



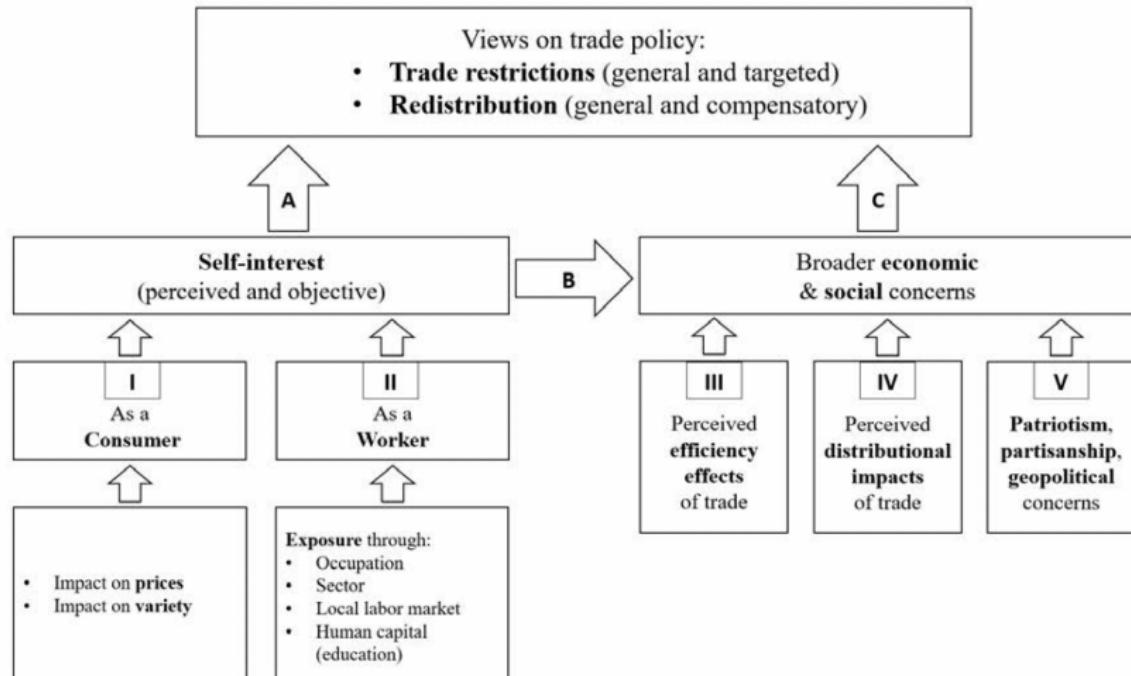
Causal Inference for IR and IPE with Substantive Applications

