Replication package for: All Eyes on Them: A Field Experiment on Citizen Oversight and Electoral Integrity

Natalia Garbiras-Díaz and Mateo Montenegro

May 3, 2022

README

Overview

This README provides guidelines to replicate the analysis from "All Eyes on Them: A Field Experiment on Citizen Oversight and Electoral Integrity" using Stata and R. It contains instructions for how to access the data (data availability statement) and also detailed instructions for replicating the results. One master do file runs all of the code to generate the data and produce the figures and tables in the main paper and Online Appendix. The estimated time to run this code is approximately 9 hours and 20 minutes.

Data Availability and Provenance Statements

Statement about Rights

We certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript. Some of the data may be accessed through "Access to/Freedom of Information" or similar requests. When necessary, we document in this README the procedure for doing so.

Summary of Availability

Although most of the data used in the analysis are publicly available and have been published in the repository, some other data **cannot** be made public. Specifically, we cannot release pre- and post-treatment survey data, in accordance with IRB restrictions and the informed consent given to survey participants when they agreed to participate in the study.

Details on each Data Source

In this section, we list the six sources we use to construct the master databases that we use in the analysis. For each one, we provide a description, the details about how to obtain the data, the folder location in the replication repository, and the analysis that require the data. The names of the files containing data from each source are listed in Table B below.

Departamento Administrativo Nacional de Estadística (DANE)

- **Description:** The National Administrative Department of Statistics (DANE, for its acronym in Spanish), Colombia's National Statistics Office, provides the following municipal-level covariates:
  - *Municipal population:* This variable is estimated using data from the country's most updated Census, conducted in 2018. The data is publicly available and can be downloaded from the DANE website at this link: https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion.

  - *Municipal GDP:* The data is publicly available and can be downloaded from the DANE website at this link: https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-nacionales-departamentales.

- **Format:** Both datasets are in comma separated values format (`.csv`).
- **Location in repository:** `Data/Raw/...`
- **Analyses requiring data:** Municipal-level balance tests, and variables included in the set used to select the covariates (using the double post-lasso covariate selection method) in the regressions described in equations (1) and (2) of the article.

Misión de Observación Electoral (MOE)

- **Description:** The Electoral Observation Mission (MOE) is an NGO that monitors elections in Colombia and also one of the two implementing partners of the intervention. The MOE shared with the authors the number of citizen reports filed through data "Pilas con el voto", its online reporting platform, for the 2015, 2018 and 2019 elections in Colombia. To ensure the protection and confidentiality of reporting individuals, all shared data were aggregated at the municipal level.
- **How to obtain the data:** Replicators may request the data by sending a request to the MOE through any of their contact channels. Visit the MOE's website at https://www.moe.org.co/contacto/, and the designated site for "Pilas con el voto" at https://www.pilasconelvoto.com/. The authors were granted permission to redistributed the aggregated data for the purpose of this replication.
- **Format:** All datasets are in comma separated values format (`.csv`).
- **Location in repository:** `Data/Raw/...`
- **Analyses requiring data:** Estimation of the effects of the intervention on reports of electoral irregularities.

Registraduría Nacional del Estado Civil

- **Description:** The *Registraduría Nacional del Estado Civil* (henceforth referred to as Registraduría), Colombia's electoral office, has provided the authors with detailed data on election results at the voting booth-level (which corresponds to the lowest level of aggregation at which election results are computed). The entity also provided detailed data on the names and individual characteristics of the universe of mayoral candidates who ran in the 2019 local elections in the country.
- **How to obtain the data:** These data are public official records and thus, according to Colombia's open data law, can be accessed through a freedom of information request (FOI). To do so, replicators must submit an official request through a designated form that can be found at the *Registraduría*'s website: https://wsp.registraduria.gov.co/contacto/solicitudes/. To file a data request, click on the link, then choose the option "Información Pública", and fill out the form indicating a brief description of the purpose of the FOI. The latter should stipulate that data must be disaggregated at the voting booth level (*mesa de votación* in Spanish).
- **Format:** The databases are in comma separated values format (`.csv`).
- **Location in repository:** `Data/Raw/...`
- **Analyses requiring data:** Data on the 2019 elections is used to estimate the effect of the intervention. Data from previous elections is used to conduct balance tests, and is included in the set used to select the covariates (using the double post-lasso covariate selection method) in the regressions described in equations (1) and (2) of the article.

