EPE - Lecture 3
Randomized Controlled Trials

Sylvain Chabé-Ferret

Toulouse School of Economics, Inra

January 2017

In a nutshell

In this lecture, we are going to study how to estimate the effect of an intervention on an outcome using randomized experiments. We are especially going to study the various types of designs and what can be recovered from them using which technique.

Key Idea

We are going to insert randomness in the allocation of the treatment. Individuals with $R_i = 1$ have a higher probability of receiving the treatment. Potential outcomes have the same distribution in both $R_i = 1$ and $R_i = 0$ groups. Comparing these groups of similar people tells us something about the causal effect of the treatment.

Key Assumption: Independence

Assumption (Independence)

We assume that the randomized allocation of the program is well done:

$$R_i \perp\!\!\!\perp (Y_i^0, Y_i^1).$$

The Designs Covered

1. Brute Force Design
2. Randomization After Self-Selection
3. Randomization After Eligibility
4. Encouragement Design

Outline

Brute Force Design: Key Idea

We randomly allocate the treatment in the overall population without any possibility for refusing the treatment or changing treatment status.

Key Assumption: Brute Force

Assumption (Brute Force)

We assume that the randomized allocation of the program is mandatory and does not interfere with how potential outcomes are generated:

$$Y_i = \begin{cases} Y_i^1 & \text{if } R_i = 1 \\ Y_i^0 & \text{if } R_i = 0 \end{cases}$$

with Y_i^1 *and* Y_i^0 *the same potential outcomes as defined in Lecture 0 with a routine allocation of the treatment.*

Identification in the Brute Force Design

Theorem (Identification in the Brute Force Design)

Under Independence and Brute Force, the WW estimator identifies the Average Effect of the Treatment (ATE):

$$\Delta^Y_{WW} = \Delta^Y_{ATE},$$

with:

$$\begin{aligned}
\Delta^Y_{WW} &= \mathbb{E}[Y_i|R_i = 1] - \mathbb{E}[Y_i|R_i = 0] \\
\Delta^Y_{ATE} &= \mathbb{E}[Y_i^1 - Y_i^0].
\end{aligned}$$

Proof

$$
\begin{aligned}
\Delta^Y_{WW} &= \mathbb{E}[Y_i | R_i = 1] - \mathbb{E}[Y_i | R_i = 0] \\
&= \mathbb{E}[Y_i^1 | R_i = 1] - \mathbb{E}[Y_i^0 | R_i = 0] \\
&= \mathbb{E}[Y_i^1] - \mathbb{E}[Y_i^0] \\
&= \mathbb{E}[Y_i^1 - Y_i^0],
\end{aligned}
$$

where the first equality uses Brute Force, the second equality Independence and the last equality the linearity of the expectation operator.

Brute Force Design: Illustration

$$R_i^* \sim \mathcal{U}[0,1]$$

$$R_i = \begin{cases} 1 & \text{if } R_i^* \leq .5 \\ 0 & \text{if } R_i^* > .5 \end{cases}$$

Brute Force Design: Illustration

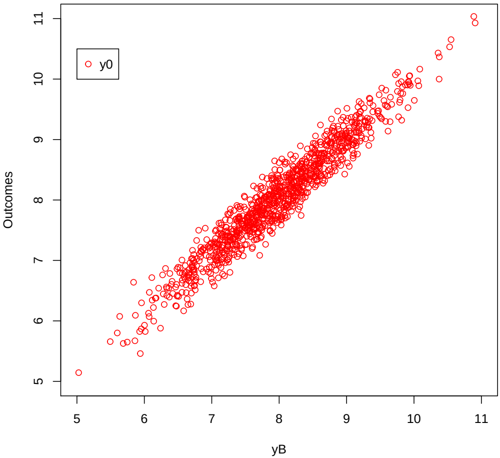

Figure: Brute Force Design

Brute Force Design: Illustration

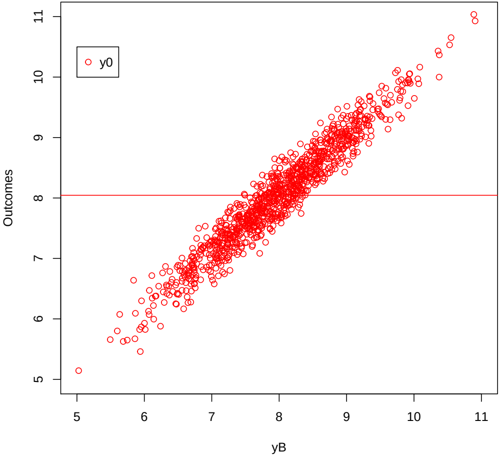

Figure: Brute Force Design

Brute Force Design: Illustration

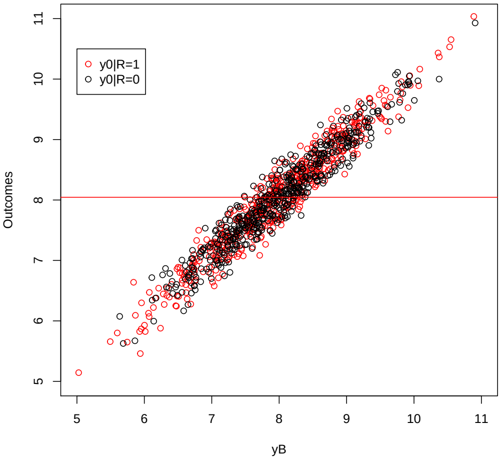

Figure: Brute Force Design

Brute Force Design: Illustration

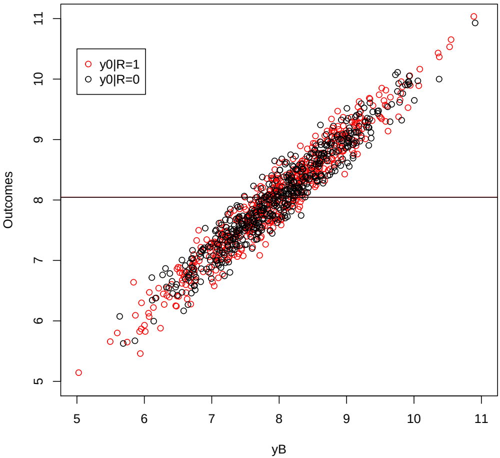

Figure: Brute Force Design

Brute Force Design: Illustration

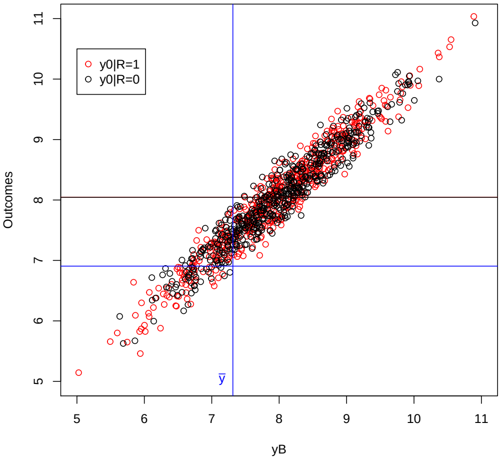

Figure: Brute Force Design

Brute Force Design: Illustration

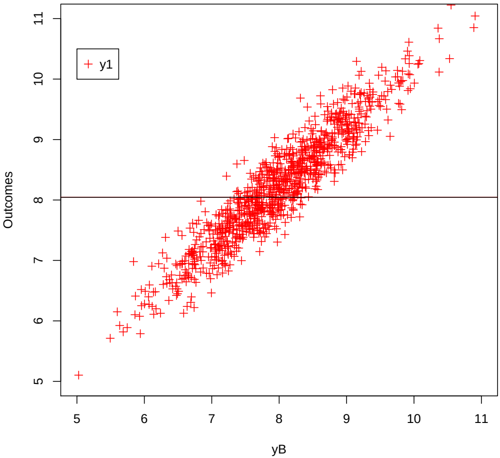

Figure: Brute Force Design

Brute Force Design: Illustration

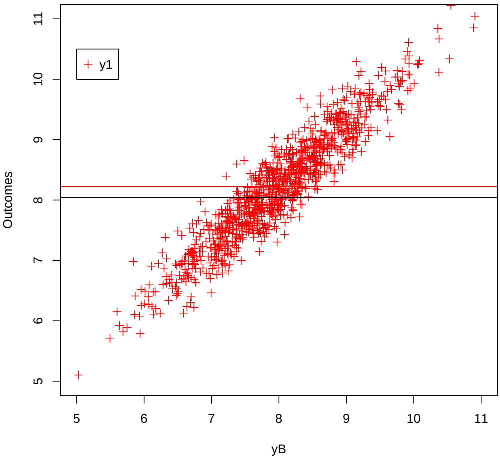

Figure: Brute Force Design

Brute Force Design: Illustration

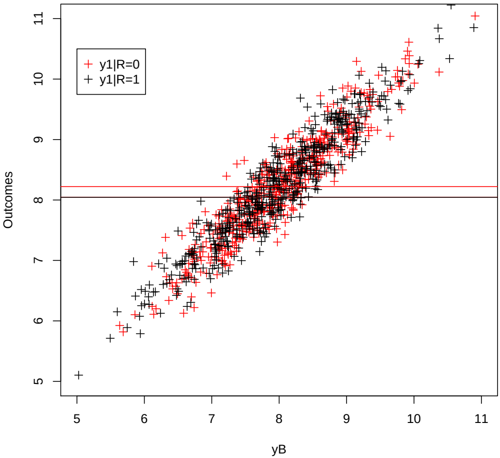

Figure: Brute Force Design

Brute Force Design: Illustration

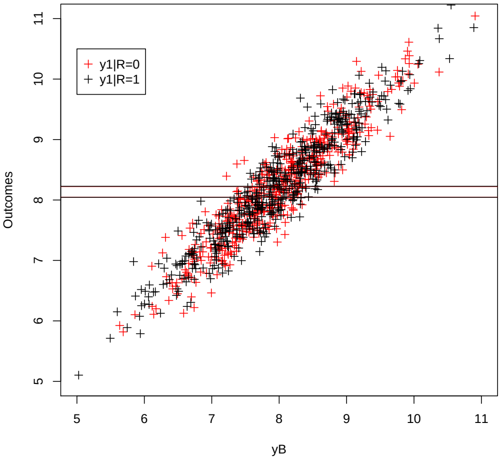

Figure: Brute Force Design

Brute Force Design: Illustration

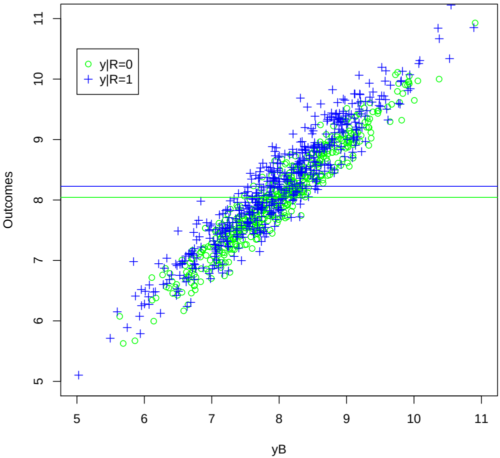

Figure: Brute Force Design

Brute Force Design: Illustration

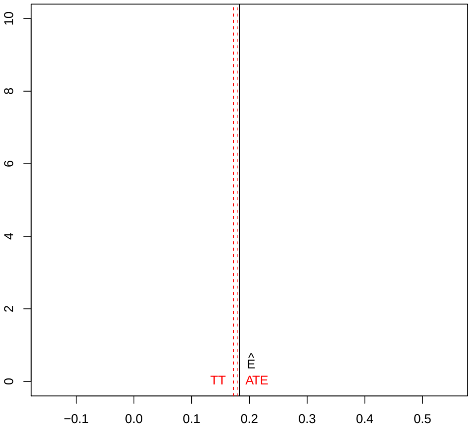

Figure: Brute Force Design

Direct Estimation Using WW

$$\hat{\Delta}^{Y}_{WW} = \frac{1}{\sum_{i=1}^{N} R_i} \sum_{i=1}^{N} Y_i R_i - \frac{1}{\sum_{i=1}^{N}(1 - R_i)} \sum_{i=1}^{N} Y_i (1 - R_i).$$

Direct Estimation Using WW: Illustration

- $\hat{\mathbb{E}}[y_i|R_i = 1] = \frac{1}{\sum R_i} \sum Y_i R_i =$8.2264295
- $\hat{\mathbb{E}}[y_i|R_i = 0] = \frac{1}{\sum (1-R_i)} \sum Y_i(1 - R_i) =$8.0436531
- $\Delta^Y_{WW} =$8.2264295$-$8.0436531$=$0.1827764

Estimation Using OLS

The OLS coefficient β in the following regression:

$$Y_i = \alpha + \beta D_i + U_i$$

is the WW estimator.

