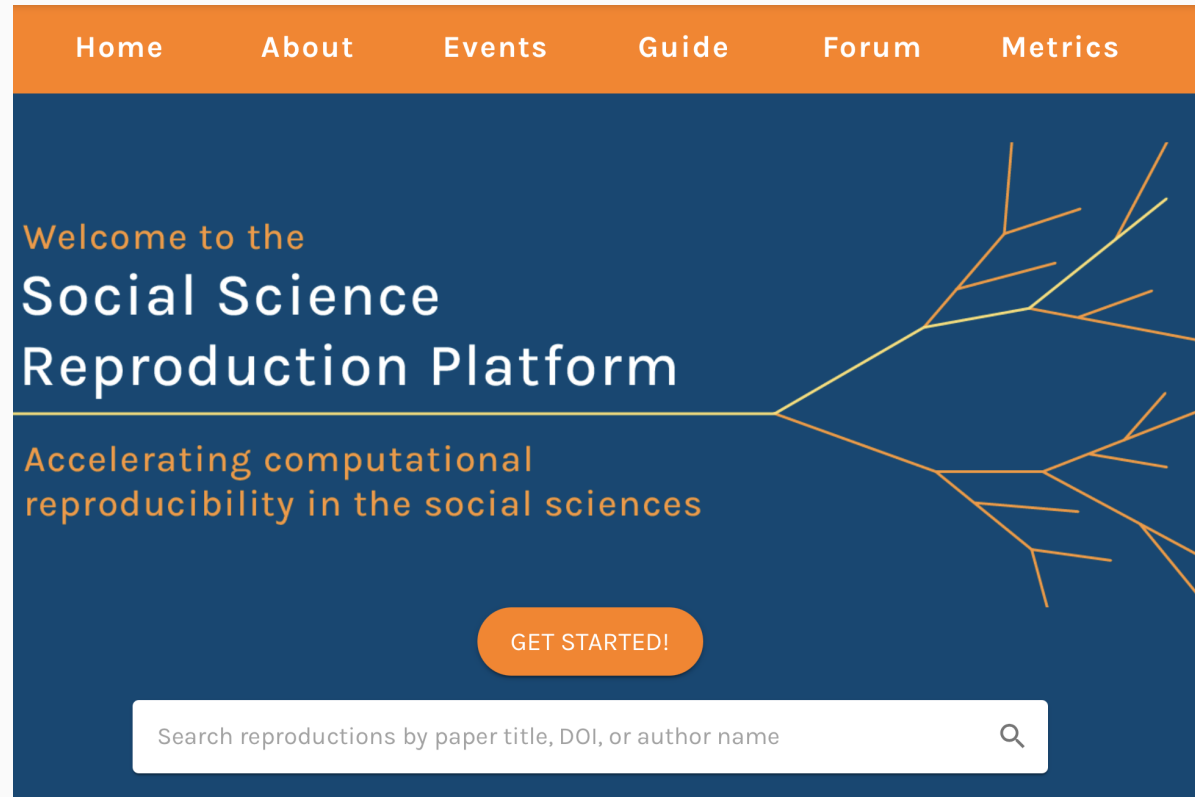# Using the Social Science Reproduction Platform in Replication Games

I4R Replication Week

Fernando Hoces de la Guardia, BITSS
October 2023

About Us: BITSS et. al.

Our goal: Standardize and preserve knowledge generated in reproduction exercises.

