For Objective Causal Inference, Design Trumps Analysis

Donald B. Rubin
Department of Statistics
Harvard University
16 March 2012

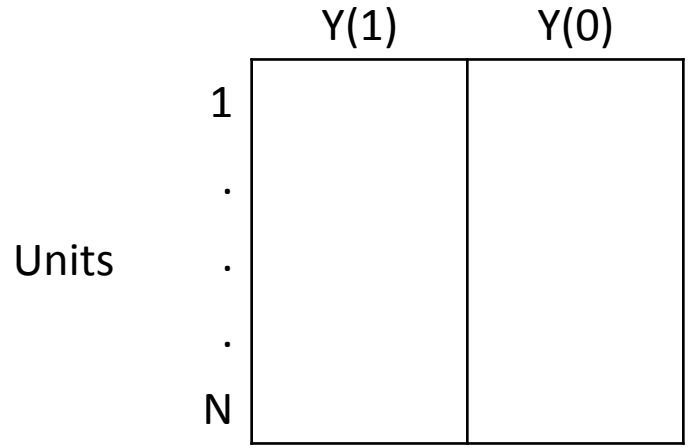
Prologue to Objective Causal Inference in Observational Studies

- My Introduction
  - Physics – Wheeler 1961
  - Experimental Design – Cochran 1968
- Clear Separation Between
  - Science = object of inference: DEFINE QUESTION FIRST
  - What is done to learn about the science
    - Intervene to measure aspects at a point in time
- Same notation/representation of science no matter how we try to learn about or measure
- Missing data always exist
  - Cannot go back in time

Potential Outcomes Approach to Causal Inference – Simplest Setting

T = 1 $\Rightarrow$ active treatment (e.g., job training)

T = 0 $\Rightarrow$ control treatment (e.g., no training)



Y(1) = outcomes if exposed to active treatment

Y(0) = outcomes if exposed to control treatment

Ave[$Y_i(1) - Y_i(0)$] = Average causal effect of active versus control treatment

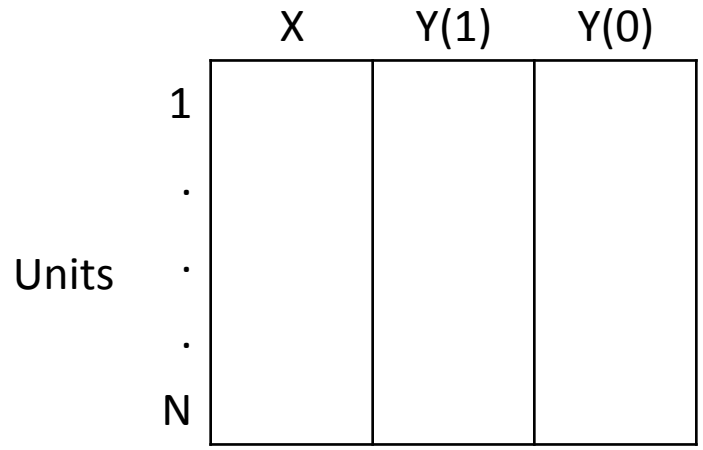
Potential Outcomes Approach to Causal Inference – Simplest Setting

- Fundamental problem of causal inference
  - For each *i*, only $Y_i(1)$ or $Y_i(0)$ can be observed

Units	Y(1)	Y(0)	T
1	✓	?	1
.	✓	?	1
.	✓	?	1
.	?	✓	0
.	?	✓	0
N	?	✓	0

- Random assignment of active versus control $\Rightarrow$ representative sample of $Y_i(1)$ will be compared to representative sample of $Y_i(0)$

Potential Outcomes Approach to Causal Inference – Simplest Setting with Covariates



- Same as before, except includes pretreatment covariates, e.g., age, sex, background education

$$\underset{i:\text{Female}}{\text{Ave}}\,[Y_i(1) - Y_i(0)] = \text{Average causal effect of active vs control treatment for females}$$

- Randomization still works for females

Randomized Trials That Are Designed Often Using Covariates

- Randomized blocks (e.g., males, females)
- Forces balance on blocking variables
- Probability of treatment versus control can depend on covariates' values, sometimes in complicated ways
- This is the template for the design and analysis of nonrandomized (e.g., observational) data
- Assignment-based approaches use assignment mechanism for inference – Fisherian & Neymanian
- Predictive approach predicts from observed values – Bayesian posterior predictive (Rubin, 1978)

Design Observational Studies to Approximate Randomized Trials

1. Hide outcome data until the design phase is complete
2. Think very carefully about decision makers and the key covariates that were used to make treatment decisions
3. If key covariates are not observed or very noisy, usually best to give up and seek better data source
4. Find subgroups (subclasses or matched pairs) in which the treatment and control groups have balance – essentially the same distribution of observed covariates
   - Not always possible to achieve balance
   - Inferences are limited to subgroups where balance is achieved
5. Protocol specified analysis

- #1 - #5 combine to create an objective design that approximates a randomized trial in each subclass that is balanced with respect to observed covariates

Illustrative Example with One Key Covariate (Cochran, 1968)

- Population: Male smokers in U.S.
- Treatment = cigar/pipe smoking
- Control = cigarette smoking
- Outcome = death rate/1000 person years
- Decision maker is the individual male smoker
- Reason for a smoking male to choose cigarettes versus cigar/pipe?
- **Age** is a key covariate for selection of smoking type for males

Subclassification to Balance Age

- To achieve balance on age, compare:
  - "young" cigar/pipe smokers with "young" cigarette smokers
  - "old" cigar/pipe smokers with "old" cigarette smokers
- Or better, compare:
  - Young, middle aged, old
  - Even more age subclasses
- Design phase, no outcome data, objective:
  - Approximates a randomized trial within subclasses
- Now look at outcome data

Reference: Rubin DB. The Design Versus the Analysis of Observational Studies for Causal Effects: Parallels With The Design of Randomized Trials. Statistics in Medicine 2007

Comparison of Mortality Rates for Two Smoking Groups in U.S.

Variable	Cigarette Smokers	Cigar/Pipe Smokers
Mortality Rates per 1000 person-years, %	13.5	17.4
Adjusted Mortality Rates using subclasses, %		
2 age subclasses	16.4	14.9
3 age subclasses	17.7	14.2
9-11 age subclasses	21.2	13.7

Source: Cochran WG. The effectiveness of adjustment of subclassification in removing bias in observational studies. Biometrics 1968; 24:295-313.

Note: 20 four-level covariates $\Rightarrow$ over million million subclasses

Propensity Score Methods

- Rosenbaum and Rubin. "The Central Role of the Propensity Score in Observational Studies." Biometrika 1983.
- Observational study analogue of randomization
- The propensity score is the probability of treatment versus control as a function of observed covariates
  - Model the reasons for treatment versus control at the level of the decision makers
  - For example, logistic regression model to predict cigarette versus cigar/pipe smoking with age, education, income, etc. as predictors
- Then subclassify (or match) on the propensity score as if it were the only covariate, e.g., 5-10 subclasses
- If correctly done, this creates balance within each subclass on **ALL** covariates used in estimating the propensity score
- Using diagnostics to assess and to document balance is critical

Example: GAO Study of Breast Conservation versus Mastectomy

- Six large and expensive randomized clinical trials had been completed showing little difference for the type of women randomized in the trials and participating clinics
- Question: Same results in general practice?
- Observational data available
  - SEER Database: covariates, treatments, post-surgery outcomes
- Design phase
  - Hide outcomes
  - Balance covariates between treatment and control
- Reasons for mastectomy versus breast conservation
  - Age, marital status, region of country, urbanization, race, size of tumor, etc.

Reference: Rubin DB. Estimated Causal Effects from Large Datasets Using Propensity Scores. Annals of Internal Medicine 1997; 127, 8(II):757-763.

Estimated 5-year Survival Rates for Node-negative Patients in Six Randomized Clinical Trials

Study	Women		Estimated Survival Rate for Women		Estimated Causal Effect
	Breast Conservation (BC)	Mastectomy (Mas)	BC	Mas	BC – Mas
	n	n	%	%	%
US-NCI†	74	67	93.9	94.7	-0.8
Milanese†	257	263	93.5	93.0	0.5
French†	59	62	94.9	96.2	-1.3
Danish‡	289	288	87.4	85.9	1.5
EORTC‡	238	237	89.0	90.0	-1.0
US-NSABP‡	330	309	89.0	88.0	1.0

†Single-center trial; ‡ Multicenter trial

Reference: Rubin DB. Estimated Causal Effects from Large Datasets Using Propensity Scores. Annals of Internal Medicine 1997; 127, 8(II):757-763.

Propensity Score Analysis Approach

- Estimate propensity scores
- Then subclassify (or match) on propensity score as if the only covariate, e.g., 5-10 subclasses
- Why does this work?
  - Creates balance in each subclass on **ALL** covariates used in estimating the propensity score
  - This balance will be achieved in large samples just like the balance that will be achieved in a large randomized clinical trial

Estimated 5-year Survival Rates for Node-Negative Patients in the SEER Database within Each of Five Propensity Score Subclasses

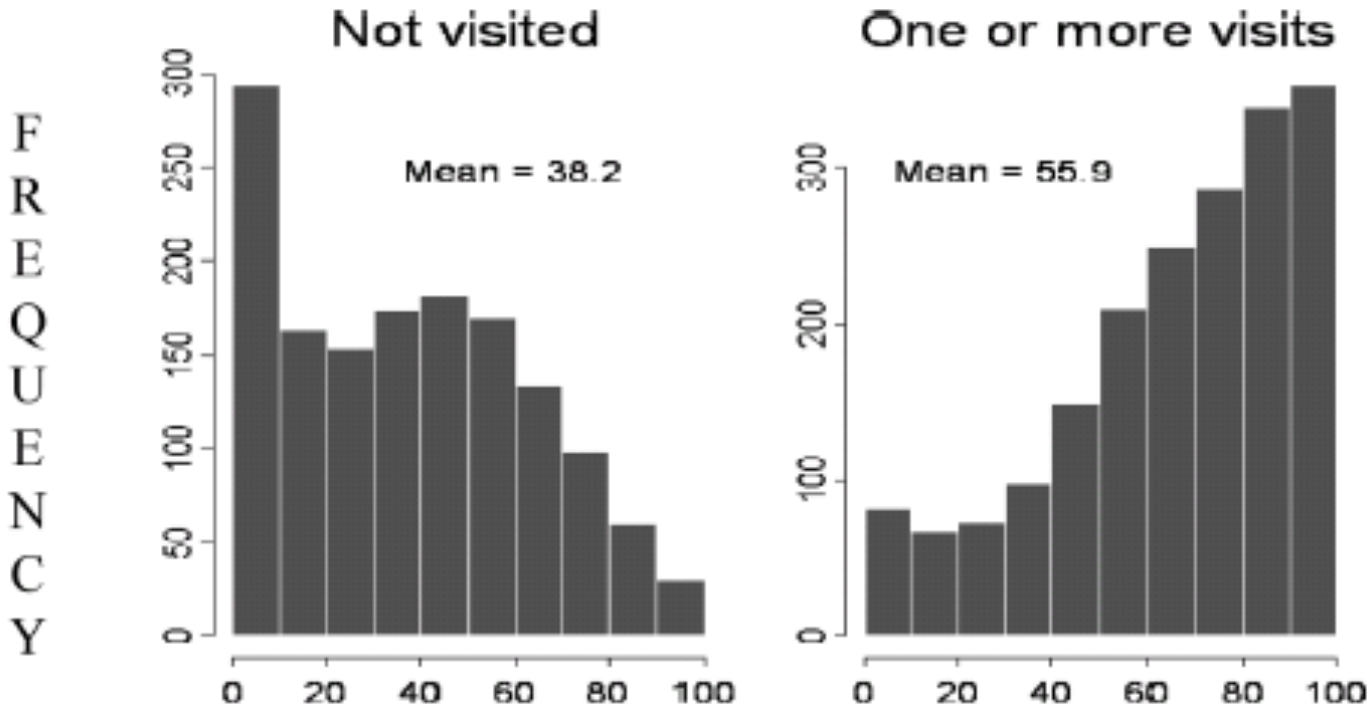
Propensity Score Subclass	Women		Estimated Survival Rate for Women		Estimated Causal Effect
	Breast Conservation (BC)	Mastectomy (Mas)	BC	Mas	BC – Mas
	n	n	%	%	%
1	56	1008	85.6	86.7	-1.1
2	106	964	82.8	83.4	-0.6
3	193	866	85.2	88.8	-3.6
4	289	978	88.7	87.3	1.4
5	462	604	89.0	88.5	0.5
Averages Across Five Subclasses			86.3	86.9	-0.6

Reference: Rubin DB. Estimated Causal Effects from Large Datasets Using Propensity Scores. Annals of Internal Medicine 1997; 127, 8(II):757-763.

