We gratefully acknowledge support from the Institute of Educational Sciences, U.S. Department of Education (through Grant R305D150007)

SYNTESIZING SINGLE-CASE STUDIES VIA MULTILEVEL MODELS: LIMITATION OF MODEL COMPLEXITY

Ke Cheng
Zhiyao Yi
John Ferron
Mariola Moeyaert
S. Natasha Beretvas
Wim Van den Noortgate

SINGLE-CASE EXPERIMENTAL DESIGNS (SCED)

- Researchers evaluate the effect of interventions by comparing repeated measurements on the same individual over time across one or more treatment conditions.
  - AB design (interrupted time-series design);
  - Withdrawal designs (e.g., ABA, ABAB, ABAC);
  - Changing criterion design (e.g., ABB’, ABB’B’’);
  - Multiple probe design;
  - Alternating treatments design (multi-element design);
  - Multiple baseline design.
SYNTHEIZING MULTIPLE SCED STUDIES

- Level 1: observations nested in participants
- Level 2: participants nested in studies
- Level 3: studies
Multilevel models allow researchers to ...

- Estimate the average treatment effect across cases and studies
- Quantify the changes in the treatment effect over time in treatment
- Estimate the variation in the treatment effect across cases within studies and across studies
- Examine potential moderators of the treatment effect

A TYPICAL MULTILEVEL MODEL

\[ Y_{ijk} = \beta_0 + \beta_1 \text{Phase} + \beta_2 \text{Time} + \beta_3 \text{Phase} \times \text{Time} + e_{ijk} \]

\[
\begin{align*}
\beta_0 &= \text{slope}_A \\
\beta_1 &= \Delta \text{slope} \\
\beta_3 &= \text{slope A}
\end{align*}
\]
A TYPICAL MULTILEVEL MODEL

\[ Y_{ijk} = \beta_{0jk} + \beta_{1jk} \text{Phase} + \beta_{2jk} \text{Time} + \beta_{3jk} \text{Phase} \times \text{Time} + e_{ijk} \]

\[ \begin{align*}
\beta_{0jk} &= \theta_{00k} + u_{0jk} \\
\beta_{1jk} &= \theta_{10k} + u_{1jk} \\
\beta_{2jk} &= \theta_{20k} + u_{2jk} \\
\beta_{3jk} &= \theta_{30k} + u_{3jk}
\end{align*} \]

\[ \begin{bmatrix}
\theta_{00k} \\
\theta_{10k} \\
\theta_{20k} \\
\theta_{30k}
\end{bmatrix} \sim N\left( \begin{bmatrix}
\sigma_{\theta0}^2 & 0 & 0 & 0 \\
0 & \sigma_{\theta1}^2 & 0 & 0 \\
0 & 0 & \sigma_{\theta2}^2 & 0 \\
0 & 0 & 0 & \sigma_{\theta3}^2
\end{bmatrix} \right) \]

TYPICAL MULTILEVEL MODEL

- the trends are non-existent or linear
- the errors within a case are homogenous and either independent or first-order autoregressive
- the case-level errors are uncorrelated
- the study-level errors are uncorrelated

TYPICAL MULTILEVEL MODEL

Estimation REML with Kenward Roger adjusted SEs and DFs

Assuming the model is correctly specified:
• Fixed effects are unbiased (e.g., average treatment effect)
• Fixed effect inferences are accurate (e.g., 95% CIs cover 95% of the time)
• Variance components are biased (e.g., across case variance in effect)

Bayesian Estimation

Reduces bias in variance components (if priors chosen well)


WHAT IF DATA ARE MORE COMPLEX?

Estimating logistic models creates problems given the small sample size

WHAT IF DATA ARE MORE COMPLEX?

Maybe fit a quadratic trajectory to the intervention phase?

WHAT IF DATA ARE MORE COMPLEX?

Maybe fit a piecewise trajectory to the intervention phase?
WHAT IF DATA ARE MORE COMPLEX?

What if the covariance structure at level-1 is not independent and homogeneous, but rather it is:

- First-order autoregressive
- Heterogeneous across phases
- Heterogeneous across cases

What if the covariance structure at level-2 is not diagonal, but unstructured?