Centro de Estudios Económicos (CEDE)

- **Description:** The Center of Economic Development Studies (CEDE, for its acroyms in Spanish), hosted by the Department of Economics at the Universidad de Los Andes in Colombia, compiles and updates a panel dataset with general characteristics, economic, social, and other related variables on the universe of municipalities in Colombia.
- **How to obtain the data:** Data is subject to a redistribution restriction, but can be freely downloaded from https://datoscede.uniandes.edu.co/es/catalogo-de-microdata. Active students and alumni of the

university can automatically download the data using their institutional email account. To do so, click on "`Panel Municipal del CEDE`," and then log in using the email account and password. External users can request access to it. To do so, replicators must send an email to datoscede@uniandes.edu.co including a brief description of the project, and agree to the conditions of use.
- **Format:** Datasets are all Stata format (`.dta`).
- **Location in repository:** `Data/Raw/`....
- **Analyses requiring data:** Municipal-level balance tests, and variables included in the set used to select the covariates (using the double post-lasso covariate selection method) in the regressions described in equations (1) and (2) of the article.

Facebook

- **Description:** We use two types of Facebook databases: The first corresponds to publicly available data that can be downloaded directly from the Facebook API, corresponding to the potential reach at the municipal level (i.e. active users of the social media platform) in the last 30 days. The second set of data corresponds to metrics on the ad campaign in terms of performance and activity. The data is part of the standard report provided by the Ads Manager site for those advertising on Facebook.
- **How to obtain the data:** Data on potential reach at the municipal level can be hand-coded from Facebook's API. Metrics on treatment ad campaign is private and only shared with the owners of the campaign.
- **Format:** The databases are in comma separated values format (`.csv`).
- **Location in repository:** `Data/Raw/`....
- **Analyses requiring data:** (a) Potential Facebook reach is used for balance tests, and is included in the set used to select the covariates (using the double post-lasso covariate selection method) in the regressions described in equations (1) and (2) of the article. (b) Data on the performance of the ad campaign is used to construct summary statistics of the performance of the Facebook ad campaign.

Siglo Data

- **Description:** *Siglo Data* is a data analysis firm based in Colombia that specializes on monitoring and classifying news in mass media and social networks. We use election news collected by this company as part of our media-based measure of election irregularities.
- **How to obtain the data:** The news database was purchased by the authors. Those interested in purchasing this data should contact the Siglo Data sales representative through the contact form available on the company's website (https://siglodata.com/#contacto) or by sending an email to contacto@siglodata.com.
- **Format:** The database is in comma separated values format (`.csv`).
- **Location in repository:** `Data/Raw/`....
- **Analyses requiring data:** Estimation of equation (1) to test for the effect of the intervention on electoral irregularities.

PARES

- **Description:** The *Fundación Paz y Reconciliación* (PARES) is an NGO in Colombia focused on producing independent research on conflict, security, governance, and democracy, among others. In 2019 they produced a report of instances of electoral irregularities gathered from media, citizen reports and their own monitoring activity in the field. See more details about this report in Online Appendix B.
- **How to obtain the data:** The data was handcoded by the authors based on the report "Informe completo II: 'Candidatos cuestionados 2019'," which contains the results of this investigation and provides the relevant data in the appendix. Visit PARES https://www.pares.com.co/post/informe-completo-ii-candidatos-cuestionados-2019
- **Format:** The database is in comma separated values format (`.csv`).
- **Location in repository:** `Data/Raw/`....

- **Analyses requiring data:** Robustness tests of results on the effect of the intervention on the vote share of candidates most likely to engage in electoral irregularities. This NGO provides for an alaternative measure to classify candidates with a past criminal history.

