BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Guidelines for Verification of Computational Reproducibility in Economics

ACRE Team

2020-06-26

Contents

5

Introduction **7**
Beyond binary judgments 8
Stages of the exercise 9
Recording the results of the exercise 10
Reproduction Strategies 10

1 Scoping **13**
1.1 From candidate to declared paper 14
1.2 Scoping your declared paper 16
1.3 Identify your relevant timeline. 18

2 Assessment **21**
2.1 Describe the inputs. 21
2.2 Connect each output to all its inputs 26
2.3 Assign a reproducibility score. 31

3 Improvements **45**
3.1 Types of output-level improvements 45
3.2 Types of paper-level improvements 49

4 Checking for Robustness **51**
4.1 Mapping the universe of robustness checks 52
4.2 Proposing a specific robustness check 53

5 Guidance for a Constructive Exchange Between Reproducers and Original Authors **55**
5.1 For Reproducers Contacting the Authors of the Original Study . 56
5.2 For Original Authors Responding to Requests from Reproducers 65

6 Reproduction Diagrams **67**
6.1 Different Scenarios 67

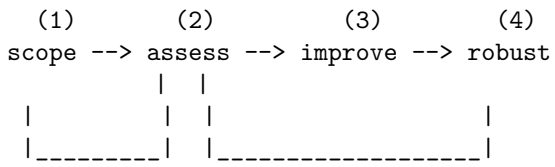
7 Additional resources **69**
7.1 Some summaries 69
7.2 Links 69

8 Contributions **71**
8.1 Contributing feedback on these guidelines 71
8.2 List of Contributors: Guidelines content and source code: 71

9 Acknowledgments **73**

10 Definitions **75**
10.1 Concepts in reproducibility 75

See a full list of contributors



(1) Scoping

(2) Assessment

(3) Improvement

(4) Robustness

(5) Extensions

Outcome-level

Paper-level

Select paper

Describe inputs

+ Raw data

+ Version control

Analytical choices

New method

Check ACRE

Reproduction diagrams

+ Analysis data

+ Documentation

Type of choice

New data

Check Rep. pkg exists

Reproduction score
+ Analysis code
+ Dynamic document
Choice value
Read paper
+ Cleaning code
+ File structure
Justify and test alternatives
Declare estimates
Debug analysis code
Debug cleaning code
Record results in Survey 1
Record results in Survey 2
Record results in Survey 3

Introduction

In 2019, the American Economic Association updated its Data and Code Availability Policy to require that the AEA Data Editor verify the reproducibility of all papers before they are accepted by an AEA journal. In addition to the requirements laid out in the policy, several specific recommendations were produced to facilitate compliance. This change in policy is expected to improve the computational reproducibility of all published research going forward, after several studies showed that rates of *computational reproducibility* in economics at large range from somewhat low to alarmingly low (Galiani et al., 2018; Chang and Li, 2015; Kingi et al., 2018).

Replication, or the process by which a study's hypotheses and findings are re-examined using different data or different methods (or both) (King, 1995) is an essential part of the scientific process that allows science to be "self-correcting." *Computational reproducibility*, or the ability to reproduce the results, tables, and other figures of a paper using the available data, code, and materials, is a precondition for replication. Computational reproducibility is assessed through the process of *reproduction*. At the center of this process is the *reproducer* (you!), a party rarely involved in the production of the original paper. Reproductions sometimes involve the *original author* (whom we refer to as "the author") in cases where additional guidance and materials are needed to execute the process.

This exercise is designed for reproductions performed in economics graduate courses or undergraduate theses, with the goal of providing a common approach, terminology, and standards for conducting reproductions. The goal of reproduction, in general, is to assess and improve the computational reproducibility of published research in a way that facilitates further robustness checks, extensions, collaborations, and replication.

This exercise is part of the Accelerating Computational Reproducibility in Economics (ACRE) project, which aims to assess, enable, and improve the computational reproducibility of published economics research. The ACRE project is led by the Berkeley Initiative for Transparency in the Social Sciences (BITSS)—an initiative of the Center for Effective Global Action (CEGA)—and Dr. Lars Vilhuber, Data Editor for the journals of the American Economic Association (AEA). This project is supported by the Laura and John Arnold Foundation.

Beyond binary judgments

Assessments of reproducibility can easily gravitate towards binary judgements that declare an entire paper "reproducible" or "non-reproducible." These guidelines suggest a more nuanced approach by highlighting two realities that make binary judgments less relevant.

First, a paper may contain several scientific claims (or major hypotheses) that may vary in computational reproducibility. Each claim is tested using different methodologies, presenting results in one or more display items (outputs like tables and figures). Each display item will itself contain several specifications. Figure 1 illustrates this idea.

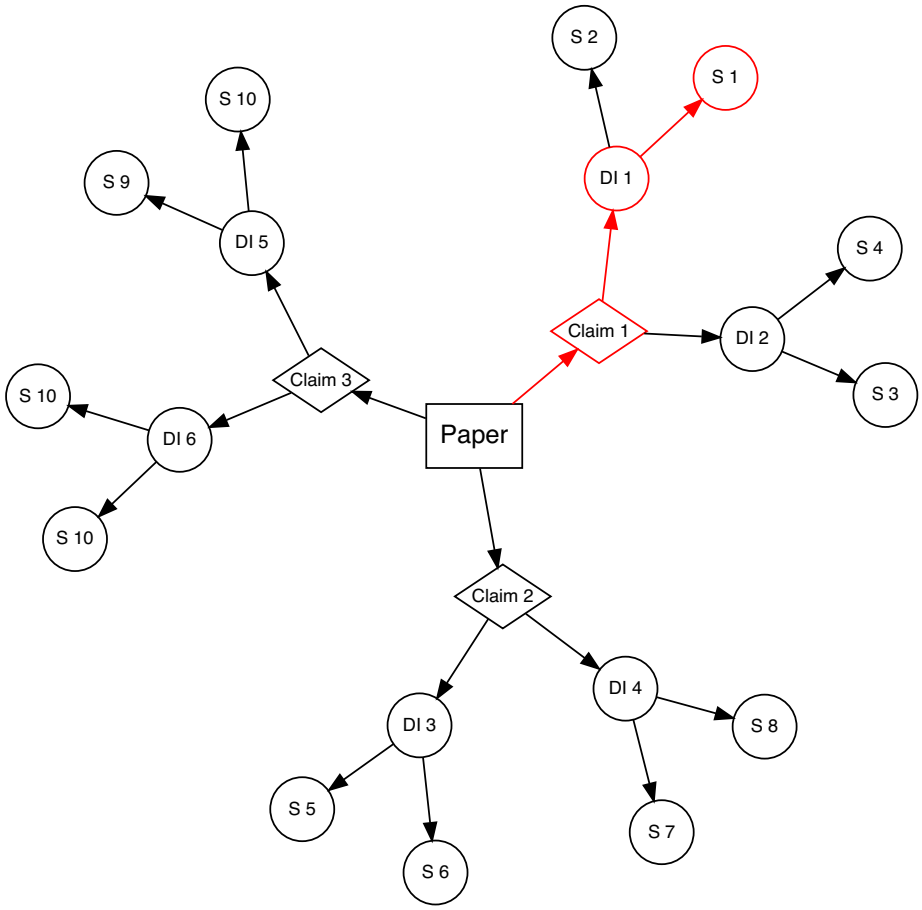


Figure 1: One paper has multiple components to reproduce. DI: Display Item, S: Specification

Second, for any given specification there are several levels of reproducibility,

ranging from the absence of any materials to complete reproducibility starting from raw data. And even for a specific claim-specification, distinguishing the appropriate level can be far more constructive than simply labeling it as (ir)reproducible.

Note that the highest level of reproducibility, which requires complete reproducibility starting from raw data, is very demanding to achieve and should not be expected of all published research — especially before 2019. Instead, this level can serve as an aspiration for the field of economics at large as it seeks to improve the reproducibility of research and facilitate the transmission of knowledge throughout the scientific community.

Stages of the exercise

This reproduction exercise is divided into four stages, corresponding to the first four chapters of these guidelines, with a fifth optional stage:

1. **Scoping**, where you (the reproducer) will define the scope of the exercise by declaring a paper and the specific output(s) on which you will focus for the remainder of the exercise;

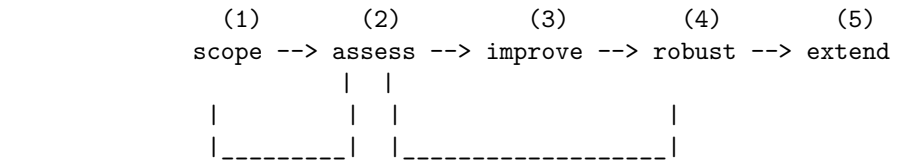
2. **Assessment**, where you will review and describe in detail the available reproduction package, and assess the current level of computational reproducibility of the selected outputs;

3. **Improvement**, where you will modify the content and/or the organization of the reproduction package to improve its reproducibility;

4. **Robustness checks**, where you will assess the quality of selected analytical choices; and

5. **Extension** (if applicable), where you may extend the current paper by including new methodologies or data. This step brings the reproduction exercise a step closer to *replication*.

Figure 2: Steps for reproduction



Suggested level of effort:
- Graduate

research:	5%	10%	5%	10%	70%
- Graduate course:	10%	25%	20%	40%	5%
- Undergrad thesis:	10%	30%	40%	20%	0%

Figure 2 depicts suggested levels of effort for each stage of the exercise depending on the context in which you are performing a reproduction. This process need not be chronologically linear. For example, you may realize that the scope of a reproduction is too ambitious and switch to a less intensive one. Later in the exercise, you can also begin testing different specifications for robustness while also assessing a paper's level of reproducibility.

Recording the results of the exercise

You will be asked to record the results of your reproduction as you progress through each stage.

In *Stage 1: Scoping*, complete **Survey 1**, where you will declare your paper of choice and the specific display item(s) and specifications on which you will focus for the remainder of the exercise. This step may also involve writing a brief 1-2 page summary of the paper (depending on your instructor or goals).

In *Stage 2: Assessment*, you will inspect the paper's reproduction package (raw data, analysis data, and code), connect the display item to be reproduced with its inputs, and assign a reproducibility score to each output.

In *Stage 3: Improvement*, you will try to improve the reproducibility of the selected outputs by adding missing files, documentation, and report any potential changes in the level of reproducibility. Use **Survey 2** to record your work at Stages 2 and 3 (you will receive access instructions for Survey 2 when you submit Survey 1).

In *Stage 4: Robustness Checks*, you will assess different analytical choices and test possible variations. Use **Survey 3** to record your work at this stage.

