500 Class 10

https://thomaselove.github.io/500-2025/

2025-03-27

Agenda for the Slides

- Some key takeaways from Posner 2001
- Discussion of Rosenbaum **Causal Inference** Chapter 9

Comparing Standard Regression, Propensity Score Matching, and Instrumental Variables Methods for Determining the Influence of Mammography on Stage of Diagnosis

MICHAEL A. POSNER*
Data Coordinating Center, Boston University School of Public Health, Boston, Massachusetts
E-mail: mposner@bu.edu

ARLENE S. ASH, KAREN M. FREUND, AND MARK A. MOSKOWITZ†
Health Care Research Unit, Section of General Internal Medicine, Evans Department of Medicine, Boston Medical Center, Boston, Massachusetts
E-mail: aash@bu.edu; karen.freund@bmc.org

MICHAEL SHWARTZ
Operations Management, Boston University School of Management, Boston, Massachusetts
E-mail: mshwartz@bu.edu

Goals of Posner et al. 2001

- Mammography screening and its effectiveness in detecting cancer at an earlier stage.

Compare results of three analytic approaches:

1. Standard (regression-based) adjustment for baseline risk plus a treatment indicator
2. Propensity score matching to account for selection bias through evening out covariate distributions
3. Instrumental variables to address unmeasured differences between treated and untreated patients

The Research Question

Use of mammography for screening women over age 70; as of 2001, the value hasn't been established

- Most RCTs of mammography include no women over age 70 (focus is on the 50-70 year olds)
- No RCT has reported age-specific data within the 50-70 age group so that trends can be studied
- Breast cancer incidence continues to rise beyond age 65 - 48% of new cases are > 65.

The Data

Linked Medicare - Tumor Registry Database

- Sample consisted of all women with a first diagnosis of breast cancer

  ...

  - In one of three regions (metropolitan Atlanta, state of Connecticut, or Seattle-Puget Sound)
  - whose utilization of mammography could be tracked for the 2 years prior to the diagnosis of breast cancer
  - who were either regular mammography users or mammography non-users (excluded "tweeners")

2. Data Description

The database we utilized for this cohort study is the Linked Medicare-Tumor Registry Database. The linked database was jointly created by the National Cancer Institute (NCI) and the Centers for Medicare and Medicaid Services (CMS), formerly the Health Care Financing Administration (HCFA) (Potosky, Riley and Lubitz et al., 1993). The database links Medicare data on women ages 65 and older from 1985 to 1994 with cancer registry information from the NCI's SEER program for cancers diagnosed between 1973 and 1993. The two databases overlap in three racially and socially diverse geographic areas: metropolitan Atlanta, Seattle-Puget Sound, and the state of Connecticut.

Treatment Variable

- Regular mammography users had claims for two separate bilateral mammograms within the two years prior to their breast cancer diagnosis, which were at least 10 months apart.
- Non-users were women with no mammography claims in the two years prior to their diagnosis.

Primary Outcome

Stage at diagnosis, dichotomized

- Early (in situ, or Stage I)
- Late (Stage II, III or IV)

Excluded the 7.4% of women with unstaged cancer

Covariates

- Age at diagnosis
  - Categorical: 67-69, 70-74, 75-79, 80-84, 85+
- Comorbidity (Charlson Comorbidity Index)
- Race (black vs. other)
- Income (median income of patient's zip code)
  - Dichotomized to highest 40% vs. lowest 60% of incomes within each region
- # of claims for primary-care office visits over the last two years (also categorized)

Approach 1: Risk Adjustment

Developed a logistic regression model to predict stage at diagnosis (early or late) from user status (regular user or non-user), controlling for:

- Region, Age, Race, Comorbidity, Median income [zip code], Primary care visits

Conclusion

Regular users have **2.97** times the odds of being diagnosed at an early stage relative to non-users (95% CI: 2.56, 3.45)

Approach 2: Propensity "Matching" (sort of)

Propensity model included same covariates as risk adjustment model:

- Region, Age, Race, Comorbidity, Median income [zip code], Primary care visits

Steps:

1. Split data into deciles based on propensity score
2. Within each decile, take a random sample from the larger group (users or non-users) to get the same number as in the smaller group
3. Matched sub-samples combined to yield final data set

I'd call this "Stratification" more than "Matching"

Propensity "Matching" inside Deciles

Decile	Non-Users	Users	*Matched* Non-Users	*Matched* Users
1	416	57	57	57
2	339	89	89	89
3	359	136	136	136
4	239	205	205	205
5	193	289	193	193
6	159	277	159	159
7	145	347	145	145
8	96	327	96	96
9	113	394	113	113
10	81	395	81	81

Covariate Balance Pre- and Post-"Matching"

Variable	Pre-match p	Post-match p
Age at diagnosis	0.001	0.98
Comorbidity Index	0.001	0.73
Race	0.001	0.35
Income	0.061	0.49
Primary Care Visits	0.001	0.51
Location (Region)	0.001	0.98

- And, looking at our outcome ...

