

SURVEY FIELDS & CODING PROTOCOL

STAGE 1

# Stage 1 Set-Up

### Coder name

| **Variable name in SurveyCTO** | [coder] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Coder name |
| **Definition (LABEL on SurveyCTO)** | Name of the coder |
| **Response options (open-text, numeric, date, text-CV)** | text-CV |
| **CV (Choices on SurveyCTO)** | list of coders |
| **SurveyCTO instructions for data entry mask and repeat level** | Drop down list |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Please select your name in SurveyCTO - If your name does not appear, reach out to your supervisor |
| **Descriptive example see the[section]used in the paper to extract** |  |

### Paper ID

| **Variable name in SurveyCTO** | [paperID] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Paper ID |
| **Definition (LABEL on SurveyCTO)** | The IDEAL ID of the paper |
| **Response options (open-text, numeric, date, text-CV)** | text-CV |
| **CV (Choices on SurveyCTO)** | List of papers IDs assigned to the coder. |
| **SurveyCTO instructions for data entry mask and repeat level** | Drop down |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Select the IDEAL paper ID for which you will start extraction. - If the paper ID does not appear on the list, reach out to your supervisor. |
| **Descriptive example see the[section]used in the paper to extract** |  |

### Paper title

| **Variable name in SurveyCTO** | [X\_titleConfirm] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Paper title |
| **Definition (LABEL on SurveyCTO)** | Please confirm: The title of the paper on SurveyCTO. |
| **Response options (open-text, numeric, date, text-CV)** | text-CV |
| **CV (Choices on SurveyCTO)** | **Select one** Yes No |
| **SurveyCTO instructions for data entry mask and repeat level** | Title associated with the ID appears  End survey if field = no |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Confirm the name of the paper that appears on SurveyCTO matches the name of the paper on the list assigned to you. - If you have selected no, a message will be sent to your supervisor. Please await a new corrected paper assignment. |
| **Descriptive example see the[section]used in the paper to extract** |  |

### Paper correction

| **Variable name in SurveyCTO** | [X\_correction] |
| --- | --- |
| **Field name  (LABEL on SurveyCTO)** | Correction |
| **Definition  (LABEL on SurveyCTO)** | Check if you are correcting an earlier entry |
| **Response options (open-text, numeric, date, text-CV)** | **Check box** |
| **CV  (Choices on SurveyCTO)** |  |
| **SurveyCTO instructions for data entry mask and repeat level** | If this box is checked, the last entry for this paper should be pre-loaded for all questions. |
| **Coding instructions for coders (Hint on SurveyCTO)** | If you have received a resubmission request for your previous entry and are ready to start your new entry, please check this box. |
| **Descriptive example see the[section]used in the paper to extract** |  |

### Multi-site study entry

| **Variable name in SurveyCTO** | [MulExp] |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Multiple experiments |
| **Definition**  **(LABEL on SurveyCTO)** | Check if you are entering responses for an experiment in a previously submitted paper. |
| **Response options (open-text, numeric, date, text-CV)** | **Check box** |
| **CV**  **(Choices on SurveyCTO)** | **N.A.** |
| **SurveyCTO instructions for data entry mask and repeat level** | If this box is checked, the last entry for this paper should be pre-loaded for all questions.  Additionally, if this box is checked, a new entry ID should be created for this entry, as “” |
| **Coding instructions for coders** **(Hint on SurveyCTO)** | If you are coding a multi-site or multi-experiment paper, please enter each experiment separately. Check this box if you are entering another experiment for a previously submitted paper.  Please do not check this box if this is the first response for the paper. |
| **Descriptive example** **see the[section]used in the paper to extract** | [Barrera-Osorio et al. 2011](https://www.aeaweb.org/articles?id=10.1257/app.3.2.167) report effects from two different experiments in Bogota in different parts of the city (San Cristobal and Suba).  We would like you to coder the experiments in San Cristobal and Suba separately. Only if you have submitted the entry for San Cristobal and are going to coder the Suba experiment, please check this box. |

### Request for review: fields

| **Level** |  |
| --- | --- |
| **Definition** | Would you like to request review for any of your responses? |
| **Response options** (open-text, numeric, date, text-CV select one, text-CV select all) | Text CV select all |
| **Controlled Vocabulary** | List of fields in order of appearance |
| **Cardinality (extraction)** |  |
| **Instruction for data entry mask** | After each logical group of 4-10 questions, create a multiple choice select-all question with the list of fields prior to this question. |
| **Instruction for coders** (This will go to the “Hints” for surveyCTO) | Please select all questions you were unsure about. |
| **Descriptive examples for coding** |  |

### Request for review: detail

| **Level** |  |
| --- | --- |
| **Definition** | Please provide details. |
| **Response options** (open-text, numeric, date, text-CV select one, text-CV select all) | Open text |
| **Controlled Vocabulary** |  |
| **Cardinality (extraction)** |  |
| **Instruction for data entry mask** | If any of the CHECK questions selected in the previous question |
| **Instruction for coders** (This will go to the “Hints” for surveyCTO) | For each field, explain what you were uncertain about and (if applicable) which options you were considering. |
| **Descriptive examples for coding** | For example, if a coder selects the “Estimand is full sample ITT and LATE/TOT” field in the request-for-review section, they could add an explanation about their uncertainty in this field.  “Unsure regarding the empirical specification of the estimand. The author mentions that "we view the evaluation design as quasi-experimental and use difference-in-differences to estimate program impact." Are the estimates indirectly ITT/LATE/TOT or are they only quasi-experimental estimates?” |

# **Stage 1 Fields**

### Number of experiments in the study

| **Variable name in SurveyCTO** | [expNum] |
| --- | --- |
| **Field name** **(LABEL on SurveyCTO)** | Number of experiments in the study |
| **Definition** **(LABEL on SurveyCTO)** | Number of experiments under evaluation in the paper |
| **Response options (open-text, numeric, date, text-CV)** | Numeric |
| **CV** **(Choices on SurveyCTO)** | None |
| **Coding instructions for coders** **(Hint on SurveyCTO)** | - Please indicate the number of experiments being evaluated in the paper. - An experiment is principally defined by the study population and unit of randomization, the intervention, and the randomization used to create comparable treatment arms.  -If results are reported from multiple countries, these are likely coming from different experiments.  -Normally, there is only one experiment being evaluated in a paper, but there are exceptions. Please see the example column. The experimental design section often provides information on how many experiments are being tested in the paper.  -Note that in the pilot, we are not coding studies that are lab-in-the-field experiments. If a study includes a field experiment and a lab-in-the-field experiment, we will code only code the field experiment. In that case, please enter 1 for this field.  -We are also not coding studies in which there is a design intervention rather than a policy intervention. A design intervention includes studies that randomize the order or wording of survey questions. |
| **Descriptive example** **see the[section]used in the paper to extract** | [Barrera-Osorio et al. 2011](https://www.aeaweb.org/articles?id=10.1257/app.3.2.167) report effects from two different experiments in Bogota in different parts of the city (San Cristobal and Suba). We know that there are two different experiments because the paper declares "As required by the SED, the assessment of the treatments was divided into two separate experiments located in two very similar localities in Bogota, San Cristobal, and Suba." The paper also reports that eligible populations for the tested interventions are different across the two sites: "Eligible registrants in San Cristobal, ranging from grade 6–11..."; and "The tertiary treatment was evaluated separately in an experiment in Suba, where students ranging from grade nine through eleven..."  The response to this field would be 2.  [Jeong et al. 2023](https://doi.org/10.1016/j.jdeveco.2022.102992) evaluate the differences in how question modules in a survey are ordered in order to examine the effects of survey fatigue.  In this case, the coder should enter 0.  [De Martino et al. 2015](https://doi.org/10.1177/1091142115604352) conduct a lab-in-the-field experiment on landholders and annual payment offers for environmental services.  As the experiment conducted was hypothetical, the coder should enter 0. |

### Number of experiments check

| **Variable name in SurveyCTO** | [expNumcheck] |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Number of experiments check |
| **Definition**  **(LABEL on SurveyCTO)** | You entered the number of experiments other than 1, please double check the number of experiments entered. |
| **Response options (open-text, numeric, date, text-CV)** | Text-CV, select one |
| **CV**  **(Choices on SurveyCTO)** | -Confirm that there is no eligible experiment reported in the paper.  -Confirm that there is more than one experiment reported in the paper. |
| **SurveyCTO instructions for data entry mask and repeat level** | Display this field if “Number of experiments in the study” is not 1.  The survey should stop if “Confirm that there is no eligible experiment reported in the paper” is selected.  Display the below warning message if “Confirm that there is more than one experiment reported in the paper”.   * You are entering responses for a multi-experiment paper. Please only focus on and code one experiment in this response. Start a new survey entry for any additional experiment. |
| **Coding instructions for coders** **(Hint on SurveyCTO)** | -You are seeing this question because you indicated that the number of experiments reported in the current papers was not 1. Please double check your answer.  -Go to the previous question to revise your answer if you think the number of experiments should be 1.  -Select “Confirm that there is no eligible experiment reported in the paper” if you believe the experiment reported in the paper is not an RCT or not a field RCT. This could include cases where a study only reports lab-in-the field experiments or a design intervention that randomizes the order or wording of survey questions. Note that this will be a rare case since all assigned papers have been pre-screened.  The survey will stop after you select this option. Please notify your supervisor that the paper is not eligible for coding.  -Select “Confirm that there is more than one experiment reported in the paper” when the paper evaluates more than one experiment. In that case, please enter the information for each experiment separately into different survey entries.  -Please reach out to your supervisors if you are not sure about your answer to this question. |
| **Descriptive example** **see the[section]used in the paper to extract** | [Barrera-Osorio et al. 2011](https://www.aeaweb.org/articles?id=10.1257/app.3.2.167) report effects from two different experiments in Bogota in different parts of the city (San Cristobal and Suba).  The response to this field would be “Confirm that there is more than one experiment reported in the paper”.  [Jeong et al. 2023](https://doi.org/10.1016/j.jdeveco.2022.102992) evaluate the differences in how question modules in a survey are ordered in order to examine the effects of survey fatigue.  In this case, the coder should select “Confirm that there is no eligible experiment reported in the paper” and notify the supervisor that this is a design experiment. |

### Country

| **Level** | Study |
| --- | --- |
| **Definition** | The country in which the experiment was implemented. |
| **Response options** (open-text, numeric, date, text-CV select one, text-CV select all) | Text-CV select one |
| **Controlled Vocabulary** | ISO country codes, "other", "not stated" |
| **Cardinality (extraction)** | 1..n Mandatory and repeatable |
| **Instruction for data entry mask** | -In the choices tab of SurveyCTO, make a list of all the country codes with labels containing the country name. Add an open-text field if "other" is selected.  -In the pilot, we only allow papers with one experiment and one country. |
| **Instruction for coders** (This will go to the “Hints” for surveyCTO) | -Select the country in which the intervention took place, even if the study did not cover the entire country or only mentions a region or city in the country.  -If the country is not found in the CV, select "other" and then write the name of the country.   -If you cannot find the name of the country in the main text of the paper or its supplementary materials/appendices, enter "not reported" and specify that you are "unsure" about this field in the request-for-review section of the survey.   -This information is usually found in the abstract, introduction, context, or research design sections of the main text of the paper. |
| **Descriptive examples for coding** | The intervention in [Chong et al. 2015](https://doi.org/10.1086/678766) takes place in Mexico, so the coder would select the ISO code and country name for Mexico. [see: Abstract, Introduction, Experimental Design and Implementation sections]  The intervention in [Lyall et al. 2020](https://www.cambridge.org/core/journals/american-political-science-review/article/can-economic-assistance-shape-combatant-support-in-wartime-experimental-evidence-from-afghanistan/CDD1F42DC1506A23A1AF3B9FA20F4A12) takes place in Afghanistan, so the coder would choose the ISO code and country name for Afghanistan. [see: Abstract, Introduction]  The intervention in [Muralidharan et al. 2021](https://www.aeaweb.org/articles?id=10.1257/app.20190783) takes place in Telangana, which is a state in India, so the coder would select the ISO code for India. [see: Abstract, Introduction, Setting and Intervention, and Research Methods sections]  The intervention in [Adida et al. 2020](https://doi.org/10.1177/0010414019879945) takes place in Benin, so the coder would select the ISO code for Benin. [See: Abstract] |

