Quasi-Experimental Designs for Studying QI Interventions: Strengths & Pitfalls

Rodney A. Hayward, MD
VA Ann Arbor HSR&D
Center of Excellence

The Challenge of QI Research

1. Need to evaluate system changes when managers & providers:
   - Are overwhelmed
   - Have external accountabilities
   - Are evaluated based upon what is achieved
   - Often feel undervalued
2. Large scale implementation at multiple sites
3. Maintain methodological rigor
4. Everyone wants immediate results

Why Is It So Difficult to Do Effectiveness Studies Using RCT Designs?

- Patient self selection (consent)
- Ethical and political issues
- Often want to change the entire system

Study Question

What are the costs and benefits of implementing a comprehensive diabetes case-management system?

Why not an RCT?

- Randomizing patients does not answer the question and is probably not feasible.
- Randomizing enough clinics is probably prohibitively expensive.

An RCT to Examine Wide-Scale, Real-World DM Case-Management

Pros:

- Gold standard for efficacy evaluation

Cons:

- Would need to either:
  - Randomize > 20-40 sites, or
  - Randomize within sites

Aren't Quasi-Experiments a Weak Methodology?

- Most published QE's are terrible, but
- Can produce strong evidence

Inherent Shortcoming of QE's

Can never eliminate the possibility of confounding

(Threat to Internal Validity)

What is a Quasi-Experiment?

- An intervention is studied but

1. Receipt of the intervention is non-random,

or

2. Randomized units are too few to be analyzed as random effects (<10-20 per intervention group)

Common Quasi-Experimental Designs

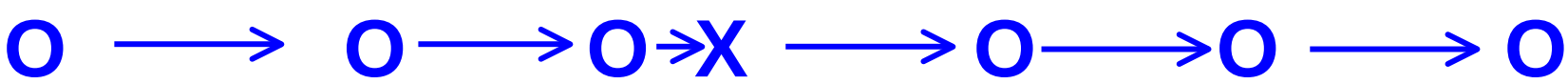
- Pre-Post comparison
- Time-series designs
- Non-random (non-equivalent) comparison groups
- Preference allocation

Simple Pre-Post Design

Intervention Group:

Time Series Analysis

Intervention Group:



Non- randomized control groups

Intervention group

O → X ⟶ O

O → X ⟶ O

Comparison Group

O ⟶ O

O ⟶ O

Common Pitfalls that Can Usually Be Addressed

- Regression towards the mean
- Confounding due to temporal or site-specific effects/events

