Econometrics Final Exam

Counterfactual Study Guide

Part II: Causal Inference - What Else Could She Ask?

Based on patterns from all 3 practice exams

RCTs • Matching • IV • RD • DiD

December 7, 2025

Your Strength: Average 85/100 (B)
Strong conceptual understanding!

Prepared to maximize your exam performance

Contents

1 Part II Exam Structure Pattern

All three Part II practice exams follow this consistent structure:

1. **5–6 questions total:** Each worth 15–20 points
2. **Topics covered every exam:**
   - Randomized Controlled Trials (RCTs)
   - Matching methods
   - Instrumental Variables (IV)
   - Regression Discontinuity (RD)
   - Difference-in-Differences (DiD)
3. **Question format:** Scenario-based
   - Assess validity of proposed design
   - Identify threats to identification
   - Suggest improvements or alternatives
4. **Skills tested:**
   - Conceptual understanding of methods
   - Knowledge of assumptions
   - Ability to calculate treatment effects
   - Intuition about when methods work/fail

Your Performance Pattern

Strengths:

- Excellent grasp of when methods are appropriate
- Strong intuition about validity threats
- Good at identifying main issues

Room for improvement:

- More precision on technical formulas
- Explicit statement of ALL assumptions
- Direction of bias when assumptions violated

2 Method 1: Randomized Controlled Trials (RCTs)

What She ASKED	What She COULD ASK
Check balance table - need controls?	Interpret p-values in balance - threshold?
Design RCT for Airbnb photos	Design RCT for different context
Randomization: country/city/accommodation?	Individual vs cluster randomization trade-offs
Worried about take-up?	Selection vs compliance vs spillover

2.1 Core Concepts You Must Know

2.1.1 ATE and ATT Formulas

CRITICAL - MEMORIZE THIS

Average Treatment Effect (ATE):

$$\text{ATE} = E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)]$$

With randomization:

$$\boxed{\text{ATE} = E[Y|D=1] - E[Y|D=0] = \bar{Y}_1 - \bar{Y}_0}$$

Average Treatment Effect on Treated (ATT)

$$\text{ATT} = E[Y_i(1) - Y_i(0)|D_i = 1]$$

With randomization: ATE = ATT
Without randomization: ATE $\neq$ ATT (selection bias)

2.1.2 Balance Checks

- **Purpose:** Verify randomization worked - treatment and control groups are similar on observables
- **What to check:** All observable covariates (demographics, pre-treatment characteristics)
- **How to check:** Compare means, look at p-values from t-tests
- **Decision rule:** If $p > 0.05$, groups are balanced (no significant difference)
- **If imbalanced:** Need to control for those covariates in regression

2.1.3 Common Threats to RCTs

Watch Out For

1. **Selection bias:** If treatment assignment not truly random
2. **Take-up/compliance:** Not everyone assigned to treatment takes it
   - Intent-to-Treat (ITT): Effect of assignment
   - Treatment-on-Treated (TOT): Effect of actual treatment
3. **Spillover effects:** Control group affected by treatment group
4. **Attrition:** People drop out of study non-randomly
5. **Hawthorne effects:** People change behavior because they're being studied

3 Method 2: Matching Methods

What She ASKED	What She COULD ASK
Calculate ATE and ATT from table	Calculate with different weighting
Assumptions needed for matching	Which assumptions testable vs not?
Is PSM convincing for job training?	When PSM works vs when need IV/RD?
Common support: $P(D=1\|X) \in (0,1)$	What if common support violated?

3.1 Direct Matching Formulas

CRITICAL - MEMORIZE THIS

ATE (Average Treatment Effect):

$$\hat{\alpha}_{ATE} = \sum_j (\bar{Y}_{j1} - \bar{Y}_{j0}) \times \frac{N_j}{N}$$

Average effect across all groups, weighted by group size.

ATT (Average Treatment Effect on Treated):

$$\hat{\alpha}_{ATT} = \sum_j (\bar{Y}_{j1} - \bar{Y}_{j0}) \times \frac{N_{j1}}{N_1}$$

Average effect for treated, weighted by treated group size.

3.2 Key Assumptions

Conditional Independence (Unconfoundedness)

$$(Y_i(0), Y_i(1)) \perp D_i | X_i$$

Meaning: Given X, treatment assignment is as good as random

NOT testable - requires assuming you've observed all confounders

Common Support (Overlap)

$$0 < P(D=1|X=x) < 1 \text{ for all } x$$

Meaning: For each value of X, there exist both treated and control units

TESTABLE - can check empirically

3.3 Propensity Score Matching

- **Propensity score:** $p(X) = P(D=1|X)$

- **Purpose:** Reduce dimensionality - match on single number instead of many X variables
- **Rosenbaum-Rubin theorem:** If $(Y(0), Y(1)) \perp D|X$, then $(Y(0), Y(1)) \perp D|p(X)$
- **When it works:** When you've observed all confounders
- **When it fails:** Unobserved confounders (ability, motivation) $\Rightarrow$ need IV

Common Misconception

PSM does NOT solve unobserved confounding!
It only controls for observed variables. If important confounders are unobserved, you still have selection bias.