Original individual surveys

- **Description:** The authors collected two original individual surveys to complement the analysis of the effects of the intervention. A first one was conducted prior to launching the intervention (referred to as "pre-treatment survey"), and a second one was conducted afterwards (referred to as "post-treatment survey"). Section C in the articule and Online Appendix F provide detailed descriptions on the recruitment of participants for both the pre- and post-treatment surveys. Survey responses were recorded on Qualtrics, an online polling software.
- **How to obtain the data:** The data used in the article contains sensitive and potentially identifiable information, so access is restricted. It is only available for academic purposes, and access must be granted through the following process. First, researchers must receive approval from their university's IRB and complete a Data Use Agreement between their university and ICPSR, which details the steps applicants will take to protect the confidentiality of the data. Once this procedure is completed and uploaded to ICPSR, ICPSR will review the application and decide whether or not to grant access to the data. The data must be requested and downloaded separately from the repository that contains them (Garbiras-Díaz and Montenegro, 2022). For any questions regarding these data, please send an e-mail to the corresponding co-author at mateomontenegro@gmail.com.
- **Format:** The databases are in comma separated values format (`.csv`).
- **Location in repository:** Restricted-use repository.
- **Analyses requiring data:** (a) estimating the candidate-specific heterogenous effects in equation (2).(b) Candidate-level balance tests, and variables included in the set used to select the covariates (using the double post-lasso covariate selection method) in the regressions described in equation (2) of the article.

Dataset list

The descriptions of all of the variables included in the analysis are detailed in file `Variable_Codebook.xslx`. The following table lists all the raw datasets in the `Data/Raw` folder of the replication folder.

Table A: List of Datasets

File name	Source	Notes	Description
anexo-2019-provisional-valor-agregado-municipio-2011-2019.csv	DANE	Public	Municipal GDP
facebook_ad_metrics.csv	Facebook Ad Manager	Private	Advertisment metrics
fb_pre_users.csv	Coded based on Facebook API	Public	Facebook users per municipality
identifier_crosswalk.csv	Registraduria	Public	Crosswalk between DANE and Registraduria identifiers
media_irregularities.csv	Propietary data	Private	News about irregularities
moe_reportes_24_a_28_octubre2019.csv	MOE	Private	MOE citizen reports during intervention period
moe_reportes_alcaldia_2015_short_window.csv	MOE	Private	MOE citizen reports in 2015 during days matching the intervention period
moe_reportes_alcaldia_2015.csv	MOE	Private	MOE citizen reports in Mayoral elections 2015
moe_reportes_congreso_2018.csv	MOE	Private	MOE citizen reports in Congressional elections 2018
moe_reportes_despues_28_octubre2019.csv	MOE	Private	MOE citizen reports after intervention period
PANEL_CARACTERISTICAS_GENERALES(2020).dta	CEDE	Private	Municipal characteristics
PANEL_CONFLICTO_Y_VIOLENCIA2017.dta	CEDE	Private	Homicide Variable
PARES_candidate_past_malfeasance.csv	Hand-coded based on PARES report	Public	History of past malfeasance by candidates
post_svy_irregularity_outcomes.csv	Post-Treatment Survey	Private	Post-treatment survey outcomes
post_svy_number_respondents.csv	Post-Treatment Survey	Private	Number of respondents in post-treatment survey
post_svy_respondent_vars.csv	Post-Treatment Survey	Confidential	Characteristics of respondents in post-treatment survey
Potencial por mesa Elecciones Congreso 2018.csv	Registraduria	Private	Registered voters in Congressional elections 2018
Potencial por puesto Elecciones Alcaldia 2019.csv	Registraduria	Private	Registered voters in Mayoral elections 2019
pre_svy_bootstrap_data.csv	Pre-treatment survey	Confidential	Individual responses from pre-treatment survey used for bootstrap
pre_svy_candidate_variables.csv	Pre-treatment survey	Private	Candidate-level variables from pre-treatment survey
pre_svy_number_respondents.csv	Pre-treatment survey	Private	Number of respondents in pre-treatment survey
pre_svy_respondent_vars.csv	Pre-treatment survey	Confidential	Characteristics of respondents in pre-treatment survey
proyecciones-poblacion-Municipal_2018-2026.csv	DANE	Public	Municipal population
Registraduría - Datos Personales Candidatos.csv	Registraduria	Public	Candidate Demographics
RESULTADOS_ELECTORALES_2014_PRESIDENCIA_SEGUNDA_VUELTA.csv	Registraduria	Public	Presidential election results 2014
RESULTADOS_ELECTORALES_2015_ALCALDIA.csv	Registraduria	Public	Mayoral election results 2015
RESULTADOS_ELECTORALES_2018_SENADO_DE_LA_REPUBLICA.csv	Registraduria	Public	Senate election results 2018
RESULTADOS_ELECTORALES_2019_ALCALDIA.csv	Registraduria	Public	Mayoral election results 2019
treatment_assignment.csv	Authors	Private	Treatment assignment in experiment

