Causal Effect Moderation with Application to Level of Care Matching for Adolescent Substance Abuse Treatment

by **Daniel Almirall** (Duke Biostatistics & Durham VAMC)

with Daniel McCaffrey & Andrew Morral (RAND), and Susan Murphy (UMich)

(speaker: Joe Schafer (Penn State))

Thursday, May 29, 2008

Society for Prevention Research, San Francisco

Outline

1. Background, Motivation
2. Causal Effect Moderation
3. The Problem of Confounding
4. A Method for Assessing Causal Effect Moderation
5. Illustration With Data
6. Discussion

Background: Providing Context

Consider The American Society of Addiction Medicine's Patient Placement Criteria (ASAM-PPC) For Adolescents

What is it? A program designed to assist in level of care (LOC) and/or treatment planning for adolescents that are substance abusers

Who uses it? Intended for substance abuse, mental health, child welfare, & juvenile justice system providers; program administrators; educators

How is it used? Based on a needs assessment, along six dimensions

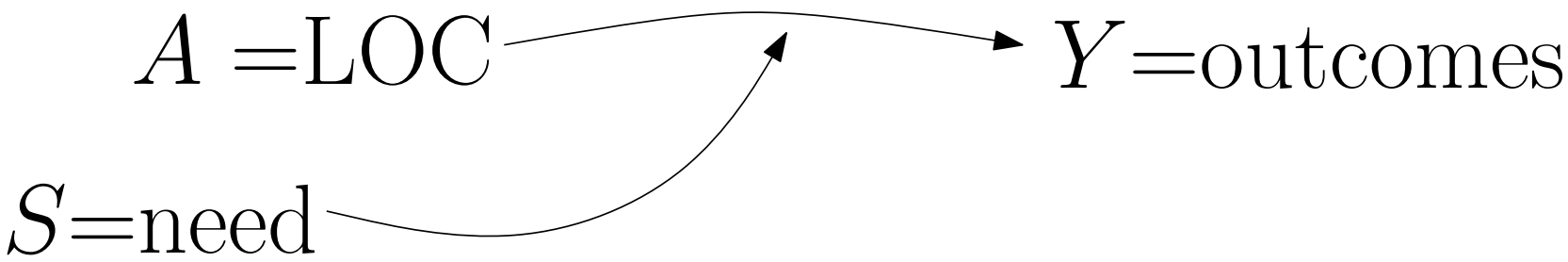
Six Dimensions of Need: (1) Withdrawal Potential, (2) Biomedical Conditions, Complications, (3) Behavioral or Cognitive Complications, (4) Readiness to Change, (5) Relapse Potential, (6) Recovery Environment

Background: Primary Objective

To evaluate the usefulness of the ASAM-PPC for matching adolescent substance abusers to the optimal level of care (LOC).

- ◇ A proper evaluation requires an experimental study (costly, not feasible)
- ◇ We actually consider a less ambitious objective: To explore which of the six ASAM-PPC dimensions of need, if any, are more important for matching adolescent substance abusers to the appropriate level of care (LOC).
- ◇ How do we think about addressing this question using observational data?

Causal Effect Moderation: Heuristically



Causal Effect Moderation: Formally

$$\mu(s) \equiv E(Y(1) - Y(0) \mid S = s),$$

the conditional average causal effect of inpatient $a = 1$ versus outpatient $a = 0$ on Y, given measured of need S at the value s.