### Sub-national location

| **Level** | Study |
| --- | --- |
| **Definition** | The subnational location where the experiment took place. |
| **Response options** (open-text, numeric, date, text-CV select one, text-CV select all) | Open-text |
| **Controlled Vocabulary** |  |
| **Cardinality (extraction)** | 1..1 Mandatory and non-repeatable |
| **Instruction for data entry mask** |  |
| **Instruction for coders** (This will go to the “Hints” for surveyCTO) | - Enter the largest geographic location within a country where the experiment took place.  - The location is often not randomly drawn from the country, for example, six states in Mexico were included in the study for some criteria or unspecified reasons.  - For multi-site studies, the location(s) should differentiate the current experimental site from other experiments reported in the same paper. Please only enter the location for the experiment you are entering information for in this entry. |
| **Descriptive examples for coding** | [Barrera-Osorio et al. 2011](https://www.aeaweb.org/articles?id=10.1257/app.3.2.167) report effects from two different experiments in Bogota in different parts of the city (San Cristobal and Suba).  If a coder is entering the information for the experiment in San Cristobal, they should only enter “San Cristobal”.  [Gaikwad and Nellis](https://nikhargaikwad.com/resources/Gaikwad-Nellis-2021_Exclusion.pdf) (2021) report effects from the same experiment in two cities in India, Delhi and Lucknow. In this case, the coder should enter “Delhi and Lucknow.” |

### Intervention assignment strategy

| **Variable name in SurveyCTO** | [intAssign] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Intervention assignment strategy |
| **Definition (LABEL on SurveyCTO)** | The strategy used for assigning interventions to study arms |
| **Response options (open-text, numeric, date, text-CV)** | text-CV |
| **CV (Choices on SurveyCTO)** | **Select one** Parallel Factorial Crossover Adaptive Other, specify |
| SurveyCTO instructions for data entry mask and repeat level | Display a text field, if “Other, specify” is selected. |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Parallel: This is the most common strategy used in randomized control trials. Each intervention is assigned to only one arm.  - Factorial: This design is used when evaluating the impact of two or more interventions alone and in combination with each other. At least one intervention is assigned to more than one study arm.  - Crossover: This is used when each study arm receives different interventions (including no intervention) in different phases of the study. Also select this option for phase-in or stepped-wedge designs, where the roll-out of the intervention is randomized and every unit ultimately receives the program. If the study endline occurs prior to units receiving interventions beyond what they were initially assigned to, **do not select this option.**   - Adaptive: In adaptive designs, the rule by which interventions are randomly assigned can change in the course of the trial, based on the experimental data. For example, in a trial conducted in multiple waves, the number of units assigned to each treatment arm may change across waves based on results in prior waves. Alternatively, in a longitudinal cross-over design, the next intervention to which an experimental unit is "switched" (or the time of switching) may depend on the outcome under the current intervention.  - Other:If the assignment strategy does not fit in any of the above categories, select this option  - Information on the assignment strategy is found in the experimental/study design or methods sections of a paper. -When available, the coder may also consult how an intervention is described in a study participant flow diagram.  - Note that authors may use the word phased-in while describing rollout of a program, even if the intervention assignment strategy is parallel or factorial.  - Note also that some studies may feature a phase-in design or analyze a pilot program, in which a broader population ultimately receives the intervention(s). However, if the study endline is conducted prior to the phase-in of the rest of the units or before the program is expanded, the study will still be included if it has a parallel or factorial intervention assignment strategy. |
| **Descriptive example** | *Parallel:*In [Barrera-Osorio et al, 2022](https://www.sciencedirect.com/science/article/pii/S0304387822000049#tbl1) , authors evaluate the performance-based reward program by randomizing primary schools into three **distinct** groups -- recognition, in-kind performance reward, and control [see: Sample and experimental design]. The response to this field is *parallel*  *Factorial:* In [Andrew et al, 2018](https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002556) researchers randomized towns into four groups. The first received the psychosocial stimulation **only** (PS), the second received multiple micronutrient supplementation **only** (MN), the third received **both**interventions (PS and MN), and the fourth received neither (Control). The response to this field is *factorial* design because one arm receives the combination intervention PS + MN.  *Crossover:* In [Lopez et al, 2022](https://pubs.aeaweb.org/doi/pdfplus/10.1257/app.20190722) authors vary in which days doctors received a doctor-specific intervention, and in which days patients received a patient-specific intervention. As doctors (and patients) can cross between the control and treatment groups, the response to this field is *crossover* design [see: Data collection, figure 2].  In [Miguel and Kremer, 2004 authors vary the timing in which three groups of randomly selected schools receive](https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1468-0262.2004.00481.x) school-based deworming. As the control group crosses over into the treatment group by the end of the study, the response to this field is *crossover*. |

### Number of interventions

| **Variable name in SurveyCTO** | [intNum] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Number of interventions |
| **Definition (LABEL on SurveyCTO)** | Number of distinct interventions in the study |
| **Response options (open-text, numeric, date, text-CV)** | Numeric entry |
| **CV (Choices on SurveyCTO)** | None |
| SurveyCTO instructions for data entry mask and repeat level | If “Crossover” is selected for “Intervention assignment strategy”, display a warning message after the current field.  “Reminder: Please make sure you have counted and included the timing of interventions in the number of interventions?” |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Count the distinct interventions in the paper under evaluation.  - Some arms receive a package of interventions, unless any intervention in the package is evaluated separately, do not split the package of interventions into separate interventions.  - If a common type of intervention is assigned in varying intensity to different study arms, count each intensity level as a unique intervention.  - In crossover design, count each roll-out timing as a different intervention.   - Include any intervention beyond the status quo administered to a group designated as a main control or comparison group.  - This information is mostly found in the experimental/study design or methods sections of a paper. If there are study participant flow diagrams, these may illuminate the distinct interventions administered to treatment groups. |
| **Descriptive example see the[section]used in the paper to extract** | [Barrera-Osorio et al., 2022](https://www.google.com/url?q=https://doi.org/10.1016/j.jdeveco.2022.102820&sa=D&source=editors&ust=1741978839857373&usg=AOvVaw2VpvARWP3EzhnzQWeauG_d) have two distinct interventions. "Out of this sample, 140 schools were randomly assigned to each of the two treatment arms – recognition or in-kind performance rewards – and 140 schools were randomly assigned to the control." For this field, the response would be 2 interventions. [see: Intervention, sample and experimental design]  [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/abs/pii/S0304387818303808%23sec2&sa=D&source=editors&ust=1741978839857525&usg=AOvVaw2w2P_twdxQy1d0fwAa14CX)have four separate arms and four unique intervention components. Out of the four interventions, one was common to all arms though not part of the status quo in child care centers outside the study sample. The response for this field would be 4 interventions: 1. Learning materials and supplies, 2. teacher training and mentoring, 3. teacher incentives, and 4. parenting training. [see: Interventions]  [Egger et al., 2022](https://www.google.com/url?q=https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA17945&sa=D&source=editors&ust=1741978839857648&usg=AOvVaw34YVzpfr1zHPn3UuVzwREd) The two interventions are "cash transfer, high saturation" and "cash transfer, low saturation". The cash transfer provided to the household is the same in both interventions, but in one arm a larger share of households receive the transfer, so the intensity of treatment at the village level is different. This means there are two different (village level) interventions.". [see: Figure A1]  [Miguel and Kremer (2004)](http://emiguel.econ.berkeley.edu/wordpress/wp-content/uploads/2021/03/Paper__Worms.pdf) have a crossover design, in which the deworming treatment is phased in to different schools in different years: “Group 1 schools received free deworming treatment in both 1998 and 1999, Group 2 schools in 1999, while Group 3 schools began receiving treatment in 2001” (page 165). There are three “interventions” in this study, one is the program intervention: “free deworming treatment”. The other two are “timing interventions” specific to the crossover design: treatment in 1998 and treatment in 1999.  Note that “treatment in 2001” happened after the study period, so should not be included in the interventions. |

### Intervention label

| Variable name in SurveyCTO | [intLabel] |
| --- | --- |
| Field name (LABEL on SurveyCTO) | Intervention label |
| Definition (LABEL on SurveyCTO) | A short, author used label for the intervention |
| Response options (open-text, numeric, date, text-CV) | Open-text |
| CV (Choices on SurveyCTO) | None |
| SurveyCTO instructions for data entry mask and repeat level | Repeat for each intervention |
| Coding instructions for coders (Hint on SurveyCTO) | - Use the short name of the intervention verbatim as described in the paper and used by the author in tables and figures.  - The intervention label will be used later to form the names of the study arms and map the unit of randomization and stratification variables to each intervention.  -Information to derive the name is mostly found in the experimental/study design or methods sections of a paper. The coder may also consult how an intervention is described in tables (for example, if a single intervention has been assigned to a treatment group, then authors may use a brief name to describe this intervention in tables presenting treatment effects). Labels for treatment groups in participant flow diagrams may also provide starting points for brief names for interventions. |
| Descriptive example see the [section] used in the paper to extract | [Barrera-Osorio et al., 2022](https://doi.org/10.1016/j.jdeveco.2022.102820) have two interventions. Section 2.1 (Performance-based reward program, p2) states “Rewards took the form of either goods (in-kind) or recognition, depending on the treatment arm to which the teacher’s school was assigned. The value of the reward was determined on an absolute scale, without relative performance comparisons to other teachers.”  One intervention is “Public recognition of high-performing teachers”, and the other one is “In-kind reward for high-performing teachers”.  [Ozler et al, 2018](https://www.sciencedirect.com/science/article/abs/pii/S0304387818303808#sec2) In this study, section 2.3 (Interventions, p451) describes the interventions. The 4 distinct interventions are – (1) “Provision of play and learning materials (intervention common to all arms)”, (2) “Training and mentoring of teachers”, (3) “Teacher incentives”, and (4) “Parenting training”. |

### The total number of study arms including control

| **Variable name in SurveyCTO** | [armNum] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | The total number of study arms including control |
| **Definition (LABEL on SurveyCTO)** | The number of study arms, i.e., subgroups of participants that receive none, one, or several specific interventions |
| **Response options (open-text, numeric, date, text-CV)** | Numeric |
| **CV (Choices on SurveyCTO)** | None |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Enter the total number of study arms created by the randomized assignment of interventions. A study arm is a subgroup of experimental units that receive the same (set of) interventions. Include the control group(s) in the total number of study arms.  - In a factorial design, include groups of experimental units that received more than one intervention as separate arms. That is, if there are two interventions (A & B) and participants are assigned to either A alone, B alone, the combination of A & B, or a control group, then this would count as 4 study arms.  - In a crossover design, each arm should include at least one intervention and a timing indicated in the intervention field. For example, intervention A implemented at timing X should be counted as one arm while intervention A implemented at timing Y is another arm.   - This information is mostly found in the intervention details, randomization or methods section of the paper. If participant flow diagrams are available, please consult them to see how many arms are in the study. Tables that present the treatment effects can illuminate the different arms as well. |
| **Descriptive example see the[section]used in the paper to extract** | [Barrera-Osorio et al., 2022](https://www.google.com/url?q=https://doi.org/10.1016/j.jdeveco.2022.102820&sa=D&source=editors&ust=1741978839860678&usg=AOvVaw3Zk3HFZ_-s9ZhQJ6VSAkhP) . "Out of this sample, 140 schools were randomly assigned to each of the two treatment arms – recognition or in-kind performance rewards – and 140 schools were randomly assigned to the control." There are 3 treatment arms in this study: 2 treatment arms and 1 control arm.  [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/abs/pii/S0304387818303808%23sec2&sa=D&source=editors&ust=1741978839860849&usg=AOvVaw1HK2U98eQHhGk4QcdnvFyG) has four separate arms and different subsets of 3 interventions are assigned to treatment arms. In this study, there are four arms -*T1*. Comparison Group: Provision of play and learning materials, *T2*. T1 + Training and mentoring of teacher, *T3*. T2 + Teacher incentives, *T4*. T2 + Parenting training [see: Interventions]. The response for this field is 4.  [Miguel and Kremer (2004)](http://emiguel.econ.berkeley.edu/wordpress/wp-content/uploads/2021/03/Paper__Worms.pdf) has three different arms: one arm that receives the deworming treatment in 1998 and 1999, one arm that receives the deworming treatment in 1999, and one comparison arm that receives the treatment after the data is collected.. The response for this field is 3. |