Carlos Felipe Balcazar

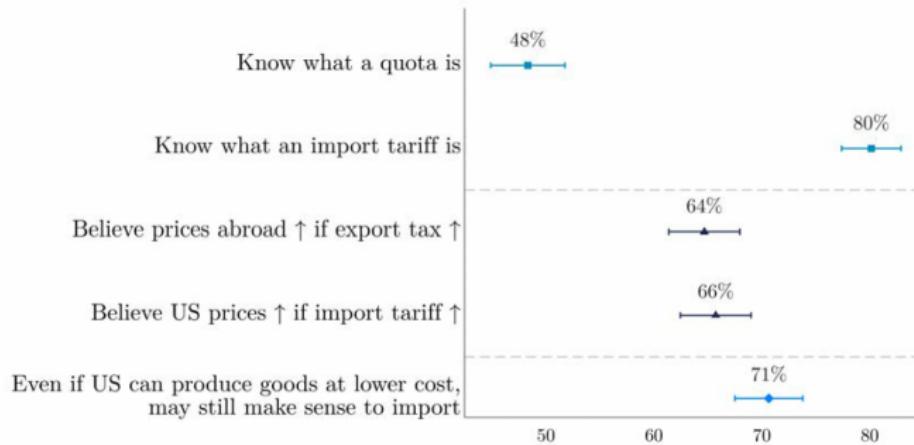
MacMillan Center

March, 2024

Views on trade policy

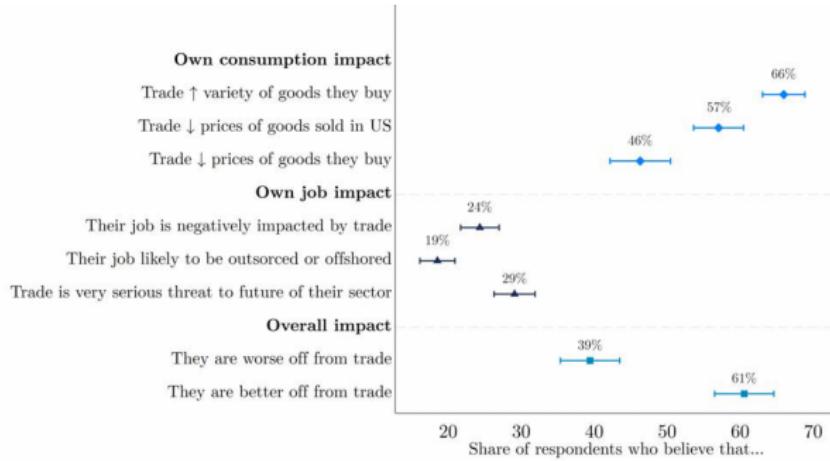


Views on trade policy



- ▶ Voters don't have a trade model in their heads.
 - ▶ But useful to think about behavior.
- ▶ Trade-off: income v. lower prices (consumption); framing matters.

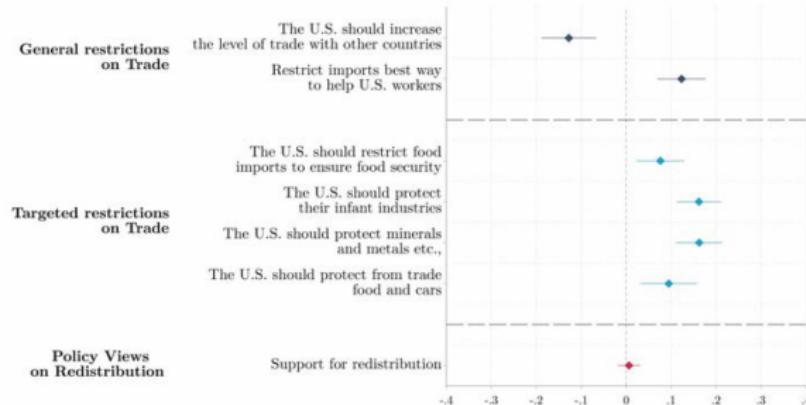
Views on trade policy



- ▶ Voters don't have a trade model in their heads.
 - ▶ But useful to think about behavior.
- ▶ Trade-off: income v. lower prices (consumption); framing matters.

Views on trade policy

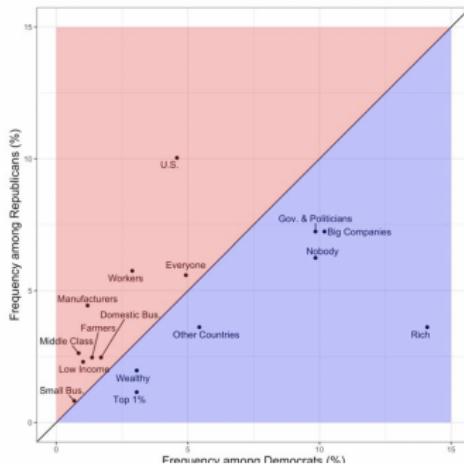
(A) CORRELATION BETWEEN WILLINGNESS TO SIGN PETITION FOR STRONGER TRADE BARRIERS AND POLICY VIEWS



- ▶ Voters don't have a trade model in their heads.
 - ▶ But useful to think about behavior.
- ▶ Trade-off: income v. lower prices (consumption); framing matters.

Views on trade policy

WHICH GROUPS OF PEOPLE DO YOU THINK WOULD GAIN IF TRADE BARRIERS SUCH AS TARIFFS WERE INCREASED?



- ▶ Voters don't have a trade model in their heads.
 - ▶ But useful to think about behavior.
- ▶ Trade-off: income v. lower prices (consumption); framing matters.

Views on trade policy

There are often both **winners** and **losers** from trade.



When there is more trade, all **households** who **consume** the imported goods can gain from it.

The benefits from increased trade can be perceived by a **large group**, throughout the country.



The losers from trade are generally a **smaller group**, often concentrated in one place or industry.