www.socialsciencereproduction.org

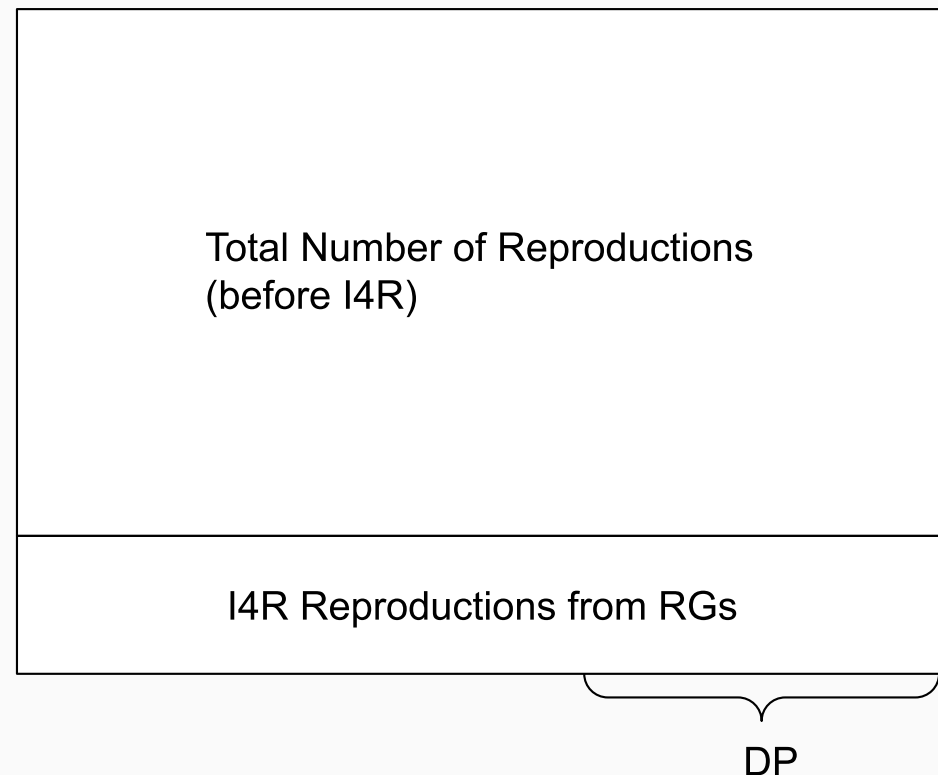# What Type of Reproduction Excercise?

Every semester, graduate students around **the world** take an Empirical/Applied [...] Economics course (e.g., Labor, Development). A typical assignment consists of reproducing the results of a paper and, possibly, testing the robustness of its results.

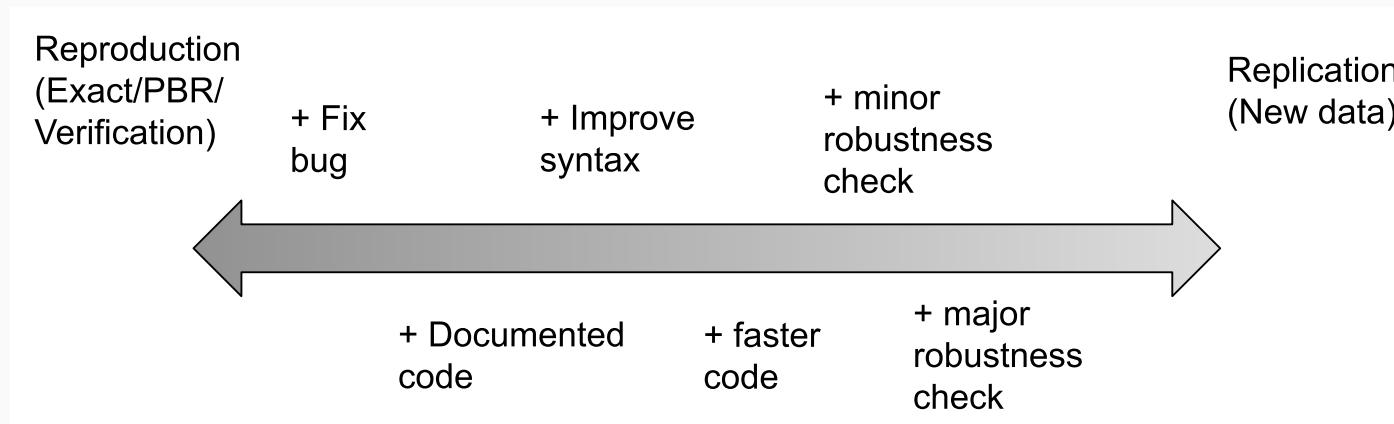| Stage	New Knowledge
Scope (select and verify)	Data and code exist?
Assess	Degree of reproducibility for specific part of the paper
Improve	E.g. fixed paths, libraries, added missing files, etc.
Test robustness	Results are robust to additional specifications

Key challenge: Standardization

Complementarities With I4R

- Contributions of I4R:

  1. Increase the total number of reproductions with Replication Games (200!)
  2. Institutionalized an outlet to published constructive summaries for an important fraction of these exercises: Discussion Paper Series (80!).