Reproduction Strategies

Generally, a reproduction will begin with a thorough reading of the study being reproduced. However, subsequent steps may follow from a *reproduction strategy*. For example, a reproduction may closely follow the order of the steps outlined above. This might entail the reproducer first choosing a set of results whose pproduction they are interested in assessing or understanding, completely reproducing these results to the extent possible, and then making modifications to the reproduction package. Another potential strategy could be for the reproducer to develop potential robustness checks or extensions while reading the study, which would lead to the definition of a set of results to be assessed via

reproduction. Yet another reproduction strategy may be for the reproducer to seek out a paper that uses a particular dataset to which they have access or an interest in using, reproducing the results that use that dataset as an input, then probing the robustness of the results to various data cleaning decisions.

The various uses of reproduction makes the number of potential reproduction strategies quite large. In choosing or designing a reproduction strategy, it is helpful to clearly identify the goal of the reproduction. In all of the examples laid out in the paragraph above, the order in which the steps of the reproduction exercise are taken is at least partially determined by what the reproducer hopes to get from the exercise. The structure provided in these guidelines, together with a clear reproduction goal, can facilitate the implementation of an efficient reproduction strategy.

Chapter 1

Scoping

In this stage, you will define the scope of your exercise by declaring a paper and the specific output(s) on which you will focus. You might first consider multiple papers without analyzing them more closely (we refer to these as **candidate papers**) before moving forward with your **declared paper**.

It is likely that you will choose a declared paper based on whether or not you can locate its reproduction package. A **reproduction package** is the collection of materials that make it possible to reproduce a paper. This package may contain data, code, or documentation. If you are unable to independently locate the reproduction package for your paper, you can ask the paper's author for it (find guidance on this in Chapter 6) or simply choose another candidate paper. If you still want to explore the reproducibility of a paper with no reproduction package, these guidelines provide instructions for requesting materials from authors to create a public reproduction package, or if this proves unsuccessful, for building your reproduction package from scratch.

To avoid duplicating the efforts of others who may be interested in reproducing one of your candidate papers, **we ask that you record your candidate papers in the ACRE database** (currently under development).

Note that in this stage, *you are not expected to review the reproduction materials in detail*, as you will dedicate most of your time to this in later stages of the exercise. If materials are available, you will read the paper and declare the scope of the reproduction exercise. You can expect to spend between 1-3 days in this Scoping stage, though this may vary based on the length and the complexity of the paper, and the availability of reproduction materials.

Use Survey 1 to record your work in this stage.

1.1 From candidate to declared paper

At this point of the exercise, you are *only validating the availability* of (at least) one reproduction package and not assessing the quality of its content. Follow the steps below to verify that a reproduction package is available, and stop whenever you find it (this may mean mean that you have found your declared paper).

1. Check whether previous reproduction attempts have beeen recorded in the ACRE Database for the paper (more on the ACRE Database in the next section).
2. Check the journal or publisher's website, looking for materials named "Data and Materials," "Supplemental Materials," "Reproduction/Replication Package/Materials," etc.

3. Look for links in the paper (review the footnotes and appendices).

4. Review the personal websites of the paper's author(s).
5. Contact the author(s) to request the reproduction package using this email template. In this and future interactions with authors, we encourage you to follow our guidance outlined in Chapter 5.
6. Deposit the reproduction package in a trusted repository (e.g., Dataverse, Open ICPSR, Zenodo, or the Open Science Framework) under the name `Original reproduction package for - Title of the paper`. You will be asked to provide the URL of the repository in Survey 1.

In case you need to contact the authors, make sure to *allocate sufficient time for this step* (we suggest at least three weeks before the date you plan to start the reproduction). Instructors should also plan to accordingly (e.g., if the ACRE exericse is expected to take place in the middle of the semester, students should review candidate papers and (if applicable) contact the authors in the first few weeks of the semester).

Review the decision tree (Figure #) below for a more detailed overview of this process. Remember, *if at any step of the process you decide to abandon the paper, make sure to record the candidate paper in the ACRE database* before moving on to another candidate paper. Once you have obtained the reproduction package, the *candidate paper* becomes your *declared paper* and you can move forward with the exercise! Do not invest time in doing a detailed read of any paper until you are sure that it is your declared paper.

1.1.1 Candidate paper entries in the ACRE Database

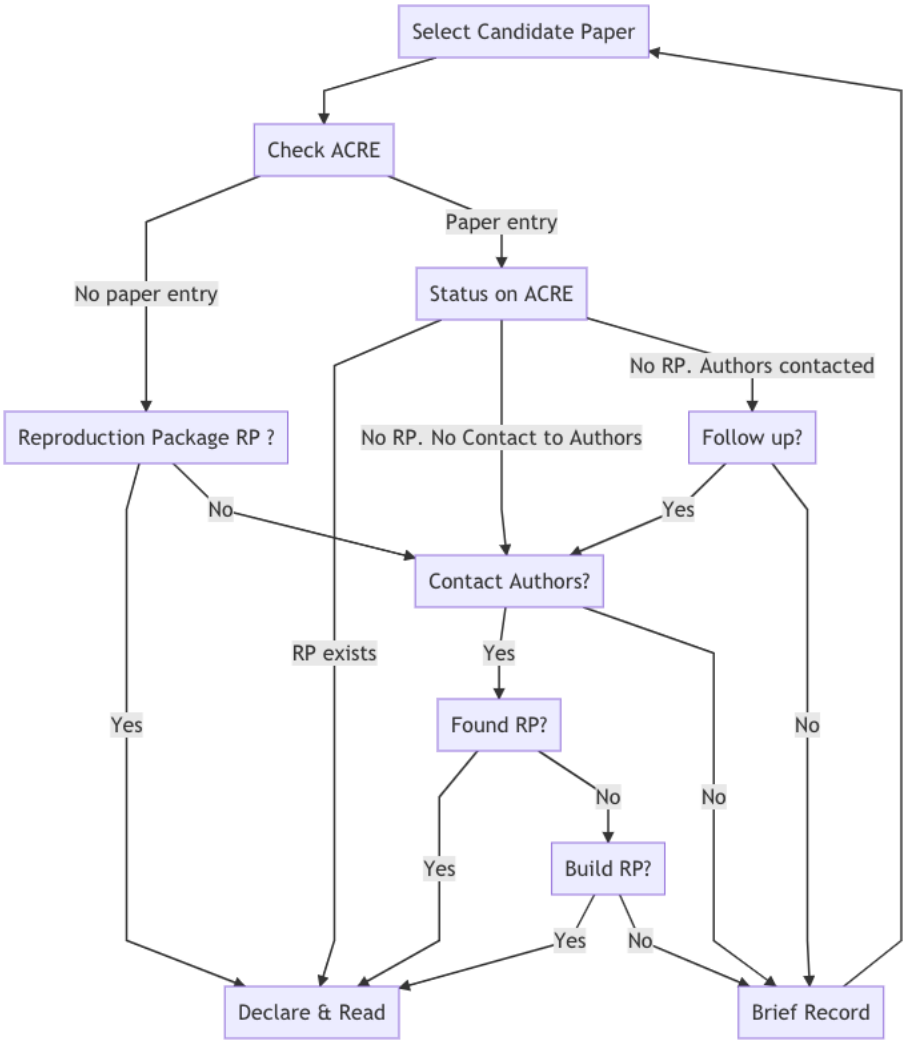
If the ACRE database contains previous reproduction attempts of the paper, you will see a report card with the following information:

> **Box 1:** Summary Report Card for ACRE Paper Entry
> **Title:** Sample Title

Authors: Jane Doe & John Doe
Original Reproduction Package Available: URL/No [What does this mean? Add some context]. [If "No"] **Contacted Authors?:** Yes/No
[If "Yes(contacted)"] **Type of Response:** Categories (6).
Additional Reproduction Packages: Number (eg., 2)
Authors Available for Further Questions for ACRE Reproductions: Yes/No/Unknown
Open for reproductions: Yes/No [Same as above: what does this mean? Add more context].

If after taking steps 1-5 above (or for some other reason) you are unable to locate the reproduction package, record your candidate paper (and if applicable, the outcome of your correspondence with the original authors) in the ACRE database following the example above.

View Decision Tree To Select Paper (Emma: add title and solve bug with svg)



1.2 Scoping your declared paper

Once you have identified your declared paper, get familiarized with it and choose the specific output(s) on which you will focus for the remainder of the exercise.

1.2.1 Read and summarize the paper

Depending on how much time you have, we recommend that you write a short (1-2 page) summary of the paper. This will help remind you of the key elements to focus on for the reproduction, and demonstrate your understanding of the paper (for yourself and others like your instructor or advisor).

When reading or summarizing the paper, try to answer the following questions:

- Would you classify the paper's scientific claims as mainly focused on estimating a causal relationship, estimating/predicting a descriptive statistic of a population, or something else?
- How many scientific claims (descriptive or causal) are investigated in the paper?
- What is the population for which the estimates apply?
- What is the population that is the focus of the paper as a whole?
- What are the main data sources used in the paper?
- How many display items are there in the paper (tables, figures, and inline results)?

- What is the main statistical or econometric method used to examine each claim?
- What is the author's preferred specification (or yours, if the authors are not clear)?
- What are some robustness checks for the preferred specification?

1.2.2 Record scope of the exercise

By now you should have a fairly good understanding of the paper's content. You do not, however, need to have spent any time reviewing the reproduction package in detail.

At this point, you should clearly specify which part of the paper will be the main focus of your reproduction. Focus on specific estimates, represented by a unique combination of claim-display item-specification as represented in figure 1. If you plan to scope more than one claim, *we strongly recommend starting with just one* and recording your results. You can then initiate another record in ACRE later for the second (or third, fourth, etc.) claim to reproduce using the materials and knowledge you developed in the first exercise. You can, however, reproduce more than one claim if you are already familiar with the paper.

In the Assessment stage, the reproduction will be centered around the display item(s) that contain the specification you indicate at this point.

Declare specific main estimates to reproduce.

Identify a scientific claim and its corresponding preferred specification, and record its magnitude, standard error, and location in the paper (page, table #, and table row and column). If the authors did not explicitly chose a particular estimate, you will be asked to select one. In addition to the preferred estimate, reproduce up to five estimates that correspond to alternative specifications of the preferred estimate.

Declare possible robustness checks for main estimates (optional).

After reading the paper, you might wonder why the authors did not conduct a specific robustness test. If you think that such analysis could have been done *within the same methodology* and *using the same data* (e.g., by including or excluding a subset of the data like "high-school dropouts" or "women"), please specify a robustness test that you would like to conduct before starting the Assessment stage.

These are the elements you will need for the Scoping stage. **You now have all the elements necessary to complete Survey 1.**

1.3 Identify your relevant timeline.

Before you begin working on the three main stages of the reproduction exercise (Assessment, Improvement, and Robustness), it is important to manage your own expectations and those of your instructor or advisor. Be mindful of your time limitations when defining the scope of your reproduction activity. These will depend on the type of exercise chosen by your instructor or advisor and may vary from a weeklong homework assignment, to a longer class project that may take a month to complete or a semester-long project (an undergraduate thesis, for example).