Computational requirements

Software Requirements

- Stata (code was last run with version 15, with all packages updated on 02/04/2022). The following packages are required to run the programs. They are all installed by running the "*aer_prep_installing_stata_programs.do*" do-file.
  - `egenmore`
  - `outreg`
  - `reclink`
  - `estout`
  - `mgof`
  - `moremata`
  - `lassopack`
  - `mat2txt`
  - `ritest`
  - `distinct`
- R (code was last run with version 4.0.0, with all packages updated on 02/04/2022). The following package is required to run the programs. It is installed by line 19 of the "*aer_balance_plots.R*" script.
  - `readxl`

Controlled Randomness

We rely on a random number generator (set at line 14 of the "*aer_master.do*") for the following analysis:

- Computation of randomization inference p-values.
- Bootstrapping p-values following the procedure explained in Section G of the Online Appendix (to account for the variance in estimating some of the candidate-level variables).
- Simulating the p-values of the forensic tests used in the paper to account for finite samples (as described in Section D of the Online Appendix).

Memory and Runtime Requirements

Summary The approximate time needed to reproduce all of the analyses on a standard 2019 computer is 9 hours and 20 minutes. The estimated time to reproduce only the parts of the code that do not use confidential data is approximately 6 hours and 20 minutes.

Details The code was last run on a **8-core laptop with a 2.3 GHz Intel Core i9 processor and MacOS version 11.6**.

Description of replication folder

The replication folder is organized as follows:

- `Code:` this folder contains the scripts that run the code to replicate the analysis. One single master do file runs all the necessary scripts that produce all the figures, tables and analysis in the paper. These scripts are organized in the following subfolders:
  - `Preparation:` The scripts in this folder creates the two main database used in the analysis: one at the municipal-level and one at the candidate-level.
  - `Analysis:` The scripts in this folder produce the numbers, figures, and tables in the main article and Online Appendix.
- `Data:` This folder provides the input datasets that are required to produce the analysis. This folder contains three subfolders:

  - `Raw:` Contains the raw datasets from the main sources describe in the previous sections.
  - `Intermediate:` Datasets in this folder correspond to ancillary files that are produced in the cleaning, processing and anlysis of the datasets contained in the `Raw` folder.
  - `Final:` The datasets in this folder are the two master databases used to reproduce all of the tables and figures in the main article and the robustness tests described in the Online Appendix.
- `Output:` this folder contains all the figures (in PDF format) and the input (as .tex files) to produce the tables that are included in the paper. The code will store the results in this folder.

Description of code

In this section, we describe the programs that are included in the folders contained in the `Code` folder of the replication folder.

1. `aer_master.do`: Runs in corresponding order all of the Stata code. It includes a preamble that installs all of the required commands and programs required to run the code.