Variable	Pre-match p	Post-match p
Stage of Cancer	0.001	0.001

Table 1. Propensity score matching results

	Pre-Matching			Post-Matching		
	Non-User	User	p-value	Non-User	User	p-value
Total Sample	2140	2516		1274	1274	
Decile 1	416	**57**		57	**57**	
Decile 2	339	**89**		89	**89**	
Decile 3	359	**136**		136	**136**	
Decile 4	239	**205**		205	**205**	
Decile 5	**193**	289		**193**	193	
Decile 6	**159**	277		**159**	159	
Decile 7	**145**	347		**145**	145	
Decile 8	**96**	327		**96**	96	
Decile 9	**113**	394		**113**	113	
Decile 10	**81**	395		**81**	81	
Age at Diagnosis	77.2	74.5	0.001	75.49	75.31	0.389
Age at Diagnosis						
67–69	14.3%	20.1%		16.2%	16.9%	
70–74	26.9%	35.6%		30.7%	29.7%	
75–79	24.3%	26.2%		27.8%	27.6%	
80–84	17.1%	13.6%		16.9%	17.1%	
85+	17.4%	4.6%	0.001	8.5%	8.7%	0.975
Charlson Comorbidities						
Not Hospitalized	27.0%	29.7%		27.4%	28.7%	
Hosp, No Comorbidies	48.0%	54.2%		51.1%	50.8%	
At Least One Comorbidity	24.9%	16.1%	0.001	21.5%	20.6%	0.726

Race						
Black	6.3%	3.1%		6.3%	5.4%	
Non-Black	93.8%	96.9%	0.001	93.7%	94.6%	0.350
Income (Median of Zip Code)	$42,030	$41,137	0.061	$41,073	$40,650	0.488
Regional Income						
Top 40%	40.2%	43.3%		41.8%	40.3%	
Lower 60%	59.8%	56.7%	0.036	58.2%	59.7%	0.444
Primary Care Visits	4.9	10.5	0.001	7.15	8.08	0.004
Primary Care Visits						
None	37.4%	6.5%		13.4%	12.8%	
1–3	22.3%	15.1%		24.3%	24.4%	
4–12	27.1%	47.3%		41.1%	39.2%	
13+	13.2%	31.1%	0.001	21.3%	23.6%	0.506
Location						
Seattle	27.0%	41.2%		32.0%	32.1%	
Atlanta	20.9%	14.6%		20.0%	20.3%	
Connecticut	52.1%	44.2%	0.001	48.0%	47.7%	0.984
Stage						
Early (TNM 0 or I)	58.9%	81.1%		58.3%	81.6%	
Late (TNM II, III, or IV)	41.1%	19.0%	0.001	41.7%	18.5%	0.001

Results from Propensity Analysis

- Most extreme propensity scores were examined, and were close to the others, so no pairs were excluded on that basis.
- Balance dramatically improved (in terms of significance) for all variables.

Conclusion

Regular users have **3.24** times the odds of being diagnosed at an early stage relative to non-users.

- 95% CI for odds ratio: (2.69, 3.88)

  [The propensity] approach estimates the impact of being a user of mammography for the population whose measured covariates conform to the matched sample ... This result being so close to that of the standard model provides some reassurance that the standard model has adjusted correctly for any differences in measured covariates between the user and non-user groups.

Approach 3: Instrumental Variables

Which variable to use as the instrument? We need:

1. An association between the instrument and the exposure (must predict user status)
2. **AND** a lack of correlation between the instrument and the unmeasured covariates that are associated with the outcome.
    - no residual predictive power on stage at diagnosis, after controlling for the other covariates in the model

Region as the Instrument

Trichotomous variable (Atlanta, Seattle, Connecticut)

1. Is there an association between region and use of mammography?
   - Literature suggests that there is.
   - These data seem to back the claim up.

Region as Instrument?

2. Is there no correlation between region and the unobserved covariates associated with the outcome (once we've adjusted for observed covariates in the model)?
    - Cannot test this statistically.
    - "Seems reasonable" that outcome for someone using mammography in one region shouldn't differ from outcome for someone of similar characteristics using mammography in another.

The Detailed Argument

- We have to agree that we would expect that a woman with certain characteristics (age, race, etc.) receiving regular screening in Seattle would have the same likelihood of early stage disease diagnosed from mammography had she lived in Atlanta or Connecticut.
- If this is not true, implies that follow-up after a positive mammogram differs by region.