### Mapping interventions to arms

| **Variable name in SurveyCTO** | [armMap] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Mapping interventions to arms |
| **Definition (LABEL on SurveyCTO)** | Interventions received by each study arms |
| **Response options (open-text, numeric, date, text-CV)** | text-CV, select all |
| **CV  (Choices on SurveyCTO)** | List of interventions in field: [Descriptive name of intervention]  Add an option “None” at the beginning of the list |
| SurveyCTO instructions for data entry mask and repeat level | Repeat for each arm |
| **Coding instructions for coders (Hint on SurveyCTO)** | - For each arm, select the name of the intervention that is assigned from the drop-down list, starting with the control arm(s).  -If the control arm received no intervention (status quo), choose "None".  - For arms that receive a combination of 2 interventions (including timing), select both interventions.  - This information is mostly found in the intervention details, randomization or methods section of the paper. If participant flow diagrams are available, please consult them to see all the arms in the study. |
| **Descriptive example see the[section]used in the paper to extract** | [Barrera-Osorio et al., 2022.](https://www.google.com/url?q=https://doi.org/10.1016/j.jdeveco.2022.102820&sa=D&source=editors&ust=1741978839861822&usg=AOvVaw29WAZ6NZ_1NSKMVKn6PbPw) "Out of this sample, 140 schools were randomly assigned to each of the two treatment arms – recognition or in-kind performance rewards – and 140 schools were randomly assigned to the control. " So for this field: First, select "None" Next, select "Public recognition of teacher high performance" Finally, select "In-kind rewards for teacher high performance"  [Knauer et al. (2020)](https://www.sciencedirect.com/science/article/pii/S0885200619300031) feature a factorial design in which each successive arm receives an additional intervention or two than the other arms.  So, for this field:  First, for arm 1, the coder would select “None.”  For arm 2, the coder would select “Storybooks.”  For arm 3, the coder would select “Storybooks + DRT + SMS”  For arm 4, the coder would select “Storybooks + DRT + SMS + Booster Training”  And finally for arm 5 the coder would select “Storybooks + DRT + SMS + Booster Training + Home Visits.”  [Miguel and Kremer (2004)](http://emiguel.econ.berkeley.edu/wordpress/wp-content/uploads/2021/03/Paper__Worms.pdf) have a crossover design, in which the deworming treatment is phased in to different schools in different years. For this field:  For arm 1, the coder would select “Free deworming treatment” + “treatment in 1998” + “treatment in 1999”  For arm 2, the coder would select “Free deworming treatment” + “treatment in 1999”  For arm 3, the coder would select “None” |

### Unit of randomization

| **Variable name in SurveyCTO** | [unitRand] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Unit of randomization |
| **Definition (LABEL on SurveyCTO)** | The unit of treatment assignment, also .referred to as the level of clustering. |
| **Response options (open-text, numeric, date, text-CV)** | text-CV |
| **CV (Choices on SurveyCTO)** | **Select all that apply** See CV A in tab Long CV |
| SurveyCTO instructions for data entry mask and repeat level | If individual, organization or family selected, add a follow up with the sub options from tab long CV  If Other is selected, textbox follow up with the name of unit  [If only one variable is reported, skip the mapping of unit of randomization to interventions] |
| **Coding instructions for coders (Hint on SurveyCTO)** | Select the unit(s) of randomization for the study. If the unit of randomization differs across interventions, select all units of randomization used.  - Unit of randomization indicates the level at which the assignment of the intervention to study arms was done and at which the intervention is delivered.  - If the selected units are "individual", "organization", "family" or "geographic unit", please also choose the subtype of unit in the follow-up field.  - Geographic units include all spatial units or physical locations.  -Select ‘other’ and fill out the textbox if the unit of randomization does not exactly match the pre-specified units in the CV.  -This information is mostly found in the randomization or methods sections of a paper. The text should indicate that the unit was used to assign the treatment. If there are study participant flow diagrams, these may illuminate the units to which treatment was allocated. |
| **Descriptive example see the[section]used in the paper to extract** | [Ozler et al, 2018](https://doi.org/10.1016/j.jdeveco.2018.04.004) is a cluster randomized trial, in which Community-Based Childcare Centers (CBCCs) were randomized into control, where the children received a learning kit, or the three treatment arms in which children also received learning kits and a combination of different interventions. The unit of randomization is the CBCC since that is the level at which any of the treatments were allocated [see: Study design and sample selection].   There is only one unit of randomization in this experiment. The response for this field would be "Other". In the follow-up field prompting the specific answer, the response would be "Childcare Center".  [Guiteras et al, 201](https://www.science.org/doi/10.1126/science.aaa0491)4 is a cluster randomized trial, in which communities were first randomized to receive a community motivation and health information campaign, or an information campaign combined with subsidies for the purchase of hygienic latrines, or a supply-side market access intervention linking villagers with suppliers and providing information on latrine quality and availability, or no interventions. Second, within the subsidy communities, eligible households were randomized to receive subsidy vouchers through household-level lotteries.   There are two units of randomization in this experiment: community and household. The response to this field would be (selecting both) "Geographic unit" and "Household", then select "Community" for the follow-up field containing the sub-type for "Geographic unit". There is no follow-up field for the choice of "Household". |

### Mapping units of randomization to interventions

| **Variable name in SurveyCTO** | [unitRandMap] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Mapping units of randomization to interventions |
| **Definition (LABEL on SurveyCTO)** | The unit of randomization for each intervention |
| **Response options (open-text, numeric, date, text-CV)** | Text-CV, select one |
| **CV (Choices on SurveyCTO)** | List of units selected in [Unit of randomization] |
| SurveyCTO instruction for data entry mask and repeat level | Skip if only one variable is selected in field [unit of randomization] |
| **Coding instructions for coders (Hint on SurveyCTO)** | If there is only one unit of randomization used for treatment assignment, this question will be skipped. When there is more than one unit of randomization (e.g. schools are assigned to a teacher training program and then families within schools are assigned to a parental support program), each unit will have to be mapped to an intervention.   This information is mostly found in the randomization or methods sections of a paper or in a study participant flow diagram. |
| **Descriptive example see the[section]used in the paper to extract** | In [Leaver et al, 2021](https://www.aeaweb.org/articles?id=10.1257/aer.20191972), there are 5 interventions. The units of randomization in the study are -- district-subject-family and schools. Here, since there is more than one unit of randomization, we need to map each intervention to its unit of randomization  Advertisement of fixed-wage contract - unit of randomization: District-by-subject-family pairs Advertisement of pay for performance contract - unit of randomization: District-by-subject-family pairs Advertisement of both fixed-wage and pay for performance contracts - unit of randomization: District-by-subject-family pairs Implementation of pay for performance contract - unit of randomization: School Implementation of fixed-wage contract - unit of randomization: School |

### Randomization method

| **Variable name in SurveyCTO** | [randmethod] |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Method used for randomization at the experiment level. |
| **Definition**  **(LABEL on SurveyCTO)** | Indicates the method used in the experiment to randomly assign randomization units to study arms. |
| **Response options (open-text, numeric, date, text-CV)** | Text CV, select one |
| **CV**  **(Choices on SurveyCTO)** | -Simple random assignment or fixed-share random assignment (without blocking)  -Block randomization/permutation blocking (without strata)  -Stratification (blocked or other)  -Matched pair design  -Re-randomization  -Covariate adaptive randomization  -Other method or not specified |
| SurveyCTO instructions for data entry mask and repeat level |  |
| **eCoding instructions for coders**  **(Hint on SurveyCTO)** | - Indicate the randomization technique used to assign any interventions to experimental units in the study.  Instead of simple randomization, most researchers use methods to balance the number of randomization units across study arms (blocking or other methods of fixed-share assignment); stratification (including matched-pair designs), re-randomization, and covariate adaptive sampling additionally balance subgroups of the sample (strata) across arms.  -**Simple randomization** **or fixed-share random assignment.** Simple randomization refers to assigning each randomization unit to one of the treatment arms with a fixed probability, e.g. choosing the treatment arm for each unit with a coin flip or by drawing a ball (with replacement) from a 50:50 urn. With simple randomization, the size of the treatment arms can randomly vary.  Fixed-share random assignmentrefers to randomly assigning a fixed share of the sample to each treatment arm, e.g. by randomly selecting one half of the sample for treatment or by drawing each unit’s assignment “without replacement” from a 50:50 urn.  Use this option for any fixed-share assignment method other than blocking.  Sometimes, the paper may not be clear about which of the two methods was used exactly, e.g. we randomly assigned 48 units to treatment and 49 units to control. In that case, you should select this option unless other randomization methods listed below were used.  -**Block randomization or permutation blocking** is a method of fixed-share random assignment in which the sample is divided into “blocks”, and each block is (randomly) assigned a treatment arm permutation. For example, with blocks of two units, the randomly assigned permutations are either C/T or T/C (but not C/C or T/T). This method is particularly suitable if units arrive sequentially. The description of the randomization procedure will use phrases such as “blocked assignment” or “randomization within blocks.” Use this option if blocking is used for fixed-share assignment, but there is no stratification on covariates. When block randomization is used, please do not select “Simple randomization or fixed-share random assignment (without blocking)”.  -**Stratification (blocked or other)** covers all stratification methods other than pure matched pair designs. If stratification is used in any part of randomization, select this option regardless of other methods used.  Stratification serves to balance specific characteristics of the sample (covariates) across treatment arms. The sample is divided into subgroups (“strata”) and fixed-share assignment are used within each stratum (the exact method may not be explicitly mentioned).  In most papers, stratification will appear in the description of the randomization procedure and will use phrases such as "stratified by..” or “balanced the assignment by..” (e.g. “by gender”) or "randomly assigned *within”* (e.g. “half of the households within each village were randomly chosen to receive…”). When blocked stratified randomization is used, the strata may themselves bethe blocks (e.g. “block-randomized 5 sites per region...”).  Note that stratification is sometimes also used in sampling, to ensure that each stratum is represented in the study Mark stratification here ONLY if the paper uses stratified randomization.  Also note that, when randomization is done in stages, the second stage randomization may be stratified by the randomization results of the first stage. Please select this option if this is the case.    -**Matched pair design** is a specific type of stratified randomization. Here, randomization units are grouped into pairs (“matched”) based on baseline characteristics, and within each pair, one unit is randomly assigned to the treatment and the other to the control. The term is typically explicitly used, but if the researchers describe the design as stratification where all strata are of size 2, please choose this option. If pair matching is used in addition to other stratification methods, please use “Stratification (blocked or other)”.  -**Covariate adaptive randomization** is a method for balancing the sample across arms that applies if units of randomization arrive sequentially, such as patients at a clinic. It varies the assignment probability of each unit based on the covariate distribution among units that arrived previously. It typically attempts to minimize any covariate imbalance in random assignments.  -**Re-randomization** is another method to achieve balance. It consists of repeating a specific randomization procedure (e.g. simple or fixed-share random assignment) and selecting the “most balanced” assignment. The authors may conduct a pre-specified number of randomizations or re-randomize until a pre-specified balance criterion is met. Re-randomization is sometimes combined with stratification. In this case, select “Stratification (blocked or other)”.  -**Other or not specified** Please use this category if the exact method of assignment is not specified, or if another assignment method is used, for example, adaptive sampling with objectives other than covariate balancing. |
| **Descriptive example**  **see the[section]used in the paper to extract** | In [Andrew et al, 2018](https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002556) "Randomization was done at the level of a town and after a stratification by region" [Randomization and masking]. This indicates that towns were randomized in equal proportion to be assigned treatment or control arms  Therefore, the response for this field would be "Stratification (blocked or other)".  In [Stewart et al. (2020)](https://pubmed.ncbi.nlm.nih.gov/32006028/), the authors report that they “block-randomized 5 sites per treatment arm per region” into the five study arms. The correct answer is “Stratification (blocked or other)”. Note that here the strata (the regions) are used as blocks. |



### Number of stratification variables

| **Variable name in SurveyCTO** | [strataNum] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Number of stratification variables |
| **Definition (LABEL on SurveyCTO)** | Number of variables used to create strata |
| **Response options (open-text, numeric, date, text-CV)** | Numeric entry |
| **CV (Choices on SurveyCTO)** |  |
| **SurveyCTO instructions for data entry mask and repeat level** | Display this question if answer to “Randomization method” is “Stratified randomization” or “Stratified block randomization (including matched pair designs)”. |
| **Coding instructions for coders (Hint on SurveyCTO)** | Please indicate the number of variables used to create the strata |
| **Descriptive example see the[section]used in the paper to extract** | In [Berman et al. (2019)](https://www.sciencedirect.com/science/article/pii/S016726811930318X), the authors note that they stratify by province, share of respondents in the baseline survey that report at least occasional access to electricity, and the share of respondents reporting that the district governor carries the most responsibility for keeping elections fair. The correct answer here would be 3.  In [Dupas (2011),](https://www.aeaweb.org/articles?id=10.1257/app.3.1.1) the author notes that the randomization procedure is stratified by teacher training status. The correct answer here would be 1. |

### Stratification variables

| **Variable name in SurveyCTO** | [strataLabel] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Stratification variables |
| **Definition (LABEL on SurveyCTO)** | The variables used to create the strata |
| **Response options (open-text, numeric, date, text-CV)** | Open-text |
| **CV (Choices on SurveyCTO)** | None |
| SurveyCTO instructions for data entry mask and repeat level | -Display this question if answer to “Randomization method” is “Stratified randomization” or “Stratified block randomization (including matched pair designs)”.  -Use the “number of stratification variables” to repeat this question. |
| **Coding instructions for coders (Hint on SurveyCTO)** | - List each of the stratification variables out in the order in which they appear in the paper.  - Use indicators for being assigned to an intervention (e.g. indicator for being assigned to provider incentive) if randomization at a lower level is stratified by groups generated by higher-level randomization of one or more other interventions.   - If local terms are used to define the strata - for example, woreda, oblast, union, or taluka - please retain the local term rather than using the author's translation of them into English.   This information is mostly found in the randomization or methods section of the paper. Phrases such as "stratified by...", "randomly assigned within..." are common when describing stratification variables. |
| **Descriptive example see the[section]used in the paper to extract** | Since the response to 10 for [Andrew et al, 2018](https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002556) is Strata. Using text from the paper "Randomisation was done at the level of the cluster (town), after stratification by region. Within each of the 3 regions, 8 towns were randomly allocated to each of the 3 treatment groups and the control group using computer-generated random numbers" [Randomization and masking].  The response to this field would be "region" since that is the variable that makes up the strata as indicated in the paper.  [In Freeman et al. (2022)](https://journals.plos.org/globalpublichealth/article?id=10.1371/journal.pgph.0000056), authors use a stratified randomization design so the response to the previous question asking if the intervention was stratified is "Strata". The authors note "a stratified random design at the woreda-level was used to assign an equal number of study kebeles to either the Andilaye intervention or the control group receiving no intervention".   The response to this question is "Woreda", as it is the local term used consistently throughout the paper. |