Estimation Using OLS: Illustration

```
reg.y.R.ols <- lm(y R)
```

$\hat{\Delta}^{y}_{WW}$ =0.1827764

Sampling Noise with WW in Brute Force Design: Illustration

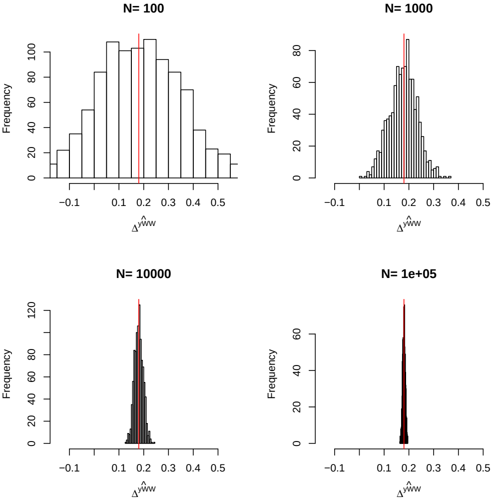

Figure: Distribution of the *WW* estimator in a Brute Force design over replications of samples of different sizes

Inference with WW In Brute Force Design

- True 99% sampling noise (from the simulations) is 0.2738818
- 99% sampling noise estimated using default OLS standard errors is 0.2942178
- 99% sampling noise estimated using heteroskedasticity robust OLS standard errors is 0.2939548

Reducing Sampling Noise By Conditioning on Covariates

Parametrically:

$$Y_i = \alpha + \beta D_i + \gamma' X_i + U_i$$

Or nonparametrically using matching.

Estimation Using OLS with Covariates: Illustration

```r
reg.y.R.yB.ols <- lm(y R + yB)
```

$\Delta^{y}_{WW} =$0.1931461

Sampling Noise with WW Conditioning on Covariates: Illustration

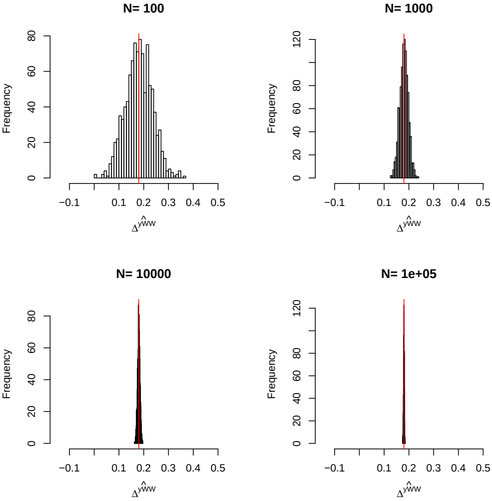

Figure: Distribution of the *WW* estimator in a Brute Force design over replications of samples of different sizes

Inference with WW Conditioning on Covariates

- True 99% sampling noise (from the simulations) is 0.0881573
- 99% sampling noise estimated using default OLS standard errors is 0.0904044
- 99% sampling noise estimated using heteroskedasticity robust OLS standard errors is 0.0902479

Conditioning on Covariates: Warning

Mention which covariates you are going to use in your preanalysis plan (should be part of the power study)

Outline

Randomization After Self-Selection: Key Idea

We randomly allocate the treatment after individuals have expressed their interest in it by self-selecting to receive it.

Self-Selection: New Notation

We are going to have two steps for receiving the treatment: Eligibility (E_i) and self-selection (D_i):

$$
\begin{aligned}
E_i &= \mathbb{1}[y_i^B \leq \bar{y}] \\
D_i &= \mathbb{1}[\underbrace{\bar{\alpha} + \theta\bar{\mu} - C_i}_{D_i^*} \geq 0 \wedge E_i = 1] \\
C_i &= \bar{c} + \gamma\mu_i + V_i \\
V_i &\sim \mathcal{N}(0, \sigma_V^2)
\end{aligned}
$$

Randomization After Self-Selection: Illustration

$$R_i^* \sim \mathcal{U}[0,1]$$

$$R_i = \begin{cases} 1 & \text{if } R_i^* \leq .5 \wedge D_i = 1 \\ 0 & \text{if } R_i^* > .5 \wedge D_i = 1 \end{cases}$$