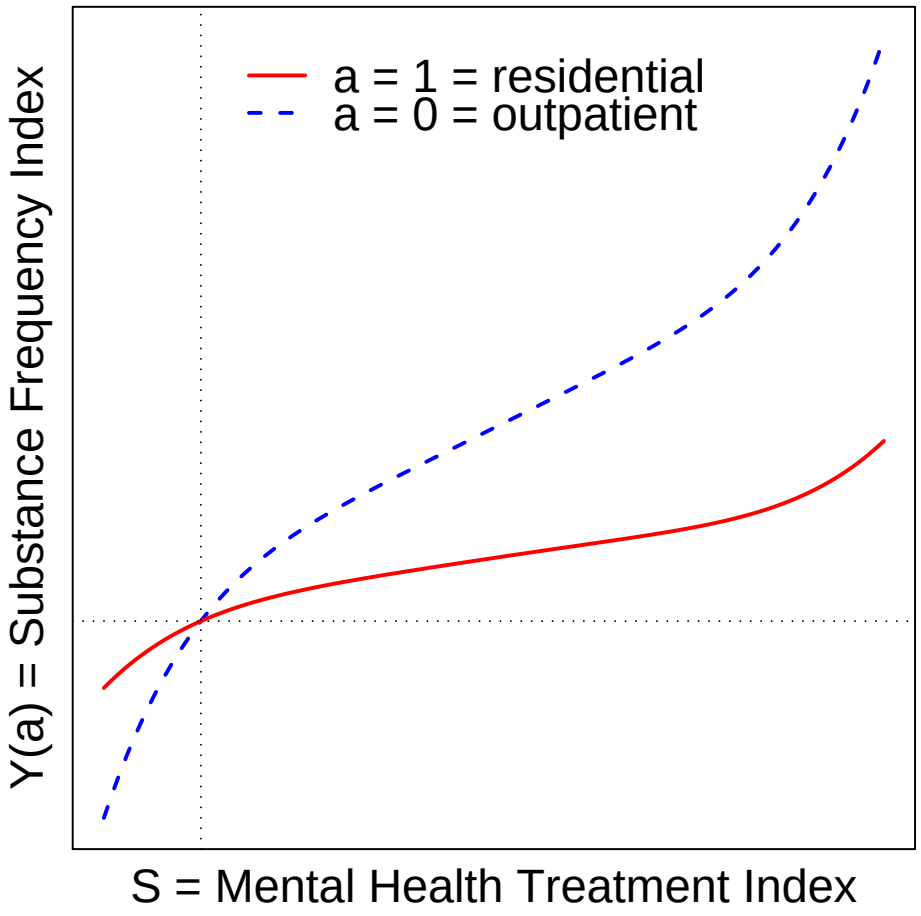
Recall:

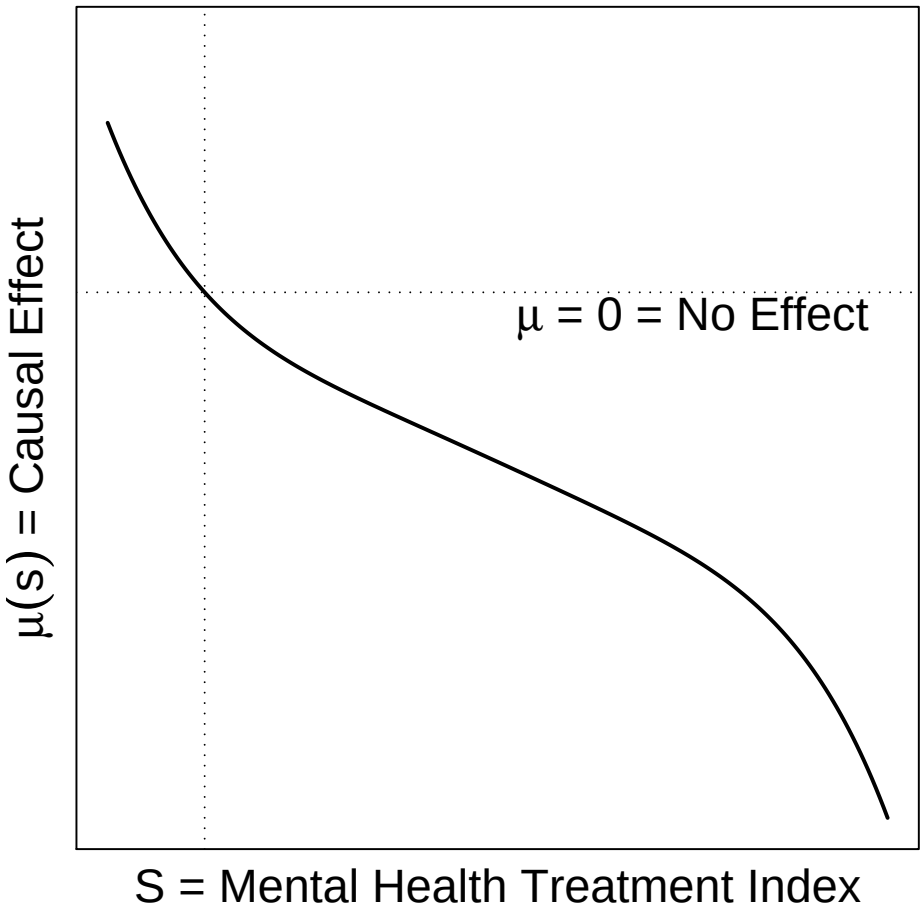
- $Y(1)$ is the outcome had the adolescent received inpatient txt
- $Y(0)$ is the outcome had the adolescent received outpatient txt
- S is a measure of need (think 6 ASAM-PPC dimensions)