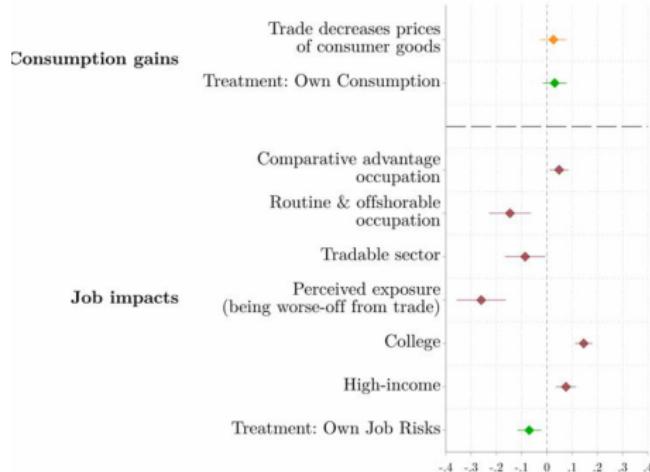
However, their losses can be very large, and therefore more **visible**.



- ▶ Voters don't have a trade model in their heads.
 - ▶ But useful to think about behavior.
- ▶ Trade-off: income v. lower prices (consumption); framing matters.

Views on trade policy

(A) SUPPORT FOR MORE FREE TRADE AND LOWER TRADE RESTRICTIONS



- ▶ Voters don't have a trade model in their heads.
 - ▶ But useful to think about behavior.
- ▶ Trade-off: income v. lower prices (consumption); framing matters.

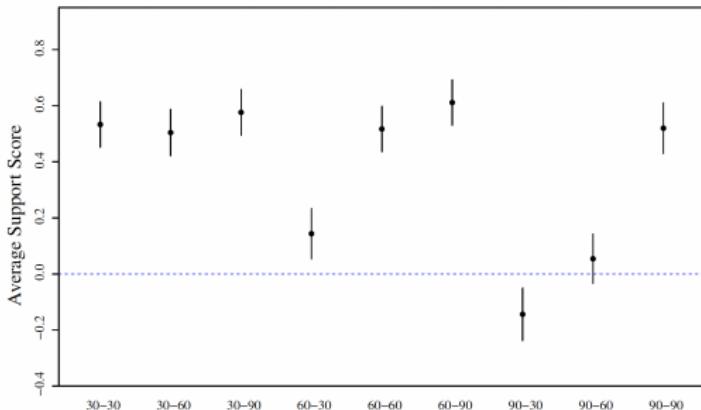
Preference formation in the backlash against globalization

The text of the first experiment is as follows:

The US is considering negotiating a trade agreement with one of its trading partners. The trade agreement will decrease the average tariffs—that is, the tax charged by the American government on foreign goods entering the United States—by [30, 60, or 90] percent. In return the trade partner will decrease their tariffs on imports from the US by [30, 60, or 90] percent.

- ▶ Informational treatment involves priming of some kind.
- ▶ Consistent with “elite-cues” theoretical approach.
 - ▶ But scope conditions are often undefined
- ▶ Foreign framing often creates backlash.

Preference formation in the backlash against globalization



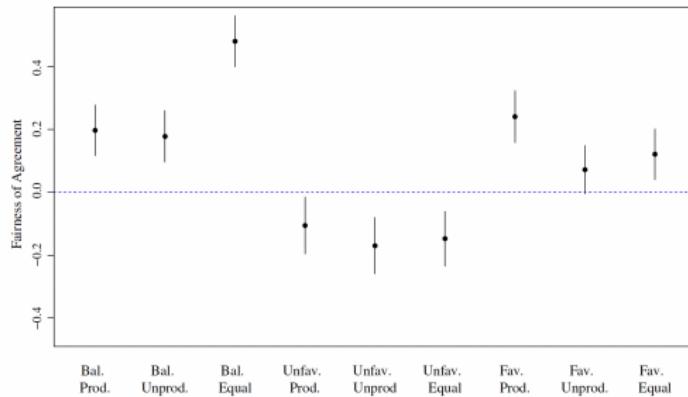
- ▶ Informational treatment involves priming of some kind.
- ▶ Consistent with “elite-cues” theoretical approach.
 - ▶ But scope conditions are often undefined
- ▶ Foreign framing often creates backlash.

Preference formation in the backlash against globalization

The United States is part of a free trade agreement in which the United States and the other country reduced its tariffs—that is, the tax charged on foreign goods and services when they cross borders. The agreement has led to a 2 to 1 trade imbalance favoring the United States in which the United States exports twice as much as it imports from the other country. Studies show that American workers are twice as productive as the workers in the other country.