Randomization After Self-Selection: Illustration

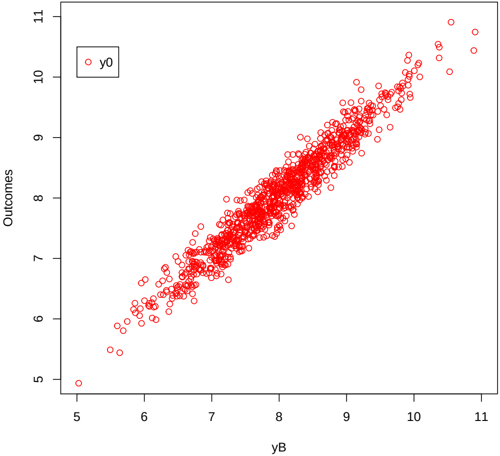

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

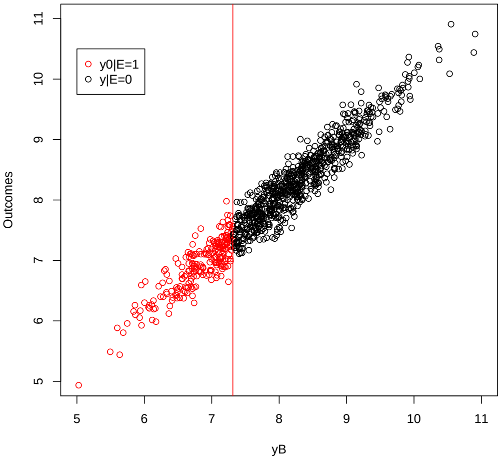

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

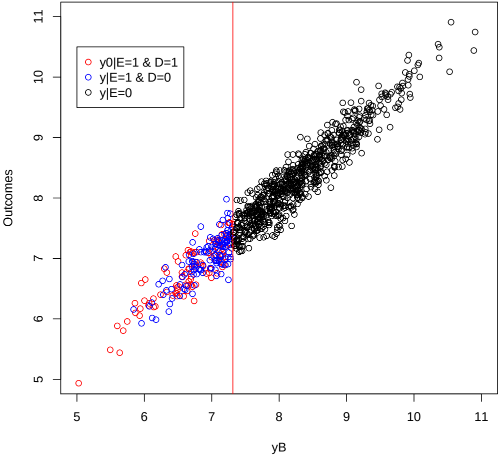

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

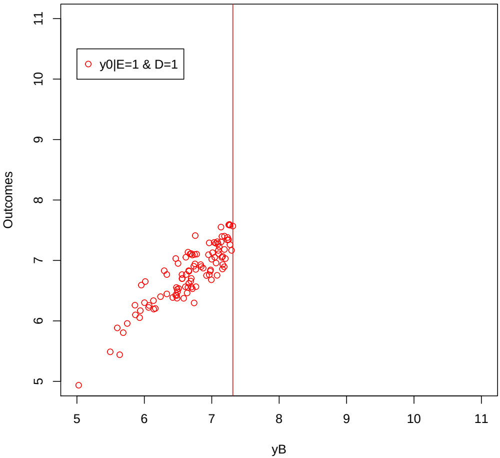

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

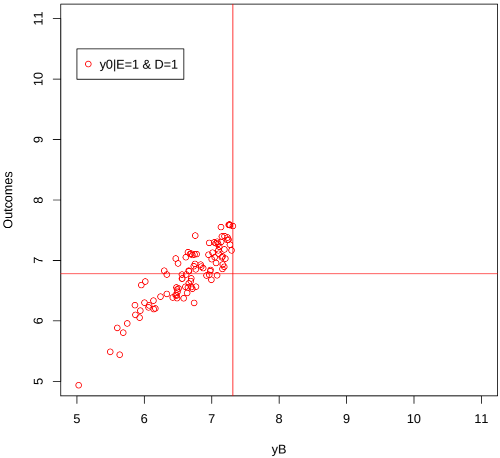

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

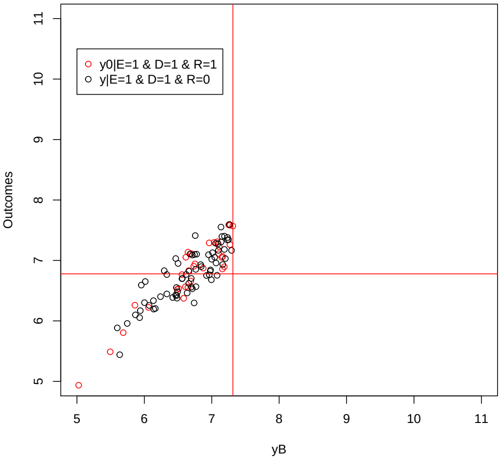

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

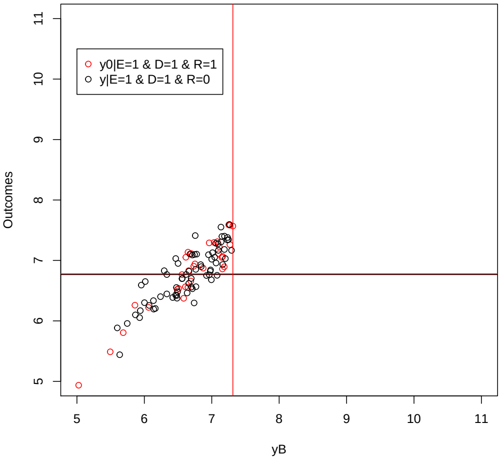

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

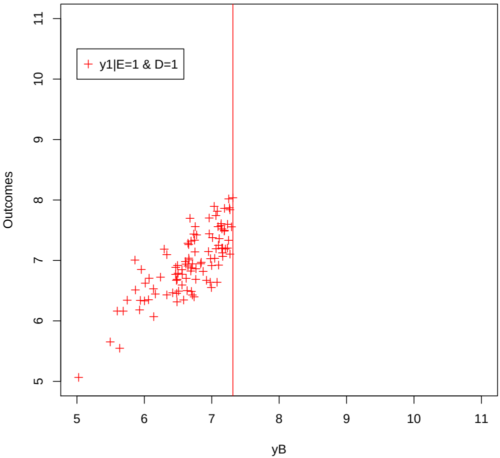

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

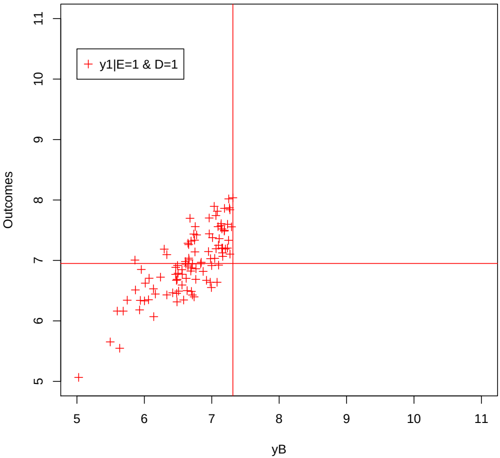

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

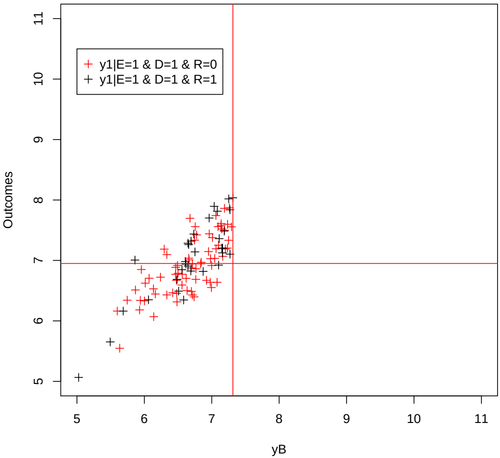

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

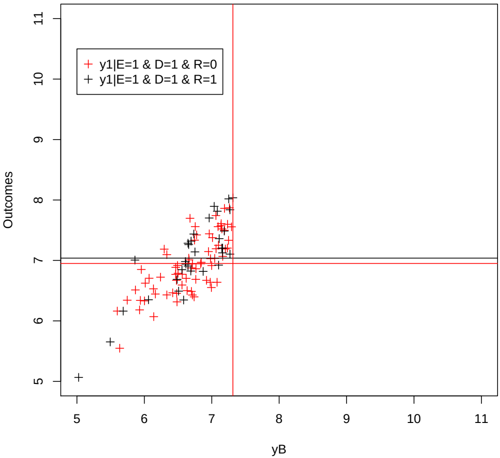

Figure: Randomization After Self-Selection

Randomization After Self-Selection: Illustration

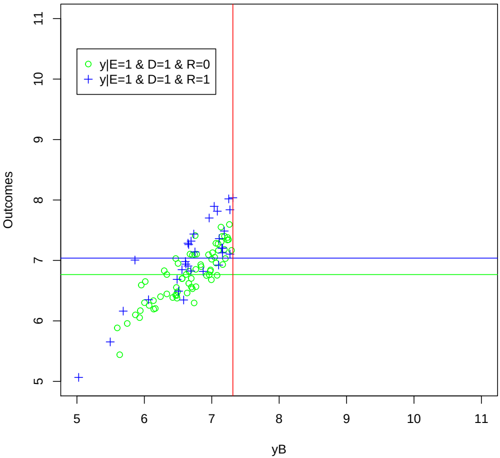

Figure: Randomization After Self-Selection

The Value of TT in our Illustration

$$\Delta^y_{TT} = \bar{\alpha} + \theta\mathbb{E}[\mu_i|\mu_i + U_i^B \leq \bar{y} \wedge \bar{\alpha} + \theta\bar{\mu} - \bar{c} - \gamma\mu_i - V_i \geq 0]$$

To compute the expectation of a doubly censored normal, I use the package tmvtnorm.

$$(\mu_i, y_i^B, D_i^*) = \mathcal{N}\left(\bar{\mu}, \bar{\mu}, \bar{\alpha} + (\theta - \gamma)\bar{\mu} - \bar{c}, \left(\begin{array}{ccc} \sigma^2_\mu & \sigma^2_\mu & -\gamma\sigma^2_\mu \\ \sigma^2_\mu & \sigma^2_\mu + \sigma^2_U & -\gamma\sigma^2_\mu \\ -\gamma\sigma^2_\mu & -\gamma\sigma^2_\mu & \gamma^2\sigma^2_\mu + \sigma^2_V \end{array}\right)\right)$$

The value of Δ^y_{TT} in our illustration is: 0.1700198.

Randomization After Self-Selection: Illustration

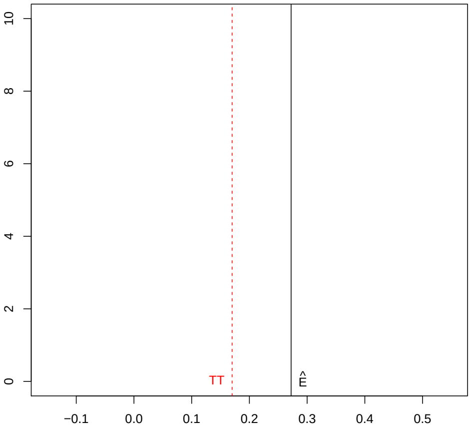

Figure: Randomization After Self-Selection

Key Assumption: Independence Among Self-Selected

Assumption (Independence Among Self-Selected)

We assume that the randomized allocation of the program is well done:

$$R_i \perp\!\!\!\perp (Y_i^0, Y_i^1) | D_i = 1.$$

Key Assumption: No Effect of Randomization

Assumption (No Effect of Randomization)

We assume that the randomized allocation of the program does not alter self-selection into the program and potential outcomes:

$$Y_i = \begin{cases} Y_i^1 & \text{if } R_i = 1 \\ Y_i^0 & \text{if } R_i = 0 \end{cases}$$

with Y_i^1 and Y_i^0 the same potential outcomes as defined in Lecture 0 with a routine allocation of the treatment.

Identification With Randomization After Self-Selection

Theorem (Identification With Randomization After Self-Selection)

Under Independence Among Self-Selected and No Effect of Randomization, the WW estimator among the self-selected identifies TT:

$$\Delta^Y_{WW|D=1} = \Delta^Y_{TT},$$

with:

$$\Delta^Y_{WW|D=1} = \mathbb{E}[Y_i|R_i = 1, D_i = 1] - \mathbb{E}[Y_i|R_i = 0, D_i = 1].$$

Proof

$$
\begin{aligned}
\Delta^Y_{WW|D=1} &= \mathbb{E}[Y_i|R_i = 1, D_i = 1] - \mathbb{E}[Y_i|R_i = 0, D_i = 1] \\
&= \mathbb{E}[Y_i^1|R_i = 1, D_i = 1] - \mathbb{E}[Y_i^0|R_i = 0, D_i = 1] \\
&= \mathbb{E}[Y_i^1|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 1] \\
&= \mathbb{E}[Y_i^1 - Y_i^0|D_i = 1],
\end{aligned}
$$

where the second equality uses No Effect of Randomization, the third equality Independence Among Self-Selected and the last equality the linearity of the expectation operator.

Direct Estimation Using WW

$$\hat{\Delta}^{Y}_{WW|D=1} = \frac{1}{\sum_{i=1}^{N} D_i R_i} \sum_{i=1}^{N} Y_i D_i R_i - \frac{1}{\sum_{i=1}^{N} D_i (1 - R_i)} \sum_{i=1}^{N} D_i Y_i (1 - R_i)$$

Direct Estimation Using WW: Illustration

- $\hat{\mathbb{E}}[y_i|R_i = 1, D_i = 1] = \frac{1}{\sum D_i R_i} \sum Y_i D_i R_i =$7.0387098
- $\hat{\mathbb{E}}[y_i|R_i = 0] = \frac{1}{\sum D_i(1-R_i)} \sum Y_i D_i (1 - R_i) =$6.766485
- $\hat{\Delta}^y_{WW|D=1} =7.0387098-6.766485=$0.2722249

Estimation Using OLS

The OLS coefficient β in the following regression:

$$Y_i = \alpha + \beta D_i + U_i$$

estimated on the sample of self-selected individuals $(D_i = 1)$ is the WW estimator.

Estimation Using OLS: Illustration

```
reg.y.R.ols.self.select <- lm(y[Ds==1] R[Ds==1])
```

$\hat{\Delta}^{y}_{WW|D=1}$ =0.2722249

Sampling Noise with WW in Randomization After Self-Selection: Illustration

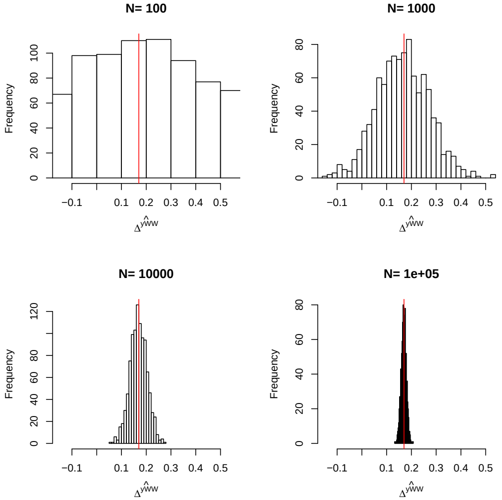

Figure: Distribution of the *WW* estimator with randomization after self-selection over replications of samples of different sizes

Inference with WW in Randomization After Self-Selection

- True 99% sampling noise (from the simulations) is 0.5478831
- 99% sampling noise estimated using default OLS standard errors is 0.5655959
- 99% sampling noise estimated using heteroskedasticity robust OLS standard errors is 0.6532778

Reducing Sampling Noise By Conditioning on Covariates

Parametrically:

$$Y_i = \alpha + \beta D_i + \gamma' X_i + U_i$$

Or nonparametrically using matching.

