IDEAL

Impact Data and Evidence Aggregation Library

Session: Outcomes

Sergio Puerto and Jennie Barker

June 10, 2025

INTRODUCTION

This session discusses the fields related to outcomes

These fields, split across Stage 1 and 3 of the survey, provide critical information for how to interpret treatment effects.

Agenda

1. Number of outcomes
2. Outcome name
3. Outcome unit of analysis
4. Unit of analysis inclusion criteria

1. Outcome variable definition
2. Outcome variable type and measurement

1 Stage 1 Outcome Fields

1. Number of Outcomes

- **Outcome:** variable for which treatment effects are estimated in the main text of the paper
  - We are interested in the number of outcomes in each exhibit, not necessarily all of the outcomes mentioned by the authors but not in an exhibit
    - Count all outcomes at the study level
    - Then map outcomes at the table level

1. Number of Outcomes in an Exhibit

- **What counts as an outcome?**
  - **Include:**
    - If an index and its components are listed, include each separately.
    - If an outcome is measured as marginal effects only
    - If exact point estimates are only in the text and not in exhibit

Note: Outcomes and Heterogeneous Treatment Effects

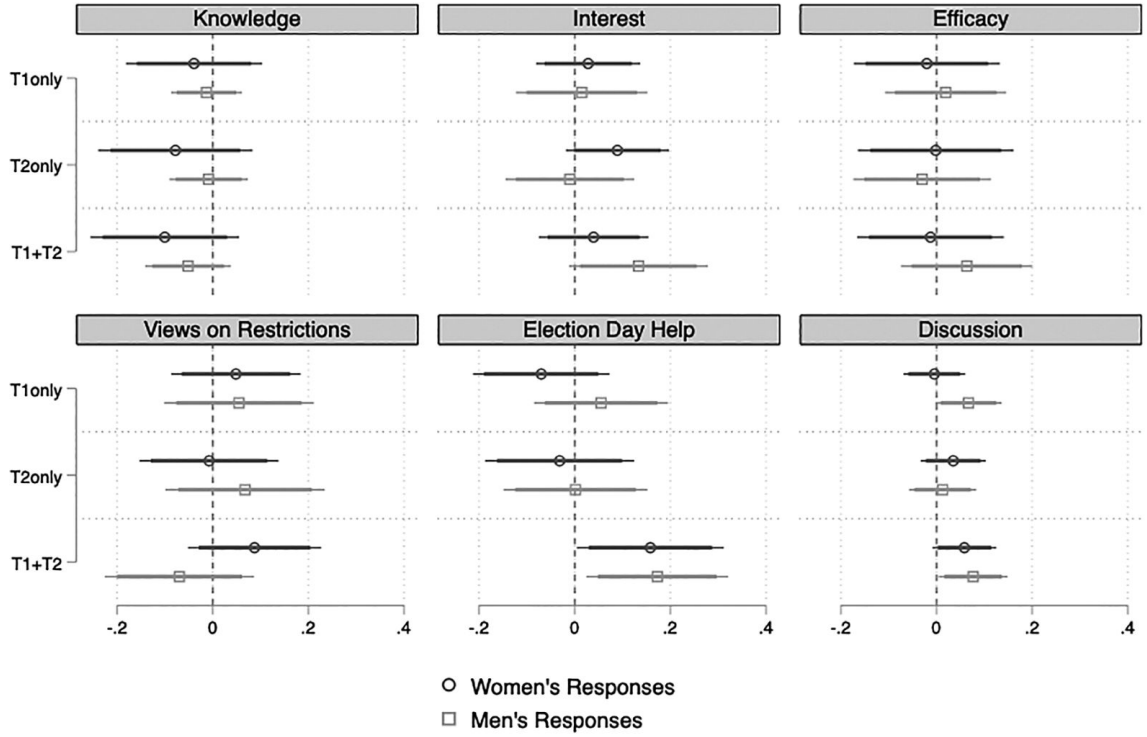
- For exhibits that meet the criteria to be included in the pilot that include heterogeneous treatment effects (e.g., for subgroups within a population):
  - Include as **separate outcomes** if outcome is reported for multiple groups
    - Female farmer knowledge
    - Male farmer knowledge
- HTE will be included:
  - If those are the only treatment effects reported
  - If it is the main research question

