Power Analysis for Multilevel Randomized Cost Effectiveness Trials

Authors and Affiliations:

Wei Li
Department of Educational, School & Counseling Psychology (ESCP)
College of Education, University of Missouri
Phone: 517-488-9089
Email: wl73d@missouri.edu

Nianbo Dong
University of Missouri
14 Hill Hall, Columbia, MO 65211
Phone: (573) 882-1495
dong.nianbo@gmail.com

Rebecca Maynard
Graduate School of Education
University of Pennsylvania
3700 Walnut Street
Philadelphia, PA 19104
rmaynard@upenn.edu
Power Analysis for Multilevel Randomized Cost Effectiveness Trials

Background:
Cost-effectiveness analysis (CEA) is a method of comparing the cost and effectiveness for alternative programs, policies, and practices with similar goals (Levin & Belfield, 2015; Muennig, & Bounthavong, 2016). It has been used most widely in assessments of health care interventions and medical practices (Petti, 1999; Robinson, 1993; Sullivan, et al., 2005). Recently, with the increasing investment in assessing the effectiveness of educational programs using rigorous methods that often include randomizing students, classrooms, or schools to treatment or control condition, there is growing interest in synthesizing and comparing the results across educational interventions. When comparing findings across interventions, it is important to consider not only the relative effectiveness in changing the focal outcome, but to also take account of the relative cost of achieving a unit change in the focal outcomes. Other things equal, decisions about changes in programs, policies or practices would take account of the relative cost-effectiveness of alternatives (Levin & McEvan, 2001). As a result, it is increasingly common for impact evaluations of educational interventions to also include rigorous cost analysis components in order to support subsequent cross-study comparisons of the cost-effectiveness of alternative strategies for achieving target educational outcomes (Hollands et al., 2016).

An important part of the design phase of any prospective impact analysis, including cost-effectiveness impact studies, is a power analysis to determine the optimal size and allocation of the study sample, give the study objectives. Just as with a traditional impact evaluation, it is important to conduct a priori power computation when designing randomized trials aimed at estimating the difference in cost-effective trials (commonly referred to as randomized cost effectiveness trials or RCETs). The objective of this analysis is to ensure that the sample design offers a “good enough” chance (e.g., $p \geq 0.80$) to detect when one of the educational strategies is more cost-effective that another.

Methods of conducting power analysis for RCETs in the health and medical sciences have been developed (e.g., Willan & Briggs, 2006). However, there is not a similar literature to guide the design of such studies in education, which has some unique features, such as the nested nature of units of analysis in education trials (Levin & Belfield, 2015). For example, often it is necessary in education trials to account for the fact that students are nested within classrooms, classrooms are nested within schools, and schools are nested within districts. Recent work in health economics (Manju, Candel, & Berger, 2014, 2015) utilized two-level (e.g., students nested in schools) hierarchical linear models (HLMs) to resolve the potential nesting effects and provided formulas to calculate power for two-level RCETs. However, their approaches cannot be adopted by educational researchers easily for three reasons: (1) their methods assumed individual level cost information is available, while in education such information is usually missing (Levin & Belfield, 2015); (2) their models do not account for the effects of covariates, which are commonly available and can impact quite considerably the required sample sizes in education studies (Bloom, Richburg-Hayes, & Black, 2007; Hedges & Hedberg, 2007); and (3) there is no software to support easy estimation of the sample size requirements under the various study designs that are common in education.

Purpose:
This project is designed to address these limitations of prior studies and contributes to the cost-effectiveness analysis and power analysis literatures by extending them to accommodate various
assumptions about covariates and alternative sample designs and analytic modelling assumptions. In particular, our research plan incorporates three complementary objectives: (1) advance statistical theory underlying RCETs by developing new statistical power formulas, (2) assess and substantiate these derivations through simulation, and (3) implement the formulas in our existing software program, PowerUp! (Dong & Maynard, 2013), to provide an accessible and user-friendly tool for researchers planning RCETs.

Research Design:

We will develop and implement new formulas for estimating the statistical power to RCETs with four alternative two-level designs (see Table 1). In this proposal we show the results for design 1 as an example. Consider a two-level blocked individual random assignment design where level-1 units (e.g., students) are randomly assigned to treatment or control conditions within level-2 units (e.g., schools). When the level-1 information of effectiveness and cost is available, two-level HLM could be used to estimate the mean differences of effectiveness and cost between the treatment and control groups, namely

\[ e_{ij} = \gamma_{00}^e + \gamma_{10}^e T_{ij} + u_{0j}^e + u_{1j}^e T_{ij} + \varepsilon_{ij}^e, \]  

\[ c_{ij} = \gamma_{00}^c + \gamma_{10}^c T_{ij} + u_{0j}^c + u_{1j}^c T_{ij} + \varepsilon_{ij}^c, \]  

and

\[ \begin{pmatrix} u_{0j}^e \\ u_{0j}^c \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{ec}^2 & \tau_{tec}^2 \\ \tau_{tec}^2 & \tau_{ec}^2 \end{pmatrix} \right), \]  

\[ \begin{pmatrix} \varepsilon_{ij}^e \\ \varepsilon_{ij}^c \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{ec}^2 & \sigma_{ec}^2 \\ \sigma_{ec}^2 & \sigma_{ec}^2 \end{pmatrix} \right), \]  

where \( e_{ij} \) is the effectiveness measure (e.g., student achievement) for individual \( i \) in cluster \( j \); \( c_{ij} \) is the cost for individual \( i \) in cluster; and \( T_{ij} \) is a treatment indicator variable (\( T_{ij} = 0.5 \) for treatment, \( T_{ij} = -0.5 \) for control).