Estimation Using OLS with Covariates: Illustration

```
reg.y.R.yB.ols.self.select <- lm(y[Ds==1]   R[Ds==1] +
yB[Ds==1])
```

$\hat{\Delta}^{y}_{WW(X)|D=1}=$0.2415448

Sampling Noise with WW in Randomization After Self-Selection: Illustration

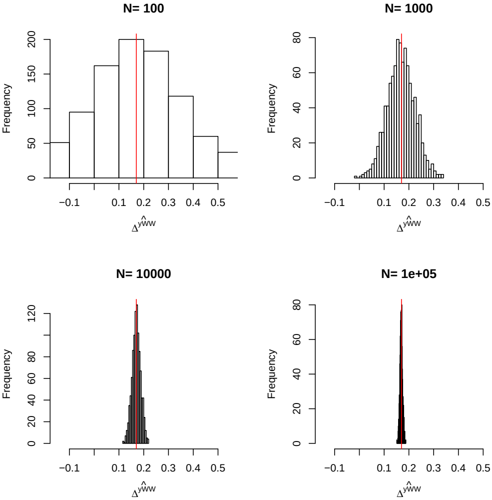

Figure: Distribution of the *WW* estimator with randomization after self-selection over replications of samples of different sizes

Inference with WW in Randomization After Self-Selection

- True 99% sampling noise (from the simulations) is 0.2947862
- 99% sampling noise estimated using default OLS standard errors is 0.3016481
- 99% sampling noise estimated using heteroskedasticity robust OLS standard errors is 0.3509391

Conditioning on Covariates: Warning

- Mention which covariates you are going to use in your preanalysis plan (should be part of the power study)
- If feasible, try to stratify your randomization on these variables

Outline

Randomization After Eligibility: Key Idea

We randomly allocate the treatment after eligibility has been determined but before individuals self-select to receive it. Among the eligibles, only the randomized-in have an opportunity to self-select into the treatment.

Randomization After Eligibilty: Illustration

$$R_i^* \sim \mathcal{U}[0,1]$$

$$R_i = \begin{cases} 1 & \text{if } R_i^* \leq .5 \wedge E_i = 1 \\ 0 & \text{if } R_i^* > .5 \wedge E_i = 1 \end{cases}$$

$$D_i = \mathbb{1}[\bar{\alpha} + \theta\bar{\mu} - C_i \geq 0 \wedge E_i = 1 \wedge R_i = 1]$$

Randomization After Eligibility: Illustration

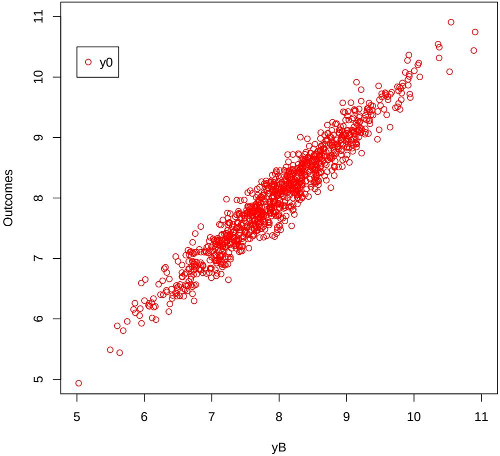

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

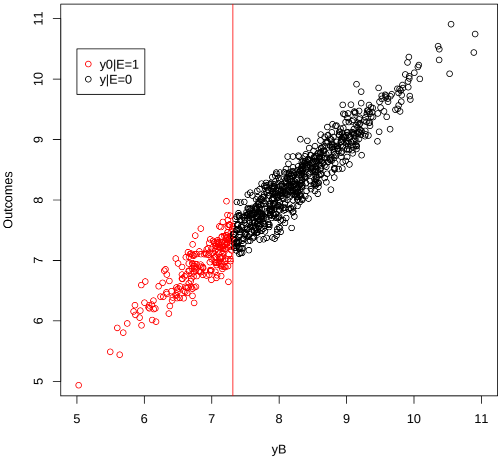

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

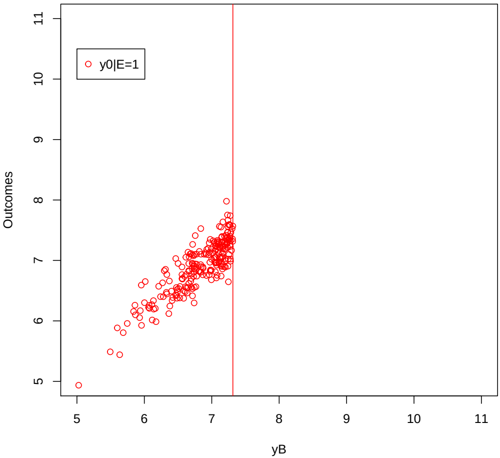

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

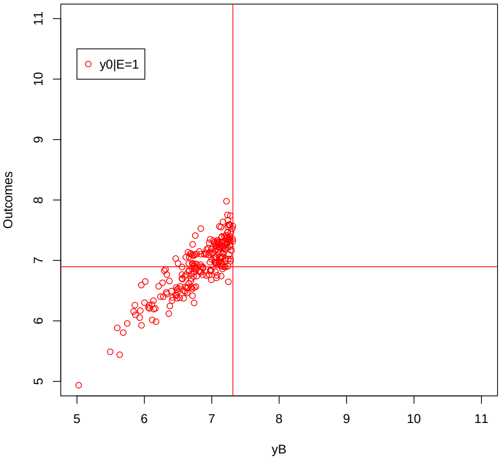

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

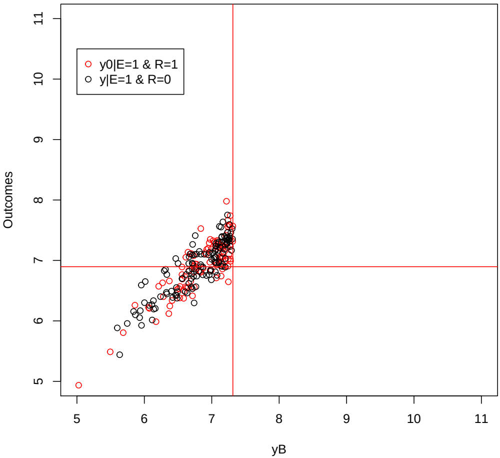

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

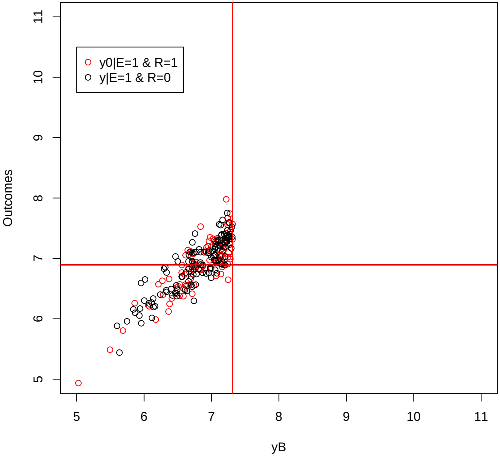

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

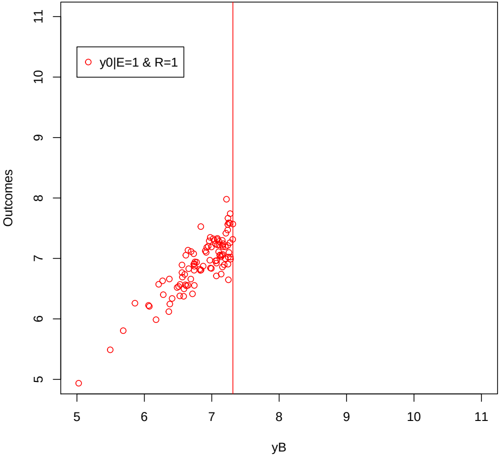

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

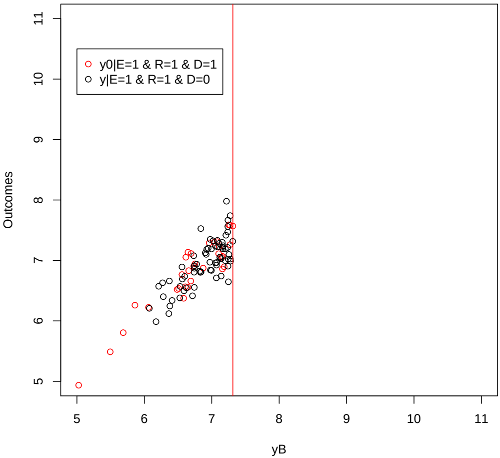

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

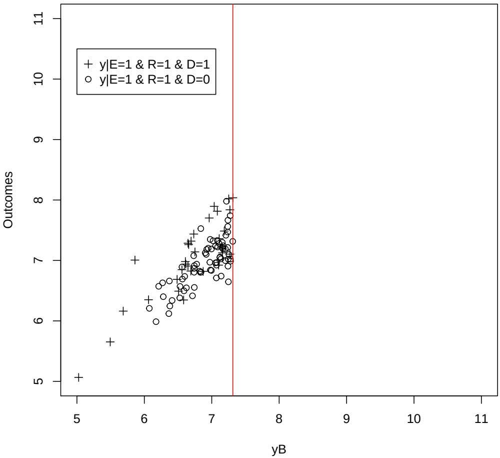

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

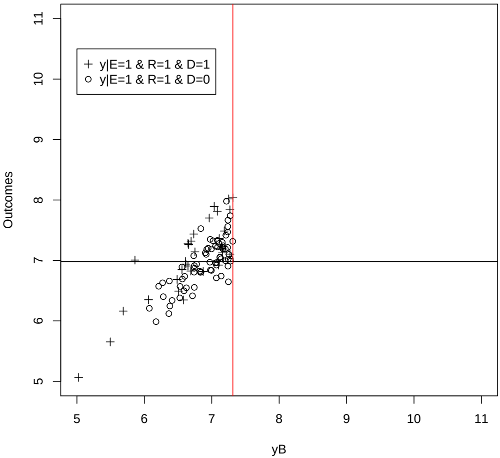

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

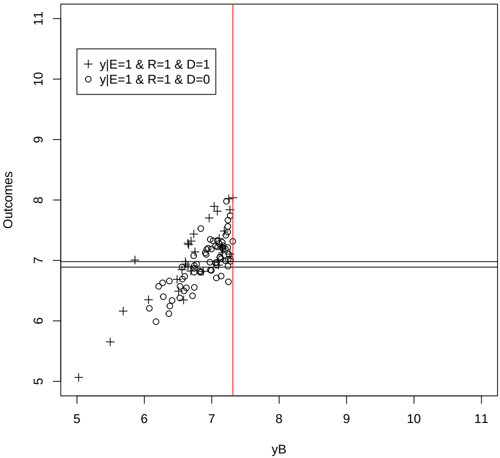

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

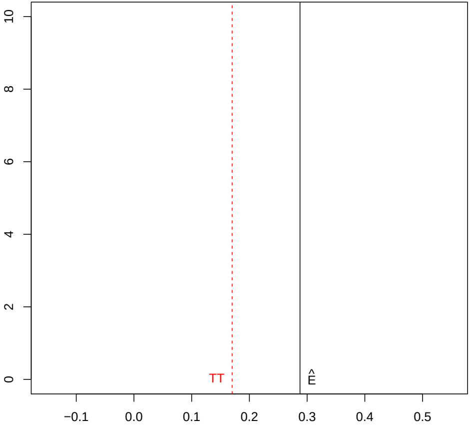

Figure: Randomization After Self-Selection

Key Assumption: Independence Among Eligibles

Assumption (Independence Among Eligibles)

We assume that the randomized allocation of the program is well done:

$$R_i \perp\!\!\!\perp (Y_i^0, Y_i^1) | E_i = 1.$$

Key Assumption: No Effect of Randomization

Assumption (No Effect of Randomization)

We assume that the randomized allocation of the program does not alter self-selection into the program and potential outcomes:

$$Y_i = \begin{cases} Y_i^1 & \text{if } D_i = 1 \\ Y_i^0 & \text{if } D_i = 0 \end{cases}$$

with Y_i^1 and Y_i^0 the same potential outcomes as defined in Lecture 0 with a routine allocation of the treatment.