Table 1 shows an example distribution of time across three different reproduction formats. The Scoping and Assessment stages are expected to last roughly the same amount of time across all formats (lasting longer for the semester-long activities, and acknowledgin that less experienceed researchers, such as undergraduate students, may need more time). Differences emerge in the distribution of time for the last two main stages: Improvements and Robustness. For shorter exercises, we recommend avoiding any possible improvements to the raw data (or cleaning code). This will limit how many robustness checks are possible (for example, by limiting your ability to reconstruct variables according to slightly different definitions), but it should leave plenty of time for testing different specifications at the analysis level.

Emma: please write this table using R and KableExtra

2 weeks (~10 days)

1 month (~20 days)

1 semester (~100 days)

analysis data

raw data

analysis data

raw data

analysis data

raw data

Scoping

10% (1 day)

5% (1 day)

5% (5 days)

Assessment

35%

25%

15%

Improvement

25%

0%

40%

20%

30%

Robustness

25%

5%

25%

25%

```
library(tidyverse)
library(knitr)
library(kableExtra)
temp_eval <- TRUE
options(tinytex.verbose = TRUE)
```

Chapter 2

Assessment

In this stage, you will review and describe in detail the available reproduction materials, and assess levels of computational reproducibility for the selected outputs, as well as for the overall paper. This stage is designed to record as much of the learning process behind a reproduction as possible to facilitate incremental improvements, and allow future reproducers to pick up easily where others have left off.

First, you will provide a detailed description of the reproduction package. Second, you will connect the outputs you've chosen to reproduce with their corresponding inputs. With these elements in place, you can score the level of reproducibility of each output, and report on paper-level dimensions of reproducibility.

In the *Scoping* stage, you declared a paper, identified the specific claims you will reproduce, and recorded the main estimates that support the claims. In this stage, you will identify all outputs that contain those estimates. You will also decide if you are interested in assessing the reproducibility of that entire output (e.g., "Table 1"), or will assess only a pre-specified estimates (e.g., "rows 3 and 4 of Table 1"). Additionally, you can include other outputs of interest.

Use *Survey 2* to record your work as part of this step.

Tip: We recommend that you first focus on one specific output (e.g., "Table 1"). After completing the assessment for this output, you will have a much easier time translating improvements to other outputs.

2.1 Describe the inputs.

This section explains how to list *all* input materials found or referred to in the reproduction package. First, you will identify data sources and connect them with their raw data files (when available). Second, you will locate and provide a

brief description of the analytic data files. Finally, you will locate, inspect, and describe the analytic code used in the paper.

The following terms will be used in this section:

- **Cleaning code:** A script associated primarily with data cleaning. Most of its content is dedicated to actions like deleting variables or observations, merging data sets, removing outliers, or reshaping the structure of the data (from long to wide, or vice versa).
- **Analysis code:** A script associated primarily with analysis. Most of its content is dedicated to actions like running regressions, running hypothesis tests, computing standard errors, and imputing missing values.

2.1.1 Describe the data sources and raw data.

In the paper you chose, find references to all *data sources* used in the analysis. A data source is usually described in narrative form. For example, if in the body of the paper you see text like "…for earnings in 2018 we use the Current Population Survey…", the data source is "Current Population Survey 2018". If it is mentioned for the first time on page 1 of the Appendix, its location should be recorded as "A1". Do this for all the data sources mentioned in the paper.

Data sources also vary by unit of analysis, with some sources matching the same unit of analysis used in the paper (as in previous examples), while others are less clear (e.g., "our information on regional minimum wages comes from the Bureau of Labor Statistics." This should be recorded as "regional minimum wages from the Bureau of Labor Statistics").

Next, look at the reproduction package and map the *data sources* mentioned in the paper to the *data files* in the available materials. Record their folder locations relative to the main reproduction folder[1]. In addition to looking at the existing data files, we recommend that you review the first lines of all code files (especially cleaning code), looking for lines that call the datasets. Inspecting these scripts may help you understand how different data sources are used, and possibly identify any files that are missing from the reproduction package.

Record this information in this standardized spreadsheet (download it or make a copy for yourself), using the following structure:

Raw data information

data_source

page

data_files

[1] a relative location takes the form of `/folder_in_rep_materials/sub_folder/file.txt`, in contrast to an absolute location that takes the form of `username/documents/projects/repros/folder_in_rep_materials/sub_folder/file.txt`

known_missing

directory

“Current Population Survey 2018”

A1

cepr_march_2018.dta

/data/

“DHS 2010 - 2013”

4

nicaraguaDHS_2010.csv; boliviaDHS2010.csv; nicaraguaDHS_2011.csv; nicaraguaDHS_2012.csv; boliviaDHS_2012.csv; nicaraguaDHS_2013.csv; boliviaDHS_2013.csv

boliviaDHS_2011.csv

/rawdata/DHS/

“2017 SAT scores”

4

Not available

/data/to_clean/

…

…

…

…

…

Raw data information:

data_source	page	data_files	known_missi
"Current Population Survey 2018"	A1	cepr_march_2018.dta	
"DHS 2010 - 2013"	4	nicaraguaDHS_2010.csv; boliviaDHS_2010.csv; nicaraguaDHS_2011.csv; nicaraguaDHS_2012.csv; boliviaDHS_2012.csv; nicaraguaDHS_2013.csv; boliviaDHS_2013.csv	boliviaDHS_
"2017 SAT scores"	4	Not available	

```
| ...                    | ...  | ...                                                  |
|-----------------------|------|-----------------------------------------------------|--
```

Note: lists if files in the data_files and known_missing columns should have entries separated by a semi-colon to for the spreadsheet to be compatible with the ACRE Diagram Builder.

2.1.2 Describe the analytic data sets.

List all the analytic files you can find in the reproduction package, and identify their locations relative to the main reproduction folder. Record this information in the standardized spreadsheet.

As you progress through the exercise, add to the spreadsheet a one-line description of each file's main content (for example: `all_waves.csv` has the simple description `data for region-level analysis`). This may be difficult in an initial review, but will become easier as you go along.

The resulting report will have the following structure:

Analysis data information

analysis_data

location

description

final_data.csv

/analysis/fig1/

data for figure1

all_waves.csv

/final_data/v1_april/

data for region-level analysis

…

…

…

```
Analysis data information:
   |----------------|-----------------------|--------------------------------|
   | analysis_data  | location              | description                    |
   |----------------|-----------------------|--------------------------------|
   | final_data.csv | /analysis/fig1/       | data for figure1               |
   |----------------|-----------------------|--------------------------------|
   | all_waves.csv  | /final_data/v1_april/ | data for region-level analysis |
   |----------------|-----------------------|--------------------------------|
```

...	...	...

2.1.3 Describe the code scripts.

List all code files that you found in the reproduction package and identify their locations relative to the master reproduction folder. Review the beginning and end of each code file and identify the inputs required to successfully run the file. Inputs may include data sets or other code scripts that are typically found at the beginning of the script (e.g., `load`, `read`, `source`, `run`, `do`). For each code file, record all inputs together and separate each item with ";". Outputs may include other datasets, figures, or plain text files that are typically at the end of a script (e.g., `save`, `write`, `export`). For each code file, record all outputs together and separate each item with ";". Provide a one-line description of what each code file does. Record all of this information in the standardized spreadsheet, using the following structure:

Code files information

file_name

location

inputs

outputs

description

primary_type

output_table1.do

/code/analysis/

analysis_data01.csv

output1_part1.txt

produces first part of table 1 (unformatted)

analysis

data_cleaning02.R

/code/cleaning

admin_01raw.csv

analysis_data02.csv

removes outliers and missing vals from raw admin data

cleaning

...

…

…

…

…

…

```
Code files information:
   |--------------------|------------------|----------------------|-------------------
   | file_name          | location         | inputs               | outputs
   |--------------------|------------------|----------------------|-------------------
   | output_table1.do   | /code/analysis/  | analysis_data01.csv | output1_part1.txt
   |                    |                  |                      |
   |                    |                  |                      |
   |--------------------|------------------|----------------------|-------------------
   | data_cleaning02.R  | /code/cleaninig/ | admin_01raw.csv      | analysis_data02.cs
   |                    |                  |                      |
   |                    |                  |                      |
   |--------------------|------------------|----------------------|-------------------
   | ...                | ...              | ...                  | ...
   |--------------------|------------------|----------------------|-------------------
```

As you gain an understanding of each code script, you will likely find more inputs and outputs – we encourage you to update the standardized spreadsheet. Once finished with the reproduction exercise, classify each code file as *analysis* or *cleaning*. We recognize that this may involve subjective judgment, so we suggest that you conduct this classification based on each script's main role.

Note: If a code script takes multiple inputs and/or produces multiple outputs they should be listed as semicolon separated lists in order to be compatible with the ACRE Diagram Builder.

2.2 Connect each output to all its inputs

Using the information collected above, you can trace your output-to-be-reproduced to its primary sources. Email the standardized spreadsheets from above (sections 2.1.1, 2.1.2 and 2.1.3) to the ACRE Diagram Builder at acre@berkeley.edu. You should receive an email within 24 hours with a reproduction diagram tree that represents the information available on the workflow behind a specific output.

2.2.1 Complete workflow information

If you were able to identify all the relevant components in the previous section, the ACRE Diagram Builder will produce a tree diagram that looks similar to the one below.

```
table1.tex
       |___[code] analysis.R
           |___analysis_data.dta
               |___[code] final_merge.do
                   |___cleaned_1_2.dta
                   |   |___[code] clean_merged_1_2.do
                   |       |___merged_1_2.dta
                   |           |___[code] merge_1_2.do
                   |               |___cleaned_1.dta
                   |               |   |___[code] clean_raw_1.py
                   |               |       |___raw_1.dta
                   |               |___cleaned_2.dta
                   |                   |___[code] clean_raw_2.py
                   |                       |___raw_2.dta
                   |___cleaned_3_4.dta
                       |___[code] clean_merged_3_4.do
                           |___merged_3_4.dta
                               |___[code] merge_3_4.do
                                   |___cleaned_3.dta
                                   |   |___[code] clean_raw_3.py
                                   |       |___raw_3.dta
                                   |___cleaned_4.dta
                                       |___[code] clean_raw_4.py
                                           |___raw_4.dta
```

This diagram, built with the information you provided, is already an important contribution to understanding the necessary components required to reproduce a specific output. It summarizes key information to allow for more constructive exchanges with original authors or other reproducers. For example, when contacting the authors for guidance, you can use the diagram to point out specific files you need. Formulating your request this way makes it easier for authors to respond and demonstrates that you have a good understanding of the reproduction package.

