Replication Archive for "Competition and Civilian Victimization"

Michael Gibilisco Brenton Kenkel Miguel R. Rueda

Contents

1 Introduction **1**

2 Files Included **1**
2.1 Data files 1
2.2 Source code 2
2.3 Other 4

3 Software Versions and Packages **4**

4 Codebooks for Tabular Data Files **4**
4.1 `colombia_latlong.csv` 5
4.2 `code_muni_cede.dta` 5
4.3 `group_origins.csv` 5
4.4 `military_bases.csv` 5
4.5 `violence_t_export.dta` 6
4.6 `violence_t_year_export.dta` 7

1 Introduction

This is the replication archive for the article "Competition and Civilian Victimization", forthcoming in the *Journal of Conflict Resolution*. If you have questions about the replication archive, please contact Brenton Kenkel at brenton.kenkel@gmail.com.

2 Files Included

2.1 Data files

- `colombia_latlong.csv`: Latitudes and longitudes of Colombian municipalities. Used by `figurea2-a.r`.

- `code_muni_cede.dta`: Colombian municipality codes from CEDE. Used by `figurea1.r`, `figurea2-a.r`, and `figurea2-b.r`.
- `group_origins.csv`: Location data on groups' areas of early influence. Used by `figurea2-b.r`.
- `military_bases.csv`: Location data on military bases. Used by `figurea2-a.r`.
- `mpio.shp` (with `mpio.dbf` and `mpio.shx`): Shapefiles for Colombian municipalities. Used by `figurea1.r`, `figurea2-a.r`, and `figurea2-b.r`.
- `violence_t_export.dta`: Main analysis data file. Used by most analysis scripts.
- `violence_t_year_export.dta`: Municipality-year analysis data. Used by `figurea3-prep.do`, `figurea4-prep.do`, and `tablea7.do`.

2.2 Source code

All files with `.do` extensions should be run in STATA. All files with `.r` extensions should be run in R.

- `figure1.r`: Generate **Figure 1** (Strategic interdependence distinguishing selective and non-selective victimization).
  - Must be run after `tablea2.do`, which outputs the required data file `boot_res_region_t.dta`.
  - Outputs image file `figure1.pdf`.
- `figure2.r`: Generate **Figure 2** (Strategic interdependence adding the ELN as a third group).
  - Must be run after `tablea3.do`, which outputs the required data file `boot_res_elnb.dta`.
  - Outputs image file `figure2.pdf`.
- `figurea1.r`: Generate **Figure A1** (Geographical variation in victimization strategies).
  - Outputs image files `figurea1-a.png` and `figurea1-b.png`.
- `figurea2-a.r`: Generate **Figure A2(a)** (Locations used in covariate specifications: Army bases, 2000).
  - Outputs image file `figurea2-a.png`.
- `figurea2-b.r`: Generate **Figure A2(b)** (Locations used in covariate specifications: Armed groups' early areas of influence).
  - Outputs image file `figurea2-b.png`.
- `figurea3-prep.do`: To be run before `figurea3.r`.
  - Outputs data file `boot_res_region_t_year.dta`.
- `figurea3.r`: Generate **Figure A3** (Strategic interdependence distinguishing selective and non-selective victimization(municipality-year)).
  - Must be run after `figurea3-prep.do`, which outputs the required data file `boot_res_region_t_year.dta`.
  - Outputs image file `figurea3.pdf`.
- `figurea4-prep.do`: To be run before `figurea4.r`.
  - Outputs data file `boot_res_eln_year.dta`.
- `figurea4.r`: Generate **Figure A4** (Strategic interdependence adding the ELN as a third group (municipality-year)).