Two-Stage Model for Instrumental Variables Approach

1. Predict user status using covariates and the instrument(s).
   - Obtain predicted probability of mammography use for each subject.
2. Predict stage at diagnosis (early or late) using the usual covariates (not including the instrument) along with the predicted probability of mammography use (instead of actual user status).

Instrumental Variable Results

- Precision will be drastically reduced from what we've seen in the previous analyses.
  - Replacing 0/1 user status with a prediction that can vary across (0, 1).

Conclusion

Regular users have **3.01** times the odds of being diagnosed at an early stage relative to non-users.

- 95% CI for odds ratio: (1.09, 8.34)

Comparison of Approaches

We start with the **standard analysis**, a logistic regression predicting stage at diagnosis that includes as independent variables a set of covariates to adjust for differences in baseline risk plus an indicator variable for whether the woman used screening. Next, we employ **propensity score matching**, which evens out the distribution of measured baseline characteristics across groups, and is more robust to model misspecification than the standard analysis. Lastly, we conduct an **instrumental variable** analysis, which addresses unmeasured differences between the users and non-users.

Approach	*OR*	95% CI
Risk Adjustment	2.97	2.56, 3.45
Propensity "Matching"	3.24	2.69, 3.88
Instrumental Variable	3.01	1.09, 8.34

OR = odds of regular users being diagnosed at an early stage relative to non-users

Table 2. Odds ratios (and 95% Confidence Intervals) from the three models to predict early stage at diagnosis

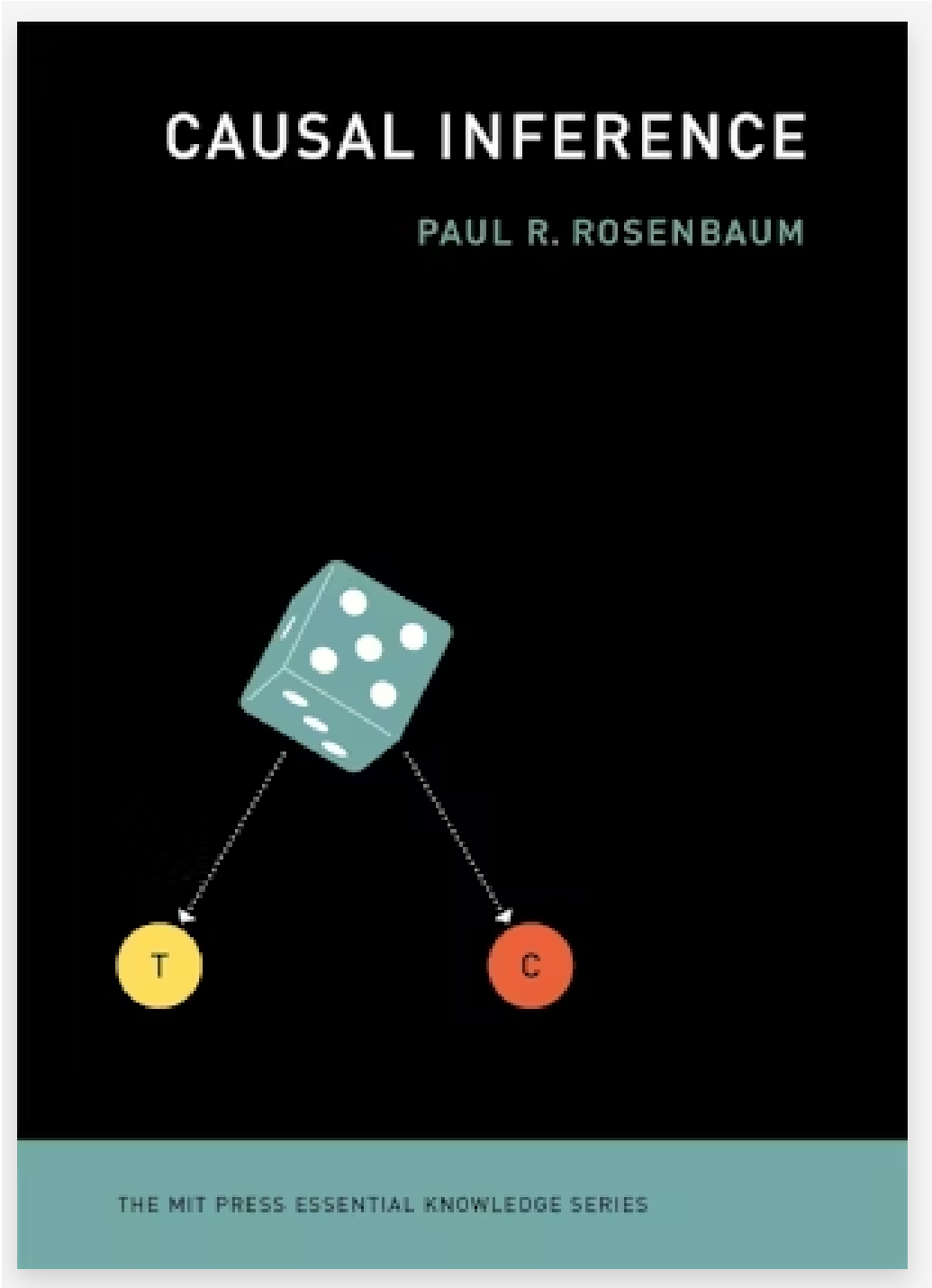
Variable	Standard Model	Propensity Score Matching	Instrumental Variable Analysis
Age, 67–69	1.00 (ref)	1.00 (ref)	1.00 (ref)
Age, 70–74	0.92 (0.75, 1.13)	1.07 (0.81, 1.40)	0.92 (0.75, 1.13)
Age, 75–79	0.93 (0.75, 1.15)	0.97 (0.74, 1.28)	0.93 (0.75, 1.16)
Age, 80–84	0.83 (0.66, 1.05)	0.92 (0.68, 1.25)	0.82 (0.63, 1.07)
Age, 85+	1.02 (0.79, 1.32)	1.31 (0.89, 1.93)	1.02 (0.70, 1.49)
Black	0.67 (0.50, 0.92)	0.70 (0.48, 1.02)	0.70 (0.51, 0.98)
Not Hospitalized	1.00 (ref)	1.00 (ref)	1.00 (ref)
No Comorbid	0.54 (0.45, 0.63)	0.56 (0.45, 0.69)	0.55 (0.47, 0.65)
Some Comorbid	0.48 (0.40, 0.59)	0.54 (0.41, 0.70)	0.50 (0.40, 0.62)
High Reg Inc	1.23 (1.08, 1.41)	0.36 (1.13, 1.64)	1.22 (1.07, 1.41)
Connecticut	1.00 (ref)	1.00 (ref)	–
Seattle	1.23 (1.06, 1.43)	1.29 (1.05, 1.59)	–
Atlanta	1.17 (0.97, 1.41)	1.07 (0.84, 1.37)	–
No PC Visits	1.00 (ref)	1.00 (ref)	1.00 (ref)
Seldom (1–3 vis)	0.77 (0.62, 0.94)	0.82 (0.61, 1.12)	0.76 (0.55, 1.04)
Often (4–12 vis)	0.97 (0.80, 1.17)	0.98 (0.74, 1.31)	0.95 (0.58, 1.58)
Very Often (13+)	0.79 (0.64, 0.98)	0.93 (0.67, 1.27)	0.80 (0.45, 1.45)
User of Mammo.	**2.97 (2.56, 3.45)**	**3.24 (2.69, 3.88)**	**3.01 (1.09, 8.34)**

