Training Policy Makers in Econometrics

Sultan Mehmood, Shaheen Naseer, Daniel L. Chen

Causal Effects of "Mastering Metrics"?

The credibility revolution has triggered a paradigm shift

(Angrist & Pischke, 2010; Abadie & Cattaneo 2018; Banerjee & Duflo, 2020)

New focus on causal inference research designs in social sciences

Randomized trials are presented as the ideal research design

(Leamer 1983; Lalonde, 1986; Angrist Pischke, 2010)

Causal Effects of "Mastering Metrics"?

The credibility revolution has triggered a paradigm shift

(Angrist & Pischke, 2010; Abadie & Cattaneo 2018; Banerjee & Duflo, 2020)

New focus on causal inference research designs in social sciences

Randomized trials are presented as the ideal research design

(Leamer 1983; Lalonde, 1986; Angrist Pischke, 2010)

Causal Effects of "Mastering Metrics"?

The credibility revolution has triggered a paradigm shift

(Angrist & Pischke, 2010; Abadie & Cattaneo 2018; Banerjee & Duflo, 2020)

New focus on causal inference research designs in social sciences

Randomized trials are presented as the ideal research design

(Leamer 1983; Lalonde, 1986; Angrist Pischke, 2010)

Motivation

- Policymakers demand causal evidence (Hjort et al. 2021, AER)
- But they are unlikely to change their policy choices in response to new evidence
- Large body of evidence is accumulating that policymakers are highly averse to shifting their beliefs and policy choices (Baekgaard et al.2019; Banuri et al. 2019; Vivalt and Coville 2021)
- Sticking to priors and being inattentive to evidence may stymie the implementation of good policies and hurt economic development (Kremer, Rao, Schilbach 2019)

Motivation

- Policymakers demand causal evidence (Hjort et al. 2021, AER)
- But they are unlikely to change their policy choices in response to new evidence
- Large body of evidence is accumulating that policymakers are highly averse to shifting their beliefs and policy choices (Baekgaard et al.2019; Banuri et al. 2019; Vivalt and Coville 2021)
- Sticking to priors and being inattentive to evidence may stymie the implementation of good policies and hurt economic development (Kremer, Rao, Schilbach 2019)

Motivation

- Policymakers demand causal evidence (Hjort et al. 2021, AER)
- But they are unlikely to change their policy choices in response to new evidence
- Large body of evidence is accumulating that policymakers are highly averse to shifting their beliefs and policy choices (Baekgaard et al.2019; Banuri et al. 2019; Vivalt and Coville 2021)
- Sticking to priors and being inattentive to evidence may stymie the implementation of good policies and hurt economic development (Kremer, Rao, Schilbach 2019)

Motivation

- Policymakers demand causal evidence (Hjort et al. 2021, AER)
- But they are unlikely to change their policy choices in response to new evidence
- Large body of evidence is accumulating that policymakers are highly averse to shifting their beliefs and policy choices (Baekgaard et al.2019; Banuri et al. 2019; Vivalt and Coville 2021)
- Sticking to priors and being inattentive to evidence may stymie the implementation of good policies and hurt economic development (Kremer, Rao, Schilbach 2019)

Motivation

- Policymakers demand causal evidence (Hjort et al. 2021, AER)
- But they are unlikely to change their policy choices in response to new evidence
- Large body of evidence is accumulating that policymakers are highly averse to shifting their beliefs and policy choices (Baekgaard et al.2019; Banuri et al. 2019; Vivalt and Coville 2021)
- Sticking to priors and being inattentive to evidence may stymie the implementation of good policies and hurt economic development (Kremer, Rao, Schilbach 2019)

Motivation

- How can policymakers be made more receptive to evidence?
- Will training them in concepts associated with the credibility revolution make them more likely to shift their beliefs?
- Will it induce them to change their policy choices?

Motivation

- How can policymakers be made more receptive to evidence?
- Will training them in concepts associated with the credibility revolution make them more likely to shift their beliefs?
- Will it induce them to change their policy choices?

Motivation

- How can policymakers be made more receptive to evidence?
- Will training them in concepts associated with the credibility revolution make them more likely to shift their beliefs?
- Will it induce them to change their policy choices?

Motivation

- How can policymakers be made more receptive to evidence?
- Will training them in concepts associated with the credibility revolution make them more likely to shift their beliefs?
- Will it induce them to change their policy choices?

What is the Causal Effects of Causal Thinking?

- Does it change willingness-to-pay for causal evidence?
- Does it affect responsiveness to causal evidence?

What is the Causal Effects of Causal Thinking?

- Does it change willingness-to-pay for causal evidence?
- Does it affect responsiveness to causal evidence?

What is the Causal Effects of Causal Thinking?

- Does it change willingness-to-pay for causal evidence?
- Does it affect responsiveness to causal evidence?

Improving State Capacity

Personnel economics of the state (Finan, Olken, and Pande 2017)

selection

incentives

monitoring

Assumes civil servants are fixed and not malleable

but a growing body of causal evidence suggests they can be taught

Improving State Capacity

Personnel economics of the state (Finan, Olken, and Pande 2017)

selection

incentives

monitoring

Assumes civil servants are fixed and not malleable

but a growing body of causal evidence suggests they can be taught

Improving State Capacity

Personnel economics of the state (Finan, Olken, and Pande 2017)

selection

incentives

monitoring

Assumes civil servants are fixed and not malleable

but a growing body of causal evidence suggests they can be taught

Improving State Capacity

Personnel economics of the state (Finan, Olken, and Pande 2017)

selection

incentives

monitoring

Assumes civil servants are fixed and not malleable

but a growing body of causal evidence suggests they can be taught

Paradigm Shift in Science

Causa effects of paradigm shifts using a field experiment

We used training of the paradigm as its instrument

After receiving a book, an intense training workshop based on SEL

Summarize

Apply

Structured discussion

Present

Videos

Economics and financial literacy (Lusardi and Mitchell 2014)

Paradigm Shift in Science

Causa effects of paradigm shifts using a field experiment

We used training of the paradigm as its instrument

After receiving a book, an intense training workshop based on SEL

Summarize

Apply

Structured discussion

Present

Videos

Economics and financial literacy (Lusardi and Mitchell 2014)

Paradigm Shift in Science

Causa effects of paradigm shifts using a field experiment

We used training of the paradigm as its instrument

After receiving a book, an intense training workshop based on SEL

Summarize

Apply

Structured discussion

Present

Videos

Economics and financial literacy (Lusardi and Mitchell 2014)

Paradigm Shift in Science

Causa effects of paradigm shifts using a field experiment

We used training of the paradigm as its instrument

After receiving a book, an intense training workshop based on SEL

Summarize

Apply

Structured discussion

Present

Videos

Economics and financial literacy (Lusardi and Mitchell 2014)

Causal Effects of "Mastering Metrics"?

We separate demand for education from impacts of education

Simplified Becker-Degroot-Marshak mechanism (Berry, et al. 2020)

Choose mastering metrics vs. self-help book

Lottery with high or low probability

Outcome of lottery allows estimating causal effects

Causal Effects of "Mastering Metrics"?

We separate demand for education from impacts of education

Simplified Becker-Degroot-Marshak mechanism (Berry, et al. 2020)

Choose mastering metrics vs. self-help book

Lottery with high or low probability

Outcome of lottery allows estimating causal effects

Causal Effects of "Mastering Metrics"?

We separate demand for education from impacts of education

Simplified Becker-Degroot-Marshak mechanism (Berry, et al. 2020)

Choose mastering metrics vs. self-help book

Lottery with high or low probability

Outcome of lottery allows estimating causal effects

Causal Effects of "Mastering Metrics"?

We separate demand for education from impacts of education

Simplified Becker-Degroot-Marshak mechanism (Berry, et al. 2020)

Choose mastering metrics vs. self-help book

Lottery with high or low probability

Outcome of lottery allows estimating causal effects

Causal Effects of "Mastering Metrics"?

We separate demand for education from impacts of education

Simplified Becker-Degroot-Marshak mechanism (Berry, et al. 2020)

Choose mastering metrics vs. self-help book

Lottery with high or low probability

Outcome of lottery allows estimating causal effects

Preview

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

After 6 months, treated individuals' ratings on the importance of quantitative analysis increase by 50%

Writings reflect an increase in perceived importance of causal inference and understanding of causal concepts

2) improved test scores on public policy and research methods

Performance in national research methods and public policy examsimproves by 0.5-0.8 sigma

3) willingness to pay for quantitative data, only for RCTs ($\downarrow$ correlations)

WTP for commissioning RCTs using public funding increases by 300% and decreases by 50% for correlations

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)

Treated ministers are twice as likely to choose a (deworming) policy for which there is RCT evidence.

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Preview

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

After 6 months, treated individuals' ratings on the importance of quantitative analysis increase by 50%

Writings reflect an increase in perceived importance of causal inference and understanding of causal concepts

2) improved test scores on public policy and research methods

Performance in national research methods and public policy examsimproves by 0.5-0.8 sigma

3) willingness to pay for quantitative data, only for RCTs (↓ correlations)

WTP for commissioning RCTs using public funding increases by 300% and decreases by 50% for correlations

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)

Treated ministers are twice as likely to choose a (deworming) policy for which there is RCT evidence.

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Preview

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

After 6 months, treated individuals' ratings on the importance of quantitative analysis increase by 50%

Writings reflect an increase in perceived importance of causal inference and understanding of causal concepts

2) improved test scores on public policy and research methods

Performance in national research methods and public policy examsimproves by 0.5-0.8 sigma

3) willingness to pay for quantitative data, only for RCTs ($\downarrow$ correlations)

WTP for commissioning RCTs using public funding increases by 300% and decreases by 50% for correlations

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)

Treated ministers are twice as likely to choose a (deworming) policy for which there is RCT evidence.

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Preview

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

After 6 months, treated individuals' ratings on the importance of quantitative analysis increase by 50%

Writings reflect an increase in perceived importance of causal inference and understanding of causal concepts

2) improved test scores on public policy and research methods

Performance in national research methods and public policy examsimproves by 0.5-0.8 sigma

3) willingness to pay for quantitative data, only for RCTs (↓ correlations)

WTP for commissioning RCTs using public funding increases by 300% and decreases by 50% for correlations

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)

Treated ministers are twice as likely to choose a (deworming) policy for which there is RCT evidence.

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Preview

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

After 6 months, treated individuals' ratings on the importance of quantitative analysis increase by 50%

Writings reflect an increase in perceived importance of causal inference and understanding of causal concepts

2) improved test scores on public policy and research methods

Performance in national research methods and public policy examsimproves by 0.5-0.8 sigma

3) willingness to pay for quantitative data, only for RCTs ($\downarrow$ correlations)

WTP for commissioning RCTs using public funding increases by 300% and decreases by 50% for correlations

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)

Treated ministers are twice as likely to choose a (deworming) policy for which there is RCT evidence.

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Preview

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

After 6 months, treated individuals' ratings on the importance of quantitative analysis increase by 50%

Writings reflect an increase in perceived importance of causal inference and understanding of causal concepts

2) improved test scores on public policy and research methods

Performance in national research methods and public policy examsimproves by 0.5-0.8 sigma

3) willingness to pay for quantitative data, only for RCTs (↓ correlations)

WTP for commissioning RCTs using public funding increases by 300% and decreases by 50% for correlations

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)

Treated ministers are twice as likely to choose a (deworming) policy for which there is RCT evidence.

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Pakistan Deputy Ministers

- Policy experts who advise the President, Prime Minister, cabinet ministers, governors and police chiefs.
- "key wheels on which the entire engine of the state runs" (Central Superior Services, 2019)
- Select 1.5% of test-takers
  - 14,521 candidates => 213 qualified
- Mandatory attendance and high-stakes training academy
  - Training workshop counted 5% towards requirements
  - Affects career trajectories, transfers, promotions

Pakistan Deputy Ministers

- Policy experts who advise the President, Prime Minister, cabinet ministers, governors and police chiefs.
- "key wheels on which the entire engine of the state runs" (Central Superior Services, 2019)
- Select 1.5% of test-takers
  - 14,521 candidates => 213 qualified
- Mandatory attendance and high-stakes training academy
  - Training workshop counted 5% towards requirements
  - Affects career trajectories, transfers, promotions

Pakistan Deputy Ministers

- Policy experts who advise the President, Prime Minister, cabinet ministers, governors and police chiefs.
- "key wheels on which the entire engine of the state runs" (Central Superior Services, 2019)
- Select 1.5% of test-takers
  - 14,521 candidates => 213 qualified
- Mandatory attendance and high-stakes training academy
  - Training workshop counted 5% towards requirements
  - Affects career trajectories, transfers, promotions

Pakistan Deputy Ministers

- Policy experts who advise the President, Prime Minister, cabinet ministers, governors and police chiefs.
- "key wheels on which the entire engine of the state runs" (Central Superior Services, 2019)
- Select 1.5% of test-takers
  - 14,521 candidates => 213 qualified
- Mandatory attendance and high-stakes training academy
  - Training workshop counted 5% towards requirements
  - Affects career trajectories, transfers, promotions

Timeline

- October 20, 2020 baseline survey
  - choose one of the two books (70-30 lottery)
  - Mastering Metrics (Joshua Angrist and Jörn-Steffen Pischke)
  - Mindsight: The New Science of Personal Transformation (Daniel J. Siegel)

Timeline

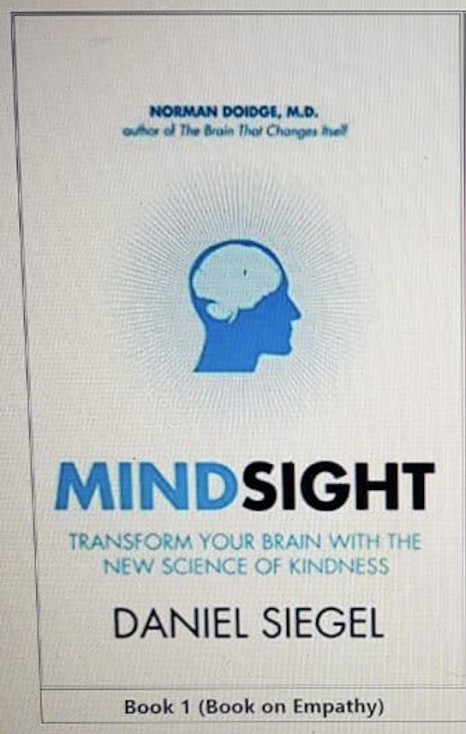
- October 20, 2020 baseline survey
  - choose one of the two books (70-30 lottery)
  - Mastering Metrics (Joshua Angrist and Jörn-Steffen Pischke)
  - Mindsight: The New Science of Personal Transformation (Daniel J. Siegel)

Timeline

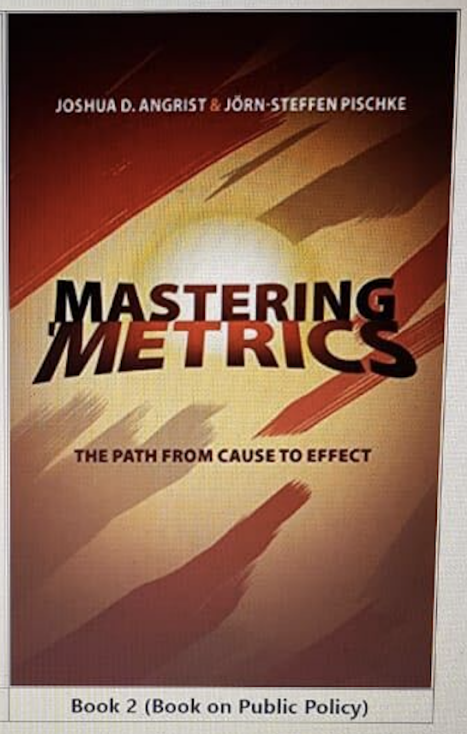
- October 20, 2020 baseline survey
  - choose one of the two books (70-30 lottery)
  - Mastering Metrics (Joshua Angrist and Jörn-Steffen Pischke)
  - Mindsight: The New Science of Personal Transformation (Daniel J. Siegel)

Book Choice (lottery)

As stated earlier in the session, you will receive a free book. Please tell us which book you prefer.



