Statistics and Modeling from the 2020 Election Night

Clint W. Stevenson
Vice President of Data Science and Statistics

LinkedIn: https://www.linkedin.com/in/clint-w-stevenson
Twitter: @StatisticalR
Web: https://statistical-research.com
Email: cstevenson@edisonresearch.com

This presentation can be found on my personal website:

https://statistical-research.com/wp-content/uploads/2021/04/ASA-CO-WY-2020-04-09.pdf
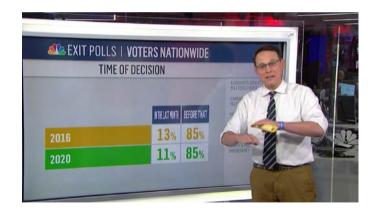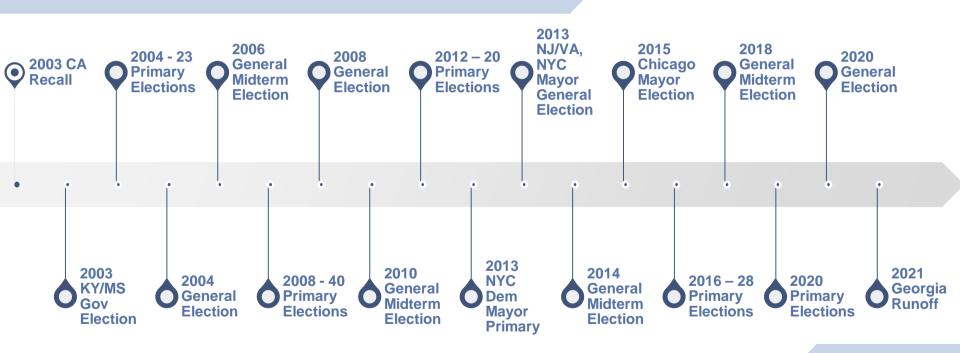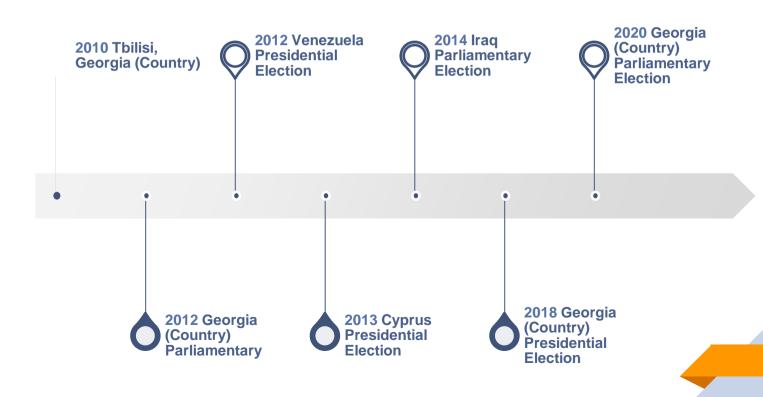
1