Definition:

S is a moderator of the effect of LOC on $Y(a)$ if $\mu(s)$ if non-constant in s.

Causal Effect Moderation: In Context





Causal Effect Moderation: A Model

Recall that we are interested in studying the function

$$\mu(s,a) = E(Y(a) - Y(0) \mid S = s) = a \times E(Y(1) - Y(0) \mid S = s).$$

How does this function relate to the conditional mean $E(Y(a) \mid S)$?

$$\begin{aligned} E(Y(a) \mid S = s) &= E(Y(0) \mid S = 0) \\ &\quad + \Big(E(Y(0) \mid S = s) - E(Y(0) \mid S = 0)\Big) \\ &\quad + E(Y(a) - Y(0) \mid S = s) \\ &= \eta_0 + \phi(s) + \mu(s,a) \\ &= \eta_0 + \eta_1 s + a \times (\beta_1 + \beta_2 s). \end{aligned}$$

If $\beta_2 = 0$, then no evidence of effect moderation by S.

Causal Effect Moderation: Generally

General Implications? General Relevance?

Theoretical Implication: Understanding the heterogeneity of the effects of causes may suggest new scientific (etiologic) hypotheses to be tested.

Practical Implication: Identifying types, or subgroups, of individuals for which treatment is not effective may suggest altering the treatment to suit the needs of that particular type of individual.

Sociologist Yu Xie's <u>Variability Principle</u>: We really want $\delta_i = Y_i(1) - Y_i(0) \ \forall \ i$. We "settle" for groupings of effects (here, groupings defined by S)

We learn more about δ_i using $\mu(s)$, than with $E(\delta) = E(Y(1) - Y(0))$.

ASAM-PPC Application: Observed Data

- $N = 1984$ adolescents undergoing treatment for substance abuse problems observed in either an inpatient/residential level of care (LOC $= A = 1$) setting or an outpatient setting (LOC $= A = 0$).
- $Y =$ SFI is a measure of frequency of substance abuse 12 months after beginning treatment. Other outcomes also available.
- GAIN Instrument: we have 64 measures of treatment need measured prior to the start of treatment. Each variable has been assigned, a priori, to correspond with one of the six ASAM-PPC dimensions of need.
- We have another 22 auxiliary covariates (e.g., demographic)

Data: 64 measures of need, the S's.

Intoxication/Withdrawal Potential: CEI, CWI, HSCWD, RECUSE2, SFI7P, DLYUSE, WKHR

Biomedical Conditions: FAMHIST, HLTHPRB, HPI3P, P3, PHTI

Emotional/Behavioral Conditions: ARRTOT, BCI, CJSI, CRNTCJ, CVI, DCI2, EPI7P, HBPY, IAI5P, IMDI, M1C2, M5, MHTI, RACMDP

Readiness for Change: HIRES, LOMOT, NFIP, NPI, NRECPRB, TMI, TRI

Relapse Potential: CSUDPTX, NDL90, POI, SATI, SDIM, SDIY, SEI, SPIM, SPIY, SPR, UND15, DLYUSE, WKHR

Environment: E5, EMAI, EMPI, ERI, GSI, GSSI, GVI, HIVICT, HMLSRUN, LRI, MAXCE, OSSI, PSSI, RERI12P, SRI, TAI5P, TPI, VRI, WKYFMP

Can We Use OLS Regression?

