Electoral precinct-level database for Mexican municipal elections

Bruno Calderón-Hernández[1,†], Horacio Larreguy[1,†*], John Marshall[2,†], and José Luis Pérez Castellanos[1,†]

[1]Instituto Tecnológico Autónomo de México, Economics, Mexico City, 01080, Mexico
[2]Columbia University, Political Science, New York, 10027, United States
[*]Corresponding author(s): Horacio Larreguy (horacio.larreguy@itam.mx)
[†]these authors contributed equally to this work

ABSTRACT

This paper introduces a database of electoral precinct-level election returns for Mexican municipal elections between 1994 and 2019. This database includes: (i) electoral precinct-level votes for each electoral coalition, the coalitions of the incumbent mayor and incumbent state governor, and the four most popular political parties; (ii) electoral precinct-level valid and total votes, the number of registered voters, and turnout; (iii) the partisan composition and municipal-level votes of the incumbent and runner-up electoral coalitions from the previous election; and (iv) the partisan composition of the state-level incumbent governor. This paper outlines the organization of this data, its sources, and key variables, and describes the processes used to standardize the data. This database has the potential to support the cross-sectional and longitudinal study of local Mexican elections over two decades using fine-grained precinct-level electoral returns that enable panel and regression discontinuity analyses.

Background and summary

A major change in the empirical study of comparative politics and political economy over recent decades has been the shift from across- to within-country analyses[1]. Comparisons within and across sub-national governments have both illuminated the workings of local government[1] and enabled researchers to more powerfully test broader theories by leveraging large numbers of sub-national units within more homogeneous contexts than cross-national analyses permit[2].

Analyzing local election results can shed light on several key aspects of electoral consequences and political behavior. For example, Ferraz and Finan[3] use municipal-level electoral outcomes and the random release of audit reports to show that audit outcomes significantly affect incumbents' electoral performance in Brazil. Moreover, they show that effects are stronger in municipalities where a local radio was present, suggesting an important role of local media in enhancing electoral accountability. Using precinct-level electoral outcomes and fine-grained data on local media coverage, Larreguy et al.[4] further provide causal evidence of local media's role in holding Mexican municipal governments accountable following the release of similar audit reports.

Mexico's federal structure, containing 31 states and Mexico City, and 2,476 municipalities, offers considerable promise for analyzing subnational elections since multi-party political competition was normalized in the 1990s. However, the systematic study of Mexican municipal elections faces two important challenges. First, state electoral institutes vary considerably, both across states and especially over time, in the quantity and quality of municipal election data they make publicly available online. Second, it is hard to construct measures of incumbent support because Mexican mayors could not seek re-election until 2018 and increasingly-pervasive cross-party municipal coalitions often change from election to election, but election returns are generally reported at the electoral coalition level without disaggregating between constituent political parties.

This paper presents a dataset that circumvents these two challengers. Building on Larreguy[5], Marshall[6], and Enríquez et al.[7]'s data collection, we compile and standardize electoral precinct-level returns for municipal elections from 1994 to 2019 in 31 Mexican states (excluding the delegations of Mexico City). Our data set includes 456,051 precinct-level electoral returns for 15,581 municipal elections, 62,341 unique electoral precincts, and 2,033 unique municipalities.

The dataset encompasses electoral precinct-level voting data for all electoral coalitions in every municipal election and further computes votes for the coalitions of the party of the incumbent mayor, the party of the state incumbent governor as well as the four most popular political parties—the PAN (from its Spanish acronym *Partido Acción Nacional*, National Action Party), PRI (*Partido Revolucionario Institucional*, Institutional Revolutionary Party), PRD (*Partido de la Revolución*

Democrática, Party of the Democratic Revolution), and MORENA (*Movimiento Regeneración Nacional*, National Regeneration Movement). The election returns were either obtained from state electoral institute websites or via freedom of information requests from these institutes. The data on incumbent electoral coalition votes involved manually identifying the party of the winning candidate in all elections where the electoral coalitions of the party of the incumbent mayor at the previous election were not the same or a subset of their coalitions in the current election. The data set also includes electoral precinct-level valid and total votes, and voter turnout whenever voter registration data is available.

Additionally, the dataset includes electoral data at higher levels of aggregation to facilitate analyses commonly used in local election research. These include lagged municipal-level voting data for the party or coalition of the incumbent mayor and runner-up and the partisan composition of their coalitions, which facilitate the implementation of municipal-level difference-in-differences and regression discontinuity designs. The dataset also provides information on the state-level incumbent governor's name and partisan coalition to permit analyses of heterogeneity and alignment between municipal and state governments.

The database is a valuable resource for researchers interested in Mexico's local electoral dynamics. Among other applications, it allows for a comprehensive analysis of political support and competition at the lowest geographical level at which election data is available—the electoral precinct—and municipal-level party incumbency effects.

Mexico's electoral background

Mexico's modern political landscape has been shaped by four major political parties: the PRI, PRD, PAN, and MORENA. Each party emerged from different historical contexts and ideological currents, contributing to the dynamic political environment in Mexico today.

Institutional Revolutionary Party

The PRI, originally founded as the PNR (*Partido Nacional Revolucionario*, National Revolutionary Party) in 1929 by Plutarco Elías Calles, was established to bring stability to Mexico following the turmoil of the Mexican Revolution (1910-1920)[8]. The revolution had left Mexico deeply fractured, with various factions vying for control. The PNR's initial purpose was to unify these political forces under a single institutional framework, preventing further violence and ensuring peaceful power transfer among the revolutionary elites[9]. Over time, the party evolved and rebranded itself, first as the PRM (*Partido de la Revolución Mexican*, Party of the Mexican Revolution) in 1938 under President Lázaro Cárdenas, and later as the PRI in 1946. For seven decades, the PRI maintained a near-monopoly on power by blending democratic and authoritarian practices, often controlling elections and using clientelism to hold on to power[10]. The PRI's national dominance ended in 2000 when it lost the presidency to PAN's Vicente Fox, marking the beginning of Mexico's modern multi-party system. However, other parties had slowly accumulated regional strongholds and took control of the federal legislature in 1997.