Cheema et al. 2023

Figure 4 includes **12** outcomes

- Variables are indices, but components only in appendix
- Women and men's responses as separate outcomes
- Point estimates are in appendix

FIGURE 4. Knowledge, Attitudes, and Behavior ITT by Respondent Gender and Treatment Category

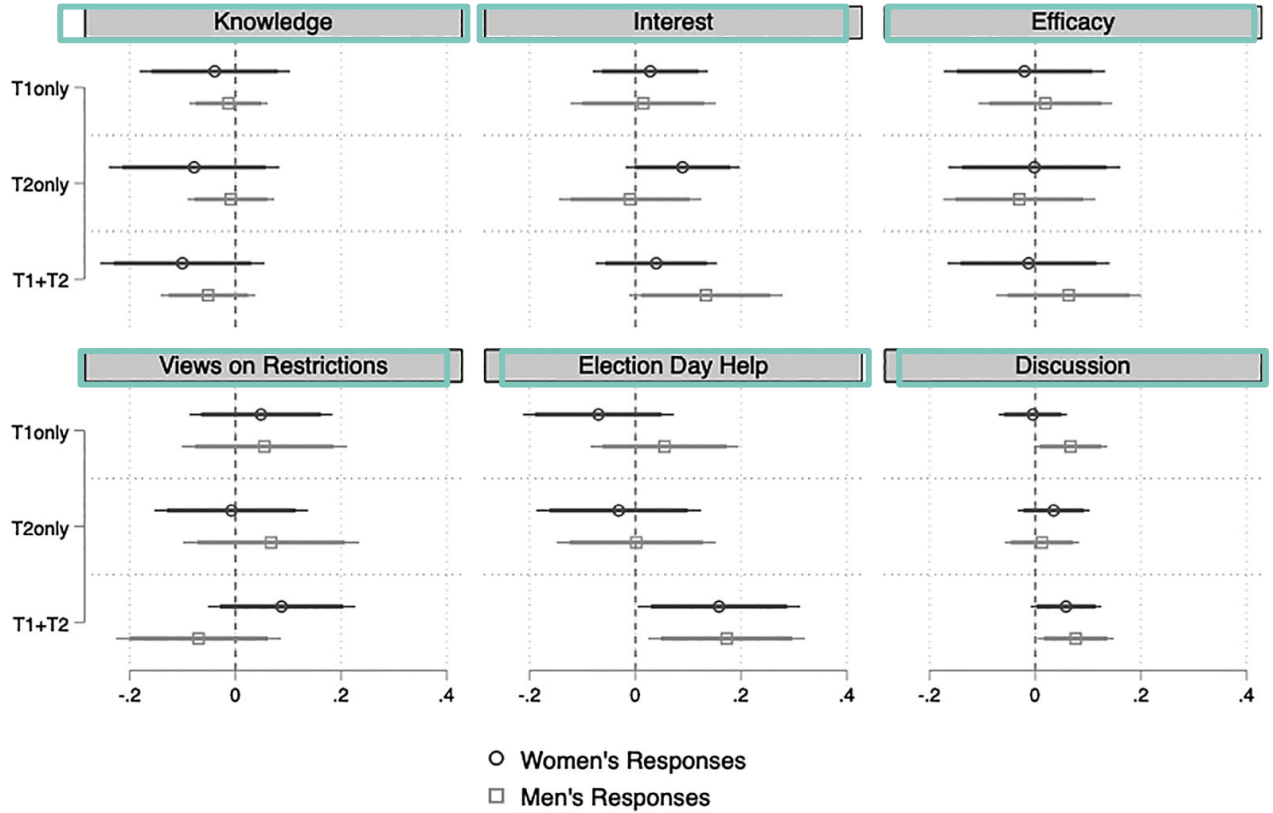


(v) election day help from men, and (vi) political discussion between household members. The results are shown in Figure 4. Appendix C.1 shows the corresponding regression table, and Appendix C.3 shows results on the individual components of index-based measures.

