Identifying Thresholds of Quality in Early Child Care and Education: A Non-Parametric Approach

Greg Welch & Ji Hoon Ryoo
Nebraska Center for Research on Children, Youth, Families and Schools
Overview

• Background
• Data Sources
• Analytic Approaches
  – General Additive Modeling
  – Spline Analyses
  – Framework for this project
• Examples
• Summary
Background

○ Led by CYFS Faculty Affiliates
  – Julia Torquati
  – Helen Raikes

○ Substantively speaking…
  ○ Evidence that better quality generally predicts more optimal child outcomes
Aims

• **Aim 1**: To determine whether there are *specific thresholds* of quality that predict children’s developmental outcomes

• **Aim 2**: To determine whether thresholds of quality vary as a function of:
  – (1) age of child in child care;
  – (2) ethnicity of teacher/provider;
  – (3) ethnic match/mismatch of teacher/provider and child;
  – (4) income level of child
Background

- Linear vs. Non-Linear models
  - Linear the norm
    - Implies relationship same across all points
  - Non-linear are necessary
    - The strength of association between predictors and outcomes varies across the measurement scale such that particular levels of quality are significantly more strongly associated with children’s development.
## Early Head Start Sample

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Family Child Care</th>
<th>Center Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>151</td>
<td>372</td>
</tr>
<tr>
<td>24</td>
<td>159</td>
<td>416</td>
</tr>
<tr>
<td>36</td>
<td>134</td>
<td>500</td>
</tr>
<tr>
<td>60</td>
<td>80</td>
<td>1043</td>
</tr>
</tbody>
</table>
QUINCE Sample

- Children ages 20 months – 5 years
- Assessed fall and spring
  - T2 focus of this analysis

<table>
<thead>
<tr>
<th>Family Child Care</th>
<th>Center Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>650</td>
<td>652</td>
</tr>
</tbody>
</table>
Analytic Approaches

• General Additive Modeling
  – Empirically driven
  – Non-parametric approach
  – Identify relationship between variables
    • Linear vs. Non-linear
  – Identify possible thresholds
    • Based on shape of observed relationship
Analytic Approaches

• Spline Models
  – Identify empirical thresholds obtained from GAM via the model comparison between linear and non-linear associations
  – Investigate the change of associations within each interval of quality of care classified by empirical thresholds
  – Predict associations based on the spline model
Introduction to GAM

• Generalized Additive Model (GAM)
  – Each term is estimated using a univariate smoother
  – The estimate explains how the dependent variable changes with the corresponding independent variables

\[ \eta = s_0 + \sum_{i=1}^{p} s_i(X_i) \]

where \( s_i \) are smooth functions defines the additive component.

• The smoothers are estimated by using the backfitting algorithm and the local scoring algorithm
  – The algorithms are implemented in PROC GAM in SAS and gamm package in R
Spline Analyses

• For identification of thresholds, the following two models are compared
  
  Linear model \[ Y_i = \beta_0 + \beta_1 \cdot X_i + e_i \]
  
  Spline model \[ Y_i = \beta_0 + \beta_1 \cdot X_i + \beta_2 \cdot T_i^1 + \beta_3 \cdot T_i^2 + e_i \]

• To investigate the effect of moderator, the following two models are compared
  
  Spline model \[ Y_i = \beta_0 + \beta_1 \cdot X_i + \beta_2 \cdot T_i^1 + \beta_3 \cdot T_i^2 + e_i \]
  
  Model with Moderator \[ Y_i = \beta_0 + \beta_1 \cdot X_i + \beta_2 \cdot T_i^1 + \beta_3 \cdot T_i^2 + \sum_{k=1}^{K} C_i^k + e_i \]
Framework for Identifying Thresholds

• Thresholds were identified
  – Empirically via GAM
  – Based on a priori research/knowledge
Example 1

• Begin by implement GAM approach
  – Non-parametric approach to investigating the relationship between quality measures of child care (ITERS or ECERS, FDCRS, and CIS) and child outcomes

• Want to determine:
  – Relationship linear or non-linear?
  – If non-linear, are thresholds implied?
Example 1

- To examine the relationship between ITERS quality measure and sustained attention outcome measure at 14 month in the EHS sample
- Using PROC GAM in SAS
  
  ```
  proc gam data=Jihoon.ehs plots=components(clm);
  model b1v3catn = spline(T1ITSTTL)/
    anodev = refit method = gcv epsscore = 1e-6;
  output out = ehs14 p;
  run;
  ```
- It produces parameter estimates, estimated degree of freedom, and smoothing plot(s)
Example 1 – GAM Plot

- ITERS and Sustained Attention
Example 1 – SAS Output

The GAM Procedure
Dependent Variable: b1v3catn
Smoothing Model Component(s): spline(T1ITSTTL)

Summary of Input Data Set
Number of Observations                 335
Number of Missing Observations        2666
Distribution                      Gaussian
Link Function                     Identity

Regression Model Analysis
Parameter Estimates

| Parameter | Estimate | Standard Error | t Value | Pr > |t| |
|-----------|----------|----------------|---------|-------|-----|
| Intercept | 5.15073  | 0.21544        | 23.91   | <.0001|
| Linear(T1ITSTTL) | -0.04984 | 0.04601       | -1.08   | 0.2795|

Smoothing Model Analysis
Analysis of Deviance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline(T1ITSTTL)</td>
<td>2.13902</td>
<td>8.274623</td>
<td>7.8755</td>
<td>0.0226</td>
</tr>
</tbody>
</table>
**Example 1 – Spline Result**

- Testing empirical thresholds - 3.3 and 5.0
  - $F_2=4.628$; $p$-value=0.010

<table>
<thead>
<tr>
<th>Line Segment</th>
<th>Estimate (slope)</th>
<th>Std Err</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0.0, 3.3]</td>
<td>0.35 (0.35)</td>
<td>0.219</td>
<td>0.111</td>
</tr>
<tr>
<td>[3.3, 5.0]</td>
<td>-0.767 (-0.417)</td>
<td>0.316</td>
<td>0.016*</td>
</tr>
<tr>
<td>[5.0, 7.0]</td>
<td>0.712 (0.295)</td>
<td>0.238</td>
<td>0.003**</td>
</tr>
</tbody>
</table>
Example 2 – GAM Plot

- ECERS-R and Howes Ladd conflict with teacher rating
Example 2 – Spline Results

- Comparing possible models
- Practical?

Fitted linear model vs. spline1 model

Fitted linear model vs. spline5 model
Example 3 – GAM Plot

- Bracken school readiness and ECERS-R
Example 3 – GAM Plot

- Bracken school readiness and ECERS-R
  - by Gender
- Gender significant (p-value=0.0002)
  - Does not significantly impact GAM result
Example 4 – GAM Plots

- ECERS-R & CIS and Howes Ladd aggressive composite
Example 4 – GAM Plots

- What if we look at the relationships separately?
  - The results are almost identical, which is a property of GAM
Example 4 – Spline Results

- Compare the spline model with the suggested thresholds with a linear model on both cases

![Linear1 model vs. spline1 model](image1)

![Linear2 model vs. spline2 model](image2)
Utility of the Approaches

• GAM excellent approach to identifying thresholds
  – No a priori knowledge necessary
    • Can be useful though
  – Does not provide for inferential statements

• Spline
  – Can use a priori and/or empirical information
  – Can make inferential statements
Thank You!

Questions?

For Additional Questions/Information

gwelch2@unl.edu
jryoo2@unl.edu