Party of the Democratic Revolution

The PRD was founded in 1989 as a response to the PRI's shift toward neoliberal economic policies in the 1980s. Disillusioned by the PRI's abandonment of its revolutionary social justice ideals, Cuauhtémoc Cárdenas—the son of revered former president Lázaro Cárdenas—led a faction known as the Democratic Current within the PRI. In 1988, after breaking away, Cárdenas formed the FDN (*Frente Democratico Nacional*, National Democratic Front) and ran for president[11]. The election was marred by widespread allegations of fraud when the government abruptly halted the vote count and declared PRI candidate Carlos Salinas de Gortari the winner, a move many believed stole the presidency from Cárdenas. After the contested election, Cárdenas and his allies established the PRD, uniting various leftist groups of former communists, socialists, and smaller nationalist factions. Throughout the 1990s and early 2000s, the party achieved some regional electoral successes, most notably when Cárdenas became mayor of Mexico City in 1997, the first time the position was filled through election rather than presidential appointment[12]. The PRD also secured seats in the national legislature and governorships in several states, becoming a major force in Mexican politics.

National Action Party

The PAN was founded in 1939 as a response to the perceived excesses of the post-revolutionary state, especially under President Lázaro Cárdenas' policies of land redistribution and nationalization of industries[13]. The party's founders, including Manuel Gómez Morín, envisioned the PAN as a center-right alternative that promoted democracy, Catholic social teaching, and free-market principles. The PAN's early support came from Mexico's urban middle class, business elites, and conservative Catholics who opposed the PRI's secular and populist-oriented policies[14]. For much of the 20th century, the PAN functioned as a "loyal opposition," advocating for electoral transparency and limited government intervention in the economy. Its breakthrough came in 2000, when Vicente Fox won the presidency, ending the PRI's 71-year rule. The PAN's subsequent presidencies, especially under Felipe Calderón (2006-2012), were marked by neoliberal reforms and an aggressive military-led campaign against the country's drug cartels, known as the "War on Drugs."

National Regeneration Movement

The party MORENA was officially founded in 2014 by Andrés Manuel López Obrador (widely known as AMLO), emerging from his disillusionment with the PRD. López Obrador, the PRD's presidential candidate in both 2006 and 2012, became increasingly dissatisfied with the party's compromises, especially its alignment with neoliberal policies. Seeking to establish a new political force, he founded MORENA, which rapidly became a significant left-wing populist party in Mexico[15]. MORENA's platform has centered on combating corruption, social justice, and economic nationalism, calling for reversing many of the neoliberal reforms previously implemented in sectors such as energy. The party also emphasizes wealth redistribution through expanded social programs. These policies resonate with López Obrador's vision of Mexico's "Fourth Transformation," which he describes as a continuation of three previous transformative eras in Mexican history: the War of Independence, the Reform War, and the Mexican Revolution. MORENA's rise to power was remarkably swift, with López Obrador winning the 2018 presidential election in a landslide, securing over 53% of the vote. His victory marked a decisive shift in Mexican politics, as it was the first time in modern history that a party other than the PRI or PAN held the presidency. MORENA has since dominated Mexican politics, significantly reshaping the country's political landscape.

Electoral institutions in Mexico

The history of Mexico's electoral system stems from the Constitution of 1917, which created local institutions in charge of organizing elections, such as the Registration Board (*Junta Empadronadora*) and the Electoral Colleges (*Colegios Electorales*)[16]. However, these institutions lacked independence from the federal government. In 1946, President Manuel Ávila Camacho passed the Federal Electoral Law (*Ley Federal Electoral*), which established the Federal Electoral Surveillance Commission (*Comisión Federal de Vigilancia Electoral*) to provide greater monitoring of electoral processes. But it still lacked autonomy from the executive branch, thus failing to ensure electoral fairness or transparency[16]. During the 1970s and 1980s, various reforms were passed, such as the Law of Political Organizations and Electoral Processes (*Ley de Organizaciones Políticas y Procesos Electorales*) in 1977, which allowed greater participation of political parties beyond the PRI.

Nonetheless, there were accusations of significant fraud during the 1988 elections. The aftermath led to the creation of the IFE (*Instituto Federal Electoral*, Federal Electoral Institute) in 1990 to assure free and transparent elections[17]. Initially, the IFE still depended on the executive power, but a reform in 1996 granted the IFE complete autonomy. However, although they followed the electoral precinct demarcation conducted by the IFE in 1994, which was only marginally updated over time to account for significant population growth, state and municipal elections continued to be conducted by state electoral institutes. In 2014, the IFE expanded into the INE (*Instituto Nacional Electoral*, National Electoral Institute), which assumed responsibility for organizing and overseeing state and local elections as well as federal elections. The INE's creation sought to homogenize electoral standards across the country and consolidate a transparent and democratic electoral system[17].

Following the establishment of the INE, which prompted reforms to the constitutions of state electoral institutes, most states began to align the scheduling of their state and municipal elections with the federal elections. This shift is depicted in Figure 1, illustrating the states that conducted municipal elections each year. By 2018, the majority of states had synchronized their municipal election calendars to occur simultaneously with federal elections.