Recall Our Analysis Goal: To screen for important S's (i.e., measures of need, among the 64) that are plausible effect moderators. This will help us understand which, if any, are the important ASAM-PPC dimensions to consider when planning inpatient versus outpatient treatment.

Can we use the following regression model, for each S?

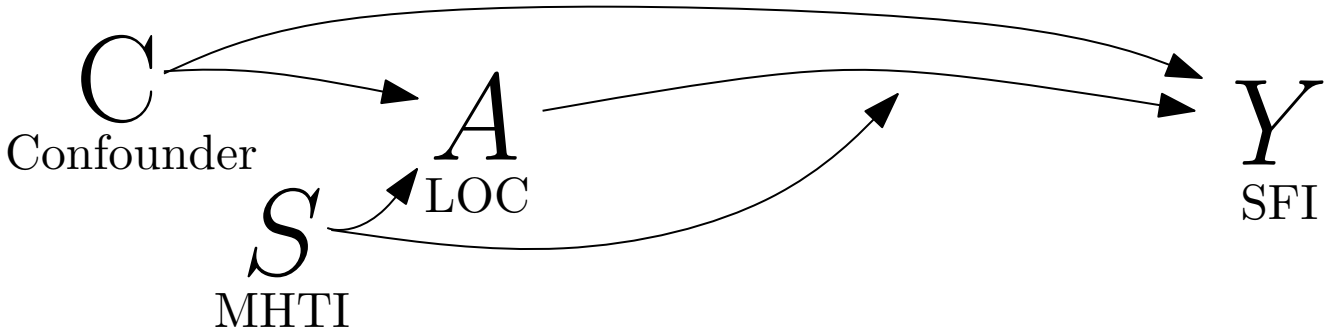
$$E(Y \mid S = s, A = a) = \eta_0^* + \eta_1^* s + a \times (\beta_1^* + \beta_2^* s).$$

And interpret $\beta_2^* = 0$ to mean that S is not a moderator?

Likely, the answer is **<u>no</u>**. Why? Confounding bias.

The Problem with Confounders

Definition. Confounders C are variables related to both the putative cause LOC and the outcome Y.



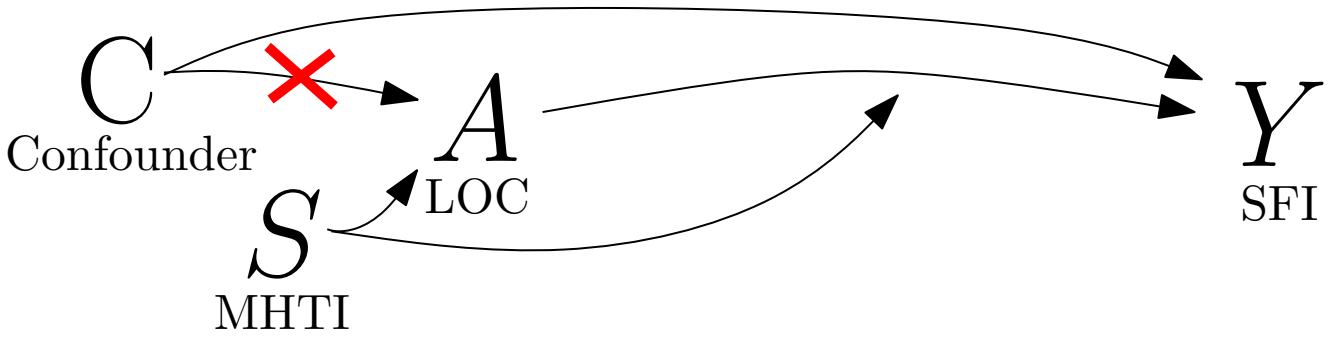
Adolescents that are worse off are likely to be sent to inpatient treatment. If not accounted for, this imbalance results in bias in the OLS regression.

