Data Article

Adoption of agricultural technologies in the developing world: A meta-analysis dataset of the empirical literature

Sacha Ruzzante*, Amy Bilton

University of Toronto, 55 St George St., Toronto, ON M5S 0C9, Canada

Article Info

Article history:
Received 11 June 2021
Revised 25 August 2021
Accepted 15 September 2021
Available online 20 September 2021

Keywords:
Adoption
Agricultural technology
Developing world
Improved varieties
Natural Resource Management
Probit
Survey

Abstract

The meta-analysis dataset presented is a convenience sample from 218 separate studies of agricultural technology adoption in Africa, Asia, and Latin America. Each study uses survey data to estimate a form of multiple regression of adoption of a technology (dependent variable) with a diverse array of predictor variables. Fifteen predictor variable categories are included in this dataset: Age, education, gender, household size, farming experience, land size, soil fertility, land slope, distance to inputs/outputs, access to credit, land tenure, livestock ownership, non-farm income, access to extension, and organization membership. Data have been cleaned and transformed to common units. A total of 384 statistical models are recorded, with a total of 2875 effect size estimates.

© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

DOI of original article: 10.1016/j.worlddev.2021.105599
* Corresponding author.
E-mail address: sachawruzzante@gmail.com (S. Ruzzante).
Social media: 🦅 (S. Ruzzante)

https://doi.org/10.1016/j.dib.2021.107384
2352-3409/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications Table

| Subject | Development |
|---------|-------------|
| Specific subject area | Agricultural economics and econometrics. Adoption of agricultural technologies in the developing world. |
| Type of data | Table |
| How data were acquired | A literature search was conducted using keyword combinations in Google Scholar and in the catalogues of several leading journals. Mendeley suggestions were also investigated. Lastly, literature review sections of the included studies were screened for citations that had not already been included. |
| Data format | Filtered |
| Parameters for data collection | The data were gathered from the academic and grey literature. Studies had to be focused on adoption of agricultural technology, had to report a regression analysis, and had to use survey data from a developing world context. Studies in English, Spanish, and French were included. |
| Description of data collection | Studies were found using three main techniques: |
| | 1) Google Scholar with combinations of search terms “technology,” “adoption,” “agriculture,” “developing world,” “logit,” “probit,” “machinery,” “conservation agriculture,” etc. |
| | 2) Searching through the catalogues of several leading journals. |
| | 3) Mendeley suggestions based on studies already added. |
| | 4) Snowball sampling by looking at cited papers. |
| Data source location | This dataset contains secondary data. The effect sizes recorded in the dataset are derived from regression models that use primary survey data from approximately 200 distinct farmer and farm household surveys. These surveys were conducted in countries in Latin America, Africa, Europe, and Asia, between 1980 and the present. |
| | The ‘Reference List’ document included with the data lists all of the publications from which data were extracted. |
| Data accessibility | https://doi.org/10.17632/8pkr9jr6v4.1 [7] |
| Related research article | Ruzzante, S., Labarta, R., & Bilton, A. (2021). Adoption of agricultural technology in the Developing World: A meta-analysis of the Empirical Literature. World Development, 146, 105599. https://doi.org/10.1016/j.worlddev.2021.105599 |

Value of the Data

- The dataset represents a significant effort to evaluate the reliability of the results of each study in the dataset. The data have been cleaned, and a record of modifications and deletions has been kept and is associated with each data point in the spreadsheet. This may be valuable for other researchers wishing to extend the meta-analysis presented in the research article associated with this Data in Brief article [1], or to investigate variables that were not fully explored in the research article.
- When conducting adoption research, it is useful to reference studies that were performed in the same region, and with related technology. Despite the large body of adoption literature, much of it is grey literature or published in journals that are not easily searchable. This dataset can be a quick resource for adoption researchers to find relevant studies.
- The data can be used to identify questions requiring further research. A quick analysis of the data shows that there are many countries, technologies, predictor variables, and statistical methodologies that have received little attention. Researchers looking for ‘holes’ in the literature can find them fairly quickly using these data.
Researchers may find these data useful to test and validate meta-analysis methodologies for social science research. The dataset contains dozens of columns that can be used to test the robustness of methodologies to deviations from normality.

1. Data Description

The data described by this article consist of three files:

1) AgTechAdoption.xlsx: This is the main data file, which includes all effect size estimates and metadata for all included studies.
2) Data Column Descriptions.pdf: This file contains detailed definitions for each column in AgTechAdoption.xlsx.
3) Reference List.pdf: This document includes formatted citations for all studies included in the dataset.

