March 30, 2023

Replication materials for

"Natural Resource Rents, Local Taxes and Government Performance: Evidence from Colombia" Review of Economics and Statistics

By Luis R. Martínez

1. Computational requirements:

Windows 10, 64 bit, 16 GB ram, Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHz processor 3.41GHz

2. Software requirements:

Stata 17, MP-Parallel Edition, 8-core network. Master do-file (00-master.do) installs a series of user-written commands. An internet connection is required to download and install these commands. Additional commands include: outreg2, reghdfe, ivreghdfe, ftools, ivreg2, splitvallabels, and parmest.

RStudio version 2022, 12.0+353. The script provided installs and imports all relevant programs. R version 4.1.3 or higher required.

3. Overview

The data and Stata do-files provided replicate the figures and tables of the paper "Natural Resource Rents, Local Taxes and Government Performance: Evidence from Colombia".

The user should create a local directory with the following structure:

./code/	- Folder with all the do-files provided (5 in total + one R script).
./data/	- Data folder.
./data/raw/	- Subfolder with all the raw data files provided (16 in total).

To run the code, user must start the file "*code/00-master.do"*, having first replaced the global *root* in line 38 with the corresponding path containing the folders list above. Execution of the master do-file calls the other files:

- *./code/10-panel-set-up.do*
- *./code/11-final-panel.do*
- *./code/20-estimation.do*
- *./code/30-appendix.do*

The master do-file creates all other folders and subfolders where the output will be saved.

The do-files also installs user-written programs which are not part of the standard package of Stata. An internet connection is required to download and install those files. The output shown in the paper is based on the latest version of the commands as of March 30, 2023. To install the additional commands, the user must set the global *install_programs* in line 72 of the master do-file equal to one.

The R script */code/31-appendix-maps.R* reproduces the maps in Figure B2. The user must manually update the directory in line 12 of the script (using forward slash for folders, /) before running the code.

Running time of all programs of the main paper (*20-estimation.do*): 30 seconds
Running time of all programs for the appendix (*30-appendix.do*): 2 minutes, 10 seconds

4. Data

The folder data/raw contains all the raw data required to construct the main data set and run the analyses. Online Appendix A provides an overview of all data sources used in the analysis.

4.1. Raw data sources:

- **01-panel-cede:** This folder contains data on municipal characteristics at the municipality-year level. The data is provided by the *Centro de Estudios sobre Desarrollo Económico* (CEDE), a research center at Universidad de los Andes. This dataset is referred to as Panel CEDE and can be downloaded from the following link: https://datoscede.uniandes.edu.co/en/. I use the 2014 version of this dataset.
- **02-expenditure:** This folder contains additional data on the allocation of natural resource royalties in 2010-11, and data on disaggregate investment between 2005 and 2011. The data was provided by the Departamento Nacional de Planeación of Colombia (DNP). Similar information is now publicly available from the following link: https://www.chip.gov.co/schip_rt/paginiciofut.htm
- **03-sgp-dnp:** This folder contains information on transfers from the national government to municipalities under the *Sistema General de Participaciones*. The information was provided by DNP and similar information is now publicly available from the following link: https://sicodis.dnp.gov.co/
- **04-royalties:** This folder contains information on municipalities' oil and coal royalties. Information on oil royalties is provided by DNP, the National Hydrocarbon Agency (ANH) and Ecopetrol. Information was accessed under the following links:[1]
    - DNP - Oil: https://sicodis.dnp.gov.co/
    - ANH - Oil: https://www.anh.gov.co/es/operaciones-y-regal%C3%ADas/regal%C3%ADas-de-producci%C3%B3n/estad%C3%ADsticas-e-informes/

[1] Information from Ecopetrol is no longer available online.

- **05-cadastral-info**: This folder contains cadastral information at the municipality-year level. Information was obtained through a formal request to Instituto Geográfico Agustín Codazzi
- **06-urosario-conflict**: This folder contains information on Colombia's internal conflict. The original data was provided by Universidad del Rosario and contains information on violent events between 1996 and 2014.
- **07-education**: This folder contains information on educational coverage at the municipality-year level. The data was provided by the (DNP) after filing a formal request for information.
- **08-social-programs:** This folder contains information on the percentage of poor population with access to subsidized health insurance. This information was provided by the Ministry of Health. It also contains information on the number of new families that received conditional cash tranfers under the program Familias en Acción. This information was provided by DNP.
- **09-aqueduct**: This folder contains information on aqueduct and sewerage coverage in 2005. Data was provided by DNP based on information from the 2005 population census. Current versions of the Panel CEDE also contain this information.
- **10-community-council**: This folder contains information on community councils held by President Alvaro Uribe between 2002 and 2010. Information was provided by Dr. Ana Tribin based on her paper:
    - Tribin, A. (2020). Chasing votes with the public budget. *European Journal of Political Economy*, 63, 101875. https://doi.org/10.1016/j.ejpoleco.2020.101875
- **11-procuraduria**: This folder contains information on disciplinary prosecutions from the Office of the Inspector General, Procuraduría General de la Nación (PGN). The data is based on the news bulletins published in the PGN website: https://www.procuraduria.gov.co/pages/Noticias.aspx. For each relevant news report, I recorded the names of the accused, their roles in the public administration, the nature of the charges, the timing of the events, and the outcome (if available).
- **12-health**: This folder contains information on multiple health outcomes.
    - The National Health Institute (INS) provided information on the Water Pollution Index (IRCA).
    - The Ministry of Health provided information on vaccination rates at the municipality-year level.
    - The Departamento Administrativo Nacional de Estadística (DANE) provided information on mortality and births. Raw data can be downloaded from the following link: https://www.dane.gov.co/index.php/estadisticas-por-tema?id=34&phpMyAdmin=3om27vamm65hhkhrtgc8rrn2g4
- **13-maps**: This folder contains the data required to reproduce the maps in Figure B3 in R.
- **14-prices**: This folder contains information on the average petroleum spot price from the IMF's International Financial Statistics (IFS).
- **15-elections**: This folder contains information on results for mayoral elections, municipal councils, and presidential elections. The data is provided by CEDE and raw data can be downloaded from the following link: https://datoscede.uniandes.edu.co/en/

- **16-banrep**: This folder contains information on the official USD to COP exchange rate. Information was provided by Colombia's central bank (*Banco de la República*). The raw data can be downloaded from the following link: https://www.banrep.gov.co/es/estadisticas/trm