Identification With Randomization After Eligibility

Theorem (Identification With Randomization After Eligibility)

Under Independence Among Eligibles and No Effect of Randomization, the Bloom estimator identifies TT:

$$\Delta^Y_{Bloom} = \Delta^Y_{TT},$$

with:

$$\Delta^Y_{Bloom} = \frac{\mathbb{E}[Y_i|R_i = 1, E_i = 1] - \mathbb{E}[Y_i|R_i = 0, E_i = 1]}{\Pr(D_i = 1|R_i = 1, E_i = 1)}.$$

Proof

Let's define D_i^* that takes value 1 is the individual self-selects in the program in routine mode and 0 otherwise. When $R_i = 1$, $D_i^* = D_i$. When $R_i = 0$, D_i^* is unobserved. The numerator of the Bloom estimator is:

$$\begin{aligned}
\mathbb{E}[Y_i|R_i = 1, E_i = 1] - \mathbb{E}[Y_i|R_i = 0, E_i = 1]& \\
&= \mathbb{E}[Y_i^1 D_i + Y_i^0(1 - D_i)|R_i = 1, E_i = 1] - \mathbb{E}[Y_i^0 D_i^* + Y_i^0(1 - D_i^*)|R_i = 0, E_i = 1] \\
&= \mathbb{E}[Y_i^1|D_i = 1]\Pr(D_i = 1|R_i = 1, E_i = 1) + \mathbb{E}[Y_i^0|D_i = 0]\Pr(D_i = 0|R_i = 1, E_i = 1) \\
&\quad - \mathbb{E}[Y_i^0|D_i^* = 1]\Pr(D_i^* = 1|R_i = 0, E_i = 1) + \mathbb{E}[Y_i^0|D_i^* = 0]\Pr(D_i^* = 0|R_i = 0, E_i = 1)
\end{aligned}$$

where the second equality uses Independence Among Eligibles. Under No Effect of Randomization, we have that the behavior in routine mode is identical to the behavior in the presence of randomization. We thus have $\mathbb{E}[Y_i^0|D_i^* = d] = \mathbb{E}[Y_i^0|D_i = d]$ and $\Pr(D_i^* = d|R_i = 0, E_i = 1) = \Pr(D_i = d|R_i = 1, E_i = 1)$. As a consequence, the numerator of the Bloom estimator is:

$$\mathbb{E}[Y_i|R_i = 1, E_i = 1] - \mathbb{E}[Y_i|R_i = 0, E_i = 1] = \mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]\Pr(D_i = 1|R_i = 1, E_i = 1).$$

