Power and Bias in Three-Form Planned Missing Designs for Longitudinal Mediation

Richard M. Kinai, Terrence D. Jorgensen, Graham G. Rifenbark, Fan Jia, Alexander M. Schoemann, Todd D. Little, Wei Wu

PLANNED MISSING DATA DESIGNS

Modern missing data methods minimize most problems associated with missing data (e.g., power loss and biased parameter estimates) and they present new opportunities such as research designs that purposefully incorporate missing data (Palmer & Royall, 2010).

Planned missing designs can minimize participant fatigue, practice effects, and dropout rates, while also reducing research costs. This has important implications for designing longitudinal studies where cost and dropout are major concerns. One popular planned missing data design with potential benefits for longitudinal studies is the 3-form design. In the 3-form design the sample is randomly split into 3 subgroups, each with its own pattern of missing data.

SIMULATION

We simulated complete data from a 3-latent variable 3-timepoint mediation model with N = 500, 200 datasets per condition. We varied:

- factor loadings (from .7 to .85),
- autoregressive paths (from .4 to .9),
- cross-lagged effects (from 0 to .4),
- within-time covariances (from .2 to .5), and
- numbers of variables with no missingness (1-6 per time point; all other variables had 1/3 missing data)

MCAR missingness was imposed according to a 3-form design (Graham et al., 2006), where the “X” set (items that all participants receive, with no missingness) contained from 1-6 items.

Models were fit using FIML estimation

THREE-FORM DESIGN

<table>
<thead>
<tr>
<th>Form</th>
<th>Common</th>
<th>Variable</th>
<th>Variable</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Set X</td>
<td>Set A</td>
<td>Set B</td>
<td>Set C</td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td>X</td>
<td>--</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>--</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

RESULTS

Parameter Estimates
- As expected, parameter estimates were unbiased, because the missingness mechanism is MCAR

Standard Errors
- Standard errors were also unbiased (compared to empirical standard errors computed as the standard deviation of the parameter estimates across replications)
- Standard errors increased as the number of items in the X block decreased (i.e., as the % missing data increased)

Power
- As the % missing data increased, power decreased very slightly (corresponding to slight increase in SEs)
- e.g., the drop in power to detect a cross-lagged path of .1 decreased from 54% with 11% missing data to 52% with 30% missing data

<table>
<thead>
<tr>
<th>No. items in X block</th>
<th>Percent missing (%)</th>
<th>Pop. cross-lag path</th>
<th>Average Est.</th>
<th>Empirical SE</th>
<th>Average Est. SE</th>
<th>Bias</th>
<th>SE Bias</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.1</td>
<td>.10</td>
<td>.10</td>
<td>.05</td>
<td>.05</td>
<td>.02</td>
<td>.03</td>
<td>.52</td>
</tr>
<tr>
<td>2</td>
<td>14.8</td>
<td>.10</td>
<td>.10</td>
<td>.05</td>
<td>.05</td>
<td>.02</td>
<td>.05</td>
<td>.54</td>
</tr>
<tr>
<td>3</td>
<td>18.5</td>
<td>.10</td>
<td>.10</td>
<td>.05</td>
<td>.05</td>
<td>.02</td>
<td>.05</td>
<td>.54</td>
</tr>
<tr>
<td>4</td>
<td>22.2</td>
<td>.10</td>
<td>.10</td>
<td>.05</td>
<td>.05</td>
<td>.02</td>
<td>.05</td>
<td>.53</td>
</tr>
<tr>
<td>5</td>
<td>25.9</td>
<td>.10</td>
<td>.10</td>
<td>.05</td>
<td>.05</td>
<td>.03</td>
<td>.08</td>
<td>.53</td>
</tr>
<tr>
<td>6</td>
<td>29.6</td>
<td>.10</td>
<td>.10</td>
<td>.05</td>
<td>.05</td>
<td>.03</td>
<td>.07</td>
<td>.54</td>
</tr>
</tbody>
</table>

DISCUSSION

- With between 11 and 30% MCAR missing through a planned missing design, there was little noticeable effect of missing data on power
- Parameter estimates and standard errors showed very little bias when FIML was used for SEM estimation
- With N = 500, the only parameters to show less than perfect power were those valued less than .4; future investigations should be done with smaller samples
- Initial investigations suggest that smaller samples may be subject to convergence failures with planned missing designs

REFERENCES


Funding generously provided by:  

July 11, 2012 • The 77th Annual Meeting of the Psychometric Society