2. Outcome Name

- **Definition:** the name **used in the exhibit** to refer to the outcome variable
    - Exhibit outcome variable names preferred over how they might be otherwise described in the text

FIGURE 4. Knowledge, Attitudes, and Behavior ITT by Respondent Gender and Treatment Category



3. Outcome unit of analysis

- **Definition:** Unit of analysis when estimating the treatment effect
  - The unit of analysis of the outcome is often distinct from the unit of randomization
  - Different outcomes in a paper, even in the same exhibit, may have different units of analyses
    - Child
    - Parent

3. Outcome unit of analysis

- The survey includes a list of units.
  - Link to CV
  - If there is not an exact match, select other and **enter the name of the unit as the authors do.**
    - E.g., 1. Individual, 1.12 Other, “Child”

Poll: Freeman et al. 2022

What unit(s) of analysis?

S5: *All household members use a latrine every time they defecate*				
• Respondent always exclusively used a latrine for defecation during last 7 days	743	53.2	729	54.1
• Head of household always exclusively used a latrine for defecation during last 7 days	529	36.5	473	33.0
• Ages 4–17 always exclusively used a latrine for defecation during last 7 days	1447	42.6	1385	35.0

4. Outcome unit of analysis inclusion criteria

- **Definition:** Inclusion criteria for the specified unit of analysis when estimating the treatment effect
  - The actual unit of observation might be different from unit of analysis and only a subsample in the unit of analysis was included in data collection, for example, children under age 5, employed members in the household.
  - This field captures the inclusion criteria within the unit of analysis.
  - This will often require reading the text sections on outcomes or samples, as space constraints may limit this information in the exhibits themselves

Pickering et al. 2019

Used antibiotics‡	2086	921 (44·2%)	2011	797 (39·6%)	0·93 (0·88 to 0·98)	0·0038
Weight-for-age Z score	3084	−1·21 (1·01)	3014	−1·25 (1·02)	Difference −0·03 (−0·10 to 0·05)	0·51
Height-for-age Z score	414	−1·36 (1·05)	385	−1·33 (1·08)	Difference 0·07 (−0·07 to 0·22)	0·30
BDT spent on treatment‡	3142	373·0 (993·0)	3063	333·0 (858·0)	Difference −0·08§ (−0·15 to −0·02)	0·016

Data are n (%) or mean (SD), unless otherwise stated. Prevalence ratios estimated by Poisson regression for binary outcomes and mean differences estimated by linear regression for continuous outcomes. All models include fixed effects for randomisation pair and month of data collection, as well as robust standard errors at the cluster (water point) level. BDT=Bangladesh taka. *1-week recall. †Observed at time of interview. ‡2-month recall. §log10-transformed difference.

Table 2: **Effect of the intervention on child diarrhoea and related illness outcomes over all follow-up survey rounds**

chlorine and vitamin C refills were depleted. Field staff did surveys with the primary caregivers of children younger than 5 years of age residing in all enrolled households to measure outcomes. Additional information on demo-

Pickering et al. 2019

Tap water quality						
Detectable total Cl (proportion)	2337	0·00 (0·04)	2009	0·83 (0·38)	0·81 (0·73 to 0·89)	<0·0001
Detectable free Cl (proportion)	2335	0·00 (0·04)	2003	0·80 (0·40)	0·78 (0·71 to 0·86)	<0·0001
Mean total Cl (ppm)	2337	0·00 (0·01)	2335	0·37 (0·32)	0·35 (0·21 to 0·40)	<0·0001
Mean free Cl (ppm)	2335	0·00 (0·01)	2003	0·33 (0·28)	0·31 (0·28 to 0·35)	<0·0001