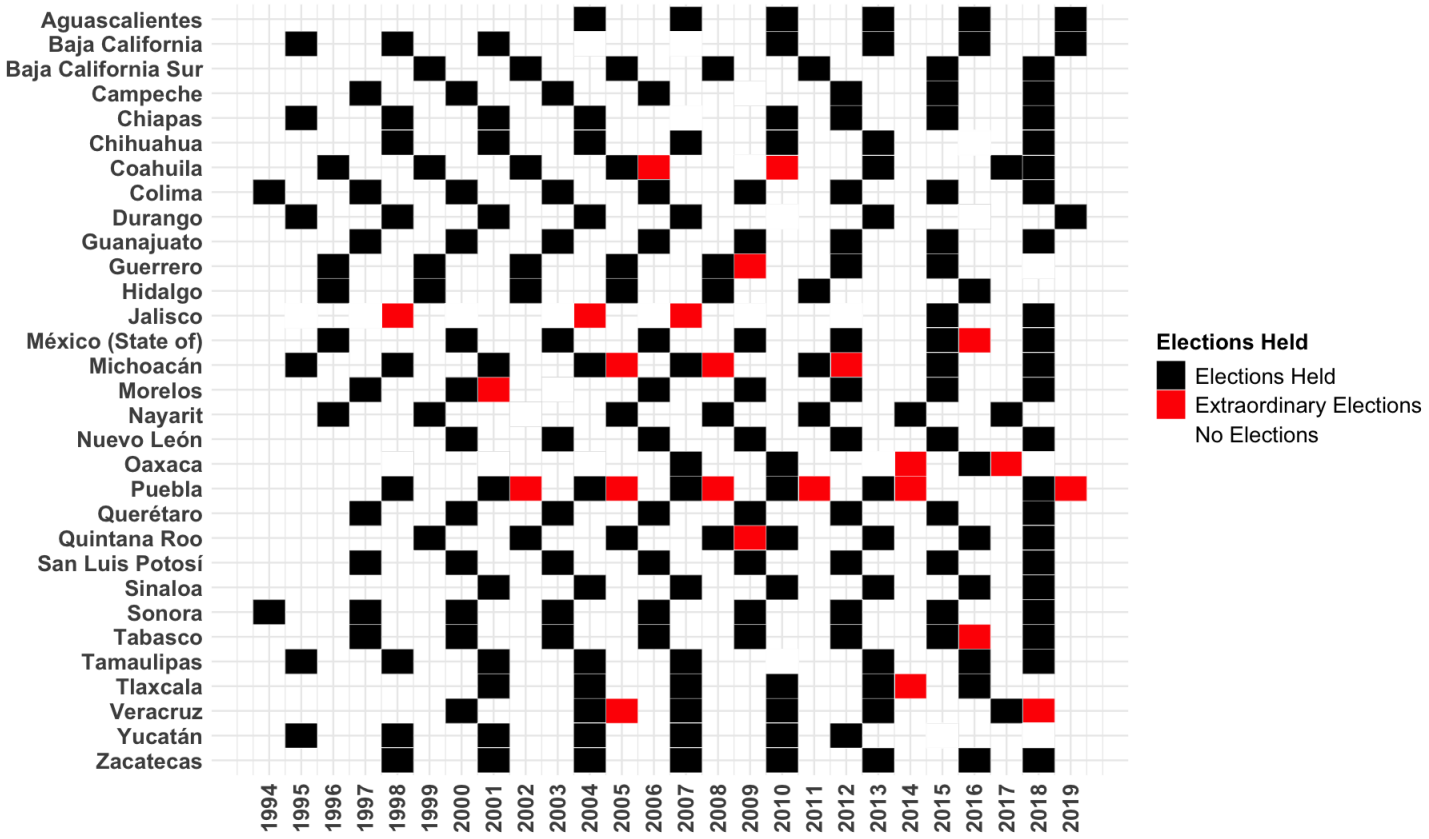

Figure 1. Municipal elections included in the database, by state and year

Electoral Coalitions

The four main political parties—PRI, PRD, PAN, and MORENA—have played pivotal roles in Mexico's political landscape. The PRI, once the dominant political force, faced increasing challenges from the rise of PAN and PRD. Despite their ideological difference, the PAN and PRD often formed electoral coalitions in local and national to help break the PRI's hegemony. After 2012, Mexico's relatively steady state-level coalitions fractured, with coalitions increasingly varying across municipalities within states and changing across elections as smaller parties jockeyed for power around candidates from leading parties. In recent years, MORENA has dominated Mexican elections, while also forming strategic electoral coalitions with smaller parties to enhance its widespread access to power.

Figure 2 shows trends in coalition formation over time, displaying the proportion of municipalities with at least one electoral coalition for each electoral year from 1997 to 2019. Despite some fluctuations, partly reflecting different states holding municipal elections on different cycles, there has been a clear upward trend in coalition formation. The notable decline in 2019 reflects only three holdout states that had not yet synchronized with the federal electoral calendars, as shown in Figure 1.

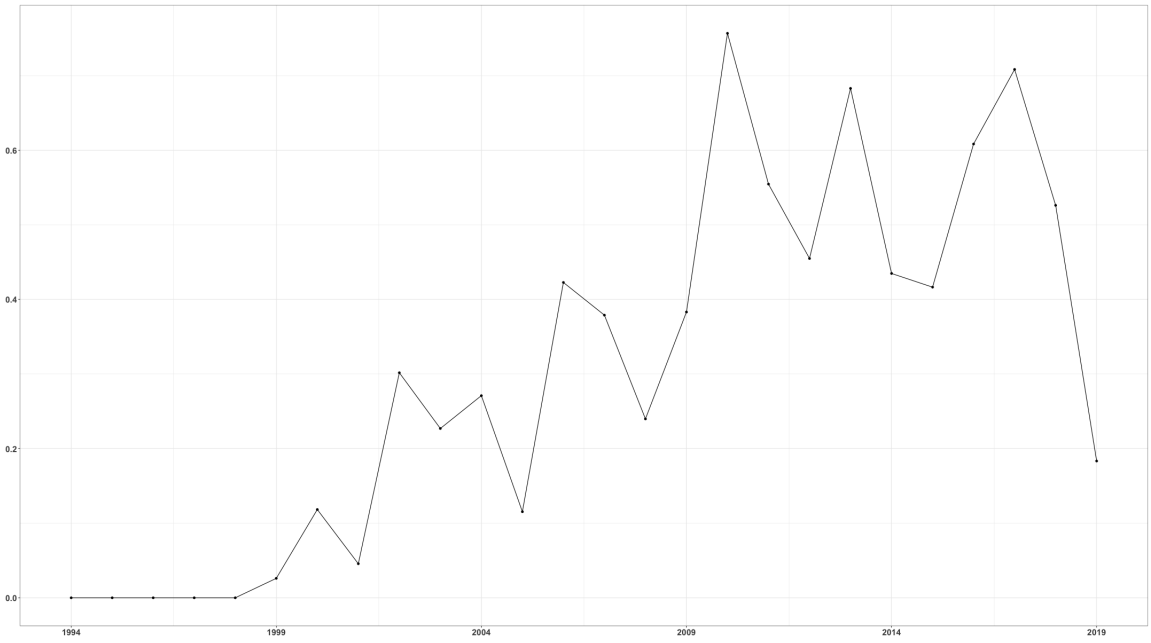

Figure 2. Proportion of municipalities with at least one electoral coalition over time.