Randomization After Eligibility: Illustration

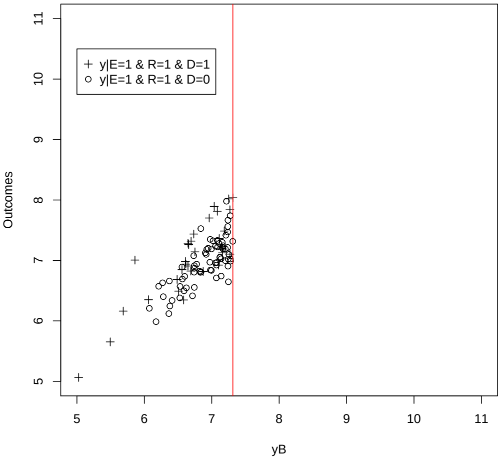

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

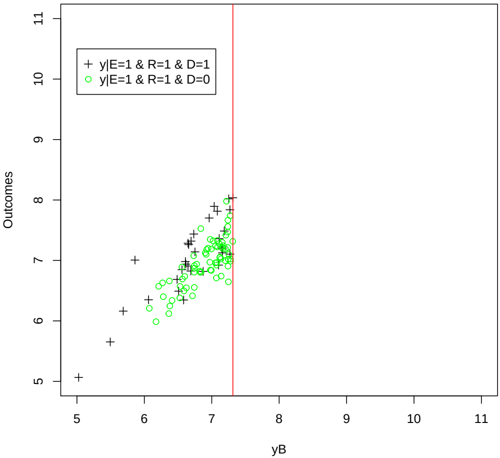

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

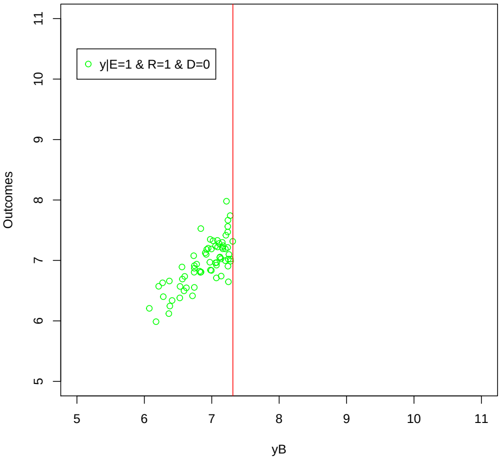

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

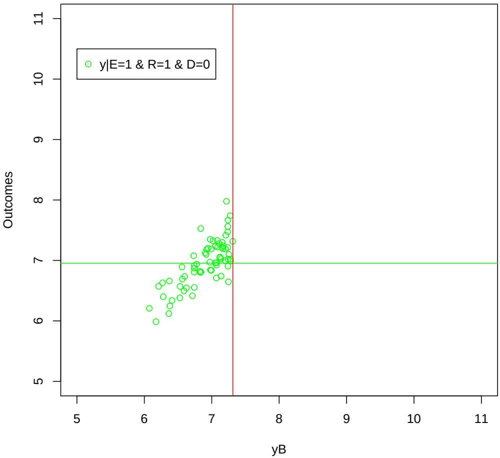

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

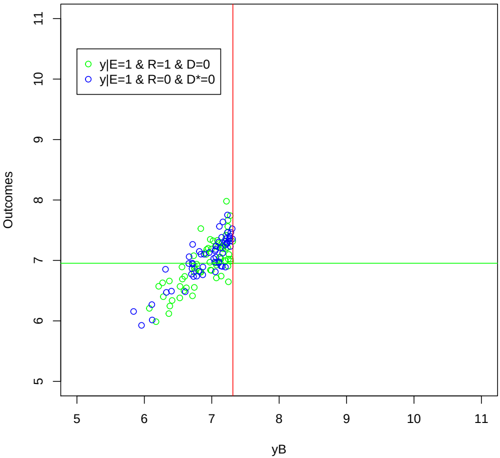

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

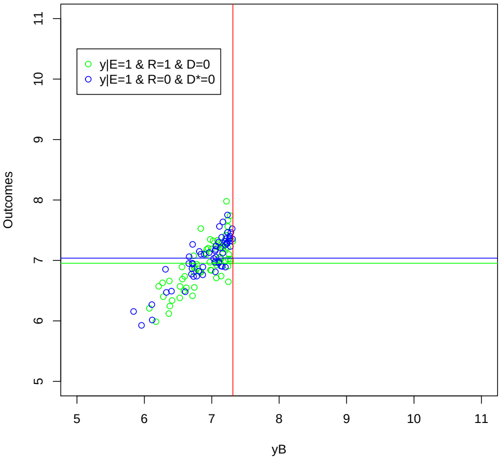

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

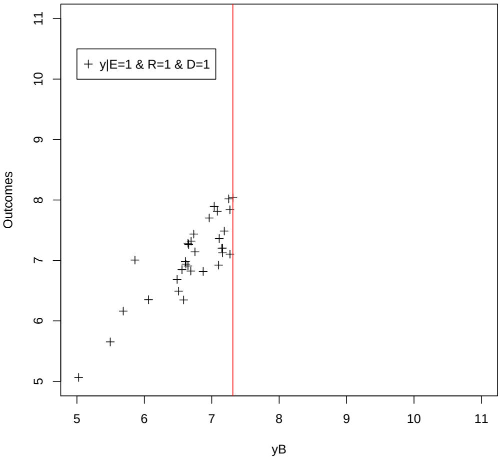

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

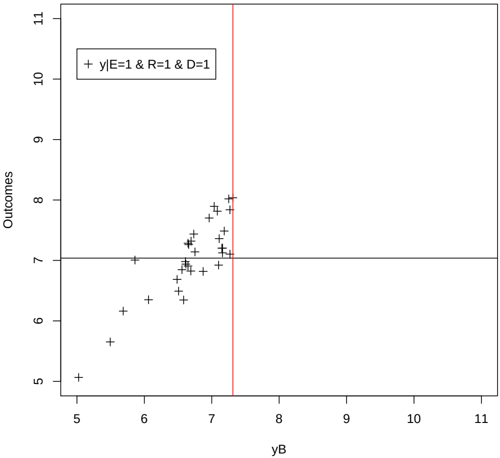

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

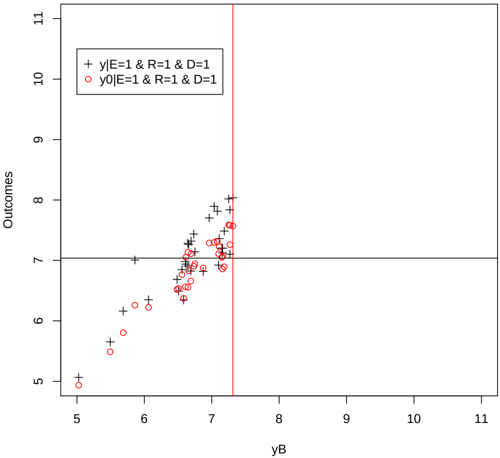

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

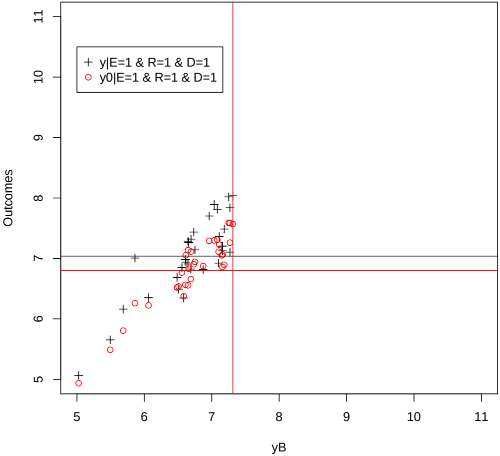

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

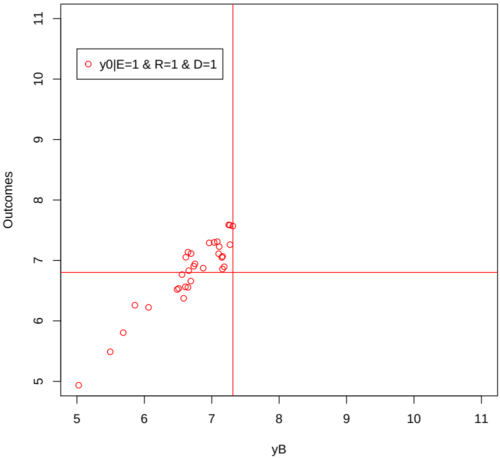

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

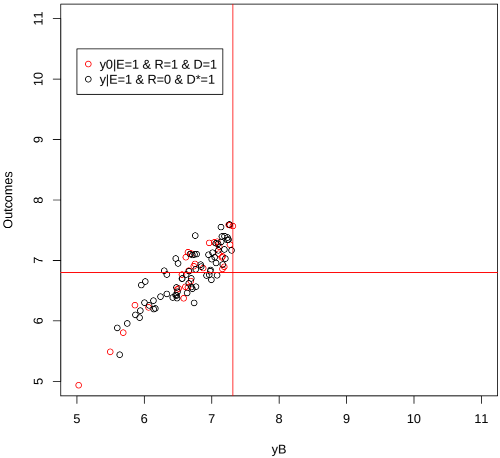

Figure: Randomization After Eligibility

Randomization After Eligibility: Illustration

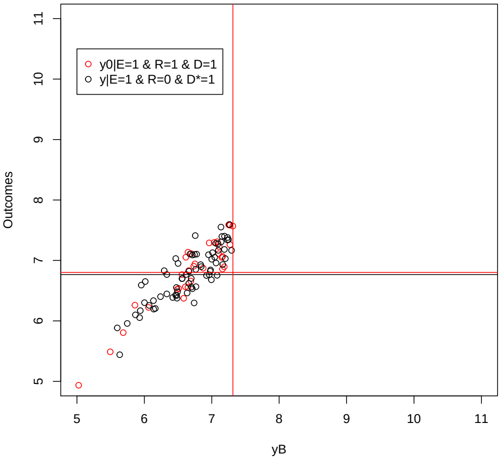

Figure: Randomization After Eligibility

Remarks

- The Bloom estimator is a special case of the Wald estimator where $\Pr(D_i = 1 | R_i = 0, E_i = 1) = 0$, that is without always takers.
- R_i act as an instrument for D_i
- The assumption of No Effect of Randomization is equivalent to an exclusion restriction

Direct Estimation Using Bloom

$$\hat{\Delta}^{Y}_{WW|D=1} = \frac{\frac{1}{\sum_{i=1}^{N} E_i R_i}\sum_{i=1}^{N} Y_i E_i R_i - \frac{1}{\sum_{i=1}^{N} E_i (1-R_i)}\sum_{i=1}^{N} E_i Y_i (1-R_i)}{\frac{1}{\sum_{i=1}^{N} E_i R_i}\sum_{i=1}^{N} D_i E_i R_i}.$$

Direct Estimation Using Bloom: Illustration

- $\hat{\mathbb{E}}[y_i|R_i = 1, E_i = 1] = \frac{1}{\sum E_i R_i} \sum Y_i E_i R_i =$6.9808686
- $\hat{\mathbb{E}}[y_i|R_i = 0, E_i = 1] = \frac{1}{\sum E_i (1-R_i)} \sum Y_i E_i (1 - R_i) =$6.8897305
- $\hat{\Pr}(D_i = 1|R_i = 1, E_i = 1) = \frac{1}{\sum_{i=1}^{N} E_i R_i} \sum_{i=1}^{N} D_i E_i R_i =$0.3168317
- $\hat{\Delta}^{y}_{Bloom} =$ (6.9808686−6.8897305)÷0.3168317=0.0911381÷0.3168317=0.2876

Estimation Using 2SLS

The 2SLS coefficient β in the following regression:

$$Y_i = \alpha + \beta D_i + U_i$$

with R_i as an instrument for D_i estimated on the sample of eligible individuals ($E_i = 1$) is the Bloom estimator.

Estimation Using 2SLS: Illustration

```
reg.y.R.2sls.elig <- ivreg(y[E==1] Ds[E==1]|R[E==1])
```

$\hat{\Delta}^{y}_{Bloom}$ =0.2876545

Sampling Noise with Bloom in Randomization After Eligibility: Illustration

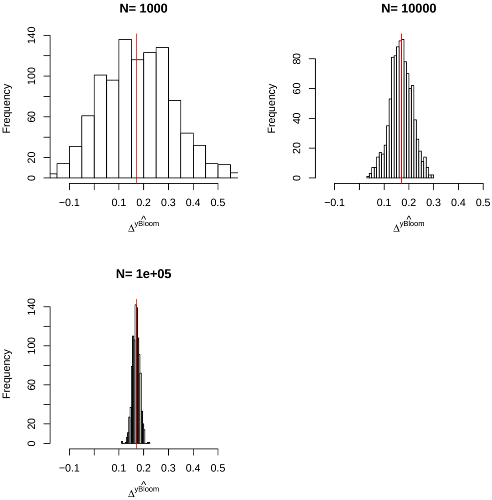

Figure: Distribution of the Bloom estimator with randomization after eligibility over replications of samples of different sizes

Inference with Bloom in Randomization After Eligibility

- True 99% sampling noise (from the simulations) is 0.7573206
- 99% sampling noise estimated using default OLS standard errors is 0.9990692
- 99% sampling noise estimated using heteroskedasticity robust OLS standard errors is 1.0174096

Reducing Sampling Noise By Conditioning on Covariates

Parametrically estimating the following equation with R_i and X_i as instruments:

$$Y_i = \alpha + \beta D_i + \gamma' X_i + U_i$$

Or nonparametrically using Frolich's Wald matching estimator.

Estimation Using 2SLS with Covariates: Illustration

```
reg.y.R.yB.2sls.elig <- ivreg(y[E==1]   Ds[E==1] + yB[E==1]
| R[E==1] + yB[E==1])
```

$\hat{\Delta}^{y}_{Bloom(X)} = 0.1328474$

Sampling Noise with Bloom in Randomization After Eligibility: Illustration

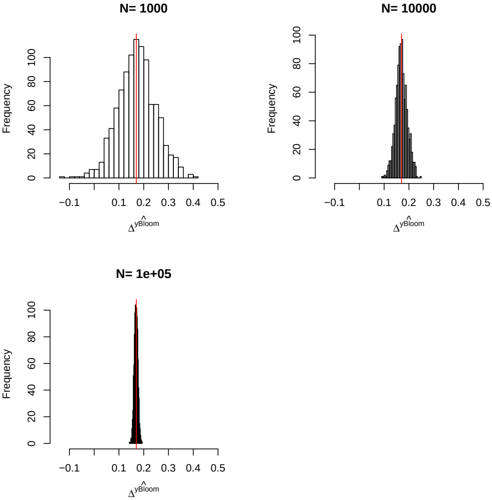

Figure: Distribution of the Bloom estimator with randomization after eligibility over replications of samples of different sizes

Inference with Bloom with Covariates in Randomization After Self-Selection

- True 99% sampling noise (from the simulations) is 0.3933989
- 99% sampling noise estimated using default OLS standard errors is 0.534024
- 99% sampling noise estimated using heteroskedasticity robust OLS standard errors is 0.5527875

Outline

Brute Force Design

Randomization After Self-Selection

Randomization After Eligibility

Encouragement Design

Exercises

Encouragement Design: Key Idea

We randomly allocate an encouragement to take the treatment (before or after eligibility). Among the eligibles, the randomized-in will be more numerous to self-select into the treatment. We can thus recover the effect of the treatment on these compliers.

Encouragement Design: Illustration

$$
\begin{aligned}
R_i^* &\sim \mathcal{U}[0,1] \\
R_i &= \begin{cases} 1 & \text{if } R_i^* \leq .5 \wedge E_i = 1 \\ 0 & \text{if } R_i^* > .5 \wedge E_i = 1 \end{cases} \\
D_i &= \mathbb{1}[\bar{\alpha} + \theta\bar{\mu} + \psi R_i - C_i \geq 0 \wedge E_i = 1]
\end{aligned}
$$

Encouragement Design: Illustration

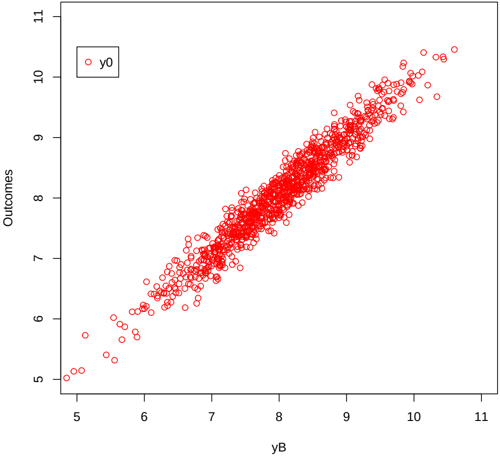

Figure: Encouragement Design

Encouragement Design: Illustration

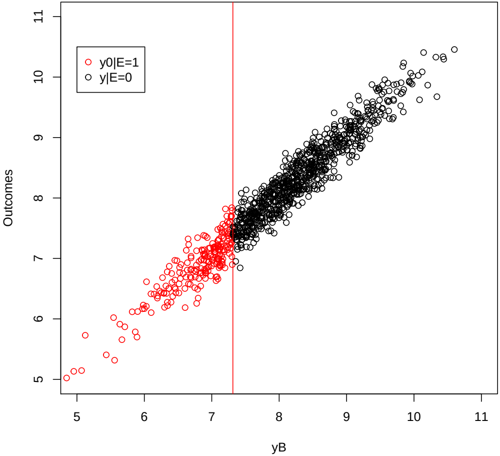

Figure: Encouragement Design

Encouragement Design: Illustration

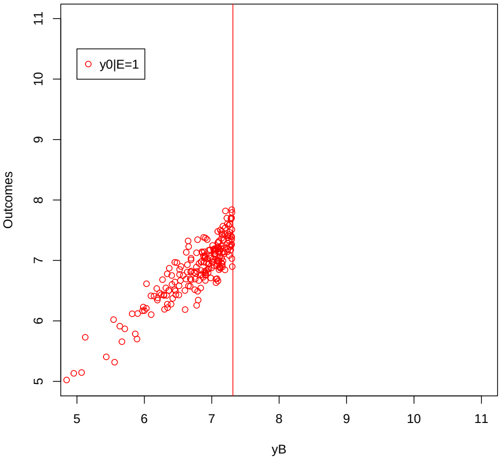

Figure: Encouragement Design

Encouragement Design: Illustration

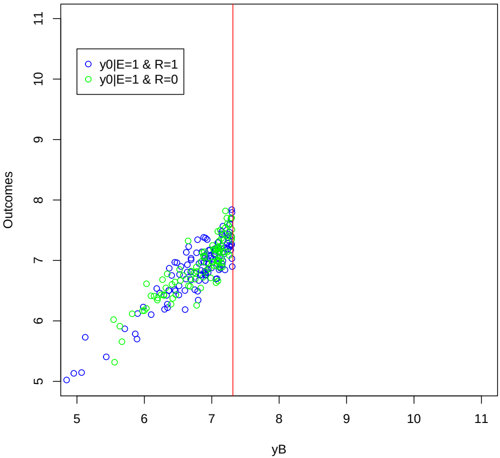

Figure: Encouragement Design

Encouragement Design: Illustration

- $\hat{\Pr}(D_i = 1 | R_i = 0, E_i = 1) =$0.4631579
- $\hat{\Pr}(D_i = 1 | R_i = 1, E_i = 1) =$0.75