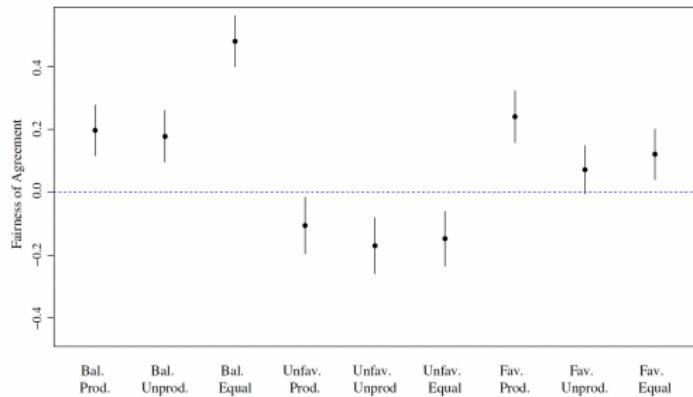
- ▶ Informational treatment involves priming of some kind.
- ▶ Consistent with “elite-cues” theoretical approach.
 - ▶ But scope conditions are often undefined
- ▶ Foreign framing often creates backlash.

Preference formation in the backlash against globalization



- ▶ Informational treatment involves priming of some kind.
- ▶ Consistent with “elite-cues” theoretical approach.
 - ▶ But scope conditions are often undefined
- ▶ Foreign framing often creates backlash.

Preference formation in the backlash against globalization



- ▶ Informational treatment involves priming of some kind.
- ▶ Consistent with “elite-cues” theoretical approach.
 - ▶ But scope conditions are often undefined
- ▶ Foreign framing often creates backlash.

Preference formation in the backlash against globalization

	Labor	Automation
Foreign	More Protection Fewer Transfers	More Regulation Fewer Transfers
Domestic	Less Protection More Transfers	Less Regulation More Transfers

Table 1: Predicted Effect of Shock Type on Responses

- ▶ Informational treatment involves priming of some kind.
- ▶ Consistent with “elite-cues” theoretical approach.
 - ▶ But scope conditions are often undefined
- ▶ Foreign framing often creates backlash.

Preference formation in the backlash against globalization

	Labor (Imported)	Automation (Exported)
Foreign	Backpedaling Policy: 63.6 Transfers: 66.9 Difference: -3.2 95% Conf. Int. [-5.8, -0.6]	Backpedaling Policy: 56.7 Transfers: 64.7 Difference: -7.9 95% Conf. Int. [-10.1, -5.7]
Domestic	Backpedaling Policy: 58.3 Transfers: 65.4 Difference: -7.2 95% Conf. Int. [-9.7, -4.6]	Backpedaling Policy: 54.4 Transfers: 66 Difference: -11.6 95% Conf. Int. [-13.8, -9.4]

- ▶ Informational treatment involves priming of some kind.
- ▶ Consistent with “elite-cues” theoretical approach.
 - ▶ But scope conditions are often undefined
- ▶ Foreign framing often creates backlash.

Shortcomings for commensurability

- ▶ Treatment uptake is related to respondent type; often undefined.
 - ▶ Ideology; location; industry; sector, etc.
 - ▶ Often ex-post theoretical.
- ▶ Design often ignores strategic responses in equilibrium.
 - ▶ Trade-offs considered in real-life environments.
 - ▶ Unclear if effect is transitory or permanent.
 - ▶ Unclear if preferences translate into policy (e.g., citizen-candidate).
 - ▶ Scope conditions are often not laid out (time/geography).
- ▶ Absence of clear equilibrium implications. Is this bad? Why?

Shortcomings for commensurability

- ▶ Treatment uptake is related to respondent type; often undefined.
 - ▶ Ideology; location; industry; sector, etc.
 - ▶ Often ex-post theoretical.
- ▶ Design often ignores strategic responses in equilibrium.
 - ▶ Trade-offs considered in real-life environments.
 - ▶ Unclear if effect is transitory or permanent.
 - ▶ Unclear if preferences translate into policy (e.g., citizen-candidate).
 - ▶ Scope conditions are often not laid out (time/geography).
- ▶ Absence of clear equilibrium implications. Is this bad? Why?

Shortcomings for commensurability

- ▶ Treatment uptake is related to respondent type; often undefined.
 - ▶ Ideology; location; industry; sector, etc.
 - ▶ Often ex-post theoretical.
- ▶ Design often ignores strategic responses in equilibrium.
 - ▶ Trade-offs considered in real-life environments.
 - ▶ Unclear if effect is transitory or permanent.
 - ▶ Unclear if preferences translate into policy (e.g., citizen-candidate).
 - ▶ Scope conditions are often not laid out (time/geography).
- ▶ Absence of clear equilibrium implications. Is this bad? Why?

