Principal Stratification and Instrumental Variables for Estimating Complier-Average Causal Effects

Joseph Hogan

Department of Biostatistics
Brown University School of Public Health

February 2016

Outline of Topics

1. Causal effects and randomization
   - Principal stratification and CACE
   - Estimating CACE
   - Assumptions
   - Instrumental variables and CACE
2. Example 1: Flu Vaccine
3. Example 2: Smoking cessation study
   - Principal stratification
   - Assumptions
   - Targets of inference
   - Compliance models
   - Analysis and results
4. Summary points

Two clinical trials

Encouragement for flu vaccine Intervene to encourage doctors to give flu shots to senior citizens. Does this reduce hospitalization?

Exercise for smoking cessation Intervene to promote vigorous exercise during smoking cessation period. Does this increase quit rates?

Causal effects and randomization

Basic set up in terms of data

$$
\begin{aligned}
Z &= \text{Randomization to treatment arm} \\
A &= \text{Compliance or receipt of } \textit{assigned} \text{ treatment} \\
Y &= \text{Outcome} \\
X &= \text{Baseline covariates}
\end{aligned}
$$

Which causal effects can be estimated?

Defining causal effects

Randomization

$$Z = 1 \text{ if randomized to treatment, 0 if placebo}$$

For each individual, can define *potential outcomes*

$$\begin{aligned} Y_1 &= \text{outcome if } Z = 1 \\ Y_0 &= \text{outcome if } Z = 0 \end{aligned}$$

Causal effect of Z on Y is

$$Y_1 - Y_0$$

Average causal effect of Z on Y is

$$\theta = E(Y_1 - Y_0)$$

Estimating causal effects

Key issue is that we only observe one potential outcome

$$Y = \begin{cases} Y_1, & \text{if } Z = 1 \\ Y_0, & \text{if } Z = 0 \end{cases}$$

If Z is randomized, then

$$\begin{aligned} E(Y_1) &= E(Y \mid Z = 1) \\ E(Y_0) &= E(Y \mid Z = 0) \end{aligned}$$

Hence $E(Y_1 - Y_0)$ can be estimated using difference in sample means

$$\overline{Y}_{Z=1} - \overline{Y}_{Z=0}$$

What about causal effect of A on Y?

Depends on what you are willing to assume.

- If everyone has some probability of receiving A, then can use structural modeling and inverse weighting to estimate average effect of A.

- If there are some who would not take A, this is an important stratification of the population. Need to focus on causal effect of A within certain subpopulations.

Complier-average causal effect

Causal effect of Z on Y among those who take the treatment prescribed by Z.

Principal stratification

Stratum	Proportion	A_0	A_1	$E(Y_1 - Y_0 \mid A_1, A_0)$
Never takers	π_{NT}	0	0	θ_{NT}
Compliers	π_{C}	0	1	θ_{C}
Defiers	π_{D}	1	0	θ_{D}
Always takers	π_{AT}	1	1	θ_{AT}

$$\text{CACE} \;=\; E(Y_1 - Y_0 \mid A_1 = 1, A_0 = 0) \;=\; \theta_{\text{C}}$$

$$\text{IT effect} \;=\; \sum_{S \in \{\text{NT,C,D,AT}\}} \theta_S \pi_S$$

Relating observed to latent data

Z	A	(A_0, A_1)	Stratum
0	0	(0,0) or (0,1)	NT or C
0	1	(1,1) or (1,0)	AT or D
1	0	(0,0) or (1,0)	NT or D
1	1	(0,1) or (1,1)	C or AT

Complier-average causal effect

CACE is:

- Causal effect of being randomized to treatment, among those who took assigned treatment
- *Randomized* comparison between those who do and do not receive treatment

Potential problems with CACE

- Causal effect only applies to a subgroup
- The subgroup cannot be identified
- Untestable assumptions needed to estimate CACE from data

Estimation of CACE

To estimate CACE from data

- Need to impose untestable assumptions
- Sometimes, inference must be expressed in terms of *bounds* or *sensitivity analysis*

Main types of assumptions

- Randomization
- Monotonicity
- Exclusion restrictions

Assumption: Randomization

Randomization

$Z \perp\!\!\!\perp \{Y_0, Y_1, A_0, A_1\}$

- Treatment is randomly allocated, thereby balancing individual characteristics.
- Enables valid estimate of causal effect of Z on Y and Z on A:

$$\widehat{E}(Y_1 - Y_0) = \overline{Y}_{Z=1} - \overline{Y}_{Z=0}$$

$$\widehat{E}(A_1 - A_0) = \overline{A}_{Z=1} - \overline{A}_{Z=0}$$

Assumption: Monotonicity

Monotonicity

$A_1 \geq A_0$

- Individuals are at least as likely to receive treatment when randomized to $Z = 1$, as compared to $Z = 0$.
- This assumption rules out the possibility of 'defiers'
- Makes it possible to estimate proportion in each principal stratum