Putative moderators of interest (the S's) may also be confounders.

OLS is appropriate if S is the <u>only confounder</u> (the only reason for the imbalance) is S.

Solution: A Weighting Method

Weighting. Weighted LS regression using weights that remove (or reduce) the impact of C on A.

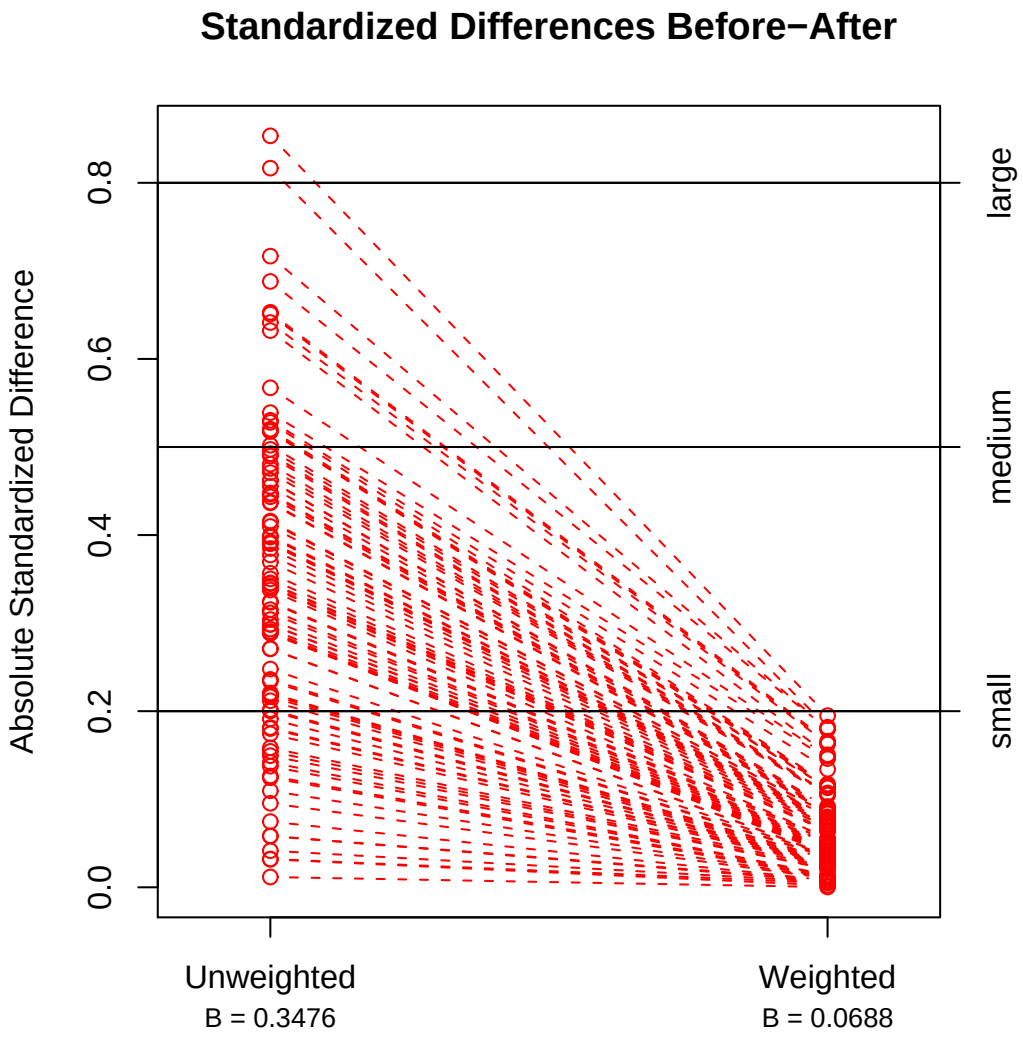


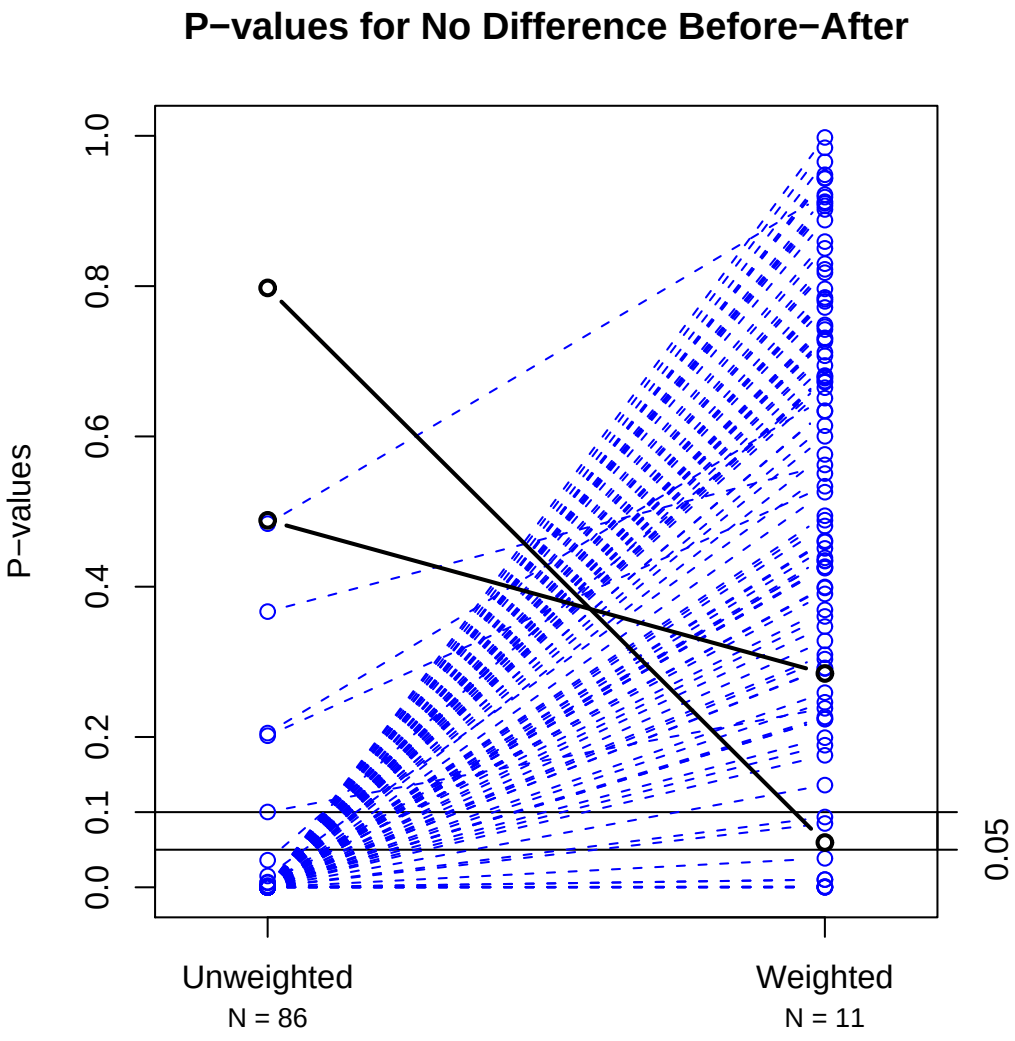
$$w(S, C, A) = A \times \frac{p(S)}{p(S, C)} + (1 - A) \times \frac{1 - p(S)}{1 - p(S, C)}.$$

$S =$ the moderator of interest; $C =$ confounding/auxiliary variables

The numerator is optional: helps with efficiency and is conceptually intuitive.