Edison Research Elections

Edison Research Elections

2003 CA Recall

2004 - 23 Primary Elections

2006 General Midterm Election

2008 General Election

2012 – 20 Primary Elections

2013 NJ/VA, NYC Mayor General Election

2015 Chicago Mayor Election

2018 General Midterm Election

2020 General Election

2003 KY/MS Gov Election

2004 General Election

2008 - 40 Primary Elections

2010 General Midterm Election

2013 NYC Dem Mayor Primary

2014 General Midterm Election

2016 – 28 Primary Elections

2020 Primary Elections

2021 Georgia Runoff

International Exit Polling

2010 Tbilisi, Georgia (Country)

2012 Venezuela Presidential Election

2014 Iraq Parliamentary Election

2020 Georgia (Country) Parliamentary Election

2012 Georgia (Country) Parliamentary

2013 Cyprus Presidential Election

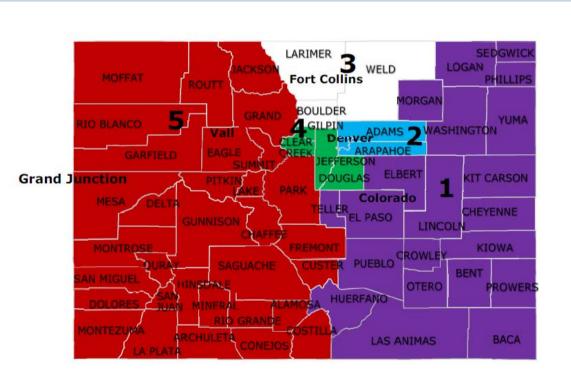2018 Georgia (Country) Presidential Election

2

Exit Poll and Election Sample Design

Geographic Stratification

Colorado

- ■ 1 - East
- ■ 2 - Denver/Eastern Suburbs
- □ 3 - Northern Suburbs
- ■ 4 - Southern Suburbs
- ■ 5 - Colorado Rockies

Sampling Methods

Election Day

- Precincts using a stratified PPS Sample
- Historical vote results for every sample precinct
- Political stratification
- Systematic sample of voters within a precinct.

Early and Absentee

- Registration-Based Sample
- Random Digit Dialing
- In-Person early vote centers

Data Sources: Interviewers

In-Person Election Day

> 700 locations

~ 33% of the vote

In-Person Early

175 voting centers

~24% of the vote

Early Telephone Interviews

Phone interviews absentee/by-mail voters

~43% of the vote.

Data Sources: Actual Results

Sample Precincts

Edison collects actual results from a sample of precincts. These tend to be faster than county results.

County Vote

We collect results from every county office starting at poll closing.

3

Exit Poll Field Operations

Interviewer Operations

~30,000 Interviewers

We can cover almost any location in the country

Custom Developed Operational Software

Edison Research has developed advanced applications for our field operations

- Developed geospatial algorithms to manage our sample precincts and county data collection.

- Custom operational applications to manage our election data research and interaction with county offices

- We manage and track historical vote data and how it's reported by our interviewers and reporters on Election Night.

YOUR ANSWERS ARE
CONFIDENTIAL
Please check only **ONE** response for each question.
Version 1

[A] Are you:
1 ☐ Male
2 ☐ Female

[B] Are you:
1 ☐ White
2 ☐ Black
3 ☐ Hispanic/Latino
4 ☐ Asian
5 ☐ American Indian
6 ☐ Other

[C] In today's election for mayor, did you just vote for:
1 ☐ Bill de Blasio (Dem)
2 ☐ Joe Lhota (Rep)
3 ☐ Adolfo Carrion (Ind)
4 ☐ Jack Hidary (Other)
9 ☐ Other: Who? _____
0 ☐ Did not vote

[D] When did you finally decide for whom to vote in the election for mayor?
1 ☐ Just today
2 ☐ In the last few days
3 ☐ In October
4 ☐ In September
5 ☐ Before that

[E] In which age group are you?
1 ☐ 18-24   5 ☐ 45-49
2 ☐ 25-29   6 ☐ 50-59
3 ☐ 30-39   7 ☐ 60-64
4 ☐ 40-44   8 ☐ 65 or over

[F] Which ONE of these five issues was the most important in deciding how you voted for mayor today?
(CHECK ONLY ONE)
1 ☐ Crime
2 ☐ Jobs and unemployment
3 ☐ Education
4 ☐ Housing
5 ☐ The city's finances

[G] Do you or does someone in your household belong to a labor union?
1 ☐ Yes, I do
2 ☐ Yes, someone else does
3 ☐ Yes, I do and someone else does
4 ☐ No one does

[H] Do you have any children under 18 living in your home?
1 ☐ Yes   2 ☐ No

[I] Are you of Hispanic or Latino descent?
1 ☐ Yes   2 ☐ No

PLEASE TURN THE QUESTIONNAIRE OVER ➡

New York City (G-S-V1-2013)