Diagnostics for Accessing Balance

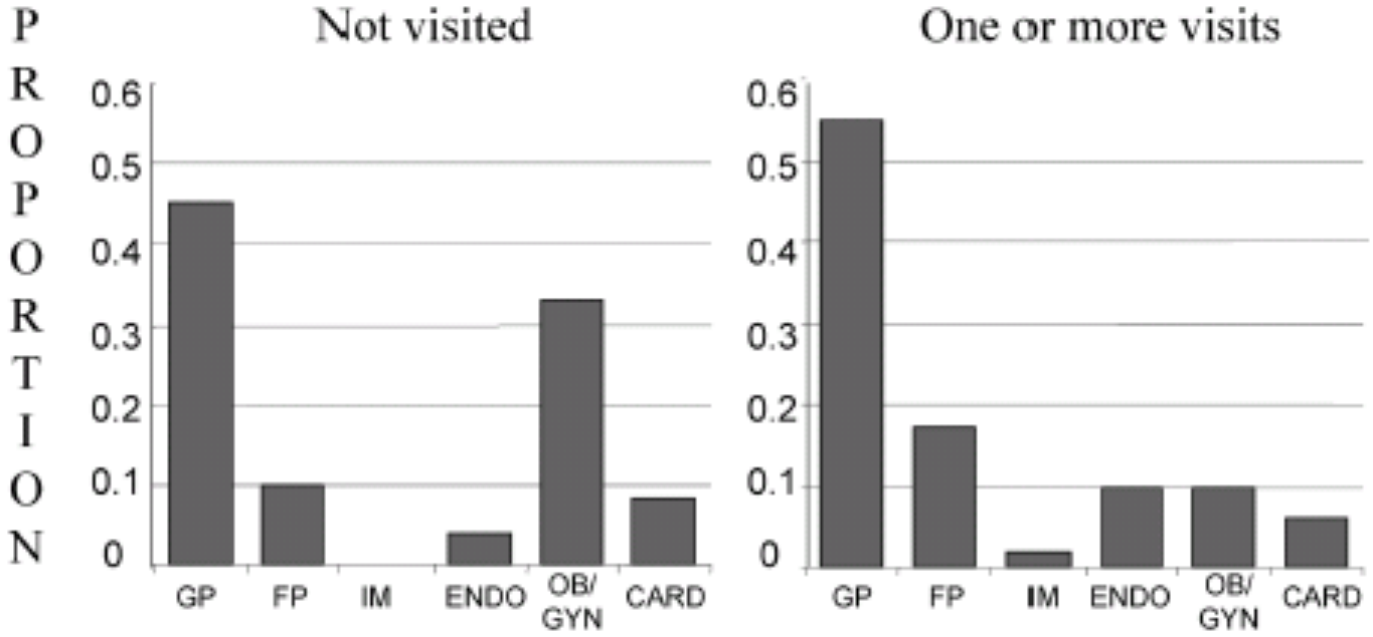
- Assessing balance simpler in large samples, just as with randomized experiments
- To illustrate diagnostics, use a marketing application that involved a weight loss drug
- Units = doctors
- Treatment = sales rep “visits” doctor to discuss
- Control = no visit
- Decision-makers = sales reps
- Key covariates = prior Rxs, medical specialty, years in practice, size of practice, etc.

Histograms for background variable: Prior Rx Score (0-100) at Baseline



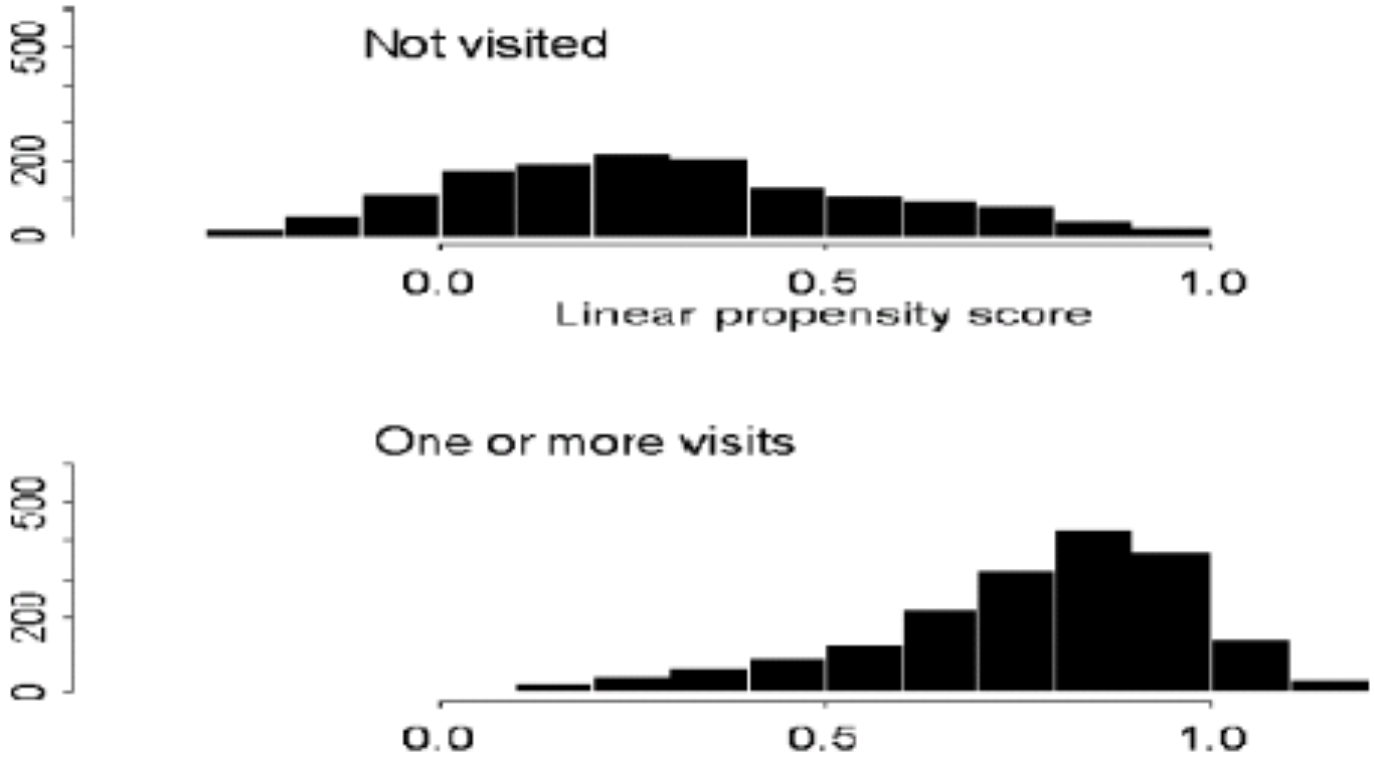
Source: Rubin DB and Waterman RP. Estimating Causal Effects of Marketing Interventions Using Propensity Score Methodology. Statistical Science 2006; 21(2):206-222.

Histograms for background variable: Specialty



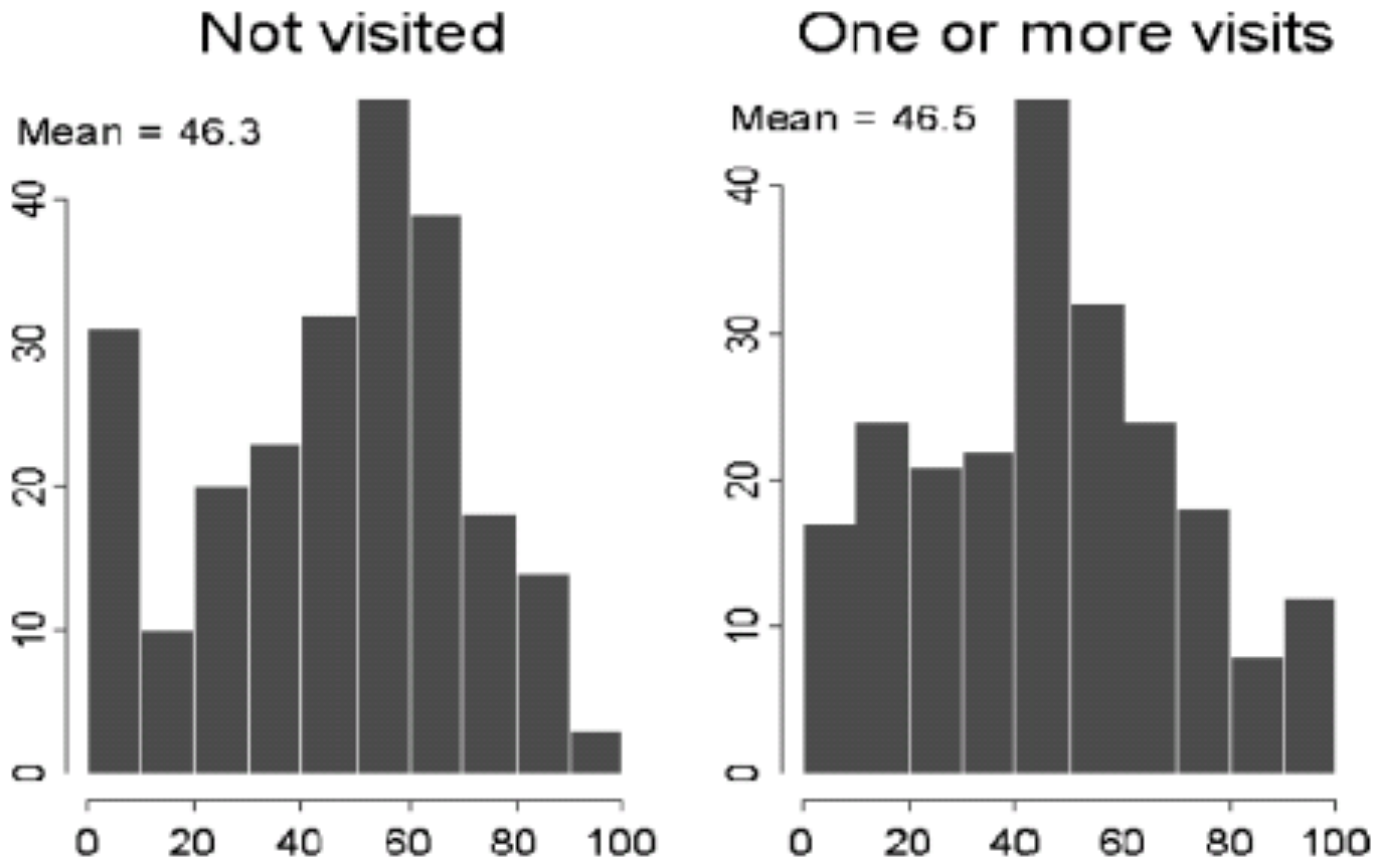
Source: Rubin DB and Waterman RP. Estimating Causal Effects of Marketing Interventions Using Propensity Score Methodology. Statistical Science 2006; 21(2):206-222.