### Stratification for study arms

| **Variable name in SurveyCTO** | [strataSame] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Stratification for study arms |
| **Definition (LABEL on SurveyCTO)** | Are the same stratification variables used for all interventions |
| **Response options (open-text, numeric, date, text-CV)** | text-CV |
| **CV (Choices on SurveyCTO)** | **Select one** Yes No |
| **SurveyCTO instructions for data entry mask and repeat level** | -Display this question if answer to “Randomization method” is “Stratified randomization” or “Stratified block randomization (including matched pair designs)”. |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Indicate whether the same (set of) stratification variable is used for assigning all interventions to study arms.   - Note for two stage randomizations, the treatment arm assignment for first stage can be the basis for stratification for the second stage.  - This is found in sections describing randomization or methods section. A flow diagram depicting the experiment may also contain helpful information for this field. |
| **Descriptive example see the[section]used in the paper to extract** | In [Ozler et al, 2018](https://doi.org/10.1016/j.jdeveco.2018.04.004), a "block randomization" was used to assign childcare centers in each district to the four study arms. "Centers were grouped based on mean height-for-age (HAZ) and Peabody Picture Vocabulary Test (PPVT - a measure of receptive vocabulary) z-scores, both of which were collected at baseline. The Ministry held a public lottery at each district capital where a representative from each center was asked to draw a colored dot from an envelope to determine that center's treatment status."   The response to this question would be "Yes", as the same set of variables were used for stratification.   In [Wolf et al. 2019](https://doi.org/10.1080/19345747.2018.1517199), there are two stages of randomization. In the first stage, the intervention assignment was stratified by district and public/private status of the school. In the second stage, the interventions were assigned within groups created by treatment assignment in the first stage. The stratification variable is indicator for teacher training/parental awareness assignment.   The response to this question would be "No", since the stratification variables are different across interventions. |

### Mapping stratification variables to interventions

| **Variable name in SurveyCTO** | [strataMap] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Mapping stratification variables to interventions |
| **Definition (LABEL on SurveyCTO)** | The list of variables used to create the randomization strata for each intervention |
| **Response options (open-text, numeric, date, text-CV)** | Text-CV, select all |
| **CV (Choices on SurveyCTO)** | Stratification variables listed in [Stratification variables] |
| **SurveyCTO instructions for data entry mask and repeat level** | -Display this question if answer to “Stratification for study arms” is “No”. |
| **Coding instructions for coders (Hint on SurveyCTO)** | -If the same (set of) stratification variable is used for all treatment assignments, then this question will be skipped.   - This information is mostly found in the randomization and methods sections of the paper. For most papers, this information will appear in the same paragraph as the one describing the stratification procedure. |
| **Descriptive example see the[section]used in the paper to extract** | [Wolf et al, 2018](https://www.tandfonline.com/doi/epdf/10.1080/19345747.2018.1517199?needAccess=true) has 4 interventions and 5 study arms:  Teacher training and coaching program Parental awareness meetings Text messages for teachers Picture-based paper flyers or texts for parents  Based on the information in the section "Randomization", in a first stage of randomization, three of the interventions (teacher training and coaching program; and parental awareness meetings) were stratified by district and public/private status of the school. The text messages for teachers and the texts/flyers for parents were assigned in a second stage of randomization. These interventions were stratified by treatment assignment in the previous stage so that stratification variables for this assignment of interventions to arms would be indicators for being assigned to the two of the arms.   So for this field, the stratification mapping for 5 study arms would be: Control - district and public/private status of the school Teacher training and coaching program - district and public/private status of the school Parental awareness meetings - district and public/private status of the school Text messages for teachers - indicator for school being part of teacher training and coaching program Picture-based paper flyers for parents - indicator for school being part of program with teacher training & coaching and parental awareness meetings |

### Description of randomization method

| **Variable name in SurveyCTO** | [randDescrip] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Description of randomization method. |
| **Definition (LABEL on SurveyCTO)** | Description of the randomization method used in the experiment. |
| **Response options (open-text, numeric, date, text-CV)** | Open-text |
| **CV (Choices on SurveyCTO)** |  |
| **SurveyCTO instructions for data entry mask and repeat level** | -Display this question if answer to “Randomization method” is “Other”. |
| **Coding instructions for coders (Hint on SurveyCTO)** | -Please provide a description of the randomization method and process used in the experiment. Please extract the information verbatim from the paper. |
| **Descriptive example see the[section]used in the paper to extract** |  |

### Number of units of analysis in the experiment

| **Variable name in SurveyCTO** | unitAnaNum |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Number of units of analysis in the experiment |
| **Definition**  **(LABEL on SurveyCTO)** | Number of units of analysis at which the treatment effects are estimated. |
| **Response options (open-text, numeric, date, text-CV)** | Numeric entry |
| **CV (Choices on SurveyCTO)** | Integer greater than 0 |
| **SurveyCTO instructions for data entry mask and repeat level** |  |
| **Coding instructions for coders**  **(Hint on SurveyCTO)** | - Count the number of unique units of analysis at which treatment effects are estimated in the experiment.  - Multiple outcomes and treatment effects can be estimated at the same unit of analysis. The same unit of analysis should only be counted once.  - If treatment effects for the unit of analysis are only reported in an appendix or supplementary materials, please do not count them for the purposes of this field.  - For heterogeneous treatment effects (which meet the criteria to be included in IDEAL), please include the corresponding units of analysis in the count. |
| **Descriptive example**  **see the[section]used in the paper to extract** | [Ashraf et al., 2010](https://www.jstor.org/stable/41038767) include treatment effects for the full sample in Tables 2, 3, 4 and 5.  The unit of analysis in Table 2 is “Household” for the outcome “Household purchased Clorin (dummy)”.  The unit of analysis in Table 3 is also “Household” for the two outcomes: “Water currently treated with Clorin” and “Drinking water contains free Clorin”.  The two outcomes from Table 3 are also in Table 4.  The unit of analysis in Table 5 is “Household” for two outcomes: “Bottle exhausted?” and “Use Clorin for non-drinking water purposes”.  Therefore, there is only ONE (1) unit of analysis in this experiment. |

### Unit of analysis variable

| **Variable name in SurveyCTO** | unitAnaLabel |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Unit of analysis variable |
| **Definition**  **(LABEL on SurveyCTO)** | Unit of analysis at which the treatment effects are estimated. |
| **Response options (open-text, numeric, date, text-CV)** | Open-text |
| **CV**  **(Choices on SurveyCTO)** |  |
| **SurveyCTO instructions for data entry mask and repeat level** | -Use the answer to “Number of units of analysis in the experiment” to repeat this question.  -This field along with the next field (“Unit of analysis category”) should be grouped together and repeated for each unit of analysis. For example, if a coder reports 3 units of analysis. The questions should be asking 1) unit of analysis variable, and 2) unit of analysis category, and then repeat for units of analysis 2-3. |
| **Coding instructions for coders**  **(Hint on SurveyCTO)** | - Enter the unit of analysis variable as it is in the paper. For example, for pregnant women visiting health facilities, the unit of analysis might be referred as woman or patient in different papers, please write down the exact unit used in the paper.   - This information can be found in the results, data, and table notes in the paper, or in the supplementary materials. |
| **Descriptive example**  **see the[section]used in the paper to extract** | For example, Ozler et al. 2018 include treatment effects estimated using various units of analysis. The treatment effects on child assessments and behavioral problems (Tables 3&4) were estimated at the child level, so the unit of analysis of those outcomes is child.  Table 5 includes impacts on parenting quality, for which the unit of analysis is “primary caregiver” (see section 2.4.2, page 453).  For impacts on CBCC outcomes in Table 6, the unit of analysis is Community-Based Childcare Center (CBCC). |

### Unit of analysis category

| **Variable name in SurveyCTO** | unitAnaCV |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Outcome unit of analysis category |
| **Definition**  **(LABEL on SurveyCTO)** | Category of unit of analysis at which treatment effects are estimated. |
| **Response options (open-text, numeric, date, text-CV)** | Text-CV, select one |
| **CV (Choices on SurveyCTO)** | See **CV A** in appendix B of Cavanagh et al (2023) at [link](https://documents1.worldbank.org/curated/en/099945502062327217/pdf/IDU081c960a8049b504197099ff0d12be0b95375.pdf) |
| **SurveyCTO instructions for data entry mask and repeat level** | -Use the answer to “Number of units of analysis in the experiment” to repeat this question. |
| **Coding instructions for coders**  **(Hint on SurveyCTO)** | - Select the unit of analysis for the treatment effect using the controlled vocabulary (CV).  - This information can be found in the results, data, and table notes in the paper, or in the supplementary materials. |
| **Descriptive example**  **see the[section]used in the paper to extract** | For example, Ozler et al. 2018 include treatment effects estimated using various units of analysis. The treatment effects on child assessments and behavioral problems (Tables 3&4) were estimated at the individual child level, so the unit of analysis of those outcomes is child.  From the pre-specified list of options, a coder would first choose “1. Individual” as the broad category and then choose “Child” as the category.  Table 5 includes impacts on parenting quality, for which the unit of analysis is primary caregiver. Similarly, a coder would first choose “1. Individual” as the broad category and then choose “1.11 Parent” as the unit of analysis category.  For impacts on CBCC outcomes in Table 6, the unit of analysis is Community-Based Childcare Center (CBCC). A coder would first choose “2. Organization or legal entity” and then “2.9 Other organization or legal entity” to type “Childcare center”. |

### Number of outcome variables in the study

| **Variable name in SurveyCTO** | [outNum] |
| --- | --- |
| **Field name** **(LABEL on SurveyCTO)** | Number of outcome variables in the study |
| **Definition** **(LABEL on SurveyCTO)** | Number of outcomes variables reported in the study for which treatment effects are estimated |
| **Response options (open-text, numeric, date, text-CV)** | Numeric entry |
| **CV** **(Choices on SurveyCTO)** | None |
| **SurveyCTO instructions for data entry mask and repeat level** |  |
| **Coding instructions for coders** **(Hint on SurveyCTO)** | - Enter all the outcome variables reported in the study for which treatment effects are estimated.  - If the outcome variable is an index and the treatment effects of the components are reported in the same exhibit, include both the index and all the components as separate outcomes. If treatment effects for the index components are only reported in an appendix or supplementary materials, do not count these as outcomes for the purposes of this field.  - Include outcome variables that are reported in an exhibit, even if the exact point estimates can only be found in supplementary materials.  - For a table with heterogeneous treatment effects (which meets the criteria to be included), if the same outcome is reported for multiple subgroups, please count them as different outcomes. For example, outcome A – female and outcome A – male should be counted as two outcomes.  - Do not include auxiliary outcomes that do not appear in the exhibit (but note that this does not include primary outcomes that were only moved to an appendix because they do not show an effect).  - Outcomes that are only measured as marginal effects should also be included. |
| **Descriptive example** **see the[section]used in the paper to extract** | In [Guiteras et al, 2015](https://www.google.com/url?q=https://www.science.org/doi/10.1126/science.aaa0491%23supplementary-materials&sa=D&source=editors&ust=1741978839872973&usg=AOvVaw22asPlBldVKJtkx07z2nfE), the published manuscript does not show any table in the main paper. Figures 1 & 2 report the treatment effects, however, we can not obtain the precise statistics such as point estimates and standard errors directly from the figures. From the notes of Figure 1, "Figure displays the sum of the estimated coefficients and the control group means found in columns (2) and (6) of table S2 and column (2) of table S3. (A) Any latrine access; (B) hygienic latrine access; (C) open defecation among adults", we learn that the estimated coefficients can be found in tables S2 and S3 in the supplementary materials.  Figure 1 includes three outcomes: "(A) Any latrine access; (B) hygienic latrine access; (C) open defecation among adults".  Figure 2 includes three outcomes: "(A) Any latrine ownership; (B) hygienic latrine ownership; (C) open defecation among adults."  [Ashraf et al., 2010](https://www.google.com/url?q=https://www.jstor.org/stable/41038767&sa=D&source=editors&ust=1741978839873091&usg=AOvVaw2bS-6pN9JMcbNOQnVKkiWl) include treatment effects for the full sample in Tables 2, 3, 4 and 5.  The outcome in Table 2 is “Household purchased Clorin (dummy)”. There is only one outcome.  There are two outcomes in Table 3. They are “Water currently treated with Clorin” and “Drinking water contains free Clorin”.  The two outcomes from Table 3 are also in Table 4.  The title of Table 5 includes Heterogeneity, however, the table reports the full sample estimates for two outcomes: “Bottle exhausted?” and “Use Clorin for non-drinking water purposes”. Note that only full-sample treatment effects and their outcomes should be included. |