What if the covariance structure at level-3 is not diagonal, but unstructured?

PURPOSE OF OUR STUDY

To compare alternative multilevel models for analyzing a series of multiple baseline studies that are characterized by multiple complexities:

- treatment phase trajectories that are non-linear
- within-case (level-1) errors that are autocorrelated and heterogeneous across phases and across cases
- level-2 and level-3 errors that may have an unstructured covariance structure
OVERVIEW OF METHODS

- Monte Carlo Study

We generated data for a series of multiple-baseline studies where the underlying model had multiple complexities.

We analyzed each generated data set with alternative models that differed in complexity.

We compared the models on convergence rates, parameter bias, and confidence interval coverage.

REVIEWED PUBLISHED APPLIED STUDIES TO MOTIVATE DATA GENERATION MODEL


Data were generated based on a 3-level model

\[ Y_{ijk} = 0.20 + u_{jk} + v_k + \frac{0.5}{1 + e^{-1.5(t_{imec})}} \times \text{Phase} + e_{ijk} \]

\( e_{ijk} \) Distributed first order autoregressive (\( \rho = .2 \))

Baseline variance of first case: \( \sigma^2_{A1} = .0025 \)

Heterogeneous across phases: \( \frac{\sigma^2_4}{\sigma^2_B} = 2 \)

Heterogeneous across cases: \( \frac{\sigma^2_4}{\sigma^2_1} = 4 \)

\( u_{jk} \) Distributed normal with variance, \( \sigma^2_u = .0025 \)

\( v_k \) Distributed normal with variance, \( \sigma^2_v = .0025 \)
• Design Factors

• Level 1: number of observation per participant (16 & 32)
• Level 2: number of participant per study (4 & 8)
• Level 3: number of study (10 & 30)

<table>
<thead>
<tr>
<th>Conditions</th>
<th># of Observation</th>
<th># of Participant</th>
<th># of Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>32</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>32</td>
<td>8</td>
<td>30</td>
</tr>
</tbody>
</table>

Number of replication in each condition: 2000
MODELS FOR ANALYZING

Model 1
• Quadratic with simple errors structure:
  Level-1 error:
  • independence;
  • homogeneity.
  Level-2 error:
  • not correlated.
  Level-3 error:
  • not correlated.

Model 2
• Quadratic with complex errors structure:
  Level-1 error:
  • autocorrelated;
  • Heterogeneous (across phases and participants).
  Level-2 error:
  • correlated.
  Level-3 error:
  • correlated.

Model 3
• Piecewise with simple errors structure:
  Level-1 error:
  • independence;
  • homogeneity.
  Level-2 error:
  • not correlated.
  Level-3 error:
  • not correlated.

Model 4
• Piecewise with complex errors structure:
  Level-1 error:
  • autocorrelated;
  • Heterogeneous.
  Level-2 error:
  • correlated.
  Level-3 error:
  • correlated.
### Model 1: Quadratic with simple errors structure

<table>
<thead>
<tr>
<th>Observation</th>
<th>Participant</th>
<th>Study</th>
<th>Effect Est</th>
<th>Bias</th>
<th>Relative Bias</th>
<th>Coverage</th>
<th>Std. E</th>
<th>Convergence Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>4</td>
<td>10</td>
<td>0.505</td>
<td>0.005</td>
<td>0.009</td>
<td>.921</td>
<td>0.008</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>0.506</td>
<td>0.004</td>
<td>0.009</td>
<td>.841</td>
<td>0.004</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.505</td>
<td>0.005</td>
<td>0.009</td>
<td>.918</td>
<td>0.005</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>0.506</td>
<td>0.004</td>
<td>0.009</td>
<td>.735</td>
<td>0.003</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>10</td>
<td>0.456</td>
<td>0.044</td>
<td>0.088</td>
<td>0</td>
<td>0.005</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>0.456</td>
<td>0.044</td>
<td>0.088</td>
<td>0</td>
<td>0.003</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.456</td>
<td>0.044</td>
<td>0.088</td>
<td>0</td>
<td>0.004</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>0.456</td>
<td>0.044</td>
<td>0.088</td>
<td>0</td>
<td>0.002</td>
<td>1</td>
</tr>
</tbody>
</table>

