Utilizing Hierarchical Linear Modeling in Evaluation: Concepts and Applications

C.J. McKinney, M.A.

Antonio Olmos, Ph.D.

Kate DeRoche, M.A.

Mental Health Center of Denver

Baltimore, MD
Overview

Hierarchical Linear Regression Concepts

Building a Model

Some Applications in Evaluation with Examples
  • Program/Policy Impact (Spline Model)
  • Probability of Incidence (Logistic Regression)

Very Brief Review of Misuses of HLM

Discussion/Questions
Theoretical Rationale
Social and Behavioral researchers typically study situations where higher level factors affect lower level outcomes.

• For Example:
  
  • An individual worker’s error rate may be affected by the departments average workload per employee, which may be affected by the division’s product demand

  • Evaluation of student reading improvement in Colorado middle schools.
Hierarchical linear regression models typically provide better estimates of relationships than more conventional regression models.

- Estimates:
  - Higher accuracy
  - Lower standard error

Due to taking into account hierarchical nature of the data.
Hierarchical Linear Regression Concepts
Hierarchical linear regression models use the coefficients of regression at the lower levels as outcomes in regression at the higher levels.

- Described as “Regression on Regression”
- Older hierarchical linear modeling papers referred to the technique as random coefficient models for this reason.
Student Level Effect

Reading Scores = Intercept + Grade

Typical SLR Model School Level Effect

Intercept = Intercept + SES

Public

Higher Level Effects

Intercept = Intercept + SES

Public

Public

Public
Errors are estimated at each level and for each cross-level effect.
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Intercept</th>
<th>Slope</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LEVEL 1:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>= Grade_{intercept} + Grade_{slope} + ( \epsilon_{11} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LEVEL 2:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade_{intercept}</td>
<td>= SES_{intercept1} + SES_{slope1} + ( r_{21} )</td>
<td>= SES_{intercept2} + SES_{slope2} + ( r_{22} )</td>
<td></td>
</tr>
<tr>
<td>Grade_{slope}</td>
<td>= SES_{intercept2}</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LEVEL 3:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES_{intercept1}</td>
<td>= Public_{intercept1} + Public_{slope1} + ( u_{31} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES_{slope1}</td>
<td>= Public_{intercept2} + Public_{slope2} + ( u_{32} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES_{intercept2}</td>
<td>= Public_{intercept3} + Public_{slope3} + ( u_{33} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES_{slope2}</td>
<td>= Public_{intercept4} + Public_{slope4} + ( u_{34} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hierarchical linear regression models have the same assumptions as conventional regression models

- Linear relationship
- Errors have Normal distribution
- Errors have Equal Variances
- Errors are Independent
Questions???
Building an HLM model
DECA Scores

Score 1: Measures teacher perception of child’s protective factors through 3 subgroups:

• Initiative
• Self-Control
• Attachment

Score 2: Measure of teacher’s concern for the child’s behavior.

Scores taken at program admission (beginning of school year) and after program discharge (end of school year) N = 536
DECA Scores

Factor of Interest:
Pre/Post Scoring

Developmental Factors:
Age and Gender
Unconditional Model

**LEVEL 1:** Score $= \pi_0 + \varepsilon$

**LEVEL 2:** $\pi_0 = \beta_{00} + r_0$
Unconditional Model

Deviance: 5502 with 2 parameters

Total Variance

\[ \varepsilon + r_0 = 30.2 + 42.7 = 72.9 \]

Variance accounted for by children: 59%

\[ \frac{r_0}{\varepsilon + r_0} = 0.59 \]
**LEVEL 1:** Score = \( \pi_0 + \pi_1 \text{(Post)} + \epsilon \)

**LEVEL 2:**
\[
\pi_0 = \beta_{00} + \beta_{01}(\text{Female}) + \beta_{02}(\text{Age}) + r_0 \\
\pi_1 = \beta_{10} + r_1
\]

Deviance: 5426 with 3 parameters

\[
\chi^2_0 = 5502 - 5426 = 76 \Rightarrow \text{with df = 1, } p < 0.005
\]

**Full better than Unconditional**

Reduction of Error: 
\[
= 1 - (\epsilon + r_0)_{\text{full}}/(\epsilon + r_0) = 1 - 56.7/72.9 \\
= 1 - 0.78 = 0.22 = \boxed{22\%}
\]
Overall Model Fit

- $R^2 = 0.85$
Parameter Estimates

**POST:**  At the end of the program the level of protective factors increases on average by **2.6** points over the level at the beginning of the program.