Assumption: Exclusion restrictions

Exclusion restrictions

$$E(Y_1 - Y_0 \mid A_0 = A_1) \quad = \quad 0$$

Equivalently

$$\begin{aligned} \theta_{\text{AT}} &= 0 \\ \theta_{\text{NT}} &= 0 \end{aligned}$$

- Randomization only affects outcome via receipt of treatment
- Put another way: there is no direct effect of the randomization Z

Effect of Exclusion Restrictions, Monotonicity

Assumptions constrain strata probabilities and causal effects

Stratum	Proportion	A_0	A_1	$E(Y_1 - Y_0 \mid A_1, A_0)$
Never takers	π_{NT}	0	0	0
Compliers	π_{C}	0	1	θ_{C}
Defiers	0	1	0	θ_D
Always takers	π_{AT}	1	1	0

Can therefore show

$$
\begin{aligned}
\pi_{\text{C}} &= \overline{A}_{Z=1} - \overline{A}_{Z=0} \\
\pi_{\text{AT}} &= P(A = 1 \mid Z = 0) \\
\pi_{\text{NT}} &= P(A = 0 \mid Z = 1)
\end{aligned}
$$

Instrumental variables estimator of CACE

When these assumptions hold, the IV estimator is a consistent estimator of CACE, treating Z as the IV

Key assumptions for an IV:

- Correlated with receipt of treatment
- Uncorrelated with potential outcomes (randomizer)

The IV estimator is

$$\widehat{\theta}_{C} = \frac{\overline{Y}_{Z=1} - \overline{Y}_{Z=0}}{\overline{A}_{Z=1} - \overline{A}_{Z=0}} = \frac{\text{IT effect of } Z \text{ on } Y}{\pi_{C}}$$

Other representations of the IV estimator

It is a ratio of covariances

$$\widehat{\theta}_{\mathrm{C}} = \frac{\mathrm{cov}(Y, Z)}{\mathrm{cov}(A, Z)}$$

Using a heuristic version of the chain rule, can also think of it as dY/dA in a causal metric,

$$\widehat{\theta}_{\mathrm{C}} = \frac{dY/dZ}{dA/dZ}$$

Encouragement Designs and Flu Vaccine

Hirano, Imbens, Rubin and Zhou. Biostatistics 2000; 1:69–88.

$$
\begin{aligned}
Z &= \text{Physician receipt of letter to encourage flu shot} \\
A &= \text{Patient decides to receive flu shot} \\
Y &= \text{Hospitalization} \\
X &= \text{Baseline covariates}
\end{aligned}
$$

Summary Statistics

Table 1. *Summary statistics, flu data (sample size 2893)*

		Means			Means		
	Grand mean	No letter $Z_i^{\text{obs}} = 0$	Letter $Z_i^{\text{obs}} = 1$	t-stat.	No flu shot $D_i^{\text{obs}} = 0$	Flu shot $D_i^{\text{obs}} = 1$	t-stat.
Letter (Z_i^{obs})	0.514	0	1	—	0.475	0.631	−7.5
Flu Shot (D_i^{obs})	0.250	0.190	0.307	−7.3	0	1	—
Hospitalization (Y_i^{obs})	0.085	0.092	0.078	1.4	0.085	0.084	0.1
Age (X_{i1}^{obs})	65.2	65.0	65.4	−0.8	64.7	66.8	−4.1
COPD (X_{i2}^{obs})	0.283	0.290	0.277	0.8	0.264	0.343	−4.0

Summary Statistics

Table 1. *Summary statistics, flu data (sample size 2893)*

		Means			Means		
	Grand mean	No letter $Z_i^{\text{obs}} = 0$	Letter $Z_i^{\text{obs}} = 1$	t-stat.	No flu shot $D_i^{\text{obs}} = 0$	Flu shot $D_i^{\text{obs}} = 1$	t-stat.
Letter (Z_i^{obs})	0.514	0	1	—	0.475	0.631	−7.5
Flu Shot (D_i^{obs})	0.250	0.190	0.307	−7.3	0	1	—
Hospitalization (Y_i^{obs})	0.085	0.092	0.078	1.4	0.085	0.084	0.1
Age (X_{i1}^{obs})	65.2	65.0	65.4	−0.8	64.7	66.8	−4.1
COPD (X_{i2}^{obs})	0.283	0.290	0.277	0.8	0.264	0.343	−4.0

