Interpretation and robustness

Felipe Balcazar

NYU

March, 2023

NEW YORK UNIVERSITY

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ units.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one *percent*, leads to an increase in Y in $\hat{\beta}_1$ units.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one *percent*, leads to an increase in Y in $\hat{\beta}_1$ units.
- **log-linear:** $\log \hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ *percent*.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one *percent*, leads to an increase in Y in $\hat{\beta}_1$ units.
- **log-linear:** $\log \hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ *percent*.
- **log-log:** $\log \hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one *percent*, leads to an increase in Y in $\hat{\beta}_1$ *percent*.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one *percent*, leads to an increase in Y in $\hat{\beta}_1$ units.
- **log-linear:** $\log \hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one unit, leads to an increase in Y in $\hat{\beta}_1$ *percent*.
- **log-log:** $\log \hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log D_{it} + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in D in one *percent*, leads to an increase in Y in $\hat{\beta}_1$ *percent*.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in GDP in one unit, leads to an increase in Dem in $\hat{\beta}_1$ units.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one unit, leads to an increase in *Dem* in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one *percent*, leads to an increase in *Dem* in $\hat{\beta}_1$ units.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one unit, leads to an increase in *Dem* in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one *percent*, leads to an increase in *Dem* in $\hat{\beta}_1$ units.
- **log-linear:** $\log \hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one unit, leads to an increase in *Dem* in $\hat{\beta}_1$ *percent*.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one unit, leads to an increase in *Dem* in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one *percent*, leads to an increase in *Dem* in $\hat{\beta}_1$ units.
- **log-linear:** $\log \hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one unit, leads to an increase in *Dem* in $\hat{\beta}_1$ *percent*.
- **log-log:** $\log \hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one *percent*, leads to an increase in *Dem* in $\hat{\beta}_1$ *percent*.

Interpretation on the basis of measurement

- **linear-linear:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one unit, leads to an increase in *Dem* in $\hat{\beta}_1$ units.
- **linear-log:** $\hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one *percent*, leads to an increase in *Dem* in $\hat{\beta}_1$ units.
- **log-linear:** $\log \hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 GDPit + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one unit, leads to an increase in *Dem* in $\hat{\beta}_1$ *percent*.
- **log-log:** $\log \hat{Dem}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \log GDPit + \hat{\delta} X_{it} + \hat{\mu}_i + \hat{\gamma}_t$
  - $\hat{\beta}_1$: An increase in *GDP* in one *percent*, leads to an increase in *Dem* in $\hat{\beta}_1$ *percent*.

Interpretation on the basis of interaction effects

$$Dem_{it} = \alpha + \beta_1 GDP_{it} + \beta_2 Oil_{it} + \beta_3 GDP_{it} \times Oil_{it} + \delta X_{it} + \mu_i + \gamma_t + \varepsilon_{it}.$$

- β_1 is the effect of the treatment conditional on $Oil = 0$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = 0, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = 0, X)$$

- $\beta_1 + \beta_3 Oil$ is the effect of the treatment conditional on $Oil = C$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = C, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = C, X)$$

- β_3 is the additional effect of the treatment for Oil when goes from 0 to C.

Interpretation on the basis of interaction effects

$$Dem_{it} = \alpha + \beta_1 GDP_{it} + \beta_2 Oil_{it} + \beta_3 GDP_{it} \times Oil_{it} + \delta X_{it} + \mu_i + \gamma_t + \varepsilon_{it}.$$

- β_1 is the effect of the treatment conditional on $Oil = 0$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = 0, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = 0, X)$$

- $\beta_1 + \beta_3 Oil$ is the effect of the treatment conditional on $Oil = C$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = C, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = C, X)$$

- β_3 is the additional effect of the treatment for Oil when goes from 0 to C.