2. `Preparation:`

   - `aer_prep_installing_stata_programs.do`: This code installs of the Stata programs necessary to run the rest of the replication code.
   - `aer_prep_election_controls.do`: This code prepares the political covariates from the 2014, 2015 and 2018 elections.
   - `aer_prep_facebook_ad_metrics.do`: This code prepares the measures of engagement and reach of the Facebook ads.
   - `aer_prep_pre_svy_variables.do`: This code prepares the measures reflecting the likelihood that candidates will engage in irregularities and their popularity from the pre-treatment survey, as well as the number of respondents to the survey, their demographic characteristics, and their individual responses about candidate irregularity likelihood.
   - `aer_prep_post_svy_variables.do`: This code prepares the irregularity outcomes from the post-treatment survey, as well as the number of respondents and their demographic characteristics.
   - `aer_prep_treatment_indicators.do`: This code prepares the treatment indicators and strata fixed effects for each municipality.
   - `aer_prep_socioeconomic_covs.do`: This code prepares the municipal socieconomic covariates coming from the CEDE dataset.
   - `aer_prep_region_indicator_covs.do`: This code prepares the geographic region indicators for each municipality.
   - `aer_prep_population.do`: This code prepares the 2018 population variables.
   - `aer_prep_moe_reports.do`: This code prepares the outcomes and covariates of citizen reports to the MOE in 2019, 2018 and 2015.
   - `aer_prep_media_irregularity_outcomes.do`: This code prepares the media-based measures of irregularities.
   - `aer_prep_incumbent_party.do`: This code prepares the indicators for whether a candidates' party was incumbent using the strict and loose definitions described in the paper.
   - `aer_prep_gdp2016.do`: This code prepares the covariate of the 2016 municipal GDP.
   - `aer_prep_forensic_outcomes_2019.do`: This code computes the forensic tests and (simulated) p-values for the second and last digits of the 2019 election counts at the voting booth level.
   - `aer_prep_forensics_2015.do`: This code computes the forensic tests and (simulated) p-values for the second digits of the 2015 election counts at the voting booth level.
   - `aer_prep_facebook_users.do`: This code prepares the covariate of the number of Facebook users per municipality.
   - `aer_prep_election_outcomes_2019.do`: This code prepares the political outcome and control variables from the 2019 elections at both municipality and candidate level.
   - `aer_prep_candidate_past_malfeasance.do`: This code prepares the indicator for whether candidates have a history of past malfeasance acording to the NGO PARES.

- `aer_prep_candidate_demographic_covs_2019.do`: This code prepares the candidate demographic variables (age and sex).

3. `Analysis`:

- `aer_balance_plots.R`: This code generates plots displaying the p-values for the main balance checks.
- `aer_build_main_datasets.do`: This code executes all the programs preparing and cleaning the data used in the paper, labels the variables, and builds the main datasets.
- `aer_figures.do`: This code produces the figures in the main text and the appendix.
- `aer_tables.do`: This code produces the tables in the main text and the appendix.
- `aer_wildbootstrap.do`: This code produces the bootstrap p-values described in Appendix D to account for the variance in estimating the candidate-level variables.
- `aer_wildbootstrap_no_controls.do`: This code produces the bootstrap p-values described in Appendix D to account for the variance in estimating the candidate-level variables for Table A20 (without controls).

Instructions to Replicators

1. If you **do not** have access to confidential data:

- Download all of the data and code files in folder with names mimicking the ones in the repository.
- Change the main directory path in line 21 of the `Code/aer_main.do` do-file.
- Run the `Code/aer_main.do` do-file.
- Run the `Code/Analysis/aer_balance_plots.R` script.

2. If you **do** have access to confidential data:

- Download all of the data and code files in folder with names mimicking the ones in the repository. All confidential raw data should be in the `Data/Raw` directory.
- Change the main directory path in line 20 of the `Code/aer_main.do` do-file.
- Change the value of the global "confidential" to one on line 43 of the `Code/aer_main.do` do-file.
- Run the `Code/aer_main.do` do-file.
- Run the `Code/Analysis/aer_balance_plots.R` script.

Details

- Please note that the above instructions should be followed in order to reproduce the results correctly.
- If running independent parts of the code, it might be needed to run all the programs preceding any particular one.
- The construction of Figure A3 was not scripted. To replicate it, please follow these instructions:
  1. Download the shapefile of Colombian municipal boundaries from the DANE website, https://geoportal.dane.gov.co/servicios/descarga-y-metadatos/descarga-mgn-marco-geoestadistico-nacional/, by clicking on option "Nivel Geográfico Municipio."
  2. Merge the shapefile using the 5-digit municipal identifier with the final data set, "Replication/Data/Final/main_data_municipal_level.dta."
  3. Display the municipalities in the experimental sample (i.e., those for which variable `_assignment` is not missing).

List of tables and programs

The provided code reproduces all tables, figures, and numbers provided in text in the paper.[1] Table B lists all the Figures and Tables that are included in the `Output` folder of the replication folder.