Summary Statistics

Table 1. *Summary statistics, flu data (sample size 2893)*

		Means			Means		
	Grand mean	No letter $Z_i^{\text{obs}} = 0$	Letter $Z_i^{\text{obs}} = 1$	t-stat.	No flu shot $D_i^{\text{obs}} = 0$	Flu shot $D_i^{\text{obs}} = 1$	t-stat.
Letter (Z_i^{obs})	0.514	0	1	—	0.475	0.631	−7.5
Flu Shot (D_i^{obs})	0.250	0.190	0.307	−7.3	0	1	—
Hospitalization (Y_i^{obs})	0.085	0.092	0.078	1.4	0.085	0.084	0.1
Age (X_{i1}^{obs})	65.2	65.0	65.4	−0.8	64.7	66.8	−4.1
COPD (X_{i2}^{obs})	0.283	0.290	0.277	0.8	0.264	0.343	−4.0

Flu shot example

Monotonicity: There are no individuals who would receive the flu shot with no encouragement, and refuse it if encouraged.

- Probably plausible.

Exclusion restrictions: Can be placed on AT, NT, or both.

- In practical terms: encouragement has no effect on those who will always *refuse/choose* (NT/AT) to receive the shot.
- Possibly not plausible – encouragement may come with other advice about avoiding flu. May therefore have direct effect on hospitalization.

Principal strata membership probabilities

$\Pr(C_i = c)$	0.119	(0.014)
$\Pr(C_i = n)$	0.692	(0.008)
$\Pr(C_i = a)$	0.189	(0.007)

Causal treatment effects

Table 4. *Summary statistics: posterior distributions*

Excl. Res. Never-takers ⟶	Yes		Yes		No		No	
Excl. Res. Always-takers ⟶	Yes		No		Yes		No	
Estimand	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
ITT_c	−0.082	(0.068)	−0.037	(0.078)	−0.196	(0.147)	−0.168	(0.161)
ITT_n	0	0	0	0	0.022	(0.026)	0.025	(0.027)
ITT_a	0	0	−0.053	(0.032)	0	0	−0.058	(0.033)
ITT	−0.010	(0.008)	−0.014	(0.008)	−0.009	(0.007)	−0.013	(0.008)

Causal treatment effects

Table 4. *Summary statistics: posterior distributions*

Excl. Res. Never-takers $\longrightarrow$	Yes		Yes		No		No	
Excl. Res. Always-takers $\longrightarrow$	Yes		No		Yes		No	
Estimand	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
ITT_c	−0.082	(0.068)	−0.037	(0.078)	−0.196	(0.147)	−0.168	(0.161)
ITT_n	0	0	0	0	0.022	(0.026)	0.025	(0.027)
ITT_a	0	0	−0.053	(0.032)	0	0	−0.058	(0.033)
ITT	−0.010	(0.008)	−0.014	(0.008)	−0.009	(0.007)	−0.013	(0.008)

Causal treatment effects

Table 4. *Summary statistics: posterior distributions*

Excl. Res. Never-takers ⟶	Yes		Yes		No		No	
Excl. Res. Always-takers ⟶	Yes		No		Yes		No	
Estimand	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
ITT_c	−0.082	(0.068)	−0.037	(0.078)	−0.196	(0.147)	−0.168	(0.161)
ITT_n	0	0	0	0	0.022	(0.026)	0.025	(0.027)
ITT_a	0	0	−0.053	(0.032)	0	0	−0.058	(0.033)
ITT	−0.010	(0.008)	−0.014	(0.008)	−0.009	(0.007)	−0.013	(0.008)

Conclusions

- ITT effect may overstate causal effect of vaccine
- Reason: there may be a direct effect of the intervention (letter) among always- or never-takers
  - Always takers are probably at higher risk for hospitalization. Hence letter may hasten timing of the shot, or lead to other preventive measures
  - For never takers, probably ok to use exclusion restriction
- Exclusion restrictions do reduce uncertainty – should be imposed with care