General theoretical considerations for commensurability

- ▶ Are experiments the gold standard? What are we learning?
- ▶ Behavioral trade-off? Cheap talk or costly signaling? Why?
 - ▶ When is it informative about behavior; novelty?
- ▶ Is theoretical prediction counterintuitive? Why?
 - ▶ Higher likelihood of capturing eq. behavior.
- ▶ Are the implications feasible? (Constraints.)
 - ▶ Relevant if policy implications are drawn.
- ▶ Evaluate a clear behavioral mechanism (e.g., lab experiments).
 - ▶ Monetary reward that is contingent on eq. behavior.
- ▶ Baseline beliefs can be valuable for assessing Bayesian updating.

General theoretical considerations for commensurability

- ▶ Are experiments the gold standard? What are we learning?
- ▶ Behavioral trade-off? Cheap talk or costly signaling? Why?
 - ▶ When is it informative about behavior; novelty?
- ▶ Is theoretical prediction counterintuitive? Why?
 - ▶ Higher likelihood of capturing eq. behavior.
- ▶ Are the implications feasible? (Constraints.)
 - ▶ Relevant if policy implications are drawn.
- ▶ Evaluate a clear behavioral mechanism (e.g., lab experiments).
 - ▶ Monetary reward that is contingent on eq. behavior.
- ▶ Baseline beliefs can be valuable for assessing Bayesian updating.

General theoretical considerations for commensurability

- ▶ Are experiments the gold standard? What are we learning?
- ▶ Behavioral trade-off? Cheap talk or costly signaling? Why?
 - ▶ When is it informative about behavior; novelty?
- ▶ Is theoretical prediction counterintuitive? Why?
 - ▶ Higher likelihood of capturing eq. behavior.
- ▶ Are the implications feasible? (Constraints.)
 - ▶ Relevant if policy implications are drawn.
- ▶ Evaluate a clear behavioral mechanism (e.g., lab experiments).
 - ▶ Monetary reward that is contingent on eq. behavior.
- ▶ Baseline beliefs can be valuable for assessing Bayesian updating.

General theoretical considerations for commensurability

- ▶ Are experiments the gold standard? What are we learning?
- ▶ Behavioral trade-off? Cheap talk or costly signaling? Why?
 - ▶ When is it informative about behavior; novelty?
- ▶ Is theoretical prediction counterintuitive? Why?
 - ▶ Higher likelihood of capturing eq. behavior.
- ▶ Are the implications feasible? (Constraints.)
 - ▶ Relevant if policy implications are drawn.
- ▶ Evaluate a clear behavioral mechanism (e.g., lab experiments).
 - ▶ Monetary reward that is contingent on eq. behavior.
- ▶ Baseline beliefs can be valuable for assessing Bayesian updating.

General theoretical considerations for commensurability

- ▶ Are experiments the gold standard? What are we learning?
- ▶ Behavioral trade-off? Cheap talk or costly signaling? Why?
 - ▶ When is it informative about behavior; novelty?
- ▶ Is theoretical prediction counterintuitive? Why?
 - ▶ Higher likelihood of capturing eq. behavior.
- ▶ Are the implications feasible? (Constraints.)
 - ▶ Relevant if policy implications are drawn.
- ▶ Evaluate a clear behavioral mechanism (e.g., lab experiments).
 - ▶ Monetary reward that is contingent on eq. behavior.
- ▶ Baseline beliefs can be valuable for assessing Bayesian updating.

General theoretical considerations for commensurability

- ▶ Are experiments the gold standard? What are we learning?
- ▶ Behavioral trade-off? Cheap talk or costly signaling? Why?
 - ▶ When is it informative about behavior; novelty?
- ▶ Is theoretical prediction counterintuitive? Why?
 - ▶ Higher likelihood of capturing eq. behavior.
- ▶ Are the implications feasible? (Constraints.)
 - ▶ Relevant if policy implications are drawn.
- ▶ Evaluate a clear behavioral mechanism (e.g., lab experiments).
 - ▶ Monetary reward that is contingent on eq. behavior.
- ▶ Baseline beliefs can be valuable for assessing Bayesian updating.