[J] Is your opinion of Bill de Blasio:
1 ☐ Favorable
2 ☐ Unfavorable

[K] Is your opinion of Joe Lhota:
1 ☐ Favorable
2 ☐ Unfavorable

[L] Do you currently have any children in New York City public schools?
1 ☐ Yes
2 ☐ No

[M] Do you approve or disapprove of the way Michael Bloomberg is handling his job as mayor?
1 ☐ Strongly approve
2 ☐ Somewhat approve
3 ☐ Somewhat disapprove
4 ☐ Strongly disapprove

[N] Do you think the condition of New York City's economy is:
1 ☐ Excellent
2 ☐ Good
3 ☐ Not so good
4 ☐ Poor

[O] Are you:
1 ☐ Protestant   5 ☐ Jewish
2 ☐ Catholic   6 ☐ Muslim
3 ☐ Mormon   7 ☐ Other
4 ☐ Other Christian   8 ☐ None

[P] No matter how you voted today, do you usually think of yourself as a:
1 ☐ Democrat
2 ☐ Republican
3 ☐ Independent
4 ☐ Something else

[Q] On most political matters, do you consider yourself:
1 ☐ Liberal
2 ☐ Moderate
3 ☐ Conservative

[R] What was the last grade of school you completed?
1 ☐ No high school diploma
2 ☐ High school graduate
3 ☐ Some college/assoc. degree
4 ☐ College graduate
5 ☐ Postgraduate study

[S] 2012 total family income:
1 ☐ Under $30,000
2 ☐ $30,000 - $49,999
3 ☐ $50,000 - $74,999
4 ☐ $75,000 - $99,999
5 ☐ $100,000 - $199,999
6 ☐ $200,000 or more

[T] Are you gay, lesbian or bisexual?
1 ☐ Yes   2 ☐ No

Please fold questionnaire and put it in the box. Thank you.

Major Changes to the 2020 Voting Process

Partisan Voting

Huge partisan differences in the voting method used by voters

Early Voting

- more than two-thirds of votes in 2020 were cast before Election Day –
- ~43% by mail
- ~24% in person at early voting centers compared to ~42% who voted before Election Day in 2016

Increase in Turnout

158 Million in 2020.
~20 million more than 2016

Vote Delays

This led to delays in reporting results in certain states and large swings in the results as they were reported over time

COVID-19 Protocols

- Exit Poll Interviewers wore mask at all times except during breaks
- Exit poll materials were placed on a folding table
- Gloves, hand sanitizer and sanitizing wipes were made available
- Interviewers were required to wipe down the folding table and use hand sanitizer before touching any exit poll materials
- Interviewers approached voters from a 6 foot distance
- Participants were instructed to use a single-use golf pencil from the table
- A garbage bag was provided for voters and interviewers to dispose of gloves, single-use golf pencils and sanitizing wipes
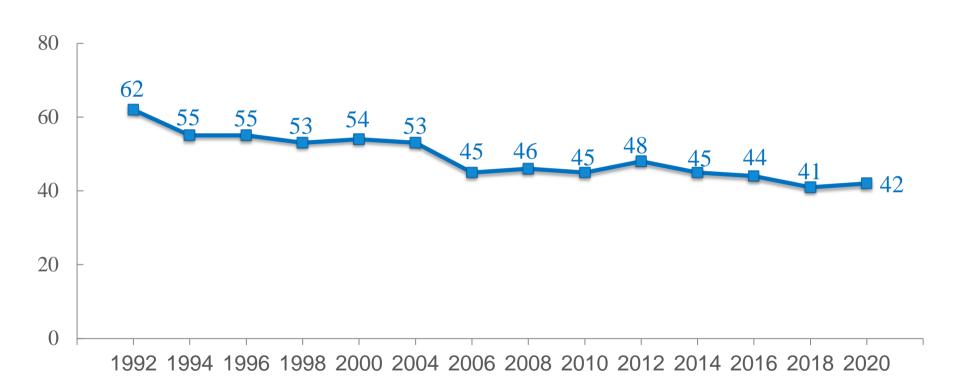
Managing Response Rates