  - Must be run after `figurea4-prep.do`, which outputs the required data file `boot_res_eln_year.dta`.
  - Outputs image file `figurea4.pdf`.
- `table1-and-a1.do`: Generate **Table 1** (Frequency of victimization patterns) and **Table A1** (Summary Statistics).
- `table2.do`: Generate **Table 2** (Estimates of the payoff parameters in the victimization model).
  - Outputs data file `boot_res_region_v.dta`.
- `table3-prep.do`: To be run before `table3.r`.
  - Outputs data file `phats.csv`.
- `table3.r`: Generate **Table 3** (Counterfactual estimates of the average probability of victimization).
  - Must be run after `table3-prep.do`, which outputs the required data file `phats.csv`.
  - Also must be run after `table2.do`, which outputs the required data file `boot_res_region_v.dta`.
- `tablea2.do`: Generate **Table A2** (Selective and Non-Selective Victimization).
  - Outputs data file `boot_res_region_t.dta`.
- `tablea3.do`: Generate **Table A3** (Three-Player Game Estimates).
  - Outputs data file `boot_res_elnb.dta`.
- `tablea4-entry25.do`: Generate lines 7–8 of **Table A4** (Robustness of the strategic spillover parameter estimates: Alternative entry criteria: 25% percentile).
- `tablea4-entry75.do`: Generate lines 5–6 of **Table A4** (Robustness of the strategic spillover parameter estimates: Alternative entry criteria: 75% percentile).
- `tablea4-first-stage-prep.r`: To be run before `tablea4-first-stage.r`.
  - Outputs data file `results-loo-first-stage.csv`.
- `tablea4-first-stage.r`: Generate lines 11–12 of **Table A4** (Robustness of the strategic spillover parameter estimates: Alternative first-stage model).
  - Must be run after `tablea4-first-stage-prep.r`, which outputs the required data file `results-loo-first-stage.csv`.
- `tablea4-single-entrant.r`: Generate lines 1–4 of **Table A4** (Robustness of the strategic spillover parameter estimates: Including single-entrant municipality-periods).
  - Must be run after `table3-prep.do`, which outputs the required data file `phats.csv`.
- `tablea4-threshold.do`: Generate lines 9–10 of **Table A4** (Robustness of the strategic spillover parameter estimates: Alternative threshold of victimization).
- `tablea5-intensity-control.do`: Generate **Table A5, Model 3** (Initial intensity of violence as control).
  - Outputs data file `boot_res_region_v_inten.dta`.
- `tablea5-interaction.do`: Generate **Table A5, Model 2** (Time invariant controls interacted with period 2002-2005 dummy).
  - Outputs data file `boot_res_region_inter.dta`.
- `tablea5-state-effects.do`: Generate **Table A5, Model 1** (State effects).
  - Outputs data file `boot_res_deptod.dta`.
- `tablea6.do`: Generate **Table A6** (Strategic and Bivariate Normal Models Comparison).

- `tablea7.do`: Generate **Table A7** (Strategic Victimization (municipality-year)).
  - Outputs data file `boot_res_region_v_year.dta`.

2.3 Other

- `Makefile`: Makefile to generate all output via GNU Make. Run `make all` at the command line to run all scripts in the necessary order. Requires command line access to STATA and R. Will take 2+ days to run on circa-2021 hardware.
- `README.pdf`: This file.
- `README.txt`: Markdown source code for this file.

3 Software Versions and Packages

All STATA analyses were run in STATA 17.0. All R analyses were run in R 4.1.2 using the following packages:

- broom 0.7.9
- car 3.0.12
- caret 6.0.90
- dplyr 1.0.7
- foreach 1.5.1
- foreign 0.8.81
- Formula 1.2.4
- ggmap 3.0.0
- ggplot2 3.3.5
- ggpubr 0.4.0
- haven 2.4.3
- iterators 1.0.13
- maptools 1.1.2
- matrixStats 0.61.0
- maxLik 1.5.2
- plyr 1.8.6
- randomForest 4.6.14
- rgdal 1.5.27
- rgeos 0.5.8
- rootSolve 1.8.2.3
- stringi 1.7.5
- stringr 1.4.0
- tidyverse 1.3.1
- xtable 1.8.4

4 Codebooks for Tabular Data Files

Entries are numbered in column order.

4.1 `colombia_latlong.csv`

The first four lines of this file must be removed before importing.

1. **Código Departamento:** Department numeric identifier.
2. **Código Municipio:** Municipality numeric identifier.
3. **Código Centro Poblado:** Population center numeric identifier.
4. **Nombre Departamento:** Department name.
5. **Nombre Municipio:** Municipality name.
6. **Nombre Centro Poblado:** Population name.
7. **Tipo Centro Poblado:** Category of population center.
8. **Longitud:** Longitude.
9. [not used]
10. **Latitud:** Latitude.
11. **Distrito:** District (if applicable).
12. **Tipo de Municipio:** Category of municipality.
13. **Area Metropolitana:** Metropolitan area name.

4.2 `code_muni_cede.dta`

1. **departamento:** Department name.
2. **dpto_code:** Department numeric identifier.
3. **municipio:** Municipality name.
4. **muni_code:** Municipality numeric identifier.

4.3 `group_origins.csv`

1. **Group:** Categorical: AUC, ELN, or FARC.
2. **municipio:** Municipality name.
3. **departamento:** Department name.
4. **muni_code:** Municipality numeric identifier.
5. **Source 1:** Data source.
6. **Five year period:** Beginning of five-year period used to indicate early influence.

4.4 `military_bases.csv`

1. **force:** Categorical: Airforce, Navy, or Army.
2. **muni_code:** Municipality numeric identifier.
3. **municipio:** Municipality name.
4. **departamento:** Department name.
5. **year:** Year of observation.
6. **name:** Name of base.
7. **Notes:** Miscellaneous notes.
8. **Website:** Data source.
9. **Counterinsurgency:** Binary indicator for counterinsurgency involvement (not used in analysis).

