The Real Steal? Electoral Forensics and the 2020 Election

Jonathan D. Simon

Introduction

Electoral Forensics is a field of which I consider myself a pioneer, certainly as it applies to the computerized voting era in the U.S. It is work I once only half-jokingly referred to as cleaning up after the election circus with a statistical pooper-scooper. With votes having been processed and counted in the pitch-dark of cyberspace, we scour data for patterns and clues that might shed light on whether the counts were performed honestly and accurately.

The 2020 election provided a trove of provocative data: disparities, anomalies, and bizarre patterns that exceeded what we have come to recognize as the red-flag norm. But *none* of the data supports former President Donald Trump's claims, and virtually *all* of it points in the precise opposite direction: in both the presidential and critical down-ballot contests (US House, Senate, state legislatures), the *red* shift (when vote counts come out to the right of poll-based or other baselines) was egregious and pervasive.

When the dust settled, Republicans won *27 out of 27* US House contests rated as "tossups" by *The New York Times*; Joe Biden's projected popular vote victory of nearly 13 million votes was cut practically in half; several Senate seats were "flipped" red relative to exit and/or tracking poll projections; and even at the state legislative level, expected Democratic gains turned into GOP gains.

All the anomalies and disparities worked to the benefit of Republican candidates (including Trump) and to the detriment of Democrats. *All*.

I have been doing this work for a long time. In 2004, I provided the primary data used to question the results, having been the only person to capture (by printing more than 300 pages of them out) the *unadjusted* exit poll crosstabs on that election night. Ever since that year, I have been working full-time to restore public, observable vote counting to our elections—and to impress the urgency of this basic reform upon those with the authority to make the necessary changes. Much of that work has consisted of presenting forensic analyses pointing to the likely exploitation of our voting systems' vulnerabilities to fraud.

Now I see electoral forensics has taken a dark turn. Our field of endeavor has been hijacked and weaponized by Donald Trump and his backers in the "Stop the Steal" movement: wild claims of theft and fraud have been trumpeted and echoed with *no* evidentiary backing, hard or soft; worse, these claims were "substantiated" with waved sheets of paper full of random and meaningless numbers, a stunt worthy of Joe McCarthy. Not only has this insistence that Trump won (because, well, he couldn't possibly *lose,* OK?) given rise to a massive GOP voter-suppression push and further undermined what trust remains in the electoral process, it has also served to bring electoral forensics—however rigorous, conscientious, and objective—into full-on disrepute.

Officeholders and media alike, whose legitimacy and credibility are grounded in a functional democracy and a trustworthy electoral process, have circled the wagons against Trump's half-cynical, half-crackpot assault and

also against *any* perceived forensic challenge, however solidly supported. We find our work less welcomed than ever, which is saying a lot.

This paper examines the data. It highlights and analyzes House and Senate contests, as well as the already controversial vote for president, and brings to bear (pp. 9-12), for the first time, a powerful and telling non-poll-based baseline that answers the standard dismissal that "the polls must have been off." You can't have polling bias where there is no polling!

This analysis is just a beginning. Like all other work in statistical forensics, it is not absolute *proof*, but it does point the way, like an X-ray scout film, toward a deeper investigation—including exhumation of voter-marked ballots and other critical election records—of both the vote casting and the counting processes in this election. It also takes its place as the latest in a long series of such troubling analyses dating back to 2004 and the inception of the Computerized Voting Era in America.

There is little question that there are thumbs on the electoral scales—some quite visible, and we suspect others hidden from view—and that these thumbs are heavy enough to make a profound difference in our electoral results, political balance of power, and national direction. There is little question that these distortions *matter* at this time in which we are facing seismic political forces and existential divisions.

If we don't expose the thumbs and get them off the electoral scales, our national prospects will be dim at best. We must, therefore, not allow Donald Trump's malignant weaponization of election "integrity" and undermining of authentic electoral forensics work to force a retreat from our efforts to present solid evidence that our electoral process remains corruptible and in urgent need of genuine repair.

Background: Election 2020 and its aftermath

It is no secret that, overall, Democratic candidates performed badly in the 2020 election (E2020), surprisingly if not shockingly below expectations at virtually every level of the ballot below the very top. This poor performance has put loss of congressional majorities very much in play for the 2022 election (E2022) and left the Democrats in significantly worse shape than expected at the state level, where critical rules of the electoral game are being "reformed" by the GOP to further dim Democratic prospects at both state and national levels.

