The Multi-Site RCT

- A fleet of RCTs!
- Each conducted in a different social setting
- Since 2002, IES has funded 175 randomized trials;
- Vast majority are multi-site trials (Spybrook, 2013)

Opportunities

- Can assess generalizability of impact
  - Gauge variation in effect
  - Model heterogeneity
  - Does the intervention equalize outcomes?

But we must formulate appropriate analyses.

This is trickier than it might seem!

Some Recent MS Trials

Study	Levels	Assigned Units	Sites	Fixed or Random sites
National Head Start Eval.	2	Children	300+ Program Sites	Random
Moving to Opportunity	2	Families	5 cities	Fixed
Boston Charter School Lotteries	2	Children	Lottery pools	Random
Tennessee STAR	3	Teachers	Schools	Random
Double-Dose Algebra	2	Children	Schools	Random

Theoretical Model

Within Sites

$$Y_{ij} = U_{0j} + B_j T_{ij} + r_{ij}, \qquad r_{ij} \sim (0, \sigma_T^2)$$

Between Sites

$$U_{0j} = \mu_0 + u_{0j}, \qquad u_{0j} \sim (0, \tau_0^2)$$

$$B_j = \beta + b_j, \qquad b_j \sim (0, \tau_b^2)$$

$$Cov(u_{0j}, b_j) = \tau_{0b}$$

4. Unbalanced Designs and Targets of Inference

Two kinds of imbalance

- Unequal *n* per site
- Unequal propensity score (probability of assignment to the treatment group)

Targets of Inference

Generalize to a Population of Sites

$$\beta_{sites} = \frac{1}{J^*}\sum_{j=1}^{J^*} B_j$$

$$\tau^2_{b\,sites} = \frac{1}{J^*}\sum_{j=1}^{J^*} (B_j - \beta)^2$$

$$\tau_{0b\,sites} = \frac{1}{J^*}\sum_{j=1}^{J^*} (U_{0j} - \mu_0)(B_j - \beta)$$

Generalize to a population of Persons

$$\beta_{persons} = \sum_{j=1}^{J^*} N_j B_j / \sum_{j=1}^{J^*} N_j$$

$$\tau^2_{b\,persons} = \sum_{j=1}^{J^*} N_j (B_j - \beta)^2 / \sum_{j=1}^{J^*} N_j$$

$$\tau_{0b\,persons} = \sum_{j=1}^{J^*} N_j (U_{0j} - \mu_0)(B_j - \beta) / \sum_{j=1}^{J^*} N_j$$

Identification

$Combine\ level\ 1\ and\ level\ 2$

$$Y_{ij} = \mu_0 + \beta T_{ij} + u_{0j} + b_j T_{ij} + e_{ij}$$

$$E(Y_{ij} \mid T_{ij}) = \mu_0 + \beta T_{ij} + E(u_{0j} \mid T_{ij}) + T_{ij} E(b_j \mid T_{ij})$$

$Worry\ about\ E(u_{0j} \mid \overline{T}_{\cdot j}), E(b_j \mid \overline{T}_{\cdot j})$

e.g., charter school lotteries

	Parameters to be estimated	Properties
a)Fixed effects	β	Biased if precision related to *B*
b)Centering HLM	*B,Var(B)*	Bias in β *if precision is related to B (less so than fixed effects)*
c) Control propensity Score	*B,Var(B)??*	Similar to Centering for β,
d)Weighting ("IPTW") with HLM	$B, Var(B), Cov(B, U_0)$	Removes bias (but may be imprecise!)

a) Fixed Effects Model

$$Y_{ij} = \beta T_{ij} + \alpha_j + e_{ij}$$

$$\hat{\beta} = \frac{\sum_{j=1}^{J} n_j \bar{T}_{\cdot j}(1 - \bar{T}_{\cdot j})\hat{B}_j}{\sum_{j=1}^{J} n_j \bar{T}_{\cdot j}(1 - \bar{T}_{\cdot j})}$$