Methods

We sourced electoral precinct-level data for 31 states of Mexico (excluding Mexico City), for which precinct-level data is available between 1994 and 2019. No centralized electoral institution in Mexico aggregates electoral data for state and municipal elections. Rather, state electoral institutes conduct state and municipal elections independently. While the INE was tasked to oversee all state and municipal elections starting in 2014, it has not systematically collected and aggregated their electoral data. We aggregated the data collected by Larreguy[5], Marshall and[6], and Enríquez et al.[7] from each state electoral institute's website for recent elections and through freedom of information requests for earlier elections. This data came in various file formats, including comma-separated values (CSV), Excel (xlsx), and portable document format (PDF) requiring optical character recognition (OCR).

We implemented a comprehensive and reproducible R pipeline that systematically handles the electoral precinct-level data for all 31 Mexican states (excluding Mexico City) to ensure consistency across datasets from different elections and states. First, we harmonized data that initially came in varied formats, renaming the variables in each dataset to follow consistent naming conventions across different datasets, adding a unique municipality identifier based on the municipality names according to the INEGI (*Instituto Nacional de Estadística y Geografía*, National Institute of Statistics and Geography) municipality catalog, and aggregating electoral returns to the electoral precinct level wherever data was reported at the polling booth level.

Second, we identified each electoral coalition in a given municipality and computed the total votes for that coalition in each electoral precinct. This process was complicated because Mexican voters can often cast a vote for either a single party within an electoral coalition, for any combination of such parties, or for all the parties in the coalition. This then required summing up the votes cast for all individual political parties within the electoral coalition, for all subsets of such parties, or for the electoral coalition itself. This process was manual since electoral coalitions vary not only across elections, but also across states, and even across municipalities within the same election and state. We also summed up all the votes cast for parties or electoral coalitions to compute the number of valid votes; further adding the number of null ballots and votes for unregistered candidates yields the total number of votes cast.

Third, to construct turnout rates, we combined the total number of votes with the number of registered voters. While total votes are reported by most state electoral institutes (or can be computed by summing across vote tallies), the number of registered voters is only sometimes included. We filled in missing voter registration data via a freedom of information request for the number of registered voters at every available point in time to the INE. The voter register is the same for federal, state, and municipal elections, allowing us to fill in missing observations with the information from the closest date to the corresponding election. This very occasionally results in turnout rates that exceed 100%; we recommend dropping these precincts from analyses.

Fourth, we computed the vote total and vote share for the incumbent party or coalition. To do this, we started by merging

the incumbent mayor's partisan coalition and the name of the winning candidate from Magar's Municipal and State Election Data[18] and the SNIM (*Sistema Nacional de Información Municipal*, National Municipal Information System) dataset[19]. Then, if the set of parties in the incumbent coalition was equal to, or a strict subset of, an electoral coalition running in the following election, we assigned this coalition's vote as the incumbent party's vote. For example, where the incumbent mayor was elected representing a coalition between the PAN and the PRI and the next election has a new coalition between the PAN, the PRD, and the PRI, we computed the incumbent party vote as the vote for the PAN-PRD-PRI coalition. Otherwise, we manually researched the individual party affiliation of the incumbent mayor within the electoral coalition, and classified the incumbent party vote as the vote for the electoral coalition the party of the incumbent ran under in the following election. For example, if the incumbent mayor was elected representing a PAN-PRD coalition, but the PAN and PRD ran separately at the next election, we investigate whether the incumbent mayor is more closely affiliated with the PAN or PRD and designate that party's vote as the incumbent party vote. This involved manually searching local and state newspapers as well as online official government sources, party websites, and social media posts of official party accounts. We were unable to identify reliable sources for only 118 municipality elections, resulting in 0.58% of the incumbent party vote data being unavailable. We analogously included the partisan coalition of the governor at the state level, name of the winning candidate for governor, and corresponding governor election year from Magar's Municipal and State Election Data[18], and computed the incumbent governor's coalition vote share at the electoral precinct level in the municipal elections.

Fifth, we also merged in the municipal-level vote share for the incumbent coalition at the previous election, the partisan composition of the runner-up coalition and name of the runner-up candidate, and the runner-up's municipal-level vote share at the previous election from Magar's Municipal and State Election Data[18]. Moreover, we coded the vote share for the electoral coalitions involving the four main parties: PRI, PAN, PRD, and MORENA.

Data Records

The Mexican municipal elections electoral precinct-level database is available through the Open Science Framework repository and a dedicated GitHub page. This section outlines the structure of the repositories and the contents of the files in them.

Overview of Data Files

The data repository is organized into several folders, R scripts, and original data files, which we explain in detail next.

- `Data`: This folder contains the following databases needed to run the script files to generate the intermediate and final data:
  - `Raw_Electoral_Data`: This folder contains all the raw electoral data. `process_raw_electoral_data.R` runs the files in this folder.
  - `collapsed_database_manual_cases`: This folder contains the `state_collapsed_edited.csv` file containing the manually-sourced information about the incumbent mayors whose electoral coalitions were not equal to, or a subset of, the coalition their party belongs to in the current election. Each case is substantiated by at least one reliable source.
  - `extraordinary_elections`: This folder contains the `correct_extra_elec_final.csv` file which lists the ordinary and extraordinary elections we remove as well as the `diff_year_extra_elec.csv` and `diff_year_extra_elec_flag.csv` files, which are needed to process the handling of extraordinary elections taking place in different calendar years.
  - `incumbent_data`: This folder contains the incumbent data from Magar's Municipal and State Election Data[18] and the SNIM dataset[19]. `incumbent_manipulator.R` runs the files in this folder.
  - `municipal_magar_data_splitcoal`: This folder contains Magar's Municipal and State Election Data[18], which is used for the technical validation by `graphs_correlations.R` and the creation of the municipal level incumbent and runner-up by `FINAL_db_formation.R`.
- `Processed Data`: This folder organizes the processed data for each of Mexico's 31 states produced by the script files.
  - `state sub-folders`: Each state-specific sub-folder contains the output from each data processing script, which serves as the input for the next script in the data pipeline. All intermediate and final datasets are thus stored within the same structure, with the final, fully-processed database for each state saved in its respective sub-folder (see below).