And yet – even though November 3, 2020, was a bad day for a lot of Democrats not named Joe Biden, and for the party from its neck down – the impression among many is of a great Democratic *victory*, and among still others of an *illegitimate* Democratic victory. Donald Trump took control of the narrative and his "Stop The Steal" challenge to Biden's victory and legitimacy became *the story*, not only sucking up all the oxygen in the room but triggering a reflexive circling of the wagons among non-MAGA politicians and media alike, who all now sing in praise of our "perfect election" ("the most secure in history") and incorruptible processes. Most ironically, not only has Trump so far somehow managed to dance at will on the traditional Third Rail of election forensics – insisting there was outcome-altering fraud virtually everywhere he lost and it was close – but the voltage on that rail seems to have been tripled for everyone else, no matter how diligent and scrupulous their work.

Lost in this surrealistic turn of events is the reality that, while Trump's ongoing challenges emerged from a virtual evidentiary vacuum, there is much evidence suggesting that further investigation of a number of key down-ballot Democratic *defeats* is warranted – and suggesting as well that the Democrats head into E2022 and

E2024 at their peril, with various red thumbs poised over the electoral scales and ready once again to impose their weight – an even heavier weight, should that prove necessary.

This situation is not without historical context. Since the accelerated computerization of voting and vote counting with the passage and implementation of the Help America Vote Act (HAVA) in 2002, multiple indicators have signaled a recurrent "red shift," in which concealed, computerized votecounts in contests bearing national significance move mysteriously to the right of both poll-based and non-poll-based expectations. Extensive and often quite sophisticated analyses of such patterns – comparing computerized counting with hand counting, votecounts with exit and tracking poll results, competitive with noncompetitive contests, voteshare by precinct size, etc. – have consistently corroborated the red-shift meta-pattern. Little serious attention has generally been paid to these findings, however, despite the fact that they jibe with acknowledged concerns about the vulnerability of our electoral process to manipulation by outsiders and insiders alike.

Bearing this in mind, let's now examine several of the forensic red flags thrown up by E2020 and assess their role in what befell Democrats up and down the ballot. I think it is fair to say that, if the parties were switched and a losing GOP were looking at these same numbers, there would be both legal and PR hell to pay and a robust *down-ballot* "Stop The Steal" movement afoot. We will examine, in turn, the "big picture" of E2020 and a few more granular data sets that illustrate what might have transpired on a microcosmic level.

The National Picture

Among the numerous unpleasant surprises E2020 held in store for Democrats was a net loss of 11 U.S. House seats, when substantial gains had been predicted (FiveThirtyEight.com forecast a 16-seat net gain); loss of several U.S. Senate seats assigned a 50%+ Democratic win percentage; and failure to regain control of several state legislative chambers, as had been expected; indeed, a net loss of state chambers. To this list may be added, for forensic purposes at least, a presidential popular vote margin of victory cut nearly in half relative to both tracking and exit polls. Biden's "negative coattails" stood out as virtually unprecedented and, though various "organic" explanations were offered up for the bizarre overall pattern, the one safe conclusion that may be drawn is that that pattern presented *either* systemic polling failure *or* systemic targeted suppression and/or mistabulation of votes.