Interpretation on the basis of interaction effects

$$Dem_{it} = \alpha + \beta_1 GDP_{it} + \beta_2 Oil_{it} + \beta_3 GDP_{it} \times Oil_{it} + \delta X_{it} + \mu_i + \gamma_t + \varepsilon_{it}.$$

- β_1 is the effect of the treatment conditional on $Oil = 0$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = 0, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = 0, X)$$

- $\beta_1 + \beta_3 Oil$ is the effect of the treatment conditional on $Oil = C$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = C, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = C, X)$$

- β_3 is the additional effect of the treatment for Oil when goes from 0 to C.

Interpretation on the basis of interaction effects

$$Dem_{it} = \alpha + \beta_1 GDP_{it} + \beta_2 Oil_{it} + \beta_3 GDP_{it} \times Oil_{it} + \delta X_{it} + \mu_i + \gamma_t + \varepsilon_{it}.$$

- β_1 is the effect of the treatment conditional on $Oil = 0$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = 0, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = 0, X)$$

- $\beta_1 + \beta_3 Oil$ is the effect of the treatment conditional on $Oil = C$:

  $$E(Dem_{it}|GDP_{it} = 1, Oil_{it} = C, X) - E(Dem_{it}|GDP_{it} = 0, Oil_{it} = C, X)$$

- β_3 is the additional effect of the treatment for Oil when goes from 0 to C.

The crux of the empirical researcher

- When we use observational data we may not observe all confounders.
  - We may be able to control for observable/measurable ones.
  - We can also use the tools we've learned: diff-in-diff, instrumental variables, RDD, and so on.
- Is this sufficient? Oftentimes it's not!
  - Robustness to small and sensible manipulations?
  - Robustness to potential confounders?
  - Robustness to alternative explanations?
- What's the purpose of addressing the previous issues?
  - To at least provide robust conclusions about the direction of the causal effect.
  - To show the estability of the magnitude/significance of the effect.

The crux of the empirical researcher

- When we use observational data we may not observe all confounders.
  - We may be able to control for observable/measurable ones.
  - We can also use the tools we've learned: diff-in-diff, instrumental variables, RDD, and so on.
- Is this sufficient? Oftentimes it's not!
  - Robustness to small and sensible manipulations?
  - Robustness to potential confounders?
  - Robustness to alternative explanations?
- What's the purpose of addressing the previous issues?
  - To at least provide robust conclusions about the direction of the causal effect.
  - To show the estability of the magnitude/significance of the effect.

The crux of the empirical researcher

- When we use observational data we may not observe all confounders.
  - We may be able to control for observable/measurable ones.
  - We can also use the tools we've learned: diff-in-diff, instrumental variables, RDD, and so on.
- Is this sufficient? Oftentimes it's not!
  - Robustness to small and sensible manipulations?
  - Robustness to potential confounders?
  - Robustness to alternative explanations?
- What's the purpose of addressing the previous issues?
  - To at least provide robust conclusions about the direction of the causal effect.
  - To show the estability of the magnitude/significance of the effect.

The crux of the empirical researcher

- When we use observational data we may not observe all confounders.
  - We may be able to control for observable/measurable ones.
  - We can also use the tools we've learned: diff-in-diff, instrumental variables, RDD, and so on.
- Is this sufficient? Oftentimes it's not!
  - Robustness to small and sensible manipulations?
  - Robustness to potential confounders?
  - Robustness to alternative explanations?
- What's the purpose of addressing the previous issues?
  - To at least provide robust conclusions about the direction of the causal effect.
  - To show the estability of the magnitude/significance of the effect.

The crux of the empirical researcher

- When we use observational data we may not observe all confounders.
  - We may be able to control for observable/measurable ones.
  - We can also use the tools we've learned: diff-in-diff, instrumental variables, RDD, and so on.
- Is this sufficient? Oftentimes it's not!
  - Robustness to small and sensible manipulations?
  - Robustness to potential confounders?
  - Robustness to alternative explanations?
- What's the purpose of addressing the previous issues?
  - To at least provide robust conclusions about the direction of the causal effect.
  - To show the estability of the magnitude/significance of the effect.