### Outcome name

| **Variable name in SurveyCTO** | [outLabel] |
| --- | --- |
| **Field name** **(LABEL on SurveyCTO)** | Outcome name |
| **Definition** **(LABEL on SurveyCTO)** | The name used in the exhibit to refer to the outcome measure. |
| **Response options (open-text, numeric, date, text-CV)** | Select one |
| **CV** **(Choices on SurveyCTO)** | None |
| **SurveyCTO instructions for data entry mask and repeat level** | -Use the answer to “Number of outcome variables in the table” to repeat this question.  -This field along with the next two (i.e. “outcome unit of analysis” and “outcome unit of analysis inclusion criteria”) should be grouped together and repeated for each outcome in the table. For example, if a coder reports 5 outcomes in the table. The questions should be asking 1) outcome name, 2) outcome unit of analysis, and 3) outcome unit of analysis inclusion criteria for Outcome 1, and then repeat for outcomes 2-5. |
| **Coding instructions for coders** **(Hint on SurveyCTO)** | - Enter the complete name of every outcome as it appears in the table. |
| **Descriptive example** **see the[section]used in the paper to extract** | In [Guiteras et al, 2015](https://www.google.com/url?q=https://www.science.org/doi/10.1126/science.aaa0491%23supplementary-materials&sa=D&source=editors&ust=1741978839874201&usg=AOvVaw29qwavFtz1pdrwXeeWJw3k), figures 1 & 2 report the treatment effects of five unique outcomes. The outcome names can be found in the notes below Figure 1 and Figure 2: -Any latrine access, -hygienic latrine access, -open defecation among adults, -any latrine ownership, and -hygienic latrine ownership.   In the [Ashraf et al. (2010)](https://www.google.com/url?q=https://www.jstor.org/stable/41038767&sa=D&source=editors&ust=1741978839874338&usg=AOvVaw1ZIWCLOm-diEyrljnjF7W9) example, the outcome names can be found in Tables 2, 3, 4 and 5, which are: -Household purchased Clorin (dummy) -Water currently treated with Clorin -Drinking water contains free Clorin -Bottle exhausted? -Use Clorin for non-drinking water purposes. |

### Outcome unit of analysis

| **Variable name in SurveyCTO** | [ outUnitSum] |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Outcome unit of analysis |
| **Definition**  **(LABEL on SurveyCTO)** | Unit of analysis for the outcome at which the treatment effect is estimated. |
| **Response options (open-text, numeric, date, text-CV)** | Text-CV, select one |
| **CV (Choices on SurveyCTO)** | The list of answers from “Unit of analysis variable” (fields: unitAnaLabel and unitAnaCV) |
| **SurveyCTO instructions for data entry mask and repeat level** | -Use the answer to “Number of outcome variables in the study” to repeat this question. |
| **Coding instructions for coders**  **(Hint on SurveyCTO)** | - Select the corresponding unit of analysis for the outcome at which the treatment effect is estimated from the list of units of analysis entered earlier. |
| **Descriptive example**  **see the[section]used in the paper to extract** | For example, Ozler et al. 2018 include treatment effects estimated using various units of analysis. The treatment effects on child assessments and behavioral problems (Tables 3&4) were estimated at the individual child level, so the unit of analysis of those outcomes is child.  A coder would choose “Child” for outcomes in Tables (3&4) as their unit of analysis.  Table 5 includes impacts on parenting quality, for which the unit of analysis is primary caregiver. For each outcome in Table 5, select “Primary caregiver” as the unit of analysis. |







### Number of rounds of data collection in the study

| **Variable name in SurveyCTO** | [roundNum] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Number of rounds of data collection in the study |
| **Definition (LABEL on SurveyCTO)** | Number of rounds of data collection including baseline in the study |
| **Response options (open-text, numeric, date, text-CV)** | Numeric entry |
| **CV (Choices on SurveyCTO)** | None |
| **SurveyCTO instructions for data entry mask and repeat level** |  |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Enter the total number of rounds of data collection reported in the paper, including baseline. A round collects data from the same data source at a given time. In other words, a round signals when different outcomes are measured.  - For studies in which data is only collected by surveys of subjects in the sample, the number of rounds typically aligns with the number of surveys (e.g., baseline, midline, endline = 3 rounds).  -If more than one survey sample collected by the authors is used to measure outcomes but they are collected around the same time (e.g. a survey of mothers and a survey of teachers), this should count as the same round.  - However, any administrative data (election results, census data, etc.) that authors use should be considered a separate data source and thus necessitate a separate round. For example, administrative data collected at the endline should be separate from survey data collected at the endline. If a paper includes a baseline survey, midline survey, endline survey, administrative data collected at midline, and administrative data collected at endline, the number of rounds would be 5 rounds.  - For administrative data, sometimes authors do not indicate when they collected the data. In this case, the number of rounds should go by data source (e.g., census data and election data = 2 rounds).  - Do not include the rounds collected for the same study but not used or reported in the paper.  - Do not include any focus group discussions or qualitative surveys here.  This can be found in the "data collection" or "data sources" section for papers. Sometimes, papers have explanatory diagrams on timeline and implementation which coders can consult to get the different points of time in which data was collected. |
| **Descriptive example see the[section]used in the paper to extract** | [The Freeman et al, 2022](https://www.google.com/url?q=https://journals.plos.org/globalpublichealth/article?id%3D10.1371/journal.pgph.0000056%23sec002&sa=D&source=editors&ust=1741978839875405&usg=AOvVaw2UK-XvR_qtpObeJYBkSRvw) study collects household surveys and observation-based data at baseline, midline, and endline. In each round, both survey and observation data were collected at the same time [see: Data collection]. Additionally, Figure 2 illuminates the points in time when each of the data collection rounds was conducted.  The response for this field would be 3.  [Pande & Field, 2008](https://www.google.com/url?q=https://www.jstor.org/stable/40282659?origin%3DJSTOR-pdf&sa=D&source=editors&ust=1741978839875524&usg=AOvVaw2oGRjhEQxO-Rp7rQE5I0xq) only use online endline data on loans and repayment in the current paper.  The response for this field would be 1.  [Muralidharan et al. 2021](https://www.google.com/url?q=https://www.jstor.org/stable/27087097&sa=D&source=editors&ust=1741978839875576&usg=AOvVaw3uqDXhwtjoPVKRehbI94-V) use administrative data from three sources: 1. Register of landlords, 2. a record of check distribution maintained by the MAOs, and 3. bank records of check encashment [see B. data]. Table 1 suggests that the register data was collected between September and December 2017. Appendix C indicates that the authors received "the up-to-date MAO and bank-based datasets at three points in time: once in July, once in August and once in September 2018. Therefore, there are seven rounds of data collection in the study.  The response for this field is 7.  De Hoyos et al. 2021 include the following data:   1. Student assessments: 2013, 2014, 2015 2. Student survey: 2013, 2015 3. Teacher survey: 2013, 2014, 2015 4. Principal survey: 2014, 2015 5. National assessments: 2016 6. Internal efficiency: 2013, 2014, 2015, 2016, 2017   Based on data source and time, there are 3 rounds of survey data, as they are conducted at the same time of year, 4 rounds of assessment data (school and national), and 5 rounds of administrative data on internal efficiency. The total would be 12 rounds of data  Round 1: Student and teacher surveys in 2013  Round 2: Teacher and principal surveys in 2014  Round 3: Student, teacher, and principal surveys in 2015  Round 4: Student assessment 2013  Round 5: Student assessment 2014  Round 6: Student assessment 2015  Round 7: National assessment 2016  Round 8: Internal efficiency 2013  Round 9: Internal efficiency 2014  Round 10: Internal efficiency 2015  Round 11: Internal efficiency 2016  Round 12: Internal efficiency 2017 |

### Round name

| **Variable name in SurveyCTO** | [roundLabel] |
| --- | --- |
| **Field name  (LABEL on SurveyCTO)** | Round name |
| **Definition  (LABEL on SurveyCTO)** | Name for each round of data collection |
| **Response options (open-text, numeric, date, text-CV)** | Open-text |
| **CV (Choices on SurveyCTO)** | None |
| **SurveyCTO instructions for data entry mask and repeat level** | Use the answer to “Number of rounds in the study” to generate the set of repeated questions. |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Label the round of data collection with descriptive names, such as "baseline", "midline", "midline phone survey", "endline", "one-year follow-up", "three-year follow up" etc.   - Retain any descriptive labels used by the study authors, such as "18-month follow up" or "23-month follow up."  - Use the data source to differentiate datasets collected at the same time, for example, phone survey at follow-up and census data at follow-up.  This can be found in the "data collection" or "data sources" section for papers. Sometimes, papers have explanatory diagrams on timeline and implementation which coders can consult to get the different points of time in which data was collected. Tables with treatment effects can also illuminate the different rounds of data collection. |
| **Descriptive example see the[section]used in the paper to extract** | [Freeman et al, 2022](https://www.google.com/url?q=https://journals.plos.org/globalpublichealth/article?id%3D10.1371/journal.pgph.0000056%23sec002&sa=D&source=editors&ust=1741978839876602&usg=AOvVaw22D5DsZ_bObH9V-bfKWkht) study collect household surveys and observations-based data at baseline, midline, and end line. [see: Data collection] Additionally, Figure 2 illuminates the points in time when each of the data collection rounds was conducted.  The response for this field would be: Baseline Midline Endline  [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/pii/S0304387818303808%23sec2&sa=D&source=editors&ust=1741978839876712&usg=AOvVaw1OerWT1RYMDok2A8YYswGU) use three rounds of data collection [see: Data sources] The response for this field would be: Baseline 18-month follow-up 36-month follow-up  For  [Muralidharan et al. 2021](https://www.google.com/url?q=https://www.jstor.org/stable/27087097&sa=D&source=editors&ust=1741978839875576&usg=AOvVaw3uqDXhwtjoPVKRehbI94-V), which draw on three administrative data sources and have seven total rounds, the round names would be:  Register of landlords September – December 2017  Bank-based dataset July 2018  MAO record dataset July 2018  Bank-based dataset August 2018  MAO record dataset August 2018  Bank-based dataset September 2018  Mao record dataset September 2018  [The Freeman et al, 2022](https://www.google.com/url?q=https://journals.plos.org/globalpublichealth/article?id%3D10.1371/journal.pgph.0000056%23sec002&sa=D&source=editors&ust=1741978839875405&usg=AOvVaw2UK-XvR_qtpObeJYBkSRvw) study collects household surveys and observation-based data at baseline, midline, and endline. In each round, both survey and observation data were collected at the same time [see: Data collection]. Additionally, Figure 2 illuminates the points in time when each of the data collection rounds was conducted.  The response for this field would be 3.  [Pande & Field, 2008](https://www.google.com/url?q=https://www.jstor.org/stable/40282659?origin%3DJSTOR-pdf&sa=D&source=editors&ust=1741978839875524&usg=AOvVaw2oGRjhEQxO-Rp7rQE5I0xq) only use online endline data on loans and repayment in the current paper.  De Hoyos et al. 2021 include the following data:   1. Student assessments: 2013, 2014, 2015 2. Student survey: 2013, 2015 3. Teacher survey: 2013, 2014, 2015 4. Principal survey: 2014, 2015 5. National assessments: 2016 6. Internal efficiency: 2013, 2014, 2015, 2016, 2017   Based on data source and time, there are 3 rounds of survey data, as they are conducted at the same time of year, 4 rounds of assessment data (school and national), and 5 rounds of administrative data on internal efficiency. The total would be 12 rounds of data  Round 1: Student and teacher surveys in 2013  Round 2: Teacher and principal surveys in 2014  Round 3: Student, teacher, and principal surveys in 2015  Round 4: Student assessment 2013  Round 5: Student assessment 2014  Round 6: Student assessment 2015  Round 7: National assessment 2016  Round 8: Internal efficiency 2013  Round 9: Internal efficiency 2014  Round 10: Internal efficiency 2015  Round 11: Internal efficiency 2016  Round 12: Internal efficiency 2017 |