1. Notice that some of these require access to restricted-use data.

Table B: List of Figures and Tables in paper and Online Appendix

Figure	Script name	Line code	Output file	Location	Notes
4	aer_figures.do	57-167	graph_any_effect_d_reportMOE_any_nw.pdf, graph_any_effect_reportMOE_any_nw.pdf	Main text	
5	aer_figures.do	176-287	graph_any_effect_d_media_irreg_nw.pdf, graph_letter_effect_d_media_irreg_nw.pdf	Main text	
6	aer_figures.do	296-406	graph_any_effect_sig95_max_second_nw.pdf, graph_letter_effect_sig95_max_second_nw.pdf	Main text	
A1	aer_figures.do	426-253	graph_A1.pdf	Appendix	
A4	aer_figures.do, aer_balance_plots.R	462-508	graph_balance_1.pdf, graph_balance_2.pdf	Appendix	

Table	Script name	Line code	Output file	Location	Notes
2	aer_tables.do	60-130	table2.tex	Article	
3	aer_tables.do	139-334	MOEreport_table_A.tex, MOEreport_table_B.tex, MOEreport_table_C.tex	Article	
4	aer_tables.do	342-539	irregularities_table_A.tex, irregularities_table_B.tex, irregularities_table_C.tex	Article	
5	aer_tables.do	551-767	elections_table_A.tex, elections_table_B.tex, elections_table_C.tex, wildbootrap_pvals.txt	Article	Code fragment replicating bootstrap p-values requires access to restricted data.
A1	aer_tables.do	783-834	table_a1.tex	Appendix	
A2	aer_tables.do	843-934	Files in the "Balance" folder not starting by "balance_can," "balance_svy or "balance_post_svy"	Appendix	
A3	aer_tables.do	945-1050	balance_svy_birth_year.tex, balance_svy_female.tex, balance_svy_high_school.tex, balance_svy_high_school_more.tex	Appendix	Requires access to restricted data.
A4	aer_tables.do	1058-1208	lateMOEreports_table_A.tex, lateMOEreports_table_B.tex, lateMOEreports_table_C.tex	Appendix	
A5	aer_tables.do	1217-1369	media_inclMOE_table_A.tex, media_inclMOE_table_B.tex, media_inclMOE_table_C.tex	Appendix	

Continued on next page

Table B – continued from previous page

Table	Script name	Line code	Output file	Location	Notes
A6	aer_tables.do	1380-1532	media_bytype_table_A.tex, media_bytype_table_B.tex, media_bytype_table_C.tex	Appendix	
A7	aer_tables.do	1544-1697	media_app_table_A.tex, media_app_table_B.tex, media_app_table_C.tex	Appendix	
A8	aer_tables.do	1706-1741	correlation_media_forensics_table.tex	Appendix	
A9	aer_tables.do	1750-1823	correlation_media_forensics_bytype_tableA.tex, correlation_media_forensics_bytype_tableB.tex, correlation_media_forensics_bytype_tableC.tex, correlation_media_forensics_bytype_tableD.tex, correlation_media_forensics_bytype_tableE.tex, correlation_media_forensics_bytype_tableF.tex, correlation_media_forensics_bytype_tableG.tex	Appendix	
A10	aer_tables.do	1832-1984	forensics_app_table_A.tex, forensics_app_table_B.tex, forensics_app_table_C.tex	Appendix	
A11	aer_tables.do	1994-2145	forensics_lastdig_table_A.tex, forensics_lastdig_table_B.tex, forensics_lastdig_table_C.tex	Appendix	
A12	aer_tables.do	2154-2253	Files in the "Balance" folder starting by "balance_can"	Appendix	
A13	aer_tables.do	2263-2295	correlation_cuestionado_table.tex	Appendix	
A14	aer_tables.do	2305-2356	correlation_can_variables_table.tex	Appendix	
A15	aer_tables.do	2365-2532	elections_app_table_A.tex, elections_app_table_B.tex, elections_app_table_C.tex	Appendix	
A16	aer_tables.do	2541-2693	elections_mun_table_A.tex, elections_mun_table_B.tex, elections_mun_table_C.tex	Appendix	
A17	aer_tables.do	2703-2863	table_a17.tex	Appendix	
A18	aer_tables.do	2873-2987	MOEreport_table_nocontrols_A.tex, MOEreport_table_nocontrols_B.tex, MOEreport_table_nocontrols_C.tex	Appendix	