Table 3: **Water quality at the point-of-collection (shared taps) and in household stored drinking water, by treatment status, over all follow-up survey rounds**

chlorine and vitamin C refills were depleted. Field staff did surveys with the primary caregivers of children younger than 5 years of age residing in all enrolled households to measure outcomes. Additional information on demo-

2 Stage 3 Outcome Fields

1. Outcome variable definition

- **Definition:** A short definition of the outcome variable
  - What is the outcome variable designed to quantify?
    - Be explicit about the meaning of the outcome value
      - Behavior = improved/better behavior or more behavior problems?
  - Be as brief as possible

Garbiras-Diaz and Montenegro (2022)

TABLE 3—IMPACTS ON REPORTS

	Reports (= 1) (1)	N. reports (2)	High quality reports (= 1) (3)	High quality n. reports (4)
Panel A. Pooled treatment				
[*T*] Any treatment	0.106	0.366	0.088	0.188
	(0.035)	(0.100)	(0.030)	(0.053)
	[0.010]	[0.005]	[0.005]	[0.003]

Results = whether any report was issued in a given municipality

N. reports = total number of reports issued in a given municipality

High quality reports = whether MOE classify a report as medium or high level of trustworthiness

High quality n. Reports = total number of reports classified by MOE as medium or high trustworthiness

2. Outcome Variable Measurement and Type

Binary	Index	Cardinal	Ordinal	Categorical
Two possible values	Multiple independently measured variables that assess different dimensions or units	Levels and differences between two values are meaningful	Scale, rank or position, but difference in values does not have numerical interpretation	Values indicate categories, not positions, and cannot be ordered

First: Binary outcome variable?

Definition

Two possible values

- Indicator/dummy
  - Yes/no
- Two categories
  - Male/female
- Estimated using linear probability, logit, or probit

Example: Freeman et al. 2022

Prevalence
High Anxiety [f]
High Depression [f]
High Emotional distress [f]
Poor well-being [g]

[f] Each of the above scores was dichotomized, with scores greater than 1.75 indicating a positive status for any of the three outcomes

[g] The above score was dichotomized with scores below 13 indicating poor well-being.

Binary outcome variable label

Definition	**Example: Freeman et al. 2022**
Meaning of outcome variable when value is equal to 1	● "...scores greater than 1.75 indicating a positive status" ○ High anxiety: when value is equal to 1, **a respondent has high anxiety**

Then: Index outcome variable?

Definition	**Example: Wolf et al. 2019**
Index/score - Multiple independently measured variables - Assess different dimensions of a concept	**“School readiness composite:** For the primary impact analysis, scores on the four domains were combined with equal weight to create a total ‘school readiness’ score.” (21)

Index outcome variable components

Definition

Select the components used to create the index outcome variable

- **Note** that some times these are not reported in the paper, so you should select 'None.'

Example: Wolf et al. 2019

Table 5. Impacts on children's school readiness outcomes.

	b	(*SE*)	*p* value
School readiness composite			
TT	0.025	(0.010)	.010**
TTPA	0.004	(0.010)	.691
Post-hoc estimates by domain			
Early numeracy			
TT	0.020	(0.009)	.031*
TTPA	−0.005	(0.010)	.609
Early literacy			
TT	0.022	(0.011)	.045*
TTPA	−0.006	(0.013)	.670
Social-emotional			
TT	0.035	(0.013)	.010**
TTPA	0.014	(0.013)	.286
Executive function			
TT	0.020	(0.014)	.148
TTPA	0.007	(0.014)	.589

Index outcome variable description

Definition	**Example: Wolf et al. 2019**
Describe the components, aggregation method, or any other details of the index outcome	"School readiness composite: For the primary impact analysis, **scores on the four domains were combined with equal weight** to create a total 'school readiness' score." (21) → Early numeracy, early literacy, socio-emotional, and executive function were combined with equal weight.