The main data file AgTechAdoption.xlsx contains effect size estimates for certain predictor variables on adoption of agricultural technologies in the developing world, drawn from 384 statistical models published in 218 studies. The data includes metadata about each study, including characteristics of the technology, information about the geographic context, details of the statistical methodology.

Effect sizes for 15 categories of predictor were recorded. These categories are:

- Age
- Education
- Gender
- Household size
- Farming experience
- Land size
- Soil fertility
- Land slope
- Distance to inputs/outputs
- Access to credit
- Land tenure
- Livestock ownership
- Non-farm income
- Access to extension
- Organization membership

Most categories include several different variable definitions, which are recorded in separate columns. In total the spreadsheet includes 138 columns, of which 81 contain effect size estimates and 56 contain study metadata. The specific definition of each column is described in the associated file “Data Column Descriptions”.

Fig. 1 below shows the number of statistical models recorded by region and type of technology. African regions are well represented across all technology categories, while studies in South Asia and Latin America tend to focus on improved varieties of crops.

Fig. 2 shows the number of effect size estimates recorded in the dataset, disaggregated by the predictor variable category and the technology. Improved Varieties provide the largest number of effect sizes for most predictor variable categories. Studies of Natural Resource Management technologies are also common. Age, education, gender, household size, and land size are the most commonly investigated predictor variable categories.
Fig. 1. Regional and technological distribution of statistical models included in dataset. The colour is related to the number of statistical models in each square.

2. Experimental Design, Materials and Methods

2.1. Inclusion criteria

Studies needed to meet several criteria to be included in the dataset:

1) Must analyse the adoption of an agricultural technology.
2) Must report a multi-variable regression analysis of some sort.
3) Must be a study taking place in a developing world context. For the purposes of this dataset, studies conducted in countries with a Human Development Index (HDI) less than 0.8 were automatically included. Studies from countries with an HDI greater than 0.8 (eg. Chile and Greece) if several conditions, such as a small average land size and low levels of modern technology adoption were satisfied.
4) Only studies published in English, Spanish and French were considered.

2.2. Search methods

Studies were gathered using a combination of techniques:

1) Combinations of search terms “technology,” “adoption,” “agriculture,” “developing world,” “logit,” “probit,” “machinery,” “conservation,” etc. were entered into Google Scholar. All search results were sequentially screened for eligible studies until the search returned only ineligible studies or duplicates.
**Fig. 2.** The bars show the number of effect size estimates recorded for each predictor variable, and for each technology. For most predictor variables the largest number of effect sizes are in the Improved Varieties category, while studies of Natural Resource Management technologies are also common.
2) Studies were added to a Mendeley catalogue. Mendeley’s proprietary software returned several suggested studies that were added to the dataset.

3) Literature reviews in included studies were screened for studies that had not yet been added.

4) Lastly, the databases of several top journals were searched, including World Development, Journal of Rural Studies, Agricultural Economics, and Food Policy.

2.3. Screening procedures

Study abstracts were screened prior to retrieving full records. In total, 351 documents were retrieved. Of those, 218 were included in the dataset. At this stage, the reasons for excluding studies included (i) The study did not report a regression model with adoption as the independent variable, (ii) The study used a unique regression model that was not considered comparable to the models used by most studies in the dataset (for example, Spatial Autoregressive Models), (iii) The study location was in a highly developed country (although most of these cases were caught during the screening of the abstracts, some studies did make it through that stage), (iv) The study presents an analysis that is later duplicated or improved for inclusion in another publication (for example, a working paper that is later published in a journal). It is important to note that at this stage the quality of each article was not formally assessed. Studies with apparent inaccuracies were left in the dataset; these data were flagged and in some cases removed, as described in the following section.

2.4. Data cleaning

Data cleaning was performed, and several accuracy issues were flagged:

1) Some studies reported statistical models with coefficients that contradicted the expected relationship from descriptive statistics. For example, the coefficient on cooperative membership in [2] is negative, although descriptive statistics suggest a positive relationship. In this case the author of the study responded to a request for clarification, confirming that the model coefficients were correct, but that there was an error in the descriptive statistics [3].

2) Unrealistically large and significant coefficients, indicating complete or quasi-complete separation. For example, the coefficient for gender on the adoption of conservation tillage is given as \( \beta = 19.6 \), \( SE = 0.92 \) in [4], but the descriptive statistics show that no women adopted conservation tillage.

3) Studies that reported standard errors or t-statistics that did not agree with the (*, **, ***) notation used.

Attempts were made to resolve these discrepancies, including contacting the study authors. In most cases this was not successful, due to non-response, defunct email addresses, or lost data. In cases where the discrepancies could not be resolved, the data was either excluded from the dataset (extreme cases), or the entry is preceded by an underscore (_), indicating there are irregularities in the data.

