J-PAL
ABDUL LATIF JAMEEL POVERTY ACTION LAB

DATA AND CODE AVAILABILITY POLICY[1]
SEPTEMBER 2020

The Abdul Latif Jameel Poverty Action Lab (J-PAL) has been committed to making research more transparent for over a decade by developing a registry for randomized evaluations and publishing data from research studies conducted by our affiliates. As of September 2020, this data and code availability policy is in effect for all studies that have received or will receive funding from J-PAL research initiatives. We strongly advise all affiliated researchers to follow this policy for studies that are not funded by J-PAL.

To complement this policy, J-PAL provides guides for researchers and their research teams on how to prepare (de-identify) and publish data, and can provide support to affiliates if needed.

PUBLISH DATA IN A TRUSTED DIGITAL REPOSITORY

All data materials associated with the study must be made publicly available in a trusted digital repository and in a manner that protects required confidentiality of Personally Identifiable Information. We encourage researchers to publish in the J-PAL Dataverse. However, researchers may alternatively choose to publish their data materials in other repositories,[2] such as but not limited to the Harvard Institute for Quantitative Social Sciences (IQSS) Dataverse, the Inter-university Consortium for Political and Social Research (ICPSR), the Yale Institution for Social and Policy Studies (ISPS) Data Archive.

WHEN SHOULD MATERIALS BE PUBLISHED

Researchers should make data publicly available when any of the following conditions are met:
- When an academic paper is accepted, the data and code used to generate the paper's findings should be available within 60 days of the paper's acceptance, or by the deadline of the journal in which the paper is accepted, whichever comes earlier;
- Within three years of the completion of data collection—or earlier if required by a funder—all primary data, metadata, codebooks, program code, and supporting information should be made available. If academic publications are still pending at this time, researchers may request an extension.
- J-PAL will ask funded researchers to submit a data publication plan in their final narrative report.

WHAT DATA MATERIALS ARE REQUIRED TO BE PUBLISHED?

Data materials include all primary data, metadata, codebooks, program code, and supporting information collected during the course of the study.

Researchers collecting original data must:
- Where applicable, provide replication materials for published papers, including software code, codebooks, and any other documentation necessary to precisely reproduce all published results; and
- Make available all primary data, including all variables, treatment conditions, and observations, regardless of whether they were used to produce published results;
- Provide an account of the procedures used to collect, preprocess, clean, or generate the clean dataset (i.e., complete and annotated cleaning code);

---

[1] This policy has been revised from the June 2015 policy, accessed here.
[2] J-PAL can help researchers determine that a particular repository is "trusted".

- Provide research materials (e.g. survey instruments) and a description of procedures necessary to conduct an independent replication of the research.

Researchers using administrative data or data received from a third party must:
- Make available all primary data, including all variables, treatment conditions, and observations, as permitted by the data use agreement;
- Provide precise information as to the source of the data (data citations, exact URLs, etc.);
- Where applicable, provide replication materials for published papers, including software code and any other documentation necessary to allow an informed researcher to precisely reproduce all published results.

Researchers re-using data available from public repositories must:
- Provide precise information as to the source of the data (data citations, exact URLs, etc.);
- Provide software code and any other documentation necessary to reproduce all published results.

RESTRICTED DATA AND RESTRICTED ACCESS REPOSITORIES

If some or all study data cannot be shared for legal, ethical, or proprietary reasons, researchers must inform J-PAL in the final narrative report. In such cases, J-PAL may waive the requirement to publish the data materials, provided researchers:
- Explain and justify the restrictions to the data materials;
- To the extent possible, provide access to the exempted data through a restricted access repository, subject to appropriate usage and access rights;
- Comply with all aspects of this policy for which the constraints do not apply; and
- Agree to provide detailed information on how, where, and under what conditions an independent researcher can access the data (e.g. in a ReadMe file accompanying the available data).

DATA AVAILABILITY STATEMENT

A data availability statement that describes how to access the data and code should be included in the published paper. The statement should cite the repository where the data materials have been published. If the data are restricted, researchers must include in the paper an explanation of the restrictions and all necessary information required for a reader or reviewer to apply for access to the data and the conditions under which access will be granted. See here for guidance on writing a data availability statement.

SHARING RIGHTS AND PROTECTION OF HUMAN SUBJECTS

J-PAL expects that researchers make reasonable efforts to include data publication provisions both when seeking Institutional Review Board approval or when negotiating data access with third parties. Researchers are responsible for ensuring they have legal and ethical authority to publish data and they have followed appropriate procedures to minimize the risk of identification of individual human subjects.

QUESTIONS OR ASSISTANCE

The J-PAL research team has published guides on de-identification and data publication which can help researchers and research staff comply with this policy. The J-PAL research team also provides data publication services to researchers affiliated with or funded by J-PAL. Assistance includes but is not limited to preparing the data and code, replicating results, tables, and uploading datasets for publication. Requests for assistance, or other questions and comments about this policy, can be sent to data@povertyactionlab.org.