THE DESIGN (VERSUS THE ANALYSIS) OF EVALUATIONS FROM OBSERVATIONAL STUDIES: PARALLELS WITH THE DESIGN OF RANDOMIZED EXPERIMENTS

DONALD B. RUBIN

My perspective on inference for causal effects:

In randomized experiments

In observational studies

The "RCM" (Holland, 1986): 1974, 1975, 1976, 1977, 1978, 1979, 1980

Two essential parts, one optional part(*)

1. Define estimands using potential outcomes – conceptual; pre data

Observable results of real or hypothetical manipulations (treatments)

"No causation without manipulation" 1975

Formal notation due to Neyman (1923) in randomized experiments

Intuitive idea old, but general use of formal notation only from 1974

$$X, Y(0), Y(1)$$

Prior to 1974 everyone used Y_{obs} in observational studies

$$Y_{obs,i} = W_i Y_i(1) + (1-W_i) Y_i(0)$$

2. Posit assignment mechanism

- Reasons for missing and observed values of $Y(0)$ and $Y(1)$

$$\Pr(W|X, Y(0), Y(1))$$

- Special cases much discussed

  Randomized experiments – Neyman, Fisher (1925)

  Supply-Demand – Haavelmo (1944)

  Optimizing Behavior – Roy (1951)

- General formulation only from 1975

*3. Full probability model to predict the missing potential outcomes

$$\Pr(X, Y(0), Y(1))$$

Bayes formalized in 1975, 1978

#2 and #3 $\Rightarrow$ Posterior distribution of all causal estimands

Here, focus on the second step: design in observational studies.

- Template is generalized randomized experiment:

$$\Pr(W, X, Y(0), Y(1)) = \Pr(W|X)$$

$$\propto \prod \Pr(W_i = 1 | X_i) = \prod \text{ propensity scores}$$

$$0 < \text{propensity scores} < 1$$

- Design of randomized experiments is without $Y(0)$, $Y(1)$
- Approximate randomized experiment in observational studies
- Blocking, matching subclassification – just use X
- Old history of matching in statistics – summarized in 2006.
- Structure an observational study as a "broken"

  randomized experiment, without using any values of potential outcomes
- Try to reconstruct underlying experiment

FDA analogy

- Pharmas do randomization over and over until they get a favorable result?
- Or use regression over and over on one sample until drug looks good?
- Instead, they carefully try to design study, followed by
    - Specification of primary analysis & secondary supporting analyses
    - And must live with answers that follow from proposed analyses

Should we accept lower standards in social science?

- Objectivity can be attained in observational studies
- Outcome-free design before any analysis with observed outcomes

Very Recent Personal Examples:

- Tobacco litigation
- Value Added Assessment
- Peer effects of smoking
- Job training in Germany Treffer
- Illustrate with first example

  "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." Health Services & Outcomes Research Methodology (2002).

Table 1

Variables Used in Propensity Model	Description
Seatbelt	5 levels of reported seat belt use
Arthritis	Whether reported suffering from arthritis
Census Division	9 census regions
Champ Insurance	
Diabetes	Doctor ever told having diabetes
Down time	6 levels of reported Down time
Dump time	6 levels of reported Dump time
Employment	Indicating employment status each quarter
English	English is a primary language
Retirement	Indicator for Retirement status
Number of Friends	7 levels measuring the number of friends
Membership in Clubs	6 levels measuring memberships in clubs
Education	Completed years of Education
HMO coverage	Indicating HMO coverage each quarter
High blood pressure	Doctor ever told having high blood pressure
Industry Code	14 Industry codes
Age	Age of the respondent
Labor Union	Indicator for a member of labor union
Log Height	Natural Logarithm of Height
Log Weight	Natural Logarithm of Weight
Marital Status	Marital status in each quarter
Medicaid	On medicaid (each quarter)
Medicare	On medicare (each quarter)
Occupation	Occupation code (13 levels)
Public Assistance	Other Public assistance program (each quarter)
Friends over	Frequency of having Friends over (7 levels)
Physical Activity	Indicator Variable for Physically active
Population density	3 levels
Poverty Status	6 levels
Pregnant 1987	Pregnancy status in 1987 (women)
Private Insurance	Other Private Insurance (each quarter)
Race	4 levels
Rated Health	5-point Self rating of health status
Home ownership	Indicator for Owning home
Rheumatism	Indicator for Suffering from rheumatism
Share Life	Indicator variable for having somebody to share their life

