Bounding Causal Effects in Survey Experiments with Noncompliance or Inattention

Matthew Tyler

Rice University (Houston, USA)

EPSA 2025 Madrid

La Carta de Tapas

- Survey experiments: effect of a stimulus

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)
- Ideally: estimate subgroup effect among respondents who **would be** attentive in both conditions (always-attentives)

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)
- Ideally: estimate subgroup effect among respondents who **would be** attentive in both conditions (always-attentives)
- Practical constraints:

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)
- Ideally: estimate subgroup effect among respondents who **would be** attentive in both conditions (always-attentives)
- Practical constraints:
  - We only observe attention in one condition

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)
- Ideally: estimate subgroup effect among respondents who **would be** attentive in both conditions (always-attentives)
- Practical constraints:
  - We only observe attention in one condition
  - Attention checks can be fooled (measurement error)

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)
- Ideally: estimate subgroup effect among respondents who **would be** attentive in both conditions (always-attentives)
- Practical constraints:
  - We only observe attention in one condition
  - Attention checks can be fooled (measurement error)
- This paper

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)
- Ideally: estimate subgroup effect among respondents who **would be** attentive in both conditions (always-attentives)
- Practical constraints:
  - We only observe attention in one condition
  - Attention checks can be fooled (measurement error)
- This paper
  - Sharp bounds for the effect among always-attentives, accounting for attention check measurement error

La Carta de Tapas

- Survey experiments: effect of a stimulus
- Nonprobability samples: lots of inattentive respondents
- **Dangers of inattentive respondents:**
  - Null/weird effects among inattentive respondents
  - Differential attention ($\approx$ differential attrition)
- Ideally: estimate subgroup effect among respondents who **would be** attentive in both conditions (always-attentives)
- Practical constraints:
  - We only observe attention in one condition
  - Attention checks can be fooled (measurement error)
- This paper
  - Sharp bounds for the effect among always-attentives, accounting for attention check measurement error
  - New computational approach to partial identification + confidence intervals

Variables

- Treatment: $D \in \{0, 1\}$

Variables

- Treatment: $D \in \{0, 1\}$
- Categorical/binary outcome: $Y \in \{y_1, \ldots, y_K\}$

Variables

- Treatment: $D \in \{0, 1\}$
- Categorical/binary outcome: $Y \in \{y_1, \ldots, y_K\}$
- Attention check/screener: $S \in \{0, 1\}$

Variables

- Treatment: $D \in \{0, 1\}$
- Categorical/binary outcome: $Y \in \{y_1, \dots, y_K\}$
- Attention check/screener: $S \in \{0, 1\}$
- **Actual attention**: $A \in \{0, 1\}$

Variables

- Treatment: $D \in \{0, 1\}$
- Categorical/binary outcome: $Y \in \{y_1, \ldots, y_K\}$
- Attention check/screener: $S \in \{0, 1\}$
- **Actual attention**: $A \in \{0, 1\}$
    - $A = 1$ if the respondent exerts the effort required to pass the screener without issue

Variables

- Treatment: $D \in \{0, 1\}$
- Categorical/binary outcome: $Y \in \{y_1, \ldots, y_K\}$
- Attention check/screener: $S \in \{0, 1\}$
- **Actual attention**: $A \in \{0, 1\}$
  - $A = 1$ if the respondent exerts the effort required to pass the screener without issue
  - $A = 0$ otherwise

Variables

- Treatment: $D \in \{0, 1\}$
- Categorical/binary outcome: $Y \in \{y_1, \ldots, y_K\}$
- Attention check/screener: $S \in \{0, 1\}$
- **Actual attention**: $A \in \{0, 1\}$
  - $A = 1$ if the respondent exerts the effort required to pass the screener without issue
  - $A = 0$ otherwise
- Note: **"attention" is defined based on screener**—changing the screener will change the meaning of "attentive"

Variables

- Treatment: $D \in \{0, 1\}$
- Categorical/binary outcome: $Y \in \{y_1, \ldots, y_K\}$
- Attention check/screener: $S \in \{0, 1\}$
- **Actual attention**: $A \in \{0, 1\}$
  - $A = 1$ if the respondent exerts the effort required to pass the screener without issue
  - $A = 0$ otherwise
- Note: **"attention" is defined based on screener**—changing the screener will change the meaning of "attentive"
- Key: the screener S is observed, but actual attention A is not!