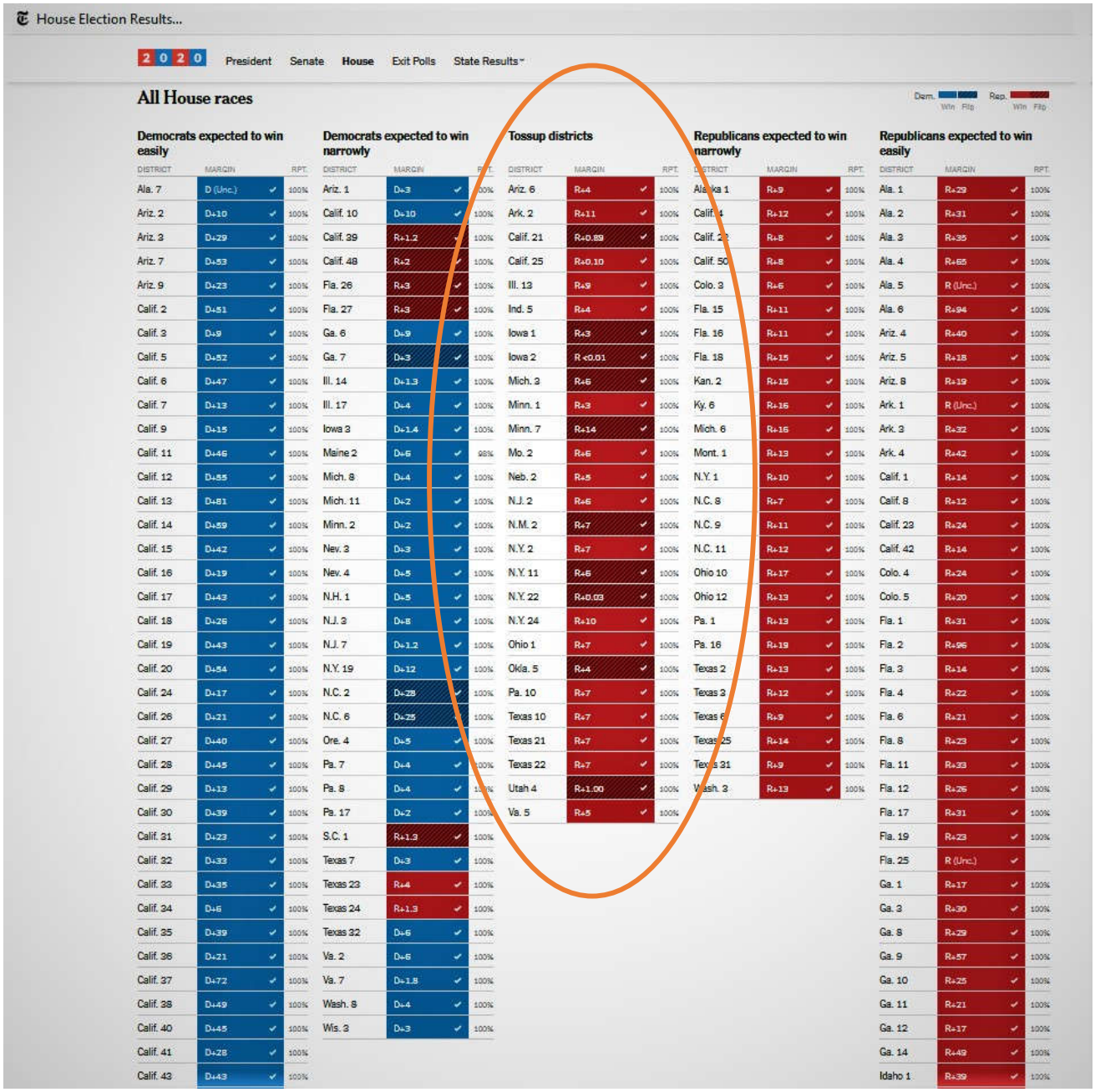
The House

The first systemic anomaly to emerge on post-mortem was that, of 27 US House contests rated as "tossups" by the *Cook Political Report* and *The New York Times*, Republicans were victorious *in all 27*. Assuming tossup equates to a 50-50 chance, a penny flip, the probability of such an outcome would be less than one in 100 million. Of course, we are *not* flipping a penny and other, non-random factors come into play, including the most obvious, which is the possibility of systemic sampling bias (selection and/or response) in the polling that contributed to the toss-up ratings (note that U.S. House contests are not individually exit polled, so this was all pre-election tracking polling). Nonetheless, the 27-for-27 table-run dramatically defied expectations – it is safe to say that if it had gone the other way, the GOP would not simply have shrugged and moved on.

The table on page 4 below illustrates that odd US House result, as well as showing that another seven contests rated D-win ("Democrats expected to win narrowly") also went Republican, while *none* of the 26 contests rated

R-win went the other way. This pattern is highly reminiscent of the distribution of close House contests in E2010 and E2014, with the significant difference that those were *midterms* during a Democratic presidency (in which attrition for the party occupying the White House is historically expected) rather than down-ballot races associated with a sweeping popular-vote victory by the Democratic ticket-header.

Virtually all these contests, being highly competitive, were heavily polled by a variety of different polling outfits situated across the political spectrum (though most were regarded as politically neutral), using a variety of methodologies and sampling techniques. The result was a data-rich rolling aggregate on which Cook and the *Times*, along with such entities as Nate Silver's FiveThirtyEight, based their predictions and ratings (for reference, the GOP took 15 of 16 US House contests rated as tossups by FiveThirtyEight).



Election 2020 US House contests categorized by predicted result (center column shows tossups; outside columns not fully shown) / NYTimes.com

We can also view the competitive House contests in terms of the red (or blue) shift and identify which ones "flipped" from predicted outcomes and by how much. The table below presents the 18 contests that flipped from blue (prediction) to red (outcome) and the single contest, in Georgia, that flipped from red to blue. The CDs are ranked from highest to lowest red shift; the right-hand column shows the expected Democratic win percentage and the left-hand column the principal equipment in use. The asymmetry is clear enough and the correlation of Election Systems & Software (ES&S) equipment with the most egregious red shifts worth noting.