Region	4 levels of region of the country
MSA	4 levels indicating types of Metropolitan Statistical Area
Risk	General risk taking attitude (5 levels)
Uninsured	Indicator for lack insurance (each quarter)
Veteran	Indicator for Veteran status
Incapler	Survey Weight in NMES database
Agesq	Age*Age
Educat.sq	Education*Education
Age_wt	Age*logwt
Age_educt	Age*Education
Age_ht	Age*Loght
Educat_wt	Education*Logwt
Educat_ht	Education*loght
Loght_logwt	Loght*logwt
Loghtsq	Loght*loght
Logwtsq	Logwt*logwt

TABLE 2

ESTIMATED PROPENSITY SCORES ON THE LOGIT SCALE FOR "SMOKERS" VERSUS NEVER SMOKERS IN FULL NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Treated Group	B	R	≤1/2	>1/2 and ≤ 4/5	>4/5 and ≤ 5/4	>5/4 and ≤ 2	>2
Male Current N = 3,510	1.09	1.00	3	9	57	26	5
Male Former N = 3,384	1.06	0.82	2	15	61	15	7
Female Current N = 3,434	1.03	0.85	1	15	59	23	2
Female Former N = 2,657	0.65	1.02	.5	7	85	7	.5

TABLE 3

ESTIMATED PROPENSITY SCORES FROM FULL NMES ON THE LOGIT SCALE FOR "SMOKERS" VERSUS NEVER SMOKERS IN MATCHED NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Treated Group	B	R	$\leq$1/2	>1/2 and $\leq$ 4/5	>4/5 and $\leq$ 5/4	>5/4 and $\leq$ 2	>2
Male Current N = 3,510	.08	1.16	1	3	90	6	0
Male Former N = 3,384	.04	0.99	1	1	94	3	1
Female Current N = 3,434	.04	0.94	1	1	93	5	0
Female Former N = 2,657	.06	1.02	0	2	91	7	0

TABLE 4

ESTIMATED PROPENSITY SCORES ON THE LOGIT SCALE FOR "SMOKERS" VERSUS NEVER SMOKERS IN MATCHED NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Treated Group	B	R	≤1/2	>1/2 and ≤ 4/5	>4/5 and ≤ 5/4	>5/4 and ≤ 2	>2
Male Current N = 3,510	0.39	1.33	0	4	88	8	0
Male Former N = 3,384	0.32	1.33	0	1	95	3	1
Female Current N = 3,434	0.35	1.18	1	1	92	6	0
Female Former N = 2,657	0.31	1.09	0	2	91	7	0

TABLE 5A

PROPENSITY SUBCLASSIFICATION ANALYSES ON THE LOGIT SCALE FOR CURRENT VS. NEVER-SMOKER MALES IN MATCHED NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Number of Subclasses	B	R	$\leq 1/2$	>1/2 and $\leq$ 4/5	>4/5 and $\leq$ 5/4	>5/4 and $\leq$ 2	>2
2	0.18	1.36	0	2	98	0	0
4	0.10	1.25	0	1	99	0	0
6	0.09	1.30	0	0	100	0	0
8	0.08	1.16	0	0	100	0	0
10	0.07	1.12	0	0	100	0	0