Posner et al. Conclusions (1/2)

In summary, all three analyses - the standard regression, the propensity score matching, and the instrumental variable analysis using region as the instrument - produced very similar results. The similarity of these results helps strengthen the credibility of the standard regression analysis. There is little model mis-specification, either from measured variables, as seen via the propensity score matching, nor from unmeasured variables (that meet the instrumental variable criteria), as seen via the instrumental variable analysis.

Posner et al. Conclusions (2/2)

We recommend that investigators analyzing administrative databases or other observational studies consider the sources of bias that may affect their results. ... It is important to look beyond the standard analysis and to consider propensity score matching when there is concern about group differences in measured covariates and instrumental variable analysis when there is concern about differences in unmeasured covariates.


CAUSAL INFERENCE
PAUL R. ROSENBAUM
T
C
THE MIT PRESS ESSENTIAL KNOWLEDGE SERIES

Rosenbaum Chapter 9

- Are Small Daily Doses of Alcohol Beneficial?
- Oncologists vs. Cardiologists
- A dissenting voice from a new tactic: Mendelian Randomization
- The Answer might be Complex; A Traditional Toxin; Total Mortality
- Is Part or All of the Supposed Heart Benefit Simply a Mistake?

So what do we do?

- What was the most important thing?
- What was the muddiest, most confusing thing?

Reminders for Next Week (for Class 11)

1. OSIA presentations, group 3
2. Skim Lehr et al. 2019 (or 2020 - the date is a little unclear)
3. Next week will also be our final opportunity to ask and answer questions about the **Project** before our presentations begin in Classes 12-14.