### Number of exhibits with treatment effects

| **Variable name in SurveyCTO** | [tableNum] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Number of exhibits |
| **Definition (LABEL on SurveyCTO)** | The number of tables or figures in the paper that report treatment effects estimated using ITT or authors' preferred estimand. |
| **Response options (open-text, numeric, date, text-CV)** | Numeric |
| **CV (Choices on SurveyCTO)** | None |
| **SurveyCTO instructions for data entry mask and repeat level** |  |
| **Coding instructions for coders (Hint on SurveyCTO)** | -Count the tables and figures in the main text that report treatment effects for **the full evaluation sample**, using **ITT or any authors’ preferred estimand**.  In the IDEAL project, the main quantity of interest (or estimand) is the intention-to-treat (ITT) effect using the entire experimental sample. IDEAL also collects treatment effects estimated using authors’ preferred estimand other than ITT.  Full sample refers to the full sample relevant for the outcomes included in the estimation of treatment effects; it is meant to contrast with subsamples created to estimate heterogeneous treatment effects. Note treatment effect estimates are available only for children because the outcome variable was only measured for a sample of children (e.g. stunting or child development), this counts as a full sample estimate. Likewise, if a table only includes estimates for healthcare providers because the outcome variable was measured only for this group (e.g. quality of care), this still counts as a full sample estimate even if other tables are concerned with different populations (e.g. patients). On the other hand, if an estimation is restricted to a certain gender or wealth quintile or any other subsample to demonstrate heterogeneity of treatment effects, this would not be included in IDEAL unless there were no full-sample treatment effects reported in the paper.  -Count figures even if the main text figure does not report exact estimates of treatment effects or their precision but rather only includes this information in an appendix or supplementary materials. However, do not include the exhibit if there are no estimates and precision values that accompany it in the paper or its appendix.  - Count tables that present treatment effects as group means and standard deviations if there is a point estimate and a formal test of treatment effects (e.g., t-test) and precision statistics reported. The treatment effect can be estimated later.  - Count treatment effects only reported in the text but not in any of the exhibits as a pseudo table and label it as “Text only”. When reporting, please group all those treatment effects in one table although they may appear in different parts of the table.   - Do not include tables that report quasi-experimental estimates (e.g. using treatment assignment as an instrumental variable) unless this is the preferred specification of the authors.  - Do not include tables that only report heterogeneous treatment effects for only a subgroup or a subsample, **EXCEPT**   1. The paper only reports heterogeneous treatment effects, OR 2. Heterogeneous treatment effects are the primary research questions.  * Note that if a table only includes estimates for children because the outcome variable was only measured for a sample of children (e.g. stunting or child development), this counts as a full sample estimate. Likewise, if a table only includes estimates for healthcare providers because the outcome variable was measured only for this group (e.g. quality of care), this still counts as a full sample estimate even if other tables are concerned with different populations (e.g. patients). * Do not include tables reporting only robustness checks. |
| **Descriptive example see the[section]used in the paper to extract** | In [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/pii/S0304387818303808%23sec2&sa=D&source=editors&ust=1741978839877854&usg=AOvVaw0r1hlYDqiEDzbkVzvN5JZT), the paper has a total of 13 tables and 2 figures. Of all the exhibits, 11 tables report treatment effects (i.e. Tables 3 though 13). However, Table 7 reports a robustness check and Table 8 reports quasi-experimental results using treatment assignment as an instrumental variable. These two tables should not be included. Thus, this paper has 9 tables with treatment effects for the full sample.  The response for this field is 9.  [Leaver et al. 2011](https://www.google.com/url?q=https://www.aeaweb.org/articles?id%3D10.1257/aer.20191972&sa=D&source=editors&ust=1741978839877995&usg=AOvVaw3JEMMj_HUciNyTtiWVXqO8) include 4 figures and 6 tables. Figure 1 and Tables 1 and 2 report experimental design and baseline characteristics. Figures 2, 3 and 4, and Tables 3, 4 and 5 include treatment effects for the full evaluation sample. Table 6 reports quasi-experimental results. Therefore, there are 6 tables or figures with full sample treatment effects in the paper.  The response for this field is 6.  [Riley 2024](https://www.google.com/url?q=https://benny.aeaweb.org/articles?id%3D10.1257/aer.20220717%26%26from%3Df&sa=D&source=editors&ust=1741978839878079&usg=AOvVaw3kvJDgd348IJqtkV2Oo3bv)has 6 tables and 2 figures. Figures 1 and 2 present take-up and balance. Table 2 only reports heterogeneous effects, and the rest 5 tables include at least one set of treatment effects for the full sample.  The response for this field is 5.  In [Ara et al. 2019](https://pubmed.ncbi.nlm.nih.gov/29660858/), the outcome variables, Median duration of EBF and Median duration of any breastfeeding should be included in Table 2 because formal p-values are reported in the text.  In [Kondylis et al. 2016,](https://drive.google.com/drive/folders/1BpqDshPqgg4mni12CkLFmp2w46k9Nr6V) the authors report treatment effects in Tables 6-10. However, they only report heterogeneous treatment effects by the gender of the farmer. In this case, we would include tables 6-10 for this paper. |

### Exhibit label

| **Variable name in SurveyCTO** | [tableLabel] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Exhibit label |
| **Definition (LABEL on SurveyCTO)** | The label of each table with full sample treatment effects reported in the study |
| **Response options (open-text, numeric, date, text-CV)** | Open-text |
| **CV (Choices on SurveyCTO)** | None |
| **SurveyCTO instructions for data entry mask and repeat level** | Use the “number of exhibits” to generate this repeated group of questions. |
| **Coding instructions for coders (Hint on SurveyCTO)** | - List the table numbers as they appear in the main text that include any full sample treatment effects.  -This is the first of a set of repeated fields and only enter one label at once.  - If the tables include letters or words such as "TABLE 1" or "TABLE 1A", please retain the exact label.  - For treatment effects only reported in the text but not in any exhibit, please use the table “Text only”. All those treatment effects should be grouped in this “Text only” exhibit, no matter where they appear in the paper.   - Do not include the caption of the table such as "18-month follow up impact", "Impact on secondary outcomes" |
| **Descriptive example see the[section]used in the paper to extract** | For [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/pii/S0304387818303808%23sec2&sa=D&source=editors&ust=1741978839879725&usg=AOvVaw3uonZegPBJF8zrGVfNdIAx), the labels are: Table 3, Table 4, Table 5, Table 6, table 7, Table 9, Table 10, Table 11, Table 12, and Table 13. First only enter "Table 3" in this field and answer the questions about Table 3, and repeat the process for each of the rest tables.  For [Leaver et al. 2021](https://www.google.com/url?q=https://www.aeaweb.org/articles?id%3D10.1257/aer.20191972&sa=D&source=editors&ust=1741978839879825&usg=AOvVaw0ynaQ8R2wwApA0M-BAxbqQ), using the order of appearance in the paper, the labels are: Figure 2, Figure 3, Table 3, Figure 4, Table 4, and Table 5. |

# [Stage 1] Table-by-table: For the set of fields below, coder goes through each table identified

#### Number of comparisons

| **Variable name in SurveyCTO** | [tableCompNum] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Number of comparisons |
| **Definition (LABEL on SurveyCTO)** | Number of comparisons or contrasts reported in table/figure for which eligible treatment effects are estimated. |
| **Response options (open-text, numeric, date, text-CV)** | Numeric |
| **CV (Choices on SurveyCTO)** |  |
| **SurveyCTO instructions for data entry mask and repeat level** |  |
| **Coding instructions for coders (Hint on SurveyCTO)** | - List the number of unique comparisons presented in the table. Any reported treatment effect is the result of a comparison of one group against another on an outcome, where a group can be a single study arm or a group of arms.  In a case where there are 4 experimental groups (Treatment A, Treatment B, Treatment A+B, and a Control), many comparisons may be reported. We may see treatment effect for Treatment A relative to the Control, or we may see an estimate for the difference between Treatment A and the combination Treatment A+B. Understanding which comparison is being reported in a table or figure will be important if the row and column labels and table/figure notes do not contain this information.  This information can be present either in the top row or column of a table or it may be reported in the rows as the variables for which treatment effects are being reported. The base arm for comparison is often the omitted category in a regression and/or mentioned in the footnotes of a table. Descriptions of each treatment effect in the results section can also illuminate the contrast under study for each estimate. |
| **Descriptive example see the[section]used in the paper to extract** | In table 3 from [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/pii/S0304387818303808%23sec2&sa=D&source=editors&ust=1741978839881027&usg=AOvVaw1H1Bh-0jNUQhwYecTzPDDu): - We see 3 comparisons in the first panel of rows: 1. T2 (teacher training) vs. Control 2. T3 (T2 + teacher incentives) vs. Control 3. T4 (T2 + parenting training) vs. Control - We see 1 additional comparison in the second panel of rows that pools all treatment groups into 1 group. 4. Any Treatment (T2, T3, or T4) vs. Control - The fourth panel of rows contains the precision but not treatment effects for additional comparisons: (i) T2 vs. T3 (ii) T2 vs. T4 (iii) T2 vs. T3  The response for this field would be 4.  In Table III from [Mbiti et al. 2019](https://www.google.com/url?q=https://academic.oup.com/qje/article/134/3/1627/5479257&sa=D&source=editors&ust=1741978839881140&usg=AOvVaw09dK2ObUvKtLkMijss0VNh), we see five comparisons in Panels A, B, and C: 1. Grants (α1) vs. None 2. Incentives (α2) vs. None 3. Combination (α3) vs. None 4. Combination (α3) vs. Grants+Incentives (α2+α1) 5. Combination (α3) vs. Grants (α1)  The response for this field would be 5. |

#### Evaluation arm

| **Variable name in SurveyCTO** | [armEval] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Evaluation arm |
| **Definition (LABEL on SurveyCTO)** | Evaluation arm for which treatment effects are estimated |
| **Response options (open-text, numeric, date, text-CV)** | **Select all** |
| **CV (Choices on SurveyCTO)** | List of study arms |
| **SurveyCTO instructions for data entry mask and repeat level** | -Display evaluation arm and reference arm options on the same page.  -Use the answer to “Number of comparisons” to repeat this field. |
| **Coding instructions for coders (Hint on SurveyCTO)** | - For each comparison, select the arm(s) being evaluated from the list of arms in the study  - This information can be present either in the top row or column of a table or it may be reported in the rows as the variables for which treatment effects are being reported. Descriptions of each treatment effect in the results section can also illuminate the contrast under study for each estimate.)  - If authors pool separate study arms into a single group to compare against another, please select multiple study arms to describe the evaluation arm.  - For factorial designs, if an interaction is being evaluated, select ALL the relevant arms in the interaction. |
| **Descriptive example see the[section]used in the paper to extract** | In table 3 from [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/pii/S0304387818303808%23sec2&sa=D&source=editors&ust=1741978839882297&usg=AOvVaw0ig2LT5Y4Z0ZjPtXJmROcW): - We see 3 comparisons in the first panel of rows: 1. T2 (teacher training) vs. Control 2. T3 (T2 + teacher incentives) vs. Control 3. T4 (T2 + parenting training) vs. Control - We see 1 additional comparison in the second panel of rows that pools all treatment groups into 1 group. 4. Any Treatment (T2, T3, or T4) vs. Control  Thus, there are 4 different evaluation arms here and so the responses would be: For evaluation arm 1: Select *Teacher training*only For evaluation arm 2: Select *Teacher training & teacher incentives* only For evaluation arm 3: Select *Teacher training & parenting training* only For evaluation arm 4. Select *teacher training, teacher training & teacher incentives,*and *teacher training & parenting training* In Table III from [Mbiti et al. 2019](https://www.google.com/url?q=https://academic.oup.com/qje/article/134/3/1627/5479257&sa=D&source=editors&ust=1741978839882476&usg=AOvVaw23tospCzv2u4SBISJ6Wyv8), we see five comparisons in Panels A, B, and C: 1. Grants (α1) vs. None 2. Incentives (α2) vs. None 3. Combination (α3) vs. None 4. Combination (α3) vs. Grants+Incentives (α2+α1) 5. Combination (α3) vs. Grants (α1)  There are 5 different evaluation arms here and the responses would be: For evaluation arm 1: Grants For evaluation arm 2: Incentives For evaluation arm 3: Combination For evaluation arm 4: Combination For evaluation arm 5: Combination |