Encouragement Design: Illustration

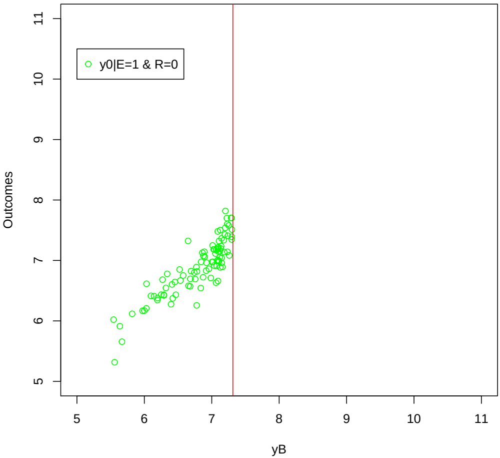

Figure: Encouragement Design

Encouragement Design: Illustration

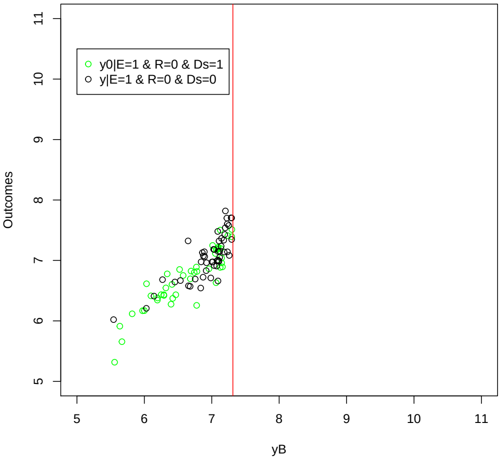

Figure: Encouragement Design

Encouragement Design: Illustration

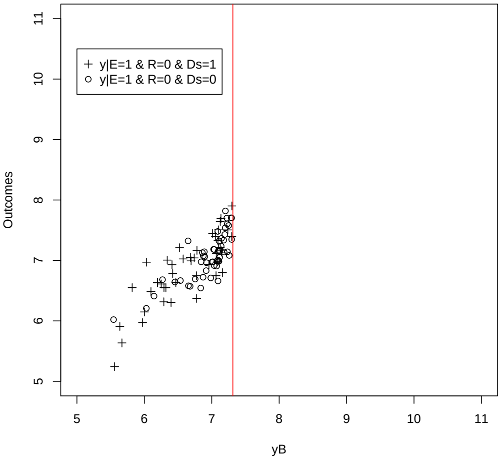

Figure: Encouragement Design

Encouragement Design: Illustration

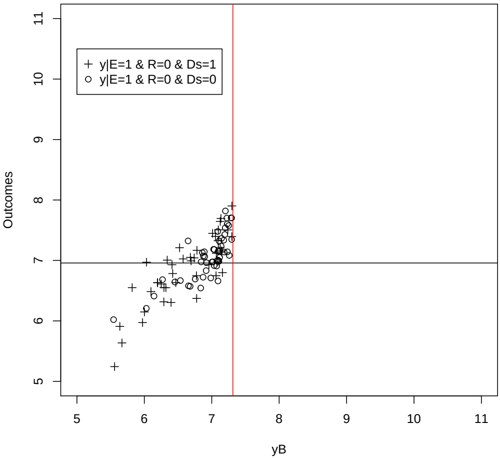

Figure: Encouragement Design

Encouragement Design: Illustration

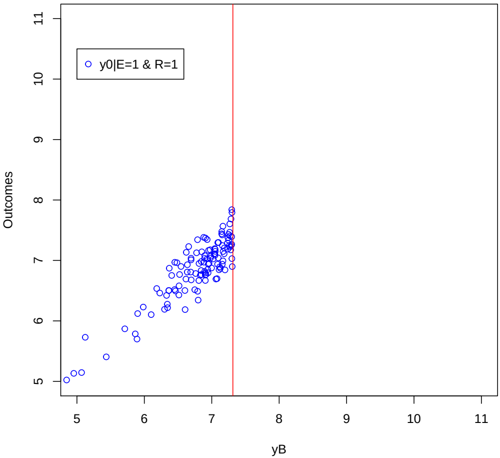

Figure: Encouragement Design

Encouragement Design: Illustration

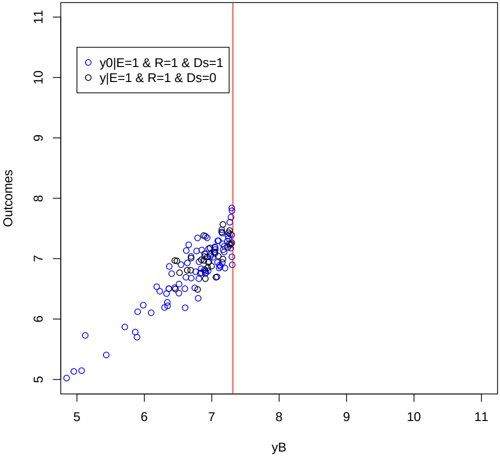

Figure: Encouragement Design

Encouragement Design: Illustration

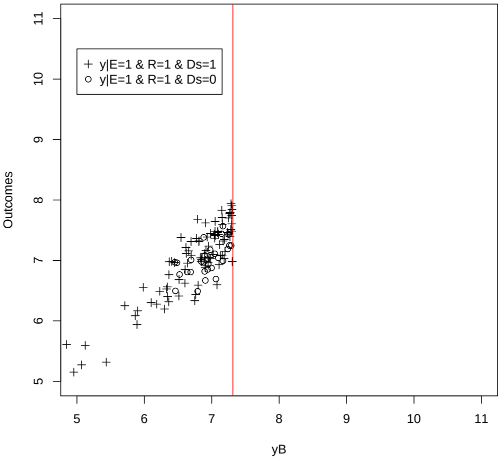

Figure: Encouragement Design

Encouragement Design: Illustration

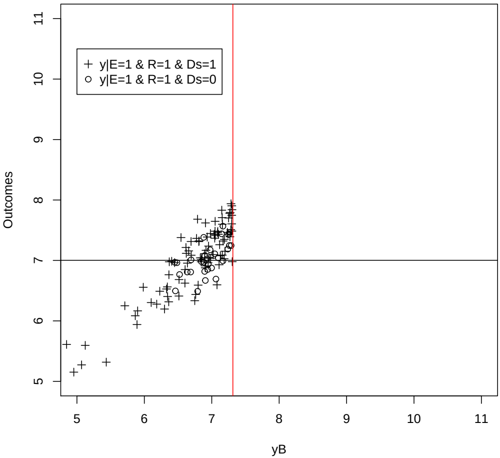

Figure: Encouragement Design

Encouragement Design: Illustration

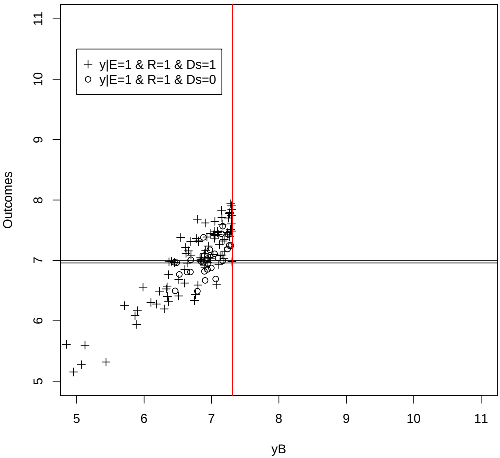

Figure: Encouragement Design

Encouragement Design: Illustration

- $\hat{\Pr}(D_i = 1 | R_i = 0, E_i = 1) =$0.4631579
- $\hat{\Pr}(D_i = 1 | R_i = 1, E_i = 1) =$0.75
- $\hat{\mathbb{E}}[y_i | R_i = 0, E_i = 1] =$6.9578739
- $\hat{\mathbb{E}}[y_i | R_i = 1, E_i = 1] =$7.0022174

Key Assumption: First Stage

Assumption (First Stage)

We assume that the encouragement does manage to increase participation:

$$\Pr(D_i = 1 | R_i = 1, E_i = 1) > \Pr(D_i = 1 | R_i = 0, E_i = 1).$$

Key Assumption: Monotonicity

Assumption (Monotonicity)

We assume that the encouragement increase participation for everyone:

$$\forall i,\ D_i^1 \geq D_i^0.$$

Key Assumption: Exclusion Restriction

Assumption (Exclusion Restriction)

We assume that the randomized allocation of the program does not alter potential outcomes:

$$Y_i^{dr} = Y_i^{d}, \forall r, d \in \{0, 1\}.$$

Key Assumption: Independence

Assumption (Exclusion Restriction)

We assume that the randomized allocation of the program is well done:

$$(Y_i^1, Y_i^0, D_i^1, D_i^0) \perp\!\!\!\perp R_i | E_i = 1.$$

Identification in an Encouragement Design

Theorem (Identification in an Encouragement Design)

Under First Stage, Monotonicity, Exclusion Restriction and Independence, the Wald estimator identifies LATE:

$$\Delta^Y_{Wald} = \Delta^Y_{LATE},$$

with:

$$\Delta^Y_{Wald} = \frac{\mathbb{E}[Y_i|R_i = 1, E_i = 1] - \mathbb{E}[Y_i|R_i = 0, E_i = 1]}{\Pr(D_i = 1|R_i = 1, E_i = 1) - \Pr(D_i = 1|R_i = 0, E_i = 1)}.$$

Proof

Same as for IV.

