



**PRINCETON
GERRYMANDING
PROJECT**

Fixing bugs in democracy

Testimony of the Princeton Gerrymandering Project

Before a Joint Hearing of the New York State Senate Standing Committee on Elections and the Standing Committee on Local Government and the New York State Assembly Standing Committee on Election Law and the Standing Committee on Local Governments.

Regarding “Elections in the Pandemic: A Review of the 2020 Primaries”

**Statement on the Electoral Transparency Provisions of the
John R. Lewis Voting Rights Act of New York S. 7528A(Myrie)/A.10841(Walker)
Tuesday, August 11, 2020**

By Hope Johnson, Data Scientist

In order to evaluate the fairness of elections, high-quality information is essential. Clear information becomes even more important during the COVID-19 global pandemic that has disrupted the regular administration of elections and required New York to adapt to the challenge at hand. As a New Yorker, I experienced first-hand the confusion surrounding how and when to vote in the most recent primary election. Data transparency and availability are always critical to evaluating the performance of elections administration to ensure that the right to vote is secured for all New Yorkers. Data is especially critical now as states evaluate the significant changes made to adapt elections to the pandemic. Data will only become even more critical as we approach the publication of data from the 2020 Census and the new redistricting cycle.

The Princeton Gerrymandering Project (PGP) combines math, law, and data science to support state-level reform efforts to eliminate partisan gerrymandering and increase transparency in the redistricting process. We analyze maps, evaluate and suggest redistricting processes and standards, and advocate against problematic reforms. The PGP staff is made up of full- and part-time researchers with expertise in election law, computational analysis, and statistics. PGP is focused on electoral policies that increase participation and effectiveness of voters in American democracy, including efforts that expand data transparency. We offer this written testimony in favor of the John R. Lewis Voting Rights Act of New York, or the JRLVRA (S7528).

As part of our mission, we aim to make redistricting data open to all people via OpenPrecincts. OpenPrecincts removes a key information bottleneck: citizens lack the data to analyze districting plans. Precinct-level voting data is the only way to determine whether a district is drawn to unduly favor a particular group. Existing precinct data resources contain inaccuracies, lack documentation, and come in idiosyncratic formats. OpenPrecincts, on the other hand, follows a universal data schema usable by any redistricting software, including free redistricting tools.

As part of our work building out OpenPrecincts, we have worked on two parallel projects in New York: (1) collecting precinct-level election results and matching those to a map of precinct boundaries for a New York county, and (2) collecting and matching precinct-level results from the 2018 general election and precinct boundaries for the entire state. Both projects have been hampered by the types of data-related issues that the JRLVRA would remedy. We hope that this written testimony sheds some light on what it is like to work with election data from New York and how the JRLVRA, if enacted, would make New York elections data more accessible and transparent.

New York Elections Data Accessibility and Transparency

Section 17-208 of the JRLVRA would require the maintenance of voting and election data. It would mandate the creation and preservation of a repository of the data necessary to determine how voting rights function across New York. We have found that the current data does not implement best practices in a number of ways, namely, data formatting issues, data inconsistencies, and missing information.

As a part of our work on the elections in a county near New York City, we filed a Freedom for Information Act (FOIA) request for precinct-level election results. The response was in PDF format, which is not useful for analysis. We tried converting the PDF files to a format that can be read by a computer using a number of different tools but, because of the format in which the records are kept, it was impossible to extract the necessary data. Instead, we were forced to use data compiled by an open source data project (OpenElections). The existence of a “state database” as described in Section 17-208 of the JRLVRA would create a centralized repository for election data and help standardize the format in which election data is maintained, allowing citizens greater access to and understanding of the electoral process.

New York elections data is also often inconsistent. In order to create a one-to-one matching between precinct-level election results and a map of the voting precincts, both the election results and a map of the boundaries must have a consistent, unique identifier. In the current data, the precinct identifier is often inconsistent between the map and the election data. If “contemporaneous shapefiles for election districts” (§ 17-208(4)(d)) existed, the process of matching a map of precincts with precinct-level election results would be much faster and easier for the average citizen.

Beyond the absence of a unique identifier for election precincts that is consistent across boundary labels and election results, we found countless spelling errors of candidate names in official election results. For example, in the 2018 election for the attorney general of New York, there were more than ten different spellings of “Letitia James.” Inconsistencies like this make the task of preparing data for analysis arduous and time-consuming.

There is also a severe lack of data documentation and transparency. This creates a situation where only people whose jobs are dedicated to collecting and verifying this type of data are able to get access to election information. After months of contacting county officials, we were sent the “precinct consolidation report,” a document that outlines, for various election years, when votes from certain precincts were combined into the vote totals of others. This document uncovered that in some areas, the votes from a single precinct were combined into three other precincts. Unless you were able to gain access to the “compiled list,” you would think from the official election results that nobody registered to vote in those precincts cast their ballots. As another example, it took months of calling and emailing the county clerk in one New York county to discover that if the 4th digit of a precinct name is “7,” there are no registered voters in that precinct. Insider details like this prevent the general public from understanding, let alone utilizing, voting and election data in New York.

In addition to the issues of formatting and inconsistencies, a state elections database would remove the immense burden of accessing data and documentations from each individual county board of election and school district. Currently, elections data is scattered among 62 New York counties and almost 700 school districts, making a comprehensive evaluation of New York elections nearly impossible for professionals, let alone members of the public.

In short, the data necessary to protect New Yorkers’ right to vote are often stored in irregular formats, contain inconsistencies, and lack proper documentation. This is all to say, the obstacles to accessing the data necessary to evaluate the fairness of New York elections are enormous if not insurmountable without specialists devoting significant time and resources. In addition to its other considerable improvements to protect communities of color, the John R. Lewis Voting Rights Act of New York (specifically Section 17-208) would be a significant improvement on the transparency and fairness of New York elections.