Histograms for summarized background variables: Linear Propensity Score



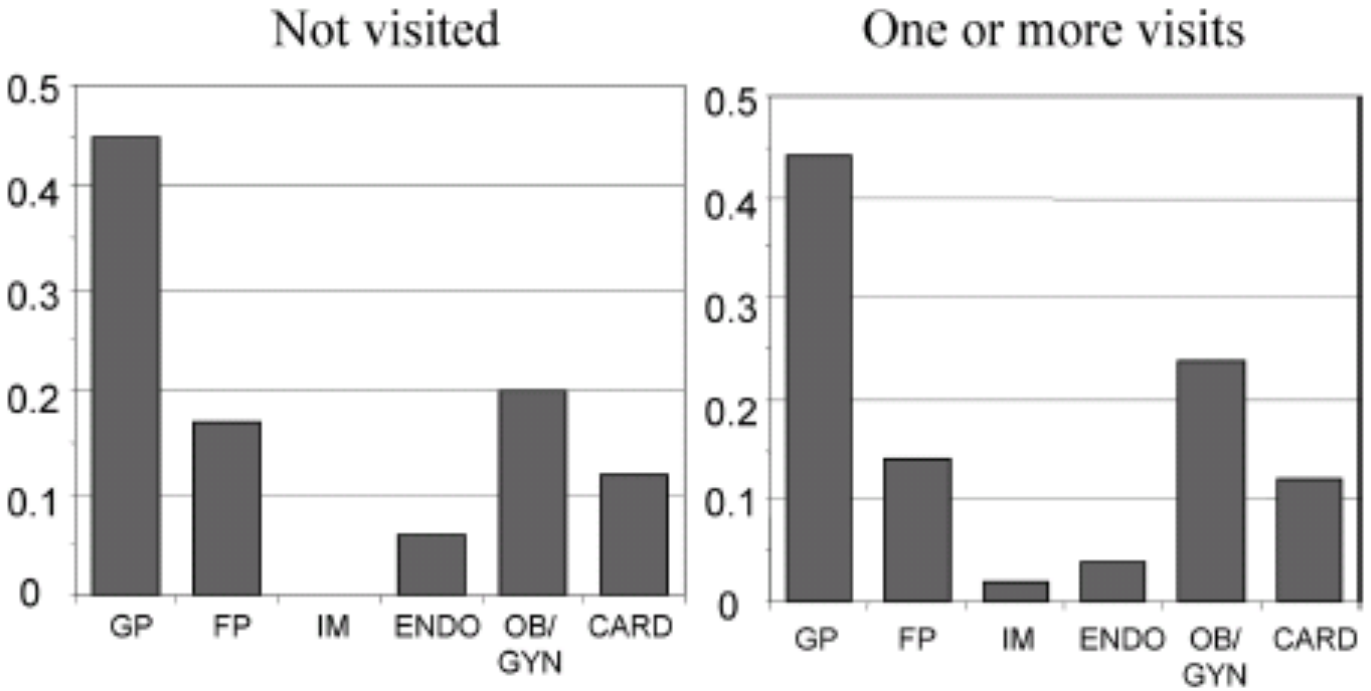
Source: Rubin DB and Waterman RP. Estimating Causal Effects of Marketing Interventions Using Propensity Score Methodology. Statistical Science 2006; 21(2):206-222.

Histograms for a variable in a subclass of propensity scores: Prior Rx Score



Source: Rubin DB and Waterman RP. Estimating Causal Effects of Marketing Interventions Using Propensity Score Methodology. Statistical Science 2006; 21(2):206-222.

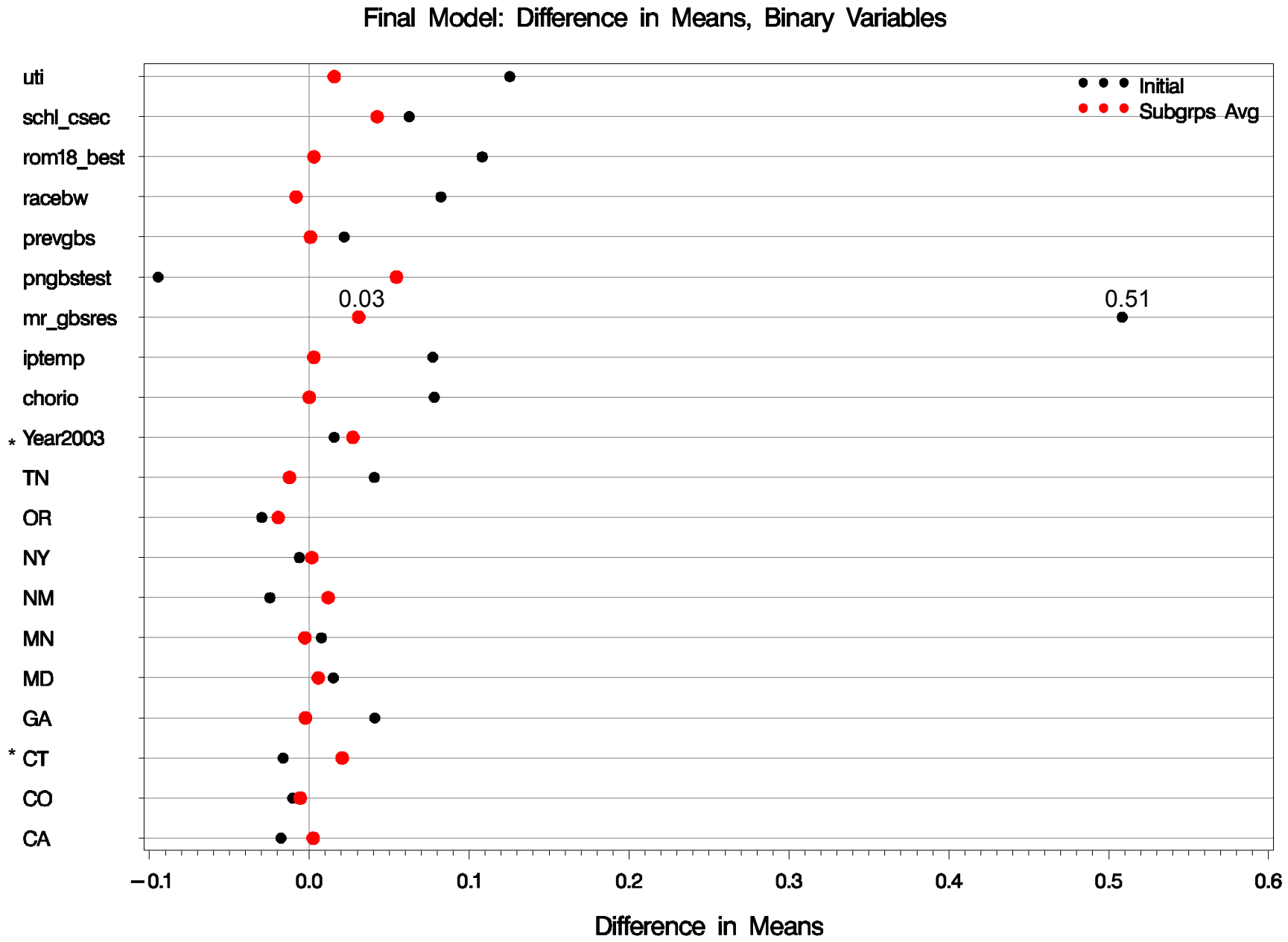
Histograms for a variable in a subclass of propensity scores: Specialty

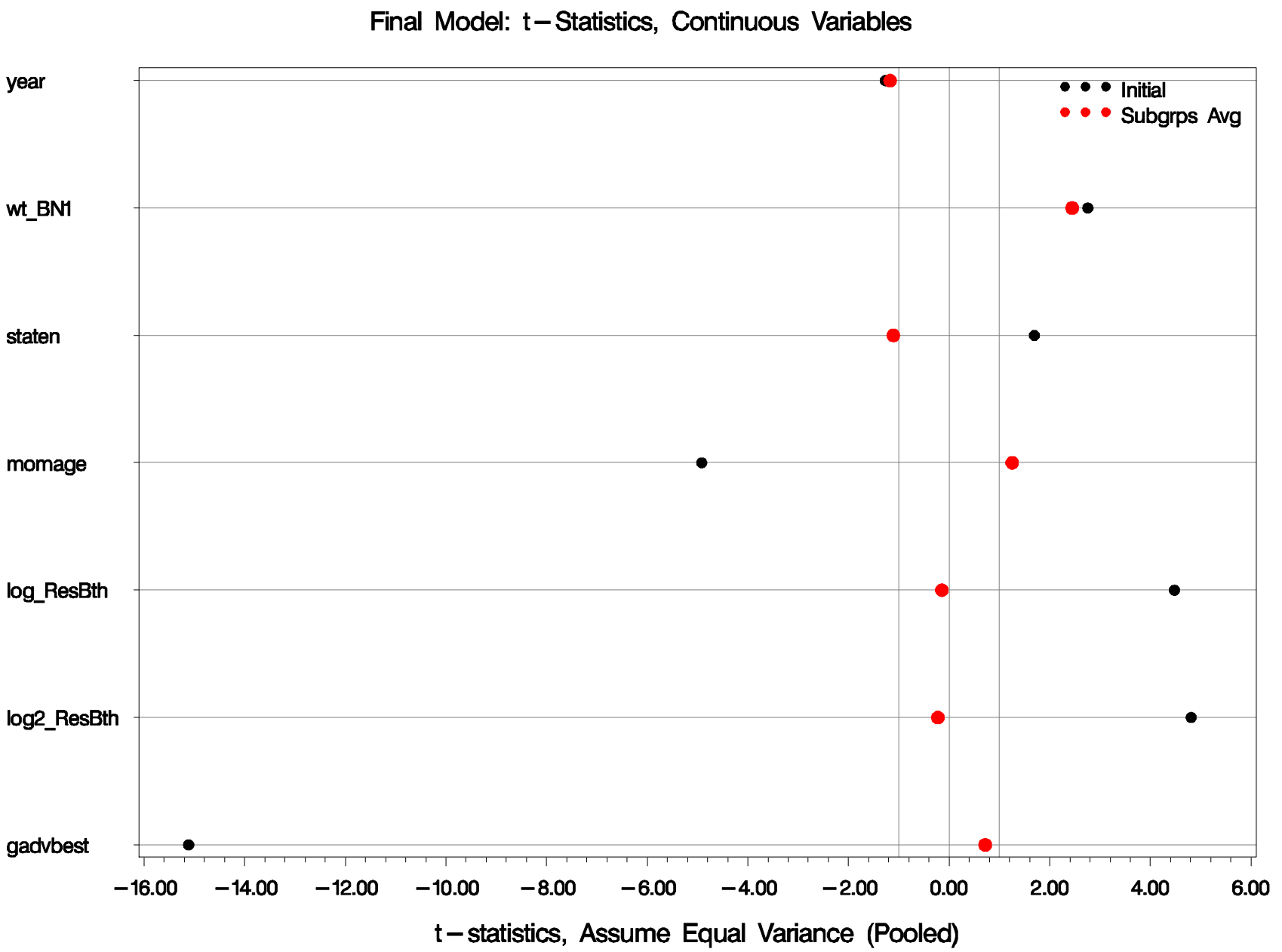


Source: Rubin DB and Waterman RP. Estimating Causal Effects of Marketing Interventions Using Propensity Score Methodology. Statistical Science 2006; 21(2):206-222.