    * state_process_raw_data.csv: file produced by the first script, process_raw_electoral_data.R
    * state_vote_manipulation.csv: file produced by the second script, vote_manipulation.R
    * state_incumbent_manipulator.csv: file produced by the third script, incumbent_manipulator.R
    * state_vote_calculator.csv: file produced by the fourth script, incumbent_vote_calculator.R
    * state_final.csv: file produced by the fifth script, final.R.
  - coalition_dic.csv: This csv file contains all the coalitions contained in our dataset.
  - duplicate_cases.csv: This CSV file contains data necessary to execute the final script that merges data from all 31 states, addressing minor inconsistencies in the final incumbent information. These inconsistencies arise from differences between data sources, which sometimes include candidates who served for interim periods (e.g., a year or two) between official election cycles.

- Final Data: This folder contains the final database that merges all 31 final state databases.
  - all_states_final.zip: The final product of the script pipeline of all states, contains the compressed csv file.

- Correlation Data: This folder contains all the data used in technical validation in section.
  - turnout_magar: This sub-folder contains all the relevant data from Magar's Municipal and State Election Data[18] and SNIM[19] used by magar_mun.R to create the corresponding data contained in the generated data sub-folder to be used by the graphs_correlation.R script.
  - generated_data: This sub-folder contains magar_turnout.csv used by the graphs_correlations.R script.

- Scripts: This folder contains the R scripts that perform the data cleaning, variable coding, and technical validation analysis. Within the State Scripts sub-folder, each state's files need to be run in the order of this list, and the resulting csv file from each of the scripts is used by the next one until the final database for each state is produced. Finally, after running it for all 31 states, FINAL_db_formation.R consolidates all the 31 final datasets into one.
  - process_raw_electoral_data.R: This script imports precinct-level data for municipal elections in the 31 states of Mexico (excluding Mexico City) for which it is available between 1994 and 2019. It standardizes variable names and, when applicable, aggregates polling-booth-level data to the electoral precinct level. This script produces state_process_raw_data.csv.
  - vote_manipulation.R: This script cleans the data originating from the previous script, producing the file named state_vote_manipulation.csv stored in the respective state's sub-folder in Processed Data. This script selects the following relevant variables: municipality and state names, state code (2-digit code uniquely identifying states), municipality code (5-digit code uniquely identifying municipalities), electoral precinct code (4-digit code uniquely electoral precincts within a state), election year, valid votes cast, total votes cast, registered voters, and all votes cast for specific parties and electoral coalitions. For some states, this script includes the use of correct_extra_elec_final.csv, diff_year_extra_elec.csv and diff_year_extra_elec_flag.csv discussed earlier.
  - incumbent_manipulation.R: This script processes and merges information at the municipal level about the name and party or electoral coalition of the incumbent mayors and runner-ups from several sources, including Magar's Municipal and State Election Data[18] and the SNIM dataset[19]. Each dataset provides information about municipal incumbents across various election years and states. The script keeps the relevant variables, standardizes the variable names across data sets, and then appends all the incumbent information into a single dataset. The resulting dataset is then merged into the electoral precinct-level dataset resulting from vote_manipulation.R. This script produces a file named state_incumbent_manipulator.csv stored in the respective state's sub-folder in Processed Data.

- `incumbent_vote_calculator.R`: This script computes the electoral precinct-level votes for the parties of the incumbent mayor and runner-up by linking the identity of such parties or their electoral coalitions to the precinct-level vote data; the resulting file is named `state_vote_calculator.csv` and stored in the respective state's sub-folder in Processed Data. The script first standardizes party names for variables capturing both individual party and coalition votes. It then merges a dataset that includes manually sourced incumbent party information for the subset of elections where the new incumbent coalitions were not equal to, or a superset of, their coalitions in the previous elections. It then creates a variable capturing the votes cast for the incumbent party's electoral coalition and an analogous variable for the runner-up party, whenever the information is available; no analogous manual process was done for the challenger coalitions.

- `final.R`: This script generates the final version of the electoral dataset for each state, named `state_final.csv` and stored in the respective state's sub-folder in Processed Data. It computes the votes for the main political parties—PRI, PAN, MORENA, and PRD—or their electoral coalitions. The script also computes the votes for the state incumbent party or its municipal electoral coalitions. The script further integrates manually-sourced incumbent information through the `state_name_collapsed_edited` files.

- `validation.R`: This script is necessary to produce any data file required for the final data. It runs a series of validation exercises per state to check each state's final script.

- `correlation`: This sub-folder contains all relevant scripts for producing the data used in the technical validation by the `graphs_correlations.R` script.

  - `magar_turnout.R`: This script generates the `magar_turnout.csv` within the `generated data` folder in the `Correlation Data` folder using Magar's Municipal and State Election Data[18] in the `magar_turnout` sub-folder.

  - `graphs_correlations.R`: This script produces the technical validation analyses, using `magar_turnout.csv`.

- `FINAL_db_formation.R`: This script compiles and merges the final datasets for the 31 Mexican states (excluding Mexico City) into a single dataset named `all_states_final.zip` and stored in Final Data. The script also computes the precinct-level vote shares for the incumbent mayor and runner-up's coalitions, as well as for the major political parties (PRI, PAN, MORENA, PRD), both as a share of valid votes and the number of registered voters.

- `README.md`: This file contains detailed documentation for replicating the data processing workflow using the provided scripts. It includes a comprehensive description of the data structure, a step-by-step guide to running the scripts, and an explanation of the purpose and functionality of each script involved in the process.

Variable description

Table 1 presents the identifying variables for municipalities and electoral precincts. Tables 2 and 3 respectively describe the electoral variables corresponding to the municipal and state levels as well as electoral-precinct level.

Table 1. Identifying variables for municipalities and electoral precincts

Variable	Description
state	Name of the state where the election took place.
mun	Name of the municipality where the election took place.
state_code	A numerical code assigned by INEGI that uniquely identifies each state. For the first 9 states (including Mexico City), this is a single digit (1–9), while for the remaining 23 states, it is a two-digit code (10–32).
mun_code	A numerical code assigned by INEGI that uniquely identifies each municipality. For 4-digit codes, the first digit corresponds to states 1 through 9; for 5-digit codes, the first two digits correspond to states 10 through 32, with the remaining digits representing the specific municipality.
precinct	A 1- to 5-digit code used to uniquely identify electoral precincts within each state.
year	The year in which the election was held.