4.5 `violence_t_export.dta`

1. **victim_farc:** Indicator for systematic civilian victimization by the FARC in this municipality-period. (Dependent variable in the main analysis.)
2. **victim_paras:** Indicator for systematic civilian victimization by the paramilitaries in this municipality-period. (Dependent variable in the main analysis.)
3. **nbi_t:** Fraction of the population with unsatisfied basic needs. ("Poverty" in regression tables.)
4. **royalties_t:** Share of municipal revenue from oil exploitation. ("Oil royalties.")
5. **coca_t:** Area of the municipality where coca is grown. ("Coca area.")
6. **share_left:** Vote share for left-wing party in prior national election. ("Liberal party vote share.")
7. **army_dist:** Distance from municipality to the nearest army base. ("Distance army base.")
8. **dmr:** Distance from municipality to the Magdalena river. ("Distance Magdalena river.")
9. **evlp:** Variation in liberal party's election share in 1974–1994 presidential elections. ("Variation liberal party vote share.")
10. **time:** Binary indicator for the 2002–2005 time period. ("Period 2002–2005.")
11. **muni_code:** Municipality numeric identifier.
12. **action_paras2:** Categorical indicator for type of civilian victimization by paramilitaries in this municipality-period. 1 = No victimization, 2 = Selective victimization, 3 = Non-selective victimization. (Dependent variable in the selective/non-selective analysis reported in Table A2.)
13. **action_farc2:** Categorical indicator for type of civilian victimization by the FARC in this municipality period. 1 = No victimization, 2 = Selective victimization, 3 = Non-selective victimization. (Dependent variable in the selective/non-selective analysis reported in Table A2.)
14. **gcaribe:** Binary indicator for municipality located in the Caribbean Region.
15. **gpacifica:** Binary indicator for municipality located in the Pacific Region.
16. **gorinoquia:** Binary indicator for municipality located in the Orinoquía Region.
17. **gamazonia:** Binary indicator for municipality located in the Amazon Region.
18. **coddepto:** Department numeric identifier.
19. **gini_i:** Gini coefficient. ("Gini.")
20. **victim_eln:** Indicator for systematic civilian victimization by the ELN in this municipality-period. (Dependent variable in the three-groups analysis reported in Table A3.)
21. **victim25_farc:** Indicator for systematic civilian victimization by the FARC in this municipality-period, using the 25th percentile alternative threshold. (Dependent variable in the "Alternative threshold of victimization" analysis reported in Table A4.)
22. **victim25_paras:** Indicator for systematic civilian victimization by the paramilitaries in this municipality-period, using the 25th percentile alternative threshold. (Dependent variable in the "Alternative threshold of victimization" analysis reported in Table A4.)
23. **victim_farc2:** Indicator for systematic civilian victimization by the FARC in this municipality-period, using the 75th percentile alternative entry criterion. (Dependent variable in the "Alternative entry criteria, 75% percentile" analysis reported in Table

A4.)

24. **victim_paras2:** Indicator for systematic civilian victimization by the paramilitaries in this municipality-period, using the 75th percentile alternative entry criterion. (Dependent variable in the "Alternative entry criteria, 75% percentile" analysis reported in Table A4.)
25. **victim_farc3:** Indicator for systematic civilian victimization by the FARC in this municipality-period, using the 25th percentile alternative entry criterion. (Dependent variable in the "Alternative entry criteria, 25% percentile" analysis reported in Table A4.)
26. **victim_paras3:** Indicator for systematic civilian victimization by the paramilitaries in this municipality-period, using the 25th percentile alternative entry criterion. (Dependent variable in the "Alternative entry criteria, 25% percentile" analysis reported in Table A4.)
27. **total_in:** Sum of all violence incidents involving the FARC, paramilitaries, and ELN at the start of the period. ("Initial violence incidents" in Table A5, Model 3.)
28. **farc_dist:** Distance by road from the municipality to the closest area of early FARC influence. ("Distance group's place of origin" for the FARC.)
29. **paras_dist:** Distance by road from the municipality to the closest area of early paramilitary influence. ("Distance group's place of origin" for the paramilitaries.)
30. **eln_dist:** Distance by road from the municipality to the closest area of early ELN influence. ("Distance group's place of origin" for the ELN.)
31. **lpobl_tot_t:** Total population of the municipality, logged. ("ln(Population).")

4.6 `violence_t_year_export.dta`

Except where noted, all variables are the same as in `violence_t_export.dta` above, except coded at the municipality-year level rather than the municipality-period.

1. **victim_farc**
2. **victim_paras**
3. **nbi_t**
4. **royalties_t**
5. **coca_t**
6. **share_left**
7. **army_dist**
8. **dmr**
9. **evlp**
10. **year:** Year of observation.
11. **muni_code**
12. **action_paras2**
13. **action_farc2**
14. **gcaribe**
15. **gpacifica**
16. **gorinoquia**
17. **gamazonia**

18. **coddepto**
19. **gini_i**
20. **victim_eln**
21. **farc_dist**
22. **paras_dist**
23. **eln_dist**
24. **lpobl_tot_t**
25. **year_d1:** Binary indicator for observation in 1998.
26. **year_d2:** Binary indicator for observation in 1999.
27. **year_d3:** Binary indicator for observation in 2000.
28. **year_d4:** Binary indicator for observation in 2001.
29. **year_d5:** Binary indicator for observation in 2002.
30. **year_d6:** Binary indicator for observation in 2003.
31. **year_d7:** Binary indicator for observation in 2004.