Marketing Example: Achieved Balance

- Within each narrow subclass of propensity scores, the treatment and control groups will be as balanced as if randomly divided
- Claim: This holds for all subclasses in which there are both treated and control subjects, and holds for all covariates that were used to estimate the propensity score
- Works best when the propensity score subclasses have large sample sizes and are relatively narrow
- Five to ten propensity score subclasses often fully adequate to balance all covariates
- No outcome data used in the design stage


Final Model: Difference in Means, Binary Variables
uti
schl_csec
rom18_best
racebw
prevgbs
pngbstest
mr_gbsres
iptemp
chorio
* Year2003
TN
OR
NY
NM
MN
MD
GA
* CT
CO
CA
Initial
Subgrps Avg
0.03
0.51
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Difference in Means


Final Model: t – Statistics, Continuous Variables
year
wt_BN1
staten
momage
log_ResBth
log2_ResBth
gadvbest
Initial
Subgrps Avg
–16.00
–14.00
–12.00
–10.00
–8.00
–6.00
–4.00
–2.00
0.00
2.00
4.00
6.00
t – statistics, Assume Equal Variance (Pooled)

Simple Noncompliance, Instrumental Variables, and Bayesian Generalizations

- Template for other observational studies involves more complex randomized experiment
- Illustrate with completely randomized experiment with noncompliance with assigned treatment
- Return later to combined analysis with observational study design

Sommer and Zeger Vitamin A Data

Row	True Compliance Type	Treatment Assignment	Treatment Received	Y_{obs}	Number of Children
1	?	0	0	0	11514
2	?	0	0	1	74
3	N	1	0	0	2385
4	N	1	0	1	34
5	C	1	1	0	9663
6	C	1	1	1	12
					23682

Reference: Sommer and Zeger (1991). On Estimating Efficacy from Clinical Trials. Statistics in Medicine.

Results of Three Standard MoM Analyses

Method	Estimate	Calculation	Row Comparison
ITT	-0.0026	$= \frac{12+34}{9663+2385+12+34} - \frac{74}{11514+74}$	3, 4, 5, & 6 vs. 1 & 2
As-treated	-0.0065	$= \frac{12}{9663+12} - \frac{34+74}{11514+2385+34+74}$	5 & 6 vs. 1, 2, 3, &4
Per protocol	-0.0052	$= \frac{12}{9663+12} - \frac{74}{11514+74}$	5 & 6 vs. 1 & 2

Reference: Sommer and Zeger (1991). On Estimating Efficacy from Clinical Trials. Statistics in Medicine.

MoM CACE Analysis

$ACE = p_N \cdot NACE + p_C \cdot CACE$

$-0.0025 = 0.2 \cdot NACE + 0.8 \cdot CACE$

$-0.0025 = 0.8 \cdot CACE \rightarrow CACE = -0.0025/0.8 = -0.0031$

Bayesian Analysis of Sommer & Zeger Data

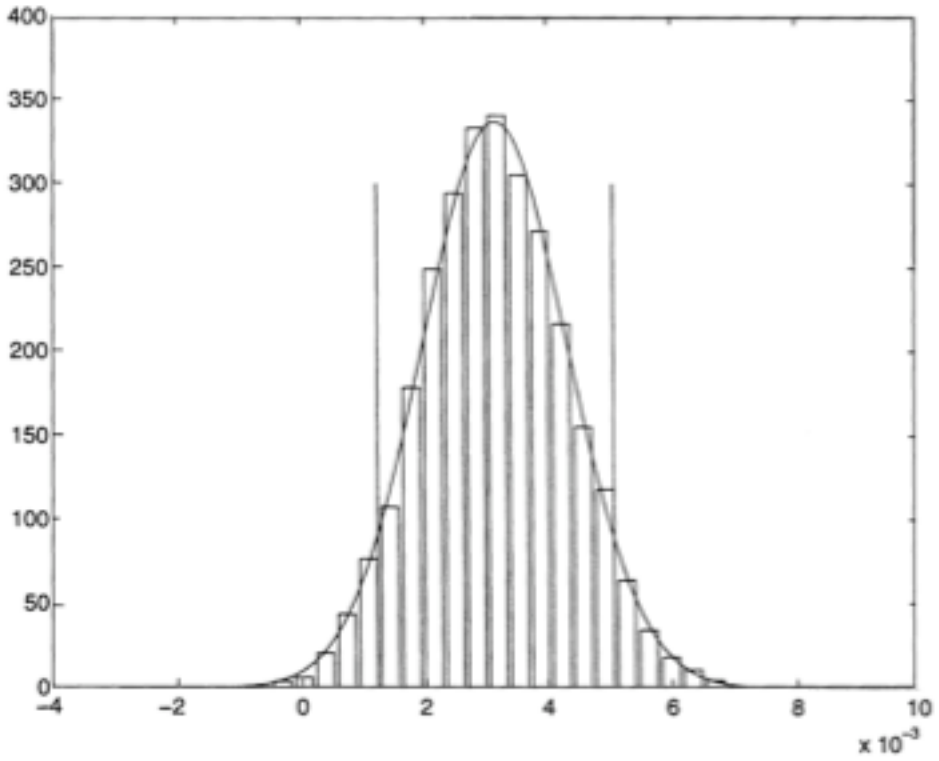


FIG. 3. *Histogram of CACE with exclusion restriction (data from Table 3).*

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Bayesian Analysis of Sommer & Zeger Data, Marginal Posterior Distributions with and without Exclusion Restriction

Estimand	Exclusion restriction	Mean	Standard deviation	Median	5^{th} percentile	95^{th} percentile
CACE	No	3.1	2.5	3.2	-0.9	7.0
$ITT_Y^{(n)}$	No	0.5	10.1	0.2	-14.1	17.5
CACE	Yes	3.1	1.2	3.1	1.2	5.1

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Bayesian Analysis of Sommer & Zeger Data

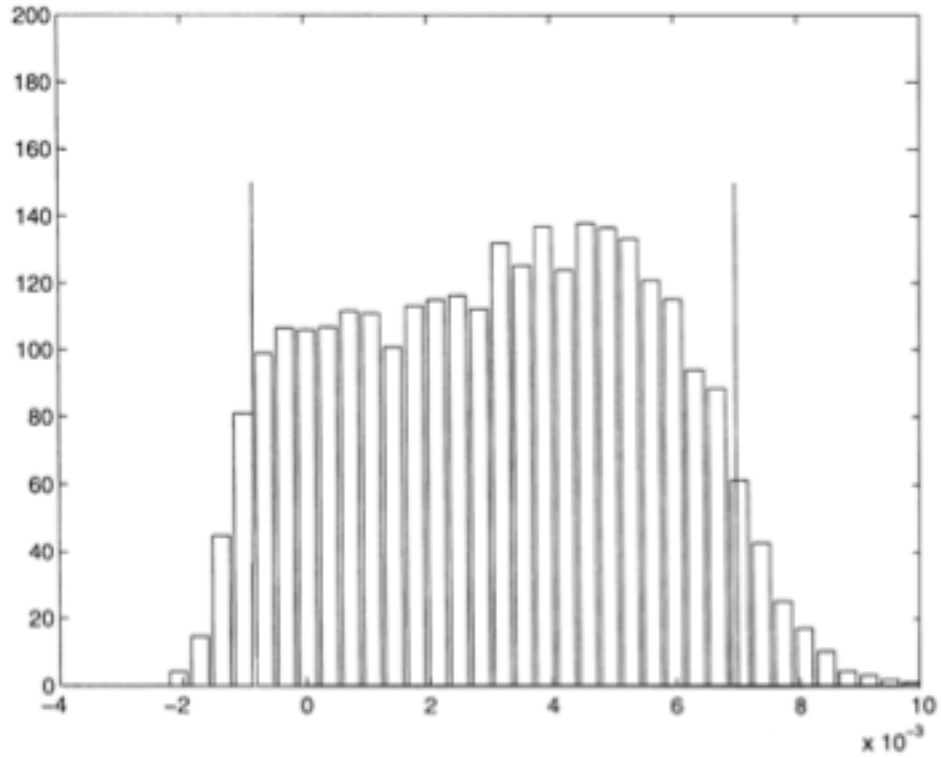


FIG. 1. *Histogram of CACE without exclusion restriction (data from Table 3).*

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Bayesian Analysis of Sommer & Zeger Data

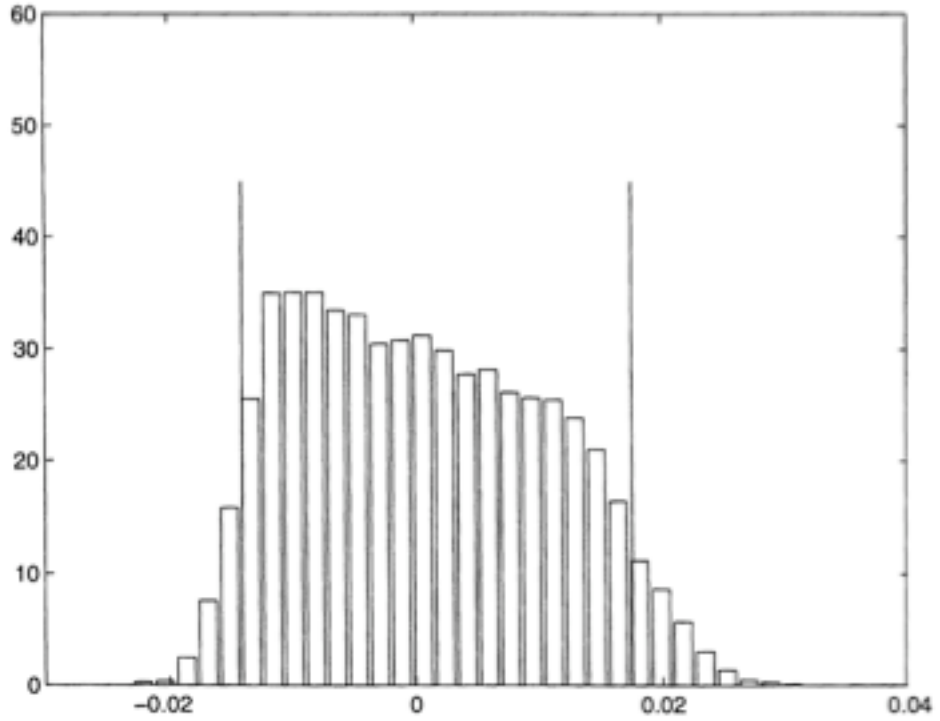


FIG. 2. *Histogram of* $\mathrm{ITT}_Y^{(n)}$ *without exclusion restriction (data from Table 3).*

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Bayesian Analysis of Sommer & Zeger Data