- There is evidence of interviewer-respondent interactions that affect both the response rates by demographic group and the exit poll estimates

- Interviewer – respondent interaction:

  - Interviewer/Voter Age Gaps

  - Interviewer/Voter Racial Differences

- Edison weights the exit poll to account for non-response

4

Exit Poll Statistics

Historical Response Rates

Non-Response

Edison has developed weighting procedures to account for respondent non-response

- ➢ Age

- ➢ Race

- ➢ Sex

Statistical Estimation

Precinct Estimation

Edison uses detailed information to predict results within each precinct and overall.

Ratio Estimators

Other models are more simple and we build it on ratios using historical vote.

County Estimation

County vote is used to model and better predict results.

Geospatial Estimation

Geography is an important factor when estimating results and designing the sample.

Regression

We have many models that use several hundred potential predictors. Requiring variable reduction methods for more manageable models.
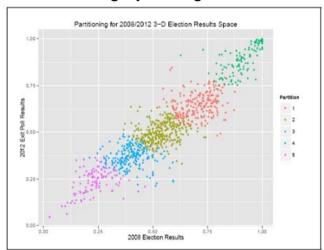
Statistical Estimation

Random Forests

For some of our models we have so many variables that we incorporate variable reduction techniques (e.g. VSURF).

Clustering

We run clustering algorithms to identify groups of precinct that tend to be similar.
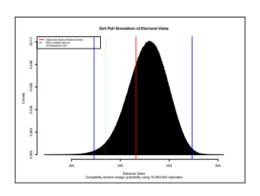
Clustering by Voting Behavior

Partitioning for 2008/2012 3-D Election Results Space

Other Experimental Approaches

Bayesian Estimation

We do varying degrees of Bayesian estimation and other MCMC methods.

Simulations

We simulate election outcomes to better tune our models and develop new models to handle different situations.

2020 Changes

With the changes in the way people vote we conducted a lot of research and development in 2020
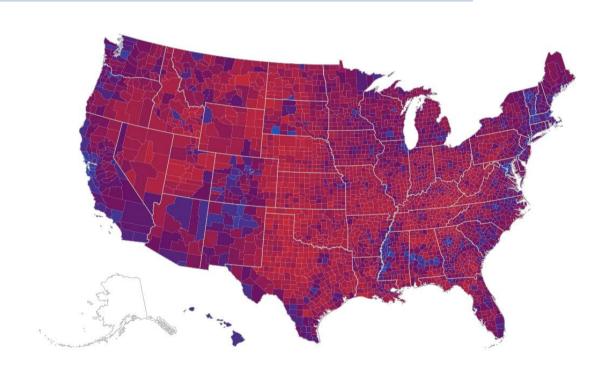
- New models to estimate turnout due to changes in early/absentee voting

- Changes in the way we estimate candidates using early voter polls.

- Changes to in-person field operations.

5

County Vote Models

Every County is Represented

County Vote Collection

In addition to the exit poll, beginning in 2018, Edison began to independentaly collect actual county vote from each of the county voting offices throughout the country.

Estimating County Vote

County Estimates

Edison collects vote results from every county in the country. We estimate candidate results in real-time