Pre-analysis plans

Advantages

Pre-Analysis Plan Checklist

<i>Item</i>	<i>Brief description</i>
Primary outcome variable	The key variable of interest for the study. If multiple variables are to be examined, one should know how the multiple hypothesis testing will be done.
Secondary outcome variable(s)	Additional variables of interest to be examined.
Variable definitions	Precise variable definitions that specify how the raw data will be transformed into the actual variables to be used for analysis.
Inclusion/Exclusion rules	Rules for including or excluding observations, and procedures for dealing with missing data.
Statistical model specification	Specification of the precise statistical model to be used, hypothesis tests to be run.
Covariates	List of any covariates to be included in analysis.
Subgroup analysis	Description of any heterogeneity analysis to be performed on the data.
Other issues	Other issues include data monitoring plans, stopping rules, and interim looks at the data.

- ▶ Hypothesis clearly laid out; needs clear theory! Parsimony!
- ▶ Less latitude for fishing; p-hacking.
- ▶ Main outcome(s); mediators; robustness; placebos.
- ▶ Proper power analysis that considers HTE.

Advantages

Pre-Analysis Plan Checklist

<i>Item</i>	<i>Brief description</i>
Primary outcome variable	The key variable of interest for the study. If multiple variables are to be examined, one should know how the multiple hypothesis testing will be done.
Secondary outcome variable(s)	Additional variables of interest to be examined.
Variable definitions	Precise variable definitions that specify how the raw data will be transformed into the actual variables to be used for analysis.
Inclusion/Exclusion rules	Rules for including or excluding observations, and procedures for dealing with missing data.
Statistical model specification	Specification of the precise statistical model to be used, hypothesis tests to be run.
Covariates	List of any covariates to be included in analysis.
Subgroup analysis	Description of any heterogeneity analysis to be performed on the data.
Other issues	Other issues include data monitoring plans, stopping rules, and interim looks at the data.

- ▶ Hypothesis clearly laid out; needs clear theory! Parsimony!
- ▶ Less latitude for fishing; p-hacking.
- ▶ Main outcome(s); mediators; robustness; placebos.
- ▶ Proper power analysis that considers HTE.

Advantages

Pre-Analysis Plan Checklist

<i>Item</i>	<i>Brief description</i>
Primary outcome variable	The key variable of interest for the study. If multiple variables are to be examined, one should know how the multiple hypothesis testing will be done.
Secondary outcome variable(s)	Additional variables of interest to be examined.
Variable definitions	Precise variable definitions that specify how the raw data will be transformed into the actual variables to be used for analysis.
Inclusion/Exclusion rules	Rules for including or excluding observations, and procedures for dealing with missing data.
Statistical model specification	Specification of the precise statistical model to be used, hypothesis tests to be run.
Covariates	List of any covariates to be included in analysis.
Subgroup analysis	Description of any heterogeneity analysis to be performed on the data.
Other issues	Other issues include data monitoring plans, stopping rules, and interim looks at the data.

- ▶ Hypothesis clearly laid out; needs clear theory! Parsimony!
- ▶ Less latitude for fishing; p-hacking.
- ▶ Main outcome(s); mediators; robustness; placebos.
- ▶ Proper power analysis that considers HTE.

Advantages

Pre-Analysis Plan Checklist

<i>Item</i>	<i>Brief description</i>
Primary outcome variable	The key variable of interest for the study. If multiple variables are to be examined, one should know how the multiple hypothesis testing will be done.
Secondary outcome variable(s)	Additional variables of interest to be examined.
Variable definitions	Precise variable definitions that specify how the raw data will be transformed into the actual variables to be used for analysis.
Inclusion/Exclusion rules	Rules for including or excluding observations, and procedures for dealing with missing data.
Statistical model specification	Specification of the precise statistical model to be used, hypothesis tests to be run.
Covariates	List of any covariates to be included in analysis.
Subgroup analysis	Description of any heterogeneity analysis to be performed on the data.
Other issues	Other issues include data monitoring plans, stopping rules, and interim looks at the data.

- ▶ Hypothesis clearly laid out; needs clear theory! Parsimony!
- ▶ Less latitude for fishing; p-hacking.
- ▶ Main outcome(s); mediators; robustness; placebos.
- ▶ Proper power analysis that considers HTE.

Empirical limitations of conjoint experiments

Interpreting effects

- ▶ Avg. extensive and intensive margin of preferences.
- ▶ Interpretation depends on group-level treatment uptake.
 - ▶ + effect: failure if + correlation between direction and intensity.
 - ▶ - effect: failure if - correlation between direction and intensity.
 - ▶ Theoretically pre-specify HTEs.
- ▶ Limited external validity.
 - ▶ Sampling matters; post-stratification often necessary.
- ▶ Survey weights are relevant.
 - ▶ Oversampling and unmodeled heteroskedasticity.
- ▶ Non-random attrition often occurs.
 - ▶ Limited internal validity; worst case bounds.