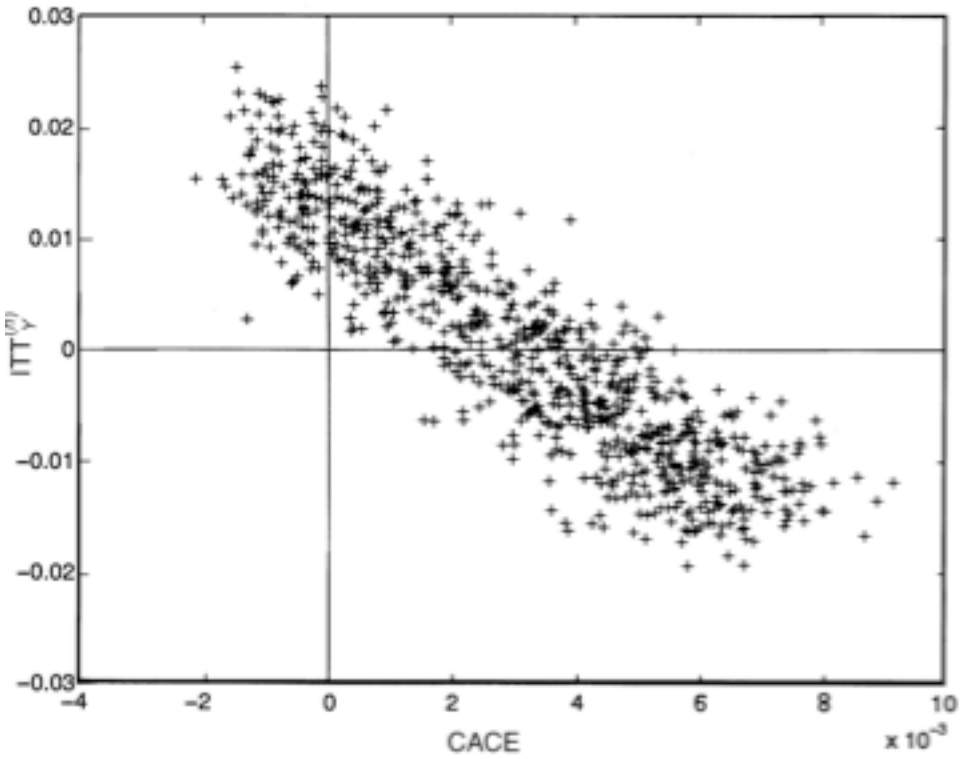


FIG. 4. *Joint posterior distribution of CACE and* $\mathrm{ITT}_Y^{(s)}$ *(data from Table 3).*

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Hypothetical Example Illustrating Frequentist Superiority of Bayes over IVE (MoM) and MLE, Population Parameters with Exclusion Restrictions and Monotonicity

T	$P(C_i = t \mid \pi)$	$D_i(0)$	$D_i(1)$	$Y_i \mid C_i = t, Z_i = 0, \pi$	$Y_i \mid C_i = t, Z_i = 0, \pi$
c	0.25	0	1	$N(0.1, 0.16)$	$N(0.9, 0.49)$
n	0.45	0	0	$N(1.0, 0.25)$	$N(1.0, 0.25)$
a	0.30	1	1	$N(0.0, 0.36)$	$N(0.0, 0.36)$

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Hypothetical Example Illustrating Frequentist Superiority of Bayes over IVE (MoM) and MLE, One Sample

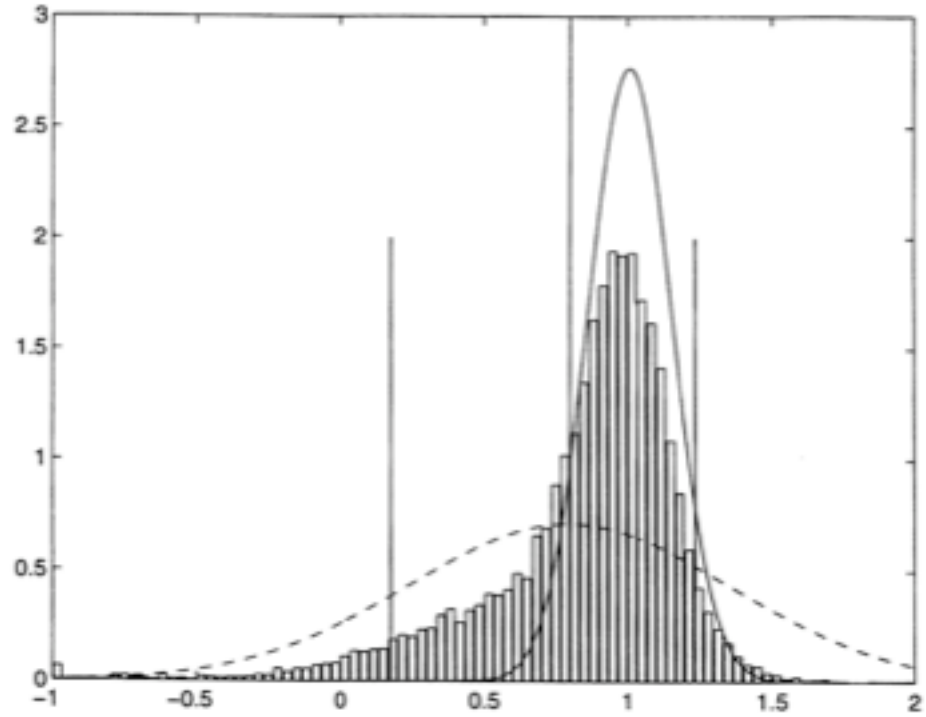


FIG. 5. *Estimates of the posterior distribution of CACE under exclusion restriction and monotonicity condition (data analyzed in Table 6): histogram is based on simulation, solid line is normal approximation based on mle, dashed line is normal approximation based on* $\widehat{\text{IVE}}$.

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Hypothetical Example Illustrating Frequentist Superiority of Bayes over IVE (MoM) and MLE, Frequentist Evaluation under Monotonicity and Exclusion Restrictions

Estimator	Mean bias	Median bias	Root mean squared error	Median absolute error	90% interval	
					Coverage rate	Median width
Posterior mean	-0.10	-0.07	0.48	0.30	0.91	1.61
Posterior median	-0.08	-0.06	0.51	0.32		
MLE	-0.14	-0.12	0.51	0.31	0.74	1.11
IVE	0.55	0.13	2.31	0.54	0.91	2.78

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Jin and Rubin (JASA, 2008) - JR

- JR reanalyzed data from EF using principal stratification (Frangakis & Rubin, Biometrics, 2002)
- Randomized Treatment (active drug) versus Control (placebo) and meaured:
    - Compliance (pill count) - intermediate outcome
    - Cholesterol reduction - primary outcome
- Critical attributes of Analysis Here:
  - Principal stratification and the various problems that it can address
  - Formulation using hypothetical experiment
  - Bayesian method of analysis, which easily incorporates scientific understanding

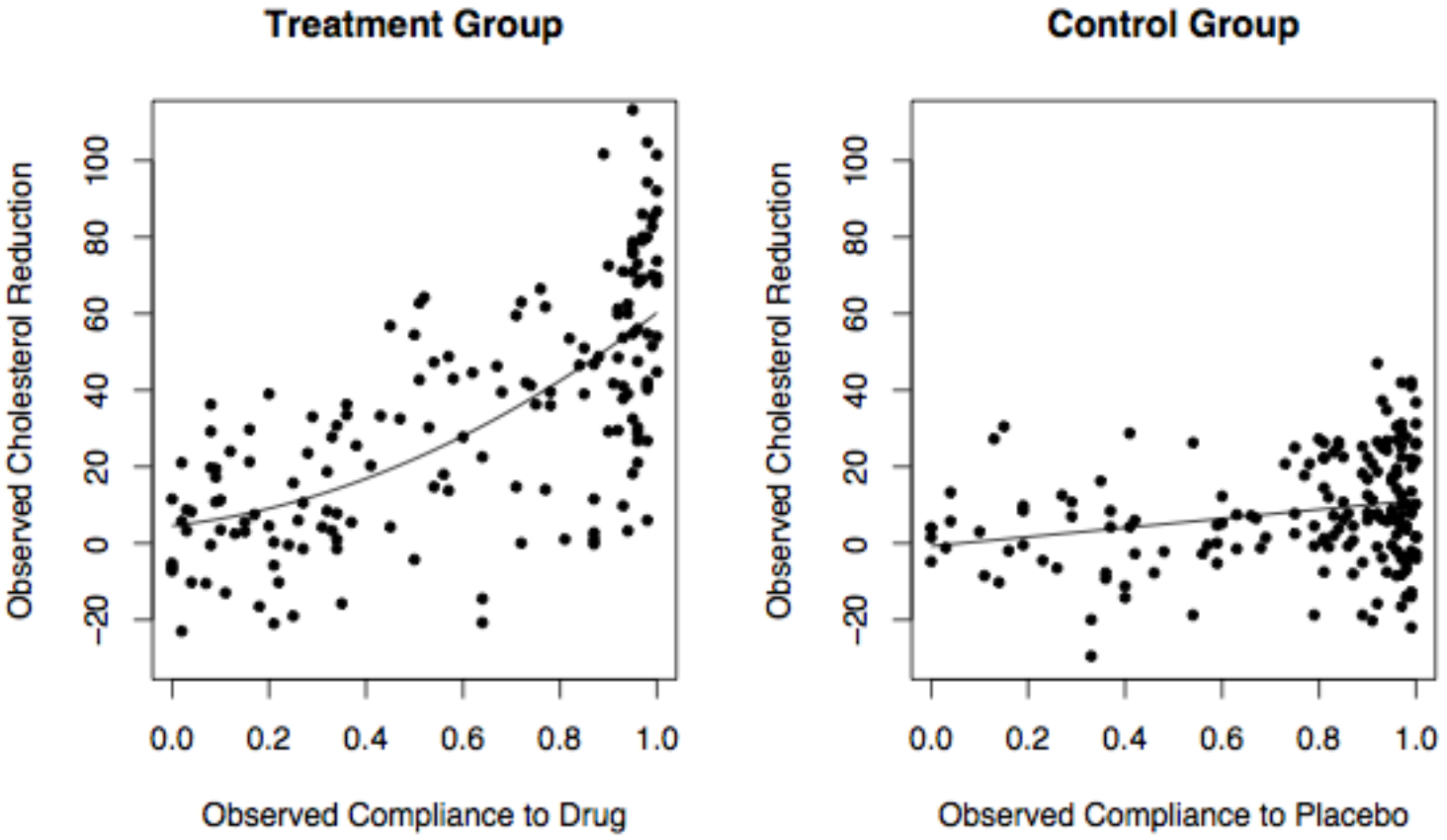
The EF Data

- 164 men were randomized to the treatment group and assigned the drug, $Z_i = T$
- 171 men were randomized to the control group and assigned placebo, $Z_i = C$
- For each patient, cholesterol levels were measured before and after taking the drug or placebo
- The outcome variable, $Y_i(T)$ or $Y_i(C)$, was the decrease in cholesterol level: the only variable used by EF or JR, besides treatment assigned and dose taken

Complications with the EF Data

- Partial and Extended Noncompliance
  - Most patients in the treatment group took only a proportion of the assigned drug: $D_i(T) \in [0,1]$
  - Most patients in the control group took only a proportion of the assigned placebo: $d_i(C) \in [0,1]$
  - By design, $D_i(C) = 0$ and $d_i(T) = 0$
- Thanks to Brad Efron for sharing data

Relationship Between Observed Cholesterol Reduction and Observed Compliance



Figures from Efron and Feldman, 1991

How to estimate dose-response?