**AGE:**  Consistent with developmental theory the level of protective factors increases by **1.4** points per year of age.

**Gender:**  Consistent with developmental theory the level of protective factors is greater for females than males by **3.8** points on average.

** - Significantly different from 0 with $p < 0.001$
Questions???
Hierarchical Linear Modeling Examples

MHCD Recovery Marker Inventory
Recovery Marker Inventory

• Consist of a scaled score from six indicators of mental health recovery.
  • Employment
  • Learning/Training
  • Housing
  • Active Growth / Orientation
  • Symptom Interference
  • Service Participation
Scores were converted from raw form to an ability score utilizing an Item Response Theory (IRT) Partial Credit model for ability estimation.

- An increase in the ability score, indicates an increase the overall factors that support mental health recovery.
Recovery Marker Inventory

- Markers are collected every 2 months on each consumer by the case managers and clinicians.
Estimated Changes in Recovery Marker Scores Over Time

- Ability-to-Recover Score
- Evaluation Period (2 Months)

- HITT Mood and Other
- CTT Mood and Other
- OP Mood and Other
- HITT Thought
- CTT Thought
- OP Thought
HLM Evaluation Applications
Program/Policy Impact

Typical Questions:

- Does the program affect performance?
- Which programs are more effective?
- What factors affect performance?
The philosophy of MHCD is that all individuals with mental illness can and many do recover.

This resulted in a policy within our adult treatment teams that as individuals recover they can be moved to a lower level team, which opens spaces for more individuals at the higher levels of treatment.

A question arose as to whether these changes in the level of treatment affected their recovery level?
<table>
<thead>
<tr>
<th>TO TEAM</th>
<th>EHITT</th>
<th>HITT 1</th>
<th>HITT 2</th>
<th>HIATT</th>
<th>HITT 3</th>
<th>ILT</th>
<th>SLT</th>
<th>CTT 1</th>
<th>CTT 2</th>
<th>CTT 3</th>
<th>OPT</th>
<th>OBRA</th>
<th>EXEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHITT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HITT 1</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>17</td>
<td>35</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>HITT 2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>13</td>
<td>2</td>
<td>86</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>HIATT</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>68</td>
<td>12</td>
<td>6</td>
<td>44</td>
<td>3</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>HITT 3</td>
<td>0</td>
<td>15</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>ILT</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>13</td>
<td>119</td>
<td>13</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>SLT</td>
<td>0</td>
<td>7</td>
<td>3</td>
<td>22</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>17</td>
<td>11</td>
<td>69</td>
<td>4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>CTT 1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>30</td>
<td>3</td>
<td>0</td>
<td>11</td>
<td>5</td>
<td>43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CTT 2</td>
<td>0</td>
<td>15</td>
<td>54</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>10</td>
<td>72</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>CTT 3</td>
<td>0</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>3</td>
<td>9</td>
<td>26</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>41</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>OPT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>OBRA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EXEC</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Increased Treatment Intensity**
- N=264
- 21.43%

**Same Treatment Intensity**
- N=107
- 8.69%

**Decreased Treatment Intensity**
- N=787
- 63.88%

**Transferred to OBRA**
- N=20
- 1.62%

**Inactive Consumer**
- N=52
- 4.22%

**Total N** = 1232

**Average time in previous team** = 490 Days
The Model

Team Change Point

TIME

No Effect

Effect
The Model

**LEVEL 1:**  
 RMIScore  
  
  \[ \pi_0 + \pi_1 (\text{Time}) + \pi_2 (\text{Time}^2) + \pi_3 (\text{CHGInt}) + \pi_4 (\text{PostTime}) + \pi_5 (\text{PostTime}^2) + \varepsilon \]  

**LEVEL 2:**

\[ \pi_0 = \beta_{00} + \beta_{01}(\text{MOOD}) + \beta_{02}(\text{THOUGHT}) + r_0 \]

\[ \pi_1 = \beta_{10} + \beta_{11}(\text{MOOD}) + \beta_{12}(\text{THOUGHT}) + r_1 \]

\[ \pi_2 = \beta_{20} + \beta_{21}(\text{MOOD}) + \beta_{22}(\text{THOUGHT}) + r_2 \]

\[ \pi_3 = \beta_{30} + \beta_{31}(\text{MOOD}) + \beta_{32}(\text{THOUGHT}) + r_3 \]

\[ \pi_4 = \beta_{40} + \beta_{41}(\text{MOOD}) + \beta_{42}(\text{THOUGHT}) + r_4 \]

\[ \pi_5 = \beta_{50} + \beta_{51}(\text{MOOD}) + \beta_{52}(\text{THOUGHT}) + r_5 \]

**Where:**

- **Time**  
  The time period the outcomes were obtained in number of months since admission.