Table 2. Municipal- and state-level variables

Variable	Description
incumbent_party_candidate	The name of the candidate from the incumbent party or coalition who won the previous municipal election.
incumbent_party	The partisan composition of the coalition of the incumbent mayor who won the previous municipal election.
incumbent_party_component	The individual party or parties within the incumbent electoral coalition that contribute to the overall incumbent vote.
runnerup_party	The partisan composition of the coalition of the runner-up that came second in the previous municipal election.
state_incumbent_party	The political party or electoral coalition that held power at the state level during the municipal election.
mun_incumbent_party_vote	The number of votes received by the incumbent party or coalition in the previous municipal election at the municipal level.
mun_runnerup_party_vote	The number of votes received by the runner-up party or coalition in the previous municipal election at the municipal level.
mun_winning_margin	The winning margin of the incumbent party or electoral coalition, or the difference in its vote share and that of the runner-up party or electoral coalition, in previous municipal elections.

Table 3. Electoral-precinct-level variables

Variable	Description
researched_incumbent_party	Manually researched incumbent party.
source_researched_incumbent_party	The source used to define the incumbent party.
incumbent_party_vote	The number of votes received in a given electoral precinct by the incumbent party or coalition in the municipal election.
share_incumbent_party_valid_vote	The percentage of votes cast in a given electoral precinct for the incumbent party or coalition as a proportion of valid votes.
share_incumbent_party_registered_voters	The percentage of votes cast in a given electoral precinct for the incumbent party or coalition as a proportion of registered voters.
runnerup_party_vote	The number of votes received in a given electoral precinct by the runner-up party or electoral coalition in the municipal election. Unlike the incumbent vote data, the runner-up data is still a work in progress since we are missing the manual work to identify the runner-up party whenever the current runner-up's electoral coalition is a superset of the coalition in the current election.
share_runnerup_party_valid_vote	The percentage of votes cast in a given electoral precinct for the runner-up party or coalition as a proportion of the valid votes.
share_runnerup_party_registered_voters	The percentage of votes cast in a given electoral precinct for the runner-up party or coalition as a proportion of registered voters.
state_incumbent_party_vote	The number of votes received in a given electoral precinct in the municipal election by the state incumbent party or the electoral coalition it ran under.
share_state_incumbent_party_valid_vote	The percentage of votes cast in a given electoral precinct in the municipal election for the state incumbent party or the electoral coalition it ran under as a proportion of valid votes.
share_state_incumbent_party_registered_voters	The percentage of votes cast in a given electoral precinct in the municipal election for the state incumbent party or the electoral coalition it ran under as a proportion of registered voters.
PRI_vote	The number of votes received in a given electoral precinct by the PRI or its electoral coalition.
share_PRI_valid_vote	The percentage of votes cast in a given electoral precinct for the PRI or its electoral coalition as a proportion of valid votes.
share_PRI_registered_voters	The percentage of votes cast in a given electoral precinct for the PRI or its electoral coalition as a proportion of registered voters.
PAN_vote	The number of votes received in a given electoral precinct by the PAN or its electoral coalition.
share_PAN_valid_vote	The percentage of votes cast in a given electoral precinct for the PAN or its electoral coalition as a proportion of valid votes.
share_PAN_registered_voters	The percentage of votes cast in a given electoral precinct for the PAN or its electoral coalition as a proportion of registered voters.
PRD_vote	The number of votes received in a given electoral precinct by the PRD or its electoral coalition.
share_PRD_valid_vote	The percentage of votes cast in a given electoral precinct for the PRD or its electoral coalition as a proportion of valid votes.
share_PRD_registered_voters	The percentage of votes cast in an electoral precinct for the PRD or its electoral coalition as a proportion of registered voters.
MORENA_vote	The number of votes received in a given electoral precinct by MORENA or its electoral coalition.
share_MORENA_valid_vote	The percentage of votes cast for MORENA or its electoral coalition as a proportion of valid votes.
share_MORENA_registered_voters	The percentage of votes cast in an electoral precinct for MORENA or its electoral coalition as a proportion of registered voters.
valid	The number of valid votes cast in a given electoral precinct, excluding invalid or spoiled ballots.
total	The sum of all votes cast in a given electoral precinct, including valid, invalid, and spoiled ballots.
registered_voters	The number of registered voters in a given electoral precinct.
turnout	The percentage of registered voters in a given electoral precinct that cast a vote in the election.

Technical validation

The validation process included various exercises, including handling electoral coalitions when computing incumbent party votes, dealing with cases of extraordinary municipal elections, and assessing the correlation of aggregate votes and turnout with alternative official data sources.

Incumbent party vote share calculation

Incumbent party vote computation required information about the incumbent party or the electoral coalition under which it ran in the past election, before then identifying under which new electoral coalition the incumbent ran in the corresponding election. We first use information on incumbent parties and their electoral coalitions from Magar's Municipal and State Election Data[18] and the SNIM[19]. If the incumbent mayor did not previously run as part of any electoral coalition, the incumbent party vote was computed as the vote of the mayor's party or the electoral coalition under which it ran in the contemporaneous election. There were 12,107 such municipal elections, corresponding to 71.5% of the electoral-precinct level observations in the data. If the incumbent mayor previously ran as part of any electoral coalition, and neither Magar's Municipal and State Election Data[18] nor the data from the SNIM[19] included information about the party of the incumbent mayor, we proceed as follows. If the new electoral coalition of the incumbent mayor's party was a superset of its coalition in the previous election, we considered the vote for its new coalition as the incumbent party vote. There were 827 such municipal elections, corresponding to 8.6% of the electoral-precinct level observations in the data. Otherwise, we conducted extensive manual research primarily through official local and state newspapers, and online government and party websites, as well as social media posts, to determine the incumbent mayor's party. The incumbent party vote was then computed as the vote of the incumbent mayor's party if it no longer ran under an electoral coalition or that of its new electoral coalition. There were 2,057 such municipal elections, corresponding to 15.7% of the electoral-precinct level observations in the data. We were unable to identify incumbent information for 880 municipal elections, resulting in 3.8% electoral-precinct level observations with no incumbent party vote data.

Extraordinary Elections

Tables 4 and 5 report the instances of extraordinary elections, which election data we chose to keep in the database, and the reason behind such a choice. We kept the election we thought best represented the electoral preferences of the electorate at the electoral precinct level.