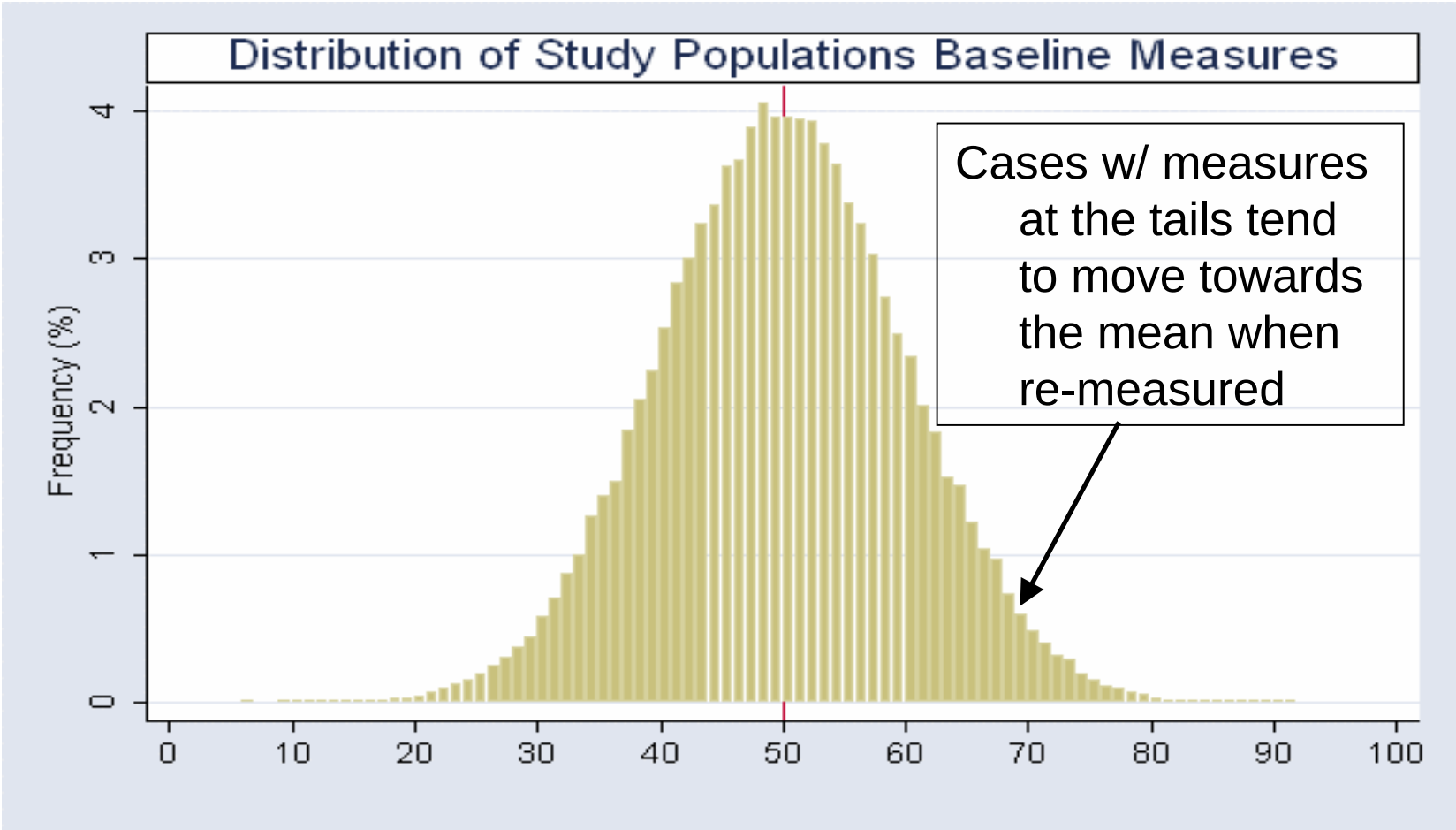
Simple Pre-Post Design

Intervention Group:

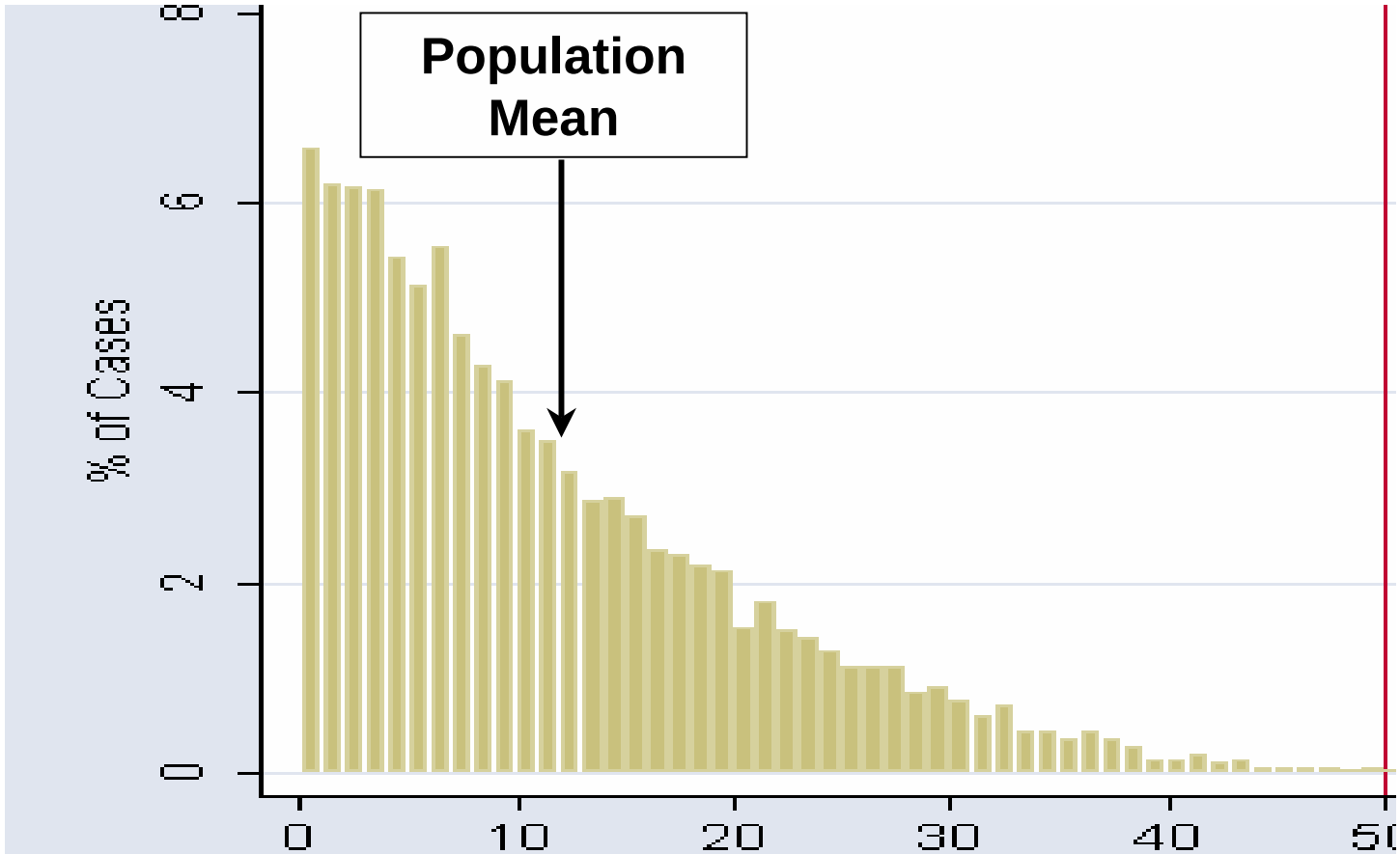
An Evaluation of An Intervention (Pre-Post)

