Summary and Action Items for: Report for 2018 by the AEA Data Editor

By Lars Vilhuber*

Keywords: reproducibility; replicability; science of science

I. Goals

- Greater transparency for all data underlying AEA articles, including confidential data, while systematically incorporating third-party repositories
- Greater reliability of computations where possible

II. Key points

- Work on an improved "Data and Code Availability Policy", associated infrastructure, and guidance is progressing
    - New policy plus interpretation drafted
    - Draft guidance available
    - Infrastructure = data and code repository: contract being negotiated, technical improvements planned
    - Infrastructure = better metadata about where data and code are stored: developed, user interface being developed
    - Infrastructure = undergraduate lab, augmented with graduate students, tasked with pre-acceptance verification is working
- Working selectively with a few editors on pre-publication (pre-acceptance) assessments (primarily AEJ:Policy at this point)
- Issues in migration /future deposits:
    - Historical archives need to be put under a license appropriate for data and code (no current license; copyright retained)
    - Future deposits need a default license, and acceptable alternate licenses

III. Outlook

- The new "AEA Data and Code Repository" will be made available to new submissions, and historical submissions migrated, in 2019Q2/2019Q3.
- The Data Editor, in coordination with the Secretary and the Editors, will disseminate and start implementing the new "Data and Code Availability Policy"

* Vilhuber: Cornell University, lars.vilhuber@cornell.edu.

- The Data Editor, in coordination with the Secretary and legal Counsel, will determine an appropriate set of licenses for own ("AEA Data and Code Repository") and third-party repositories
- Enhanced but user-friendly metadata will identify locations, persistence, and licenses of all data and code
- Scaling pre-acceptance verification to more journals, using non-Cornell students

IV. Action items for the EC

ACTION: The EC is **requested to approve**: All historical archives shall be licenses under a mixed open-source license (CC-BY 4.0 for databases and text, modified BSD for code). This is a change to the overall policy of the AEA as to intellectual property.

Notes:
- Creative Commons Attribution CC-BY 4.0 is the gold-standard for open licenses, allowing all re-use, without conditions other than attribution (citing). Other licenses impose restrictions - ShareAlike, Non-Commercial, Non-Derivative.
- Absence of license means that re-use is not approved in general, every re-use would require the AEA to provide permission. In theory, all replication attempts currently are "unauthorized" in the absence of a license
- The Creative Commons organization itself suggests NOT to use CC licenses for software
- Modified BSD license is the suggested license for open re-use of software (code), but others are also acceptable (as per `https://opensource.org/licenses`).

ACTION: The EC is **requested to approve**: All future supplements in the hosted repository shall have **(a)** copyright by the authors, not the AEA **(b)** requirement to make databases available under some open license sufficient for replication, and **(c)** all code available under an open-source license sufficient to allow for replication

Notes:
- This is a change for the AEA in terms of who holds the copyright
- It is in line with best practices for replicable science