where

$$\hat{B}_j = \bar{Y}_{1j} - \bar{Y}_{0j}$$

Potential Bias if heterogeneous impact, and unbalanced design

$$E(\hat{\beta} \mid B, T) = E\left\{ \frac{\sum_{j=1}^{J} n_j \overline{T}_{\cdot j}(1 - \overline{T}_{\cdot j}) B_j}{\sum_{j=1}^{J} n_j \overline{T}_{\cdot j}(1 - \overline{T}_{\cdot j})} \right\}$$

$$A \ precision - weighted \ \ average$$

b) Centering the Predictor AND the Outcome ("*FIRC*" Model)

$$Y_{ij} - \bar{Y}_{j} = B_{j}(T_{ij} - \bar{T}_{\cdot j}) + e_{ij}$$

$$B_{j} = \beta + b_{j} \qquad b_{j} \sim N(0, \tau_{b}^{2})$$

Robustness (compared to fixed effects)

$$E(\hat{\beta} \mid \beta, T, \sigma^2, \tau_b^2) = E\left\{ \frac{\sum_{j=1}^{J} \left[\tau_b^2 + \frac{\sigma^2}{n_j \bar{T}_{\cdot j}(1-\bar{T}_{\cdot j})} \right]^{-1} B_j}{\sum_{j=1}^{J} \left[\tau_b^2 + \frac{\sigma^2}{n_j \bar{T}_{\cdot j}(1-\bar{T}_{\cdot j})} \right]^{-1}} \right\}$$

As heterogeneity increases, τ_b^2 *increases, reliance on* $n_j \bar{T}_{\cdot j}(1-\bar{T}_{\cdot j})$ *decreases.*

Estimation of heterogeneity

$$\tau_b^{(m)2} = \frac{\sum_{j=1}^{J}\left[\tau_b^{(m-1)2} + \frac{\sigma^{2(m-1)}}{n_j\bar{T}_{\cdot j}(1-\bar{T}_{\cdot j})}\right]^{-2}[(\hat{B}_j - \beta^{(m-1)})^2 - V_j]}{\sum_{j=1}^{J}\left[\tau_b^{(m-1)2} + \frac{\sigma^{2(m-1)}}{n_j\bar{T}_{\cdot j}(1-\bar{T}_{\cdot j})}\right]^{-2}}$$

As heterogeneity increases, τ_b^2 *increases, reliance on* $n_j\bar{T}_{\cdot j}(1-\bar{T}_{\cdot j})$ *decreases.*

c) Control the Propensity Score

$Level\ 1$

$$Y_{ij} = U_{0j} + B_j T_{ij} + e_{ij}$$

$Level\ 2$

$$U_{0j} = \mu_0 + \gamma \overline{T}_{\cdot j} + u_{0j}$$

$$B_j = \beta + b_j$$

Very similar to centering for average treatment effect,

no good estimate of τ_b^2

Variance Estimation

$$vec(\boldsymbol{\tau}^{(m)}) = \left\{ \sum_{j=1}^{J} \left[\left(\boldsymbol{\tau}^{(m-1)} + \mathbf{V}_j^{(m-1)} \right)^{-1} \otimes \left(\boldsymbol{\tau}^{(m-1)} + \mathbf{V}_j^{(m-1)} \right)^{-1} \right] \right\}^{-1}$$

$$* \sum_{j=1}^{J} \left[\left(\boldsymbol{\tau}^{(m-1)} + \mathbf{V}_j^{(m-1)} \right)^{-1} \otimes \left(\boldsymbol{\tau}^{(m-1)} + \mathbf{V}_j^{(m-1)} \right)^{-1} \right] vec(\mathbf{d}_j^{(m-1)} \mathbf{d}_j^{(m-1)T})$$

$$d_j = \begin{pmatrix} \hat{U}_{0j} - \gamma_0^{(m-1)} - \gamma_1^{(m-1)} \overline{T}_j \\ \hat{B}_j - \beta^{(m-1)} \end{pmatrix}$$

Summary so far

Four commonly used options

* limit what we can estimate
* produce inconsistent estimates of what we can estimate.