2.2.2 Incomplete workflow information

In many cases, some of the components of the workflow will not be easily identifiable (or missing) in the reproduction package. Here the Diagram Builder will return a partial reproduction tree diagram. For example, if the files `merge_1_2.do`, `merge_3_4.do`, and `final_merge.do` are missing from the previous diagram, the ACRE Diagram Builder will produce the following diagram:

```
cleaned_3.dta
       |___[code] clean_raw_3.py
           |___raw_3.dta


   table1.tex
```

```
        |___[code] analysis.R
            |___analysis_data.dta

    cleaned_3_4.dta
        |___[code] clean_merged_3_4.do
            |___merged_3_4.dta

    cleaned_1.dta
        |___[code] clean_raw_1.py
            |___raw_1.dta

    cleaned_2.dta
        |___[code] clean_raw_2.py
            |___raw_2.dta

    cleaned_4.dta
        |___[code] clean_raw_4.py
            |___raw_4.dta

    cleaned_1_2.dta
        |___[code] clean_merged_1_2.do
            |___merged_1_2.dta
    Unused data sources: None.
```

In this case, you can still manually combine this partial information with your knowledge from the paper and own judgement to produce a "candidate" tree diagram (which might lead to different reproducers recreating different diagrams). This may look like the following:

```
table1.tex
       |___[code] analysis.R
           |___analysis_data.dta
               |___MISSSING CODE FILE(S) #3
                   |___cleaned_3_4.dta
                   |       |___[code] clean_merged_3_4.do
                   |           |___merged_3_4.dta
                   |               |___MISSSING CODE FILE(S) #2
                   |                   |___cleaned_3.dta
                   |                   |       |___[code] clean_raw_3.py
                   |                   |           |___raw_3.dta
                   |                   |___cleaned_4.dta
                   |                           |___[code] clean_raw_4.py
                   |                               |___raw_4.dta
                   |___cleaned_1_2.dta
                           |___[code] clean_merged_1_2.do
                               |___merged_1_2.dta
                                   |___MISSSING CODE FILE(S) #1
```

```
|___cleaned_1.dta
|       |___[code] clean_raw_1.py
|           |___raw_1.dta
|
|___cleaned_2.dta
        |___[code] clean_raw_2.py
            |___raw_2.dta
```

To leave a record of the reconstructed diagrams, you will have to amend the input spreadsheets using placeholders for the missing components. In the example above, you should add the following entries to the code description spreadsheet:

Adding rows to code spreadsheet

file_name

location

inputs

outputs

description

primary_type

...

...

...

...

...

...

missing_file1

unknown

cleaned_1.dta, cleaned_2.dta

merged_1_2.dta

missing code

unknown

missing_file2

unknown

cleaned_3.dta, cleaned_4.dta

merged_3_4.dta

missing code

unknown

missing_file3

unknown

merged_3_4.dta, merged_1_2.dta

analysis_data.dta

missing code

unknown

```
Adding rows to code spreadsheet:
   |------------------|------------------|---------------------|---------------------
   | file_name        | location         | inputs              | outputs
   |------------------|------------------|---------------------|---------------------
   | ...              | ...              | ...                 | ...
   |------------------|------------------|---------------------|---------------------
   | missing_file1    | unknown          | cleaned_1.dta;      | merged_1_2.dta
   |                  |                  | cleaned_2.dta       |
   |------------------|------------------|---------------------|---------------------
   | missing_file2    | unknown          | cleaned_3.dta;      | merged_3_4.dta
   |                  |                  | cleaned_4.dta       |
   |------------------|------------------|---------------------|---------------------
   | missing_file3    | unknown          | merged_3_4.dta;     | analysis_data.dta
   |                  |                  | merged_1_2.dta      |
   |------------------|------------------|---------------------|---------------------
```

As in the cases with complete workflows, these diagrams (fragmented or reconstructed trees) provide important information for assessing and improving the reproducibility of specific outputs. Reproducers can compare reconstructed trees and/or contact original authors with highly specific inquiries.

For more examples of diagrams connecting final outputs to initial raw data, see here.

2.2.3 Unused Data Sources

It is possible that not all data included in a replication package are actually used in code scripts in the reproduction package. This would be the case if, for example, the raw data and analysis data are included, but not the script that generates the analysis data. As a concrete example, consider what the original diagram above would look like if the only code included in the reproduction package were analysis.R:

```
table1.tex
       |___[code] analysis.R
```

```
        |___analysis_data.dta

Unused data sources:
raw_1.dta
raw_2.dta
raw_3.dta
raw_4.dta

Unused analysis data:
cleaned_1.dta
cleaned_2.dta
cleaned_3.dta
cleaned_4.dta
merged_1_2.dta
merged_3_4.dta
cleaned_1_2.dta
cleaned_3_4.dta
```

In this case, there are many data files that were listed in the raw data and analytic data spreadsheets that are not used by any code script in the replication package.

2.3 Assign a reproducibility score.

Once you have identified all possible inputs and have a clear understanding of the connection between the outputs and inputs, you can start to assess the output-specific level of reproducibility.

Take note of the following concepts in this section:

- **Computationally Reproducible from Analytic data (CRA):** The output can be reproduced with minimal effort starting from the *analytic* datasets.
- **Computationally Reproducible from Raw data (CRR):** The output can be reproduced with minimal effort from the *raw* datasets.
- **Minimal effort:** One hour or less is required to run the code, not including computing time.

2.3.1 Levels of Computational Reproducibility for a Specific Output

Each level of computational reproducibility is defined by the availability of data and materials, and whether or not the available materials faithfully reproduce the output of interest. The description of each level also includes possible im-

provements that can help advance the reproducibility of the output to a higher level. You will learn in more detail about the possible improvements.

Note that the assessment is made *at the output level* – a paper can be highly reproducible for its main results, but suffer from low reproducibility for other outputs. The assessment includes a 10-point scale, where 1 represents that, under current circumstances, reproducers cannot access any reproduction package, while 10 represents access to all the materials and being able to reproduce the target outcome from the raw data.

- **Level 1 (L1):** No data or code are available. Possible improvements include adding: raw data (+AD), analysis data (+RD), cleaning code (+CC), and analysis code (+AC).

You will have detected papers that are reproducible at Level 1 as part of the Scoping stage (unsuccessful candidate papers). Make sure to take record them in Survey 1.

- **Level 2 (L2):** Code scripts are available (partial or complete), but no data are available. Possible improvements include adding: raw data (+AD) and analysis data (+RD).
- **Level 3 (L3):** Analytic data and code are partially available, but raw data and cleaning code are not. Possible improvements include: completing analysis data and/or code, adding raw data (+RD), and adding analysis code (+AC).
- **Level 4 (L4):** All analytic data sets and analysis code are available, but code does not run or produces results different than those in the paper (not CRA). Possible improvements include: debugging the analysis code (DAC) or obtaining raw data (+RD).
- **Level 5 (L5):** Analytic data sets and analysis code are available. They produce the same results as presented in the paper (CRA). The reproducibility package may be improved by obtaining the original raw data sets.

This is the highest level that most published research papers can attain currently. Computational reproducibility *from raw data* is required for papers that are reproducible at Level 6 and above.

- **Level 6 (L6):** Cleaning code is partially available, but raw data is not. Possible improvements include: completing cleaning code (+CC) and/or raw data (+RD).
- **Level 7 (L7):** Cleaning code is available and complete, but raw data is not. Possible improvements include: adding raw data (+RD).
- **Level 8 (L8):** Cleaning code is available and complete, and raw data is partially available. Possible improvements include: adding raw data (+RD).

- **Level 9 (L9):** All the materials (raw data, analytic data, cleaning code, and analysis code) are available. The analysis code produces the same output as presented in the paper (CRA). However, the cleaning code does not run or produces different results that those presented in the paper (not CRR). Possible improvements include: debugging the cleaning code (DCC).
- **Level 10 (L10):** All the materials are available and produce the same results as presented in the paper with minimal effort, starting from the analytic data (yes CRA) or the raw data (yes CRR). Note that Level 10 is aspirational and may be very difficult to attain for most research published today.

The following figure summarizes the different levels of computational reproducibility (for any given output). For each level, there will be improvements that have been made ( ) or can be made to move up one level of reproducibility (-).

Levels of Computational Reproducibility
(P denotes “partial”, C denotes “complete”)

Availability of materials, and reproducibility

Analysis Code

Analysis Data

CRA

Cleaning Code

Raw Data

CRR

P

C

P

C

P

C

P

C

L1: No materials

–

–

–
–
–
–
–
–
–
–

L2: Only code

–
–
–
–
–
–
–
–

L3: Partial analysis data & code

–
–
–
–
–
–
–

L4: All analysis data & code

–
–
–
–
–
–

L5: Reproducible from analysis

–
–
–
–
–

L6: Some cleaning code

–
–
–

–

L7: All cleaning code

–

–

–

L8: Some raw data

–

–

L9: All raw data

–

L10: Reproducible from raw data

Levels of Computational Reproducibility
(P denotes "partial", C denotes "complete")

	Availability of materials, and reproducibility									
	Analysis Code		Analysis Data		CRA	Cleaning Code		Raw Data		CRR
	P	C	P	C		P	C	P	C	
L1: No materials	-	-	-	-	-	-	-	-	-	-
L2: Only code			-	-	-	-	-	-	-	-
L3: Partial analysis data & code.				-	-	-	-	-	-	-
L4: All analysis data & code					-	-	-	-	-	-
L5: Reproducible from analysis						-	-	-	-	-
L6: Some cleaning code							-	-	-	-
L7: All cleaning code								-	-	-
L8: Some raw data									-	-
L9: All raw data										-
L10:Reproducible from raw data										

You may disagree with some of the levels outlined above, particularly wherever subjective judgment may be required. If so, you are welcome to interpret the levels as unordered categories (independent from their sequence) and suggest

improvements using the "Edit" button above (top left corner if you are reading this document in your browser).

Adjusting Levels To Account for Confidential/Proprietary Data

A large portion of published research in economics uses confidential or proprietary data, most often government data from tax records or service provision and what is generally referred to as *administrative data*. Since administrative and proprietary data are rarely publicly accessible, some of the reproducibility levels presented above only apply once modified. The underlying theme of these modifications is that when data cannot be provided, you can assign a reproducibility score based on the level of detail in the instructions for accessing the data. Similarly, when reproducibility cannot be verified based on publicly available materials, the reproduction materials should demonstrate that a competent and unbiased third party (not involved in the original research team) has been able to reproduce the results.