Table 4. Extraordinary elections in the same year of ordinary elections

Year	State	Municipality	Chosen	Reason
2015	Chiapas	Tapilula	Normal	Original election resulted in a tie
2013	Chihuahua	Coyame del Sotol	Normal	Original election resulted in a tie
2016	Hidalgo	Omitlán de Juárez	Extraordinary	Religious figure campaigning interference
2009	Jalisco	San Cristóbal de la Barranca	Normal	Winner was disqualified due to a technical error when registering the party
2009	Jalisco	Gómez Farías	Normal	Winner was disqualified due to a technical error when registering the party
2003	México	Atenco	Extraordinary	Uprising in the rural sector prevented voting during the normal election
2015	Michoacán	Sahuayo	Normal	No data for extraodinary election
2018	Nuevo León	Monterrey	Extraordinary	Inconsistencies and possible tampering of votes and the handling of electoral data
2018	Oaxaca	San Juan Ihualtepec	Extraordinary	No data for extraodinary election
2013	Tlaxcala	Apetatitlán de Antonio Carvajal	Extraordinary	Religious figure campaigning interference
2016	Zacatecas	Zacatecas	Normal	No data for extraodinary election

Variable correlation

To further validate our data, we correlate the variables in the data set with related variables from other official data sources. First, we correlate electoral-precinct-level turnout in our data set with the exact same outcomes but provided by the INE[20] for municipal elections in 2015, 2016, 2017, and 2018. The correlation is 0.979 (49,675 observations). Second, we aggregated our turnout variable at the municipal level by summing total votes and registered voters, then calculating the turnout as the ratio of these sums. We then correlate these municipal-level figures with those from Magar's Municipal and State Election Data[18]. Third, we similarly aggregated the vote shares of the PRI, PAN, PRD, and MORENA or the coalitions under which they run to the municipal level and correlated them against analogous figures from Magar's Municipal and State Election Data[18]. Overall, the correlation coefficients reported in Table 6 are very high, ranging between 0.814 and 0.94, which supports the validity of our data. These correlations are not perfect, however, since sometimes electoral-court rulings invalidate the electoral outcomes of certain electoral precincts.

Table 5. Extraordinary elections in the subsequent year of ordinary elections

Year	State	Municipality	Chosen	Reason
2006	Coahuila	Abasolo	Normal	Original victorious candidate died
2010	Coahuila	Juárez	Extraordinary	Exceeded campaign funds in original election
2010	Coahuila	Lamadrid	Extraordinary	Government official participated as representative of victorious party
2009	Guerrero	Malinaltepec	Extraordinary	Original ballots did not include the logo of the coalition formed by two parties
1998	Jalisco	Juchitlán	Normal	Original election resulted in a tie
2004	Jalisco	Tamazula de Gordiano	Extraordinary	Irregularities in the the original electoral process
2007	Jalisco	Tuxcueca	Extraordinary	Irregularities in the the original electoral process
2016	México	Chiautla	Extraordinary	Involvement of the church in the original electoral process
2005	Michoacán	Tumbiscatio	Extraordinary	Irregularities in the original electoral process
2008	Michoacán	Yurécuaro	Extraordinary	Irregularities in the original electoral process
2012	Michoacán	Morelia	Extraordinary	Irregularities in the original electoral process
2001	Morelos	Ocuituco	Extraordinary	Irregularities in the original electoral process
2014	Oaxaca	Tlacamama	Extraordinary	Burning of voting booths
2017	Oaxaca	Santa María Xadani	Extraordinary	Irregularities in the original electoral process
2002	Puebla	Molcaxac	Extraordinary	Irregularities in the original electoral process
2005	Puebla	Santa Inés Ahuatempan	Normal	Original election resulted in a tie
2008	Puebla	General Felipe Ángeles	Extraordinary	Irregularities in the original electoral process
2011	Puebla	Ixcamilpa de Guerrero	Extraordinary	Opposition coalition challenged the official results
2011	Puebla	San Jerónimo Tecuanipan	Normal	Original election resulted in a tie
2011	Puebla	Tlaola	Extraordinary	Irregularities in the original electoral process
2014	Puebla	Acajete	Extraordinary	Irregularities in the original electoral process
2014	Puebla	Cuapiaxtla de Madero	Extraordinary	Irregularities in the original electoral process
2019	Puebla	Ahuazotepec	Extraordinary	Irregularities in the original electoral process
2019	Puebla	Cañada Morelos	Extraordinary	Irregularities in the original electoral process
2019	Puebla	Mazapiltepec de Juárez	Extraordinary	Irregularities in the original electoral process
2019	Puebla	Ocoyucan	Extraordinary	Irregularities in the original electoral process
2019	Puebla	Tepeojuma	Extraordinary	Irregularities in the original electoral process
2009	Quintana Roo	Tulum	Extraordinary	Due to it being a newly created municipality, extraordinary elections were held as the first official ones
2016	Tabasco	Centro	Extraordinary	Irregularities in the original electoral process
2014	Tlaxcala	Acuamanala de Miguel Hidalgo	Extraordinary	Irregularities in the original electoral process
2005	Veracruz	Landero y Coss	Extraordinary	Opposition party challenged the official results
2018	Veracruz	Camarón de Tejeda	Extraordinary	Exceeded campaign funds in original election
2018	Veracruz	Emiliano Zapata	Extraordinary	Irregularities in the original electoral process
2018	Veracruz	Sayula de Alemán	Extraordinary	Exceeded campaign funds in original election

Table 6. Municipal-Level Turnout and Vote Shares for PRI, PAN, PRD, and MORENA

	Turnout	PRI vote share	PAN vote share	PRD vote share	MORENA vote share
	(1)	(2)	(3)	(4)	(5)
Turnout[18]	0.937***				
	(0.009)				
PRI vote share[18]		0.814***			
		(0.004)			
PAN vote shares[18]			0.894***		
			(0.003)		
PRD vote share[18]				0.940***	
				(0.002)	
MORENA vote share[18]					0.915***
					(0.002)
Observations	14,450	14,450	14,450	14,450	14,450
R^2	0.436	0.705	0.874	0.916	0.903

Note: Regressors are from Magar's Municipal and State Election Data[18]. All the data is aggregated at the municipal level. Vote shares are over the number of registered voters.
*p<0.1; **p<0.05; ***p<0.01

Usage Notes

This database provides precinct-level election returns for Mexican municipal elections from 1994 to 2019. Researchers using this dataset should consider several points to maximize its utility for electoral studies and ensure appropriate use of the data. First, this dataset can support a wide range of empirical research, particularly studies focusing on local electoral dynamics, including incumbency effects and electoral competition. The dataset includes information on votes for the four major political parties (PRI, PAN, PRD, and MORENA) as well as the incumbent and runner-up coalitions. The inclusion of lagged municipal-level votes and state-level incumbent identities also facilitates regression discontinuity designs and heterogeneity analyses.