4 Method 3: Instrumental Variables (IV)

What She ASKED	What She COULD ASK
IV conditions: relevance & independence	Add exclusion & monotonicity
Calculate IV from distribution table	Calculate with means or regression
Identify compliance types	Which groups can we estimate for?
Why instrument needed? (self-selection)	Weak instruments or exclusion violations

4.1 IV Estimator

CRITICAL - MEMORIZE THIS

IV Estimator (Wald estimator):

$$\hat{\beta}_{IV} = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}$$

$$= \frac{\text{Reduced Form}}{\text{First Stage}}$$

where Z is the instrument, D is the endogenous treatment, Y is the outcome.

4.2 IV Assumptions (Constant Effects)

Relevance (First Stage)

$$\text{Cov}(Z, D) \neq 0$$

Meaning: Instrument must affect the treatment
Test: First-stage F-statistic > 10 (rule of thumb)
TESTABLE

Independence (Exogeneity)

$$\text{Cov}(Z, \varepsilon) = 0$$

Meaning: Instrument uncorrelated with error term
NOT testable (unless overidentified)

4.3 IV Assumptions (Heterogeneous Effects)

For heterogeneous treatment effects, we need **two additional** assumptions:

Exclusion Restriction

$$\text{Cov}(Z_i, Y_i(d)) = 0 \text{ for all } d$$

Meaning: Z only affects Y through its effect on D (no direct effect)
NOT testable

Monotonicity (No Defiers)

$$D_i(1) \geq D_i(0) \text{ for all } i$$

Meaning: No one does the opposite of what the instrument suggests
NOT testable

4.4 Compliance Types

Type	$D_i(1)$	$D_i(0)$
Compliers	1	0
Always-takers	1	1
Never-takers	0	0
Defiers	0	1

- **Compliers:** Take treatment when encouraged ($Z = 1$), don't when not ($Z = 0$)
- **Always-takers:** Always take treatment regardless of Z
- **Never-takers:** Never take treatment regardless of Z
- **Defiers:** Do the opposite (ruled out by monotonicity)

4.5 LATE (Local Average Treatment Effect)

CRITICAL - MEMORIZE THIS

What IV estimates:

$$\text{LATE} = E[Y_i(1) - Y_i(0) | D_i(1) > D_i(0)]$$

This is the effect **for COMPLIERS ONLY**.

Cannot estimate effects for: always-takers or never-takers

Key Limitation

IV does NOT estimate the ATE (average treatment effect for everyone).
It estimates the LATE (local average treatment effect for compliers).

If compliers are very different from the population, external validity is limited.

5 Method 4: Regression Discontinuity (RD)

What She ASKED	What She COULD ASK
Sharp vs fuzzy RD - which to use?	When treatment prob jumps vs changes
Two conditions: continuity & discontinuity	Why need continuity of potential outcomes?
Graph acceptance/rejection regions	Bandwidth selection - trade-offs
Running variable is revenue ratio	Manipulation test & implications

5.1 RD Conditions

Condition 1: Continuity of Potential Outcomes

$$\lim_{z \to z_0^-} E[Y_i(d)|Z_i = z] = \lim_{z \to z_0^+} E[Y_i(d)|Z_i = z]$$

Meaning: Units just below and above the cutoff are similar (only jump is from treatment)

Partially testable - can check continuity of covariates

Condition 2: Discontinuity in Treatment

$$\lim_{z \to z_0^-} E[D_i|Z_i = z] \neq \lim_{z \to z_0^+} E[D_i|Z_i = z]$$

Meaning: Treatment probability jumps at the cutoff

TESTABLE - can verify visually and formally

5.2 Sharp vs Fuzzy RD

Sharp RD

$$P(D_i = 1|Z_i = z) = \begin{cases} 0 & \text{if } z < z_0 \\ 1 & \text{if } z \geq z_0 \end{cases}$$

Treatment jumps from 0 to 1 **exactly** at cutoff.

Fuzzy RD

Treatment probability increases at cutoff but **not from 0 to 1**.

Example: Some below cutoff get treated, some above don't.

Estimation: Use IV approach with $Z \geq z_0$ as instrument for D

5.3 RD Estimand

CRITICAL - MEMORIZE THIS

What RD estimates:

$$\tau = \lim_{z \to z_0^+} E[Y_i|Z_i = z] - \lim_{z \to z_0^-} E[Y_i|Z_i = z]$$

This is a **LOCAL** effect - only for units near the cutoff.

External validity limited: Effect may not generalize to units far from cutoff.