Interpreting effects

- ▶ Avg. extensive and intensive margin of preferences.
- ▶ Interpretation depends on group-level treatment uptake.
 - ▶ + effect: failure if + correlation between direction and intensity.
 - ▶ - effect: failure if - correlation between direction and intensity.
 - ▶ Theoretically pre-specify HTEs.
- ▶ Limited external validity.
 - ▶ Sampling matters; post-stratification often necessary.
- ▶ Survey weights are relevant.
 - ▶ Oversampling and unmodeled heteroskedasticity.
- ▶ Non-random attrition often occurs.
 - ▶ Limited internal validity; worst case bounds.

Interpreting effects

- ▶ Avg. extensive and intensive margin of preferences.
- ▶ Interpretation depends on group-level treatment uptake.
 - ▶ + effect: failure if + correlation between direction and intensity.
 - ▶ - effect: failure if - correlation between direction and intensity.
 - ▶ Theoretically pre-specify HTEs.
- ▶ Limited external validity.
 - ▶ Sampling matters; post-stratification often necessary.
- ▶ Survey weights are relevant.
 - ▶ Oversampling and unmodeled heteroskedasticity.
- ▶ Non-random attrition often occurs.
 - ▶ Limited internal validity; worst case bounds.

Interpreting effects

- ▶ Avg. extensive and intensive margin of preferences.
- ▶ Interpretation depends on group-level treatment uptake.
 - ▶ + effect: failure if + correlation between direction and intensity.
 - ▶ - effect: failure if - correlation between direction and intensity.
 - ▶ Theoretically pre-specify HTEs.
- ▶ Limited external validity.
 - ▶ Sampling matters; post-stratification often necessary.
- ▶ Survey weights are relevant.
 - ▶ Oversampling and unmodeled heteroskedasticity.
- ▶ Non-random attrition often occurs.
 - ▶ Limited internal validity; worst case bounds.

Interpreting effects

- ▶ Avg. extensive and intensive margin of preferences.
- ▶ Interpretation depends on group-level treatment uptake.
 - ▶ + effect: failure if + correlation between direction and intensity.
 - ▶ - effect: failure if - correlation between direction and intensity.
 - ▶ Theoretically pre-specify HTEs.
- ▶ Limited external validity.
 - ▶ Sampling matters; post-stratification often necessary.
- ▶ Survey weights are relevant.
 - ▶ Oversampling and unmodeled heteroskedasticity.
- ▶ Non-random attrition often occurs.
 - ▶ Limited internal validity; worst case bounds.

Thinking about design

- ▶ With enough power and outcomes a result may arise by chance.
 - ▶ Families of outcomes need adjustment (Bonferroni; Holm step-down).
- ▶ Simple randomization may be insufficient.
 - ▶ Unequal treatment assignment; needs doubly robust estimation.
- ▶ Time/location convenience samples have limited external validity.
 - ▶ Sampling and post-stratification often necessary.
- ▶ Evidence of common support and balance are relevant.
- ▶ Cognitive biases need to be ruled out.
 - ▶ Experimenter effects; social desirability; confirmation bias; cognitive limitations, etc.

Thinking about design

- ▶ With enough power and outcomes a result may arise by chance.
 - ▶ Families of outcomes need adjustment (Bonferroni; Holm step-down).
- ▶ Simple randomization may be insufficient.
 - ▶ Unequal treatment assignment; needs doubly robust estimation.
- ▶ Time/location convenience samples have limited external validity.
 - ▶ Sampling and post-stratification often necessary.
- ▶ Evidence of common support and balance are relevant.
- ▶ Cognitive biases need to be ruled out.
 - ▶ Experimenter effects; social desirability; confirmation bias; cognitive limitations, etc.

Thinking about design

- ▶ With enough power and outcomes a result may arise by chance.
 - ▶ Families of outcomes need adjustment (Bonferroni; Holm step-down).
- ▶ Simple randomization may be insufficient.
 - ▶ Unequal treatment assignment; needs doubly robust estimation.
- ▶ Time/location convenience samples have limited external validity.
 - ▶ Sampling and post-stratification often necessary.
- ▶ Evidence of common support and balance are relevant.
- ▶ Cognitive biases need to be ruled out.
 - ▶ Experimenter effects; social desirability; confirmation bias; cognitive limitations, etc.