Total Number of Reproductions
(before I4R)

I4R Reproductions from RGs
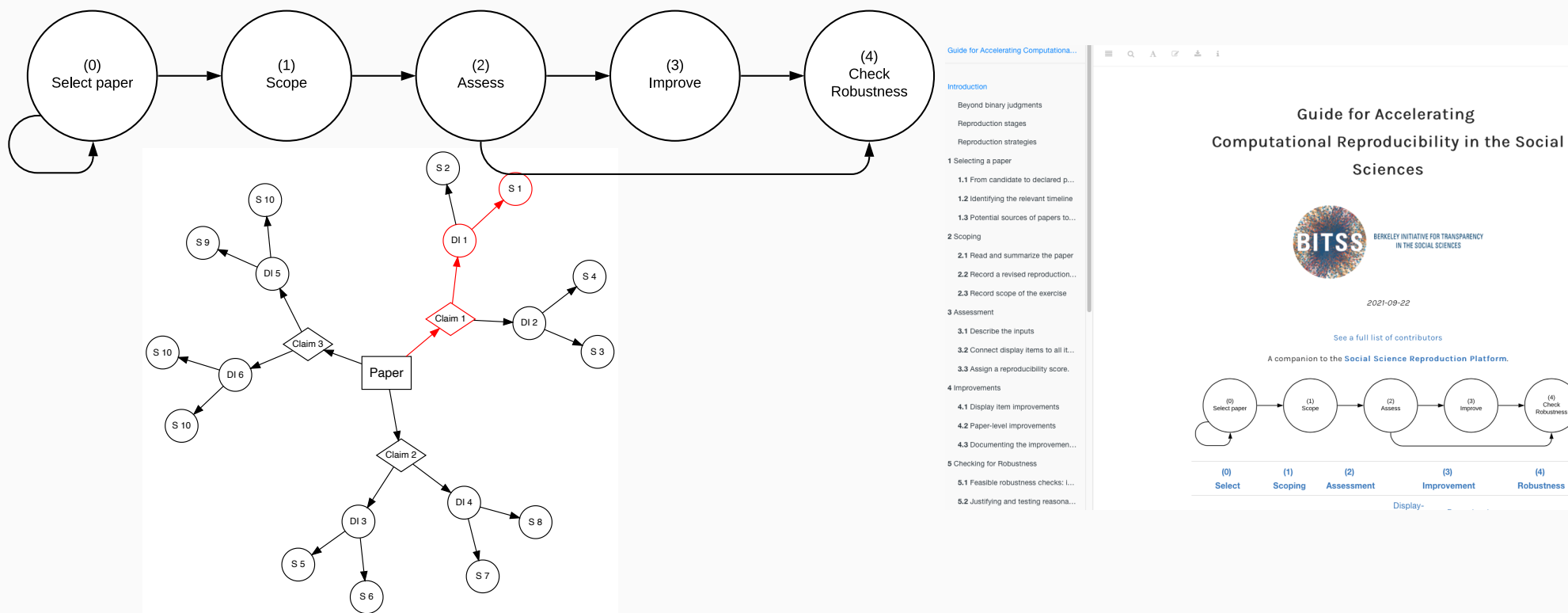
DP

Complementarities With I4R

- SSRP can help as a starting point in the Reproduction-Replication continuum

- How can the SSRP can help Replication Game participants:

  - As a companion and record to the reproduction section of the exercise.

| Reproduction (Exact/PBR/ Verification) | + Fix bug | + Improve syntax | + minor robustness check | Replication (New data) |

| + Documented code | + faster code | + major robustness check |

- Aspiration: every RG participants/team enters a reproduction in the SSRP at the end of a RG day to standardize and preserve the non-replication part of its exercise.