Outcome measurement tool

- **Definition:** Name of the tool used to measure the outcome variable
  - Name *and* citation
    - Proprietary tool
    - Tool used by other authors (cited)
      - **Adapted** tools are included

Outcome measurement tool

- **Wolf et al. 2019:**
  - School readiness composite and its components are measured using IDELA by Save the Children

Child school readiness outcomes. School readiness was directly assessed in four domains: early literacy, early numeracy, social-emotional skills, and executive function. The instrument used was the International Development and Early Learning Assessment (IDELA), developed by Save the Children (Pisani, Borisova, & Dowd, 2015). The tool was translated into three local languages (Twi, Ewe, and Ga). Surveys were translated and then back-translated by a different person to check for accuracy. Any discrepancies

Outcome variable standardization

- An outcome is standardized if it is converted from the original raw scale to a specific standard scale or z-score using mean and SD
  - **Internally standardized:** using mean and SD of any group in study sample
  - **Externally standardized**: using distribution of sample outside of study (e.g., Bayley Scales)
  - Provide information on procedure, statistics and sample
    - Often for only internal standardization

Poll

How is the following variable from Ganimian et al. 2023 standardized?

A. Internally
B. Externally
C. Not standardized

round 16 months after program rollout (March-April 2018). The test instruments were designed to minimize ceiling and floor effects and produce a distribution with broad support. Baseline test scores are standardized ($\mu = 0$, $\sigma = 1$) in the full sample, and endline scores are standardized relative to the control group distribution. Appendix C provides more details on test construction, characteristics, and administration.

Finally: Outcome variable type

- If the variable is **non-binary, not an index, AND is not standardized**, then you will receive a question asking about other variable types that we discussed:
  - **Cardinal**: levels and differences can be interpreted in a meaningful way
    - Ratios, percentages, currency, z-scores, odds ratios
  - **Ordinal**: values have scale or rank, so it can be ordered, but differences cannot be interpreted
    - Likert scale
  - **Categorical:** values are labels and cannot be ordered
    - Categorical variables will be flagged

Poll

What are the following variables from Chong et al. 2015 examples of?

A. Cardinal variables
B. Ordinal variables
C. Categorical variables

Note The dependent variables are: in columns (1)–(3) total number of votes; in columns (4)–(6) votes for the incumbent party; and in column (7) and (9) votes for any challenger party. All dependent variables are divided by number of registered voters and multiplied by 100. All specifications include municipality fixed

Poll

What are the following variables from Aker et al. 2017 examples of?

A. Cardinal variables
B. Ordinal variables
C. Categorical variables

After matching the electoral problems with our experimental locations, we coded each of the problematic locations as having had election day misconduct, campaign misconduct, or violence and intimidation. We also constructed a measure of the highest intensity of electoral problems for each problematic polling location. This measure has five categories: 1 corresponds to minor problems; 2 corresponds to nonviolent occurrences, including campaign misconduct and election day problems; 3 corresponds to occurrences leading to physical intimidation, including vandalism; 4 corresponds to occurrences resulting in injured people; and 5 corresponds to occurrences resulting in deaths. These data

Outcome variable endpoint values and labels

- For **non-standardized index variables and ordinal variables**, we want to know highest and lowest values and their labels
  - **Aker et al. 2017:**
    - Lowest possible value: 0
    - Lowest possible value label: No electoral problems
    - Highest possible value: 5
    - Highest possible value label: Occurrences resulting in homicides

Outcome variable with time reference

- **Definition:** Indicates whether the outcome variable is measured over a specific time reference period
  - Variable reference period
    - Examples:
      - Incidence in last 24 hours
      - Monthly income
  - Invariable reference periods
    - Examples
      - Height for age
      - Baseline to endline height change