Thinking about design

- ▶ With enough power and outcomes a result may arise by chance.
 - ▶ Families of outcomes need adjustment (Bonferroni; Holm step-down).
- ▶ Simple randomization may be insufficient.
 - ▶ Unequal treatment assignment; needs doubly robust estimation.
- ▶ Time/location convenience samples have limited external validity.
 - ▶ Sampling and post-stratification often necessary.
- ▶ Evidence of common support and balance are relevant.
- ▶ Cognitive biases need to be ruled out.
 - ▶ Experimenter effects; social desirability; confirmation bias; cognitive limitations, etc.

Thinking about design

- ▶ With enough power and outcomes a result may arise by chance.
 - ▶ Families of outcomes need adjustment (Bonferroni; Holm step-down).
- ▶ Simple randomization may be insufficient.
 - ▶ Unequal treatment assignment; needs doubly robust estimation.
- ▶ Time/location convenience samples have limited external validity.
 - ▶ Sampling and post-stratification often necessary.
- ▶ Evidence of common support and balance are relevant.
- ▶ Cognitive biases need to be ruled out.
 - ▶ Experimenter effects; social desirability; confirmation bias; cognitive limitations, etc.

Addressing biases in design stage

- ▶ Order of questions/treatments matters; randomize ordering.
- ▶ Needs attention checks to evaluate compliance/fatigue.
- ▶ Obfuscation is often needed to address cognitive biases.
- ▶ Measurement needs precision and creativity.
 - ▶ Outcome binary/continuous/categorical. Why?
 - ▶ Likert scales induce cognitive biases; bunching in mid/extreme.
- ▶ Effects may be short lived - follow up surveys may be necessary.
- ▶ Lack of power when not considering HTEs.
 - ▶ Theory matters!

Addressing biases in design stage

- ▶ Order of questions/treatments matters; randomize ordering.
- ▶ Needs attention checks to evaluate compliance/fatigue.
- ▶ Obfuscation is often needed to address cognitive biases.
- ▶ Measurement needs precision and creativity.
 - ▶ Outcome binary/continuous/categorical. Why?
 - ▶ Likert scales induce cognitive biases; bunching in mid/extreme.
- ▶ Effects may be short lived - follow up surveys may be necessary.
- ▶ Lack of power when not considering HTEs.
 - ▶ Theory matters!

Addressing biases in design stage

- ▶ Order of questions/treatments matters; randomize ordering.
- ▶ Needs attention checks to evaluate compliance/fatigue.
- ▶ Obfuscation is often needed to address cognitive biases.
- ▶ Measurement needs precision and creativity.
 - ▶ Outcome binary/continuous/categorical. Why?
 - ▶ Likert scales induce cognitive biases; bunching in mid/extreme.
- ▶ Effects may be short lived - follow up surveys may be necessary.
- ▶ Lack of power when not considering HTEs.
 - ▶ Theory matters!

Addressing biases in design stage

- ▶ Order of questions/treatments matters; randomize ordering.
- ▶ Needs attention checks to evaluate compliance/fatigue.
- ▶ Obfuscation is often needed to address cognitive biases.
- ▶ Measurement needs precision and creativity.
 - ▶ Outcome binary/continuous/categorical. Why?
 - ▶ Likert scales induce cognitive biases; bunching in mid/extreme.
- ▶ Effects may be short lived - follow up surveys may be necessary.
- ▶ Lack of power when not considering HTEs.
 - ▶ Theory matters!

Addressing biases in design stage

- ▶ Order of questions/treatments matters; randomize ordering.
- ▶ Needs attention checks to evaluate compliance/fatigue.
- ▶ Obfuscation is often needed to address cognitive biases.
- ▶ Measurement needs precision and creativity.
 - ▶ Outcome binary/continuous/categorical. Why?
 - ▶ Likert scales induce cognitive biases; bunching in mid/extreme.
- ▶ Effects may be short lived - follow up surveys may be necessary.
- ▶ Lack of power when not considering HTEs.
 - ▶ Theory matters!

Addressing biases in design stage

- ▶ Order of questions/treatments matters; randomize ordering.
- ▶ Needs attention checks to evaluate compliance/fatigue.
- ▶ Obfuscation is often needed to address cognitive biases.
- ▶ Measurement needs precision and creativity.
 - ▶ Outcome binary/continuous/categorical. Why?
 - ▶ Likert scales induce cognitive biases; bunching in mid/extreme.
- ▶ Effects may be short lived - follow up surveys may be necessary.
- ▶ Lack of power when not considering HTEs.
 - ▶ Theory matters!

Next class...

Natural resources, conflict and authoritarianism!