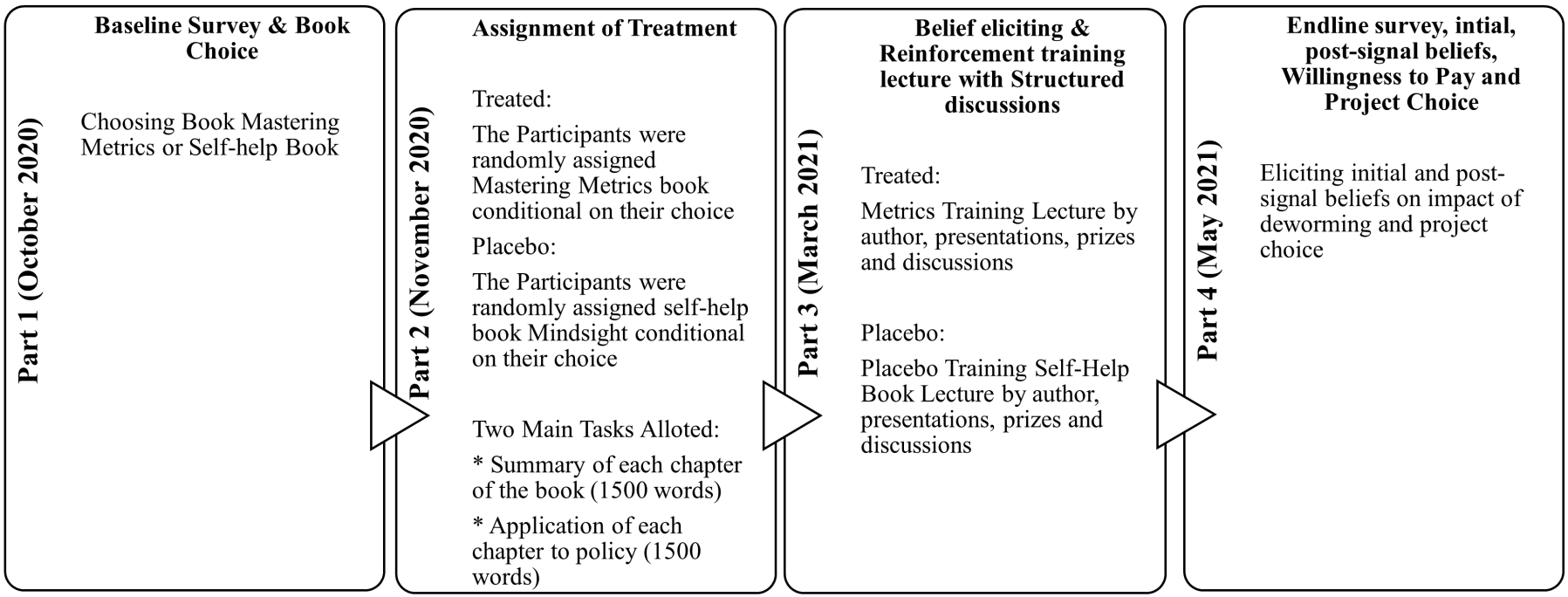
Book 1 (Book on Empathy)



Book 2 (Book on Public Policy)

We hope to send you the book of your choice. Therefore, it is very important that you provide us the mailing address where you would like to receive the book in the next two weeks. You will also receive questions about the book from the Civil Service Academy, so make sure you receive the book on time.

Timeline


Part 1 (October 2020)
Baseline Survey & Book Choice
Choosing Book Mastering Metrics or Self-help Book
Part 2 (November 2020)
Assignment of Treatment
Treated:
The Participants were randomly assigned Mastering Metrics book conditional on their choice
Placebo:
The Participants were randomly assigned self-help book Mindsight conditional on their choice
Two Main Tasks Alloted:
* Summary of each chapter of the book (1500 words)
* Application of each chapter to policy (1500 words)
Part 3 (March 2021)
Belief eliciting & Reinforcement training lecture with Structured discussions
Treated:
Metrics Training Lecture by author, presentations, prizes and discussions
Placebo:
Placebo Training Self-Help Book Lecture by author, presentations, prizes and discussions
Part 4 (May 2021)
Endline survey, intial, post-signal beliefs, Willingness to Pay and Project Choice
Eliciting initial and post-signal beliefs on impact of deworming and project choice

Work with CSA

- November 10, assignment of treatment
- December 10, deadline of first assignment
  - Main Task 1: Provide a summary of the whole book of around 1500 words.
  - Main Task 2: Provide an analysis of how you would apply the lessons learned from the book in your future job. This again should be around 1500 words.

Work with CSA

- November 10, assignment of treatment
- December 10, deadline of first assignment
  - Main Task 1: Provide a summary of the whole book of around 1500 words.
  - Main Task 2: Provide an analysis of how you would apply the lessons learned from the book in your future job. This again should be around 1500 words.

Work with CSA

- November 10, assignment of treatment
- December 10, deadline of first assignment
  - Main Task 1: Provide a summary of the whole book of around 1500 words.
  - Main Task 2: Provide an analysis of how you would apply the lessons learned from the book in your future job. This again should be around 1500 words.

Work with CSA

- November 10, assignment of treatment
- December 10, deadline of first assignment
  - Main Task 1: Provide a summary of the whole book of around 1500 words.
  - Main Task 2: Provide an analysis of how you would apply the lessons learned from the book in your future job. This again should be around 1500 words.

Work with CSA

- March 10, 2021 (zoom)
  - Belief Elicitation I
  - Reinforcement lecture of 30 minutes
    - Links in oTree to watch videos
    - Students could pick which video (behavioral measure of defier)
  - Live discussion of 30 minutes
    - a) what do you think is the main point of the lecture?
    - b) how can you apply the concepts learned in this lecture to future job?
  - Belief Elicitation II

Work with CSA

- March 10, 2021 (zoom)
  - Belief Elicitation I
  - Reinforcement lecture of 30 minutes
    - ⋆ Links in oTree to watch videos
    - ⋆ Students could pick which video (behavioral measure of defier)
  - Live discussion of 30 minutes
    - ⋆ a) what do you think is the main point of the lecture?
    - ⋆ b) how can you apply the concepts learned in this lecture to future job?
  - Belief Elicitation II

Work with CSA

- March 10, 2021 (zoom)
  - Belief Elicitation I
  - Reinforcement lecture of 30 minutes
    - ⋆ Links in oTree to watch videos
    - ⋆ Students could pick which video (behavioral measure of defier)
  - Live discussion of 30 minutes
    - ⋆ a) what do you think is the main point of the lecture?
    - ⋆ b) how can you apply the concepts learned in this lecture to future job?
  - Belief Elicitation II

Work with CSA

- March 10, 2021 (zoom)
  - Belief Elicitation I
  - Reinforcement lecture of 30 minutes
    - ⋆ Links in oTree to watch videos
    - ⋆ Students could pick which video (behavioral measure of defier)
  - Live discussion of 30 minutes
    - ⋆ a) what do you think is the main point of the lecture?
    - ⋆ b) how can you apply the concepts learned in this lecture to future job?
  - Belief Elicitation II

Work with CSA

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - (1) expert opinion about the project choice from a senior bureaucrat
    - (2) RCT assessing the impact of deworming on hourly earnings
    - (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Work with CSA

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Work with CSA

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Work with CSA

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Work with CSA

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Treatment: Mastering Metrics

- Randomized control trials and quasi-experimental methods
  - Regressions
  - Instrumental variables
  - Difference-in-Differences and RDD
- With particular focus on public policy applications
- Written for undergraduates
  - all policymakers had at least bachelor's degree

Treatment: Mastering Metrics

- Randomized control trials and quasi-experimental methods
  - Regressions
  - Instrumental variables
  - Difference-in-Differences and RDD
- With particular focus on public policy applications
- Written for undergraduates
  - all policymakers had at least bachelor's degree

Placebo: Mindsight

- Popular self-help book emphasizing "personal transformation"
  - 1641 google scholar cites since 2010
  - Amazon #39 in Health, Mind, and Body Reference
  - #264 in all Cognitive Psychology
- Contains 'recipes' to improve your mental health
  - developing an attitude so that one is not "unreasonably angry or upset"
  - "explore the essence of who we are"
  - develop a more positive outlook in life

Placebo: Mindsight

- Popular self-help book emphasizing "personal transformation"
  - 1641 google scholar cites since 2010
  - Amazon #39 in Health, Mind, and Body Reference
  - #264 in all Cognitive Psychology
- Contains 'recipes' to improve your mental health
  - developing an attitude so that one is not "unreasonably angry or upset"
  - "explore the essence of who we are"
  - develop a more positive outlook in life

Empirical Methodology

- $Y_i = \theta + \alpha$ Metrics Assigned$_i$ + β Metrics Chosen$_i$ + $W'_i\psi + \epsilon_i$
- Subscript i represent policy-maker
- Y represents a set of attitudes, beliefs, and choices
- Metrics Chosen denotes an indicator variable that switches on for **choosing** Mastering Metrics vs. Mindsight
- Metrics Assigned denotes an indicator variable that switches on for being randomly **assigned** the book
- W is a vector of individual controls from administrative data
  - written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining public service, age, education, foreign visits and occupational group dummies.
- α provides us the impact of econometrics training
- cluster standard errors at the individual level (level of randomization)

Empirical Methodology

- $Y_i = \theta + \alpha$ Metrics Assigned$_i$ + β Metrics Chosen$_i$ + $W'_i\psi + \epsilon_i$
- Subscript i represent policy-maker
- Y represents a set of attitudes, beliefs, and choices
- Metrics Chosen denotes an indicator variable that switches on for **choosing** Mastering Metrics vs. Mindsight
- Metrics Assigned denotes an indicator variable that switches on for being randomly **assigned** the book
- W is a vector of individual controls from administrative data
  - written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining public service, age, education, foreign visits and occupational group dummies.
- α provides us the impact of econometrics training
- cluster standard errors at the individual level (level of randomization)

Empirical Methodology

- $Y_i = \theta + \alpha$ Metrics Assigned$_i$ + β Metrics Chosen$_i$ + $W'_i\psi + \epsilon_i$
- Subscript *i* represent policy-maker
- Y represents a set of attitudes, beliefs, and choices
- Metrics Chosen denotes an indicator variable that switches on for **choosing** Mastering Metrics vs. Mindsight
- Metrics Assigned denotes an indicator variable that switches on for being randomly **assigned** the book
- W is a vector of individual controls from administrative data
  - written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining public service, age, education, foreign visits and occupational group dummies.
- α provides us the impact of econometrics training
- cluster standard errors at the individual level (level of randomization)

Empirical Methodology

- $Y_i = \theta + \alpha$ Metrics Assigned$_i$ + β Metrics Chosen$_i$ + $W'_i\psi + \epsilon_i$
- Subscript i represent policy-maker
- Y represents a set of attitudes, beliefs, and choices
- Metrics Chosen denotes an indicator variable that switches on for **choosing** Mastering Metrics vs. Mindsight
- Metrics Assigned denotes an indicator variable that switches on for being randomly **assigned** the book
- W is a vector of individual controls from administrative data
  - written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining public service, age, education, foreign visits and occupational group dummies.
- α provides us the impact of econometrics training
- cluster standard errors at the individual level (level of randomization)

Empirical Methodology

- $Y_i = \theta + \alpha$ Metrics Assigned$_i$ + β Metrics Chosen$_i$ + $W'_i\psi + \epsilon_i$
- Subscript *i* represent policy-maker
- Y represents a set of attitudes, beliefs, and choices
- Metrics Chosen denotes an indicator variable that switches on for **choosing** Mastering Metrics vs. Mindsight
- Metrics Assigned denotes an indicator variable that switches on for being randomly **assigned** the book
- W is a vector of individual controls from administrative data
  - written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining public service, age, education, foreign visits and occupational group dummies.
- α provides us the impact of econometrics training
- cluster standard errors at the individual level (level of randomization)

Empirical Methodology

- $Y_i = \theta + \alpha$ Metrics Assigned$_i$ + β Metrics Chosen$_i$ + $W'_i\psi + \epsilon_i$
- Subscript *i* represent policy-maker
- Y represents a set of attitudes, beliefs, and choices
- Metrics Chosen denotes an indicator variable that switches on for **choosing** Mastering Metrics vs. Mindsight
- Metrics Assigned denotes an indicator variable that switches on for being randomly **assigned** the book
- W is a vector of individual controls from administrative data
  - written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining public service, age, education, foreign visits and occupational group dummies.
- α provides us the impact of econometrics training
- cluster standard errors at the individual level (level of randomization)

Empirical Methodology

- $Y_i = \theta + \alpha$ Metrics Assigned$_i$ + β Metrics Chosen$_i$ + $W'_i\psi + \epsilon_i$
- Subscript *i* represent policy-maker
- Y represents a set of attitudes, beliefs, and choices
- Metrics Chosen denotes an indicator variable that switches on for **choosing** Mastering Metrics vs. Mindsight
- Metrics Assigned denotes an indicator variable that switches on for being randomly **assigned** the book
- W is a vector of individual controls from administrative data
  - written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining public service, age, education, foreign visits and occupational group dummies.
- α provides us the impact of econometrics training
- cluster standard errors at the individual level (level of randomization)

Balance Check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Written test scores	Interview test scores	Birth in political capitals	Income	Age	Education	Visited Foreign Country	PAS	PSP	Other groups
Metrics Assigned	0.867	1.221	0.087	-7412.042	0.232	0.082	-0.002	-0.002	-0.054	0.023
	(5.000)	(3.141)	(0.091)	(4629.089)	(0.392)	(0.087)	(0.072)	(0.045)	(0.035)	(0.057)
Metrics Chosen	-3.164	-4.326	-0.027	8366.162*	-0.229	-0.054	0.066	0.051	0.053	0.065
	(4.652)	(2.813)	(0.086)	(4595.222)	(0.391)	(0.083)	(0.069)	(0.042)	(0.037)	(0.053)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190	190	190
R-squared	0.485	0.22	0.05	0.161	0.257	0.223	0.072	0.634	0.452	0.647
Mean of dependent variable	655.585	131.085	0.324	34258.260	26.775	0.516	0.225	0.169	0.099	0.641

Differences are small, almost all p-values > 0.1

High take-up with $> 80\%$ compliance in all responses

(no attrition differences)

Balance Check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Written test scores	Interview test scores	Birth in political capitals	Income	Age	Education	Visited Foreign Country	PAS	PSP	Other groups
Metrics Assigned	0.867	1.221	0.087	-7412.042	0.232	0.082	-0.002	-0.002	-0.054	0.023
	(5.000)	(3.141)	(0.091)	(4629.089)	(0.392)	(0.087)	(0.072)	(0.045)	(0.035)	(0.057)
Metrics Chosen	-3.164	-4.326	-0.027	8366.162*	-0.229	-0.054	0.066	0.051	0.053	0.065
	(4.652)	(2.813)	(0.086)	(4595.222)	(0.391)	(0.083)	(0.069)	(0.042)	(0.037)	(0.053)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190	190	190
R-squared	0.485	0.22	0.05	0.161	0.257	0.223	0.072	0.634	0.452	0.647
Mean of dependent variable	655.585	131.085	0.324	34258.260	26.775	0.516	0.225	0.169	0.099	0.641

Differences are small, almost all p-values > 0.1

High take-up with $> 80\%$ compliance in all responses

(no attrition differences)

Balance Check

	(1) Written test scores	(2) Interview test scores	(3) Birth in political capitals	(4) Income	(5) Age	(6) Education	(7) Visited Foreign Country	(8) PAS	(9) PSP	(10) Other groups
Metrics Assigned	0.867	1.221	0.087	-7412.042	0.232	0.082	-0.002	-0.002	-0.054	0.023
	(5.000)	(3.141)	(0.091)	(4629.089)	(0.392)	(0.087)	(0.072)	(0.045)	(0.035)	(0.057)
Metrics Chosen	-3.164	-4.326	-0.027	8366.162*	-0.229	-0.054	0.066	0.051	0.053	0.065
	(4.652)	(2.813)	(0.086)	(4595.222)	(0.391)	(0.083)	(0.069)	(0.042)	(0.037)	(0.053)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190	190	190
R-squared	0.485	0.22	0.05	0.161	0.257	0.223	0.072	0.634	0.452	0.647
Mean of dependent variable	655.585	131.085	0.324	34258.260	26.775	0.516	0.225	0.169	0.099	0.641

Differences are small, almost all p-values > 0.1

High take-up with $> 80\%$ compliance in all responses

(no attrition differences)

Partial and Full Metrics Training

- Partial Training = Pre Lecture
  - Write short summary of each chapter of the book
  - Write main lessons that you learned to apply each chapter of the book's concepts in your policymaking
- Full Training = Post Lecture
  - Video Lectures by authors: Joshua Angrist and Daniel Siegelman
  - Structured discussion: "what are the key lessons?" "how will you apply these concepts in your career"
  - Presentation of the main lessons of the book
- Assignments and presentations were rated in a competitive manner with cash awards and commemorative shields awarded to top performers

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

Treated policy-makers increase 35% rating on quantitative evidence

Reinforcement video increases effect by 50%

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

Treated policy-makers increase 35% rating on quantitative evidence

Reinforcement video increases effect by 50%

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

Treated policy-makers increase 35% rating on quantitative evidence

Reinforcement video increases effect by 50%

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

No crowd-out on perceived importance of qualitative analysis

15 (25) percentage points more likely to choose RCTs before policy rollout

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

No crowd-out on perceived importance of qualitative analysis

15 (25) percentage points more likely to choose RCTs before policy rollout

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

No crowd-out on perceived importance of qualitative analysis

15 (25) percentage points more likely to choose RCTs before policy rollout

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

In high stakes writing assignment

15 percentage points more likely to write that randomized evaluations are "apple to apple comparisons"

Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

In high stakes writing assignment

15 percentage points more likely to write that randomized evaluations are "apple to apple comparisons"

