Connecting the dots: Statistics, causality, and policy

Frank Yoon, Ph.D.
Statistician
Brief outline

- **Statistics and causal inference** (I’ll talk)
  - Background and review
  - Definitions and concepts
  - Formalize *statistical intuition*

- **Case studies** (We’ll talk)
  - Charter schools, job training, health insurance
  - Weigh the evidence critically
  - Strengthen *statistical intuition*

- **A hypothetical policy study** (You’ll talk)
  - Exercise *statistical intuition*
What is statistics?

- Statistics is the discipline that concerns itself with the study of the nature of *data*.

- Elementary concepts
  - Distributions
  - Averages, variation, correlation

- Associations are described by *statistical models*.
  - Are student outcomes better in charter schools?
  - Do trained job seekers have more success than others?
  - How do health costs change after coverage expansion?
What is causality?

- Causation is the relationship between an event (*cause*) and a second event (*effect*), where the second event is understood as a consequence of the first.

- Inherently, the study causality concerns itself with logic.

- Causality is described by *causal statements*.
  - Do charter schools improve student outcomes?
  - Does job training benefit job seekers?
  - How does coverage expansion impact costs?
“Correlation is not causation”

**Statistical**
- Are student outcomes better in charter schools?
- Do trained job seekers have more success than others?
- How do health costs change after coverage expansion?

**Causal**
- Do charter schools improve student outcomes?
- Does job training benefit job seekers?
- How does coverage expansion impact costs?

A → B
Statistics and causal inference

Photo credit: http://auracle.ca/news/
A simple study
A simple study
A simple study
A simple study
A simple study
A simple study
The Rubin Causal Model

A study unit has two potential outcomes.

- What would have happened under intervention
- What would have happened under control

The causal effect is the difference between the two.

Fundamental Problem of Causal Inference

We only get to see one outcome for a given unit.
Causal framework

- “No causation without manipulation”
- How are units assigned to intervention?
  - **Randomization**
    - Independence: the gold standard
    - Addresses all sources of bias
  - **Non-randomized assignment**
    - Decisions were made about intervention
    - Observed and unobserved biases
Study sample
Study sample
Randomized experiment
Randomized experiment
Observational study
Observational study
Observational study
Observational study
Can we attribute the observed difference between the two groups to the intervention?

Randomization guarantees internal validity.

In an observational study adjustments are needed.
  – Assume no hidden biases
    • *Strongly ignorable treatment assignment*
    • How big is this assumption?
  – Remove observed biases
    • Regression
    • Matching
External validity

- What would happen outside study conditions?

versus
Bias matters

- Internal validity is needed to claim that the intervention works in the study.

- External validity is needed to claim that the intervention would work elsewhere.

- A good evidence base has
  - Lots of studies with internal validity…
  - Representing lots of real-world conditions.
What works for whom?

- The **average effect** is often reported.
  - But does it represent everyone?
What works for whom?

- **Heterogeneity**
  - In a study the impacts of intervention vary along the study units, whose characteristics matter.
Weighing the evidence

- Internal validity is a must.
- External validity makes the result more useful.
- Usually, the average doesn’t describe it all.
Weighing the evidence

Randomized study

Observational study

The real world
What is policy?
Case studies

- The effect of charter-school management organizations on student achievement
- The effect of voucher models on employment and earnings
- The effect of mental health parity on costs and service use
Charter-school management organizations (CMOs) establish and operate multiple charter schools.