- **CHGInt**  
  The direction and magnitude of the team change (indicated adjustment to slope if service change occurred).

- **PostTime**  
  Same as Time with values only for those after a team change occurred (Indicates a difference in slope after team change).

- **MOOD/THOUGHT**  
  An indicator variable related to a consumer having a mood or thought disorder.
The Results

**LEVEL 1:**

\[ \text{RMIScore} = \pi_0 + \pi_1 (\text{Time}) + \pi_2 (\text{Time}^2) + \pi_3 (\text{CHGInt}) + \pi_4 (\text{PostTime}) + \pi_5 (\text{PostTime}^2) + \varepsilon \]

**LEVEL 2:**

\[ \pi_0 = \beta_{00} + \beta_{01} (\text{MOOD}) + \beta_{02} (\text{THOUGHT}) + r_0 \]
\[ \pi_1 = \beta_{10} + \beta_{11} (\text{MOOD}) + \beta_{12} (\text{THOUGHT}) + r_1 \]
\[ \pi_2 = \beta_{20} + \beta_{21} (\text{MOOD}) + \beta_{22} (\text{THOUGHT}) + r_2 \]
\[ \pi_3 = \beta_{30} + \beta_{31} (\text{MOOD}) + \beta_{32} (\text{THOUGHT}) + r_3 \]
\[ \pi_4 = \beta_{40} + \beta_{41} (\text{MOOD}) + \beta_{42} (\text{THOUGHT}) + r_4 \]
\[ \pi_5 = \beta_{50} + \beta_{51} (\text{MOOD}) + \beta_{52} (\text{THOUGHT}) + r_5 \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B00</th>
<th>B02</th>
<th>B10</th>
<th>B20</th>
<th>Parameter</th>
<th>R0</th>
<th>R1</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
<td>4.72</td>
<td>-0.34</td>
<td>0.012</td>
<td>-0.0002</td>
<td>Estimate</td>
<td>1.04</td>
<td>0.084</td>
<td>0.001</td>
</tr>
<tr>
<td>SE</td>
<td>0.033</td>
<td>0.04</td>
<td>0.0029</td>
<td>0.00005</td>
<td>( \chi^2 (df) )</td>
<td>1789(524)</td>
<td>899(527)</td>
<td>912(527)</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>p-value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Conclusions

1. All of the parameters related to the team change period demonstrated no significant differences from the pre-change period.

2. The typical parameters related to time overall, and the mood and thought disorder parameters both provided results consistent with previous models.

3. This indicates that when a consumer is moved from one team to another, in either direction, they appear to keep the same level of recovery and rate of change in recovery.

4. This provides evidence to support the current practice of clinicians moving consumers to higher levels or lower levels of service as clinically indicated, as it does not affect the recovery supports for the consumer.
Questions???
Modeling Rate of Incidence

Typical Questions:

- Does the program affect the rate of incidence of an event?
- What factors affect the rate of incidence of an event?

Typically use a Logistic Regression Model

\[ \ln\left( \frac{\pi}{1-\pi} \right) = \beta_0 + \beta_1(x) + \varepsilon \]

Where \( \pi \) is the proportion of individuals with a specified characteristic, and \( \ln(\pi/1-\pi) \) is the log-odds or logit.
The logistic regression model can be recast into the HLM framework by simply allowing rate of incidence to vary across higher level units.
The Question: Does the prevalence of substance abuse change with the recovery ability?

- We assume that as the factors that support recovery increase, the prevalence of the abuse of substances will go down.
We also assume that the rate of change varies across each individual, making the HLM model appropriate.

On the consumer level, we looked at whether the consumer was in a drug treatment team or not.