Example 2: Commit to Quit Study

Roy, Hogan, Marcus, 2008 Biostatistics

- Intervention study on effect of supervised exercise on smoking cessation
- Key feature of this example: two active treatments
- Question that can be answered by PS: *What is the effect of an intervention among those who would accept it if offered?*
- Compare to this question: *What is the effect of an intervention if it was accepted by everyone?*

Motivation

Smoking prevalence rates among women were declining at a slower rate than among men

Fear of weight gain is a motivator for continued smoking

This is especially true among female smokers

Participants

Sedentary women aged 18 to 65 who are regular smokers

281 subjects agreed to participate

147 in wellness arm, 134 in exercise arm

Smoking status and treatment compliance recorded each week

Control group intervention

Cognitive-behavioral therapy augmented by wellness education program

Cognitive-behavioral smoking cessation therapy (CBT)

- 12-session, group-based smoking cessation program
- This is standard therapy.

Wellness education program

- 3 supervised health education lectures each week

Exercise group intervention

Cognitive-behavioral therapy augmented by exercise program

Cognitive-behavioral smoking cessation therapy (CBT)

- 12-session, group-based smoking cessation program

Vigorous exercise program

- 3 supervised exercise sessions per week

Data

$Z =$ randomization to exercise (1) or wellness (0)

$X =$ baseline covariates

$A =$ compliance indicator (1 if yes, 0 if no)

$Y =$ smoking cessation (1 if yes, 0 if no)

Basic principal strata

S	A_0	A_1	description
0	0	0	non-compliant with either treatment
1	1	0	only compliant with standard treatment
2	0	1	only compliant with new treatment
3	1	1	compliant with either treatment

Strata membership

Given Z and A, only two possible values of S

Z	A	S
0	0	0 or 2
1	0	0 or 1
0	1	1 or 3
1	1	2 or 3

Assumptions

Randomization

$Z \perp\!\!\!\perp \{Y_0, Y_1, A_0, A_1\}$

Assumptions

Treatment access restriction

Subjects in group $Z = z$ do not have access to the active treatment assigned in arm $Z = 1 - z$, for $z = 0, 1$.

Assumptions

Monotonicity

$$P(A_1 = 1 \mid A_0 = 1) \geq P(A_1 = 1 \mid A_0 = 0)$$

Those who would comply with 'wellness' are more likely to also comply with 'exercise', when compared to those who would not comply with 'wellness'

Assumptions

Exclusion Restriction

$E(Y_1 - Y_0 \mid A_0 = 0, A_1 = 0) = 0$

There is no causal effect of being randomized to treatment among those who would comply with neither intervention.

Targets of inference

$$E(Y_1 - Y_0) = E(Y_1 - Y_0 | S \in \{0, 1, 2, 3\})$$

- This is the usual ITT effect
- Average response if everyone were randomized to new treatment minus the average response if everyone were randomized to standard treatment.

Targets of inference

$$E(Y_1 - Y_0) = E(Y_1 - Y_0 | S \in \{0, 1, 2, 3\})$$

- This is the usual ITT effect
- Average response if everyone were randomized to new treatment minus the average response if everyone were randomized to standard treatment.

$$E(Y_1 - Y_0 | S = 3)$$

- Causal effect of assignment among subpopulation of compliers
- Average causal effect comparing new treatment to standard treatment among compliers

Targets of inference

$$E(Y_1 - Y_0) = E(Y_1 - Y_0 | S \in \{0, 1, 2, 3\})$$

- This is the usual ITT effect
- Average response if everyone were randomized to new treatment minus the average response if everyone were randomized to standard treatment.

$$E(Y_1 - Y_0 | S = 3)$$

- Causal effect of assignment among subpopulation of compliers
- Average causal effect comparing new treatment to standard treatment among compliers

$$E(Y_1 - Y_0 | S \in \{2, 3\})$$

- Causal effect of assignment among subpopulation that would comply with new treatment

What's different about this example: Use covariates

If S were observed, estimating the principal effects would be easy.

Given Z and A for a subject, S can only take on two possible values.

Use **covariates** to help identify which of these two strata the subject is in.

- If we have good predictors of compliance, class membership will be nearly identified for some subjects (i.e., class probability near 1)
- Essentially, weight responses by class probabilities to estimate class-specific means.

Use covariates to determine stratum membership

In order to identify the principal strata, we need to identify $[A_0, A_1 | X]$.