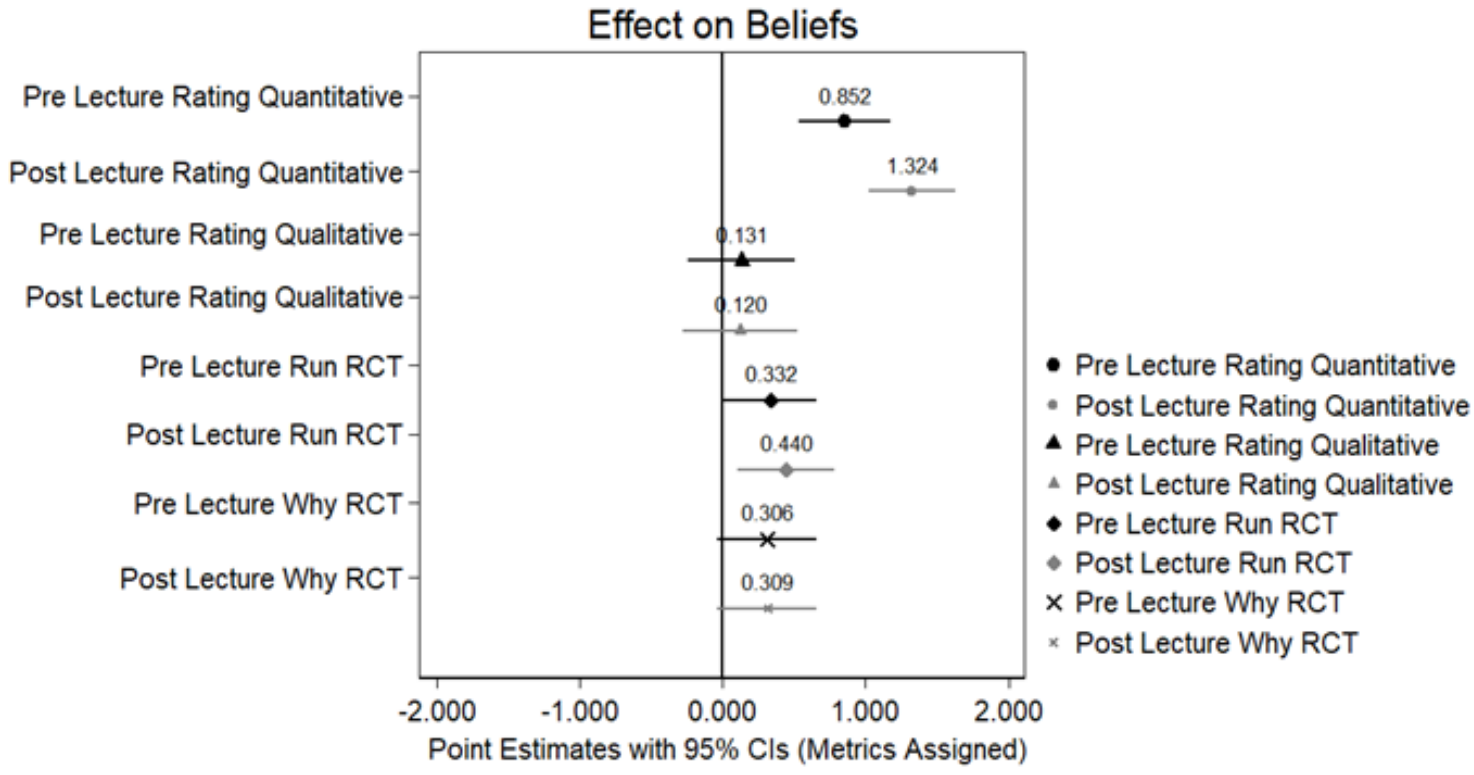
Impacts of Metrics Training on Beliefs

	Pre-Lecture Rating Quantitative	Post-Lecture Rating Quantitative	Pre-Lecture Rating Qualitative	Post Lecture Rating Qualitative	Pre-Lecture Run RCT	Post-Lecture Run RCT	Pre-Lecture Why Run RCT	Post-Lecture Why Run RCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metrics Assigned	0.912***	1.538***	0.136	0.122	0.167**	0 .220**	0.151*	0.153*
	(0.176)	(0.178)	(0.196)	(0.206)	(0.082)	(0.085)	(0.087)	(0.087)
Metrics Choose	-0.042	0.006	0.203	0.068	-0.126	-0.133	-0.126	-0.118
	(0.178)	(0.181)	(0.183)	(0.200)	(0.082)	(0.085)	(0.084)	(0.085)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	190	190	190	190	190	190	190	190
Mean of dep. var. (placebo)	2.745	2.979	2.490	2.596	0.362	0.404	0.396	0.396

In high stakes writing assignment

15 percentage points more likely to write that randomized evaluations are "apple to apple comparisons"

Impacts of Metrics Training on Beliefs - Standardized

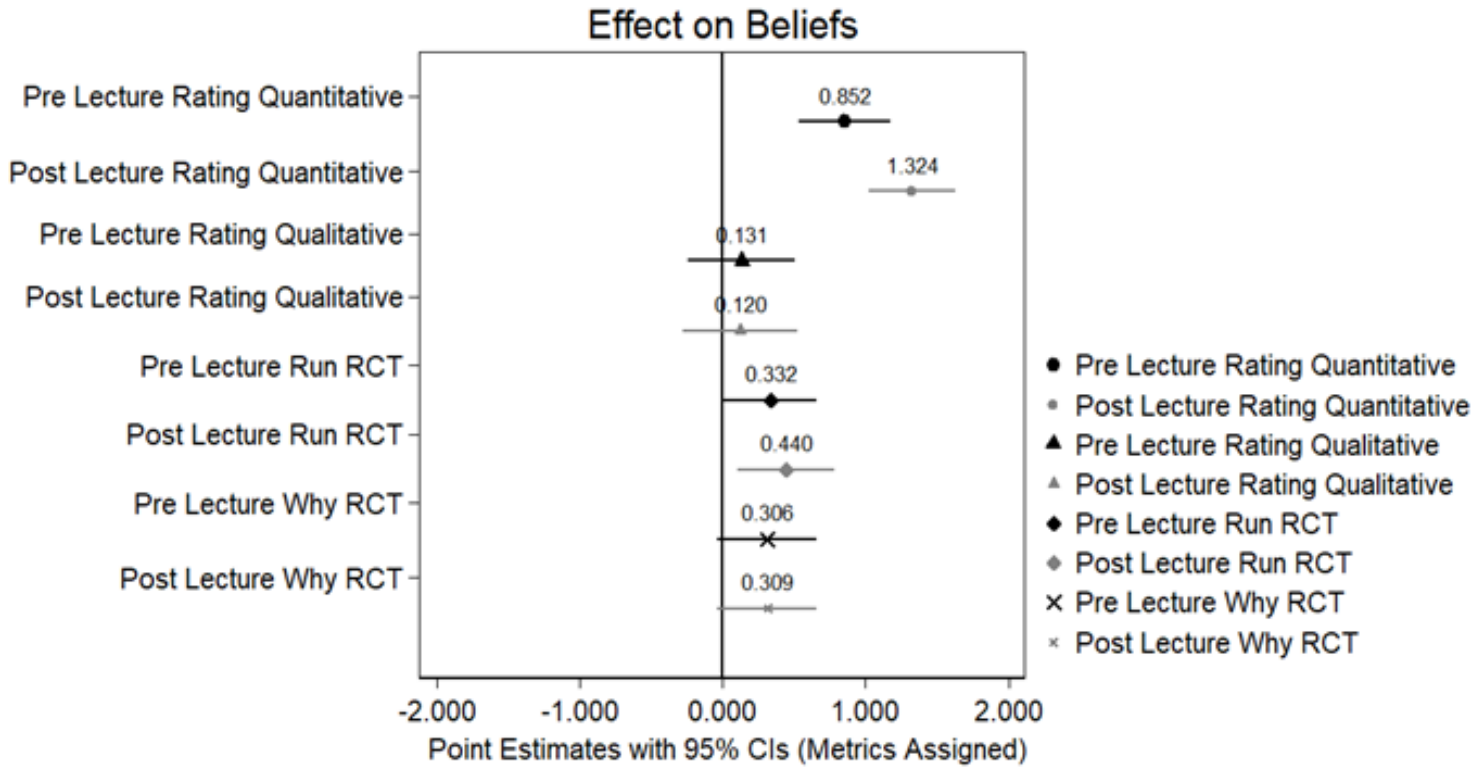


0.85-1.32 σ increase in rating assigned to quantitative evidence

0.33-0.44 σ increase in requesting randomized evaluations

0.30-0.31 σ increase answering that randomized evaluations allow for apple to apple comparisons

Impacts of Metrics Training on Beliefs - Standardized

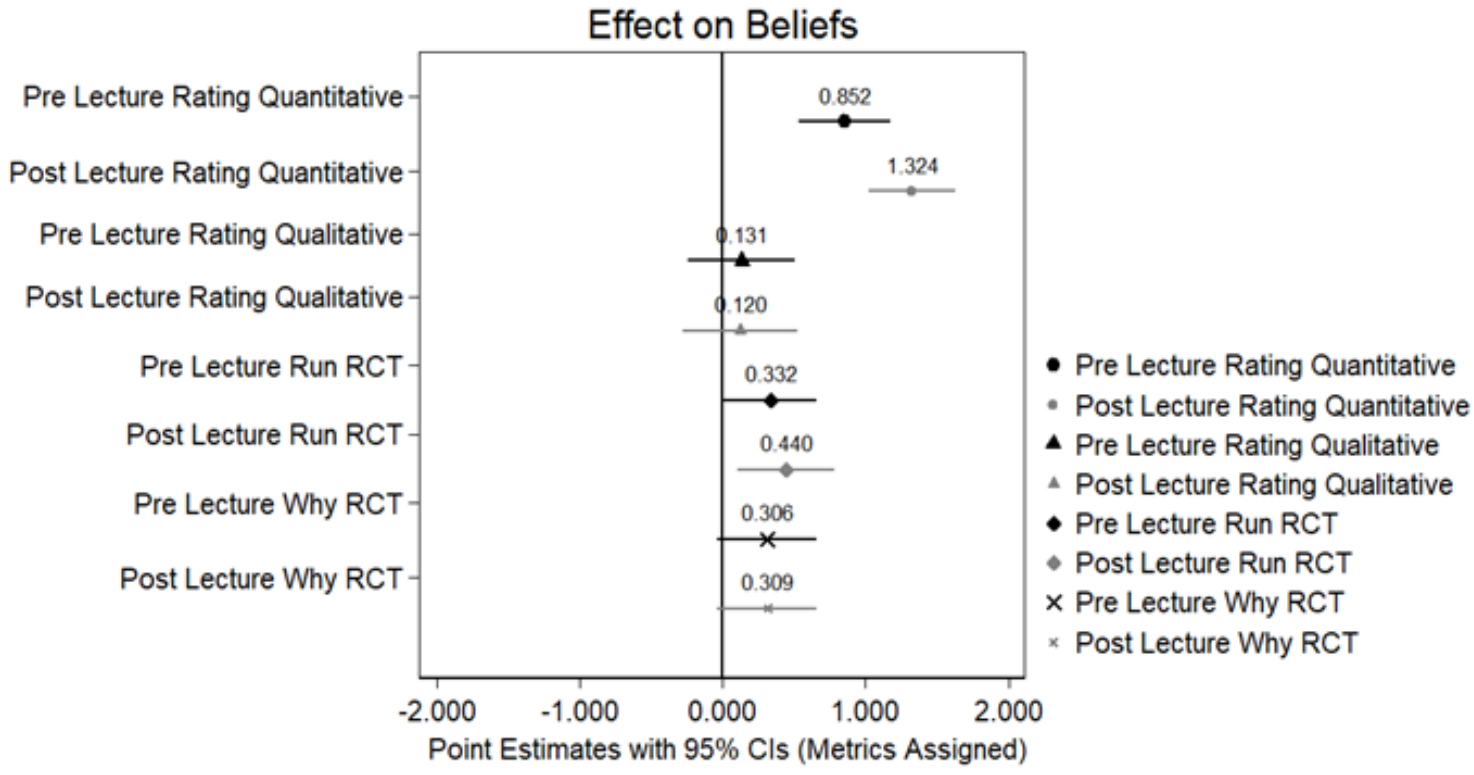


0.85-1.32 σ increase in rating assigned to quantitative evidence

0.33-0.44 σ increase in requesting randomized evaluations

0.30-0.31 σ increase answering that randomized evaluations allow for apple to apple comparisons

Impacts of Metrics Training on Beliefs - Standardized

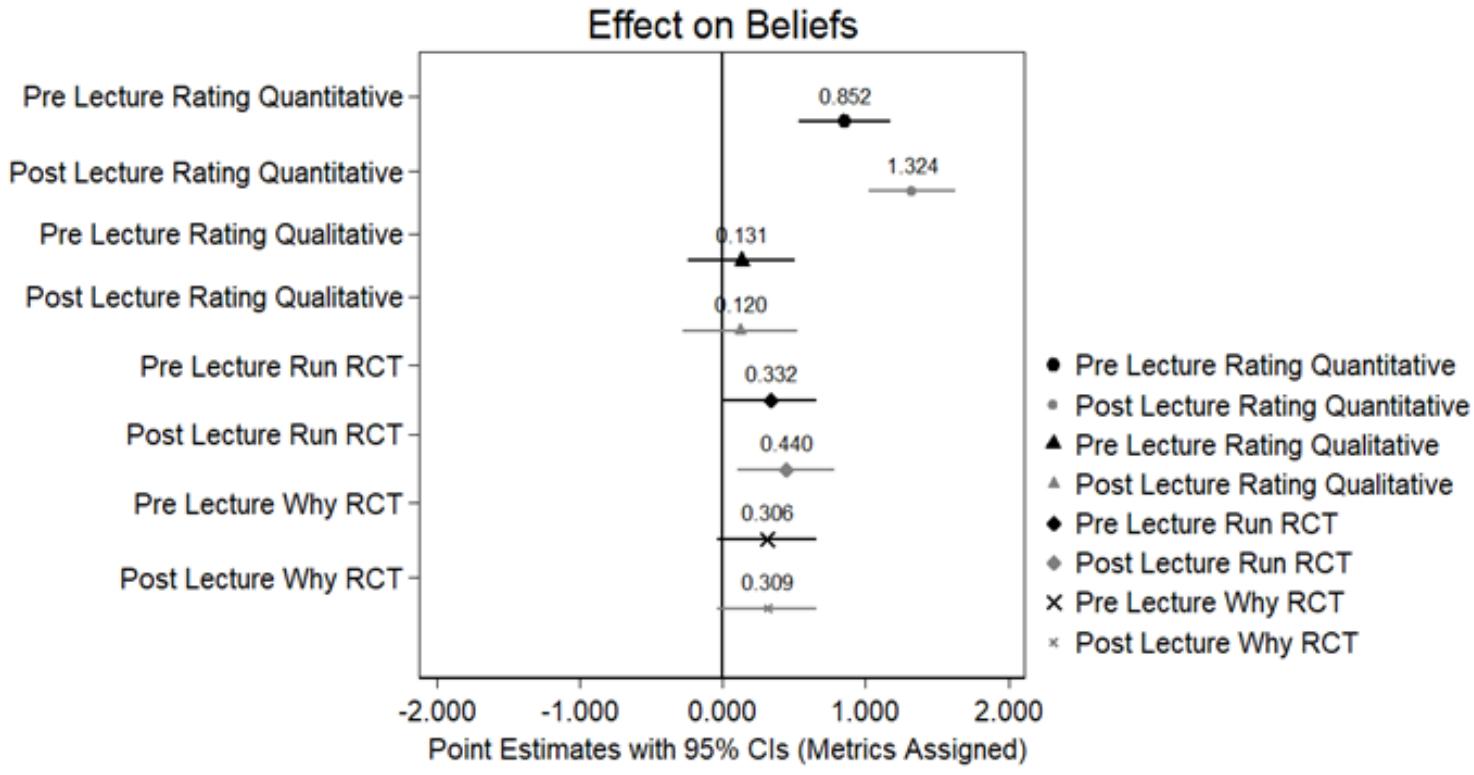


0.85-1.32 σ increase in rating assigned to quantitative evidence

0.33-0.44 σ increase in requesting randomized evaluations

0.30-0.31 σ increase answering that randomized evaluations allow for apple to apple comparisons

Impacts of Metrics Training on Beliefs - Standardized



0.85-1.32 σ increase in rating assigned to quantitative evidence

0.33-0.44 σ increase in requesting randomized evaluations

0.30-0.31 σ increase answering that randomized evaluations allow for apple to apple comparisons

Impact of Metrics Training on Policy Making Assessments

	Assessment Public Goods and Policy		*Assessment Research Methods*		*Assessment Teamwork*	
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	0.574***	0.505***	0.814***	0.799***	-0.004	0.017
	(0.165)	(0.168)	(0.173)	(0.180)	(0.189)	(0.194)
Metrics Choose	0.168	0.132	0.170	0.142	0.075	0.084
	(0.160)	(0.168)	(0.161)	(0.173)	(0.188)	(0.186)
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	190	190	190	190	190	190
R-squared	0.097	0.171	0.186	0.219	0.001	0.123

policymakers' performance in the national research methods and public policy exams (24% of total grade)

0.5σ higher in national public policy exams

0.8σ higher in research methods

no effect on teamwork assessments

Impact of Metrics Training on Policy Making Assessments

	Assessment Public Goods and Policy		*Assessment Research Methods*		*Assessment Teamwork*	
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	0.574***	0.505***	0.814***	0.799***	-0.004	0.017
	(0.165)	(0.168)	(0.173)	(0.180)	(0.189)	(0.194)
Metrics Choose	0.168	0.132	0.170	0.142	0.075	0.084
	(0.160)	(0.168)	(0.161)	(0.173)	(0.188)	(0.186)
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	190	190	190	190	190	190
R-squared	0.097	0.171	0.186	0.219	0.001	0.123

policymakers' performance in the national research methods and public policy exams (24% of total grade)

0.5σ higher in national public policy exams

0.8σ higher in research methods

no effect on teamwork assessments

Impact of Metrics Training on Policy Making Assessments

	Assessment Public Goods and Policy		*Assessment Research Methods*		*Assessment Teamwork*	
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	0.574***	0.505***	0.814***	0.799***	-0.004	0.017
	(0.165)	(0.168)	(0.173)	(0.180)	(0.189)	(0.194)
Metrics Choose	0.168	0.132	0.170	0.142	0.075	0.084
	(0.160)	(0.168)	(0.161)	(0.173)	(0.188)	(0.186)
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	190	190	190	190	190	190
R-squared	0.097	0.171	0.186	0.219	0.001	0.123

policymakers' performance in the national research methods and public policy exams (24% of total grade)