- **Levels 1 and 2** can be applied as described above.
- **Adjusted Level 3 (L3*):** All analysis code is provided, but only partial instructions on how to access the ***analysis data*** are available. This means that the authors have provided some, but not all, of the following information:
    a. *Contact information*, including name of the organization(s) that provides access to the data and contact information of at least one individual.
    b. *Terms of use*, including licenses and eligibility criteria for accessing the data, if any.
    c. *Information on data files (meta-data)*, including the name(s) and number of files, file size(s), relevant file version(s), and number of variables and observations in each file. Though not required, other relevant information may be included, including a description dataset dictionary, summary statistics, and synthetic data (fake data with the same statistical properties as the original data)
    d. *Estimated costs for access*, including monetary costs such as fees and licences required to access the data, and non-monetary costs such as wait times and specific geographical locations from where researchers need to access the data.
- **Adjusted Level 4 (L4*):** All analysis code is provided, and complete and detailed instructions on how to access the *analysis data* are available.
- **Adjusted Level 5 (L5*):** All requirements for Level 4* are met, and the authors provide a certification that the output can be reproduced from the analysis data (CRA) by a third party. Examples include a signed letter by a disinterested reproducer or an official reproducibility certificate from a certification agency for data and code (e.g., see cascad).

- **Levels 6 and 7** can be applied as described above.
- **Adjusted Level 8 (L8*):** All requirements for Level 7* are met, but instructions for accessing the ***raw data*** are incomplete. Use the instructions described in Level 3 above to assess the instructions' completeness.
- **Adjusted Level 9 (L9*):** All requirements for Level 8* are met, and instructions for accessing the ***raw data*** are complete.
- **Adjusted Level 10 (L10*):** All requirements for Level 9* are met, and a certification that the output can be reproduced from the raw data is provided.

Levels of Computational Reproducibility with Proprietary/Confidential Data
(P denotes "partial", C denotes "complete")

Availability of materials, and reproducibility

Analysis Code

Instr. Analysis Data

CRA

Cleaning Code

Instr. Raw Data

CRR

P

C

P

C

P

C

P

C

L1: No materials

–

–

–

–

–

–

–

–

–

–

L2: Only code

–

–

–

–

–

–

–

–

L3: Partial analysis data & code

–

–

–

–

–

–

–

L4*: All analysis data & code

–

–

–

–

–

–

L5*: Proof of third party CRA

–

–

–

–

–

L6: Some cleaning code

–

–

–

–

L7: All cleaning code

–

–

–

L8*: Some instr. for raw data

–

–

L9*: All instr. for raw data

–

L10*: Proof of third party CRR

Levels of Computational Reproducibility with Proprietary/Confidential Data
(P denotes "partial", C denotes "complete")

	Availability of materials, and reproducibility										
	Analysis Code		Instr. Analysis Data		CRA	Cleaning Code		Instr. Raw Data		CRR
	P	C	P	C		P	C	P	C	
L1: No materials	-	-	-	-	-	-	-	-	-	-
L2: Only code			-	-	-	-	-	-	-	-
L3: Partial analysis data & code.				-	-	-	-	-	-	-
L4*: All analysis data & code					-	-	-	-	-	-
L5*: Proof of third party CRA						-	-	-	-	-
L6: Some cleaning code							-	-	-	-
L7: All cleaning code								-	-	-
L8*: Some instr. for raw data									-	-
L9*: All instr. for raw data										-
L10*:Proof of third party CRR										

2.3.2 Reproducibility dimensions at the paper level

In addition to the output-specific assessment and improvement of computational reproducibility, several practices can facilitate reproducibility at the level of the overall paper. You can read about such practices in greater detail in the next chapter, dedicated to Stage 3: *Improvements*. In this Assessment section, you should only verify whether the original reproduction package made use of any of the following:

- Master script that runs all steps
- Readme file
- Standardized file organization

- Version control
- Open source (statistical) software

- Dynamic document

- Computing capsule (e.g. CodeOcean, Binder, etc.)

Congratulations! You have now completed the *Assessment* stage of this exercise. You have provided a concrete building block of knowledge to improve understanding of the state of reproducibility in Economics.

Please continue to the next section where you can help improve it!

```
library(tidyverse)
library(knitr)
library(kableExtra)
```

Chapter 3

Improvements

After assessing the paper's reproducibility package, you can start proposing ways to improve its reproducibility. Making improvements provides an opportunity to gain a deeper understanding of the paper's methods, findings, and overall contribution. Each contribution can also be assessed and used by the wider ACRE community, including other students and researchers using the ACRE platform.

As with the *Assessment* section, we recommend that you first focus on one specific display item (e.g., "Table 1"). After making improvements to this first item, you will have a much easier time translating those improvements to other ones.

Use Survey 2 to record your work as part of this step.

3.1 Types of output-level improvements

3.1.1 Adding raw data: missing files or metadata

Reproduction packages often do not include all original raw datasets. To obtain any missing raw data, or information about them, follow these steps:

1. Identify the missing file. During Assessment, you identified all data sources from the paper's body and appendices (column `data_source` in this standarized spreadsheet). However, some data sources (as collected by the original investigators) might be missing one or more data files. You can sometimes find the specific name of those files by looking at the beginning of the cleaning code scripts. If you find the name of the file, record it in the `known_,missing` field of the same spreadsheet as above. If not, record it as "Some/All" in the `known_,missing` field of the for each specific data source.

2. Verify whether this file (or files) can be easily obtained from the web.
    - 2.1 - If yes: obtain the missing files and add them to the reproduction package. Make sure to obtain permission from the original author to publicly share this data. See tips for communication for relevant guidance.

    - 2.2 - If no: proceed to step 3.

3. Use the ACRE database to verify whether there have been previous attempts to contact the authors regarding this paper, about this specific missing raw files.

4. Contact the original authors and politely request the original materials. Be mindful of the authors' time, and remember that the paper you are trying to reproduce was possibly published at a time when standards for computational reproducibility were different. See tips for communication for sample language on how to approach the authors for this specific scenario.

5. If the datasets are not available due to legal or ethical restrictions, you can still improve the reproduction package by providing detailed instructions for future researchers to follow, including contact information and possible costs of obtaining the raw data.

In addition to trying to obtain the raw data, you can also contribute by obtaining missing analytic data.

3.1.2 Adding missing analytic data files

Analytic data can be missing for two reasons: (i) raw data exists, but the procedures to transform it into analytic data are not fully reproducible, or (ii) some or all raw data is missing, and some or all analytic data is not included in the original reproduction package. To obtain any missing analytic data, follow these steps:

1. Identify the specific name of the missing data set. Typically this information can be found in some of the analysis code that calls the data to perform an analysis (eg `analysis_data_03.csv`).

2. Verify that the data cannot be obtained by running the data cleaning code over the raw data.

3. Use the ACRE database to verify if previous attempts have been made to contact the authors about this data.

4. Contact the authors and request the specific data set.

3.1.3 Adding missing analysis code

Analysis code can be added when analytic data files are available, but some or all methodological steps are missing from the code. In this case, follow these steps:

1. Identify the specific line or paragraph in the paper that describes the analytic step that is missing from the code (e.g., "We impute missing values to…" or "We estimate this regression using a bandwidth of…").
2. Identify the code file and the approximate line in the script where the analysis can be carried out. If you cannot find the relevant code file, identify its location relative to the main folder using the the steps in the reproduction diagram.
3. Use the ACRE database to verify if previous attempts have been made to contact the authors about this issue.
4. Contact the authors and request the specific code files.
5. If step #4 does not work, we encourage you to attempt to recreate the analysis using your own interpretation of the paper, and making explicit your assumptions when filling in any gaps.

3.1.4 Adding missing data cleaning code

Data cleaning (processing) code might be added when steps are missing in the creation or re-coding of variables, merging, subsetting of the data sets, or other steps related to data cleaning and processing. You should follow the same steps you used when adding missing analysis code (1-5).

3.1.5 Debugging analysis code

Whenever code is available in the reproduction package, you should be able to debug those scripts. There are four types of debugging that can improve the reproduction package:

- *Code cleaning:* Simplify the instructions (e.g., by wrapping repetitive steps in a function or a loop) or remove redundant code (i.e., old code that was commented out) while keeping the original output intact.
- *Performance improvement:* Replace the original instructions with new ones that perform the same tasks but take less time (e.g., choose one numerical optimization algorithm over another, but obtain the same results).
- *Environment set up:* Modify the code to include correct paths to files, specific versions of software, and instructions to install missing packages or libraries.

- *Correcting errors:* A coding error will occur when a section of the code in the reproduction package executes a procedure that is in direct contradiction with the intended procedure expressed in the documentation (i.e., paper or code comments). For example, an error will happen if the paper specifies that the analysis is performed on a population of males, but the code restricts the analysis to females only. Please follow the ACRE procedure to report coding errors.

3.1.6 Debugging cleaning code

Follow the same steps that you did to debug the analysis code, but report them separately.

3.1.7 Reporting results

Track all the different types of improvements you make and record in this standarized spreadsheet with the following structure:

Level-specific quality improvements: add data/code, debug code

output_name

imprv

description_of_added_files

lvl

table 1

+AD

ADD EXAMPLES

5

table 1

+RD

ADD EXAMPLES

5

table 1

DCC

ADD EXAMPLES

5

figure 1

+CC

6

figure 1

DAC

6

inline 1

DAC

8

…

…

…

…

```
Level-specific quality improvements: add data/code, debug code.

   | output_name | imprv | description_of_added_files        | lvl |
   |-------------|-------|-----------------------------------|-----|
   | table 1     | +AD   |        ADD EXAMPLES               |  5  |
   | table 1     | +RD   |        ADD EXAMPLES               |  5  |
   | table 1     | DCC   |        ADD EXAMPLES               |  5  |
   | figure 1    | +CC   |                                   |  6  |
   | figure 1    | DAC   |                                   |  6  |
   | inline 1    | DAC   |                                   |  8  |
   | ...         | ...   | ...                               | ... |
```

3.2 Types of paper-level improvements

There are at least six additional improvements you can make to improve a paper's overall reproducibility. These additional improvements can be applied across all reproducibility levels (including level 10).

1. Set up the reproduction package using version control software, such as Git.
2. Improve documentation by adding extensive comments to the code.
3. Integrate the documentation with code by adapting the paper into a literate programming environment (e.g., using Jupyter notebooks, RMarkdown, or a Stata Dynamic Doc).
4. If the code was written using a proprietary statistical software (e.g., Stata or Matlab), re-write it using an open source statistical software (e.g., R, Python, or Julia).
5. Re-organize the reproduction package into a set of folders and sub-folders that follow standardized best practices, and add a master script that

executes all the code in order, with no further modifications. See AEA's reproduction template.

6. Set up a computing capsule that executes the entire reproduction in a web browser without needing to install any software. For examples, see Binder and Code Ocean.