- Select 1000 people with A1c > 9%
- At 6 months, A1c decreased by 1.5 points ($p < 0.001$)

Regression Towards the Mean



Regression Towards the Mean



Why Does Regression towards the Mean Occur

- Imprecise measurement leads to both misclassification errors and over-estimation of population variance

or

- Averaging over time decreases population variance

When Will Regression towards the Mean Occur

Whenever:

1) Baseline measures are imperfectly correlated with follow-up measures (in the absence of intervention),

AND

2) You preferentially sample subjects that are above or below the mean.

How to Prevent Regression towards the Mean Biasing Results

- Do not sample based upon baseline values (you do not need to intervene on the entire sample, however)
- Use a control group, or
- Use time-series analyses (increase baseline replicates)

How Big Can Regression to the Mean Be?

$$P_{rm} = 100\ (1\text{-}r)$$

P_{rm} = % regression towards mean

r = correlation of pre-post measures

How Big Can Regression to the Mean Be?

$$P_{rm} = 100\ (1\text{-}r)$$

P_{rm} = % regression towards mean

r = correlation of pre-post measures

If r = 1, then no regression

If r =0.5, then 50% regression towards mean

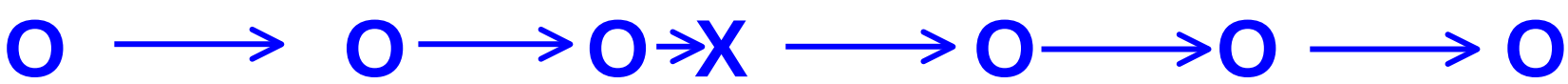
If r = 0, then complete regression

Precision of Baseline Measurement is Critical

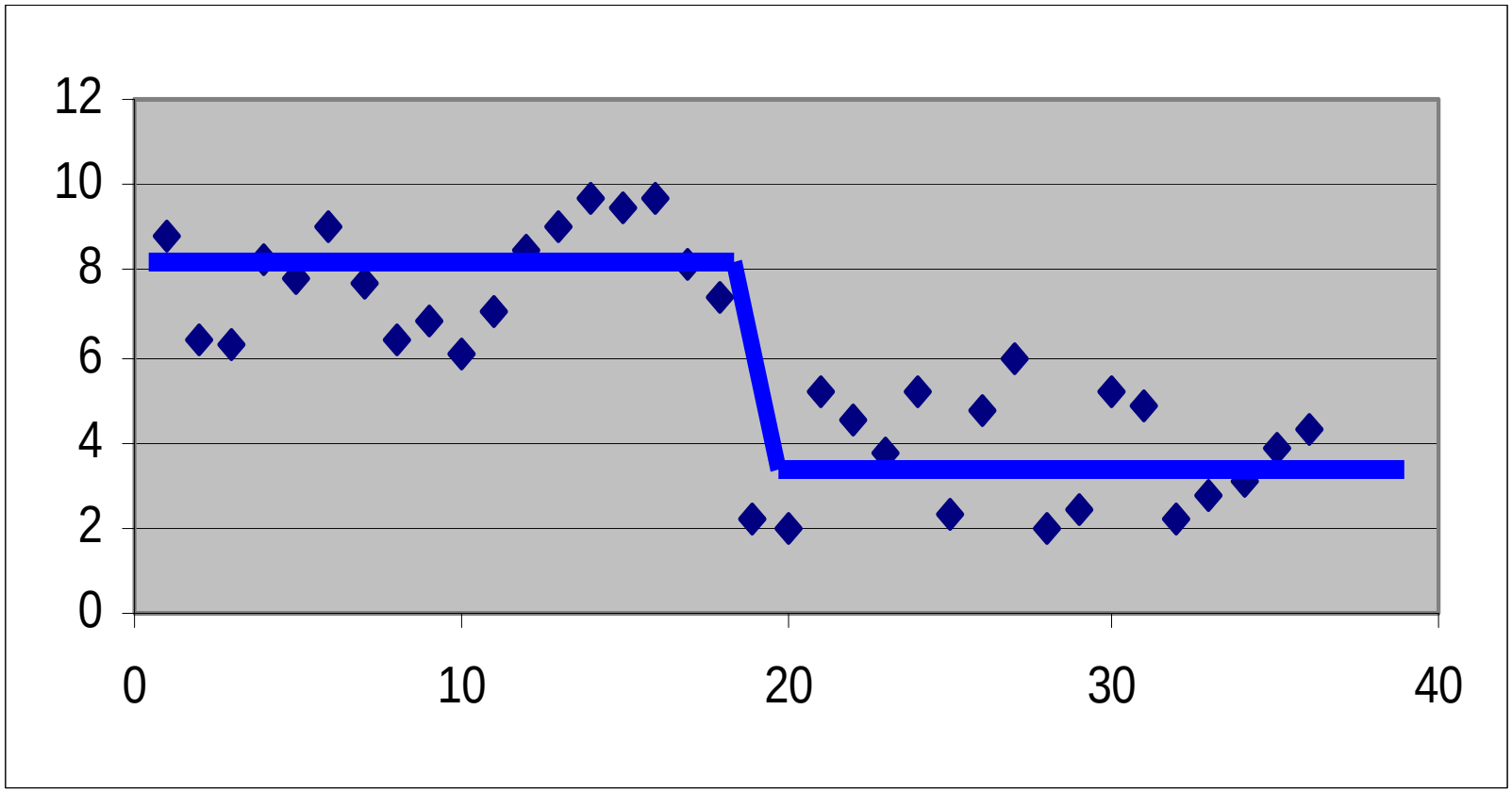
- Replicate measures to estimate reliability
- Measurements over time to estimate temporal stability or trends

Time Series Analysis

Intervention Group:



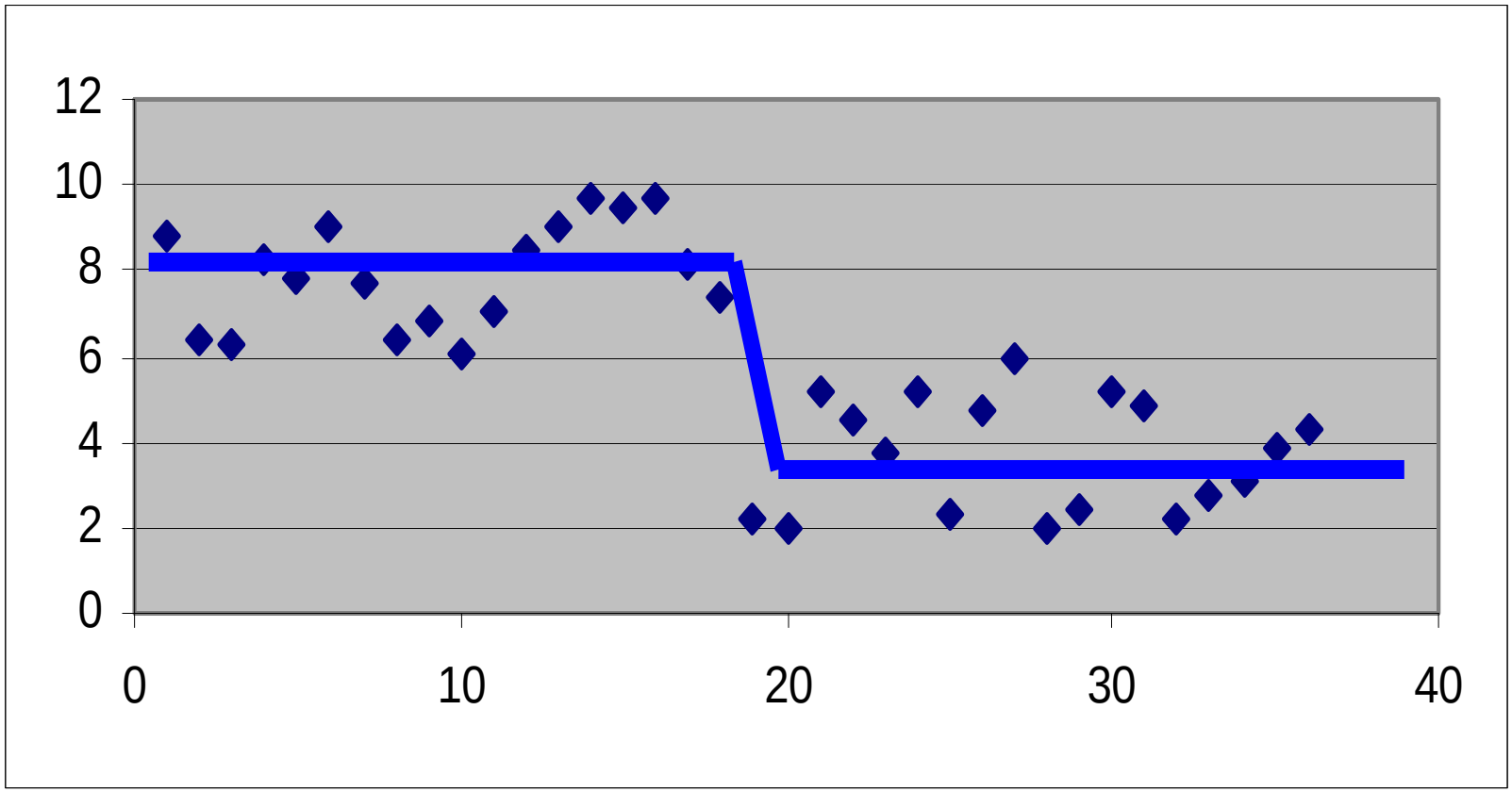
Time-series trial



Transfer & Impulse-Response Functions
(Box-Jenkins, dynamic regression, etc)

- Examines for 2 things
  - Abrupt change
  - Change in slope of temporal changes