0.5σ higher in national public policy exams

0.8σ higher in research methods

no effect on teamwork assessments

Impact of Metrics Training on Policy Making Assessments

	Assessment Public Goods and Policy		*Assessment Research Methods*		*Assessment Teamwork*	
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	0.574***	0.505***	0.814***	0.799***	-0.004	0.017
	(0.165)	(0.168)	(0.173)	(0.180)	(0.189)	(0.194)
Metrics Choose	0.168	0.132	0.170	0.142	0.075	0.084
	(0.160)	(0.168)	(0.161)	(0.173)	(0.188)	(0.186)
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	190	190	190	190	190	190
R-squared	0.097	0.171	0.186	0.219	0.001	0.123

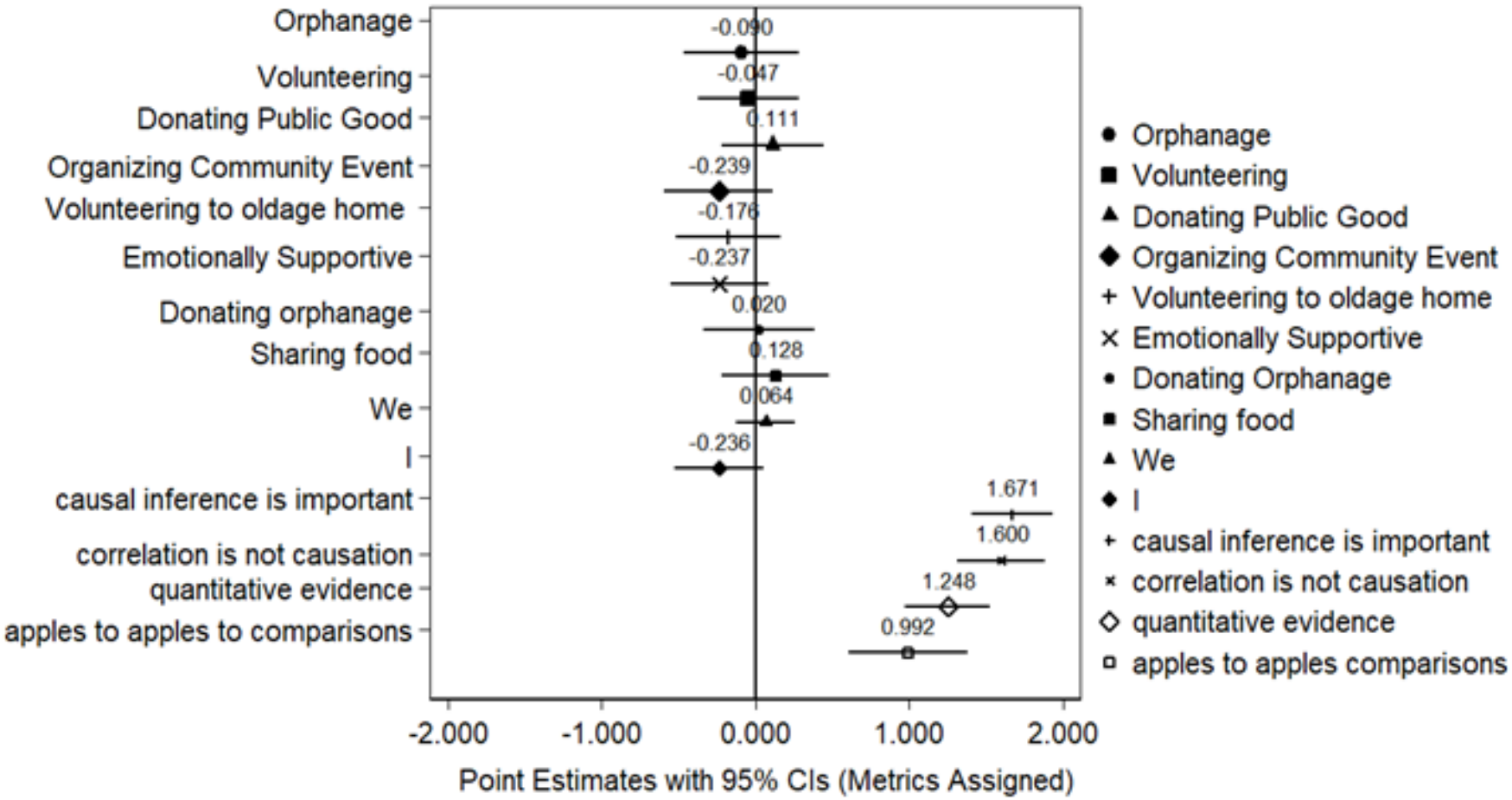
policymakers' performance in the national research methods and public policy exams (24% of total grade)

0.5σ higher in national public policy exams

0.8σ higher in research methods

no effect on teamwork assessments

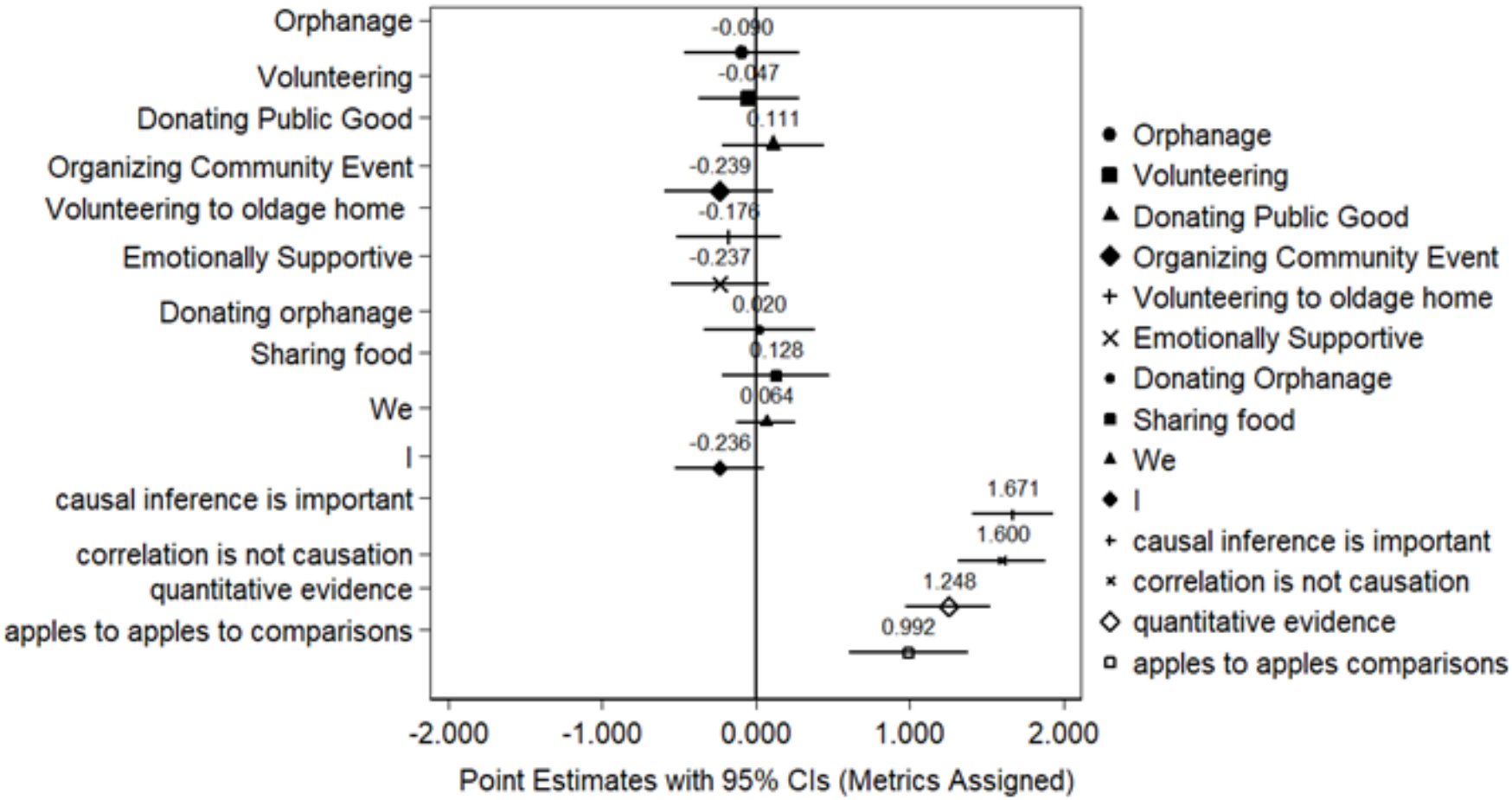
Impacts of Metrics Training on Prosociality and Causal Language



no crowdout of prosocial behavior

measured as visits to orphanages and volunteering in impoverished schools & use of prosocial phrases

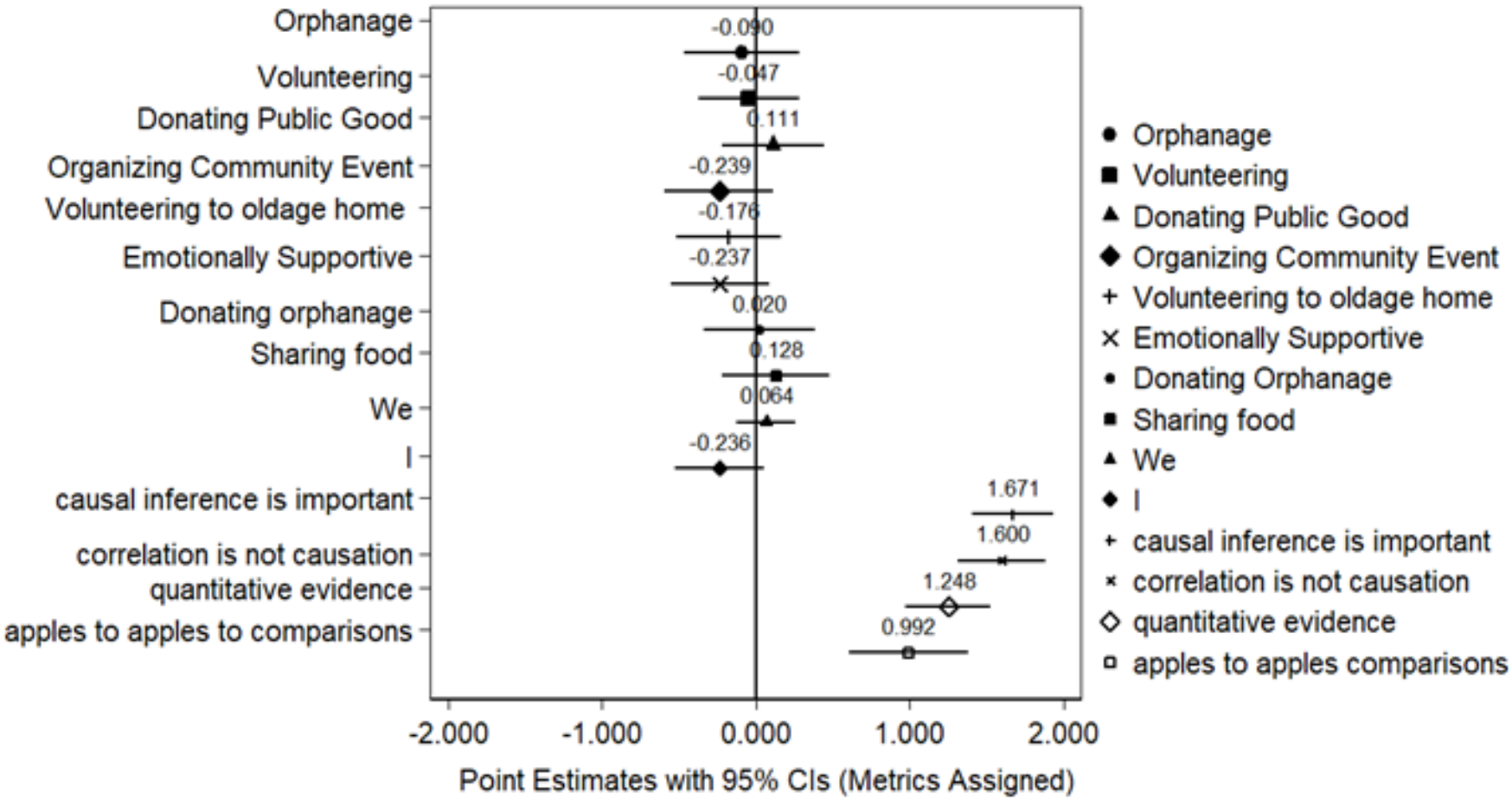
Impacts of Metrics Training on Prosociality and Causal Language



no crowdout of prosocial behavior

measured as visits to orphanages and volunteering in impoverished schools & use of prosocial phrases

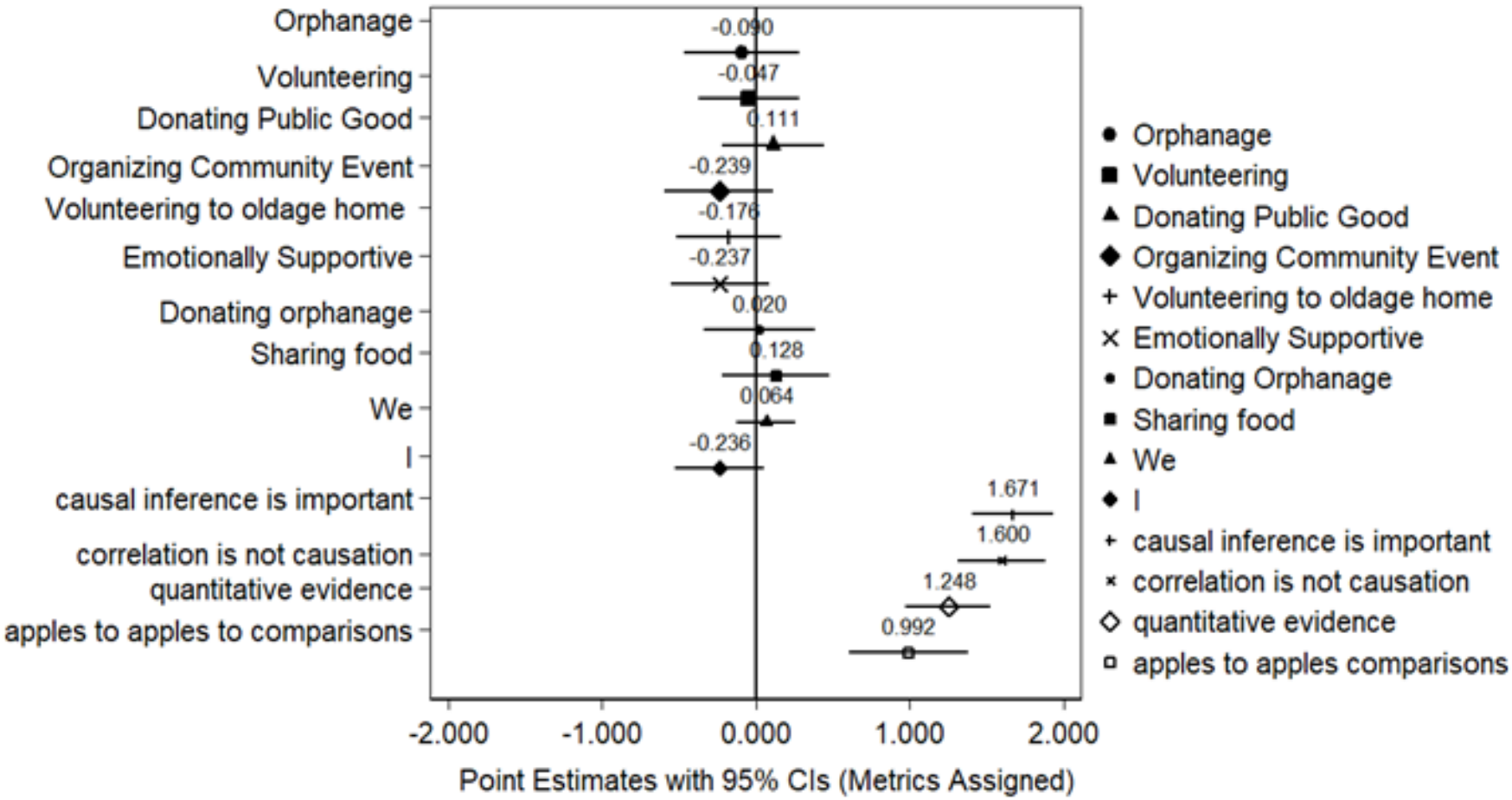
Impacts of Metrics Training on Prosociality and Causal Language