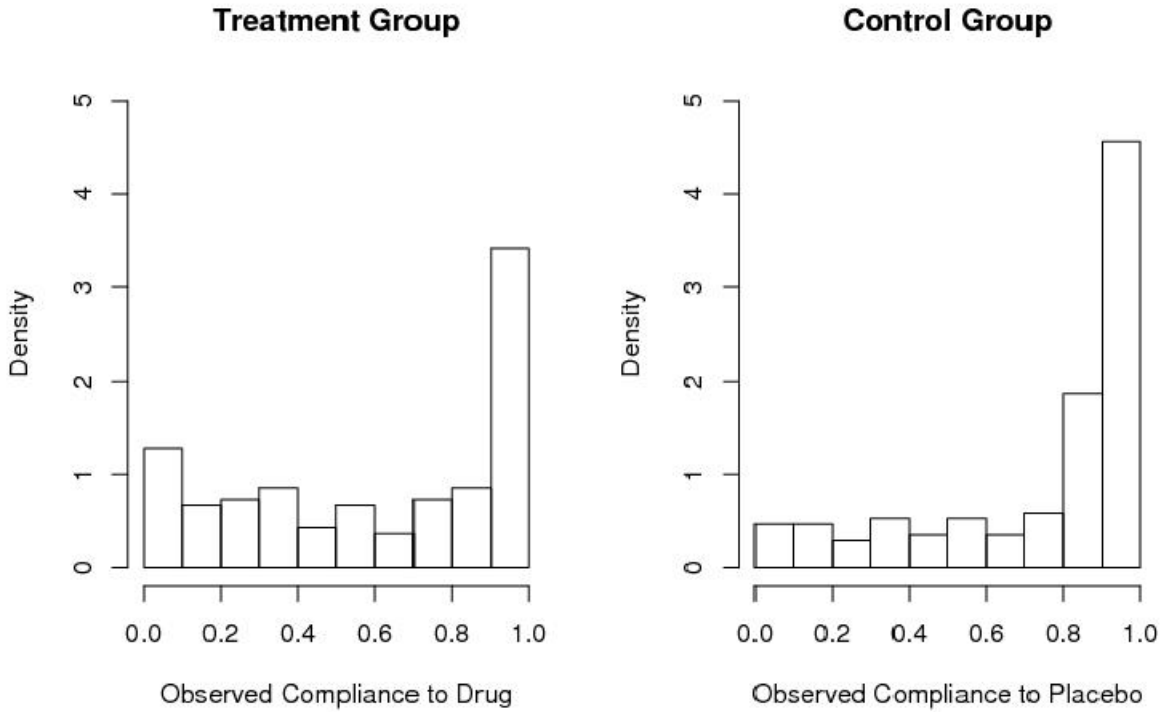
- Observed “dose-response” in both arms
- Somehow, “subtract” Y(C) versus d(C) plot from Y(T) versus D(T) plot
- EF attempted this, but written discussion (including by DBR) of article indicated debatable success
- Objective in JR was to do this “subtraction” correctly under explicit assumptions
- Here, highlight principal stratification, hypothetical experiment for dose-response, and Bayesian approach to analysis

Standard Assumptions

- Stable Unit Treatment Value Assumption (SUTVA):
  - One patient's treatment assignment does not affect other patients' potential outcomes;
  - For each patient, no hidden versions of treatment and no hidden versions of control
- Ignorable Treatment Assignment of T versus C
  - True for randomized experiment

These are accepted by both EF and JR.

Histograms of Observed Compliance



Figures from Jin and Rubin, 2008

Q-Q Plot of Observed Drug and Observed Placebo Compliance

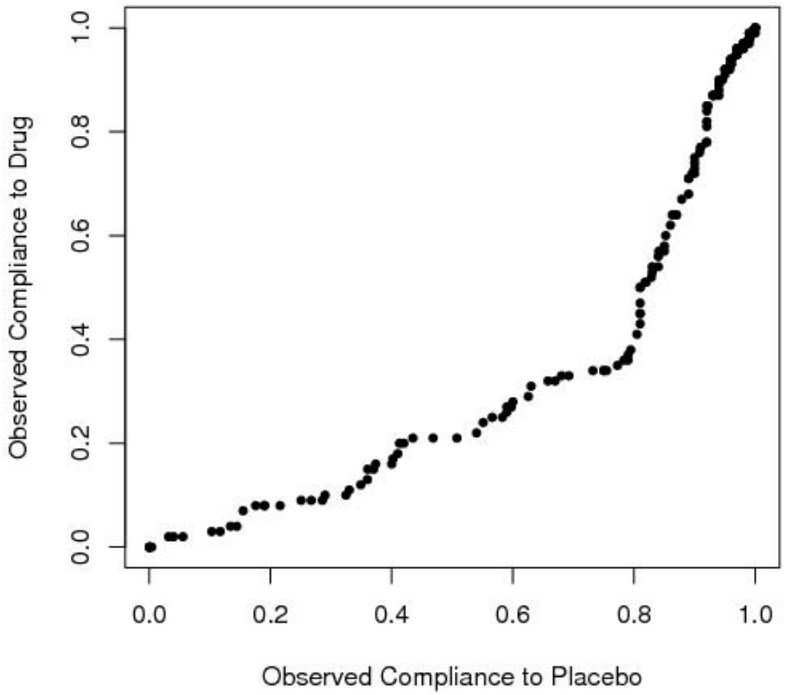


Figure from Jin and Rubin, 2008

Possible Assumptions at the Individual Level

- Perfect Blind: $D_i(T) = d_i(C)$; obviously wrong
- Equipercentile Equatable Compliances
  - Align percentiles of $D_i(T)$ and $d_i(C)$, as in Q-Q
  - So both known for all men: $D_i(T) = F(d_i(C))$
  - EF assume this, which is true in expectation
- Side-Effect Monotonicity
  - Negative: $D_i(T) \leq d_i(C)$
  - Positive: $D_i(T) \geq d_i(C)$
  - JR assume negative side effects; plausible

Meaning of d_i and D_i

- d_i: compliance to placebo indicates patient i's "psychological" compliance status, a covariate that is missing for men assigned drug
- D_i: compliance to drug reflects both patients i's psychological compliance status and his tolerance to negative side effects of the drug, etc.
- But D_i hints at possibility of estimating dose-response
- Similar comments in EF, but JR allow $D_i(T) \neq F(d_i(C))$

Estimating a Dose-Response Relationship within the "Rubin Causal Model" (Holland, 1986)

- To estimate a dose-response relationship
  - Need a hypothetical experiment where different doses of drug are randomly assigned and enforced
- Principal stratification framework (Frangakis and Rubin, 2002) – vast generalization of IVE
  - The intermediate outcome $d_i(C)$ is unaffected by treatment assignment
  - Therefore is a partially observed covariate
- For each stratum of patients with the same $d_i(C)$, the assignment of dose is stochastic and "latent ignorable" (Frangakis and Rubin, 1999)

Specific Hypothetical Experiment

- Measure d_i^* = baseline compliance for each patient when assigned full placebo dose
- Randomly divide patients into Treatment and Control
- In treatment group, stochastically assign dose $Z_{D_i} \leq d_i^*$ according to a Beta random variable
- In control group, assign full placebo and measure d_i
- We notice $d_i(C) = d_i^*$ in the control group, then "toss" d_i^* in the control group and in the treatment group
- Thus, nonignorable assignment of Z_{D_i}, but latent ignorable given d_i^*
- Also, "forget" the rule for the assignment of Z_{D_i}

Principal Stratification Framework for Dose-Response with $d_i(C)$ Defining Strata and $Z_{D_i}(T)$ Defining Dose

i	d_i^*	Z_i	Z_{Di}	$d_i(T)$	$d_i(C)$	$Y_i(T_0)$	...	$Y_i(T_D)$	...	$Y_i(T_1)$	$Y_i(C)$
1	?	T	T_0	0	?	$\star$	?	?	?	?	?
...	?	T	...	0	?	...	...	...	...	...	?
...	?	T	T_D	0	?	?	?	$\star$	?	?	?
...	?	T	...	0	?	...	...	...	...	...	?
n_T	?	T	T_1	0	?	?	?	?	?	$\star$	?
n_T+1	?	C	?	0	$\star$	?	?	?	?	?	$\star$
...	?	C	?	0	$\star$	?	?	?	?	?	$\star$
...	?	C	?	0	$\star$	?	?	?	?	?	$\star$
n	?	C	?	0	$\star$	?	?	?	?	?	$\star$

"$\star$" represents observed data, "?" represents missing data.

Treatment Assignment Mechanism

Notation

- θ = All parameters
- Assume $d_i(C) = d_i^*$ for everyone, denoted d_i

Actual Randomization of Z=T versus Z=C

- $[Z|d_i, Y_i(C), \{Y_i(Z_{D_i})\}, \theta] = [Z|\theta] \sim$ constant

Hypothetical Randomization of Dose Z_D given Z=T

- $[Z_{D_i}|d_i, Y_i(C), \{Y_i(Z_{D_i})\}, Z=T, \theta] = [Z_{D_i}|d_i, Z=T, \theta] \sim d_i \cdot \text{Beta}(\alpha_1, \alpha_2)$
- Latently ignorable given partially observed variable d_i

JR's Parametric Model Given θ

Covariate Distribution

- $[d_i|\theta] \sim \text{Beta}(\alpha_3, \alpha_4)$

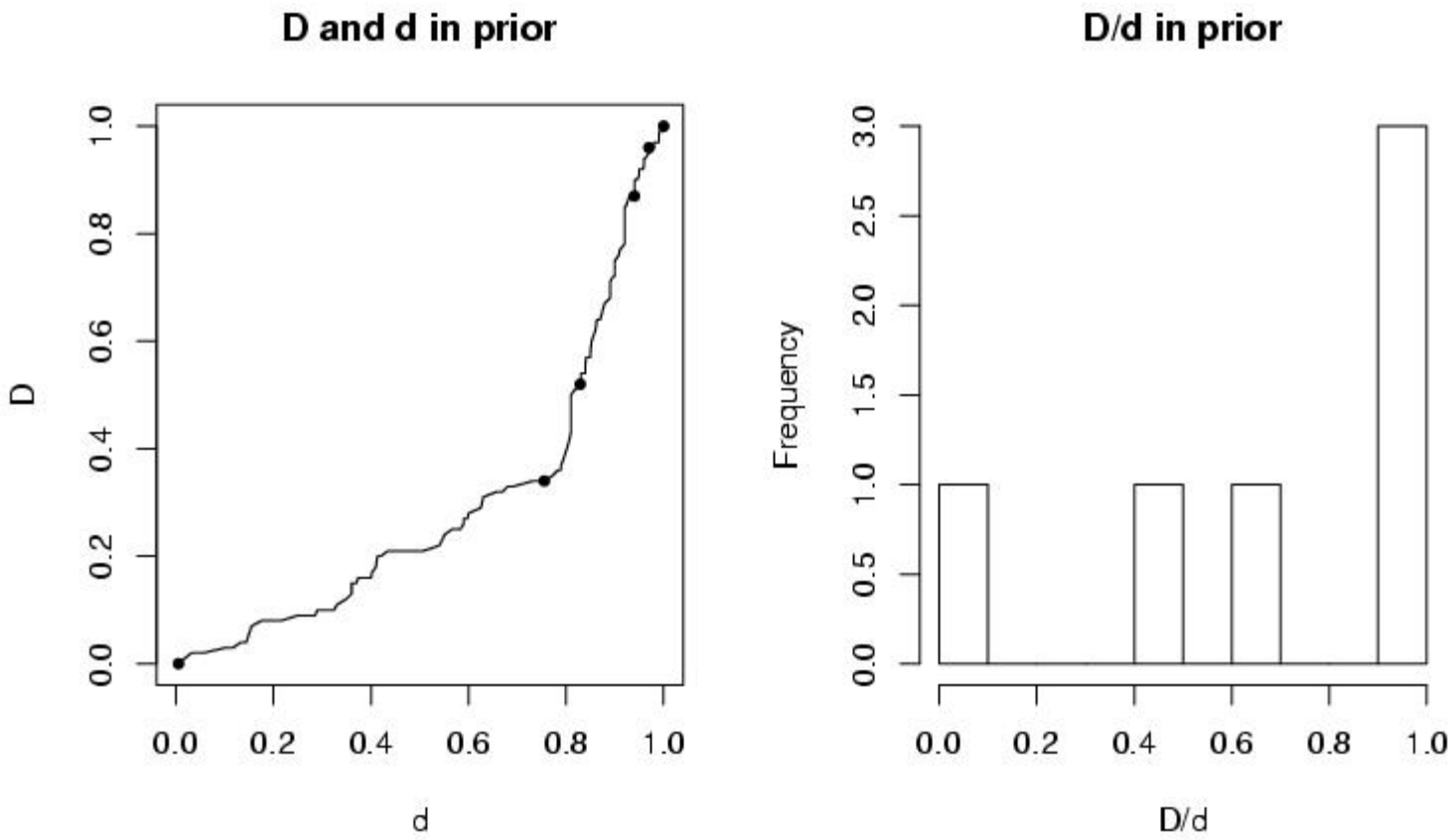
Potential Outcomes Joint Distribution – "Close to" EF's"