- **Strengths**: Detailed guidance, and flexible platform to record as much detail as possible from a reproduction exercise.
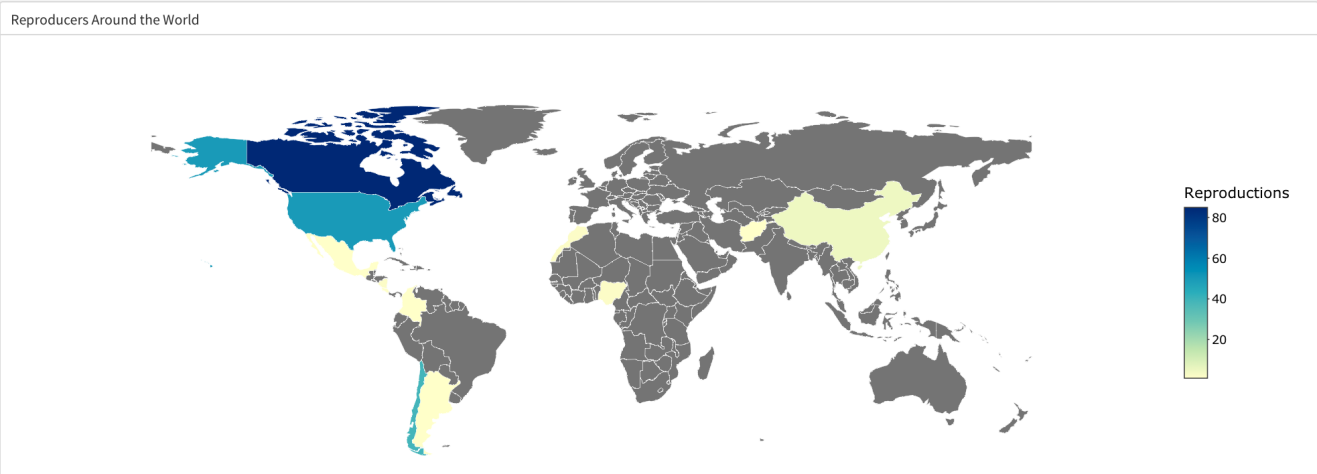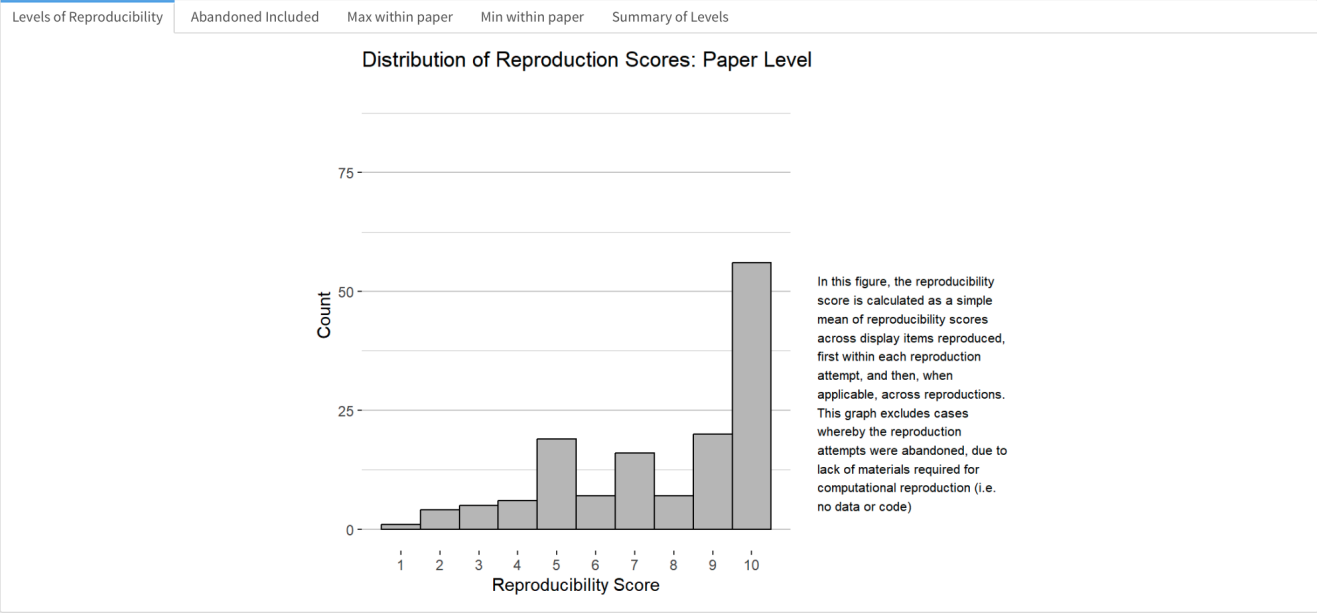

