Election Night Reporting in the 2022 Election

Samuel Baltz, MIT
John Lapinski, UPenn
Marc Meredith, UPenn
Stephanie Perry, UPenn
Stephen Pettigrew, UPenn
Jason Rhode, MIT
Charles Stewart, MIT
Jack Williams, MIT

January 5, 2023

The system of reporting vote results on election night has become increasingly fractured in recent years. The emergence of competing national vote count operations after the 2016 election has created an environment with a higher potential for errors in vote reporting. On the other hand, multiple vote counts creates the opportunity to improve on real-time data quality assurance in a way that was not possible when one vendor monopolized the vote count aggregation market. The goal of this project is to leverage this competition to create tools that can be used before, during, and after election night in order to minimize the number of vote reporting errors that happen and to provide explanations for those that still occur.

Background and goals

Through the 2016 election cycle, nearly all major media reported election result data that were collected and distributed by the Associated Press. Because of the single-sourced nature of vote counting, discrepancies in vote totals between TV or online news sources tended to be rare and inconsequential. That changed after 2016, when Edison Research began to provide the vote count data to the networks in the National Election Pool (NBC, CBS, CNN, and ABC).[1] Fox News and many local and print news organizations maintained a relationship with the AP. Around the same time, Decision Desk HQ began providing third option for real-time vote results. Official, online data feeds from states and counties have also become increasingly prevalent in recent years.[2]

We believe our project is the first time anybody has collected and compared the data directly from the feeds of these vote count sources in real time. We developed a prototype data dashboard that allows us to monitor the data and identify discrepancies between the data sources throughout election night. After Election Day, our timestamped archive of the data provided an opportunity to investigate the cause and nature of any reporting anomalies from any of the sources.

Results

We focused our data collection on Georgia and North Carolina–two states with highly competitive races, where the state provides a high-quality, public dashboard for reporting official vote results on Election Day. We set up a dedicated server to ingest and archive all county- and state-level vote reported provided by each of the three vote vendors and from the states' websites.

On Election Night, we used a custom-built data dashboard to monitor the vote counts. For any race in these two states, we used the the dashboard to compare, between data sources, the current candidate vote counts or percentages and vote totals at the state and county level. This dashboard had tools that helped automatically detect situations where one data source was noticeably out-of-step with the other sources. We then were able to use our own expertise to discern whether an anomaly was the result of an error or mistake in the vote feed, or whether that data source was updating at a slower pace than the others.

The figure shows every update we received from the AP, Edison, and state-sourced vote feeds in the Georgia and North Carolina Senate races.[3] As the graph shows, all the data sources were largely in-sync with the others. This was particularly the case in North Carolina, where the AP and Edison receive updates from a single spreadsheet on the Board of Elections website.

In Georgia, the sources were slightly less in lock-step with one another, although the differences are slight. The reason for these differences is that in addition to the state's data dashboard, each county has its own dashboard. It is not uncommon for votes often appear on

---

[1] More recently, Edison has provided the vote count data to Reuters as well as other news outlets.

[2] For more details on how votes are counted and reported see: Pettigrew and Stewart. "Protecting the Perilous Path of Election Returns: From the Precinct to the News." Ohio State Technology Law Journal. Summer 2020.

[3] Because of API issues with Decision Desk HQ, we unfortunately were not able to capture their data in real-time on election night.
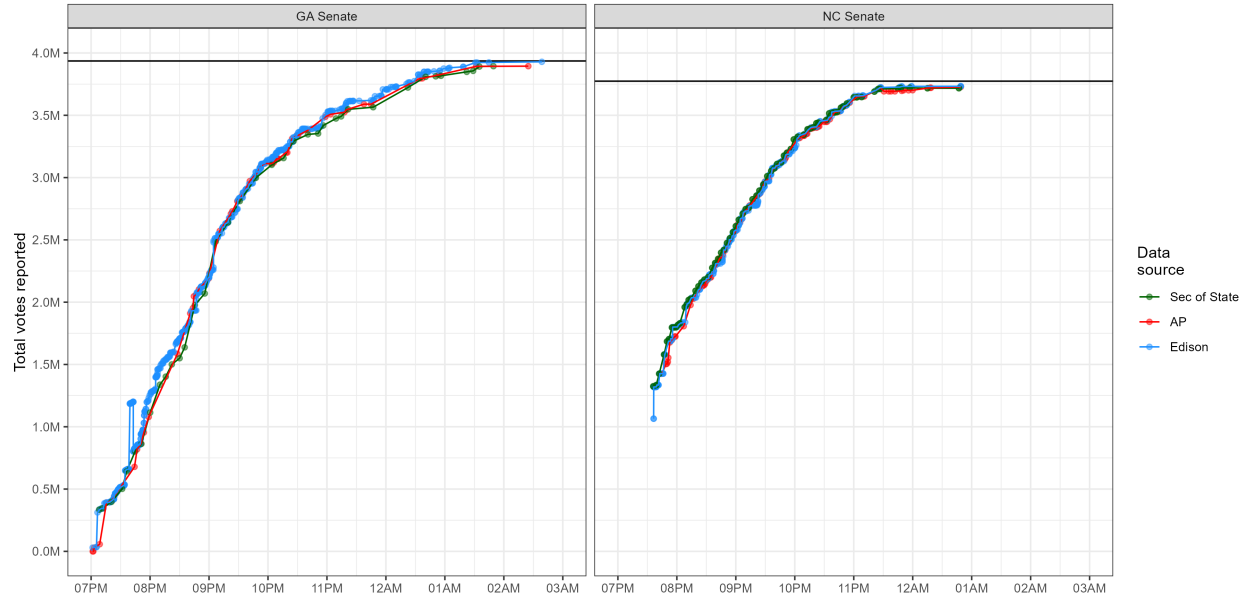
a county's dashboard before the state one. Georgia's data also provides an example of a data anomaly. At 7:39:30pm, Edison's vote feed reported nearly 400,000 votes in Camden County. The county's vote total was reverted back to zero with three minutes. Our monitoring system flagged this error to us, and had it not been fixed so quickly, we could have reported the error to Edison.

Three next steps

First, we plan to put out a public report that describes our project. In it, we will specifically detail, investigate, and explain the nature of the handful of vote reporting anomalies in 2022. We hope that this report can serve as a template for how to combat misinformation and conspiracy theory formulation in the aftermath of a competitive election.

Second, we will produce a white paper that contains best-practice recommendations to both states and vote count vendors for doing real-time quality assurance on vote count data. The insight we have gained from this project could be used to mitigate mistakes before they happen and catch them quickly when they do.

Third, we hope to expand the project in 2024 into states with a less centralized approach to vote reporting, such as Pennsylvania, Wisconsin, Michigan, or Arizona. We believe our experience in 2022 has provided a proof-of-concept that will make it easier to partner directly with election officials in those states or vote vendors so that we can provide real-time feedback about data anomalies.