- $[Y_i(C)|\ d_i, \theta] \sim N(\beta_0 + \beta\ d_i, \sigma_C^2)$
- $[Y_i(Z_{D_i})|\ Y_i(C), d_i, \theta] \sim N[Y_i(C) + \gamma_1 Z_{D_i} + \gamma_2 Z_{D_i}^2 + \gamma_3 Z_{D_i} d_i, \sigma^2_{T\cdot C}]$, Mutually conditionally independent across the Z_{D_i}, -- plausible?
  And $\gamma_1 \geq 0$, $\gamma_2 \geq 0$, $\gamma_1 + \gamma_3 \geq 0$
  - When $Z_{D_i} = 0$, expectation of $Y_i(Z_{D_i}) - Y_i(C) = 0$
  - Dose-response is monotonely increasing for this range of doses

Prior Distribution on θ

- Prior distribution on parameters of Betas is specified by adding six "fake" observations with both Z_{D_i} and d_i observed on the equal percentile equating line and nothing else observed
  - These are the minimum, 25th percentile, median, 75th percentile, maximum
  - Purpose of these observations is simply to stabilize computation and has little influence on inference (n_{fake}=6 versus n_{real}>300) and are accurate in expectation because of the randomization
- Prior distribution on the rest of θ is independent and is the standard noninformative prior proportional to $1/(\sigma_C \sigma_{T\cdot C})$

Six Prior Data Points for $\pi(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$



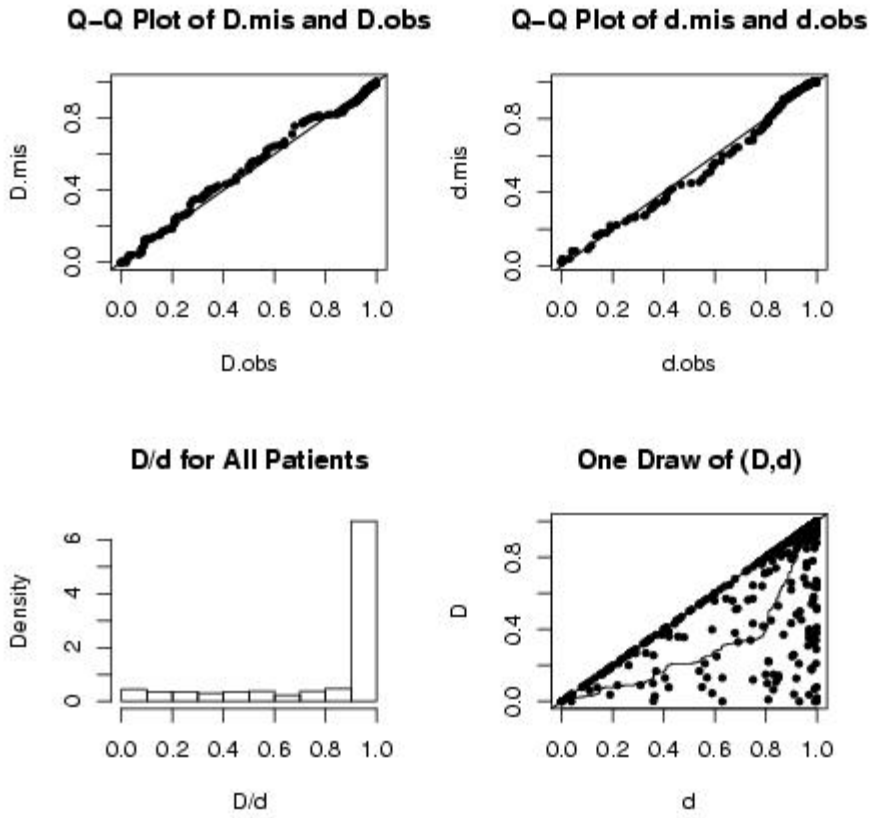
Figures from Jin and Rubin, 2008

JR's Computation

- Missing data problem, which is addressed using MCMC to draw Bayesian inferences
  - Parameters are θ
  - Key missing data are d_i for those assigned treatment and Z_{D_i} for those assigned control
  - Given θ, draw key missing data; given key missing data, draw θ; iterate until approximate convergence
  - Vast number of such draws approximates posterior distribution of dose-response as a function of principal strata defined by $d_i(C)$

Diagnostic Checks for JR's Model

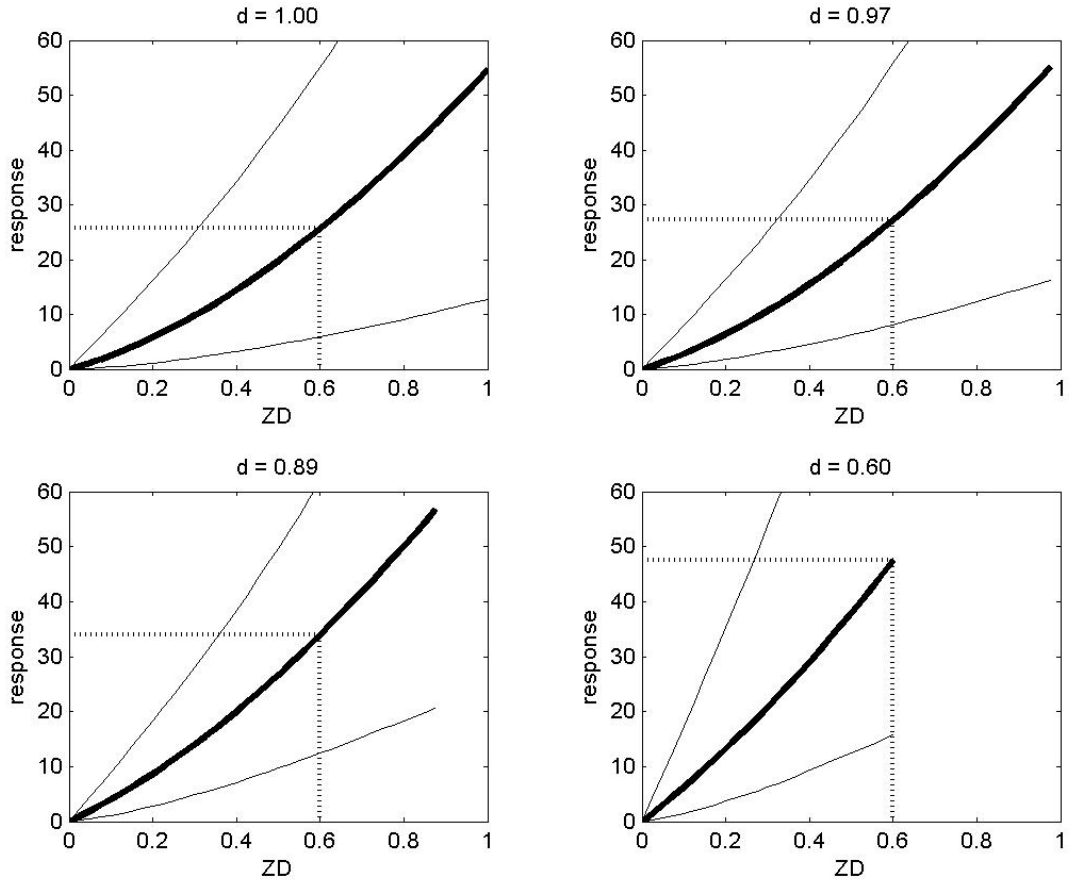
One Posterior Draw of Key Missing Data



Figures from Jin and Rubin, 2008

Dose-Response Results for Principal Strata

Maximum d, 75th d, median d, 25th d



Figures from Jin and Rubin, 2008

Discussion of the Dose Response Conclusions

- Under EF's assumptions, dose-response at each $d_i(C)$ is a point because $D_i(T) = F(d_i(C))$
- JR's dose-response results are causal under debatable assumption
  - Is "Nature's randomization" of dose given placebo compliance (i.e., the crucial latent ignorability assumption) plausible?
  - Or do we need to condition further on background medical characteristics related to possible side effects of the drug? Such sensitivity analysis is future work

Are Job-Training Programs Effective?

Donald B. Rubin

Harvard University

Presentation based on joint work with Fabrizia Mealli, Paolo Frumento, and Barbara Pacini.

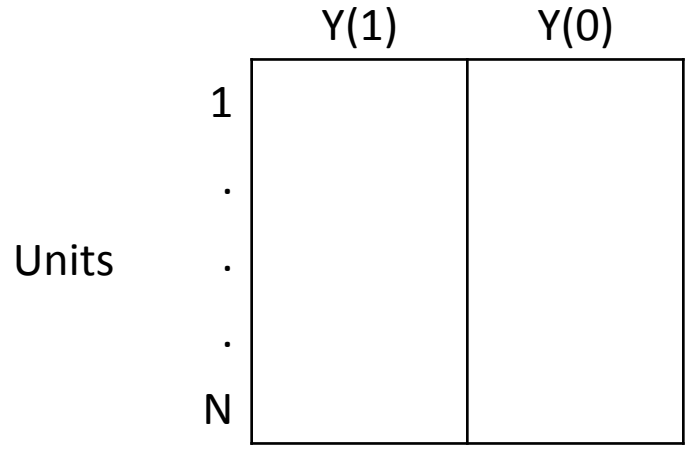
The National Job Corps Study

- A randomized study to evaluate the effects of a training program on employment and wages
  - Randomization assures fair comparison, in expectation, between treatment groups
  - Sampled youths (n=15,386) were assigned randomly to a job training program group or a control group
  - Only those assigned to the job training program group were able to enroll in Job Corps
- Post-treatment complications
  - Noncompliance (only 73% attended the offered training)
  - Truncation of wages for the unemployed
  - Missing outcomes due to nonresponse

Potential Outcomes Approach to Causal Inference – Simplest Setting

T = 1 $\Rightarrow$ active treatment (e.g., job training)

T = 0 $\Rightarrow$ control treatment (e.g., no training)



Y(1) = outcomes if exposed to active treatment

Y(0) = outcomes if exposed to control treatment

Ave[$Y_i(1) - Y_i(0)$] = Average causal effect of active versus control treatment