#### Reference arm

| **Variable name in SurveyCTO** | [armNonEval] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Reference arm |
| **Definition (LABEL on SurveyCTO)** | Reference arm in comparison, or the arm against which the treatment effect is estimated. |
| **Response options (open-text, numeric, date, text-CV)** | Select all |
| **CV (Choices on SurveyCTO)** | List of study arms |
| **SurveyCTO instructions for data entry mask and repeat level** | -Display evaluation arm and reference arm options on the same page.  -Use the answer to “Number of comparisons” to repeat this field. |
| **Coding instructions for coders (Hint on SurveyCTO)** | - For each comparison, select the reference arm for the comparison. - In most cases for a parallel design, the comparison is the control arm.  - This information can be present in the footnotes of the table such as "Control is the base arm" |
| **Descriptive example see the[section]used in the paper to extract** | The reference arm for [Ozler et al, 2018](https://www.google.com/url?q=https://pdf.sciencedirectassets.com/271688/1-s2.0-S0304387818X00029/1-s2.0-S0304387818303808/main.pdf?X-Amz-Security-Token%3DIQoJb3JpZ2luX2VjEPr%252F%252F%252F%252F%252F%252F%252F%252F%252F%252FwEaCXVzLWVhc3QtMSJHMEUCIBIXtAWC87o2rydQ72iymKULj81rLFn4VXvHUiv7y%252BDgAiEAuKwN6pYoZSavSyC0eOEvdrx9LuTKsVo3LCWPsd9En00quwUIk%252F%252F%252F%252F%252F%252F%252F%252F%252F%252F%252FARAFGgwwNTkwMDM1NDY4NjUiDBfsWdcUXKTZaxOScyqPBQS1UbyR6AHJJg1CfcLM0ktwrRC9hUnUrZduQG1Hs7QeSUP5Ctg1RIQSzvT32JI7dSEL7QWpPSCnv7OmxrgbUV1CcLDtkE%252Fg7ftbtRbgXaD6zBWaZKJmAh4HBJeXf1ZXAGHuo3fqxK9rpL6n12hGnJTCR5%252F9MKPAnwPxFd1CXzKAh4kvQaWuwXoTQjTQwJP79x2TeASNF%252BY4h7Ut8Vpq68h8U3lszNk%252FeKRkFKMPfv9F%252B%252BH8o80LXhjKvwsuwv4oryaL6wVWH2o9mR5IXYbkMScBRonIxwVLS0Un5l4cN6ThMFJgAtQlzRObBugdSJ6ZuDB0xeBNUx2gs4%252Bf%252Bv8WrnpOaxv94DDvUMpaYRQvBUigsmeVw%252Bda1oMQtrEv486ojuPkvv0wB4A%252FQYNVkBPzUEIYKTkanodISHS5bYfeeVMwCX3GU3ubksvPYNlC6z%252FkDYDWi6Etfa%252FX5F1Tter8V8jBGlKnrCYNE%252FyqynBNaVgr7iL%252FDI2%252B6lhyh975QgNhG0Y4WJ9EM6meO821t%252FhC1eU31tkbQH%252Fijdh947qXMr%252F4PoDEvvbtl3pB51zwQcuLSfEWWQ5WKU1Nwjo13N6eetT6BMqCA%252BIwnZwyUKWS1JJEC6f08PWBwqybVCuJfsa8wbMNQ9XC10iDdDrXlrgiLQzQixcMiOH%252ByYS7Iqm2OQb9AASsLjENuDDcyiXbd2mxenvbZX%252BZHsKS2D4esis%252FH%252BW9U2KOXiywfsDb1HTZl4nekcvkF6SvywhPXMbVnpqUPAhmvAhXUMxdrud8gYEbmozTEDLLyxuLNCnVQ65Z5jiHENC3m1CxQ%252BxH49lHv62KT%252FL%252BdHVTahYtuDrQbvqXxnFv6nOcw0%252FgueNTcPNQ7qgwnuasswY6sQHOYcl%252Bx140IIzljyv1YtUGu4z3ExkqoQ4wUU0O%252BgZ1LF%252BAtOqkxD37ZOE%252BUAI5SlUKKkLoLRyZpqZHuc3J8rFOAAQc7L68ukfnmBxU5m9l2%252Fcb4Ly7TmdQCpxgwhs70ejAZDfu4ekeBSZpdhnGbVEZlSXghbovRUrm94QtOFvTChMrXUUKT0XBRicoOr7BhLDIQow3FyABC2mERBFBqy8uwHsg0yxk8%252B6QwTW34lacR8k%253D%26X-Amz-Algorithm%3DAWS4-HMAC-SHA256%26X-Amz-Date%3D20240613T192501Z%26X-Amz-SignedHeaders%3Dhost%26X-Amz-Expires%3D300%26X-Amz-Credential%3DASIAQ3PHCVTYS767FPFS%252F20240613%252Fus-east-1%252Fs3%252Faws4_request%26X-Amz-Signature%3D5e7ac749915b5f79f19a99726187c3c5b933f57ac1bdd3b749fafee8b50989aa%26hash%3Db5ab2ac1b28e433cf586fc9bb42656d599d4843fdf9f9eb81c41860c218b95c3%26host%3D68042c943591013ac2b2430a89b270f6af2c76d8dfd086a07176afe7c76c2c61%26pii%3DS0304387818303808%26tid%3Dspdf-3f22ac66-d959-4bc9-8ecf-3913910ae521%26sid%3D53b3d59033d4934dd358d4f26a3193c211d3gxrqa%26type%3Dclient%26tsoh%3Dd3d3LnNjaWVuY2VkaXJlY3QuY29t%26ua%3D11145e58515358590a%26rr%3D893476d64fa13adc%26cc%3Dus%26kca%3DeyJrZXkiOiJiTERybnNWZ2V4VW5NcmxBTHdjYUpnUjZrUU96MnlhWm5nWkJIaXJPWkM4UFBDSmcwWFM5ZnhuU0EvQzhoWnFobzRPcS9wM3AvT3pId2kzYnlXRVczcE80VE8yd1ZxRkg2emxsWlo5cmNDMjhLT1IweHBUeG90SHFISTVIQ1YzSldHWE0xaWJRRlBlZlo1UXdjOVh3MkVaNXNCOVZNRFY1RHJ0ZStZcE43aUlsUkZuSDZBPT0iLCJpdiI6ImQ3NjEwNGEyNDcyN2FmMWZhYWYxMzQ4ZGRkYWRkNGY1In0%3D_1718306752902&sa=D&source=editors&ust=1741978839883529&usg=AOvVaw21yl66x9df7MRWwGDO7r5s) table 3 is the control group.  Here, coders should select the arm "Learning kits" that was created when they mapped intervention to arms.  The five reference arms for Table III from Mbiti et al. 2019 are: 1. None 2. None 3. None 4. Grants + Incentives [select two interventions] 5. Grants |

#### Estimand

| **Variable name in SurveyCTO** | [tableestFull] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Estimand estimated for treatment effects in the table. |
| **Definition (LABEL on SurveyCTO)** | Indication that the table reports estimates using intention-to-treat and/or local average treatment effect/treatment on the treated estimand. |
| **Response options (open-text, numeric, date, text-CV)** | Text-CV, select one |
| **CV (Choices on SurveyCTO)** | **Select one**  Only ITT estimates Only LATE estimates Both ITT and LATE/TOT estimates Neither |
| **Coding instructions for coders (Hint on SurveyCTO)** | -In the IDEAL project, the main quantity of interest (or estimand) is the intention-to-treat (ITT) effect using the entire experimental sample. At the same time, IDEAL also collects treatment effects estimated using authors’ preferred estimand other than ITT, such as LATE/TOT.  -In this question, please indicate all the estimands reported in the table, either they are preferred by IDEAL or the authors.   -An effect is intent-to-treat (ITT) if the authors are interested in estimating the effect on everyone assigned to receive treatment, regardless of whether or not they actually received the treatment. When there is perfect compliance - for example, say that in a population of 200, 100 people are randomly assigned to treatment and all 100 people are actually treated, the ITT is the same as the Average Treatment Effect (ATE).  -An effect is local average treatment effect (LATE) the authors are estimating the effect among those who comply with treatment assignment.  -An effect is treatment on the treated (TOT)/average treatment on the treated (ATET) if the authors are estimating the effect on those who actually take up the treatment and non-compliance is one-sided. That is, the control group cannot or does not get the treatment.  -Usage of the ITT/LATE/TOT/ATET terminology is mostly seen in the discipline of Economics. For studies from other disciplines, ascertaining whether or not an effect is meant to be an ITT/LATE/TOT/ATET effect will require inference from what is described in a section on Methods.  - Read through the methods sections, tables and table notes to find information on the estimand used for listed outcomes. The estimands may vary across outcomes and rounds of data collection in a paper.  Select -Only ITT estimates; if only ITT is reported for all of the listed outcomes -Only LATE estimates; if only LATE is reported for all of the listed outcomes -Both ITT and LATE/TOT estimates; if both ITT and LATE/TOT estimates are reported for any of the listed outcomes. If ITT is reported for one set of outcomes and LATE/TOT is reported for another set of outcomes, select this option. -Neither; if neither of the estimands is reported for any of the listed outcomes. |
| **Descriptive example see the[section]used in the paper to extract** | [Baysan, 2022](https://www.google.com/url?q=https://www.aeaweb.org/articles?id%3D10.1257/aer.20201892&sa=D&source=editors&ust=1741978839863776&usg=AOvVaw3TlIb2_VUhJ-xZ-x8hG6ml). In the section D. Implementation, the author notes "... Therefore, I estimate only the intent-to-treat (ITT) effect". There are four exhibits in the papers. Figure 1 and Table 2 report full-sample treatment effects using ITT. Table 1 and Figure 2 only present treatment effects by quartile instead of full sample, thus they are not considered in IDEAL data extraction.  The response to this question across all the four exhibits is "Only ITT estimates". Note that you need to answer this question for each table separately.   In [Yoshikawa et al., 2015](https://www.google.com/url?q=https://pubmed.ncbi.nlm.nih.gov/25706589/&sa=D&source=editors&ust=1741978839863949&usg=AOvVaw0J0i4LHKVIGuyN_acZIO84), the ITT/LATE/TOT/ATET terminology is not used. In the Data Analysis Strategy section, the estimation equation indicates that treatment effects are estimated based on the treatment assignment, i.e. being in a classroom in a FULL UBC prekindergarten. Thus, the estimand is ITT by inference. There are two exhibits in the paper including treatment effects. Table 1 reports the treatment effects for the full sample of teachers, and Table 2 presents the treatment effects for the full sample of children, both using ITT estimand.  The response to this question for each of the two tables is "Only ITT estimates"  [Sigh et al., 2018](https://www.google.com/url?q=https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6073612/&sa=D&source=editors&ust=1741978839864032&usg=AOvVaw1e3dYrxd1IyGWVPs8anvPu) do not use the ITT/LATE/TOT/ATET terminology and do not use an equation to describe the estimation. In 2.9 Statistical Analysis, they write "An analysis of the effectiveness of the intervention was based on the randomization of the product the patients were originally assigned using all available case data including patients missing follow-ups and dropouts. This analysis method handles missing data by fitting a statistical model over all available case data without introducing bias.", which implies the estimand is ITT as the analysis estimates treatment effects on patients originally assigned to receive the product. Table 4 reports the treatment effects through differences in means for all patients.  The response to this question is "Only ITT estimates"  [Linhares et al., 2022](https://www.google.com/url?q=https://journals.copmadrid.org/pi/archivos/1132_0559_inter_31_1_0021.pdf&sa=D&source=editors&ust=1741978839864137&usg=AOvVaw2Gr1uYgpYAEsgPxzLWSxPu) performed both ITT and TOT analyses, however, only TOT results are presented in the paper exhibits (Tables 3 and 4). The ITT results are omitted because the effects are not significant. The ITT estimates can not be extracted from the paper.  The response to this question is "Only LATE/TOT estimates"  [Ganimian, Mulralidharan, and Walters, 2022](https://www.google.com/url?q=https://econweb.ucsd.edu/~kamurali/papers/Working%2520Papers/TNICDS_ECE%2520(Current%2520WP).pdf&sa=D&source=editors&ust=1741978839864226&usg=AOvVaw1bC7Utz7Os1llnlSQFSG3Y)present both ITT and TOT results using two samples (i.e. HH assessments and AWC assessments). The authors explain in the Introduction, "Moreover, treatment-on-the-treated effects obtained by scaling the household sample estimates by the share of children observed at the center at the endline are close to the AWC and common sample estimates. We therefore interpret the AWC estimates as reflecting treatment effects on children who actively attended the centers, while the household estimates capture intent-to-treat-style impacts on the set of eligible children, many of whom had limited treatment exposure." In Tables 2, 5 and 6, both ITT and TOT effects on assessment scores are reported.  The response to this question across tables is "Both ITT and LATE/TOT estimates"  In [Banerjee et al., 2020](https://www.google.com/url?q=https://www.aeaweb.org/articles?id%3D10.1257/pol.2.1.1&sa=D&source=editors&ust=1741978839864319&usg=AOvVaw2487QxhsT9jvTYGh5Y1Xfv), ITT estimates are reported for all outcomes (Tables 1 - 4). In addition, LATE estimates are also presented for learning outcomes (in Table 4).  For Tables 1 -3 , the answer would be “Only ITT estimates”.  The answer to Table 4 would be “Both ITT and LATE/TOT estimates”. |