Timing of the attention check/screener

- Influential methodologists have argued that all attention checks/screeners should be measured pre-treatment

Timing of the attention check/screener

- Influential methodologists have argued that all attention checks/screeners should be measured pre-treatment
- Good intentions: avoids post-treatment conditioning problems

Timing of the attention check/screener

- Influential methodologists have argued that all attention checks/screeners should be measured pre-treatment
- Good intentions: avoids post-treatment conditioning problems
- However: ignores the potential for differential attention

Timing of the attention check/screener

- Influential methodologists have argued that all attention checks/screeners should be measured pre-treatment
- Good intentions: avoids post-treatment conditioning problems
- However: ignores the potential for differential attention
- Also ignores what we care most about: attention to the stimulus

Timing of the attention check/screener

- Influential methodologists have argued that all attention checks/screeners should be measured pre-treatment
- Good intentions: avoids post-treatment conditioning problems
- However: ignores the potential for differential attention
- Also ignores what we care most about: attention to the stimulus
- Thus, we need a method for post-treatment attention checks/screeners.

Core assumptions

- A0: n iid draws from some superpopulation.

Core assumptions

- A0: n iid draws from some superpopulation.
- A1: SUTVA.

$$Y = Y(D), \quad A = A(D), \quad S = S(D)$$

Core assumptions

- A0: n iid draws from some superpopulation.
- A1: SUTVA.

$$Y = Y(D), \quad A = A(D), \quad S = S(D)$$

- A2: D randomly assigned.

Core assumptions

- A0: n iid draws from some superpopulation.
- A1: SUTVA.

$$Y = Y(D), \quad A = A(D), \quad S = S(D)$$

- A2: D randomly assigned.
- A3: Known false positive/negative rate.

$$P[S(d) = 1 \mid A(d) = 1] = 1$$
$$P[S(d) = 1 \mid A(d) = 0] = \alpha_d$$

Sensitivity analysis: explore different values of α_d

Optional assumptions

- A4: False positive rate does not depend on Y.

$$P[S(d) = 1 \mid A(d) = 0, Y(d) = y] = \alpha_d$$

Optional assumptions

- A4: False positive rate does not depend on Y.

$$P[S(d) = 1 \mid A(d) = 0, Y(d) = y] = \alpha_d$$

- A5: Attention monotonicity. For all respondents, either

$$A(1) \geq A(0)$$

or the reverse. [See Lee (2009).]

Optional assumptions

- A4: False positive rate does not depend on Y.

$$P[S(d) = 1 \mid A(d) = 0, Y(d) = y] = \alpha_d$$

- A5: Attention monotonicity. For all respondents, either

$$A(1) \geq A(0)$$

or the reverse. [See Lee (2009).]

- A6: Fixed attention/screener.

$$A(1) = A(0), \quad S(1) = S(0)$$

(Only guaranteed for pre-treatment screeners.)

Estimand

- Average treatment effect (ATE): contaminated by inattention

Estimand

- Average treatment effect (ATE): contaminated by inattention
- Instead: look at effect among always-attentive stratum:

$$A(1) = 1, A(0) = 1$$

Estimand

- Average treatment effect (ATE): contaminated by inattention
- Instead: look at effect among always-attentive stratum:

$$A(1) = 1, A(0) = 1$$

- Average treatment effect among always-attentive (ATAC):

$$\text{ATAC} = E[Y(1) - Y(0) \mid A(0) = 1, A(1) = 1]$$

Estimand

- Average treatment effect (ATE): contaminated by inattention
- Instead: look at effect among always-attentive stratum:

$$A(1) = 1, A(0) = 1$$

- Average treatment effect among always-attentive (ATAC):

