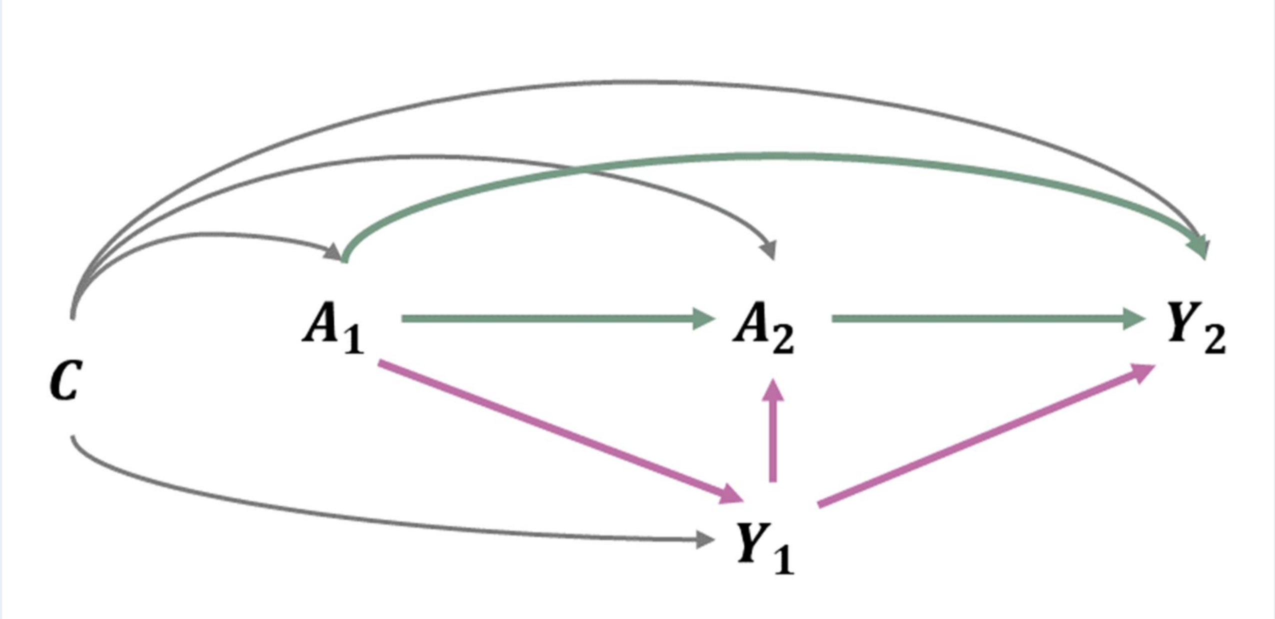
Disentangling Longitudinal Treatment Effects by Regimes: A Comparison of Selection Bias Adjustment Approaches

Hanna Kim and Jee-Seon Kim

1. Study Design and Research Questions

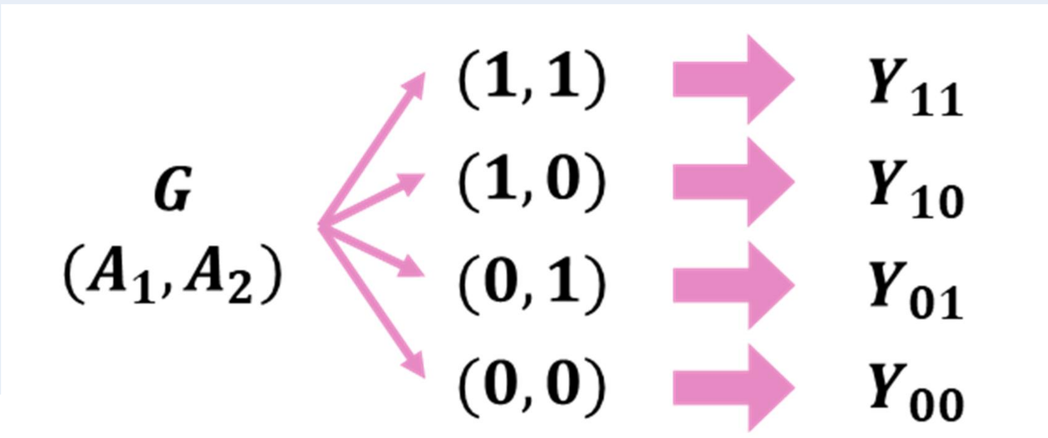
- Multiple treatment regimes occur from longitudinal treatments implemented at multiple time points.
- Contrasting specific regimes, we can determine the effectiveness of competing implementation strategies and explain treatment effect heterogeneity over time.