Strategy:

- Specify marginal model $[A_0 | X]$
- Specify marginal model $[A_1 | X]$
- Specify an association model $[A_1 | A_0, X]$ that is compatible with the two marginal models

The association model is indexed by a dependence parameter ϕ that cannot be identified by data

Interpretation of ϕ

If $\phi = 0$, then have conditional independence

- conditional independence $A_0 \perp\!\!\!\perp A_1 | X$

If $\phi = 1$, then strongly dependent

- Strongest degree of dependence that is still compatible with marginal distributions

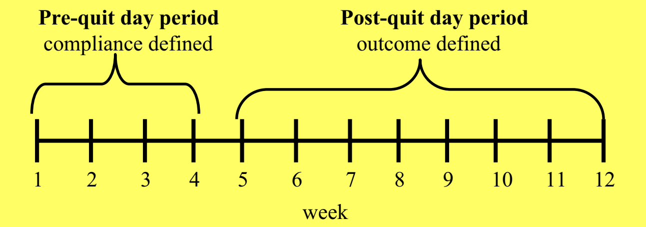
Definition of compliance and the outcome

Compliance

- $A = 1$ if attended ≥ 2 sessions every week and 3 sessions at least 2 of the weeks, during the pre-quit day period

Outcome:

- $Y = 1$ if continually abstinent



Covariates

Baseline covariates that were available included

- age, gender, race/ethnicity, education
- weight, body fat percentage, BMI
- Number of previous quit attempts, number of cigarettes per day, fagerstrom score (levels measure of dependence)

Models for stratum membership

Logistic regression model for $P(A_0 = 1|X) = P(A = 1|Z = 0, X)$:

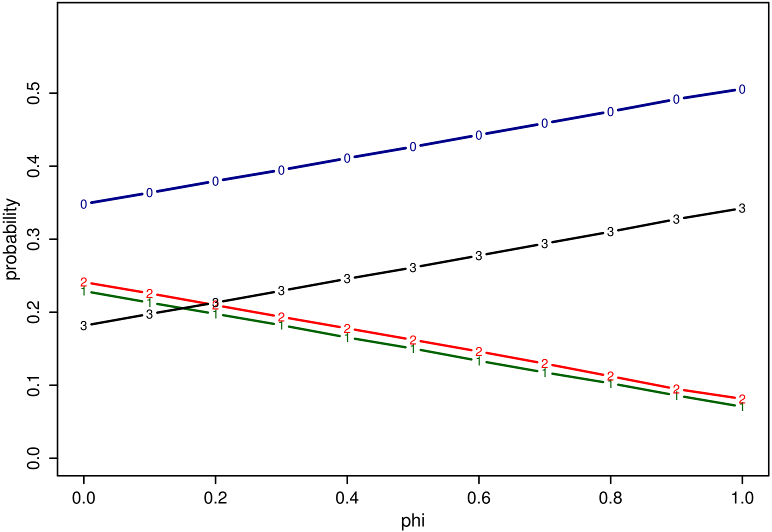
- X included `age`, `employed`, `married`, `married*employed`, `HS education` and `some college`

Logistic regression model for $P(A_1 = 1|X) = P(A = 1|Z = 1, X)$:

- X included `HS education` and `some college`

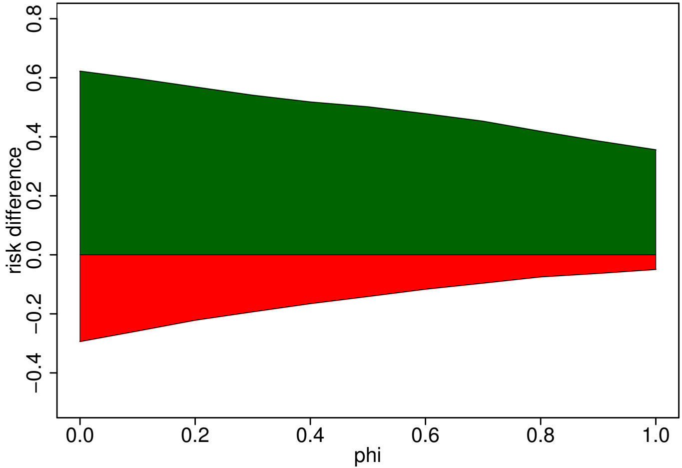
Fix value of ϕ to capture dependence between A_0 and A_1

Estimated strata probabilities as function of ϕ



Complier-average causal effect (CACE)

Effect of exercise vs wellness, among those who would accept either.



Compare to ITT effect estimate

This captures effect of *being offered* exercise vs *being offered* wellness.