Continued on next page

Table B – continued from previous page

Table	Script name	Line code	Output file	Location	Notes
A19	aer_tables.do	2996-3110	irregularities_table_nocontrols_A.tex, irregularities_table_nocontrols_B.tex, irregularities_table_nocontrols_C.tex	Appendix	
A20	aer_tables.do	3120-3254	elections_table_nocontrols_A.tex, elections_table_nocontrols_B.tex, elections_table_nocontrols_C.tex, wildbootrap_pvals_nocontrols.txt	Appendix	
A21	aer_tables.do	3263-3383	long_reg_MOEreports.tex	Appendix	
A22	aer_tables.do	3391-3512	long_reg_irregularities.tex	Appendix	
A23	aer_tables.do	3521-3656	long_reg_candidates.tex	Appendix	
A28	aer_tables.do	3665-3778	Files in the "Balance" folder starting by "balance_post_svy"	Appendix	
A29	aer_tables.do	3788-3885	balance_post_svy_female_post_svy.tex, balance_post_svy_birth_year_post_svy.tex, balance_post_svy_high_school_post_svy.tex, balance_post_svy_high_school_more_post_svy.tex	Appendix	Requires access to restricted data.
A30	aer_tables.do	3893-4044	survey_irregs_table_A.tex, survey_irregs_table_B.tex, survey_irregs_table_C.tex	Appendix	

References

Acevedo, Karina, and Iván Bornacelly. 2014. *Panel municipal del CEDE [dataset].* Accessed via Documentos CEDE 2014-26 [distributor]. https://repositorio.uniandes.edu.co/handle/1992/8510,2022-02-07. (accessed 2022-02-01).

Centro de Estudios sobre Desarrollo Económico. 2014a. *Panel municipal - Características generales [dataset].* Retrieved from Universidad de los Andes [distributor]. https://datoscede.uniandes.edu.co/es/catalogo-de-microdata. (accessed 2021-09-01).

———. 2014b. *Panel municipal - Violencia y conflicto [dataset].* Retrieved from Universidad de los Andes [distributor]. https://datoscede.uniandes.edu.co/es/catalogo-de-microdata. (accessed 2018-06-01).

Departamento Administrativo Nacional de Estadística. 2018. *Proyecciones de población [dataset].* Retrieved from Departamento Administrativo Nacional de Estadística [distributor] Censo 2018, https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion. (accessed 2022-02-01).

———. 2020a. *Geoportal - Nivel geográfico municipio [dataset].* https://geoportal.dane.gov.co/servicios/descarga-y-metadatos/descarga-mgn-marco-geoestadistico-nacional/. (accessed 2022-02-01).

———. 2020b. *PIB por departamento [dataset].* Retrieved from Departamento Administrativo Nacional de Estadística [distributor] 2021-06-25. https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-nacionales-departamentales. (accessed 2022-02-01).

Facebook Ads Manager. 2019. *Report: Metrics Ad Campaign [dataset].* Unpublished data.

Fundación Paz y Reconciliación. 2020. *Segunda entrega: candidatos cuestionados a alcaldías y gobernaciones [dataset].* 2019-10-11 https://www.pares.com.co/post/informe-completo-ii-candidatos-cuestionados-2019. (accessed 2021-08-01).

Garbiras-Díaz, Natalia, and Mateo Montenegro. 2022a. "Supplementary Data for: All Eyes on Them: A Field Experiment on Citizen Oversight and Electoral Integrity." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/E160921V1.

———. 2022b. "Supplementary Data for: All Eyes on Them: A Field Experiment on Citizen Oversight and Electoral Integrity [restricted-access]." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/E161701V1.

Misión de Observación Electoral. nd. *Citizen Reports to "Pilas con el Voto" 2015-2019 [dataset].* Unpublished data.

Registraduría Nacional del Estado Civil de Colombia. nd. *Election Results [dataset].* Unpublished data.

Siglo Data. 2019. *News Coverage of the Elections [dataset].* Unpublished data.

Acknowledgements

This README draws on the Social Science Data Editors template README available here.