

# AidGrade

## *Coding Manual*

May 31, 2013

## Table of Contents

1. Data Extraction Form Fields.....	3
1.1 General Identifying Information.....	3
1.2 Paper and Result Characteristics .....	6
1.3 Results.....	13
1.4 Notes.....	15
2 Conventions.....	15
3 Mistakes to avoid.....	16

# 1. Data Extraction Form Fields

## 1.1 General Identifying Information

### **Author**

This is the author's last name. If there are two authors, please follow the convention "LastName1 and LastName2"; if there are three, "LastName1, LastName2 and LastName3"; if there are more than three, "LastName1 et al."

### **Number of authors**

This is the number of authors.

### **Publication year**

This is the year of the publication or working paper. Only the most recent version of the papers were collected.

### **Primary purpose code, secondary purpose code.**

The primary purpose code provides a general description of the overall purpose of the study. The code should clarify the goal the study as a whole had in mind, as opposed to the trial name, which lists the specific outcome focused on for a single set of results within the study. Assign one purpose code for each project, representing the project's main goal.

Use the OECD CRS main purpose codes available in the Drive.

If the project has two target sectors, judge which is the secondary purpose and use it to fill out the field for the secondary purpose code. If it has three or more purposes, only code the main two.

For the secondary purpose code field:

9 = no secondary purpose code

All primary purpose code fields must be filled.

### **Primary activity code, secondary activity code**

This field provides a more detailed description of the project. We will use AidData's coding scheme, derived from the CRS coding scheme; this is also in the Drive.

Every project should be assigned at least one activity code. The primary activity code should always fall under the primary purpose code and the secondary activity code should always fall under the secondary purpose code. As an example, activity code "11220.03: Basic education infrastructure" falls under purpose code "11220: Primary education".

### **Publication code**

- 0 = unpublished paper or working paper that is not part of a working paper series
- 1 = working paper that is part of a working paper series
- 2 = published in an economics journal
- 3 = published in a public health journal
- 4 = published in another journal
- 9 = not an article (*e.g.* it is published in a book)

This field codes whether the paper is unpublished or published, with different codes depending on the type of journal in which the paper was published.

### **Publication name**

This is the name of the journal in which the paper was published, if it was published, or the name of the working paper series, if it is in a working paper series. Please write the name as it would be written in bibliography, such as "World Bank Policy Research Working Paper" (if it is explicitly in that working paper series; many World Bank publications are not). Never abbreviate the journal name.

- . = unpublished paper or working paper that is not part of a working paper series
- 9 = not an article

### **Organization code**

This is the type of organization or group that implemented the program, such as a government agency, NGO, or academic research group. This group is usually different from the author(s) conducting the study - the organization is the one administering the treatment in question, rather than analyzing the data on the treatment.

0 = not stated

1 = government agency/program

2 = NGO/nonprofit

3 = private sector

4 = academic research institution (including professors)

9 = other

### **Organization name**

This is the name of the organization that implemented the program, if applicable.

If it is a government program, it should be listed as “Government of X”. More detailed information, such as the name of the ministry or department running the program, instead goes under the field “Program name”.

If the organization is an NGO or nonprofit, report the name exactly as it is written in the paper itself.

Beware studies that appear to have been run by academics, because they may have been run by a local partner. If the program is truly run by academics, please list it as “Department of X, University Name”.

“..” = not stated

### **Program name**

This is the name of the program, if applicable. For example, one government program was called “Familias en Acciones”.

“..” = not stated

## **Type study paper**

This field uniquely identifies papers. Start with 1 and increase by 1 every time you hit a new paper; we will modify these values later, so the actual values taken are not important (instead, it is important they be unique within the dataset when they are supposed to be unique and not unique when they are supposed to not be unique).

## **1.2 Paper and Result Characteristics**

### **Crossover design**

0 = did not use a crossover design

1 = used a crossover design

Very few papers use a crossover design, but it is important to note those that do. If a paper did use a crossover design, it would have made a note of this in the text. Search the text for “crossover” and “cross-over”, then make sure that any results returned do refer to the study’s design.

### **Method**

Coding scheme:

1 = randomized controlled trial (RCT);

2 = matching (whether "propensity score matching" or "nearest neighbour matching" or any other kind of matching);

3 = differences-in-differences (DD or diff-in-diff). For this type of study, different groups receive the treatment at different times, but it isn't randomized which group will get it first (if it is, code it as an RCT). For example, maybe the program was rolled out geographically;

4 or 5 = regression discontinuity (RD) design. In a regression discontinuity design, there was some cut-off threshold to determine who received the treatment – for example, only people below the poverty line receive the program – and the intervention explicitly looks at only those just above and just below that cut-off threshold as the control and treatment groups.