3.2.1 Reporting improvements

You will be asked to provide this information in the Assessment and Improvement Survey.

```
library(tidyverse)
library(knitr)
library(kableExtra)
temp_eval <- FALSE
```

Chapter 4

Checking for Robustness

Once you have assessed and improved the computational reproducibility of a paper, you can assess the quality of different analytical choices by including new robustness checks in addition to those included in the original paper. We use the term **robustness checks** to describe any possible change in a computational choice, both in data analysis and data cleaning, and its subsequent effect on the main estimates of interest. The universe of robustness checks can be very large or potentially infinite. The focus should be on the set of **reasonable specifications** (Simonsohn et. al., 2018), defined as (1) sensible tests of the research question, (2) expected to be statistically valid, and (3) not redundant with other specifications in the set.

The addition of new robustness checks will depend on the current level of reproducibility. E.g., for claims supported by display items reproducible at level 0-1, it is not possible to perform any other robustness checks in addition to what is already in the paper (**?**, include a brief explanation why: because...). It may be possible to perform additional robustness checks for claims supported by display items reproducible at levels 2-4, but not using the specific estimates declared in *Stage 1: Scoping* because the display items are not computationally reproducible from analysis data (CRA). It is possible to include additional robustness checks to validate the core conclusion of a claim based on a display item reproducible at level 5. Finally, a claim associated with display items reproducible at level 6 and above allows for robustness checks that involve variable definitions and data manipulations. When checking the robustness to a new variable definition, reproducers will also have the possibility of testing how the main estimate changes under an alternative variable definition *and* an alternative core analytical choice. (**?**, please verify whether this is what you meant to say in the last 2 sentences)

Going back to our diagram that represents the multiple parts of a paper (1), the robustness section begins at the claim level. For a given claim, there will

be several specifications presented in the paper, one of which is identified by the authors (or yourself, in the absence of one designated by the authors) as the main or preferred specification. Identify which display item contains this specification and refer to the reproduction tree to identify the code files where you can potentially modify a computational choice. Using the example tree discussed in the *Assessment* stage, we can remove the data files for simplicity and obtain the following:

```
table1.tex (contains preferred specification of a given claim)
    |___[code] analysis.R
            |___[code] final_merge.do
                    |___[code] clean_merged_1_2.do
                    |       |___[code] merge_1_2.do
                    |               |___[code] clean_raw_1.py
                    |               |___[code] clean_raw_2.py
                    |___[code] clean_merged_3_4.do
                            |___[code] merge_3_4.do
                                    |___[code] clean_raw_3.py
                                    |___[code] clean_raw_4.py
```

This simplified tree gives you a list of potential files where you could test different reasonable specifications. Here we suggest two types of contributions to robustness checks: i) mapping the universe of robustness checks and ii) testing reasonable specifications. Both contributions should be recorded in the ACRE platform referring to files in a specific reproduction package.

4.1 Mapping the universe of robustness checks

Analytical choices in data cleaning code - Variable definition
- Data sub-setting
- Data reshaping (merge, append, long/gather, wide/spread)
- Others (specify as "processing - other") **Analytical choices in analysis code**
- Regression function (link function)
- Key parameters (tuning, tolerance parameters, etc.)
- Controls
- Adjustment of standard errors
- Choice of weights
- Treatment of missing values
- Imputations - Other (specify as "methods - other")

Once finished, transcribe all of the information on analytical choices into a dataset (the ACRE platform will allow for easier recording once deployed). For the `source` field type *"original"* whenever the analytical choice is identified for the first time, and `file_name-line number` every subsequent time when the same analytical choice is applied (for example if an analytic choice is identified

for the first time in line #103 and for the second time in line #122 their respective values for the `source` field should be `original` and `code_01.do-L103`, respectively).

For each analytical choice recorded, add the specific choice that the paper used, and describe what other alternatives could have been used. The resulting database should have the following structure:

entry_id	file_name	line_number	choice_type	choice_value	choice_range	Source
1	code_01.do	73	data subsetting	males	males, female,	original
2	code_01.do	122	variable definition	income = wages + capital gains	wages, capital gains, gifts	"code_01.do-L103"
3	code_05.do	123	controls	age, income, education	age, income, education, region	original
...	...	...	...	...	...	...

The advantage of this type of contribution is that you are not required to have an in-depth knowledge of the paper and its methodology to contribute. This allows you to potentially map several code files, achieving a broader understanding of the paper. The disadvantage is that you are not expected to test alternative specifications.

4.2 Proposing a specific robustness check

When performing a specific robustness test, follow these steps:

1. Search in the mapping database (previous section) (**?**, is this referring to the reproduction tree diagram?) and record the identifier(s) corresponding to the analytical choice to test (`entry_id`). If there is no entry corresponding for the specific lines, please create one.
2. Propose a specific variation to this analytical choice.
3. Discuss whether you think this variation is sensible, specifically in the context of the claim tested (e.g. does it make sense to include exclude low-income Hispanics from the sample?).
4. Discuss how this variation could affect the validity of the results (e.g. likely effects on omitted variable bias, measurement error, change in the Local Average Treatment Effects for the underlying population).
5. Confirm that test is not redundant with other tests in the paper/robustness exercise.

6. Report the results from the robustness check (new estimate, standard error, and units).

The advantage of this approach is that it allows for an in-depth inspection of a specific section of the paper. The main limitation is that justifying sensibility and validity (and non-redundancy, to some extent) requires a much deeper understanding of the topic and the methods of the paper, making it less feasible for undergraduate students or graduates with only a general interest in the paper. (**?**, what does it mean to have only a general interest in the paper?)

Chapter 5

Guidance for a Constructive Exchange Between Reproducers and Original Authors

The purpose of this chapter is to facilitate constructive and respectful communication between reproducers and original authors. Exchanges that contain charged or adversarial language can damage professional relationships and hamper scientific progress. Janz and Freese (2019) articulate two important steps reproducers can take to ensure their interactions with original authors are constructive. We provide a summary below and encourage you to follow this guidance. Remember the **golden rule of reproductions** (and replications): *treat others and their work, as you would like others to treat you and your work!*

1. Carefully and transparently plan your study.

a. Clearly state that you are conducting a reproduction of the original work.

b. Explain why you have chosen this study.
c. Explain how "far" your results must deviate from the original work before claiming that the study could not be reproduced. Engage deeply with the substantive literature to ensure that your interpretation of differences between the original and reproduction is thorough and acceptable to other authors in the field.

2. Use professional and sensitive language. Discuss potential discrepancies between your work and the original paper just like you would have done for your own work.

a. Avoid binary judgments and statements like "failed to reproduce." Clearly state which results reproduced and which did not (e.g., "we successfully reproduced X, but failed to reproduce Y") unless you uncover apparent scientific misconduct (e.g. see Broockman, Kalla and Aronow, 2015). [FH: last sentence seems broken]
b. Talk about *the study, not the author*, to avoid making it personal. Make clear what the positive contribution of the original article is. Consider sending a copy of your reproduction report to the original authors.

c. Discuss what your reproduction contributes to the literature, and refrain from claiming to give the final answer to the question.

d. For papers published five or more years ago, be mindful that norms for reproducibility have evolved since then.

e. Remember, *the goal is not to criticize previous work or hunt for errors, but to move the literature forward!*

To help put these recommendations into practice, we provide template language for common scenarios that reproducers and authors may encounter in their interactions.

While we hope that you find these useful, note that they are *only recommendations*, and you are welcome to modify them based on the context and needs of your specific project. Feel free to contact us if you need more guidance or would like to provide feedback on these materials (acre@berkeley.edu).

5.1 For Reproducers Contacting the Authors of the Original Study

Consider the following *before* you contact the original author:

1. Carefully read all footnotes, appendices, tables, captions, etc. to learn if, how, and where reproduction materials are provided. Follow this Data and Code Guidance to determine whether you have everything before you start. A few things to consider:

    - A *Readme* file, if available, would be a good place to start. All papers published in AEA journals after July 2019 should have such documentation.

    - Check whether there are any restrictions to accessing the data or code, and whether there are instructions on how to access these files for the purpose of reproduction.

2. If a reproduction package is not readily available in the location where the article is published (e.g., the journal website), check the authors' websites,

Dataverse profiles, the ICPSR Publications Related Archive, and other relevant archives and/or data repositories.

3. If steps 1 and 2 don't work, contact the corresponding author (copying the co-authors, if any), consolidating your requests into as few emails as possible. In your email, make sure to include the following details:

    - Basic information about the paper being reproduced (include title, version, date, and a DOI link (or just a URL));
    - Context for the reproduction (as part of a class exercise, thesis, etc.) and a notice that the outcome will be recorded in the ACRE reproducibility database;
    - Items from the reproduction package that are missing, as well as locations where you had (unsuccessfully) searched for them;
    - Use plan: Will the materials be used exclusively for this project? Ask for permission to share the data publicly.
    - Right to consultation and results: Will you share the outcome of the reproduction exercise with the original authors?
    - A deadline to respond (we suggest at least two weeks).

4. Follow up if you don't get a response within two weeks (or whatever deadline you set), and include any details or clarifications that were left out in your first email.

5. Record the outcome of your interaction with the original author in the ACRE database. You can qualify the outcome as one of the following:

    - A *complete* reproduction package was provided
    - An *incomplete* reproduction package was provided. You can also select one of the following reasons:
        - Data is of sensitive, confidential, or proprietary character and cannot be shared;
        - Data is of sensitive, confidential, or proprietary character, but access instructions were provided.
    - Author *refused* to share the reproduction package
    - Author *did not respond* (including after a reminder was sent) within 4 weeks after the initial request.

5.1.1 Contacting the original author(s) when there is no reproduction package

*Template email:**

Subject: Reproduction package for `["Title of the paper"]`

Dear Dr. `[Lastname of Corresponding Author]`,

I am contacting you to request a reproduction package for your paper titled `[Title]` which was published in `[Journal]` in `[year]` (vol `[volume]`, no. `[no.]`), `[link]`. A reproduction package may contain (raw and/or analytic) data, code, and other documentation that makes it possible to reproduce paper. Would you be able to share any of these items?

I am a `[graduate student/postdoc/other position]` at `[Institution]`, and I would like to reproduce the results, tables, and other figures using the reproduction materials mentioned above. I have chosen this paper because `[add context for why you want to reproduce this particular paper using neutral language (e.g., "This is a seminal paper in my field"), avoiding any statements that would put the respondent on the defensive]`. Unfortunately, I was not able to locate any of these materials on the journal website, Dataverse `[or other data and code repositories]`, or in your website.

I will record the result of my reproduction attempt on the Advancing Computational Reproducibility in Economics (ACRE) platform, which is an open-source repository to systematically source, collect and present the results of verifications of computational reproducibility of published work in economics. With your permission, I will also record the materials you share with me, which would allow access for other reproducers and avoid repeated requests directed to you. Please let me know if there are any legal or ethical restrictions that apply to all or parts of the reproduction materials so that I can take that into consideration during this exercise.