- *e.g.*, Children's vocabulary skills (Y_2) after staying in the Head Start program ($(A_1, A_2) = (1,1)$), combining Head Start with other childcare services ($(A_1, A_2) = (1,0)$ or $(0,1)$), or never attending Head Start ($(A_1, A_2) = (0,0)$) for two years
- For valid comparison, we need to adjust for baseline covariates (C) and intermediate vocabulary skills (Y_1)

2. Adjusting for Selection Bias in Multiple Treatments

- First, we conceptualize all possible patterns of treatment participation as categories of a multiple treatment (Imbens, 2000).
- *e.g.*, Head Start as a four-category treatment



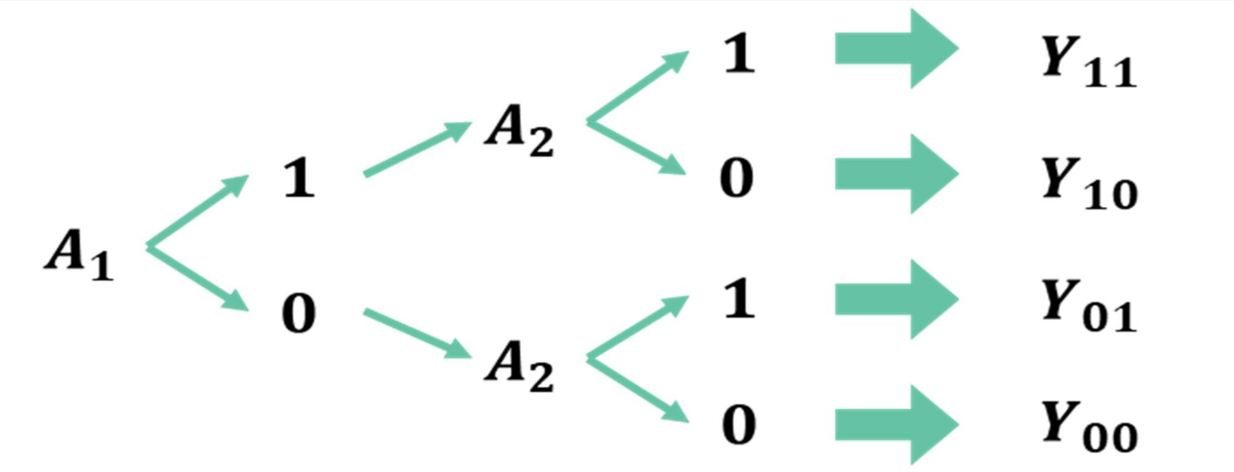
- Generalized propensity scores (GPS)

$$P(G_i = g|C) = \frac{\exp(\beta_{0g} + \beta_{1g}X')}{1 + \sum_{g=1}^{G-1} \exp(\beta_{0g} + \beta_{1g}X')}$$

- Weights computed by entropy balancing (Hainmueller, 2012)
- $E[Y_2|G = g, C = c] = \theta_0 + \theta_1 g + \theta_2 c' + \theta_3 g * c'$
- $\hat{\delta}_{ATE}^{WLS}(g_T - g_C) = \frac{1}{N}\sum_{i=1}^{N}\{\hat{E}_w[Y_2|g_T, c] - \hat{E}_w[Y_2|g_C, c]\}$
- HC3 standard errors computed for the contrasts

3. Adjusting for Selection Bias in Sequential Treatment Regimes

- On the other hand, we consider a sequence of treatments at successive time points.



- *e.g.*, Head Start as a sequence of binary treatments
- Multiplicative inverse probability of treatment weights are produced, and potential outcomes are estimated for different sequences of treatment participation.

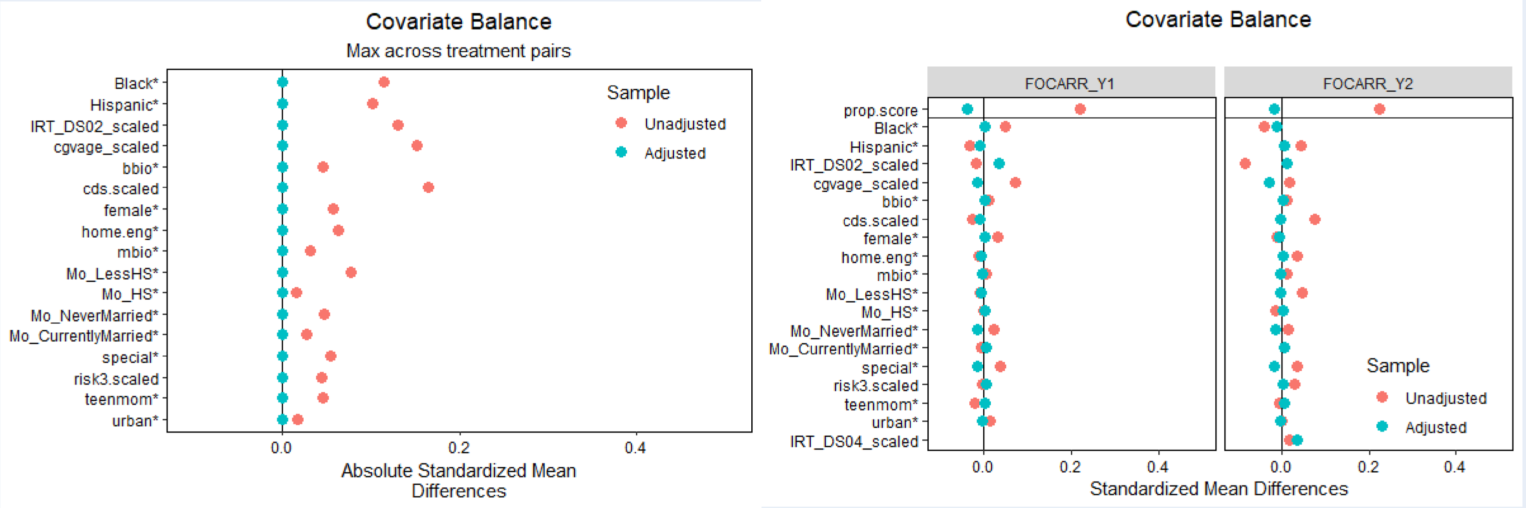
$$logit[P(A_1 = 1|C = c)] = \beta_0 + \beta_1 c'$$

