DIME Analytics
Code Review Package Checklist
v2.0

The submission package for peer code review should include all components in the checklist below. Code review participants can provide either a link to the complete package (**OneDrive** or **GitHub**), or a zipped folder - along with information on programming language(s) involved, and approximate run time. Questions should be directed to DIME Analytics.

Package Requirements

All scripts to be reviewed.

A **main script** that the reviewer can use to run all code after changing only the top-level directory. See here for a sample main script.

All de-identified data necessary to run the code, if you need the reviewer to verify computational reproducibility.

A ReadME file that contains the following information:

Software and version

Approximate run time of code

Software and version

Purpose of each folder

High-level objective of the code (e.g., this code performs regression analysis and outputs result tables)

Line of code in **main script** that should be changed to run the code in a different machine

Purpose of each script (if more than one script is used)

Outputs produced (if part of the review)

Which script creates which output

Any specific requests / guidance as to what the reviewer should focus on

README. file path and name:

Component-specific Requirements

Before submitting the package, please make sure you keep the following in mind with respect to individual components.

1. Main script requirements

You can refer to this template main do-file for more details.

> The main script runs from start to end if the reviewer changes only directory path(s) in one location of the main script. This should be laid out clearly in the ReadME file as well.
>
> The main script installs any commands required (e.g. from SSC in Stata, CRAN in R, etc.), or provides the means to do so. *(Refer to this session on Building Reproducible Environments for more information.)*
>
> In Stata, the main do file sets critical configurations such as version, matsize, and varabbrev (either directly, or through a wrapper command such as the ieboilstart command included in ietoolkit)
>
> In R, all necessary packages to run the code are loaded in the main script
>
> **Approximate time** it takes to run all code from the the main script:

2. Data requirements

> Are you able to share the data for the purpose of the review?
> If yes, is the data de-identified?
> Are all datasets necessary to run the code included in the package?

3. Output requirements

Refer to this session on Creating Reproducible Outputs for more details.

> All outputs created by the code to be reviewed are included
> The outputs created are saved and stored in the correct folders, e.g. final outputs are in the `Outputs/Final` folder