Total Expected Vote

We estimate the final end of night vote prior to Election Day. Then we implement a learning models to update that throughout the night.
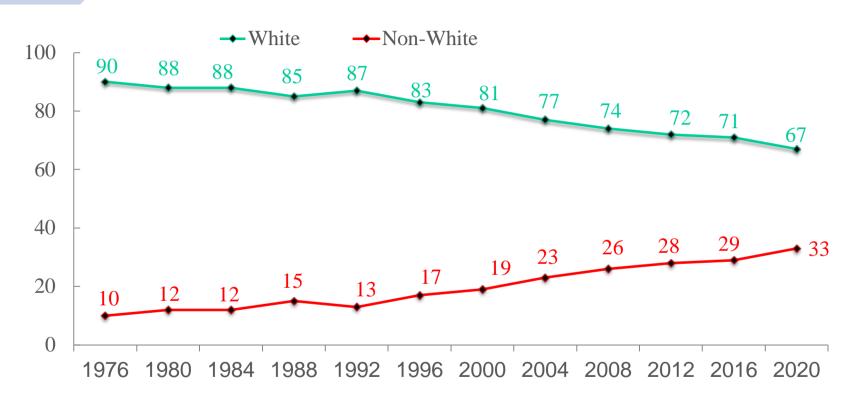
Variability Estimation

We use empirical and theoretical approaches to estimates variability of candidate vote for each county and congressional district.

6

Results

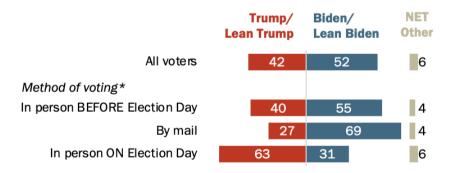# Historical Composition of the U.S. Electorate

	2020			2016		
	Total	Biden	Trump	Total	Clinton	Trump
18-24	9	65	31	10	56	34
25-29	7	54	43	9	54	38
30-39	16	51	46	17	51	39
40-49	16	54	44	19	46	49
50-64	30	47	52	30	44	52
65+	22	47	52	16	45	52
Total	100	51	47	100	48	46

Trump fares better among those who plan to vote on Election Day; Biden holds advantages among mail, early voters

% of registered voters who support …

	Trump/ Lean Trump	Biden/ Lean Biden	NET Other
All voters	42	52	6
Method of voting*			
In person BEFORE Election Day	40	55	4
By mail	27	69	4
In person ON Election Day	63	31	6

*Method of voting categories include those who have already voted.
Notes: Based on registered voters. No answer responses not shown. "Other" candidates include those who say they are voting for Jorgensen, Hawkins or someone else.
Source: Survey of U.S. adults conducted Sept. 30-Oct. 5, 2020.

PEW RESEARCH CENTER

33

Georgia President by Voting Method

	Joe Biden	Donald Trump
Absentee by Mail (26%)	64.5%	34.3%
Early Voting in Person (54%)	46.4%	52.7%
Election Day in Person (20%)	37.9%	60.0%
Total	49.5%	49.3%

34

North Carolina President by Voting Method
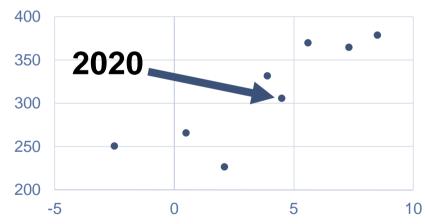
	Joe Biden	Donald Trump
Absentee by Mail (18%)	70.3%	27.9%
Early Voting in Person (65%)	46.6%	52.3%
Election Day in Person (17%)	32.8%	64.7%
Total	48.6%	49.9%

Calling the Elections

	Dem Popular Vote	Dem Electoral Vote
1992	+ 5.6	370
1996	+ 8.5	379
2000	+ 0.5	266
2004	- 2.5	251
2008	+ 7.3	365
2012	+ 3.9	332
2016	+ 2.1	227
2020	+ 4.5	306

Popular Vote vs Electoral Vote

2020

Calling the Elections

7

What Else Do We Do

We've conducted exit poll research in other places

Movie Theaters
Sports Arenas
Concert Venues

Retail Stores
Restaurants
Bars

Malls
Festivals
Events

Taxi Cabs
Doctor's offices
Office Buildings

Golf Courses
Convention
Centers Grocery
Stores

Airports
Subways
Bus Stops

Any many more locations…