$$\text{ATAC} = E[Y(1) - Y(0) \mid A(0) = 1, A(1) = 1]$$

- Not <u>point</u> identifiable: observe either $A(0)$ or $A(1)$, not both

Estimand

- Average treatment effect (ATE): contaminated by inattention
- Instead: look at effect among always-attentive stratum:

$$A(1) = 1, A(0) = 1$$

- Average treatment effect among always-attentive (ATAC):

$$\text{ATAC} = E[Y(1) - Y(0) \mid A(0) = 1, A(1) = 1]$$

- Not <u>point</u> identifiable: observe either $A(0)$ or $A(1)$, not both
- If $S = A$ always, then ATAC is boundable (Lee 2009)

Estimand

- Average treatment effect (ATE): contaminated by inattention
- Instead: look at effect among always-attentive stratum:

$$A(1) = 1, A(0) = 1$$

- Average treatment effect among always-attentive (ATAC):

$$\text{ATAC} = E[Y(1) - Y(0) \mid A(0) = 1, A(1) = 1]$$

- Not point identifiable: observe either $A(0)$ or $A(1)$, not both
- If $S = A$ always, then ATAC is boundable (Lee 2009)
- No method for bounding ATAC when $S \neq A$

New computational approach

- Parameterize joint distribution of all potential outcomes

$$\begin{aligned}\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\, & A(0) = a_0, A(1) = a_1, \\ & S(0) = s_0, S(1) = s_1, \\ & Y(0) = y_j, Y(1) = y_k \,]\end{aligned}$$

New computational approach

- Parameterize joint distribution of all potential outcomes

$$\begin{aligned}\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\, & A(0) = a_0, A(1) = a_1, \\ & S(0) = s_0, S(1) = s_1, \\ & Y(0) = y_j, Y(1) = y_k \,]\end{aligned}$$

- Example: when $K = 2$, $\text{vec}(\pi^*)$ is 64-length vector

New computational approach

- Parameterize joint distribution of all potential outcomes

$$\begin{aligned}\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\, & A(0) = a_0, A(1) = a_1, \\ & S(0) = s_0, S(1) = s_1, \\ & Y(0) = y_j, Y(1) = y_k \,]\end{aligned}$$

- Example: when $K = 2$, vec(π^*) is 64-length vector
- Assumptions A1-A6 are **linear** constraints on vec(π^*)

New computational approach

- Parameterize joint distribution of all potential outcomes

$$\begin{aligned}\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\, & A(0) = a_0, A(1) = a_1, \\ & S(0) = s_0, S(1) = s_1, \\ & Y(0) = y_j, Y(1) = y_k \,]\end{aligned}$$

- Example: when $K = 2$, vec(π^*) is 64-length vector
- Assumptions A1-A6 are **linear** constraints on vec(π^*)
- Convex optimization: minimize/maximize feasible ATAC(π).

New computational approach

- Parameterize joint distribution of all potential outcomes

$$\begin{aligned}\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\,A(0) = a_0, A(1) = a_1,\\ S(0) = s_0, S(1) = s_1,\\ Y(0) = y_j, Y(1) = y_k\,]\end{aligned}$$

- Example: when $K = 2$, vec(π^*) is 64-length vector
- Assumptions A1-A6 are **linear** constraints on vec(π^*)
- Convex optimization: minimize/maximize feasible ATAC(π).
    - Estimated bounds (EBs) ignore sampling variance.s

New computational approach

- Parameterize joint distribution of all potential outcomes

$$\begin{aligned}\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\, & A(0) = a_0, A(1) = a_1, \\ & S(0) = s_0, S(1) = s_1, \\ & Y(0) = y_j, Y(1) = y_k \,]\end{aligned}$$

- Example: when $K = 2$, $\text{vec}(\pi^*)$ is 64-length vector
- Assumptions A1-A6 are **linear** constraints on $\text{vec}(\pi^*)$
- Convex optimization: minimize/maximize feasible $\text{ATAC}(\pi)$.
    - Estimated bounds (EBs) ignore sampling variance.s
    - Confidence intervals (CIs) account for sampling variance.

