README_De_Maria_2020_Dataset

1. PROJECT

Title: Essays on the Economics of Large-Scale Land Acquisitions [PhD thesis]
Dates: 20 September 2015 – 20 September 2019
Funding organisation: University of Reading
Funding: University of Reading Research Studentship in Social Sciences 2015

2. DATASET

Title: Implementation of Large-Scale Land Acquisitions (LSLAs) and the Role of Institutions

Description: This dataset provides a unique set of variables which was used to assess the factors hindering or fostering the actual implementation of Large-Scale Land Acquisitions projects concluded since the year 2000 in different countries all over the world. In particular, the dataset contains deal-specific variables, as well as macro-economic and institutional variables measured at the country level for both host and investor countries involved in each LSLA deal. Using this data for different purposes and for learning activities is encouraged.

This is an original dataset built from Land Matrix and World Bank Data, with all the original sources of data listed below under "Source(s)" and licensed under Creative Commons Attribution 4.0 International Licence: https://creativecommons.org/licenses/by/4.0/.

Publication Year: 2020
Creator(s): Marcello De Maria
Organisation(s): University of Reading
Rights-holder(s): Marcello De Maria

Source(s):

 [1] The Land Matrix. International Land Coalition (ILC), Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), Centre for Development and Environment (CDE), German Institute of Global and Area Studies (GIGA) and Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ). Web: https://landmatrix.org/ [Retrieved on 11/04/2018].

 [2] Worldwide Governance Indicators. The World Bank. Web: www.govindicators.org [Retrieved on 11/04/2018]

 [3] World Development Indicators. The World Bank. Web: https://datacatalog.worldbank.org/dataset/world-development-indicators [Retrieved on 11/04/2018]

3. TERMS OF USE

This dataset is licensed by the rights-holder(s) under a Creative Commons Attribution 4.0 International Licence:
https://creativecommons.org/licenses/by/4.0

4. CONTENTS

File listing:

[1] LSLAs_Implementation_Dataset.csv
Description: this file contains the data in .csv format.

[2] README_De_Maria_2020_Dataset.pdf
Description: this file contains the dataset metadata in pdf format.

[3] CODEBOOK_De_Maria_2020_Dataset.pdf
Description: this file contains metadata for each variable included in the dataset in pdf format.

[4] LSLAs_Implementation_Dataset.xlsx
Description: this file contains the data in .xlsx format.

5. METHOD and PROCESSING

This original database is structured as a cross-section of LSLAs, where deal-specific information relies on the Land Matrix, while other economic and institutional variables are taken from the World Bank data repository (see "Sources").

Please note that this dataset contains missing values and that the original sources used for building this dataset are constantly changing over time, so that the sources might have changed compared to the time when the records were originally downloaded and processed.

The variable "Y" [Y] is an *ad hoc* dummy variable which takes the unit value when a deal is in production – that is, the deal is labelled both "concluded" [NEG_ST3] and "in operation" [IMPL_ST] - and zero otherwise. Deals marked as "concluded" [NEG_ST3], but without any information about the implementation status [IMPL_ST] were coded as missing values, just like deals for which both information on negotiation status [NEG_ST3] and implementation status [IMPL_ST] is missing. Overall, this dummy variable accounts for 1,424 deals in production (53.1%) over a total of 2,345 non-missing records [with 334 missing values in total.

Deal-specific variables include the size of the deal, as well as several dummies accounting for the existence of a domestic participation in the land deal, the presence of multiple investor countries at the same time and the nature of the proposed investment (food, livestock, biofuel, multiple intention, etc.). Each deal is typically observed in the negotiation year, which corresponds to the year of negotiation reported in the Land Matrix.

The size variable [SIZE] used in this dataset is measured in hectares and was built as a combination of the intended size [SIZE_INT], contract size [SIZE_CT] and production size [SIZE_PD] included in the original Land Matrix dataset. In particular, the values of the size variable are equal to the contracted size unless this information is missing, in which case the production size is taken. When both the contract and production size are unavailable, the intended size is used as a proxy to populate the size variable. The variable is also available in log [LNSIZE].

Additional control variables included in the dataset - for both target and investor countries - have been extensively used in the empirical literature on LSLAs. These include data series for the GDP, the average cereal yield per hectare (both obtained from the World Development Indicators [WDI] database of

the World Bank), as well as a suite of institutional variables retrieved from the Worldwide Governance Indicator [WGI] dataset and available on the World Bank data repository.

Missing values in given years for given countries, for both World Bank's WDI and WGI series, were estimated using simple linear interpolation and extrapolation techniques, so that values of institutional and macro-economic variables are taken at the specific time in which each deal is recorded. In case of multiple investors involved in the same deal at the same time, WDI and WGI-related variables are calculated as the simple average of all investor countries participating into a given deal, thus assuming — with no additional detail available — that each investor country contributes to the project with the same share.

The dataset was originally created in 2018 and processed using Stata®. The actual code used for creating, processing, merging and cleaning the dataset can be requested by contacting directly the creator and rights-holder - Dr. Marcello De Maria (marcello.demaria@reading.ac.uk).

Additional information will be provided in the following forthcoming paper:

[1] De Maria, M., Zanello, G. and E. Robinson (In preparation): The Role of Institutions in the Implementation of Large-Scale Land Acquisitions Projects.