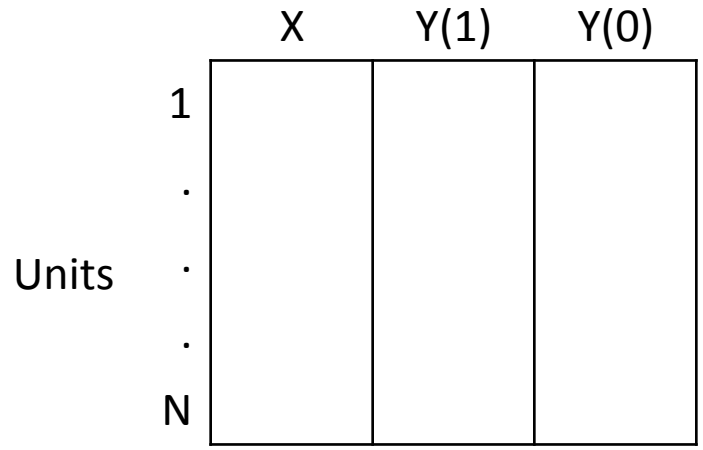
Potential Outcomes Approach to Causal Inference – Simplest Setting

- Fundamental problem of causal inference
  - For each *i*, only $Y_i(1)$ or $Y_i(0)$ can be observed

Units	Y(1)	Y(0)	T
1	✓	?	1
.	✓	?	1
.	✓	?	1
.	?	✓	0
.	?	✓	0
N	?	✓	0

- Random assignment of active versus control $\Rightarrow$ representative sample of $Y_i(1)$ will be compared to representative sample of $Y_i(0)$

Potential Outcomes Approach to Causal Inference – Simplest Setting with Covariates



- Same as before, except includes pretreatment covariates, e.g., age, sex, background education

$$\underset{i:\text{Female}}{\text{Ave}}\,[Y_i(1) - Y_i(0)] = \text{Average causal effect of active vs control treatment for females}$$

- Randomization still works for females

Potential Outcomes Approach to Causal Inference – Simple Noncompliance with Active Treatment

		D(1)	D(0)	Y(1)	Y(0)	
Units	1	1	0			compliers
	.	1	0			
	.	1	0			
	.	0	0			noncompliers
	N	0	0			

D(1) = treatment taken when assigned active treatment
D(1) = 1 ⇒ active taken, D(1) = 0 ⇒ control taken

D(0) = treatment taken when assigned control treatment
simple setting, always control ⇒ D(0) = 0

$$\underset{i:D_i(1)=1}{\text{Ave}}\,[Y_i(1) - Y_i(0)] = \text{Average causal effect for true compliers}$$

- Randomization still works for compliers

Potential Outcomes Approach to Causal Inference – Simple Noncompliance with Active Treatment: Observed Data

Units	D(1)	D(0)	Y(1)	Y(0)	T	
1	1	0	✓	?	1	complier status observed
.	1	0	✓	?	1	
.	?	0	?	✓	0	complier status missing
.	?	0	?	✓	0	
.	0	0	✓	?	1	noncomplier status observed
N	?	0	?	✓	0	noncomplier status missing

- **Compliers**
  - For individuals assigned treatment (T=1), D(1)=1 & D(0)=0
  - For individuals assigned control (T=0), D(1)=? because true compliance under treatment is unknown & D(0)=0
- **Noncompliers**
  - For individuals assigned treatment (T=1), D(1)=0 & D(0)=0
  - For individuals assigned control (T=0), D(1)=? because true compliance under treatment is unknown & D(0)=0
- **Randomization still works for compliers**

Key Idea: Principal Stratification
(Frangakis and Rubin, 2002)

- Stratify on values of post-treatment intermediate outcome
- Convert $D_i(1)$, $D_i(0)$ into stratification variable
  - True complier "c" if $D_i(1)=1$
  - Noncomplier "n" if $D_i(1)=0$
- Idea works more generally

Intermediate Outcome - Employment

- Employed (yes, no) at a given time post-treatment is an important outcome, but is also needed to define principal strata for "final" outcomes, Y, describing attributes of possible employment, such as wages, retirement plan benefits, etc., which are not well-defined if unemployed
- Principal strata are defined by employment status
  - EE = employed whether assigned to training or not
  - EU = employed if trained, unemployed if not trained
  - UE = unemployed if trained, employed if not trained
  - UU = unemployed whether assigned to training or not
- Causal effects of training on Y only well-defined for EE
- UE empty? Reservation wage issue

Causal Effects of Training within Principal Strata

- Principal strata are defined by compliance with assignment to job training and by employment status
  - c&EE, c&EU, c&UE, c&UU
  - n&EE, n&EU, n&UE, n&UU
- By assumption (exclusion restriction on employment), we rule out n&EU and n&UE
  - If assignment does not affect entry into training, assignment cannot affect employment status
  - Also assume exclusion for attributes of employment, Y
- Causal effects of T on Y are only well-defined for c&EE and n&EE principal strata (no effect on Y in n&EE by exclusion restriction)

Not Done Yet with Needed Principal Strata

- Indicators for response to survey items asking about employment status and wages, etc.
  - R(1) and R(0), each indicating respondent or not
- Do not make exclusion restriction here
  - e.g., males could have $R_i(1) \Longrightarrow$ respond if assigned training, but $R_i(0) \Longrightarrow$ not respond if assigned control
- But do assume missing at random (MAR)
  - A nuisance, not of scientific interest

Causal Effects

- Assignment to be trained on being job-trained
  Pr(c) = proportion compliers
- Assignment to be trained on being employed
  Pr(c&EU) – Pr(c&UE)
- Assignment to be trained on being employed for compliers
  [Pr(c&EU) – Pr(c&UE)]/Pr(c)
- Relative sizes of principal strata
  – c&EE, c&EU, c&UE, c&UU, n&EE, n&UU
- Distributions of X within principal strata

Causal Effects on Wages

- For the always employed

  Ave[$Y_i(1) - Y_i(0)|$c&EE or n&EE]

- For the always employed compliers

  Ave[$Y_i(1) - Y_i(0)|$c&EE]

- By exclusion, for the always employed noncompliers
  - Ave[$Y_i(1) - Y_i(0)|$n&EE] = 0

Method of Analysis

- Direct likelihood at each of three post-treatment points in time
- Search for parsimonious model to help guide policy
- Needs scientific judgement

Estimated Means of Covariates within Principal Strata Week 52

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Percent in Stratum	24	4	5	42	12	13
Female	0.4	0.3	0.3	0.4	0.4	0.5
Age at baseline	19.0	19.1	19.3	18.4	19.5	18.9
White	0.3	0.4	0.3	0.2	0.3	0.2
With a Partner	0.1	0.0	0.0	0.1	0.1	0.1
Has children	0.2	0.1	0.1	0.2	0.2	0.3
Education	0.4	0.4	0.5	0.3	0.6	0.4
Ever arrested	0.2	0.3	0.3	0.3	0.3	0.3
Mother's education	0.2	0.2	0.3	0.2	0.2	0.2
Father's education	0.3	0.3	0.3	0.3	0.3	0.3
Household income > $6000	0.6	0.6	0.5	0.5	0.6	0.5
Person income > $6000	0.1	0.1	0.0	0.1	0.2	0.1
Have job	0.3	0.3	0.3	0.2	0.3	0.2
Had job, previous year	0.8	0.8	0.7	0.6	0.8	0.6
Months in Job, previous year	5.0	5.1	4.9	2.9	5.6	3.1
Earnings, previous year	0.2	0.3	0.1	-0.2	0.4	-0.1

Estimated Means of Covariates within Principal Strata Week 130

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Percent in Stratum	31	7	5	32	13	13
Female	0.4	0.3	0.2	0.5	0.4	0.5
Age at baseline	18.9	18.9	18.9	18.4	19.5	18.9
White	0.3	0.4	0.4	0.2	0.3	0.2
With a Partner	0.1	0.1	0.1	0.0	0.1	0.1
Has children	0.2	0.2	0.2	0.2	0.2	0.3
Education	0.4	0.4	0.4	0.3	0.5	0.4
Ever arrested	0.2	0.3	0.3	0.3	0.3	0.3
Mother's education	0.2	0.2	0.3	0.2	0.2	0.2
Father's education	0.3	0.3	0.2	0.3	0.3	0.3
Household income > $6000	0.6	0.6	0.6	0.5	0.6	0.5
Person income > $6000	0.1	0.1	0.1	0.1	0.1	0.1
Have job	0.3	0.3	0.3	0.2	0.3	0.2
Had job, previous year	0.7	0.8	0.7	0.6	0.8	0.6
Months in Job, previous year	4.3	5.2	4.3	2.9	5.0	3.5
Earnings, previous year	0.1	0.3	0.2	-0.2	0.3	-0.0

Estimated Means of Covariates within Principal Strata Week 208

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Percent in Stratum	39	6	5	26	15	10
Female	0.4	0.3	0.2	0.5	0.4	0.5
Age at baseline	18.9	18.7	18.6	18.4	19.4	18.9
White	0.3	0.4	0.4	0.2	0.3	0.2
With a Partner	0.1	0.0	0.1	0.0	0.1	0.1
Has children	0.2	0.1	0.2	0.2	0.2	0.2
Education	0.4	0.4	0.4	0.3	0.5	0.4
Ever arrested	0.2	0.3	0.3	0.3	0.3	0.3
Mother's education	0.2	0.3	0.2	0.2	0.2	0.2
Father's education	0.3	0.3	0.2	0.3	0.3	0.3
Household income > $6000	0.6	0.7	0.7	0.5	0.6	0.5
Person income > $6000	0.1	0.1	0.1	0.0	0.1	0.1
Have job	0.3	0.3	0.2	0.1	0.3	0.2
Had job, previous year	0.7	0.7	0.7	0.5	0.7	0.6
Months in Job, previous year	4.5	4.4	4.1	2.6	4.9	3.5
Earnings, previous year	0.1	0.2	0.1	-0.3	0.2	-0.0

Percent within Principal Strata by Time Period

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Week 52	24	4	5	42	12	13
Week 130	31	7	5	32	13	13
Week 208	39	6	5	26	15	10

- For compliers, % EE increases in time, and % UU decreases
- For noncompliers, EE remains fairly stable
- Causal effect of training slightly increases in time, i.e., the difference between proportions in c&EU and c&UE appears to increase in time
  - Economists' "lock-in effect" during the period of training

Estimated Average Hourly Wages for Those Employed in Dollars within Principal Strata by Time Period

Principal Stratum	c&EE(1)	c&EE(0)	c&EU(1)	c&UE(0)	n&EE
Week 52	5.8	5.6	7.2	6.6	6.6
Week 130	6.7	6.5	9.3	7.5	8.0
Week 208	7.7	7.5	9.3	9.6	9.1

- Estimated causal effect on wages for always employed compliers is approximately 0.2 for all time periods
- Always employed compliers, whether trained or not, have the lower hourly wages than the sometimes employed (c&EU or c&UE) or n&EE
- Wages tend to increase in time

Final Conclusions for This Job Training Program

- In long run, for compliers, minor positive effect on employment status
- For always employed compliers, minor positive effect on wages at all time periods
- Background characteristics of individuals differ across principal strata
- Suggests need for more targeted programs
- Even if evaluation is based on randomized experiment, difficult to analyze correctly

Observed Values

- Covariates, X_i
- Treatment received, T_i
- Observed outcome:

$$Y_i^{obs} = T_i\, Y_i(1) + (1 - T_i) Y_i(0)$$

- Critical to understand the assignment mechanism
  - Reasons for $T_i = 1$ versus $T_i = 0$