Variable time period

1. **Basinga et al., 2011:** Outcome is “Younger than 23 months preventative visit in previous 4 weeks”

Invariable time period

1. **Basinga et al., 2011:** Outcome is “Any prenatal care”

Outcome variable with time reference

- **If variable, CV:**
  - 24 hours/day
  - 7 days/one week
  - 14 days/two weeks
  - 30 days
  - 4 weeks
  - One month
  - One minute/60 seconds
  - One hour/60 minutes
  - One year
  - Other, specify

Outcome variable unit of measurement

- For cardinal variables
  - **Definition:** Unit of measurement of the outcome variable
    - Why is this important?
      - How variables are measured vs. how they are estimated
      - If they are transformed (e.g., log transformation)

- Directions
  - Start with broad category and then specify
  - Transformed variables fall into the category of “unitless or other”

Outcome variable unit of measurement

Currency	Country currency	Year	Nominal, real
Time	Year, month, week, day, hour, minute, other (specify)		
Height	Cm, inch, meter, feet, height for age, height for weight, other (specify)		
Weight	Gram, kg, ounce, pound, ton, other		
Percent (0-100)			
Non-binary fraction (0-1)			
Odds ratio			
Count	Number of [ ]		
Standard deviation			
Unitless or other (specify)	Log of [ ], sine of [ ], inverse hyperbolic function of [ ]		

Pickering et al. 2019

1. **Weight for age z-score**

Standard deviation

Basinga et al. 2011

1. **Younger than 23 months—preventive visit in previous 4 weeks (%)**

Percent

Field	High anxiety (binary)	School readiness composite (index)	Vote percentage (cardinal)
Binary outcome	Yes	No	No
Binary outcome label	High anxiety	**X**	**X**
Index outcome	**X**	Yes	No
Description of index outcome	**X**	See previous slide	**X**
Outcome measurement tool	No	Yes - IDELA	No
Outcome standardization type	No	No	No
Outcome standard. details	**X**	**X**	**X**
Outcome variable type	**X**	**X**	Cardinal
Endpoint values and labels	**X**	Asked, but NA	**X**
Variable with time reference period	No	No	No
Time ref. period for measurement	**X**	**X**	**X**
Unit of measurement	**X**	**X**	Percent

Outcome variable target population

- **Definition:** Description of target population of outcome variable
  - What is the target population **if it is different** than the unit of analysis?
    - If target unit and unit of analysis are the same, then enter ‘None’

Outcome variable target population

- **Definition:** Description of target population of outcome variable
  - What is the target population **if it is different** than the unit of analysis?
    - If target unit and unit of analysis are the same, then enter ‘None’
  - Example:
    - Unit of analysis: child
    - Target population: children under 5

Summary Fields for Outcomes

- **Additional details**
  - Information related to construction or processing (imputation), validation or quality control (double entry, back checks), and measurement (exact procedures).
    - Information that **is not captured** by previous fields
      - E.g., Coding decisions

Summary Fields for Outcomes

- **Summary of outcome variable definition**
  - Revised ***prefilled*** definition based on existing answers to the following fields
    - Outcome variable definition
    - Unit of analysis
    - Unit of analysis inclusion criteria
    - Time reference period
  - Edit to make it grammatically correct and readable

Summary Fields for Outcomes

- **Basinga et al. 2011:**
  - Outcome variable definition: Child visited a health facility for preventative care
  - Unit of analysis: Child
  - Unit of analysis inclusion criteria: Child aged 23 months or younger
  - Time reference period: 4 weeks

- Example: Child visited a health facility for preventative care over a four week period (Child aged 23 months or younger) → **Any preventative care health facility visit over a 4-week period (Child under 2 years of age)**

Thank you for listening

Sergio Puerto
sergiopuerto@berkeley.edu

Jennie Barker
jlbarker@berkeley.edu

IDEAL
Impact Data and Evidence Aggregation Library