RDs can use parametric or non-parametric methods. Please code parametric RDs as 4 and non-parametric RDs as 5.

9 = None of the above. The above are the main codes, but you may encounter papers that don't appear to neatly fall into any of these categories! If you do, please code it as a 9. If you are unclear on which method was used in a paper, please email, attaching the relevant paper.

Please see the handout on IE methods to learn a bit more about each of these types of studies.

For any given paper, the authors may have used more than one type of method. In general, we want their “preferred specification”. Sometimes they state which it is explicitly, sometimes it is not explicit. If it is difficult to figure out which is the best data to pull, please ask. In general, if the authors do not state which is their preferred specification, we will consider matching < DD < parametric RD < non-parametric RD < RCT, but sometimes the DD result is more credible than the RD result (this depends on whether it is likely either estimator is picking up spurious shocks).

### **Blinded**

0 = not blinded  
1 = single-blinded  
2 = double-blinded  
9 = not stated

### **Cluster randomized**

0 = not cluster randomized  
1 = cluster randomized  
9 = not stated

To be able to tell whether a study was randomized by cluster or not, please first consider at what level the outcome variable is at – e.g. is it an effect that varies by individual? Household? Village? Then consider at what level the treatment was randomized. E.g. was it randomized by individual, household, village? If it was randomized at a higher level than the outcome variable, it was randomized by cluster.

To give a concrete example: if it was randomized by village or by school (some villages got it and others didn't, or some schools got it and others didn't), but the actual treatment effects are measured at the level of individual people, that would count as being randomized by cluster. If the effects are measured at the village or school level, that would not be randomized by cluster. Similarly, if the effects are at the individual level and the randomization was done at an individual level, then that would not be randomized by cluster.

### **Village, province, country**

This records the village or city, province or state, and country in which the intervention took place.

### **Number of months after intervention**

This records the number of months between the beginning of the treatment and the time of the data collection for the midline or end results being reported in the row. If weeks are reported, please convert that to months using  $(\text{weeks})/7*30$ .

Record "0" if the entry is referring to baseline results.

If the intervention or data collection occurred over a period of months, enter the shortest period of time between the beginning of the intervention and the data collection. Do not enter any characters other than numbers in this field.

0 = baseline

999 = not stated

### **Certainty of number of months after intervention**

0 = not clear

1 = clear

9 = not applicable: 999 was selected for the number of months after intervention



It can be difficult to tell how much time elapsed between the end of an intervention and data collection, thus the separate subjective code for whether there is uncertainty about the value given in the previous column.

### **Attrition reported**

Attrition reported takes the value “1” if the study discusses attrition, i.e. if anyone dropped out of the study. You don't actually have to record the attrition rate, just whether it was mentioned or not. If you can find a number who dropped out, code “1”, regardless of whether any analysis is done of this attrition, otherwise “0”. Some studies might deal with attrition in a complicated way, so this is intended to keep it easier to record. The point of recording this is that some people say that whether a study reports attrition is one measure of how good a quality the study is – it's part of a commonly used quality indicator.

### **Costs discussed**

Costs discussed reported takes the value “1” if the study discusses the cost per treated unit in the study itself or provides enough information about costs that the costs per treated unit in the study itself could be derived. This field takes the value “2” if the study does not discuss the cost per treated unit in the study itself but does discuss the typical cost per treated unit in another or other contexts. (If the study does both, code as “1”.) You don't actually have to record the costs, just whether they were mentioned or not. If the study did not mention costs, enter “0”.

### **Type study sample**

This field uniquely identifies sample groups. Some papers cover more than one group of people. On the other hand, if a paper is a follow-up to another paper, it might be on the same group of people, so this would be a way of showing that. As with “Type study paper”, the actual number does not matter; start with 1 and increase by 1 every time you hit a new sample group.

### **Sample description**

This field provides a fuller description of the sample, such as whether it was on a particular gender or age group.

## **Trial number**

This is a unique id number we assign to a particular set of results. It generally is different for each row of data, with the exception that if different rows regard the same results (see section on: baseline, final, and change data) they should be given the same identifier. As with “Type study paper”, the actual number does not matter; start at 1 and increase by 1 every time you hit a new set of results.

## **Trial name**

This field describes the outcome variable in words and includes the units of the reported results. It should be detailed enough for an outside reader to understand and should include definitions where applicable.

## **Baseline, final, change**

Papers may report the baseline values of the outcome variables, the final values of the outcome variables, or the change in values of the outcome variable. Baseline, final and change are binary variables that take the value “1” if the row represents baseline, final or change data and “0” otherwise. If the results include midline data, include and code it as “final”; the difference will be clear by the “no. of months after intervention” variable. It is also important to give the different rows the same identifier in column A if they are regarding the same results.