The National Study of CMO Effectiveness
- Middle-school outcomes
  - Mathematics and reading test scores after 1-3 years
  - Social studies and science test scores after 3 years
- Data sources
  - Student-level administrative records
  - 2010 Survey of CMO and District Principals
Charter-school management organizations

- Intervention: students enrolled in CMO schools
  - Selected by lottery

- Control: students enrolled in the same public school district
  - Matched on: pre-entry test scores, race, gender, free/reduced-price lunch status
Charter-school management organizations

Figure II.13. Incoming Reading and Math Scores for CMO Students and Their District Peers
Charter-school management organizations

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CMO</td>
<td>CMO</td>
</tr>
<tr>
<td>N</td>
<td>-0.33</td>
<td>-0.27</td>
</tr>
<tr>
<td>C</td>
<td>0.00</td>
<td>-0.01</td>
</tr>
<tr>
<td>H</td>
<td>0.44</td>
<td>0.36</td>
</tr>
<tr>
<td>E</td>
<td>0.43</td>
<td>0.46</td>
</tr>
<tr>
<td>D</td>
<td>-0.43</td>
<td>-0.37</td>
</tr>
<tr>
<td>Q</td>
<td>-0.24</td>
<td>-0.22</td>
</tr>
<tr>
<td>T</td>
<td>-0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>L</td>
<td>0.11</td>
<td>0.23*</td>
</tr>
<tr>
<td>P</td>
<td>-0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>M</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Charter-school management organizations

Figure IV.1. Distribution of Test-Score Effect Sizes After Two Years in Math

Figure IV.2. Distribution of Test-Score Effect Sizes After Two Years in Reading
In 2000 the Workforce Investment Act (WIA) replaced the Job Partnership Training Act.

- WIA aimed to transform the employment and training system by streamlining service delivery and providing universal access to services.
- It promoted customer choice, attempted to more fully engage businesses, and changed service provided to youth.

Training programs financed through Individual Training Accounts (ITAs).
The Workforce Investment Act

- DOL launched the ITA experiment in 1999 to provide state and local workforce agencies with systematic assessment of alternative approaches to ITAs and for estimating impacts.

- 8,000 training-eligible WIA customers in 8 sites

- Randomly assigned to 1 of 3 ITA approaches varying on intensity of counseling, authority of counselor over customer’s choice and over ITA amount
The Workforce Investment Act

- **Approach 1: Structured Customer Choice.** This most directive approach required customers to receive intensive counseling, and counselors had considerable discretion to customize the amount of the ITA investment. On one hand, counselors were expected to constrain customers by steering them to training with a high expected return, and they could reject customers’ choices that did not fit this criterion. On the other hand, counselors also had much greater discretion to set higher ITA amounts (up to a maximum of $8,000 in most sites) if they felt expensive training was a sound investment for certain customers.

- **Approach 2: Guided Customer Choice.** This approach, similar to what most workforce agencies adopted in the transition to WIA, involved mandatory counseling. However, counseling was less intensive than under the preceding approach. Counselors could not reject customers’ choices if the chosen provider was on the state’s approved list. The amount of the ITA award was fixed at $3,000 to $5,000, depending on the site.

- **Approach 3: Maximum Customer Choice.** This approach, the least structured of the three, did not require customers to participate in counseling after being found eligible for WIA-funded training, but they could request and receive it. Customers received a fixed ITA award of $3,000 to $5,000, depending on the site (as in the preceding approach). Counselors could not reject customers’ choices if the provider was on the state’s approved list.
The Workforce Investment Act

- **Maximum choice**
  - More likely to attend ITA orientation and use ITA
  - 6-7% higher participation rates
  - Entered training two weeks sooner on average
  - Less likely to participate in counseling after orientation

- **Structured choice**
  - Not fully implemented as designed – counselors were reluctant to comply!
  - $1,308 higher average earnings than maximum choice
Mental health parity

- Role of insurance: protection from catastrophic financial loss due to severe illness

- Traditionally, mental health and substance use disorder (MH/SUD) benefits more limited than general medical benefits
  - Limits disproportionately affect persons with severe illness

- In era of managed MH/SUD care
  - Are benefit limits needed to manage costs?
  - Will patients receive needed care?
Mental health parity

- In 2001 MH/SUD parity implemented in Federal Employees Health Benefit Program plans (FEHBP)
  - Improved insurance protection without increasing costs
  - Not associated with changes in access to care or MDD

