Readme file- Reproducible package

Readme File for Replication of the Paper: "The Effects of Regulatory Information Systems on Credit Allocation: Evidence from Brazil's Rural Credit Bureau"

This readme file was created to guide the reproducibility package for the analysis of the paper.

Authors: Silvio Arduini, Faruk Miguel, Claudio Moreira, Alvaro Pedraza, Claudia Ruiz-Ortega, Paulo Sampaio and Lucas Dos Santos

Overview

This paper studies how improvements in regulatory information systems affect credit allocation in rural financial markets. In 2021, the Central Bank of Brazil introduced major enhancements to the Rural Credit Bureau, first by integrating geospatial data on borrowers' plots with national environmental and social compliance registries, and then by automating real-time verification of credit applications. These reforms simultaneously increased the information available to financial institutions and established a centralized second line of defense that blocks non-compliant credit operations. Using municipality-level data, we find that after the improvement in regulatory information, credit growth slowed in areas with larger shares of protected land, particularly among cooperatives and public banks. The effect is consistent across regions and biomes, indicating the operation of a uniform national enforcement mechanism. We rule out confounding policy changes and show that the results are not driven by shifts in credit demand or bank risk exposure. The findings suggest that combining improved data integration with real-time monitoring can strengthen environmental enforcement by redirecting financial flows away from high-risk areas without broadly constraining access to rural credit.

Data Availability

This research utilizes data from public sources. The data entries in this package outline all sources, along with instructions for their download and placement in the specified directories for full replication.

Data Sources

The master dataset used in the analysis is "/Raw data/aggregate_monthly_muni3.dta". This data is generated from five external sources. These sources are not included in the reproducibility package due to size. Instead, we provide direct links to each source, from which users can download the required files. All downloaded files should be placed in the folder: /Raw data/source files/.

External sources 3, 4, and 5 provide geographic information in *shapefile* format. Because shapefiles require *QGIS* software to process, we already include the processed outputs used in the paper. These are:

- biomes.dta
- border_dummies.csv
- protected_share_by_year.csv

These files were generated using the original shapefiles downloaded from the external links indicated below and are already copied into the "/Raw data/source files/" folder to facilitate replication. (See /Readme/Protected area methodology.docx for details.)

Below we list the five external data sources, the specific datasets, and instructions for downloading the required files:

1. From Central Bank of Brazil website: "Microdatos crédito rural" https://www.bcb.gov.br/estabilidadefinanceira/tabelas-credito-rural-proagro (accessed June 1, 2025). Navigate to "1. Sicor - Tabelas de Domínio do Crédito Rural e do Proagro" and download the files 1.1 to 1.5. Clicking on the links below provides direct download.

   1.1. Programa ou Linha de Crédito (https://www.bcb.gov.br/htms/sicor/Programa.csv).
   1.2. Subprogramas (https://www.bcb.gov.br/htms/sicor/Subprogramas.csv).
   1.3. Fonte de Recursos (https://www.bcb.gov.br/htms/sicor/FonteRecursos.csv).

1.4. Empreendimento (https://www.bcb.gov.br/htms/sicor/Empreendimento.csv).
1.5. IFs do Sicor (https://www.bcb.gov.br/htms/sicor/DadosBrutos/SICOR_LISTA_IFS.csv).
1.6. Navigate to "*2. Sicor - Microdados do Crédito Rural e do Proagro (desde 01/01/2013) - Operações contratadas (Recursos Públicos e Privados)*" and download the files "Em YYYY: SICOR_OPERACAO_BASICA_ESTADO_YYYY.gz" for years (YYYY) 2013 to 2024.
1.7. Navigate to "3. Sicor – Microdados do Crédito Rural Público" and download the file "Complemento às Operações contratadas" (https://www.bcb.gov.br/htms/sicor/DadosBrutos/SICOR_COMPLEMENTO_OPERACAO_BASICA.gz).

2. From the Brazilian Institute of Geography and Statistics – IBGE (accessed June 1, 2025) files 2.1 and 2.2.

2.1. Municipalities shapefiles: https://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2024/Brasil/BR_Municipios_2024.zip
2.2. biomes per municipality: extracted from pp.18-37 of: https://biblioteca.ibge.gov.br/visualizacao/livros/liv102097.pdf (accessed June 1, 2025). These data have been compiled into *biomes.dta*, already included in "/Raw data/source files/".

3. From the Ministry of Environment and Climate Change, conservation units: https://dados.gov.br/dados/conjuntos-dados/unidadesdeconservacao, click on "Recursos", and click on "Accesar o recurso" below "Polígono CNUC 2024_10" (accessed June 1, 2025).

4. From National Foundation for Indigenous Peoples – FUNAI. Indigenous lands (shape files), downloaded from this link: Terras Indígenas no Brasil (shp) (accessed June 1, 2025)

5. From the National Institute for Colonization and Agrarian Reform – INCRA. Quilombola settlements (shape files): https://www.gov.br/insa/pt-br/centrais-de-conteudo/mapas/mapas-em-shapefile/quilombos-incra.zip/view (accessed June 1, 2025)

The shapefiles from items 2.1, 3, 4, and 5 were processed using QGIS to produce the files protected_share_by_year.csv and border_dummies.csv, which are included in /Raw data/source files/.

Statement about Rights

We certify that the authors of the manuscript are using publicly available data.

Instructions for Replicators

Code

As the reader can see from the folder structure section, there is a Code folder with three (3) Stata do-files and (1) python script:

1) 0_master_dataset.do: cleans and merges the source files, generates a file with the key variables to run the other scripts (aggregate_monthly_muni3.dta). The script automatically creates an auxiliary folder where the intermediary datasets are saved.
2) 1_create data.do: prepares a balanced panel of municipalities and financial institutions over time.
3) 2_analysis.do: produces the tables and figures of the paper.
4) QGIS_Figure2.py produces figure 2.