Time-series trial



Non- randomized control groups

Intervention group

O → X ⟶ O

O → X ⟶ O

Comparison Group

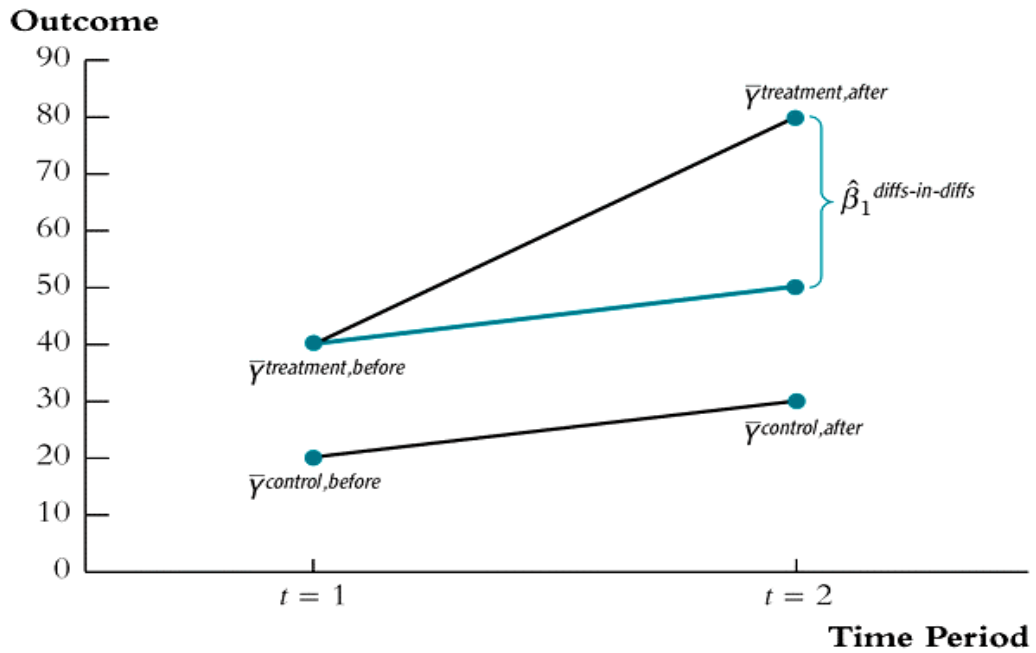
O ⟶ O

O ⟶ O

Difference-in-Differences Estimation

FIGURE 11.1 The Differences-in-Differences Estimator

The post-treatment difference between the treatment and control groups is 80 – 30 = 50, but this overstates the treatment effect because before the treatment Y was higher for the treatment than the control group by 40 – 20 = 20. The differences-in-differences estimator is the difference between the final and initial gaps, so that $\hat{\beta}_1^{\text{diffs-in-diffs}}$ = (80 – 30) – (40 – 20) = 50 – 20 = 30. Equivalently, the differences-in-differences estimator is the average change for the treatment group minus the average change for the control group, that is, $\hat{\beta}_1^{\text{diffs-in-diffs}} = \Delta\bar{Y}^{\text{treatment}} - \Delta\bar{Y}^{\text{control}}$ = (80 – 40) – (30 – 20) = 30.



JH Stock & MW Watson. Introduction to Econometrics

A Couple of Other QE Designs

- **Cross-Lagged Designs**

  (Tries to infer causality by examining temporal correlations between two variables)

- **Regression-Discontinuity Design**

  (Allows statistically robust design even if you preferentially offer the intervention to outliers)

Regression-Discontinuity Design

- Allows more statistically accurate assessment of studies that target interventions preferentially to those who are most likely to need (and benefit) from the intervention

- Key issue = Measure all subjects pre and post, but preferentially intervene on those with greater need for the intervention

Two Ways to Enhance Internal Validity

- Qualitative Evaluations for:
  - Evaluations of confounders
  - Examination of pathways

- Quantitative evaluation of pathways (intervening variables)

Designs For When Everyone Needs to Get an Intervention

- Randomize who gets the intervention ***first***
- Randomize two interventions with different outcomes

QI Research as A Partnership

- Respect
- Understanding
- Sharing Credit

Partnership Issues

- Your Partners will usually be evaluated on tangible results
- Unhappy constituents can get them fired
- They are often overwhelmed with other tasks

Study Question

What are the costs and benefits of implementing a comprehensive diabetes case-management system?

Basic Design Issues

- Unit of intervention allocation (doctor, clinic half-day, site?)
- Pre-Post, Time-series, Control sites?
- Baseline data?
- Intervening variable?
- Qualitative Evaluations?
- Partnership Issues?

Key Points

- QE's are often the optimal design for answering policy-relevant effectiveness questions

- QE's can be very strong designs, especially if:
  - Combine time-series & comparison group designs
  - Qualitative examination of micro-environment
  - Examine pathways (intervening factors)

Today's session was recorded and will be available online!

Check hsrd.webex.com for the archived version of this session!

For more information on HSRD Cyber Seminars visit
www.hsrd.research.va.gov/for_researchers/cyber_seminars/