#### Outcomes in table

| **Variable name in surveyCTO** | [tableOut] |
| --- | --- |
| **Field name**  **(LABEL on SurveyCTO)** | Outcome name |
| **Definition**  **(LABEL on SurveyCTO)** | The outcomes reported in the exhibit that contain full sample treatment effect(s) |
| **Response options (open-text, numeric, date, text-CV)** | List of outcomes |
| **CV**  **(Choices on SurveyCTO)** | List of outcomes pre-specified from list of outcomes in the study [outLabel] |
| **Coding instructions for coders** **(Hint on surveyCTO)** | - Display all the outcomes for which full sample treatment effects are reported in the table  - These may appear on the heading row or the first column of the table. |
| **Descriptive example** **see the[section]used in the paper to extract** | For [Ozler et al, 2018](https://pdf.sciencedirectassets.com/271688/1-s2.0-S0304387818X00029/1-s2.0-S0304387818303808/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEPr%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJHMEUCIBIXtAWC87o2rydQ72iymKULj81rLFn4VXvHUiv7y%2BDgAiEAuKwN6pYoZSavSyC0eOEvdrx9LuTKsVo3LCWPsd9En00quwUIk%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FARAFGgwwNTkwMDM1NDY4NjUiDBfsWdcUXKTZaxOScyqPBQS1UbyR6AHJJg1CfcLM0ktwrRC9hUnUrZduQG1Hs7QeSUP5Ctg1RIQSzvT32JI7dSEL7QWpPSCnv7OmxrgbUV1CcLDtkE%2Fg7ftbtRbgXaD6zBWaZKJmAh4HBJeXf1ZXAGHuo3fqxK9rpL6n12hGnJTCR5%2F9MKPAnwPxFd1CXzKAh4kvQaWuwXoTQjTQwJP79x2TeASNF%2BY4h7Ut8Vpq68h8U3lszNk%2FeKRkFKMPfv9F%2B%2BH8o80LXhjKvwsuwv4oryaL6wVWH2o9mR5IXYbkMScBRonIxwVLS0Un5l4cN6ThMFJgAtQlzRObBugdSJ6ZuDB0xeBNUx2gs4%2Bf%2Bv8WrnpOaxv94DDvUMpaYRQvBUigsmeVw%2Bda1oMQtrEv486ojuPkvv0wB4A%2FQYNVkBPzUEIYKTkanodISHS5bYfeeVMwCX3GU3ubksvPYNlC6z%2FkDYDWi6Etfa%2FX5F1Tter8V8jBGlKnrCYNE%2FyqynBNaVgr7iL%2FDI2%2B6lhyh975QgNhG0Y4WJ9EM6meO821t%2FhC1eU31tkbQH%2Fijdh947qXMr%2F4PoDEvvbtl3pB51zwQcuLSfEWWQ5WKU1Nwjo13N6eetT6BMqCA%2BIwnZwyUKWS1JJEC6f08PWBwqybVCuJfsa8wbMNQ9XC10iDdDrXlrgiLQzQixcMiOH%2ByYS7Iqm2OQb9AASsLjENuDDcyiXbd2mxenvbZX%2BZHsKS2D4esis%2FH%2BW9U2KOXiywfsDb1HTZl4nekcvkF6SvywhPXMbVnpqUPAhmvAhXUMxdrud8gYEbmozTEDLLyxuLNCnVQ65Z5jiHENC3m1CxQ%2BxH49lHv62KT%2FL%2BdHVTahYtuDrQbvqXxnFv6nOcw0%2FgueNTcPNQ7qgwnuasswY6sQHOYcl%2Bx140IIzljyv1YtUGu4z3ExkqoQ4wUU0O%2BgZ1LF%2BAtOqkxD37ZOE%2BUAI5SlUKKkLoLRyZpqZHuc3J8rFOAAQc7L68ukfnmBxU5m9l2%2Fcb4Ly7TmdQCpxgwhs70ejAZDfu4ekeBSZpdhnGbVEZlSXghbovRUrm94QtOFvTChMrXUUKT0XBRicoOr7BhLDIQow3FyABC2mERBFBqy8uwHsg0yxk8%2B6QwTW34lacR8k%3D&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Date=20240613T192501Z&X-Amz-SignedHeaders=host&X-Amz-Expires=300&X-Amz-Credential=ASIAQ3PHCVTYS767FPFS%2F20240613%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Signature=5e7ac749915b5f79f19a99726187c3c5b933f57ac1bdd3b749fafee8b50989aa&hash=b5ab2ac1b28e433cf586fc9bb42656d599d4843fdf9f9eb81c41860c218b95c3&host=68042c943591013ac2b2430a89b270f6af2c76d8dfd086a07176afe7c76c2c61&pii=S0304387818303808&tid=spdf-3f22ac66-d959-4bc9-8ecf-3913910ae521&sid=53b3d59033d4934dd358d4f26a3193c211d3gxrqa&type=client&tsoh=d3d3LnNjaWVuY2VkaXJlY3QuY29t&ua=11145e58515358590a&rr=893476d64fa13adc&cc=us&kca=eyJrZXkiOiJiTERybnNWZ2V4VW5NcmxBTHdjYUpnUjZrUU96MnlhWm5nWkJIaXJPWkM4UFBDSmcwWFM5ZnhuU0EvQzhoWnFobzRPcS9wM3AvT3pId2kzYnlXRVczcE80VE8yd1ZxRkg2emxsWlo5cmNDMjhLT1IweHBUeG90SHFISTVIQ1YzSldHWE0xaWJRRlBlZlo1UXdjOVh3MkVaNXNCOVZNRFY1RHJ0ZStZcE43aUlsUkZuSDZBPT0iLCJpdiI6ImQ3NjEwNGEyNDcyN2FmMWZhYWYxMzQ4ZGRkYWRkNGY1In0=_1718306752902), table 3 presents 5 outcomes. These are present in the top row of table 3 under the title "Dependent variable". Since these outcome names were collected previously in the survey, here coders should select:  -Attending CBCC: 2012-13 -Enrolled in Primary: 2012-13 -Malawi Developmental Assessment Tool: Total Score -Malawi Developmental Assessment Tool: Language Skills -Malawi Developmental Assessment Tool: Fine Motor/Perception Skills  For [Freeman et al. 2022](https://pubmed.ncbi.nlm.nih.gov/36962125/), Table 2 reports treatment effects on 10 outcomes. Each outcome is listed as a row in the Indicator column. The full list of outcomes collected for the study would appear as options, and coders should select:  -Anxiety score -Depression score -Emotional distress score -Well-being score -High Anxiety -High Depression -High Emotional distress -Poor well-being |

#### Number of periods in the table

| **Variable name in SurveyCTO** | [tableRoundNum] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Number of periods in the table |
| **Definition (LABEL on SurveyCTO)** | Number of periods over which treatment effects were estimated in the exhibit. |
| **Response options (open-text, numeric, date, text-CV)** | Numeric |
| **CV (Choices on SurveyCTO)** | Numeric |
| **Coding instructions for coders (Hint on SurveyCTO)** | - List the number of distinct periods over which the treatments effects in the table were estimated. Periods refer to the units of time for which the paper reports treatment effects.  - Most times each period only involves one round of data collection. But authors may pool several rounds of data collection into one extended period for some analysis. [See examples in the Descriptive example column]  - When authors combine rounds of data collections, the number of periods in any given table may exceed the total number of rounds of data collection.  - In general, the baseline is not a period. This is because studies do not estimate the treatment effect at the time of the baseline survey/ However, if the authors use a difference-in-differences design to estimate the treatment effect, then the baseline is included in the period. |
| **Descriptive example see the[section]used in the paper to extract** | In [Hanna et al., 2016](https://www.google.com/url?q=https://www.aeaweb.org/articles?id%3D10.1257/pol.20140008&sa=D&source=editors&ust=1741978839884727&usg=AOvVaw0SM4CnWW8T4d4kcf9dFeia), table 3 presents 5 periods used to calculate treatment effects. The first row reports a period that includes all four rounds of midline and endline surveys in estimating the treatment effects. The following rows report the treatment effects in each individual period of midline and endline surveys (one per year since the treatment was administered). Coders should report "5" for this field, 1 each for the 4 separate yearly mid- and end-line surveys and 1 for the pooled endline.  In contrast, in [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/pii/S0304387818303808?fr%3DRR-2%26ref%3Dpdf_download%26rr%3D8aeff7110aa405c4&sa=D&source=editors&ust=1741978839884820&usg=AOvVaw3PXoFNmYyRc5jOg_7X02ap), table 3,the title of the table specifies the results are for the 18-month follow up. Coders should report "1" for this field, as only one round of data collection was used in this table. |

#### Rounds of data collection in the table

| **Variable name in SurveyCTO** | [tableRound] |
| --- | --- |
| **Field name (LABEL on SurveyCTO)** | Rounds of data collection in exhibit. |
| **Definition (LABEL on SurveyCTO)** | The rounds of data collection in each exhibit. |
| **Response options (open-text, numeric, date, text-CV)** | Select all |
| **CV (Choices on SurveyCTO)** | List of round names from “Round name” |
| **Coding instructions for coders (Hint on SurveyCTO)** | - Select the rounds of data collection for each period reported in the table.  - For treatment effects estimated within a period with only one round of data collection, select solely that round. For treatment effects estimated in a period with multiple rounds of data collection, select all rounds that were pooled in that estimation during the period. |
| **Descriptive example see the[section]used in the paper to extract** | In [Ozler et al, 2018](https://www.google.com/url?q=https://www.sciencedirect.com/science/article/pii/S0304387818303808?fr%3DRR-2%26ref%3Dpdf_download%26rr%3D8aeff7110aa405c4&sa=D&source=editors&ust=1741978839885614&usg=AOvVaw1RQsroP2nS7KCFDnzoIjlt), table 3, the title of the table specifies the results are for the 18-month follow up.  So for this field, coders should select "18-month follow up"  In [Hanna et al., 2016](https://www.google.com/url?q=https://pubs.aeaweb.org/doi/pdfplus/10.1257/pol.20140008&sa=D&source=editors&ust=1741978839885698&usg=AOvVaw09fFT4SxEpNNtT3s6E8wBJ), table 3 presents treatment effects for individual rounds of data collection. For the individual rounds, coders should select the corresponding round names (e.g., 0-12 month survey, 13-24 month survey, etc.). For the pooled round, coders should select all four of the individual rounds. |

#### Number of empirical specifications in the table

| **Variable name in SurveyCTO** | [tableSpecNum] |
| --- | --- |
| **Field name** **(LABEL on SurveyCTO)** | Number of empirical specifications in each exhibit |
| **Definition** **(LABEL on SurveyCTO)** | Number of empirical specifications used to estimate treatment effects in the table. |
| **Response options (open-text, numeric, date, text-CV)** | Numeric |
| **CV** **(Choices on SurveyCTO)** | Integer greater than zero |
| **Coding instructions for coders** **(Hint on SurveyCTO)** | -Please identify and indicate the number of unique empirical specifications used to estimate treatment effects in the exhibit.  -A unique empirical specification is defined by the inclusion (or not) of strata correction, baseline outcome and other controls. [This table](https://ideal-consortium.github.io/Schema/IDEAL_Ranking_Index.html) illustrates all 8 types of specifications, and a coder should count how many of them were used to estimate treatment effect in the exhibit.   -B*aseline value of outcome* refers to the baseline value of the outcome for which the treatment effect is being estimated. Controls for baseline values of other variables are not included in *baseline value of outcome*.  - *Controls include those are observed before treatment assignment* refer to variables measured at baseline, as well as *the ones that are* measured after baseline but are considered to be static over the study period - for example, parental education in the context of a program targeting children, gender in the context of a program that is not assumed to affect gender identity, or residence in a flood-prone zone in a context of a program that is not assumed to affect choice of residence.  - A simple difference in means with no additional information on the specification used or the presence of any adjustment should be considered as "No other controls". Likewise, a regression with just the outcome of interest as the dependent variable and just an indicator for treatment status as the sole independent variable should be coded as "No other controls."  - Information on specifications can be found in both data and statistical analysis sections and table notes. In some papers, authors report the regression specification or empirical model used to estimate treatment effects and/or indicate their specification in column labels in tables. Sometimes, authors add additional details about specifications to table notes only. In other studies, this information must be inferred from the text and/or notes that accompany tables or figures. |
| **Descriptive example** **see the[section]used in the paper to extract** | In Riley, 2024, E. The empirical Strategy describes the specifications in equation (1) that strata dummies and the baseline value of the outcome (if measured at baseline, otherwise excluded) are included in the estimation of treatment effects. In Table 1, where full sample results are reported, the notes indicate "All regressions include strata dummies and include the baseline value of the outcomes."  The response to this question is: 1  In [Ashraf et al., 2010](https://www.jstor.org/stable/41038767), according to tables notes, in Table 2, two specifications are used for each outcome: with and without baseline controls (including baseline Clorin usage and water cholorination, general health behaviors and attitudes, household demographics, and locality fixed effects).  The response to this question is: 2  [Grossman and Baldassarri, 2012](https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-5907.2012.00596.x) report treatment effects using four set of specifications in Table 2 including no controls, individual controls, monitor profile (post-treatment control), and both individual controls and monitor profile. There are in total 4 specifications.  The response to this question is: 4 |

\*\*\* End of Stage 1 \*\*\*