no crowdout of prosocial behavior

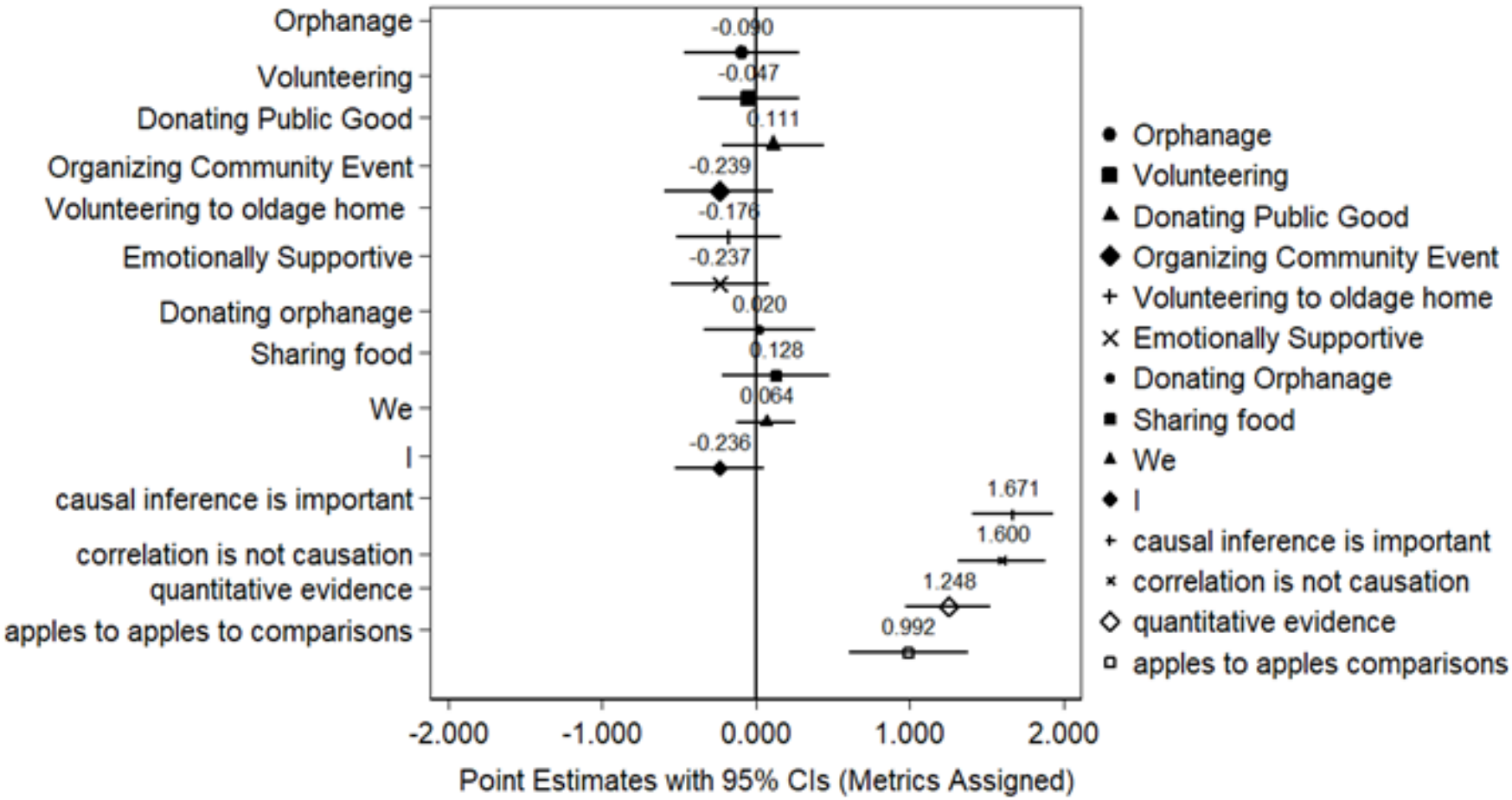
measured as visits to orphanages and volunteering in impoverished schools & use of prosocial phrases

Impacts of Metrics Training on Prosociality and Causal Language



significant increase in beliefs about quantitative evidence - 38 percentage points more likely to use phrases associated with "causal inference is important"

Impacts of Metrics Training on Prosociality and Causal Language



significant increase in beliefs about quantitative evidence - 38 percentage points more likely to use phrases associated with "causal inference is important"

Signpost

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

2) improved test scores on public goods and research methods

3) willingness to pay for quantitative data, only for RCTs ($\downarrow$ correlations)

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Signpost

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

2) improved test scores on public goods and research methods

3) willingness to pay for quantitative data, only for RCTs (↓ correlations)

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy > 10 percent (Hjort et al. 2021)

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Signpost

Training about causal thinking increases

1) perceived value of causal inference, quantitative data and RCTs

2) improved test scores on public goods and research methods

3) willingness to pay for quantitative data, only for RCTs (↓ correlations)

SHIFT IN DEMAND FOR INFORMATION

4) responsiveness to evidence from a RCT

40 percent more likely to adopt policy > 10 percent (Hjort et al. 2021)

OPENNESS TO INFORMATION

5) only for those whose priors were lower than the evidence from the RCT

LACK OF EVIDENCE OF CONFIRMATION BIAS/MOTIVATED REASONING

Signpost

Does training about causal thinking affect

1) perceived value of causal inference, quantitative data and RCTs

5) improved test scores on public goods and research methods

DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

CONFIRMATION BIAS/MOTIVATED REASONING

Signpost

Does training about causal thinking affect

1) perceived value of causal inference, quantitative data and RCTs

5) improved test scores on public goods and research methods

DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

CONFIRMATION BIAS/MOTIVATED REASONING

Signpost

Does training about causal thinking affect

1) perceived value of causal inference, quantitative data and RCTs

5) improved test scores on public goods and research methods

DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

CONFIRMATION BIAS/MOTIVATED REASONING

Signpost

Does training about causal thinking affect

1) perceived value of causal inference, quantitative data and RCTs

5) improved test scores on public goods and research methods

DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

CONFIRMATION BIAS/MOTIVATED REASONING

Demand for Information

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Metrics Training on Willingness to Pay

Panel A: Private Spending						
	Before Signal Amount Randomized Trial	*Before Signal Amount Correlational Data*	*Before Signal Amount Expert Bureaucrat*	*After Signal Amount Randomized Trial*	*After Signal Amount Correlational Data*	*After Signal Amount Expert Bureaucrat*
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	1369.290*	-1048.355*	279.115	2012.892***	-1018.698***	-1866.809
	(702.448)	(555.314)	(271.880)	(585.769)	(335.110)	(1365.653)
Metrics Choose	728.1.59	948.031*	-99.643	1646.977***	880.527***	-285.501
	(725.718)	(538.251)	(232.258)	(602.013)	(332.546)	(1396.293)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	180	180	180	180	180	180
R-squared	0.107	0.104	0.111	0.217	0.204	0.104
Mean of dep. var. (placebo)	2503.594	2267.188	430.977	1539.453	2214.07	4490.008
Panel B: Public Spending						
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	739148.720***	-55164.030***	15058.640	1301597.000*	-36507.414***	4685.422
	(224074.900)	(13629.180)	(47215.290)	(721138.800)	(12872.410)	(32512.540)
Metrics Choose	119080.700	12554.130	22428.540	-356638.500	-8874.433	15751.800
	(201589.900)	(13895.560)	(46853.630)	(650493.700)	(13992.260)	(34727.420)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	180	180	180	180	180	180
R-squared	0.234	0.170	0.069	0.092	0.150	0.611
Mean of dep. var. (placebo)	224604.700	109746.900	168464.800	928546.100	94136.720	115937.500

WTP $10 (1% of monthly salary)

WTP $5K (of public funds)

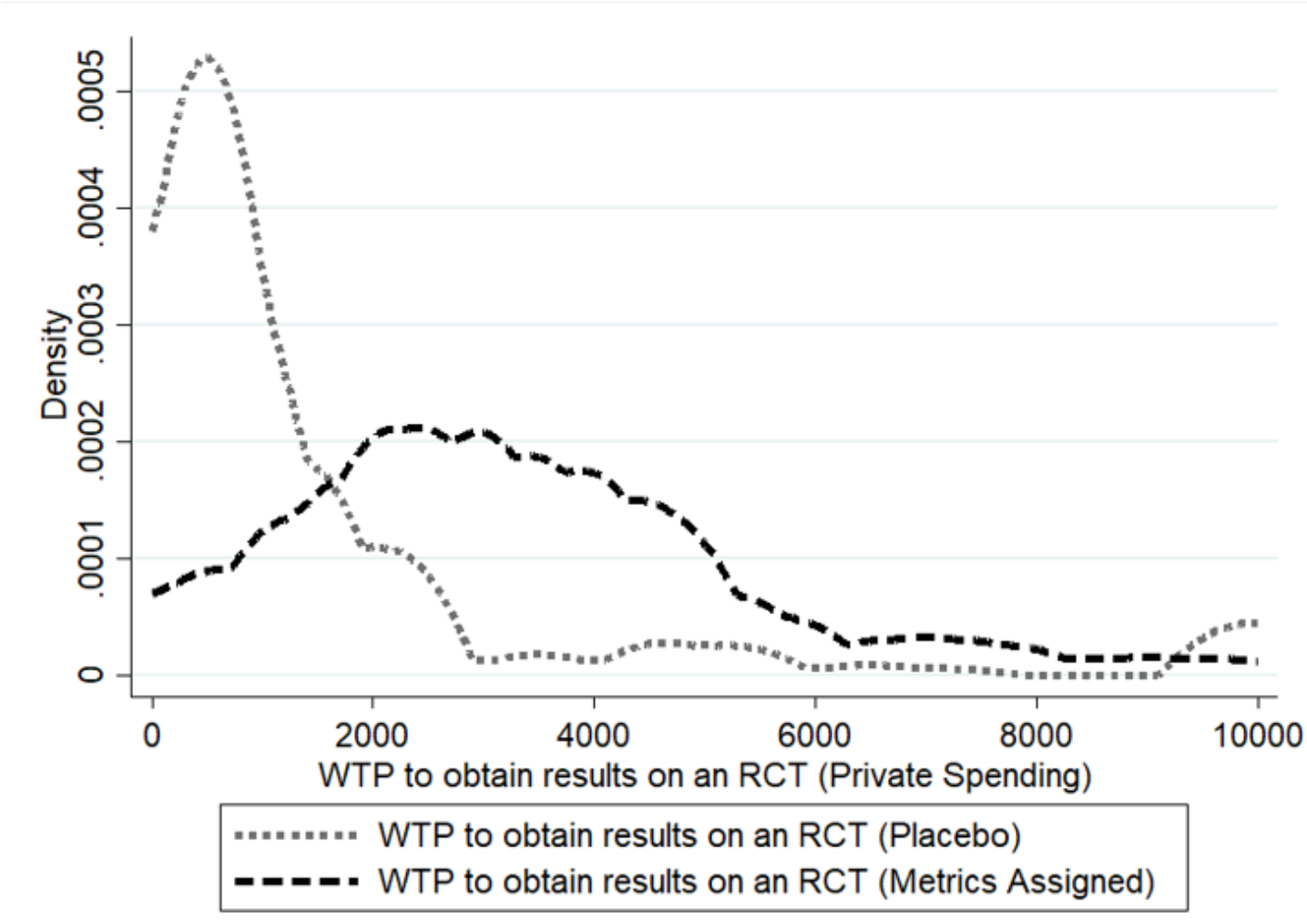
Metrics Training on Willingness to Pay

Panel A: Private Spending						
	Before Signal Amount Randomized Trial	*Before Signal Amount Correlational Data*	*Before Signal Amount Expert Bureaucrat*	*After Signal Amount Randomized Trial*	*After Signal Amount Correlational Data*	*After Signal Amount Expert Bureaucrat*
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	1369.290*	-1048.355*	279.115	2012.892***	-1018.698***	-1866.809
	(702.448)	(555.314)	(271.880)	(585.769)	(335.110)	(1365.653)
Metrics Choose	728.1.59	948.031*	-99.643	1646.977***	880.527***	-285.501
	(725.718)	(538.251)	(232.258)	(602.013)	(332.546)	(1396.293)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	180	180	180	180	180	180
R-squared	0.107	0.104	0.111	0.217	0.204	0.104
Mean of dep. var. (placebo)	2503.594	2267.188	430.977	1539.453	2214.07	4490.008
Panel B: Public Spending						
	(1)	(2)	(3)	(4)	(5)	(6)
Metrics Assigned	739148.720***	-55164.030***	15058.640	1301597.000*	-36507.414***	4685.422
	(224074.900)	(13629.180)	(47215.290)	(721138.800)	(12872.410)	(32512.540)
Metrics Choose	119080.700	12554.130	22428.540	-356638.500	-8874.433	15751.800
	(201589.900)	(13895.560)	(46853.630)	(650493.700)	(13992.260)	(34727.420)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	180	180	180	180	180	180
R-squared	0.234	0.170	0.069	0.092	0.150	0.611
Mean of dep. var. (placebo)	224604.700	109746.900	168464.800	928546.100	94136.720	115937.500

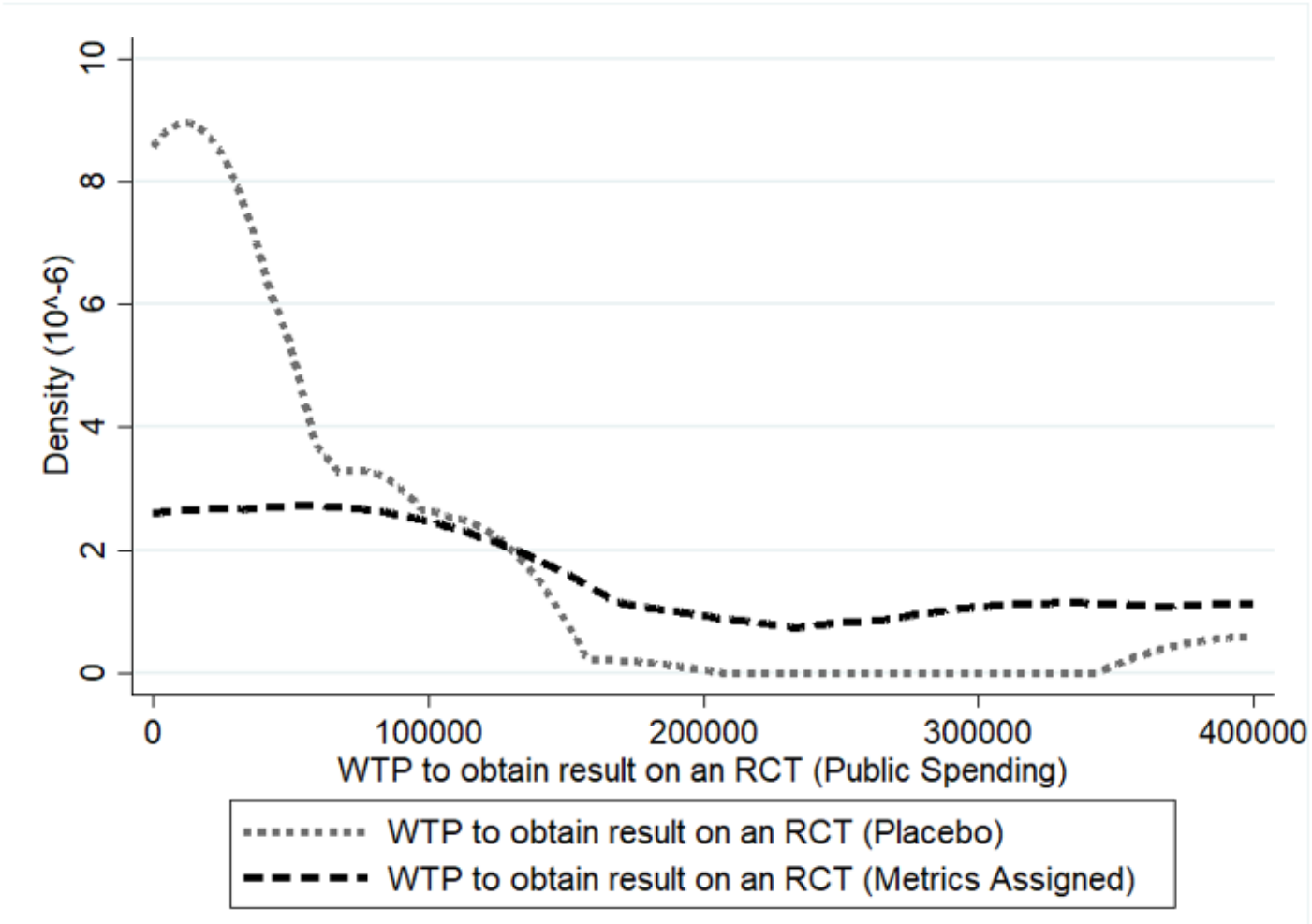
WTP \$10 (1% of monthly salary)

WTP \$5K (of public funds)