The posterior median and 95% credible interval for the $E(Y_1 - Y_0)$:

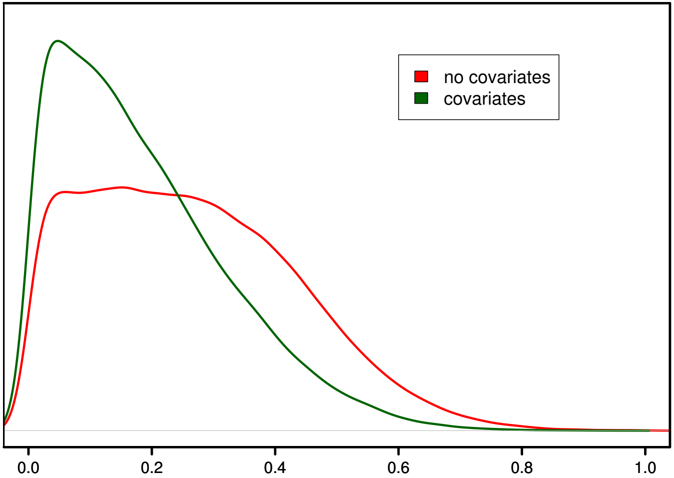
- $0.085\ (0.002, 0.170)$.

Impact of covariates

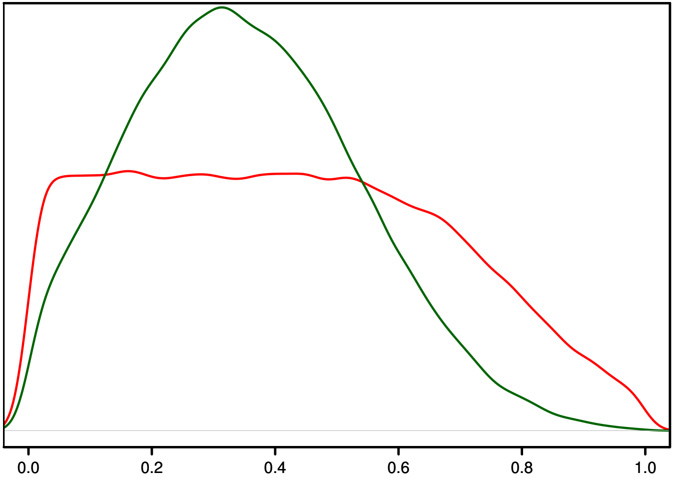
We fitted models with and without covariates, and compared the posterior distributions.

- In general, the model without covariates had wider posterior distributions.
- Covariates essentially reduced the range of plausible values that the causal parameters can take over the range of ϕ
- See plot

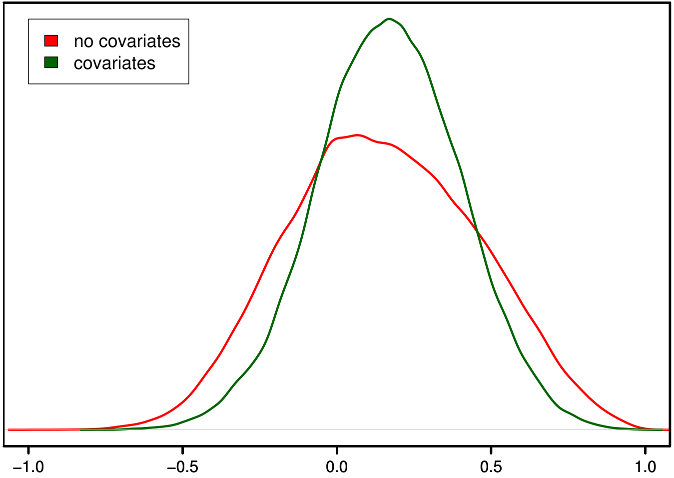
Cessation rate for compliers on wellness arm (with and without covariates)



Cessation rate for compliers on exercise arm (with and without covariates)



CACE with and without covariates



Conclusions

- Subjects randomized to CBT+exercise were more likely to quit.
- Among highly compliant subjects, exercise appeared to be beneficial.
- 43% complied with exercise
- Only 15-35% of subjects would have complied with both interventions.
- College graduates more likely to comply with exercise program.
- In simulation study and the analysis, demonstrated benefits from including covariates in the model.

Summary points

- CACE should be viewed in a larger context of principal stratification.
- It estimates causal effect among a subgroup of compliers
- PS can be used to estimate causal effect within other subgroups
- PS has been used to address the mediator problem ... relevance less clear
- Assumptions – these are critical and must be thought through carefully in context
- If using standard software, carefully evaluate assumptions