The crux of the empirical researcher

- When we use observational data we may not observe all confounders.
  - We may be able to control for observable/measurable ones.
  - We can also use the tools we've learned: diff-in-diff, instrumental variables, RDD, and so on.
- Is this sufficient? Oftentimes it's not!
  - Robustness to small and sensible manipulations?
  - Robustness to potential confounders?
  - Robustness to alternative explanations?
- What's the purpose of addressing the previous issues?
  - To at least provide robust conclusions about the direction of the causal effect.
  - To show the estability of the magnitude/significance of the effect.

The crux of the empirical researcher

- When we use observational data we may not observe all confounders.
  - We may be able to control for observable/measurable ones.
  - We can also use the tools we've learned: diff-in-diff, instrumental variables, RDD, and so on.
- Is this sufficient? Oftentimes it's not!
  - Robustness to small and sensible manipulations?
  - Robustness to potential confounders?
  - Robustness to alternative explanations?
- What's the purpose of addressing the previous issues?
  - To at least provide robust conclusions about the direction of the causal effect.
  - To show the estability of the magnitude/significance of the effect.

Robustness tests are ways to

- Robustness means two things in general:
  - The direction of the effect doesn't change.
  - The magnitude of the effect doesn't change.
  - Results are more robust when it is the latter.
  - Effects must remain statistically significant.

Robustness tests are ways to

- Robustness means two things in general:
  - The direction of the effect doesn't change.
  - The magnitude of the effect doesn't change.
  - Results are more robust when it is the latter.
  - Effects must remain statistically significant.
- Robustness tests can be of many kinds:
  - Introduce sets of confounders at a time to check for stability.
  - Drop observations with replacement.
  - Use another measure for the treatment and (or) control.
  - Sensitivity of to unmeasured confounding (wont' be covered).
  - Worst-case bounds (won't be covered).

Robustness tests are ways to

- Robustness means two things in general:
  - The direction of the effect doesn't change.
  - The magnitude of the effect doesn't change.
  - Results are more robust when it is the latter.
  - Effects must remain statistically significant.
- Robustness tests can be of many kinds:
  - Introduce sets of confounders at a time to check for stability.
  - Drop observations with replacement.
  - Use another measure for the treatment and (or) control.
  - Sensitivity of to unmeasured confounding (wont' be covered).
  - Worst-case bounds (won't be covered).
- Robustness tests can be also clever exercises! (Placebo tests)
  - Theoretical prior indicates where should find an effect and where we shouldn't.

Empirical robustness tests: stability

- Introduce sets of confounders at a time to check for stability.
  - Classify confounders in sets, e.g.: fixed effects, individual-level, municipality-level.
  - Time $\times$ area fixed effects - control s for time-variant confounders at higher level than the level of the treatment assignment.
- Drop observations/areas/sets of observations with replacement.
  - Similar to bootstrapping, but not bootstrapping.
  - One at a time, recompute, plot coefficient with SE.
- Use another measure for the treatment and (or) control.
  - They need to be good proxies for treatment and (or) control.
  - Discrete to continuous; different standardization; polynomial of order > 1.

Empirical robustness tests: stability

- Introduce sets of confounders at a time to check for stability.
  - Classify confounders in sets, e.g.: fixed effects, individual-level, municipality-level.
  - Time $\times$ area fixed effects - control s for time-variant confounders at higher level than the level of the treatment assignment.
- Drop observations/areas/sets of observations with replacement.
  - Similar to bootstrapping, but not bootstrapping.
  - One at a time, recompute, plot coefficient with SE.
- Use another measure for the treatment and (or) control.
  - They need to be good proxies for treatment and (or) control.
  - Discrete to continuous; different standardization; polynomial of order > 1.