Here is an example:

Suppose Group A receives treatment, Group B is the control. At baseline, you observe:

Group A outcome value: 10

Group B outcome value: 11

6 months later, you observe:

Group A outcome value: 15

Group B outcome value: 14

You would record all this information – in the row in which you record 15 and 14, you would record a "1" for "final", and in the row in which you record 10 and 11, "1" for "baseline". You would record these in different rows with the same identifier in column A so they can be matched later.

The paper might instead only report the values 15 and 14, and not report baseline results. Then you would just record a "1" for "final" in the same row in which you record 15 and 14.

The paper might instead only report the changes in the values:

Group A change in outcome value: 5

Group B change in outcome value: 3

Then you would record a "1" for "change" in the same row in which you record 5 and 3.

### **Intent-to-treat**

0 = TOT

1 = ITT

7 = not stated; implied TOT

8 = not stated; implied ITT

9 = not stated; no idea

We say that an estimate is an intent-to-treat (ITT) estimate if it looks at the effect on everyone assigned to receive treatment, regardless of whether or not they actually took advantage of the program. The alternative is that it estimates the treatment effect on the treated (TOT).

Note that some papers may provide an estimate that appears to be a TOT or ITT estimate, but will not clearly state the fact. These papers should be coded as "7" or "8", respectively. If it is not possible to infer from the text whether an estimate refers to TOT or ITT results, code it as "9".

If a study includes both TOT and ITT results, code each in separate rows.

### **Shared control group**

A paper might compare multiple treatment arms against the same control group. This may appear in the data set as several rows sharing the same control group data or may be harder to see as the comparison with the control group is implicit (as in an estimated treatment coefficient). Regardless of whether the comparison with the control group is implicit or explicit, if several treatment arms within a paper use the same control group, flag it here. If the paper only has a single treatment arm, code this as “9” – not applicable.

0 = no multiple treatment arms / each treatment arm has its own control group

1 = multiple treatment arms share the same control group

9 = not applicable

### **Spillover effect**

A paper might report “spillover effects”; effects that a treatment had on a group that was not treated. For example, a scholarships program might only provide scholarships to girls, but boys in the same school may be affected by it. In this case, record both sets of results, but code the spillover effect accordingly.

0 = not a spillover effect

1 = spillover effect

### **Aggregate effect**

A paper might report results for many different subgroups or many different time periods as well as all subgroups combined or all time periods combined. Record results for each subgroup if applicable, according to our conventions, and also for all the subgroups or all the time periods combined if this information is available. Use this column to record whether the result is at the most aggregate level.

For example, suppose a paper reports results separately for urban girls, urban boys, rural girls, rural boys, all rural children, all urban children, all girls, all boys, and all children. Each result would be entered in a separate row, but only the result for all children would be coded as “1” in this column. If no results for subgroups are reported, code this variable as “9”.

Aggregate effects should always be recorded where available; for subgroups, refer to the section in this document on “Conventions”.

0 = not an aggregate effect  
1 = aggregate effect  
9 = no results for subgroups reported

### 1.3 Results

There are several ways in which researchers may present their results. It is very important to pay attention to the difference between these data types. The next sections describe several ways in which the data could be reported, in turn.

If the data in question are not provided, please enter “..” to indicate their absence.

#### **Treatment coefficient, treatment standard error, number of observations, reg CI lower bound, reg CI upper bound, reg t-statistic and reg p-value**

First, researchers may provide the estimate of the treatment effect (as in the coefficient in a regression), its standard error, and the number of observations on which this estimate was based. This would result in 3 columns in the Excel file.

Please note that there are some regression results we cannot use as the variables are not on a scale that we can combine with the results of other regressions. For example, if we are looking at the effect of a school meals program on test scores, we cannot combine studies that report coefficients representing raw changes in test scores. Consider the case of tests out of different scores – a score of 10/10 is different than a score of 10/100. If the paper reports results using standardized values, however, in which the coefficients represent changes in terms of standard deviations, we can combine them.

If it's a regression and it doesn't use standardized values, we can still sometimes standardize the results if the paper gives enough information. If you see a paper with regression results that need standardization, write [codinghelp@aidgrade.org](mailto:codinghelp@aidgrade.org) a note about this with the paper attached and someone will look for the required data.

Occasionally a paper may provide a confidence interval, a t-statistic or a p-value instead of the standard error. In this case, please record the reported values in reg CI lower bound, reg CI upper bound, reg t-statistic or reg p-value. If the standard error is reported, you do not need to record these values.

**Treatment group mean, treatment group std dev, treatment group number, control group mean, control group std dev, control group number**

Second, researchers may provide the mean, standard deviation, and number of observations for each of the treatment and control group. This would result in 6 columns in the Excel file.