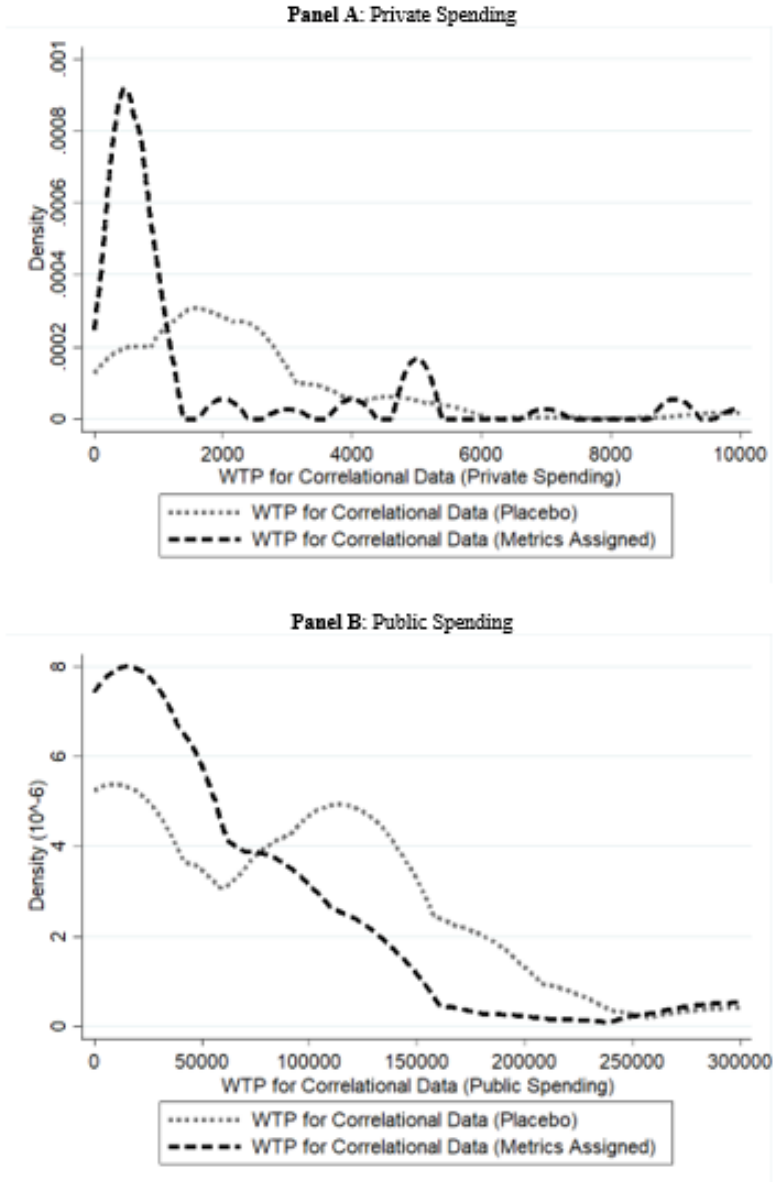
Metrics Training on WTP - RCT Study - Private



Metrics Training on WTP - RCT Study - Public

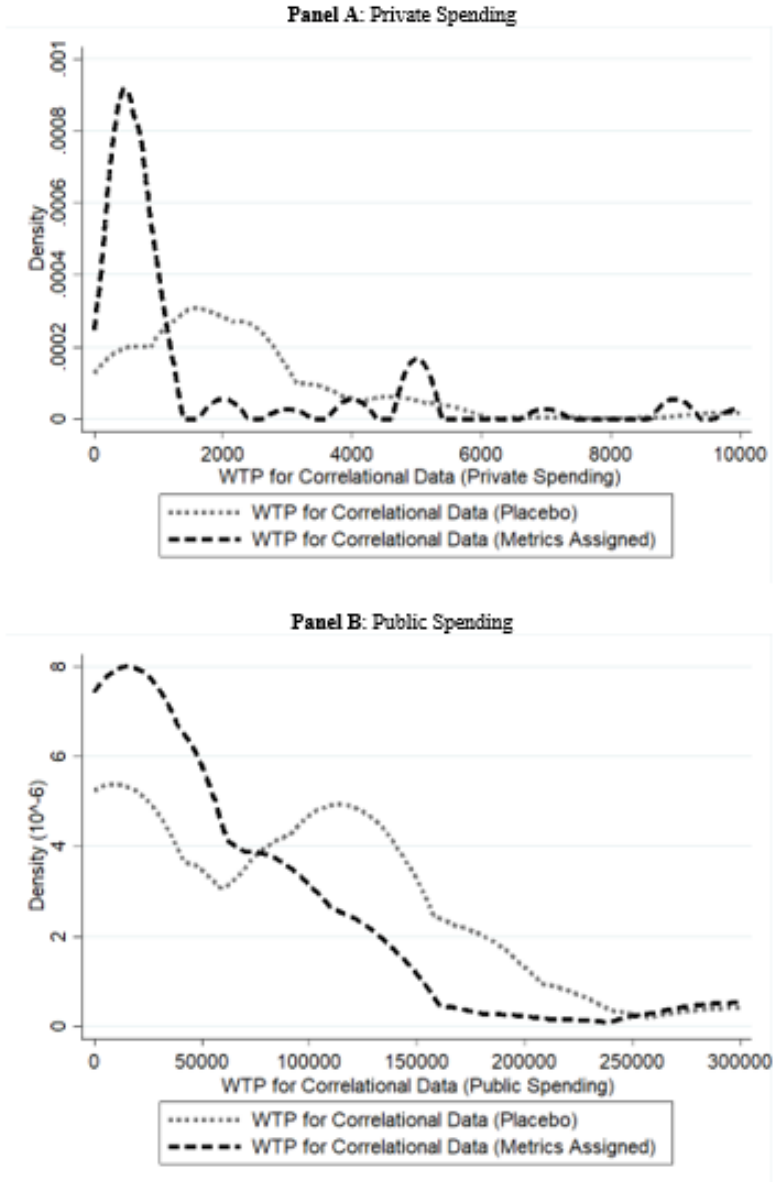


Metrics Training on WTP - Correlational Data



Lower demand for correlational data

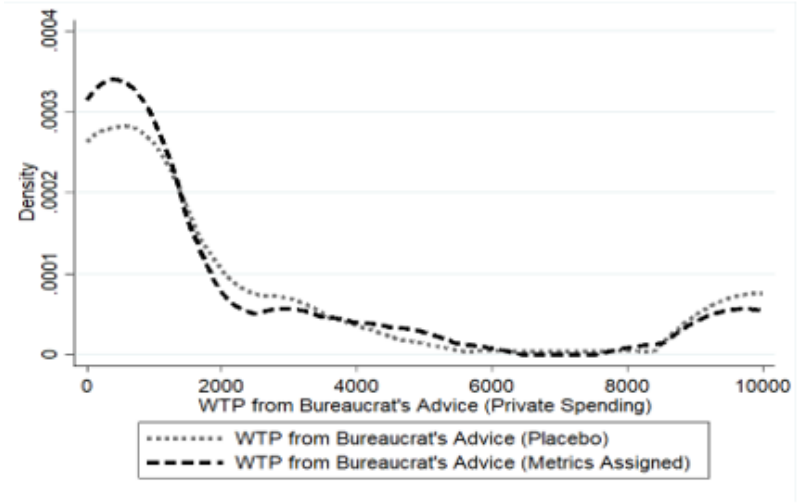
Metrics Training on WTP - Correlational Data



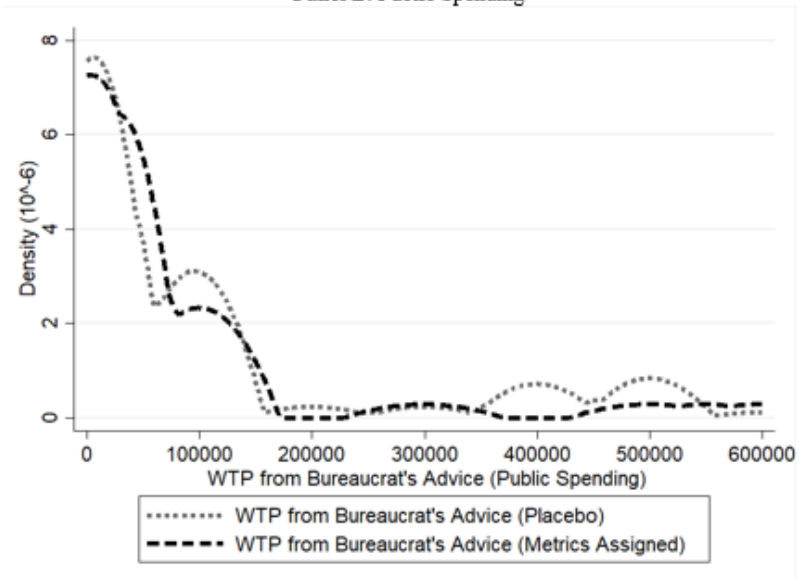
Lower demand for correlational data

Metrics Training on WTP - Advice Data

Panel A: Private Spending

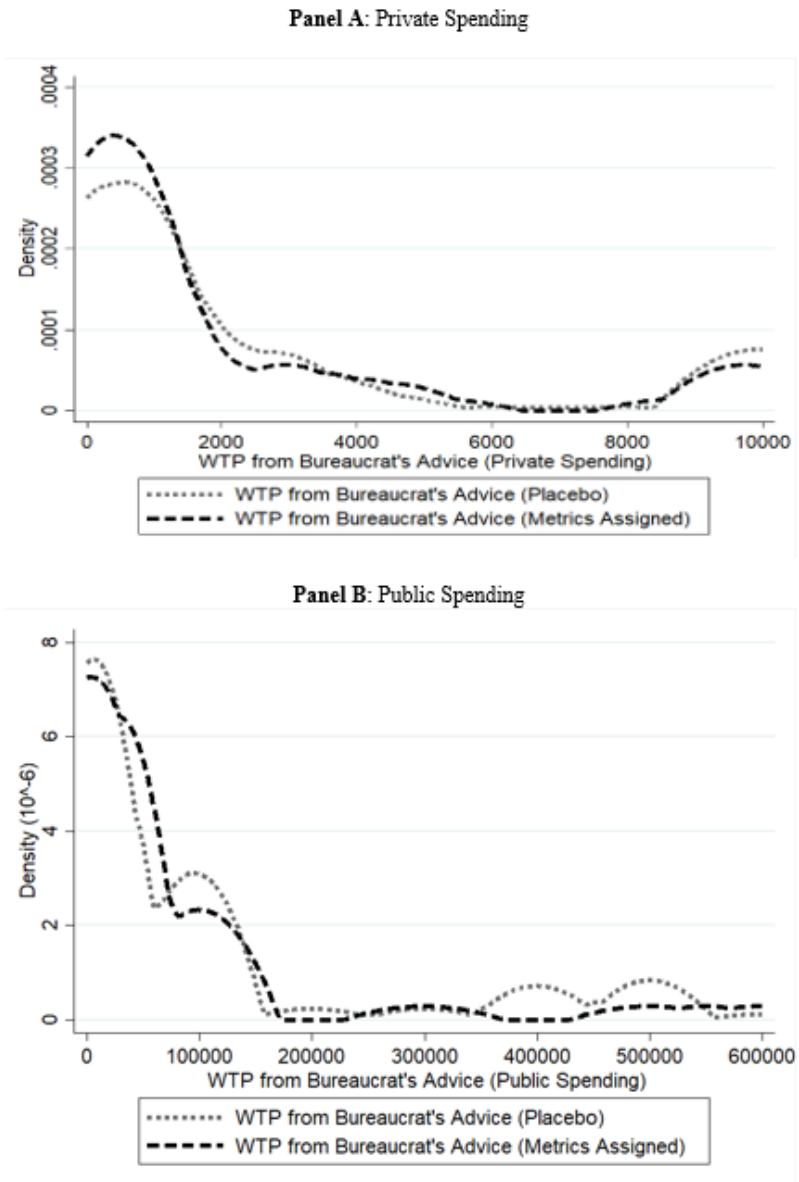


Panel B: Public Spending



No difference in demand for bureaucrat advice

Metrics Training on WTP - Advice Data



No difference in demand for bureaucrat advice

Responsiveness to Information

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Responsiveness to Information

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Responsiveness to Information

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Information - A Signal about Policy

Recent randomized evaluation finds deworming impacts on economic outcomes up to 20 years later. Individuals who received deworming experience up to 3 additional years of schooling, 14% increases in consumption expenditure, **13.01%** increases in hourly earnings, 9% in non-agricultural work hours (Source PNAS, 2020).

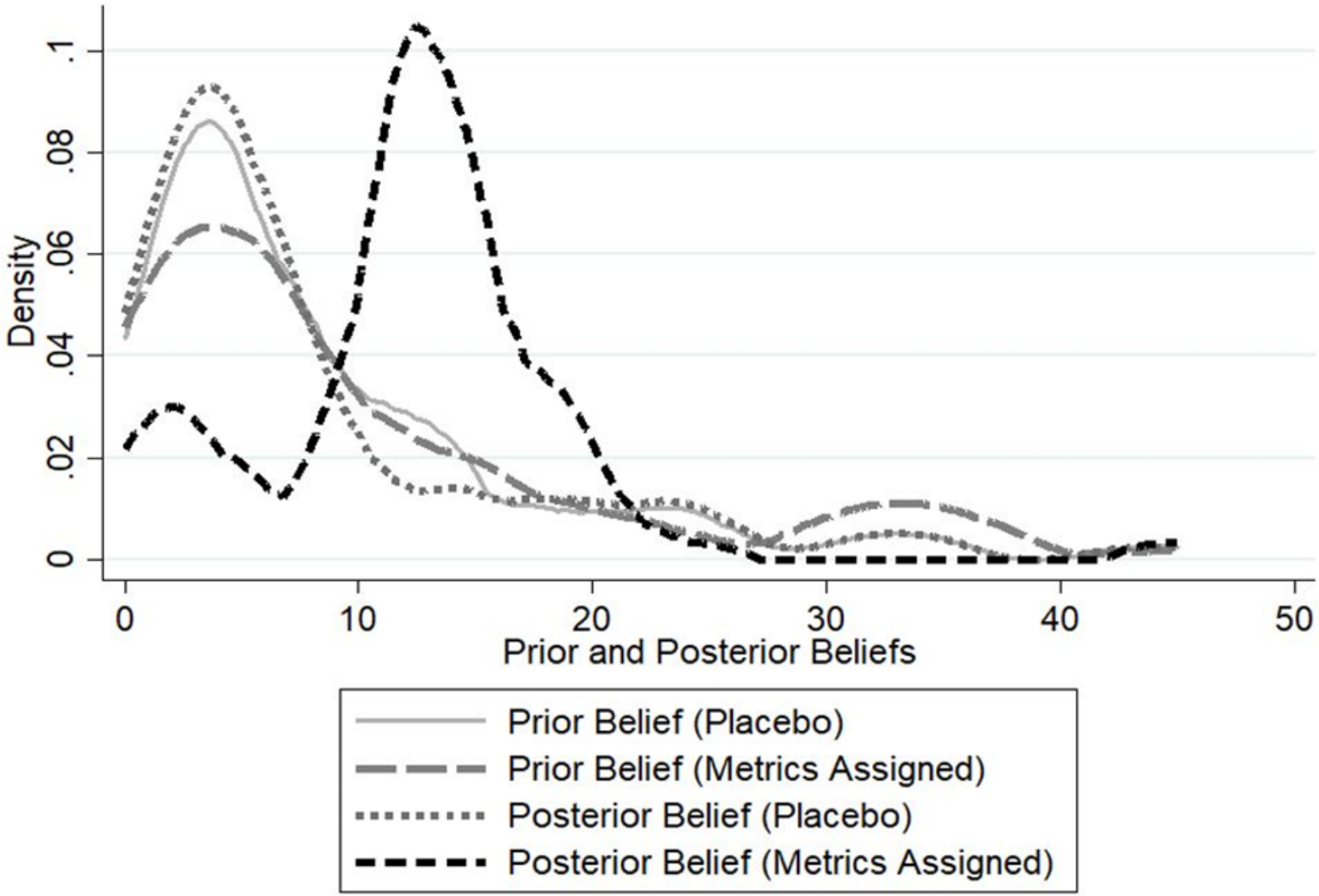
policymakers underestimate the long-run impacts of deworming

Information - A Signal about Policy

Recent randomized evaluation finds deworming impacts on economic outcomes up to 20 years later. Individuals who received deworming experience up to 3 additional years of schooling, 14% increases in consumption expenditure, **13.01%** increases in hourly earnings, 9% in non-agricultural work hours (Source PNAS, 2020).