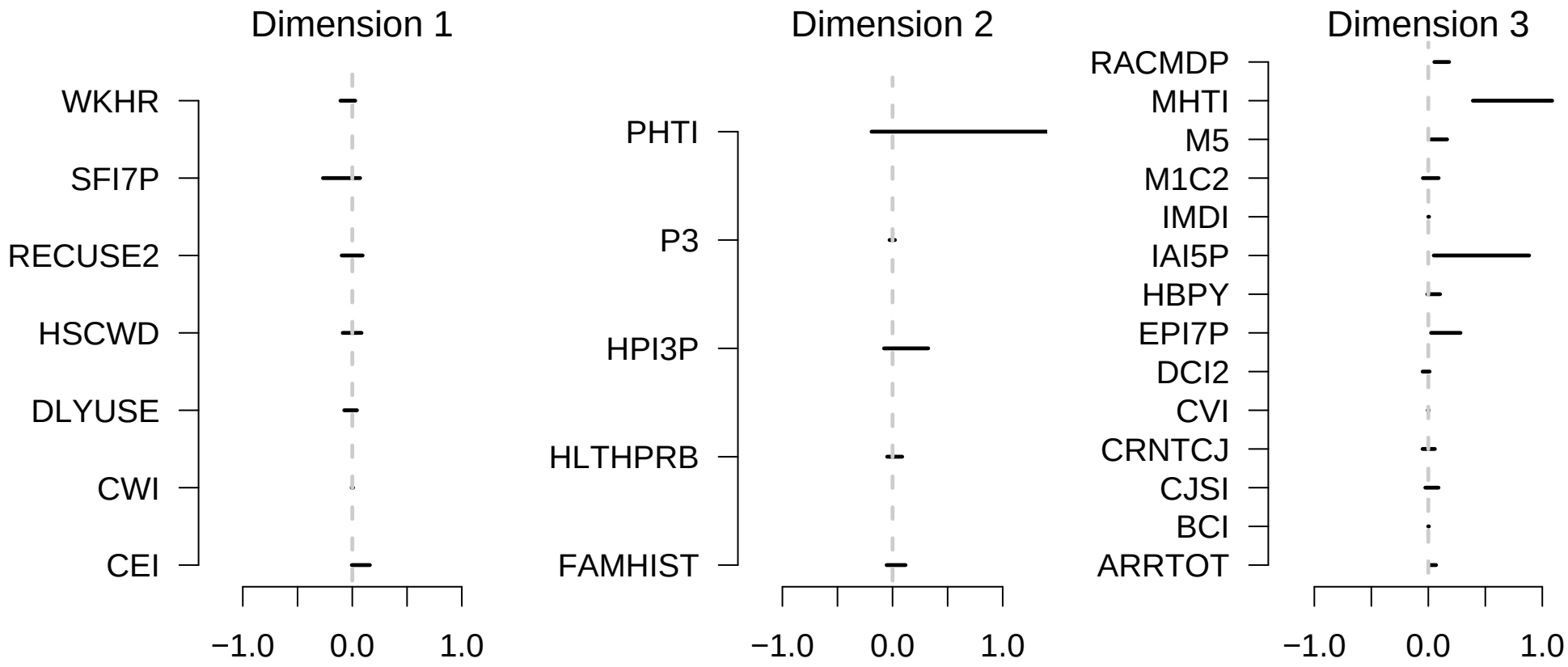
Illustration: Did weights improve balance?