**Treatment number of events, treatment number of possible events, control number of events, control number of possible events**

Third, sometimes researchers will report the number of events that occurred in each of the treatment and control group out of a number of possible events. E.g. 3 out of 100 people in the treatment group died and 10 out of 100 people in the control group died. This would yield 4 columns.

**E<sub>e</sub>, t<sub>e</sub>, e<sub>c</sub>, t<sub>c</sub>, e<sub>e</sub>/t<sub>e</sub>, e<sub>c</sub>/t<sub>c</sub>**

Fourth, the paper may present rates in a relatively disaggregated form. Rates often mention units such as "person-years". Rates are more complicated to deal with and warrant a bit more explanation. For a good description of rates, see: [http://www.mrc-bsu.cam.ac.uk/cochrane/handbook/chapter\\_9/9\\_4\\_8\\_meta\\_analysis\\_of\\_counts\\_and\\_rates.htm](http://www.mrc-bsu.cam.ac.uk/cochrane/handbook/chapter_9/9_4_8_meta_analysis_of_counts_and_rates.htm). We follow the same naming conventions for simplicity.

**Risk ratio, risk ratio CI lower bound, risk ratio CI upper bound, risk ratio standard error, risk ratio t-statistic, risk ratio p-value, hazard ratio, hazard ratio CI lower bound, hazard ratio CI upper bound, hazard ratio standard error, hazard ratio t-statistic, hazard ratio p-value, odds ratio, odds ratio CI lower bound, odds ratio CI upper bound, odds ratio standard error, odds ratio t-statistic, odds ratio p-value, rate ratio, rate ratio CI lower bound, rate ratio CI upper bound, rate ratio standard error, rate ratio t-statistic, rate ratio p-value**

Fifth, the paper may present ratios. Most often, a paper will report a point estimate and the lower and upper bound of the confidence interval around that point, but sometimes a paper may report standard errors, t-statistics or p-values instead. Please record the point estimate and the lower and upper bounds. If the confidence intervals are not reported but the standard error, t-statistic or p-value is provided, enter that instead in its respective column.

## **Page number**

Record the page number of the pdf on which you found the result entered. If the pdf uses page numbers that do not correspond with the pages as numbered within the journal article or working paper (e.g. page 3 of the pdf corresponds to page 219 in the page numbering of the nested document), please use the page numbering of the nested document.

## **1.4 Notes**

### **Notes**

This field contains a description of anything you would like to flag.

## **2 Conventions**

We have our own internal conventions.

First, say a study is done on children aged 6-11. If a researcher runs regressions on the whole sample and then on only children aged 6-7, we do not treat those latter results as completely trustworthy, because researchers generally only report sub-groups when they are significant. However, if the researchers report results for children aged 6-7, 8-9, and 10-11, we could include those data, because they would not be likely to suffer the same bias.

Similarly, we try to use results with as few control variables as possible. Control variables can be very helpful and we would like to include results that use them, but unfortunately, they can be “gamed” to ensure obtaining the results the researcher wants.

Similarly, until researchers can pre-commit to their analysis plans, we cannot trust results involving the treatment interacted with another variable. We use regressions that do not include the treatment interacted with anything else where possible.

### 3 Mistakes to avoid

Please note the following so that the data can be analyzed more easily:

- Don't try to combine different data types in the same column. Each different data type should go into a different column so that when you look at that column you know what was being reported. The idea is that someone should be able to look at the Excel and know the results of the paper.
- Please put the lower bound and the upper bound of confidence intervals in different columns.
- Please keep all columns for the same topic in the same Excel sheet as in the template. The data will then look very sparse, with many “.” cells coded for the results section.
- Please be specific about what the outcome variable was and what the treatment was, if it is not clear. E.g. if the outcome variable was whether the person had malaria, how was that defined clinically? If the outcome variable is height, what were the units?
- Please only mention percents in the "trial name" column and not attached to the data. (E.g. -17 rather than -17%.)
- Please distinguish between percents (%) and percentage points. (An example of the difference: if enrollment rates increase from 80% by 10%, enrollment rates are now 88%. If enrollment rates increase from 80% by 10 percentage points, enrollment rates are now 90%.)
- Please do not place asterisks next to the numbers. If something needs explaining, you can use the "notes" column for this without interfering with the numbers.
- Standard errors are different from standard deviations.
- When you see text that says something like "standard errors were clustered", this is not the same thing as the study being clustered. See the section on “cluster randomized” for an explanation.
- Do not use accents or other special characters such as “-“, “;“, “:“, etc.

This document tried to cover the most common issues that come up, but you will undoubtedly have other questions as you work through the papers. Do email [codinghelp@aidgrade.org](mailto:codinghelp@aidgrade.org) immediately if any questions arise, attaching the paper your question is in reference to.