In addition, two methodological issues were flagged:

1) Some studies included many (likely) highly covariant predictors. For example, [5] include three variables: (i) total land owned (ha), (ii) farmer opinion of whether land owned is sufficient, and (iii) number of plots owned. Including these variables may be beneficial for the study itself, but each variable cannot then be interpreted independently.

2) Similarly, some studies include higher-order effects (squared variables or interactions). This reduces the interpretability of the main effects.

In these cases, results were generally omitted. However, in some cases the higher-order effects were very small compared to the main effect, or an estimate of the true main effect could
be deduced by taking a weighted average over all possible values of the interaction terms. In these special circumstances the results have been included.

2.5. Calculations and adjustments to the data

For the most part, the data are presented as-is, and no calculations or adjustments were applied by the authors. However, several variables in the dataset do represent calculated data. These are described briefly below. The Data Column Descriptions document provides further details of each variable.

1) Indices: The Human Development Index (HDI), Education Index (EI), Income Index (II) and Health Index (HI) are referenced from the Subnational Human Development Database [6] for the study region and the year in which the study data were collected. In cases where the study included more than one of the regions listed by the Subnational Human Development Database, or where the study data were collected over multiple years, an average was taken.

2) In cases where studies used non-standard units for a predictor variable, the effect size values were converted. This is relevant in three cases:
   a. Land size. For example, ‘kert’ are sometimes used in Ethiopia instead of hectares, while ‘decimals’ are used in parts of Bangladesh and India. Effect sizes in acres were also converted to hectares.
   b. Distance to inputs/outputs. For example, effect sizes in miles were converted to kilometres. Effect sizes in minutes were converted to hours.
   c. For any binary predictor variable, if the definitions of the ‘1’ and ‘0’ values in a study were opposite to the definitions used in this dataset, the effect size was converted. For example, for gender, this dataset uses the convention that Male=1 and Female=0, while some studies define Male=0 and Female=1. In these cases the predictor coefficient for gender was multiplied by –1.

In all cases where data was omitted, preceded by an underscore, or adjusted, an explanatory note is attached to the spreadsheet cell. Although the authors believe most issues have been identified and resolved, they can provide no guarantee that the dataset is free of errors. The dataset is also not exhaustive. There are undoubtedly studies that were not found, and more are published every year.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT Author Statement

Sacha Ruzzante: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Funding acquisition; Amy Bilton: Conceptualization, Resources, Writing – review & editing, Supervision, Funding acquisition.

Acknowledgments

We wish to thank the authors who responded to requests for extra data and clarification while compiling this dataset. This research was supported in part by the following:
- Powering Agriculture: An Energy Grand Challenge for Development. Partners: the United States Agency for International Development (USAID), the Swedish Government, the German Federal Ministry for Economic Cooperation and Development (BMZ), Duke Energy, and the Overseas Private Investment Corporation (OPIC) under grant number AID-OAA-A-16-00005.
- University of Toronto Centre for Global Engineering, Metcalfe Family Graduate Scholarship.
- Natural Sciences and Engineering Research Council, Canada Graduate Scholarship.
- Government of Ontario, Canada, Ontario Graduate Scholarship.

References

[1] S. Ruzzante, R. Labarta, A. Bilton, Adoption of agricultural technology in the developing world: a meta-analysis of the empirical literature, World Dev. 146 (2021) 105599, doi: 10.1016/j.worlddev.2021.105599.
[2] Y.A. Yigezu, A. Mugera, T. El-Shater, A. Aw-Hassan, C. Piggin, A. Haddad, … S. Loss, Enhancing adoption of agricultural technologies requiring high initial investment among smallholders, Technol. Forecast. Soc. Change 134 (2018) 199–206.
[3] Yigezu, Y.A. (2018, November 15). Personal Communication.
[4] M. Kassie, P. Zikhali, K. Manjur, S. Edwards, Adoption of organic farming techniques: evidence from a semi-arid region of ethiopia, Enviroment for Development Discussion Paper Series, Environment for Development, Washington, DC, 2009.
[5] M. Tura, D. Aredo, W. Tsegaye, R.L. Rovere, G. Tesfahun, W. Mwangi, G. Mwabu, Adoption and continued use of improved maize seeds: case study of central ethiopia, Afr. J. Agric. Res. 5 (17) (2010) 2350–2358.
[6] J. Smits, I. Permanyer, The subnational human development database, Sci. Data 6 (2019) 190038.
[7] S. Ruzzante, A. Bilton, A Meta-Analysis Dataset of the Empirical Agricultural Technology Adoption Literature, Mendeley Data (V1) (2021), doi:10.17632/8prk9jrs4v4.1.