Some Results: Dimensions 1-3

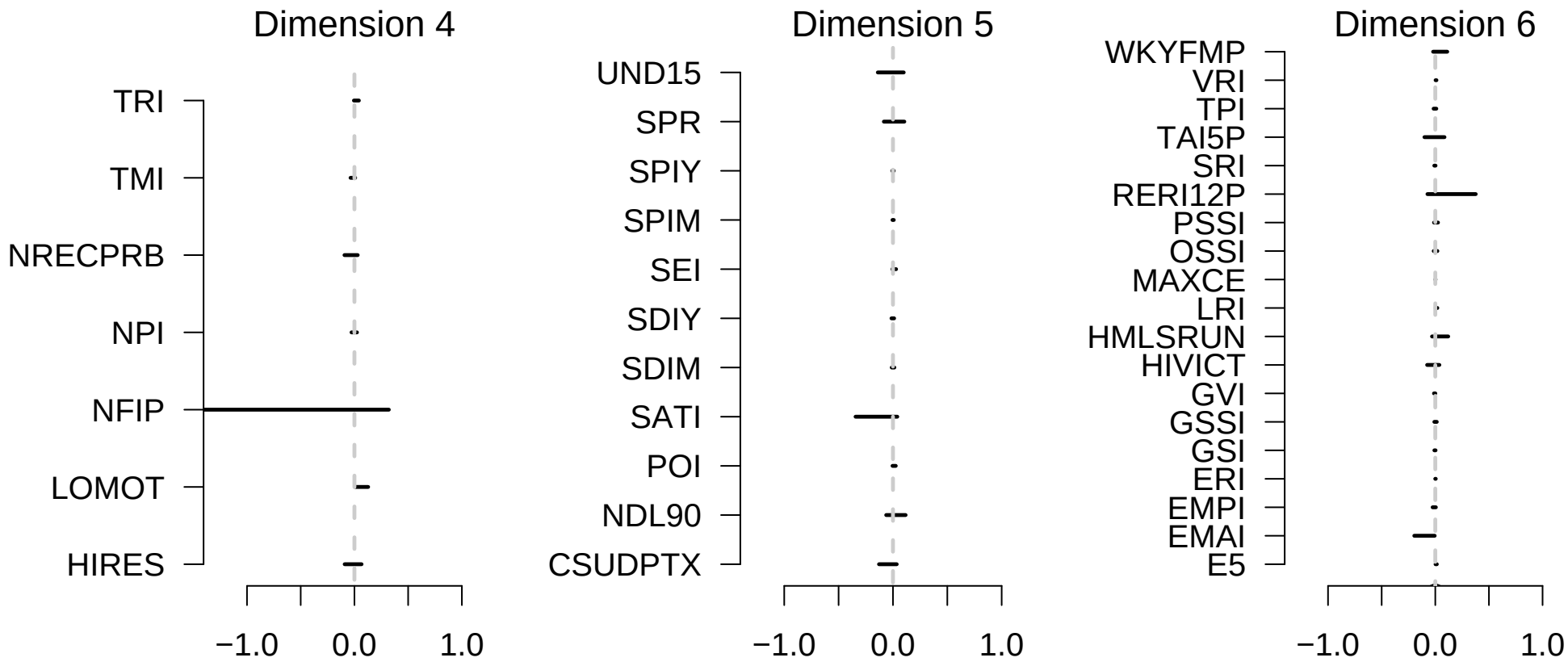
Model Used for Each S**:** $E(Y(a) \mid S = s) = \eta_0 + \eta_1 s + a \times (\beta_1 + \beta_2 s).$



Shown are point estimates for β_2 and its corresponding 90% asymptotic CIs.

Some Results: Dimensions 4-6

Model Used for Each S: $E(Y(a) \mid S = s) = \eta_0 + \eta_1 s + a \times (\beta_1 + \beta_2 s).$



Shown are point estimates for β_2 and its corresponding 90% asymptotic CIs.

Discussion

- **Messages:** Concept of Causal Effect Moderation; Distinction between Confounders and Moderators; Weighting Method to resolve confounding
- **Analysis:** In terms of $Y =$ SFI, **Dimension 3 (Emotional Conditions) was found to be the most important dimension** to consider when deciding whether to match adolescent substance abusers to outpatient vs. inpatient LOC.
- **Limitation:** But why consider each S by itself?
- **Moving Forward:** Currently exploring **how to construct features** of measures of patient need that are moderators. This is more useful to scientists than considering each moderator by itself. With this data, it would be natural to consider dimension-specific features.

Thank you!

Please email Danny with questions, concerns, and/or ideas about how to improve this research!

Daniel Almirall

Biostatistics, Duke University Medical Center

HSRD, Durham VA Medical Center

dalmiral@umich.edu