### Model 2: Quadratic with complex errors structure

Sadly...IT DIDN’T CONVERGE...
### SIMULATION RESULTS

#### Model 3  
**Piecewise with simple errors structure**

<table>
<thead>
<tr>
<th>Observation</th>
<th>Participant</th>
<th>Study</th>
<th>Effect Est</th>
<th>Bias</th>
<th>Relative Bias</th>
<th>Coverage</th>
<th>Std. E</th>
<th>Convergence Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>4</td>
<td>10</td>
<td>0.496</td>
<td>0.005</td>
<td>0.009</td>
<td>0.920</td>
<td>0.007</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.496</td>
<td>0.004</td>
<td>0.009</td>
<td>0.818</td>
<td>0.004</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.496</td>
<td>0.004</td>
<td>0.008</td>
<td></td>
<td>0.894</td>
<td>0.005</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.496</td>
<td>0.004</td>
<td>0.009</td>
<td>0.663</td>
<td>0.003</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>10</td>
<td>0.497</td>
<td>0.003</td>
<td>0.005</td>
<td>0.934</td>
<td>0.005</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.497</td>
<td>0.003</td>
<td>0.005</td>
<td>0.863</td>
<td>0.003</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.497</td>
<td>0.003</td>
<td>0.006</td>
<td></td>
<td>0.896</td>
<td>0.004</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.497</td>
<td>0.003</td>
<td>0.006</td>
<td>0.749</td>
<td>0.002</td>
<td>1</td>
</tr>
</tbody>
</table>

#### Model 4  
**Piecewise with complex errors structure**

<table>
<thead>
<tr>
<th>Observation</th>
<th>Participant</th>
<th>Study</th>
<th>Effect Est</th>
<th>Bias</th>
<th>Relative Bias</th>
<th>Coverage</th>
<th>Std. E</th>
<th>Convergence Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>4</td>
<td>10</td>
<td>0.496</td>
<td>0.004</td>
<td>0.007</td>
<td>0.96</td>
<td>0.007</td>
<td>.90</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.496</td>
<td>0.004</td>
<td>0.007</td>
<td>0.87</td>
<td>0.004</td>
<td>.99</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.497</td>
<td>0.003</td>
<td>0.007</td>
<td></td>
<td>0.93</td>
<td>0.005</td>
<td>.71</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.498</td>
<td>0.002</td>
<td>0.005</td>
<td>0.96</td>
<td>0.005</td>
<td>.95</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.498</td>
<td>0.003</td>
<td>0.006</td>
<td>0.89</td>
<td>0.003</td>
<td>.97</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.496</td>
<td>0.004</td>
<td>0.007</td>
<td></td>
<td>0.76</td>
<td>0.003</td>
<td>.84</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.498</td>
<td>0.002</td>
<td>0.005</td>
<td>0.96</td>
<td>0.005</td>
<td>.95</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0.498</td>
<td>0.003</td>
<td>0.006</td>
<td>0.89</td>
<td>0.003</td>
<td>.97</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.496</td>
<td>0.004</td>
<td>0.007</td>
<td></td>
<td>0.76</td>
<td>0.003</td>
<td>.84</td>
</tr>
</tbody>
</table>

*Simulation is still running*
Summary

Model 1: 32 Observations -- more bias, and coverage is 0.

- Piecewise trajectory has less bias and better 95% confidence interval coverage across all the conditions than quadratic trajectory in the treatment phase.

Model 2: Did not converge.

Model 3: 30 Studies -- the coverage was lower than the nominal .95.

Model 4: Will conclude it after get the final simulation results.

- For the model with a piecewise trajectory in the treatment phase and a simpler error structure, the coverage was lower than the nominal .95.

FURTHER RESEARCH DESIGN

1. Only limited conditions were examined;
2. Future research should examine:
   • Other non-linear treatment trajectories
   • Non-normally distributed error structure (level-1, level-2, and level-3)
   • Other methods of estimating the model (such as Bayesian estimation)
   • Other dependent error structures in level-1 (moving average; 2nd order autoregressive)
Thank you