- **Weakness**: Too much detail can be overwhelming. Lots of space for improvement in user experience.

Current Users of the SSRP

219 Reproductions

141 Papers

342 Claims

Levels of Reproducibility | Abandoned Included | Max within paper | Min within paper | Summary of Levels

Distribution of Reproduction Scores: Paper Level

In this figure, the reproducibility score is calculated as a simple mean of reproducibility scores across display items reproduced, first within each reproduction attempt, and then, when applicable, across reproductions. This graph excludes cases whereby the reproduction attempts were abandoned, due to lack of materials required for computational reproduction (i.e. no data or code)

Reproducers Around the World

Demo of RG-type Reproduction in the SSRP

- We will start by looking at a high profile journal: AER-I
- Starting from the latest look for a paper that has a reproduction package with data and code: June 2023 Issue (September had none!)
- First found with data and code: Leonardi & Moretti (2023)
- Let's go to socialsciencereproduction.org
    1. Create an account/Log in.
    2. Start a new reproduction.
    3. Look for the paper's DOI.
    4. Download reproduction package, and check readme file to see if there is enough data to conduct a reproduction.
    5. [Optional] Set up a repository for your potential improvements.
- Now we will see how to record this reproduction with some detail, and just the essential.

Demo: Some Detail 1/2

Scope:

1. Read the paper and define the scope of exercise: Identify the claim(s) of interest, and the display items that present evidence supporting this claim(s).
   - "We find that after 2005, the geographical concentration of restaurants increased sharply and at an accelerating rate."
   - Figure 2 (show how to enter it in table)

Assess:

1. Read the readme file.
2. [Optional] Provide a description of the materials found in the materials
3. try to run the code
   - Change directories
   - In my environment, I have `set varabbrev off` so here I need to turn them on again.
   - `ssc install outreg2`
   - After modification, the code runs, and produces same output.
4. Report.

Demo: Some Detail 2/2

[Optional] Improve:

1. Record the improvements made above

   - Added line to make it robust to no abbreviations.
   - Added line to install `outreg2`

2. Suggest specific improvements:

   - Separate the file `replication_submit.do` into smaller self contained scripts (start with line 91).

[Optional] Robustness Check:

- Lines 245-248 looks strange, lets see how the results change without them
- Report results (and add more suggestions for improvements if you want)

Demo: Just The Essentials

Scope:

1. Skip (record blank)

Assess:

1. try to run the code
   - Change directories
   - In my environment, I have `set varabbrev off` so here I need to turn them on again.
   - `ssc install outreg2`
   - After modification, the code runs, and produces same output.
2. Report:
   - Click on Assess display item, click "+Assess additional display item"
   - Add caption and answer yes to next question (5.1.1.2)
   - Then finish the assessment (materials, computation, subjective assessment)

Submit reproduction.

- **Upcoming Feature:** Minimal Reproduction Version (Mid November): The essential reproduction, without having to know how to navigate to get it.

Thank You

acre@berkeley.edu
fhoces@berkeley.edu