While QGIS_Figure2.py is independent, the other scripts should be executed in sequence, first 0_master_dataset.do, followed by 1_create_data.do, and finally 2_analysis.do. Before running the scripts, the user must update the file paths to match their local directory structure and install the following Stata packages: distinct, winsor2, reghdfe, ftools, coefplot, and estout. For QGIS_Figure2.py we suggest to simply open in QGIS python editor and run the script, no file path changes needed.

In order to properly run the first script, the user needs to download and extract source files from 1.1 to 2.1 and copy them in the "/Raw data/source files" folder. The remaining source files are already processed and provided in biomes.dta, Border_dummies.csv, and protected_share_by_year.csv.

Computational settings

The script was developed and tested using **Stata version MP 18.0** and **QGIS version 3.38.1** on a **64-bit Windows system** (MinGW environment) with **32 GB of RAM (32000 MB)**. While the script ran successfully in this environment, we expect it would also run with less RAM.

Under the described settings, the reproducibility package takes around **5** hours to run from beginning to end.

List of Exhibits

Exhibit name	Output file name	Script	Note
Figure 1. Panel A: Rural earmarked credit in Brazil (by type of financial institution)	Fig1A.gph	2_analysis.do Code section: Fig1-Panel A	Found in /Results/
Figure 1. Panel B: Rural earmarked credit in Brazil (for municipalities with and without protected land)	Fig1B.gph	2_analysis.do Code section: Fig1-Panel B	Found in /Results/
Figure 3. Panel A: Impact of reform on lending (number of financial institutions)	Fig3A.gph	2_analysis.do Code section: Fig3	Found in /Results/
Figure 3. Panel B: Impact of reform on lending (total lending volume in million Reais)	Fig3B.gph	2_analysis.do Code section: Fig3	Found in /Results/
Figure 3. Panel C: Impact of reform on lending (probability of operating for the average financial institution)	Fig3C.gph	2_analysis.do Code section: Fig3	Found in /Results/
Figure 3. Panel D: Impact of reform on lending (volume for the average financial institution in million Reais)	Fig3D.gph	2_analysis.do	Found in /Results/

		Code section: Fig3	
Figure 4. Panel A: Impact of reform on lending by type of financial institution (number of cooperatives)	Fig4A.gph	2_analysis.do Code section: Fig4	Found in /Results/
Figure 4. Panel B: Impact of reform on lending by type of financial institution (lending volume of average cooperative in million Reais)	Fig4B.gph	2_analysis.do Code section: Fig4	Found in /Results/
Figure 4. Panel C: Impact of reform on lending by type of financial institution (number of public banks)	Fig4C.gph	2_analysis.do Code section: Fig4	Found in /Results/
Figure 4. Impact of reform on lending by type of financial institution (lending volume of average public bank in million Reais)	Fig4D.gph	2_analysis.do Code section: Fig4	Found in /Results/
Figure 4. Panel E: Impact of reform on lending by type of financial institution (number of private banks)	Fig4E.gph	2_analysis.do Code section: Fig4	Found in /Results/
Figure 4. Panel F: Impact of reform on lending by type of financial institution (lending volume of average private bank in million Reais)	Fig4F.gph	2_analysis.do Code section: Fig4	Found in /Results/
Figure 5. Panel A: Regional Heterogeneity (average share of protected area by region)	Fig5A.gph	2_analysis.do Code section: Fig5	Found in /Results/
Figure 5. Panel B: Regional Heterogeneity (Earmarked credit by region)	Fig5B.gph	2_analysis.do Code section: Fig5	Found in /Results/
Table 1: Summary Statistics	Tables.xlsx, sheet(T1)	2_analysis.do Code section: Table 1	Found in /Results/
Table 2: Impact of reform on lending volume and presence of financial institutions	Tables.xlsx	2_analysis.do Code section: Table 2	Found in /Results/
Table 3: Impact of reform by type of financial institution	Tables.xlsx	2_analysis.do	Found in /Results/

Table 4: Effect heterogeneity across regions, biomes and the Amazons	Tables.xlsx	Code section: Table 3 2_analysis.do Code section: Table 4	Found in /Results/

Folder structure

C:\Users\wb381032\OneDrive - WBG\Research\Brazil Green Credit Bureau\Replication Package

- Code
  - 0_master_dataset.do
  - 1_create data.do
  - 2_analysis.do
  - QGIS_Figure2.py
- Raw data
  - Aggregate_monthly_muni3.dta
  - source files
    - biomes.dta
    - Border_dummies.csv
    - protected_share_by_year.csv
- Readme
  - Readme.docx
  - Protected area methodology.docx
- Results
  - Fig1A.gph
  - Fig1B.gph
  - Fig2.png
  - Fig3A.gph
  - Fig3B.gph

    - Fig3C.gph
    - Fig3D.gph
    - Fig4A.gph
    - Fig4B.gph
    - Fig4C.gph
    - Fig4D.gph
    - Fig4E.gph
    - Fig4F.gph
    - Fig5A.gph
    - Fig5B.gph
    - Tables.xlsx
- Temp
    - T2A.txt
    - T2A.xml
    - T2B.txt
    - T2B.xml
    - T3A.txt
    - T3A.xml
    - T3B.txt
    - T3B.xml
    - T4A.txt
    - T4A.xml
    - T4B.txt
    - T4B.xml
    - balanced_panel.dta
    - fis.dta
    - table1.csv