TABLE 5B

PROPENSITY SUBCLASSIFICATION ANALYSES ON THE LOGIT SCALE FOR FORMER VS. NEVER-SMOKER MALES IN MATCHED NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Number of Subclasses	B	R	$\leq 1/2$	$>1/2$ and $\leq 4/5$	$>4/5$ and $\leq 5/4$	$>5/4$ and ≤ 2	>2
2	0.16	1.38	0	1	98	1	0
4	0.09	1.32	0	1	98	1	0
6	0.07	1.30	0	0	99	1	0
8	0.07	1.37	0	0	99	1	0
10	0.07	1.31	0	0	99	1	0

TABLE 5C

PROPENSITY SUBCLASSIFICATION ANALYSES ON THE LOGIT SCALE FOR CURRENT VS. NEVER-SMOKER FEMALES IN MATCHED NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Number of Subclasses	B	R	$\leq 1/2$	$>1/2$ and $\leq 4/5$	$>4/5$ and $\leq 5/4$	$>5/4$ and ≤ 2	>2
2	0.14	1.26	0	1	98	1	0
4	0.08	1.44	0	1	99	0	0
6	0.06	1.69	0	1	99	0	0
8	0.05	1.69	0	0	100	0	0
10	0.05	1.70	0	0	100	0	0

TABLE 5D

PROPENSITY SUBCLASSIFICATION ANALYSES ON THE LOGIT SCALE FOR FORMER VS. NEVER-SMOKER FEMALES IN MATCHED NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Number of Subclasses	B	R	$\leq 1/2$	$>1/2$ and $\leq 4/5$	$>4/5$ and $\leq 5/4$	$>5/4$ and ≤ 2	>2
2	0.13	1.09	0	0	97	3	0
4	0.08	0.85	0	0	99	1	0
6	0.07	0.85	0	0	100	0	0
8	0.06	0.77	0	0	100	0	0
10	0.06	0.92	0	0	100	0	0

TABLE 6

WEIGHTED PROPENSITY SCORE ANALYSES BASED ON INVERSE PROBABILITIES FOR WEIGHTS IN MATCHED NMES

B = Bias, R = Ratio of "smoker" to never-smoker variances;
also displayed is the distribution of the ratio of variances
in the covariates orthogonal to the propensity score

Treated Group	B	R	Percent of covariates with specified variance ratio after adjustment for propensity score ≤1/2	>1/2 and ≤ 4/5	>4/5 and ≤ 5/4	>5/4 and ≤ 2	>2
Male Current N = 3,510	0.03	1.19	0	0	100	0	0
Male Former N = 3,384	0.08	0.22	0	0	100	0	0
Female Current N = 3,434	0.03	1.70	0	0	100	0	0
Female Former N = 2,657	0.03	0.66	0	0	100	0	0

Table 7

ESTIMATED PROPENSITY SCORES ON THE LOGIT SCALE FOR MALE CURRENT SMOKERS (N = 3510) VERSUS MALE NEVER SMOKERS IN NMES

B = Bias, R = Ratio of "smoker" to never-smoker variance;
also displayed is the distribution of the ratio of variances
in the covariates after adjusting for the propensity score

			Percent of covariates with specified variance ratio after adjustment for propensity score				
Analysis	B	R	$\leq 1/2$	>1/2 and $\leq 4/5$	>4/5 and $\leq 5/4$	>5/4 and ≤ 2	>2
Full N = 4,297	1.09	1.00	3	9	57	26	5
Matched N = 3,510	0.08	1.16	1	3	90	6	0
Matched							
K = 1	0.39	1.33	0	4	88	8	0
K = 2	0.18	1.36	0	2	98	0	0
K = 4	0.10	1.25	0	1	99	0	0
K = 6	0.09	1.30	0	0	100	0	0
K = 8	0.08	1.16	0	0	100	0	0
K = 10	0.07	1.12	0	0	100	0	0
K = ∞	0.03	1.19	0	0	100	0	0