New computational approach

- Parameterize joint distribution of all potential outcomes

$$\begin{aligned}\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\, & A(0) = a_0, A(1) = a_1, \\ & S(0) = s_0, S(1) = s_1, \\ & Y(0) = y_j, Y(1) = y_k \,]\end{aligned}$$

- Example: when $K = 2$, $\text{vec}(\pi^*)$ is 64-length vector
- Assumptions A1-A6 are **linear** constraints on $\text{vec}(\pi^*)$
- Convex optimization: minimize/maximize feasible $\text{ATAC}(\pi)$.
    - Estimated bounds (EBs) ignore sampling variance.s
    - Confidence intervals (CIs) account for sampling variance.
- Theorem 2: CIs have desired asymptotic coverage rate.

New computational approach

- Parameterize joint distribution of all potential outcomes

$$
\begin{aligned}
\pi^*(a_0, a_1, s_0, s_1, j, k) = P[\, & A(0) = a_0, A(1) = a_1, \\
& S(0) = s_0, S(1) = s_1, \\
& Y(0) = y_j, Y(1) = y_k \,]
\end{aligned}
$$

- Example: when $K = 2$, $\text{vec}(\pi^*)$ is 64-length vector
- Assumptions A1-A6 are **linear** constraints on $\text{vec}(\pi^*)$
- Convex optimization: minimize/maximize feasible $\text{ATAC}(\pi)$.
  - Estimated bounds (EBs) ignore sampling variance.s
  - Confidence intervals (CIs) account for sampling variance.
- Theorem 2: CIs have desired asymptotic coverage rate.
  - CIs much smaller than `autobounds` (Duarte et al. 2024).

Reanalysis of Dancygier and Wiedemann (2024)

- Survey experiment on support for expropriating corporate landlords in German cities.

Reanalysis of Dancygier and Wiedemann (2024)

- Survey experiment on support for expropriating corporate landlords in German cities.
- Control: paragraph about how broccoli is healthy

Reanalysis of Dancygier and Wiedemann (2024)

- Survey experiment on support for expropriating corporate landlords in German cities.
- Control: paragraph about how broccoli is healthy
- Treatment: paragraph about how corporate landlords are buying up homes as a (speculative) investment

Reanalysis of Dancygier and Wiedemann (2024)

- Survey experiment on support for expropriating corporate landlords in German cities.
- Control: paragraph about how broccoli is healthy
- Treatment: paragraph about how corporate landlords are buying up homes as a (speculative) investment
- Screener: 4-options multiple choice checking what each respondent read

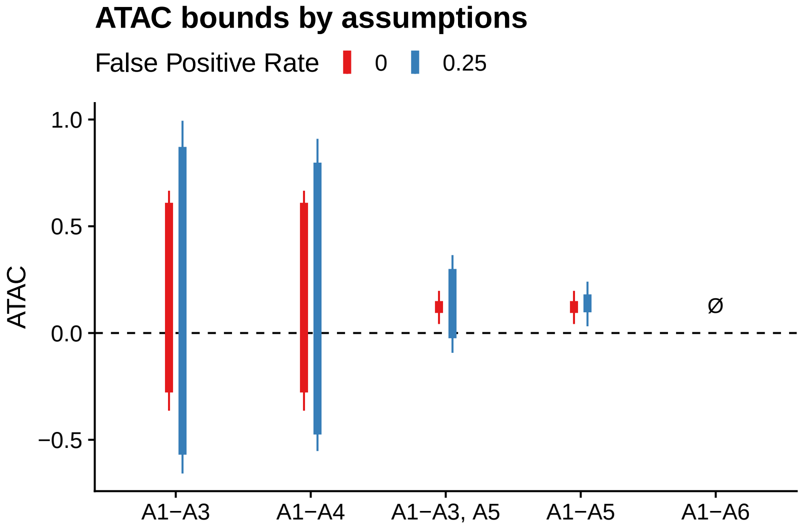
Reanalysis of Dancygier and Wiedemann (2024)