5.4 Common Threats to RD

Watch Out For

1. **Manipulation of running variable:**
   - If units can control whether they're above/below cutoff
   - Test: McCrary density test for discontinuity in density
2. **Discontinuities in other variables:**
   - Other things change at cutoff besides treatment
   - Test: Check continuity of all covariates
3. **Functional form misspecification:**
   - Wrong polynomial degree for relationship
   - Solution: Sensitivity analysis with different specifications
4. **Bandwidth selection:**
   - Too narrow: large variance, few observations
   - Too wide: bias from including dissimilar units

6 Method 5: Difference-in-Differences (DiD)

What She ASKED	What She COULD ASK
Calculate DiD from 2×2 table	DiD with multiple time periods
Draw DiD graph with parallel trends	Event study graphs - interpret
Is DiD valid if both affected by tariffs?	Triple-difference for common shocks
Parallel trends assumption	Test parallel trends (pre-treatment)

6.1 DiD Estimator

CRITICAL - MEMORIZE THIS

Difference-in-Differences Estimator:

$$\begin{aligned}\text{DiD} &= (\bar{Y}_{\text{treated,after}} - \bar{Y}_{\text{treated,before}}) - (\bar{Y}_{\text{control,after}} - \bar{Y}_{\text{control,before}}) \\ &= \Delta\bar{Y}_{\text{treated}} - \Delta\bar{Y}_{\text{control}}\end{aligned}$$

Interpretation: Change in treated minus change in control

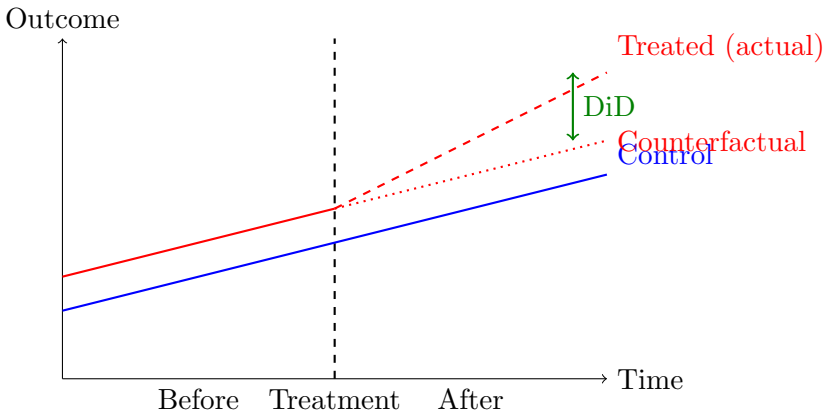
6.2 Parallel Trends Assumption

The Critical Assumption

$$E[Y_{0t} - Y_{0s}|D = 1] = E[Y_{0t} - Y_{0s}|D = 0]$$

In words: Without treatment, treated and control would have followed the same trend.

NOT testable - can't observe counterfactual
Can check with pre-treatment data - verify trends were parallel before treatment



6.3 Common Violations

Threats to Parallel Trends

1. **Common shocks:**
   - Both groups affected by same event (recession, policy)
   - Solution: Triple-difference (find group with shock but no treatment)
2. **Pre-existing differential trends:**
   - Groups already on different trajectories before treatment
   - Test: Check for parallel trends in pre-treatment period
3. **Composition changes:**
   - Who's in treated/control groups changes over time
   - Solution: Use balanced panel or control for composition
4. **Anticipation effects:**
   - Treatment group changes behavior before treatment
   - Look for: Divergence from parallel trends just before treatment

6.4 Triple-Difference (DDD)

When to Use Triple-Difference

Problem: Common shock affecting both treatment and control

Strategy: Find another group with same shock but NO treatment

Estimator:

$$\text{DDD} = \text{DiD}(\text{treated country}) - \text{DiD}(\text{comparison country})$$

7 Critical Method Comparison

7.1 When Each Method Works Best

Method	When It Works Best
RCT	When you can randomly assign treatment (ideal but not always feasible)
Matching	When you have rich data on confounders and believe you've observed everything important
IV	When you have unobserved confounders but can find a valid instrument
RD	When treatment assigned based on cutoff rule and units can't manipulate running variable
DiD	When you have before/after data for treatment and control groups with parallel trends

7.2 Assumptions Summary

Method	Main Assumption	Testable?
RCT	Random assignment	YES (balance)
Matching	Conditional independence Common support	NO YES
IV	Relevance Exclusion restriction	YES (F-stat) NO
RD	Continuity of $Y(0), Y(1)$ No manipulation	PARTIAL YES (McCrary)
DiD	Parallel trends	PARTIAL (pre-trends)

8 Common Mistakes to Avoid

Don't Make These Errors!

1. **Confusing ATE and ATT** - they're only equal with randomization
2. **Thinking PSM solves unobserved confounding** - it doesn't! Need IV
3. **Claiming IV estimates ATE** - it estimates LATE (compliers only)
4. **Forgetting to check parallel trends** - most critical for DiD
5. **Using sharp RD when fuzzy needed** - if treatment doesn't jump 0→1, it's fuzzy
6. **Not stating direction of bias** - always explain if estimate too high/low
7. **Ignoring external validity** - RD and LATE have limited generalizability

9 What Makes a Good Answer

Strategy for Scenario Questions

1. **Identify the threat:** What assumption is violated?
2. **Explain the mechanism:** HOW does this threaten validity?
3. **Direction of bias:** Will estimate be too high or too low?
4. **Suggest solution:** How could we fix this?

You're Already Strong on Part II!

Just tighten up the technical details and you'll ace it.

Good luck! ✓