Those in a drug treatment team were expected to have a higher intercept and steeper decrease in prevalence of substance abuse.
The Model

**LEVEL 1:** \[
\ln(\pi/1-\pi) = \beta_0 + \beta_1 \text{(RMIScore)} + \epsilon
\]

**LEVEL 2:**
- \[
\beta_0 = \phi_{00} + \phi_{01} \text{(DrugTeam)} + \phi_{02} \text{(Mood)} + \phi_{03} \text{(Thought)} + r_0
\]
- \[
\beta_1 = \phi_{10} + \phi_{11} \text{(DrugTeam)} + \phi_{12} \text{(Mood)} + \phi_{13} \text{(Thought)} + r_1
\]

Where:
- \[
\ln(\pi/1-\pi) \quad \text{– The log-odds of the prevalence of substance abuse.}
\]
- **RMIScore** – The recovery marker score for each outcome for indicator of substance abuse.
- **DrugTeam** – An indicator variable related to whether consumer was in drug treatment team or not.
- **MOOD/THOUGHT** - An indicator variable related to a consumer having a mood or thought disorder.
Results

**LEVEL 1:** \[ \ln(\pi/1-\pi) = \beta_0 + \beta_1 \text{(RMIScore)} + \epsilon \]

**LEVEL 2:**
- \[ \beta_0 = \varphi_{00} + \varphi_{01} \text{(DrugTeam)} + \varphi_{02} \text{(Mood)} + \varphi_{03} \text{(Thought)} + \rho_0 \]
- \[ \beta_1 = \varphi_{10} + \varphi_{11} \text{(DrugTeam)} + \varphi_{12} \text{(Mood)} + \varphi_{13} \text{(Thought)} + \rho_1 \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \psi_{00} )</th>
<th>( \psi_{01} )</th>
<th>( \psi_{10} )</th>
<th>( \psi_{12} )</th>
<th>( \psi_{13} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimate</strong></td>
<td>-1.569</td>
<td>0.96</td>
<td>-0.219</td>
<td>-0.059</td>
<td>-0.176</td>
</tr>
<tr>
<td><strong>SE</strong></td>
<td>0.136</td>
<td>0.305</td>
<td>0.036</td>
<td>0.022</td>
<td>0.023</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>&lt;0.001</td>
<td>0.03</td>
<td>&lt;0.001</td>
<td>0.009</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Results

[Graph showing the probability of incidence against RMI Score for different groups: Null, DrugTeam, Mood, Thought, DrugTeam/Mood, DrugTeam/Thought]
Conclusions

1. As expected there was an overall increase in the prevalence of substance abuse in the drug treatment teams, but a significant difference in slope was not found.

2. One item of interest is that the rate of decrease was dependant on whether the consumer had a mood or thought disorder. Both were higher than the case of a general disorder, but thought was much greater. It is thought that this results from the substance abuse being more directly related to the symptoms of thought disorders, thus as recovery increases it would be more likely for those to stop the substance abuse.

3. Overall, the negative relationship with the recovery marker score indicates that an increase in recovery supports also helps reduce the prevalence of substance abuse.
Questions???
Misuses of HLM

1. Data is not Hierarchical
   • Since HLM has become so popular, the incidence of researchers using HLM for non-hierarchical data has also increased.

2. The variance estimates are not significantly different from zero at the higher levels.
   • If the all variance components of the higher level effects is 0, this implies these are fixed/constant in the lower levels, thus HLM is not needed.
**LEVEL 1:**
Score = $\text{Grade}_{\text{intercept}} + \text{Grade}_{\text{slope}} + \varepsilon_{11}$

**LEVEL 2:**
$\text{Grade}_{\text{intercept}} = \text{SES}_{\text{intercept1}} + \text{SES}_{\text{slope1}} + r_{21}$

$\text{Grade}_{\text{slope}} = \text{SES}_{\text{intercept2}} + \text{SES}_{\text{slope2}} + r_{22}$

**LEVEL 3:**
$\text{SES}_{\text{intercept1}} = \text{Public}_{\text{intercept1}} + \text{Public}_{\text{slope1}} + u_{31}$

$\text{SES}_{\text{slope1}} = \text{Public}_{\text{intercept2}} + \text{Public}_{\text{slope2}} + u_{32}$

$\text{SES}_{\text{intercept2}} = \text{Public}_{\text{intercept3}} + \text{Public}_{\text{slope3}} + u_{33}$

$\text{SES}_{\text{slope2}} = \text{Public}_{\text{intercept4}} + \text{Public}_{\text{slope4}} + u_{34}$
3. Violations of Assumptions

- It is assumed that the error terms at all levels are equal across units and normally distributed.
- This can be difficult to assess and if not meet will result in inaccurate inferences.

Overall:

As with all statistical models, HLM models have various assumptions, and violations of these assumptions can result in inference errors and/or utilization of a more complicated than necessary model.
Discussion
Books:


Papers:


HLM in SPSS:


Consideration in use of HLM:


For more information about research on mental health and Recovery/Resiliency at MHCD, or to access our conference presentations and/or publications, please visit our website at

http://www.outcomesmhcd.com/