policymakers underestimate the long-run impacts of deworming

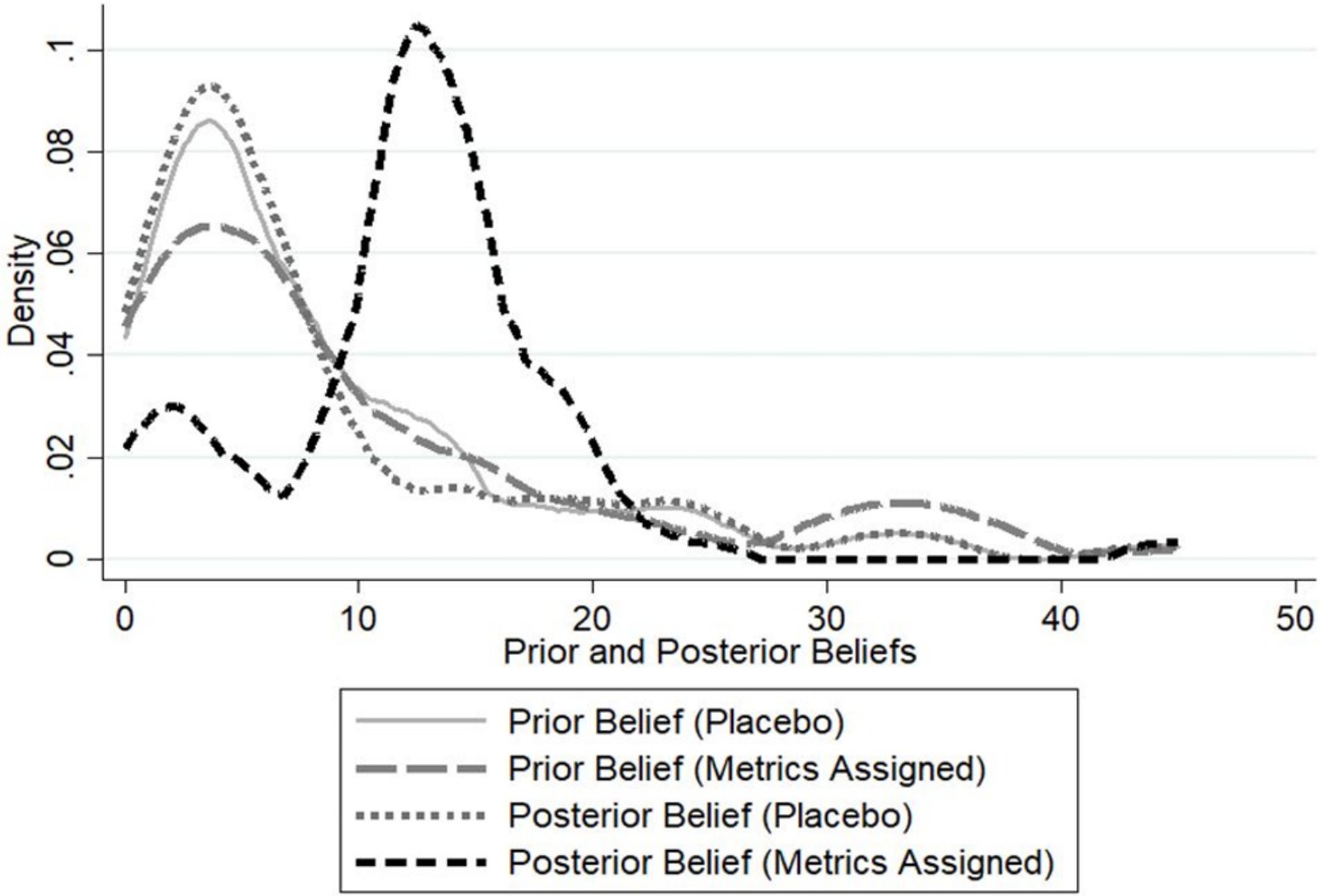
Distribution of Prior and Posterior Beliefs



both groups' priors: boost income by about 5% in the long run

vast majority of treated policymakers update priors to 13%

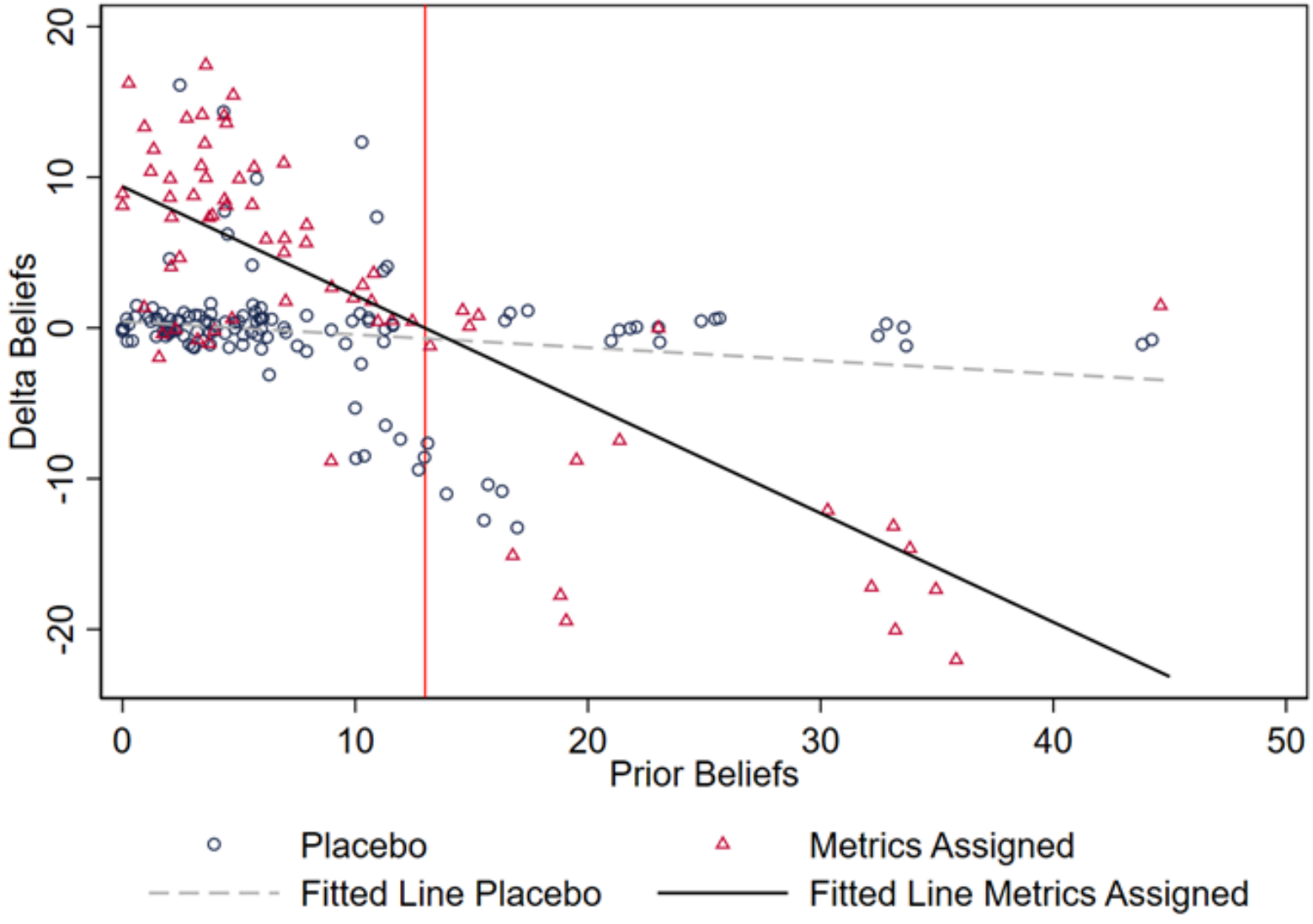
Distribution of Prior and Posterior Beliefs



both groups' priors: boost income by about 5% in the long run

vast majority of treated policymakers update priors to 13%

Initial and Shifts in Beliefs



non-treated policymakers do not update

Confirmation Bias/Motivated Reasoning

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ⋆ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ⋆ (1) expert opinion about the project choice from a senior bureaucrat
    - ⋆ (2) RCT assessing the impact of deworming on hourly earnings
    - ⋆ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Confirmation Bias/Motivated Reasoning

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Confirmation Bias/Motivated Reasoning

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

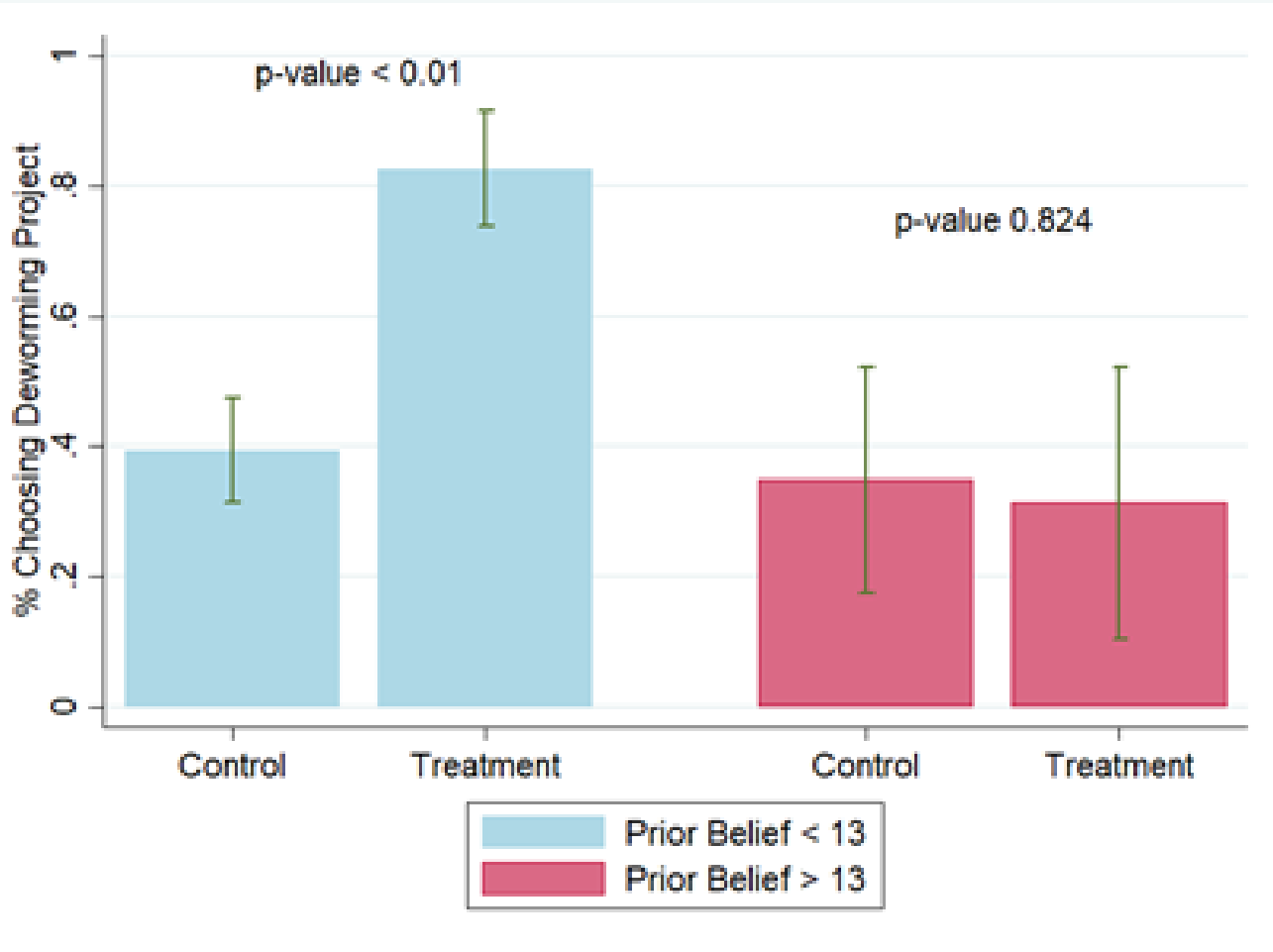
Confirmation Bias/Motivated Reasoning

- May 16, 2021
  - Prior beliefs on the impact of deworming on earnings of children 20 years after the deworming
  - Project Choice
    - ★ deworming relative to computer labs project
  - Willingness to Pay from both private and public funds for
    - ★ (1) expert opinion about the project choice from a senior bureaucrat
    - ★ (2) RCT assessing the impact of deworming on hourly earnings
    - ★ (3) Correlational data showing a relationship between deworming and hourly earnings.
  - Signal (Kremer et al. 2021)
  - Posterior beliefs
  - Updated Project Choice

Project Choice After Signal

What project would you choose 1 or 2? One being deworming and two being computer lab.

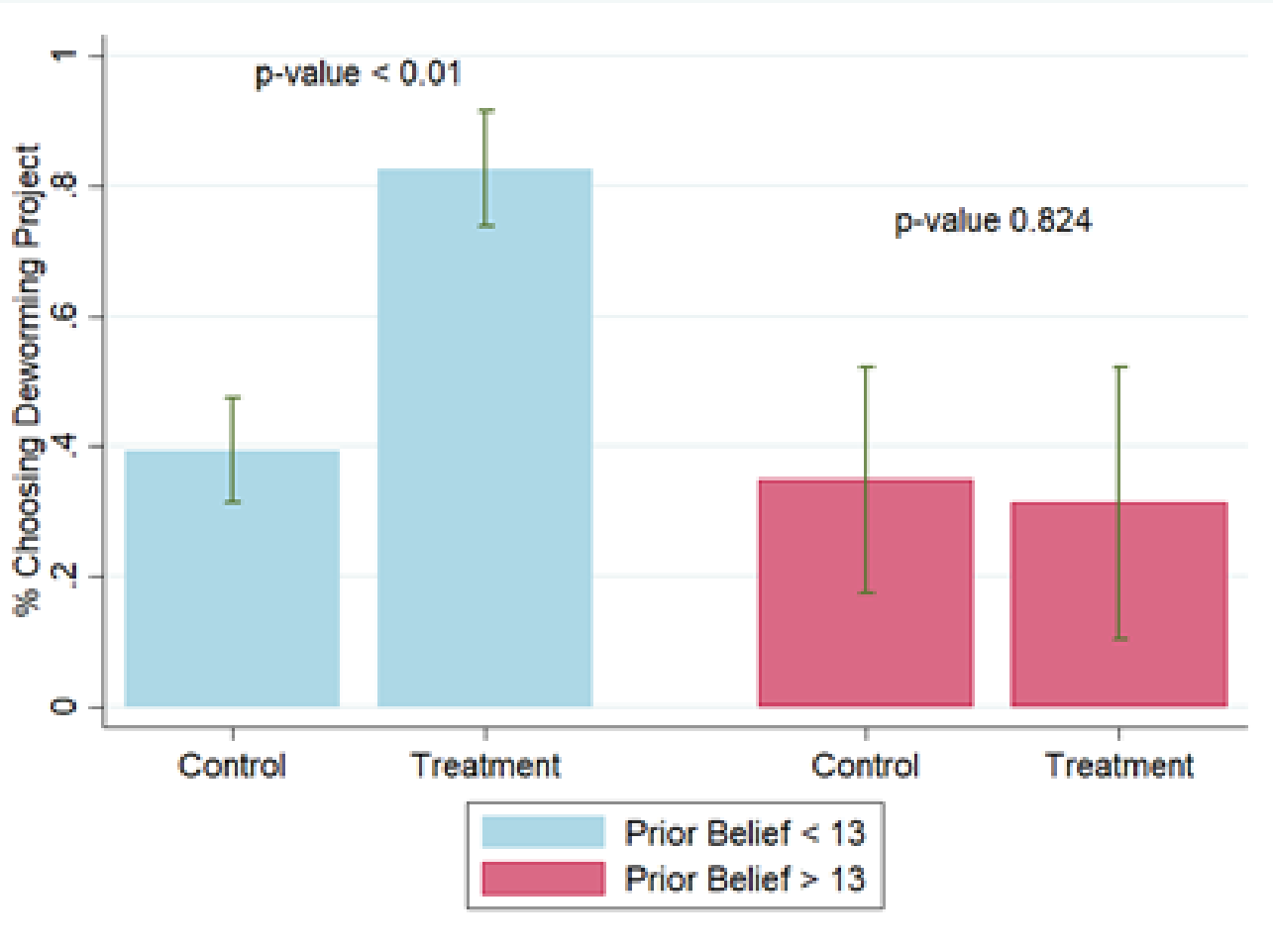
Effect of Metrics Training on Deworming Project Choice by Prior Beliefs



40 percentage points more likely to pick deworming

- Only if prior beliefs are less than evidence from RCT

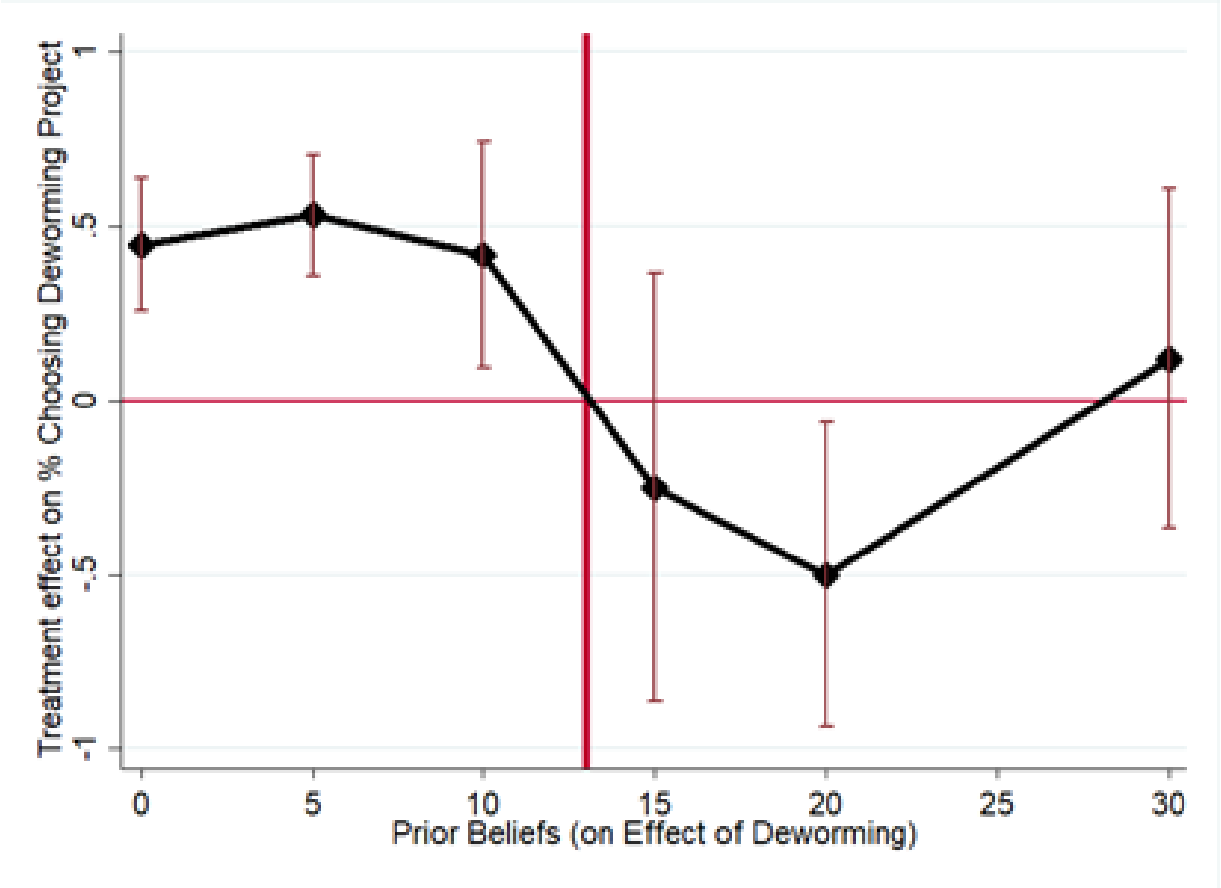
Effect of Metrics Training on Deworming Project Choice by Prior Beliefs