6.HLM with Inverse Probability of Treatment Weighting

- IPTW
- Embedding within HLM
- Always produces consistent estimates
- May be quite imprecise

Inverse Probability of Treatment Weights (Robins and Greenland, 2000)

$$If\ \ T_{ij} = 1, \qquad w_{ij} = \bar{T} / \bar{T}_j$$

$$If\ \ T_{ij} = 0, \qquad w_{ij} = (1 - \bar{T}) / (1 - \bar{T}_j)$$

so

$$w_{ij} = T_{ij} \frac{\bar{T}}{\bar{T}_j} + (1 - T_{ij}) \frac{1 - \bar{T}}{1 - \bar{T}_j}$$

Weighting Schemes for HLM

Level-2 weight	Level-1 weight	Result
$\frac{n_j}{\overline{n}}$	$\frac{T_{ij}\overline{T}}{\overline{T}_{\cdot j}} + \frac{(1-T_{ij})(1-\overline{T})}{1-\overline{T}_{\cdot j}}$	Weights site-specific estimates of impact by n_j
1	$\frac{\overline{n}}{n_j}\left(\frac{T_{ij}\overline{T}}{\overline{T}_{\cdot j}} + \frac{(1-T_{ij})(1-\overline{T})}{1-\overline{T}_{\cdot j}}\right)$	Weights site-specific estimates equally

Estimation via Weighted log Likelihood

(weighted complete-data log likelihood)

$$Y_j = X_{ij}^T\gamma + Z_{ij}^T u_j + e_{ij}, \qquad u_j \sim N(0,\tau), \quad e_{ij} \sim N(0,\sigma^2)$$

$$\ln[h(Y,u)] = \ln[f(Y \mid u)] + \ln[p(u)]$$

After weighting :

$$= \sum_{j=1}^{J}\sum_{i=1}^{n_j} l_{wij} = -\tfrac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j} w_{ij}(Y_{ij} - X_{ij}^T\gamma - Z_{ij}^T u_{rj})^2/\sigma^2 - N\log(2\pi\sigma^2)$$

$$-\tfrac{1}{2}\sum_{j=1}^{J} w_{2j} u_j^T \tau^{-1} u_j - J\log(2\pi\,|\tau|)$$

$$\text{where } \sum_{j=1}^{J}\sum_{i=1}^{n_j} w_{ij} = N \qquad \sum_{j=1}^{J} w_{2j} = J,$$

Two examples of weighting schemes

Equally weight people	**Equally weight sites**
Level-1 weight $T_{ij}\frac{\overline{T}}{\overline{T}_j}+(1-T_{ij})\frac{1-\overline{T}}{1-\overline{T}_j}$	Level-1 weight $\frac{\overline{n}}{n_j}\left(T_{ij}\frac{\overline{T}}{\overline{T}_j}+(1-T_{ij})\frac{1-\overline{T}}{1-\overline{T}_j}\right)$
Level-2 weight $n_j/\overline{n}$	Level-2 weight 1
$\hat{\beta}=\sum_{j=1}^{J}n_j\hat{B}_{1j}/\sum_{j=1}^{J}n_j$	$\hat{\beta}=\sum_{j=1}^{J}\hat{B}_j/J$
$\hat{\tau}_b^2=J^{-1}\sum_{j=1}^{J}\frac{n_j}{\overline{n}}\left[(\hat{B}_j-\hat{\beta})^2-V_j\right]$	$\hat{\tau}_b^2=J^{-1}\sum_{j=1}^{J}\left[(\hat{B}_j-\hat{\beta})^2-V_j\right]$
$\hat{\tau}_{b0}=J^{-1}\sum_{j=1}^{J}\frac{n_j}{\overline{n}}\left[(\hat{B}_j-\hat{\beta})(\hat{U}_{0j}-\hat{\mu}_0)-C_j\right]$	$\hat{\tau}_{b0}=J^{-1}\sum_{j=1}^{J}\left[(\hat{B}_j-\hat{\beta})(\hat{U}_{0j}-\hat{\mu}_0)-C_j\right]$

7. Discussion
A Class of Estimators…

We are approximating various balanced designs by maximizing the weighted log likelihood

This gives us a family of method-of-moments estimators

Hence no reliance on normality or homoscedasticity

May be quite imprecise

Extensions

- Extend easily to multi-site clustered randomized designs
- Balance propensity within covariate classes in randomized studies
- Extend to observational studies
- Extend to weighted mediation models (Hong et al, 2012)
- Application to local average treatment effects (Raudenbush, Nomi, and Reardon, 2016)