Encouragement Design: Illustration

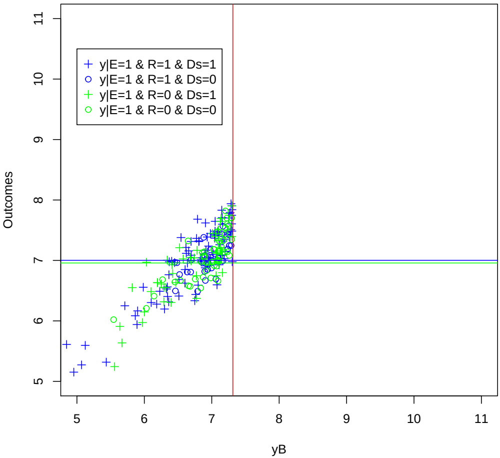

Figure: Encouragement Design

Encouragement Design: Illustration

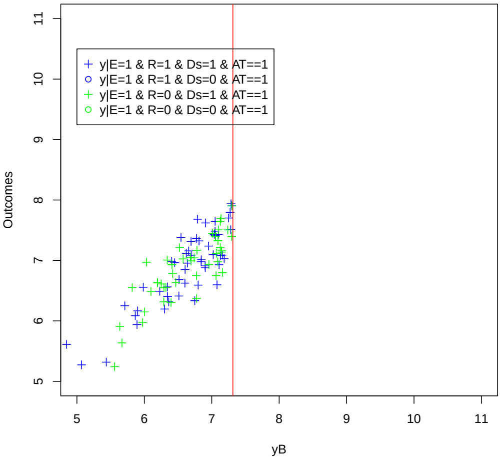

Figure: Encouragement Design

Encouragement Design: Illustration

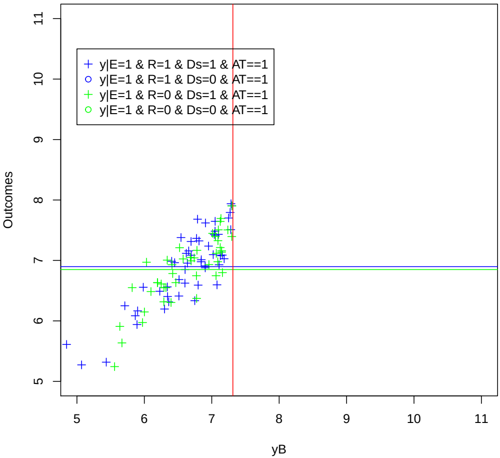

Figure: Encouragement Design

Encouragement Design: Illustration

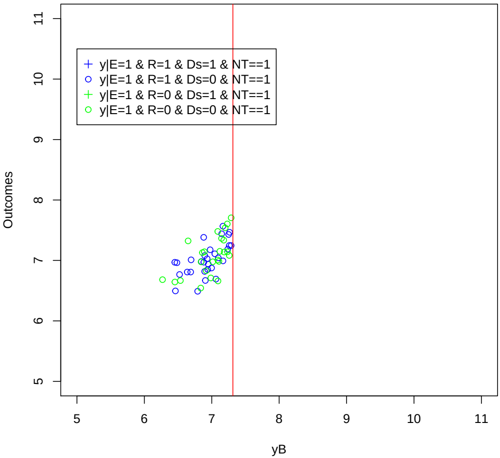

Figure: Encouragement Design

Encouragement Design: Illustration

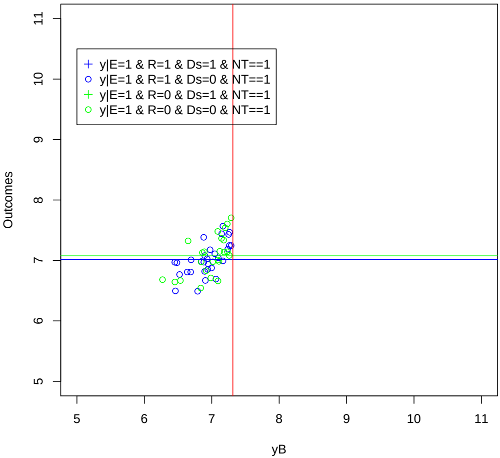

Figure: Encouragement Design

Encouragement Design: Illustration

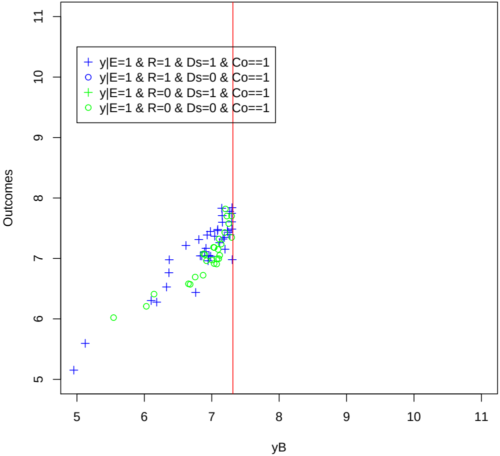

Figure: Encouragement Design

Encouragement Design: Illustration

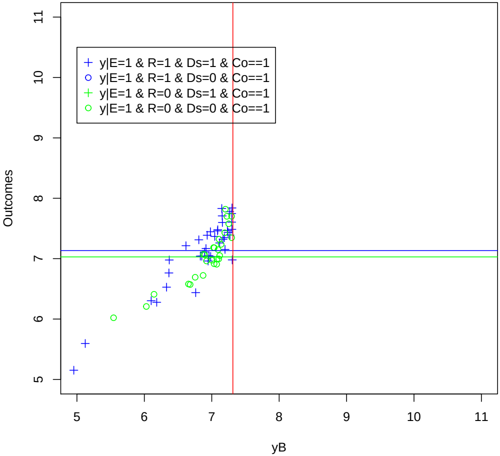

Figure: Encouragement Design

Encouragement Design: Illustration

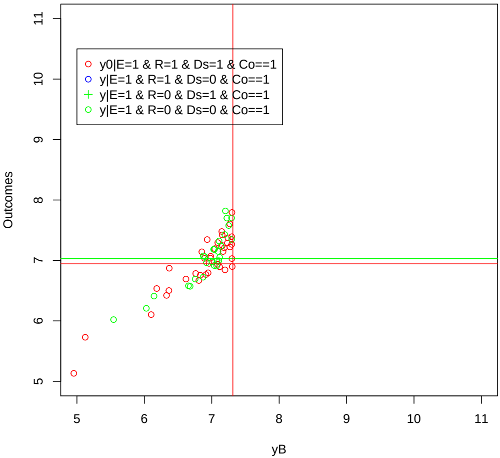

Figure: Encouragement Design

Direct Estimation Using Wald

$$\hat{\Delta}^{Y}_{Wald} = \frac{\frac{1}{\sum_{i=1}^{N} E_i R_i}\sum_{i=1}^{N} Y_i E_i R_i - \frac{1}{\sum_{i=1}^{N} E_i (1-R_i)}\sum_{i=1}^{N} E_i Y_i (1-R_i)}{\frac{1}{\sum_{i=1}^{N} E_i R_i}\sum_{i=1}^{N} D_i E_i R_i - \frac{1}{\sum_{i=1}^{N} E_i (1-R_i)}\sum_{i=1}^{N} D_i E_i (1-R_i)}.$$

Direct Estimation Using Wald: Illustration

$$\begin{aligned}\hat{\Delta}^{y}_{Wald} &= \frac{7.0022174 - 6.9578739}{0.75 - 0.4631579} \\ &= \frac{0.0443435}{0.2868421} \\ &= 0.1545919\end{aligned}$$

Estimation Using 2SLS

The 2SLS coefficient β in the following regression:

$$Y_i = \alpha + \beta D_i + U_i$$

with R_i as an instrument for D_i estimated on the sample of eligible individuals ($E_i = 1$) is the Wald estimator.

Estimation Using 2SLS: Illustration

```
reg.y.R.2sls.encourage <- ivreg(y[E==1] Ds[E==1]|R[E==1])
```

$\hat{\Delta}^{y}_{Wald} = 0.1545919$

The Value of LATE in our Illustration

$$\Delta^y_{LATE} = \bar{\alpha} + \theta\mathbb{E}[\mu_i | \mu_i + U_i^B \leq \bar{y} \wedge -\psi \leq \bar{\alpha} + \theta\bar{\mu} - \bar{c} - \gamma\mu_i - V_i \leq 0]$$

To compute the expectation of a doubly censored normal, I use the package tmvtnorm.
The value of Δ^y_{LATE} in our illustration is: 0.1731111.

Encouragement Design: Illustration

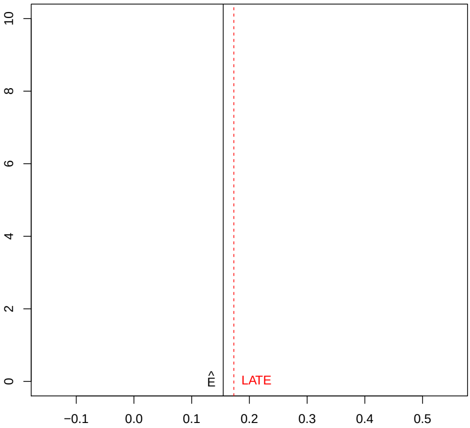

Figure: Encouragement Design

Sampling Noise with Wald in Encouragement Design: Illustration

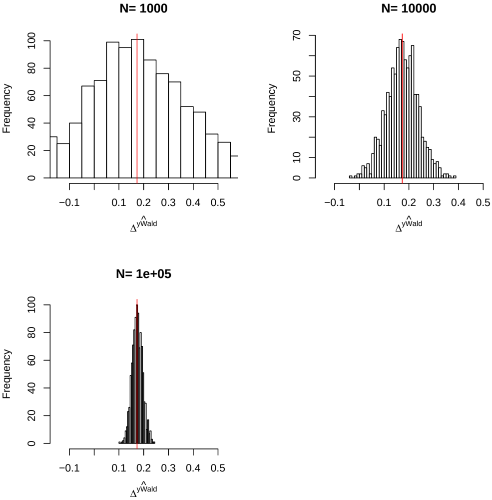

Figure: Distribution of the Wald estimator in an encouragement design over replications of samples of different sizes

Inference with Wald in Encouragement Design

- True 99% sampling noise (from the simulations) is 1.1664111
- 99% sampling noise estimated using default 2SLS standard errors is 1.3043035
- 99% sampling noise estimated using heteroskedasticity robust 2SLS standard errors is 1.2894331

Reducing Sampling Noise By Conditioning on Covariates

Parametrically estimating the following equation with R_i and X_i as instruments:

$$Y_i = \alpha + \beta D_i + \gamma' X_i + U_i$$

Or nonparametrically using Frolich's Wald matching estimator.

Estimation Using 2SLS with Covariates: Illustration

```
reg.y.R.yB.2sls.encourage <- ivreg(y[E==1]   Ds[E==1] +
yB[E==1] | R[E==1] + yB[E==1])
```

$\hat{\Delta}^{y}_{Wald(X)} =$0.1958423

Sampling Noise with Wald in Encouragement Design: Illustration

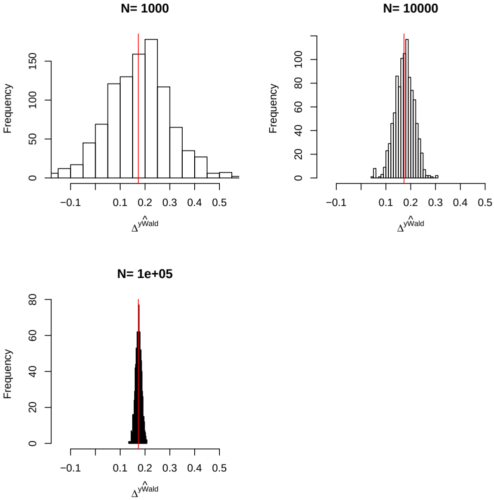

Figure: Distribution of the Wald estimator in an encouragement design over replications of samples of different sizes

Inference with Wald and Covariates in an Encouragement Design

- True 99% sampling noise (from the simulations) is 0.7053325
- 99% sampling noise estimated using default 2SLS standard errors is 0.673272
- 99% sampling noise estimated using heteroskedasticity robust 2SLS standard errors is 0.6744888

Outline

Brute Force Design

Randomization After Self-Selection

Randomization After Eligibility

Encouragement Design

Exercises

Exercises

1. Generate and analyze simulated data for the brute force, after self selection, eligibility and encouragement designs (compute WW, Bloom and Wald, with and without conditioning on covariates and estimate sampling noise)
2. Use your RCT to compute the Ww and Wald (or Bloom) estimator, conditioning and not conditioning on covariates, along with sampling noise.