40 percentage points more likely to pick deworming

- Only if prior beliefs are less than evidence from RCT

Effect of Metrics Training on Deworming Project Choice by Prior Beliefs

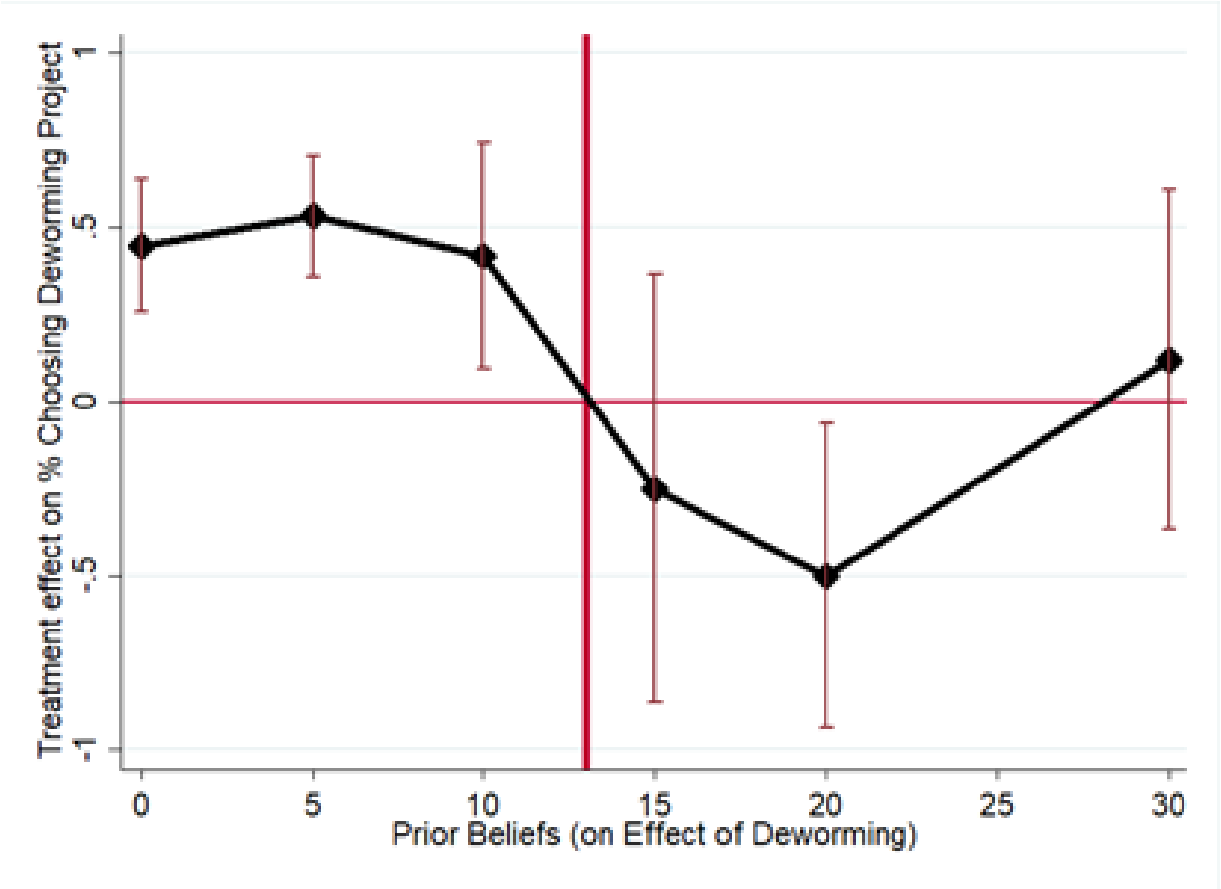


DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

UPDATING POLICY CHOICES

Effect of Metrics Training on Deworming Project Choice by Prior Beliefs

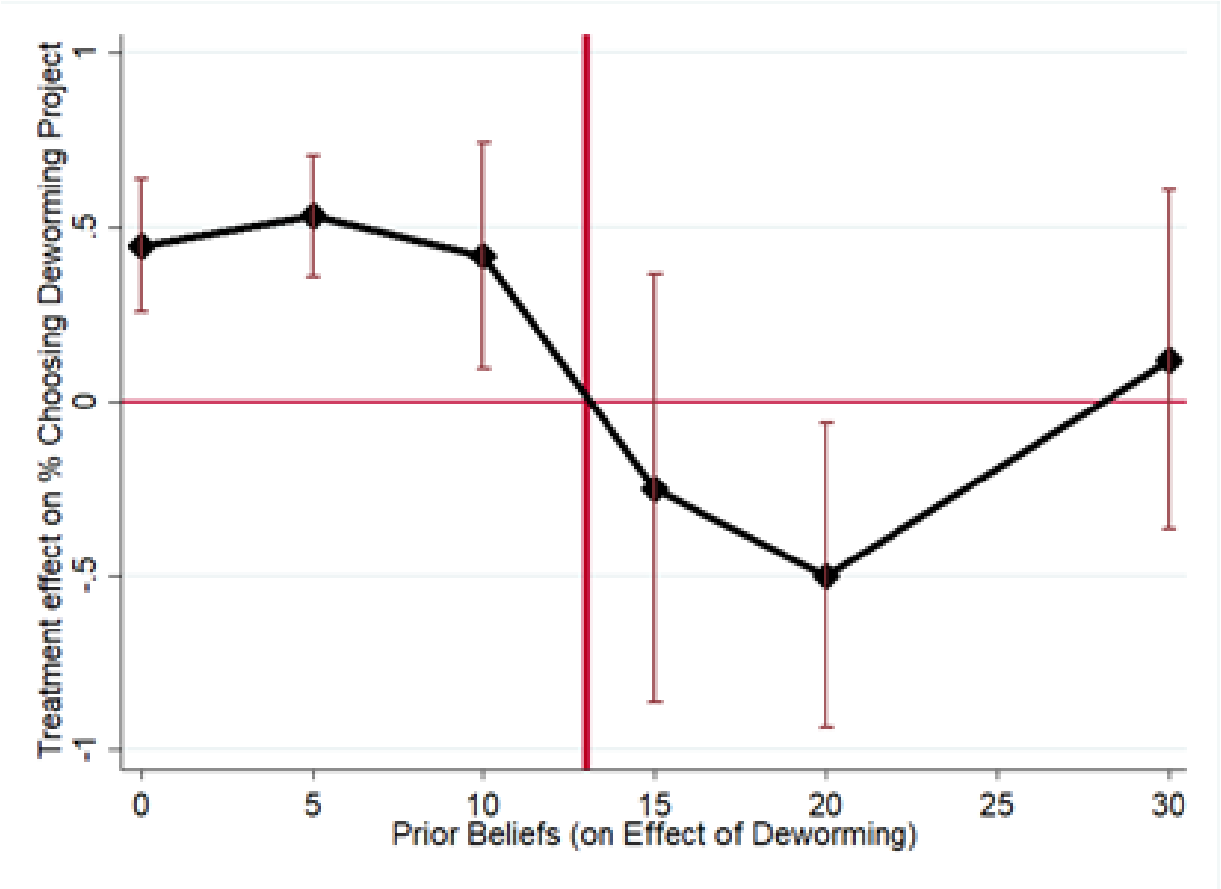


DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

UPDATING POLICY CHOICES

Effect of Metrics Training on Deworming Project Choice by Prior Beliefs

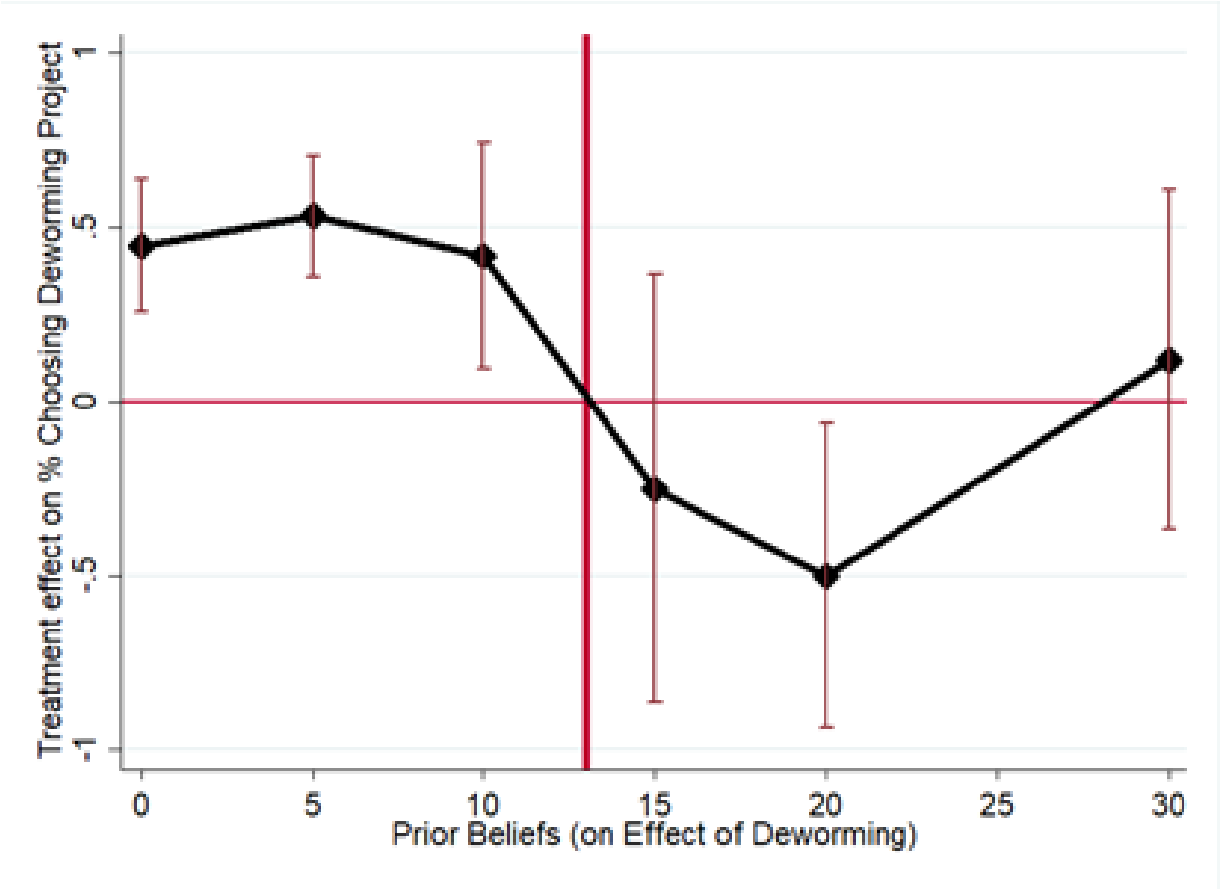


DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

UPDATING POLICY CHOICES

Effect of Metrics Training on Deworming Project Choice by Prior Beliefs



DEMAND FOR INFORMATION

RESPONSIVENESS TO INFORMATION

UPDATING POLICY CHOICES

Counting Defiers

- March 10, 2021 (zoom)
  - Belief Elicitation I
  - Reinforcement lecture of 30 minutes
    - ★ Links in oTree to watch videos
    - ★ Students could pick which video (behavioral measure of defier)
  - Live discussion of 30 minutes
    - ★ a) what do you think is the main point of the lecture?
    - ★ b) how can you apply the concepts learned in this lecture to future job?
  - Belief Elicitation II

Effect of Treatment on WTP by Defiers

	Private Spending		***Public Spending***	
	Amount Randomized Trial	*Amount Randomized Trial*	*Amount Randomized Trial*	*Amount Randomized Trial*
	(1)	(2)	(3)	(4)
Defiers X Metrics Assigned	-550.306	-888.410	-1987315.100***	-2531651.000*
	(1177.783)	(1301.347)	(667571.300)	(1367816.000)
Metrics Assigned	2342.989***	2179.581***	1986365.500***	1775415.200***
	(711.613)	(648.553)	(667569.300)	(635063.500)
Defiers X Metrics Chosen	-1377.298	-428.640	495448.220	3326010.500
	(1196.483)	(1207.932)	(606330.500)	(2437943.000)
Defiers	-160.589	-463.802	-1172597.800*	-3101800.100
	(629.255)	(857.643)	(610215.800)	(2201009.000)
Metrics Chosen	1511.444**	1690.724**	-498314.560	-641721.280
	(630.280)	(654.429)	(606327.300)	(808353.900)
Individual Controls	No	Yes	No	Yes
Observations	180	180	180	180
R-squared	0.154	0.224	0.031	0.114
Mean of dep. var. (placebo)	1539.453	1539.453	928546.1	928546.1

Defiers' WTP for RCT less or unaffected by treatment

Effect of Treatment on WTP by Defiers

	Private Spending		*Public Spending*	
	Amount Randomized Trial	*Amount Randomized Trial*	*Amount Randomized Trial*	*Amount Randomized Trial*
	(1)	(2)	(3)	(4)
Defiers X Metrics Assigned	-550.306	-888.410	-1987315.100***	-2531651.000*
	(1177.783)	(1301.347)	(667571.300)	(1367816.000)
Metrics Assigned	2342.989***	2179.581***	1986365.500***	1775415.200***
	(711.613)	(648.553)	(667569.300)	(635063.500)
Defiers X Metrics Chosen	-1377.298	-428.640	495448.220	3326010.500
	(1196.483)	(1207.932)	(606330.500)	(2437943.000)
Defiers	-160.589	-463.802	-1172597.800*	-3101800.100
	(629.255)	(857.643)	(610215.800)	(2201009.000)
Metrics Chosen	1511.444**	1690.724**	-498314.560	-641721.280
	(630.280)	(654.429)	(606327.300)	(808353.900)
Individual Controls	No	Yes	No	Yes
Observations	180	180	180	180
R-squared	0.154	0.224	0.031	0.114
Mean of dep. var. (placebo)	1539.453	1539.453	928546.1	928546.1

Defiers' WTP for RCT less or unaffected by treatment

Concluding remarks

- Econometrics training yielded
  - Persistent effects on policymakers in attitudes, beliefs, and exams (0.5-0.8 sigma)
- Greater **demand for RCTs** (300%)
  - Reduced demand for correlational information (50%)
- Enhanced **responsiveness to RCTs**
  - More likely to adopt a policy when shown RCT evidence
  - Only when priors are less than the evidence
- Receptiveness to policy evidence only with econometrics training
  - 40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)
- Updated policy choices only present with econometrics training
  - twice as likely to adopt policy with RCT evidence

Concluding remarks

- Econometrics training yielded
  - Persistent effects on policymakers in attitudes, beliefs, and exams (0.5-0.8 sigma)
- Greater **demand for RCTs** (300%)
  - Reduced demand for correlational information (50%)
- Enhanced **responsiveness to RCTs**
  - More likely to adopt a policy when shown RCT evidence
  - Only when priors are less than the evidence
- Receptiveness to policy evidence only with econometrics training
  - 40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)
- Updated policy choices only present with econometrics training
  - twice as likely to adopt policy with RCT evidence

Concluding remarks

- Econometrics training yielded
  - Persistent effects on policymakers in attitudes, beliefs, and exams (0.5-0.8 sigma)
- Greater **demand for RCTs** (300%)
  - Reduced demand for correlational information (50%)
- Enhanced **responsiveness to RCTs**
  - More likely to adopt a policy when shown RCT evidence
  - Only when priors are less than the evidence
- Receptiveness to policy evidence only with econometrics training
  - 40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)
- Updated policy choices only present with econometrics training
  - twice as likely to adopt policy with RCT evidence

Concluding remarks

- Econometrics training yielded
  - Persistent effects on policymakers in attitudes, beliefs, and exams (0.5-0.8 sigma)
- Greater **demand for RCTs** (300%)
  - Reduced demand for correlational information (50%)
- Enhanced **responsiveness to RCTs**
  - More likely to adopt a policy when shown RCT evidence
  - Only when priors are less than the evidence
- Receptiveness to policy evidence only with econometrics training
  - 40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)
- Updated policy choices only present with econometrics training
  - twice as likely to adopt policy with RCT evidence

Concluding remarks

- Econometrics training yielded
  - Persistent effects on policymakers in attitudes, beliefs, and exams (0.5-0.8 sigma)
- Greater **demand for RCTs** (300%)
  - Reduced demand for correlational information (50%)
- Enhanced **responsiveness to RCTs**
  - More likely to adopt a policy when shown RCT evidence
  - Only when priors are less than the evidence
- Receptiveness to policy evidence only with econometrics training
  - 40 percent more likely to adopt policy $>$ 10 percent (Hjort et al. 2021)
- Updated policy choices only present with econometrics training
  - twice as likely to adopt policy with RCT evidence