guidelines for calculating the d.f. for SEM can be found in Rigdon (1994) and Acock (2013).

The RMSEA formula also incorporates the sample size that is set to  $N - J$  for level-1 evaluations and  $J$  for level-2 evaluations. Balance in the number of observations per cluster is implied by the formula. I was fortunate in that the MMT study design aimed to sample the same number of patients per clinic. Moreover, missing data was minimal. As a result, there was a large degree of balance in the number of patients per clinic. More complex RMSEA formulas may be needed to adequately assess model fit in the presence of large imbalances in the number of observations per cluster.

## 7 Acknowledgments

My time was supported by National Institutes of Health grants (P30MH058107; U19HD089886). MMT data were collected through a National Institute on Drug Abuse grant (R01DA033130).

## 8 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 21-1
. net install st0633      (to install program files, if available)
. net get st0633          (to install ancillary files, if available)
```

## 9 References

- Acock, A. C. 2013. *Discovering Structural Equation Modeling Using Stata*. Rev. ed. College Station, TX: Stata Press.
- Bentler, P. M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin* 107: 238–246. <http://doi.org/10.1037/0033-2909.107.2.238>.
- Bentler, P. M., and D. G. Bonett. 1980. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* 88: 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>.
- Browne, M. W., and R. Cudeck. 1993. Alternative ways of assessing model fit. In *Testing Structural Equation Models*, ed. K. A. Bollen and J. S. Long, 136–162. Newbury Park, CA: SAGE.
- Goldstein, H., and R. P. McDonald. 1988. A general model for the analysis of multilevel data. *Psychometrika* 53: 455–467. <https://doi.org/10.1007/BF02294400>.

- Li, L., W. S. Comulada, C. Lin, J. Hsieh, S. Luo, and Z. Wu. 2017. Factors related to client satisfaction with methadone maintenance treatment in China. *Journal of Substance Abuse Treatment* 77: 201–206. <https://doi.org/10.1016/j.jsat.2017.02.012>.
- Longford, N. T., and B. O. Muthén. 1992. Factor analysis for clustered observations. *Psychometrika* 57: 581–597. <https://doi.org/10.1007/BF02294421>.
- Muthén, B. O. 1994. Multilevel covariance structure analysis. *Sociological Methods & Research* 22: 376–398. <https://doi.org/10.1177/0049124194022003006>.
- Muthén, B. O., and A. Satorra. 1995. Complex sample data in structural equation modeling. *Sociological Methodology* 25: 267–316. <https://doi.org/10.2307/271070>.
- Rigdon, E. E. 1994. Calculating degrees of freedom for a structural equation model. *Structural Equation Modeling: A Multidisciplinary Journal* 1: 274–278. <https://doi.org/10.1080/10705519409539979>.
- Ryu, E., and S. G. West. 2009. Level-specific evaluation of model fit in multilevel structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal* 16: 583–601. <https://doi.org/10.1080/10705510903203466>.
- Steiger, J. H. 1990. Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research* 25: 173–180. [https://doi.org/10.1207/s15327906mbr2502\\_4](https://doi.org/10.1207/s15327906mbr2502_4).
- Yuan, K.-H., and P. M. Bentler. 2007. Multilevel covariance structure analysis by fitting multiple single-level models. *Sociological Methodology* 37: 53–82. <https://doi.org/10.1111/j.1467-9531.2007.00182.x>.

#### About the author

W. Scott Comulada is an associate professor-in-residence in the Department of Psychiatry and Biobehavioral Sciences and in the Department of Health Policy and Management at the University of California, Los Angeles. He is also a methods core co-director for the Center for HIV Identification, Prevention and Treatment Services (CHIPTS; P30MH058107) and an analytic core project lead for an Adolescent Trials Network U19 (U19HD089886).