Let \( INB_{ij} \) represent the net monetary benefit for individual \( i \) in cluster \( j \), we have

\[ INB_{ij} = k(\gamma_{00}^e + \gamma_{10}^e T_{ij} + u_{0j}^e + u_{1j}^e T_{ij} + \varepsilon_{ij}^e) - (\gamma_{00}^c + \gamma_{10}^c T_{ij} + u_{0j}^c + u_{1j}^c T_{ij} + \varepsilon_{ij}^c) 
= (k\gamma_{00}^e - \gamma_{00}^c) + (k\gamma_{10}^e - \gamma_{10}^c) T_{ij} + (k u_{0j}^e - u_{0j}^c) + (k u_{1j}^e - u_{1j}^c) T_{ij} + (k \varepsilon_{ij}^e - \varepsilon_{ij}^c) 
= \pi_{00} + \pi_{10} T_{ij} + r_{0j} + r_{ij} T_{ij} + \delta_{ij}, \]  

where \( k \) is a positive constant that can be considered as the amount of society is willing to pay for a unit of effectiveness (Manju, Candel, & Berger, 2014; Willan, 2001), \( r_{0j} \sim N(0, \tau_{r}^2) \), \( r_{1j} \sim N(0, \tau_{r}^2) \), and \( \delta_{ij} \sim N(0, \sigma^2) \). The parameter of interest now is \( \pi_{10} \). When \( \pi_{10} > 0 \), it indicates the treatment is cost-effective, when \( \pi_{10} < 0 \), it indicates the treatment is not cost-effective. Suppose there are \( M \) level-2 units and \( n \) level-1 units within each treatment or control condition. The variance of \( \hat{\pi}_{10} \) is

\[ \text{Var}(\hat{\pi}_{10}) = \frac{1}{Mn} \left( nt_{\pi}^2 + 2\sigma^2 \right) \]  

And the non-centrality parameter (unstandardized) is

\[ \lambda = \pi_{10} \sqrt{Mn} \left( \frac{1}{nt_{\pi}^2 + 2\sigma^2} \right) \]
where \( \tau^2_T = k^2 \tau^2_{Te} + \tau^2_{Tc} - 2k\tau_{Tec} \) and \( \sigma^2 = k^2\sigma^2_e + \sigma^2_c - 2k\sigma_{ec} \).

### Following Hedges and Rhoads (2010)

Define

\[
\begin{align*}
\sigma_e &= \sigma^e_s + \sigma^e_w \\
\text{ICC}_e &= \frac{\sigma^e_s}{\sigma_e} \\
\sigma_c &= \sigma^c_s + \sigma^c_w \\
\text{ICC}_c &= \frac{\sigma^c_s}{\sigma_c} \\
\sigma_e &= r_1 \sigma_c \\
\sigma^e_c &= r_2 \sigma_c \\
\sigma^e_w &= r_3 \sigma_c 
\end{align*}
\]

where \( \sigma^e_s \) and \( \sigma^e_w \) are the covariance between the cost and effectiveness at the second and the first level. Also, following Hedges and Rhoads (2010), define

\[
\begin{align*}
\tau_{Te} &= r_4 \sigma^e_s \\
\tau_{Tc} &= r_5 \sigma^c_s \\
\tau_{Tec} &= r_6 \sigma^e_c 
\end{align*}
\]

Define the standardized effect size as

\[
\text{ES} = \frac{\pi_{10}}{\sqrt{\sigma_c}} \tag{7}
\]

And then the non-centrality parameter becomes

\[
\lambda = \sqrt{MnES} \left[ \frac{1}{(nr_1r_4-2)k^2\text{ICC}_e+(nr_5-2)\text{ICC}_c+2k(nr_1-r_2r_3)\lambda} \right] \tag{8}
\]

The power is defined as:

\[
p = 1 - H \left[ c(\alpha/2, M-1), (M-1), \lambda \right] + H \left[ -c(\alpha/2, M-1), (M-1), \lambda \right] \tag{9}
\]

where \( c(\alpha, v) \) is the level \( \alpha \) one-tailed critical value of the \( t \)-distribution with \( v \) degrees of freedom (e.g., \( c(0.05, 20) = 1.72 \)), \( H(x, \nu, \lambda) \) is the cumulative distribution function of the non-central \( t \)-distribution with \( \nu \) degrees of freedom and non-centrality parameter \( \lambda \).

### Results and Discussion

The formulas that we derive can be used for power analysis of multilevel randomized cost effectiveness trials. We will provide examples of application and suggestions of conducting randomized cost effectiveness trials in the presentation should it is accepted.

---

1 Note that \( \sigma^2_e = \sigma^e_w, \sigma^2_c = \sigma^c_w, \) and \( \sigma_{ec} = \sigma^e_{ec} \).
References:


<table>
<thead>
<tr>
<th>Design</th>
<th>Level of Clustering</th>
<th>Level of Treatment Assignment</th>
<th>Level of Cost Information Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design 1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Design 2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Design 3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Design 4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>