Researchers should cite this paper[21] and data the sources referenced[5–7]. Moreover, users of the incumbent vote data should be mindful of the changing electoral coalitions when interpreting results on incumbent vote. Additionally, users should consider that the data includes (rare) extraordinary elections in certain municipalities where the initial elections were annulled and repeated due to irregularities.

Code availability

The code used for processing the raw electoral data and generating the final dataset has been written in R (version 4.1.0). All scripts used for data cleaning and analysis are publicly available through the Open Science Framework repository and a dedicated GitHub page. The repository also contains a codebook and documentation to aid understanding of the dataset structure, processing, and analysis.

References

1. Giraudy, A., Moncada, E. & Snyder, R. Subnational research in comparative politics: Substantive, theoretical, and methodological contributions. In *Inside Countries: Subnational Research in Comparative Politics*, 3–54 (Cambridge University Press, 2019).
2. King, G., Keohane, R. O. & Verba, S. Designing Social Inquiry: Scientific Inference in Qualitative Research. *Princet. Univ. Press.* (1994).
3. Ferraz, C. & Finan, F. Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes. *Q. J. Econ.* **123**, 703–745 (2008).
4. Larreguy, H., Marshall, J. & Snyder, J. M. Publicising Malfeasance: When the Local Media Structure Facilitates Electoral Accountability in Mexico. *The Econ. J.* **130**, 2291–2327 (2020).
5. Larreguy, H. Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico. *EPSA 2013 Annu. Gen. Conf. Pap.* (2012).
6. John Marshall. Political information cycles: When do voters sanction incumbent parties for high homicide rates? (2023). Working paper.
7. José Ramón Enríquez, Horacio Larreguy, John Parshall & Alberto Pimpser. Mass Political Information on Social Media: Facebook Ads, Electorate Saturation, and Electoral Accountability in Mexico. *J. Eur. Econ. Assoc.* **22**, 1678–1722, 10.1093/jeea/jvae011 (2024).
8. Britannica, T. Editors of Encyclopaedia. Plutarco Elías Calles. Encyclopedia Britannica (2024). Retrieved from https://www.britannica.com/biography/Plutarco-Elias-Calles.
9. National Revolutionary Party (PNR) (2024). Encyclopedia of Latin American History and Culture. Retrieved September 12, 2024, from Encyclopedia.com: https://www.encyclopedia.com/humanities/encyclopedias-almanacs-transcripts-and-maps/national-revolutionary-party-pnr.
10. The Legacy of the PRI's Dominance in Mexican Politics. Mexico Histórico (n.d.). Retrieved from https://www.mexicohistorico.com/paginas/The-Legacy-of-the-PRI---s-Dominance-in-Mexican-Politics.html.
11. Britannica, T. Editors of Encyclopaedia. Cuauhtémoc Cárdenas. Encyclopedia Britannica (2024). Retrieved from https://www.britannica.com/biography/Cuauhtemoc-Cardenas-Solorzano.
12. Dan La Botz. Mexico's Party of the Democratic Revolution at 25: Disappointment Disillusion. New Politics (2014). Retrieved from https://newpol.org/mexicos-party-democratic-revolution-25-disappointment-disillusion/.
13. Britannica, T. Editors of Encyclopaedia. National Action Party (PAN). Encyclopedia Britannica (2024). Retrieved from https://www.britannica.com/topic/National-Action-Party-political-party-Mexico.

14. National Action Party (PAN) (2024). Encyclopedia of Latin American History and Culture. Retrieved September 11, 2024, from Encyclopedia.com: https://www.encyclopedia.com/humanities/encyclopedias-almanacs-transcripts-and-maps/national-action-party-pan.

15. Jessica Guerrero. MORENA: A Rebirth of Mexican Politics. Pulse News Mexico (2022). Retrieved from https://pulsenewsmexico.com/2022/01/25/morena-the-rebirth-of-old-mexican-politics/.

16. Historia del Instituto Federal Electoral. Instituto Nacional Electoral. Retrieved from https://portalanterior.ine.mx/archivos3/portal/historico/contenido/menuitem.cdd858023b32d5b7787e6910d08600a0/.

17. Lauren TerMaat. Mexico's National Electoral Institute - Explainer. Wilson Center (2023). Retrieved from https://www.wilsoncenter.org/article/mexicos-national-electoral-institute-explainer.

18. Magar, E. Recent Mexican Election Vote Returns. https://github.com/emagar/elecRetrns (2018). Last revised on March 10, 2023.

19. Sistema Nacional de Información Municipal. Sistema Nacional de Información Municipal (SNIM). http://www.snim.rami.gob.mx/ (2024). Data retrieved from multiple official sources including INEGI, CONAPO, and PNUD.

20. Instituto Nacional Electoral (INE). Sistema de Consulta de la Estadística de las Elecciones (SICEE). https://siceen21.ine.mx/downloadzip (2024). Retrieved on September 6, 2024.

21. Calderón-Hernández, B., Larreguy, H., Marshall, J. & Pérez Castellanos, J. L. Electoral precinct-level database for Mexican municipal elections. https://github.com/brunocalderonh/The-Mexican-municipal-elections-electoral-precinct-level-database (2024). Version 1.0.0, Released on September 20, 2024.

Acknowledgments

This paper is dedicated to the late Francisco Cantú, who is greatly missed and encouraged us to make this data publicly available to help other scholars, as he helped us all. Salvador Ascencio Pastora provided valuable research assistance. Larreguy gratefully acknowledges funding from the French Agence Nationale de la Recherche under the Investissement d'Avenir program ANR-17-EURE-0010.

Author contributions statement

All authors contributed to the data code and manuscript. BC and JL performed the technical validation analysis.

Competing interests

The authors declare no competing interests.