Empirical robustness tests: stability

- Introduce sets of confounders at a time to check for stability.
  - Classify confounders in sets, e.g.: fixed effects, individual-level, municipality-level.
  - Time $\times$ area fixed effects - control s for time-variant confounders at higher level than the level of the treatment assignment.
- Drop observations/areas/sets of observations with replacement.
  - Similar to bootstrapping, but not bootstrapping.
  - One at a time, recompute, plot coefficient with SE.
- Use another measure for the treatment and (or) control.
  - They need to be good proxies for treatment and (or) control.
  - Discrete to continuous; different standardization; polynomial of order > 1.

Empirical robustness tests: stability

- Sensitivity of to unmeasured confounding (wont' be covered).
  - Generate a variable that has different levels of correlation with treatment and outcome.
  - Plot it against all observable confounders - it should look extraneous/implausible.
- Manski bounds/partial identification (won't be covered).
  - Assume worst case scenario and generate the bounds.
  - Trim bounds if warranted.
- Additional: For multiple treatments/outcomes if measurement is similar:
  - Equivalence testing. Evaluate all four patterns:
    - Statistically equivalent.
    - Statistically different from.
    - Practically insignificant.
    - Inconclusive .

Empirical robustness tests: stability

- Sensitivity of to unmeasured confounding (wont' be covered).
  - Generate a variable that has different levels of correlation with treatment and outcome.
  - Plot it against all observable confounders - it should look extraneous/implausible.
- Manski bounds/partial identification (won't be covered).
  - Assume worst case scenario and generate the bounds.
  - Trim bounds if warranted.
- Additional: For multiple treatments/outcomes if measurement is similar:
  - Equivalence testing. Evaluate all four patterns:
    - Statistically equivalent.
    - Statistically different from.
    - Practically insignificant.
    - Inconclusive .

Empirical robustness tests: stability

- Sensitivity of to unmeasured confounding (wont' be covered).
  - Generate a variable that has different levels of correlation with treatment and outcome.
  - Plot it against all observable confounders - it should look extraneous/implausible.
- Manski bounds/partial identification (won't be covered).
  - Assume worst case scenario and generate the bounds.
  - Trim bounds if warranted.
- Additional: For multiple treatments/outcomes if measurement is similar:
  - Equivalence testing. Evaluate all four patterns:
    - Statistically equivalent.
    - Statistically different from.
    - Practically insignificant.
    - Inconclusive .

Theoretical robustness tests: measurement and placebos

- No empirical paper is good without a theory (or model) driving data analysis.
- Solid theory gives guidelines to define measurement of dependent variable, treatment and confounders.
  - We should observe (weaker/stronger/no) effect if measurement changes or confounder is added.
  - It can also guide the selection of an instrumental variable.
- Solid theory provides use with guidelines for computing interaction effects.
  - We should observe (weaker/stronger/no) effect for certain groups in the sample.
  - A placebo, in particular, is a test where we shouldn't observe an effect of the treatment in a gi en group in the sample.

Theoretical robustness tests: measurement and placebos

- No empirical paper is good without a theory (or model) driving data analysis.
- Solid theory gives guidelines to define measurement of dependent variable, treatment and confounders.
  - We should observe (weaker/stronger/no) effect if measurement changes or confounder is added.
  - It can also guide the selection of an instrumental variable.
- Solid theory provides use with guidelines for computing interaction effects.
  - We should observe (weaker/stronger/no) effect for certain groups in the sample.
  - A placebo, in particular, is a test where we shouldn't observe an effect of the treatment in a gi en group in the sample.

Theoretical robustness tests: measurement and placebos

- No empirical paper is good without a theory (or model) driving data analysis.
- Solid theory gives guidelines to define measurement of dependent variable, treatment and confounders.
  - We should observe (weaker/stronger/no) effect if measurement changes or confounder is added.
  - It can also guide the selection of an instrumental variable.
- Solid theory provides use with guidelines for computing interaction effects.
  - We should observe (weaker/stronger/no) effect for certain groups in the sample.
  - A placebo, in particular, is a test where we shouldn't observe an effect of the treatment in a gi en group in the sample.