- Mental Health Parity and Addiction Equity Act
  - Became Federal law in 2008
  - Regulations implemented in 2010
### Mental health parity

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>FEHBP (N=19,094)</th>
<th>Comparison (N=10,521)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>64.9</td>
<td>68.5</td>
</tr>
<tr>
<td>Employee</td>
<td>61.9</td>
<td>63.6</td>
</tr>
<tr>
<td>Geographical Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>11.4</td>
<td>16.7</td>
</tr>
<tr>
<td>South</td>
<td>64.3</td>
<td>15.9</td>
</tr>
<tr>
<td>Midwest</td>
<td>6.8</td>
<td>59.5</td>
</tr>
<tr>
<td>West</td>
<td>17.6</td>
<td>7.9</td>
</tr>
<tr>
<td>Mean age [s.d.]</td>
<td>46.2 [8.3]</td>
<td>43.6 [11.0]</td>
</tr>
<tr>
<td><strong>Psychiatric Diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar Disorder</td>
<td>2,557</td>
<td>13.4</td>
</tr>
<tr>
<td>Major Depression</td>
<td>10,412</td>
<td>54.5</td>
</tr>
<tr>
<td>Adjustment Disorder</td>
<td>6,125</td>
<td>32.1</td>
</tr>
</tbody>
</table>

**Note:** The table above compares the mental health parity of FEHBP (Federal Employees Health Benefit Program) and a comparison group, highlighting differences in various characteristics and psychiatric diagnoses.
## Mental health parity

<table>
<thead>
<tr>
<th></th>
<th>Baseline spending ($) per person</th>
<th>FEHBP Mean (s.d.)</th>
<th>Comparison Mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total spending</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>$3,347 (5,503)</td>
<td>3,536 (5,306)</td>
<td></td>
</tr>
<tr>
<td>Major depression</td>
<td>1,907 (2,943)</td>
<td>2,354 (3,055)</td>
<td></td>
</tr>
<tr>
<td>Adjustment disorder</td>
<td>807 (894)</td>
<td>931 (951)</td>
<td></td>
</tr>
<tr>
<td><strong>Out-of-pocket spending</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>$841 (1,263)</td>
<td>387 (557)</td>
<td></td>
</tr>
<tr>
<td>Major depression</td>
<td>568 (657)</td>
<td>302 (348)</td>
<td></td>
</tr>
<tr>
<td>Adjustment disorder</td>
<td>339 (334)</td>
<td>153 (199)</td>
<td></td>
</tr>
</tbody>
</table>
Mental health parity

- Difference-in-difference

![Graph showing out-of-pocket spending before and after parity]

Out-of-pocket spending

Pre-parity  Post-parity

FEHB  Comparison

Parity’s impact
### Mental health parity

<table>
<thead>
<tr>
<th>Out-of-pocket spending ($)</th>
<th>Diff-in-diff (99% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipolar Disorder</td>
<td>-149 (-217, -85)</td>
</tr>
<tr>
<td>Major Depression</td>
<td>-100 (-123, -77)</td>
</tr>
<tr>
<td>Adjustment Disorder</td>
<td>-69 (-84, -54)</td>
</tr>
</tbody>
</table>
Hypothetical case study
If you insist on strict proof (or strict disproof) in the empirical sciences, you will never benefit from experience, and never learn from it how wrong you are.

Sir Karl Popper
*The Logic of Scientific Discovery*
Many thanks!

- Senate Office of Education and Training
  - Megan Daly and Charles Revak

- The American Statistical Association
  - Steve Pierson

- Mathematica Policy Research
  - Carmen Ferro and Jennifer de Vallance
Contact me (just down the street)

Frank Yoon, Statistician
Mathematica Policy Research
1100 1st Street, NE
Washington, DC  20002

fyoon@mathematica-mpr.com
(202) 554-7518

www.mathematica-mpr.com
References