In addition to your response above, would you be available to respond future (non-repetitive) inquieries from me or other reproducers conducting an ACRE excercise? Though your cooperation with mine and/or future request would be extremely helpful, please note that you are *not required to comply.*

Since I am required to complete this project by `[date]`, I would appreciate your response by `[deadline]`.

Let me know if you have any questions. Please also feel free to contact my supervisor/instructor `[Name (email)]` for further details on this exercise. Thank you in advance for your help!

Best regards,
`[Reproducer]`

5.1.2 Contacting the original author(s) to request specific missing items of a reproduction package

Template email:

> **Subject:** Reproduction materials for `["Title of the paper"]`
>
> Dear `[Title] [Lastname of Corresponding Author]`,
>
> I am contacting you regarding reproduction materials for your paper titled `[Title]` which was published in `[Journal]` in `[year]` (vol `[volume]`, no. `[no.]`), `[link]`. I am a `[graduate student/postdoc/other position]` at `[Institution]`, and I'm working to reproduce this paper as part of a class exercise. `[Add context for why you want to reproduce this particular paper using neutral language (e.g., "This is a seminal paper in my field"), avoiding any statements that would put the respondent on the defensive]`.
>
> To help me reproduce the paper in full, I hope that you can share the following items: [`list items missing from reproduction package, preferably bulleted if more than one (e.g., raw/analytic data, code, protocols for conducting the experiment, etc.)`]. I have already searched `[locations where you searched for items, with links provided]`, but I was unable to locate the items. You can be assured that I will not share any of the materials without your permission, and I will use them exclusively for the purpose of this exercise. Let me know if there are any legal or ethical restrictions that apply to all or parts of the reproduction materials so that I can take that into consideration during this exercise.
>
> Note that I will record the outcome of my reproduction on the Accelerating Computational Reproducibility in Economics (ACRE) platform, an online catalog of reproduction projects in economics. ACRE is hosted by the Berkeley Initiative for Transparency in the Social Sciences (BITSS). Let me know if you would like me to share the outcome of my exercise with you, and whether you are interested in providing a response.
>
> Since I am required to complete this project by `[date]`, I would appreciate your response by `[deadline]`.
>
> Let me know if you have any questions. Please also feel free to contact my supervisor/instructor `[Name (email)]` for further details on this exercise. Thank you in advance for your help!
>
> Best regards,
> `[Reproducer]`

5.1.3 Asking for additional guidance when some materials have been shared

Note: Even when a corresponding author has shared a reproduction package, you may still run into challenges in interpreting or executing the materials. That shouldn't discourage you from asking the corresponding author to provide clarifications or share missing materials. As in the previous scenario described above, demonstrate that you made an honest effort to reproduce the work using the available resources and try to consolidate your requests into as few emails as possible.

Template email:

> **Subject:** Clarification for reproduction materials for `["Title of the paper"]`
>
> Dear Dr. `[Lastname of Corresponding Author]`,
>
> Thank you for sharing the materials. They have been immensely helpful for my work.
>
> Unfortunately, I ran into a few issues as I delved into the reproduction exercise, and I think your guidance would be helpful in resolving them. `[Describe the issues and how you have tried to resolve them. Describe whatever files or parts of the data or code are missing. Refer to examples 1 and 2 below for more details]`.
>
> Thank you in advance for your help.
>
> Best regards,
> `[Reproducer]`

1: An example of well described issues:

> Specifically, I am attempting to reproduce OUTPUT X (e.g., table 1, figure 3). I found that the following components are required to reproduce to reproduce OUTPUT X:

```
OUTPUT X
     [code] formatting_table1.R
         output1_part1.txt
       |     [code] output_table1.do
       |         [data] analysis_data01.csv
       |            [code] data_cleaning01.R*
       |               [data] UNKNOWN
         output1_part2.txt
             [code] output_table2.do
                 [data] analysis_data02.csv
                    [code] data_cleaning02.R
                       [data] admin_01raw.csv*
```

> I have marked with an asterix (*) the items that I could not find in the reproduction materials: **data_cleaning01.R** and **admin_01raw.csv**. After accessing these files, I will also be able to identify the name of the raw data set required to obtain output1_part1.txt. This is to let you know that I may need to contact you again if I cannot find this file (labeled as UNKNOWN above) in the reproduction materials.
>
> I understand that this request will require some work for you or somebody in your research group, but I want to assure you that I will add these missing files to the reproduction package for your paper on the ACRE platform. **Doing this will ensure that you will not be asked twice for the same missing file.**

2. An example of poorly described issues:

> Your paper does not reproduce. I have tried for several hours now, and can't get the DO files to run. Could you please share all the missing reproduction materials? Data and code sharing are a basic principle of open science, so I am confident that you will do the right thing.

5.1.4 Response when the original author has refused to share due to *undisclosed reasons*

Note: You can also use this template if a corresponding author has not submitted a response after two or more follow-up emails.

Template email:

> **Subject:** Re: Reproduction materials for '["Title of the paper"]
>
> Dear Dr. `[Last Name of Corresponding Author]`,
>
> Thank you for considering my request. I will try to reproduce the paper using the available materials, and will record the missing items accordingly on the ACRE platform. l will also post my assessment of the reproducibility of the paper in its current form based on the ACRE reproducibility scale.
>
> Let me know if you have any questions.
>
> Best regards,
> `[Reproducer]`

5.1.5 Response when the original author has refused to share due to legal or ethical restrictions of the data

Template email:

Subject: Re: Reproduction materials for `["Title of the paper"]`

Dear Dr. `[Last Name of Corresponding Author]`,

Thank you for your response and for clarifying the terms of use of the reproduction materials.

Though I understand that you are unable to share the raw data, there may be alternative steps you can take that would help me improve the reproducibility of your paper. These include:

1. Sharing the analytic version of the data (the version of the dataset that was used in the final version of your paper);

2. Providing a public description of the steps other researchers can follow to request access to the raw data or materials, including an estimate of the costs and the duration of the process. Find examples of data availability statements for proprietary or restricted-access data here; and

3. Providing access to all data and materials for which the constraints do not apply.

Based on my assessment, your paper would currently rank at `[level X]` on the ACRE reproducibility scale, however *this score can be easily improved*. Being able to provide analytic data would elevate the reproducibility of your paper to `[level Y]`. Providing public instructions on how other parties can access the data would further elevate its reproducibility to `[level Z]`.

I would be happy to help if you are interested in taking any of the steps I outlined above. Let me know if that would be helpful.

Thank you for your help!

Best regards,
`[Reproducer]`

5.1.6 Contacting the original author to share the results of your reproduction exercise

Note: Reporting the results of reproductions is probably the most contentious part of the process, particularly in instances where the reproducer is not able to fully reproduce the paper or finds significant deviations from the original work. However, if the reproduction can correctly identify the sources of such deviations, it may be viewed as an improved version of the original work.

Regardless of the outcome of the reproduction exercise, the guidance from the introduction of this chapter still stands here: *reproduce the work of others as*

you would like for others to reproduce yours, and make sure that is reflected in how you discuss any discrepancies between your and the original work.

Template email:

> **Subject:** Reproducibility Assessment of `["Title of the paper"]`
>
> Dear Dr. `[Last Name of Corresponding Author]`,
>
> Thank you for your support throughout my project as I worked to verify and advance the reproducibility of `[Paper]`. I'm writing now to share the results of my project with you and invite your feedback.
>
> The results of each step of my exercise, include i) Assessment, ii) Improvements, iii) Robustness Checks, (and iv) Extensions, if applicable).
> '[Include the following items in the body of your email:
>
> - Briefly describe which parts of the paper you tried to reproduce (e.g. a specific estimate, a table, etc.).
>
> - Within the scope of your reproduction, describe exactly which items you managed to reproduce.
>
> - Discuss the differences you observed between the results of your reproduction and the original work, and demonstrate that you did your due diligence in trying to reproduce the item. Remember that it is more constructive to discuss discrepancies, differences or deviations, rather than errors, mistakes or failures, and *always talk about the work – not the author!*
>
> - Use sensitive language when presenting discrepancies, e.g. "Unfortunately, I found X, which differs from the Y result in the original paper…". Be cognizant of any potential limitations of your work, and explain how you have tried to address them – that way you will proactively address potential criticism!
>
> - Describe how you tried to improve, the reproducibility of the paper. If some of the improvements are based on discretionary judgment (e.g. file organization or code commenting), try to explain why you think they are an upgrade over the original work. If you didn't make improvements, point out some concrete steps that the author(s) can take to improve the reproducibility of the section you reproduced.]'
>
> I look forward to your questions, comments, and suggestions on what I laid out above. As discussed previously, I will record the outcomes of my exercise, along with the improvements, on the ACRE platform.

Best regards,
`[Reproducer]`

5.1.7 Responding to hostile responses from original authors

Note: Planning your study carefully and transparently, and using professional and sensitive language are the best ways to ensure that the interaction will be beneficial to both you and the original author. However, unpleasant interactions may happen despite your best efforts, and can range anywhere from dismissive comments to bullying, discrimination, and harassment.

5.1.7.1 Dismissive comments

In cases of dismissive comments, the best course of action may be to simply thank the author for their response and continue with the exercise.

Template email:

> **Subject:** Re: Reproduction materials for '["Title of the paper"]
>
> Dear Dr. `[Last Name of Corresponding Author]`,
>
> Thank you for your response. I will work to reproduce using the available materials, and will record my results accordingly on the ACRE platform. l will also post my assessment of the reproducibility of the paper in its current form based on the ACRE reproducibility scale.
>
> Let me know if you have any questions.
>
> Best regards,
> `[Reproducer]`

5.1.7.2 Harassment and/or discrimination

The AEA and other economic societies have strict policies against harassment and discrimination. Here are some of the behaviors that the AEA Policy on Harassment and Discrimination has listed as unacceptable, and could emerge in a hostile exchange regarding a reproduction:

- Intentionally intimidating, threatening, harassing, or abusive actions or remarks (both spoken and in other media)
- Prejudicial actions or comments that undermine the principles of equal opportunity, fair treatment, or free academic exchange
- Deliberate intimidation, stalking, or following
- Real or implied threat of physical harm.

Here are a some steps you can take if you believe you have experienced bullying, discrimination or harassment:

- **File a complaint with the AEA Ombudsperson.** Any AEA member can file a complaint (you can also join the AEA solely for the purpose of filing a report). The person about whom you are making the complaint need not be an AEA member. A non-AEA member can also file a report if the act of harassment or discrimination was committed by an AEA member or in the context of an AEA-sponsored activity. Learn more about the process here.
- **File a report with your institution's office for the prevention of harassment & discrimination.** US-based institutions have internal mechanisms that allow students and faculty to seek support in cases of discrimination and harassment on the basis of race, color, national origin, gender, age, or sexual orientation/identity, including allegations of sexual harassment and sexual violence. Formal titles of this office vary across institutions, but common names include "Office for the Prevention of Harassment and Discrimination" (in institutions that are part of the University of California system), "Office of Equity and Title IX", etc.
- **Contact your institution's Ombudsperson/Ombuds Office.** If you believe that you have experienced academic bullying or other forms of disrespectful behavior that fall outside the scope of harassment and/or discrimination as described above, you should know that university ombuds officers are a confidential, impartial resource to discuss your concerns and learn about potential next steps available in your case.
- **Access mental health services at your institution.** Many universities offer short-term Counseling & Psychological Services (CAPS) for academic, career, and personal issues.
- **Ask for support from your academic supervisor.** If you are unsure on how to proceed, consult your academic supervisor on whether continuing the exercise would be appropriate.