- Survey experiment on support for expropriating corporate landlords in German cities.
- Control: paragraph about how broccoli is healthy
- Treatment: paragraph about how corporate landlords are buying up homes as a (speculative) investment
- Screener: 4-options multiple choice checking what each respondent read
- Treatment increased average expropriation support (60% $\rightarrow$ 69%**) and decreased check passage (78% $\rightarrow$ 74%**)

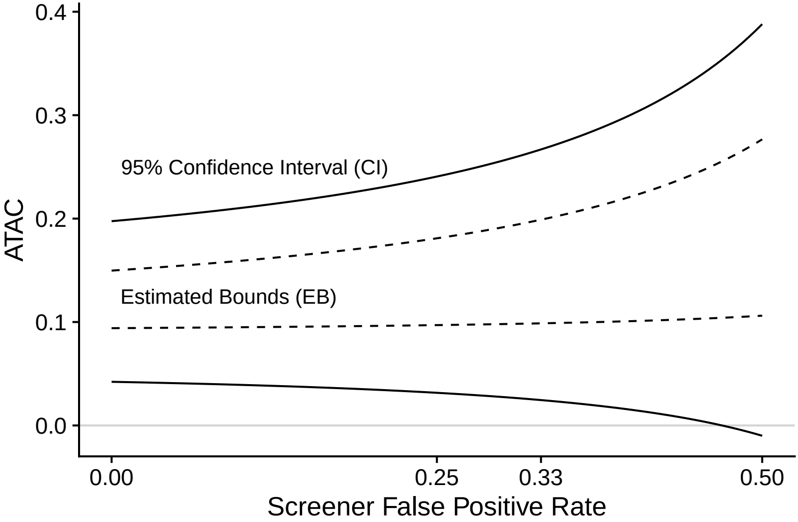
Reanalysis of Dancygier and Wiedemann (2024)

- Survey experiment on support for expropriating corporate landlords in German cities.
- Control: paragraph about how broccoli is healthy
- Treatment: paragraph about how corporate landlords are buying up homes as a (speculative) investment
- Screener: 4-options multiple choice checking what each respondent read
- Treatment increased average expropriation support (60% $\rightarrow$ 69%**) and decreased check passage (78% $\rightarrow$ 74%**)
- What's the effect among always-attentive respondents?

Assumption sensitivity



False positive rate sensitivity (A1-A5)



Conclusion

- Developed a new method for bounding causal effects in survey experiments with post-treatment inattention.

Conclusion

- Developed a new method for bounding causal effects in survey experiments with post-treatment inattention.
- Fast computational method for computing bounds; established confidence interval coverage rate.

Conclusion

- Developed a new method for bounding causal effects in survey experiments with post-treatment inattention.
- Fast computational method for computing bounds; established confidence interval coverage rate.
- Demonstrated the method's utility with a reanalysis of Dancygier and Wiedemann (2024).

Conclusion

- Developed a new method for bounding causal effects in survey experiments with post-treatment inattention.
- Fast computational method for computing bounds; established confidence interval coverage rate.
- Demonstrated the method's utility with a reanalysis of Dancygier and Wiedemann (2024).
- Future work: different sets of assumptions to achieve tighter bounds without attention monotonicity (A5).

Conclusion

- Developed a new method for bounding causal effects in survey experiments with post-treatment inattention.
- Fast computational method for computing bounds; established confidence interval coverage rate.
- Demonstrated the method's utility with a reanalysis of Dancygier and Wiedemann (2024).
- Future work: different sets of assumptions to achieve tighter bounds without attention monotonicity (A5).
- Future work: Apply the computational method to other causal inference settings.

Thank You

- Matthew Tyler @ Rice University (Houston, USA)

Thank You

- Matthew Tyler @ Rice University (Houston, USA)
- Measurement, causal inference, survey methods

Thank You

- Matthew Tyler @ Rice University (Houston, USA)
- Measurement, causal inference, survey methods
- Contact: mdtyler@rice.edu