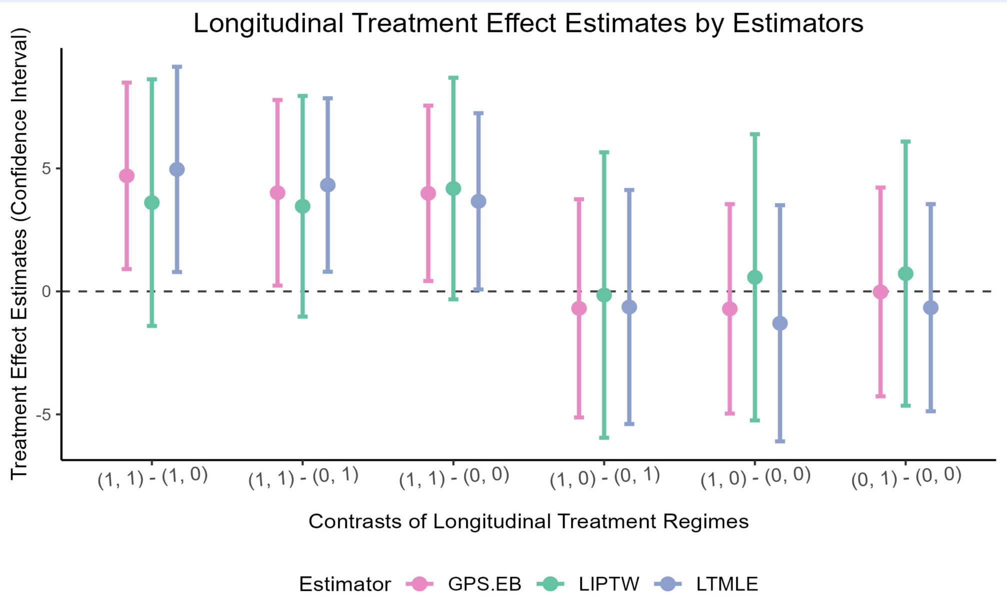
$$logit[P(A_2 = 1|C = c, A_1 = a, Y_1 = y)] = \beta_2 + \beta_3 c' + \beta_4 a + \beta_5 y$$

- Longitudinal IPTW estimation
- $w(A_1 = a_1, A_2 = a_2) = \frac{P(A_1 = a_1)}{P(A_1 = a_1|C)} \times \frac{P(A_2 = a_2|A_1 = a_1)}{P(A_2 = a_2|A_1 = a_1, C, Y_1)}$
- $\hat{\delta}_{ATE}^{LIPTW}(g_T - g_C) = \frac{\sum_{i=1}^{N} w(A_1 = a_{1T}, A_2 = a_{2T}) \times Y_i}{\sum_{i=1}^{N} w(A_1 = a_{1T}, A_2 = a_{2T})} - \frac{\sum_{i=1}^{N} w(A_1 = a_{1C}, A_2 = a_{2C}) \times Y_i}{\sum_{i=1}^{N} w(A_1 = a_{1C}, A_2 = a_{2C})}$
- Longitudinal TMLE presented for comparison

4. Real Data Analysis Results

- Head Start Impact Study (HSIS) data analysis
- Both GPS entropy balancing and longitudinal IPTW achieved covariate balance.





5. Discussion

- This study proposed two distinct ways of inverse probability weighting based on different conceptualization of longitudinal treatment regimes.
- Considering longitudinal Head Start as a multiple treatment produced smaller standard errors compared to the sequential treatment approach.
- However, time-varying covariates (e.g., Y_1) cannot be appropriately incorporated within the multiple treatment approach.
- Alternative estimators such as the LTMLE can help gain precision and double robustness.
- Further extensions may address partially or fully clustered data and evaluate the performances of multiple estimation methods with simulation studies.

Support for this research was provided by the University of Wisconsin-Madison Graduate School Fellowship and the Wisconsin Alumni Research Foundation.

Correspond to **Hanna Kim**: hanna.kim@wisc.edu
Department of Educational Psychology, University of Wisconsin-Madison