5.2 For Original Authors Responding to Requests from Reproducers

[under development]

5.2.1 Responding to a repeated request

[TO DO]

5.2.2 Acknowledging that some information is missing

[TO DO]

5.2.3 Acknowledging that some material is still embargoed for future research

[TO DO]

5.2.4 Responding to incomplete/aggressive requests from reproducer

Chapter 6

Reproduction Diagrams

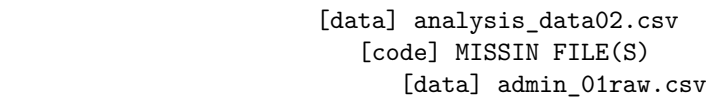
6.1 Different Scenarios

6.1.1 Complete

```
table 1
    [code] formatting_table1.R
        output1_part1.txt
      |     [code] output_table1.do
      |         [data] analysis_data01.csv
      |            [code] data_cleaning01.R
      |               [data] survey_01raw.csv
        output1_part2.txt
            [code] output_table2.do
                [data] analysis_data02.csv
                   [code] data_cleaning02.R
                      [data] admin_01raw.csv
```

6.1.2 Raw data and analytic data, but cleaning code is missing.

```
table 1
    [code] formatting_table1.R
        output1_part1.txt
      |     [code] output_table1.do
      |         [data] analysis_data01.csv
      |            [code] MISSING FILE(S)
      |               [data] survey_01raw.csv
        output1_part2.txt
            [code] output_table2.do
```


[data] analysis_data02.csv
[code] MISSIN FILE(S)
[data] admin_01raw.csv

Chapter 7

Additional resources

- **Coding errors:** A coding error will occur when a section of the code, of the reproduction package, executes a procedure that is in direct contradiction with the intended procedure expressed in the documentation (paper or comments of the code). For example an error happens if the paper specify that the analysis is perform on the population of males, but the code restricts the analysis to females only. Please follow the ACRE procedure to report coding errors.

Create a section with short summaries of great resources for comp. repro and invite reader to contribute.

7.1 Some summaries

7.1.1 Summary on reproducible workflow (Chapter 11) from Christensen et al. (2019):

- TODO

7.2 Links

TODO: Add and classify

- Project TIER
- IDB's cheatsheet for transparency, reproducibility and ethics
- Lars Vilhuber LDI's Wiki for Reproducibility. Particularly this section.
- World Bank DIME's Wiki for transparent and reproducible research.

- Dynamic documents in R, Python and Stata

- Git resources:
  - Jenny Bryan's book and video

  - Github learning lab
  - Udacity's intro

  - Git for poets

  - Combining GitHub and Dropbox

  - Atlassian intro to Git
  - Software Carpentry tutorial from the command line
- Open Science Framework (OSF)
- R for Stata users

Chapter 8

Contributions

8.1 Contributing feedback on these guidelines

The ACRE project welcomes feedback from participants and the wider social science community. If you wish to provide feedback on specific chapters or sections, click the "edit" icon at the top of this page (this will prompt you to sign into or create a GitHub account), after which you'll be able to suggest changes directly to the text. Please submit your suggestions using the "create a new branch and start a pull request" option and provide a summary of the changes you've proposed in the description of the pull request. The ACRE project team will review all suggested changes and decide whether to "push" them to the guidelines document or not. For more general feedback, please contact ACRE@berkeley.edu.

Major contributions to these guidelines will be acknowledged below. The ACRE project employs the Contributor Roles Taxonomy (CRediT). Major contributions are defined as any pushed revisions to the guideline language or source code beyond corrections of spelling and grammar.

8.2 List of Contributors: Guidelines content and source code:

(in alphabetical order) - Aleksandar Bogdanoski – Funding acquisition, Project administration, Writing (original draft), Writing (reviewing and editing) - Carson Christiano – Funding acquisition, Project administration, Writing (reviewing and editing) - Joel Ferguson – Writing (original draft), Writing (reviewing and editing) - Fernando Hoces de la Guardia – Conceptualization, Funding acquisition, Writing (original draft), Writing (reviewing and editing) - Katherine Hoeberling – Funding acquisition, Project administration, Writing (original

draft), Writing (reviewing and editing) - Edward Miguel – Conceptualization, Funding acquisition, Supervision - Emma Ng – Visualization, Writing (original draft), Writing (reviewing and editing) - Lars Vilhuber – Conceptualization, Funding acquisition, Supervision

Chapter 9

Acknowledgments

Support for the development of the ACRE Guidelines document and platform was provided by Arnold Ventures.

Chapter 10

Definitions

10.1 Concepts in reproducibility

##Concepts in reproducibility - **Analytic data** – Data used as the final input in a workflow in order to produce a statistic displayed in the paper (including appendices). - **Causal claim** – An assertion that invokes causal relationships between variables. A paper may estimate the effect of X on Y for population P, using method F. Example: "This paper investigates the impact of bicycle provision on secondary school enrollment among young women in Bihar/India, using a Difference in Difference approach." - **Data availability statement** – A description, normally included in the paper, of the terms of use for data used in the paper, as well as the procedure to obtain the data (especially important for restricted-access data). Data availaibility statements expand on and complement data citations. Find guidance on data availability statements for reproducibility here. - **Data citation** – The practice of citing a dataset, rather than just the paper in which a dataset was used. This helps other researchers find data, and rewards researchers who share data. Find guidance on data citation here. - **Data sharing** – Making the data used in an analysis widely available to others, ideally through a trusted public repository/archive. - **Descriptive/predictive claim** – A paper with such kind of a claim estimates the value of Y (estimated or predicted) for population P under dimensions X using method M. Example: "Drawing on a unique Swiss data set (population P) and exploiting systematic anomalies in countries' portfolio investment positions (method M), I find that around 8% of the global financial wealth of households is held in tax havens (value of Y)." - **Disclosure** – In addition to publicly declaring all potential conflicts of interest, researchers should detail all the ways in which they test a hypothesis, e.g., by including the outcomes of all regression specifications tested. This can be presented in appendices or supplementary material if room is limited in the body of the text. - **Intermediate data** – Data not directly used as final input for analyses presented

in the final paper (including appendices). Intermediate data should not contain direct identifiers. - **Literate programming** – Writing code to be read and easily understood by a human. This best practice can make a researcher's code more easily reproducible. - **Pre-specification** – The act of detailing the method of analysis before actually beginning data analysis. - **Processed data** – Raw data that have gone through any transformation other than the removal of PII. - **Raw data** – Unmodified data files obtained by the authors from the sources cited in the paper. Data from which personally identifiable information (PII) has been removed are *still considered raw.* All other modifications to raw data make it *processed.* - **(Trial) registry** – A database of registered studies or trials, for example the AEA RCT Registry or clinicaltrials.gov. Some of the largest registries only accept randomized trials, hence the frequent discussion of 'trial registries. *Registration* is the act of publicly declaring that a hypothesis is being, has been, or will be tested, regardless of publication status. Registrations are time-stamped. - **Replication** – Conducting an existing research project again. A subtle taxonomy exists and there is disagreement, as explained in Hamermesh, 2007 and Clemens, 2015. *Pure Replication, Reproduction, or Verification* entails re-running existing code, with error-checking, on the original dataset to check if the published results are obtained. *Scientific Replication* entails attempting to reproduce the published results with a new sample, either with the same code or with slight variations on the original analysis. - **Reproducibility** – A research paper or a specific display item (an estimate, a table, or a graph) included in a research paper is reproducible if it is possible to reproduce within a reasonable margin of error (generally 10%) using the data, code, and materials made available by the author. Computational reproducibility is assessed through the process of **reproduction**. - **Reproduction package** – A collection of all the materials associated with the reproduction of a paper. A reproduction package may contain data, code and documentation. When the materials are provided in the original publication they will be labeled as 'original reproduction package', when they provided by a previous reproducer they will be referred as 'reproducer X's reproduction package'. At this point you are only assessing the existence of one (or more) reproduction packages, you are will not be assessing the quality of its content at this stage. " - **Researcher degrees of freedom** – The flexibility a researcher has in data analysis, whether consciously abused or not. This can take a number of forms, including specification searching, covariate adjustment, or selective reporting. - **Robustness check:** – Any possible change in a computational choice, both in data analysis and data cleaning, and its subsequent effect on the main estimates of interest. In the context of ACRE, the focus should be on the set of **reasonable specifications** (Simonsohn et. al., 2018), defined as (1) sensible tests of the research question, (2) expected to be statistically valid, and (3) not redundant with other specifications in the dataset. - **Specification searching** – Searching blindly or repeatedly through data to find statistically significant relationships. While not necessarily inherently wrong, if done without a plan or without adjusting for multiple hypothesis testing, test statistics and results no longer hold their traditional meaning, can result in false positives, and thus impede replicability.

"- **Trusted digital repository** – An online platform where data can be stored such that it is not easily manipulated, and will be available into the foreseeable future. Storing data here is superior to simply posting on a personal website since it is more easily accessed, less easily altered, and more permanent." - **Version control** – The act of tracking every change made to a computer file. This is quite useful for empirical researchers who may edit their programming code often.

##Concepts in the ACRE exercise and the platform - **Candidate paper** is a paper that has been considered for reproduction, but the reproducer decided not to move forward with the analysis due to failure to locate a reproduction package. Learn more here. - **Declared paper** is the paper that the reproducer analyzes throughout the exercise.

Bibliography

Chang, A. and Li, P. (2015). Is economics research replicable? sixty published papers from thirteen journals say'usually not'. *Available at SSRN 2669564*.

Christensen, G., Freese, J., and Miguel, E. (2019). *Transparent and reproducible social science research: How to do open science*. University of California Press.

Galiani, S., Gertler, P., and Romero, M. (2018). How to make replication the norm. *Nature*, 554(7693):417–419.

King, G. (1995). Replication, replication. *PS: Political Science and Politics*, 28:444–452. See updates to this paper for how I use this paper as a class assignment now.

Kingi, H., Vilhuber, L., Herbert, S., and Stanchi, F. (2018). The reproducibility of economics research: A case study. Presented at the BITSS Annual Meeting 2018 and available at the Open Science ….