Election 2020 US House Competitive/Flipped - Equipment

		11/3/2020 Vote Share*					11/3/2020 Polling Aggregate			
		Dem	GOP	Margin	%Reporting	Poll-VC Shift	Dem	GOP	Margin	DWin%
State/District	Equipment (Able Voters)					Red = +				538-Deluxe
FL-26	ES&S - DS200; EV-BMD	48.3	51.7	-3.4	100	11.6	54.1	45.9	8.2	82
IA-1	ES&S - DS200; EV-BMD	48.7	51.3	-2.6	100	10.6	54.0	46.0	8.0	87
FL-27	ES&S - DS200; EV-BMD	48.6	51.4	-2.8	100	10.1	53.6	46.3	7.3	81
TX-23	ES&S - DS200; EV-BMD	46.6	50.6	-4.0	100	9.0	50.6	45.6	5.0	74
IA-2**	ES&S - DS200; EV-BMD	50.0	50.0	0.0	100	8.6	54.3	45.7	8.6	88
OH-1	Hart Intercivic/Clear Ballot	44.6	51.8	-7.2	100	8.6	49.4	48.0	1.4	58
NY-2	Dominion Hybrid	46.0	52.9	-6.9	100	8.5	49.0	47.4	1.6	57
NM-2	Dominion Hybrid	46.3	53.7	-7.4	100	8.4	50.5	49.5	1.0	55
NY-11	ES&S - DS200; AutoMark	46.8	53.1	-6.3	100	7.9	50.8	49.2	1.6	58
CA-48	Hart Intercivic	48.9	51.0	-2.1	100	6.1	52.0	48.0	4.0	68
CA-39	Hart Intercivic/Smartmatic	49.4	50.6	-1.2	100	5.8	52.3	47.7	4.6	74
NY-22**	Dominion	49.0	49.0	0.0	100	4.6	50.9	46.3	4.6	73
OK-5	Hart Intercivic Hybrid	47.9	52.1	-4.2	100	4.4	50.1	49.9	0.2	51
IN-5	MicrovoteVVPAT/ES&S	45.9	50.0	-4.1	100	4.2	48.4	48.3	0.1	50
SC-1	ES&S - DS200; ExpressVoteBMD	49.4	50.6	-1.2	100	3.8	51.3	48.7	2.6	64
CA-21	Dominion	49.6	50.4	-0.8	100	2.6	50.9	49.1	1.8	58
UT-4	ES&S-DS200; Premier DRE	46.7	47.7	-1.0	100	2.1	46.5	45.4	1.1	56
CA-25**	Dominion/LACounty	50.0	50.0	0.0	100	1.0	50.5	49.5	1.0	55
GA-7	Dominion Image Cast X	51.4	48.6	2.8	100	-4.2	49.3	50.7	-1.4	43

*Updated 1/22/21 **Won by R

R Win D Win

Dpoll/Rwin Rpoll/Dwin

Prepared by Jonathan Simon using predictive data from www.FiveThirtyEight.com

This is a very basic forensic overview of the anatomy of the Democratic US House debacle. Given the paucity of competitive House contests that determine majority control of the chamber, such sweeping table-runs (see also E2010, E2014) have enormous potential impact. For such to occur in the context of a *failed* Republican presidency and a convincing overall (i.e., presidential) Democratic *victory* should raise serious how-and-why questions, among them how much of a role voter suppression and/or votecount manipulation may have played.

The Presidential Race

While all the clamor regarding the Trump-Biden contest has come from those who suspect – in millions of cases, to the point of *certainty* – that Trump won and the victory (in his words, a "*sacred landslide*") was somehow stolen from him, the forensic arrow points sharply in the opposite direction. While it is true that, as in E2016, the winner, courtesy of the Electoral College, benefited from a string of narrow margins in swing states, *none* of those margins was achieved, as in E2016, in contravention of polling and predictions; in fact, all exhibited some degree of red shift, with Biden's margin *slimmer* than expected. The table shows the poll-votecount disparities in rank order.

E2020 Poll-Voteshare Disparities Ranked By State

State	Voteshare			EP-VC Shift	Poll-VC Shift
	Biden	Trump	Margin	Red = +	Red = +
North Dakota	31.8	65.1	-33.3	NA	11.6
Montana	40.4	56.7	-16.3	8.3	10.0
South Dakota	35.6	61.8	-26.2	NA	9.2
Idaho	33.1	63.8	-30.7	NA	8.7
Wyoming	26.6	69.9	-43.3	NA	8.5
West Virginia	29.7	68.6	-38.9	NA	8.2
Oklahoma	32.3	65.4	-33.1	NA	7.9
Kentucky	36.2	62.1	-25.9	8.5	7.9
Tennessee	37.4	60.7	-23.3	NA	7.6
Wisconsin	49.5	48.8	0.7	9.8	7.6
Ohio	45.2	53.3	-8.1	2.8	7.5
Hawaii	63.7	34.3	29.4	NA	7.2
Iowa	44.9	53.1	-8.2	9.1	6.7
Utah	37.6	58.1	-20.5	NA	6.7
Delaware	58.8	39.8	19.0	NA	6.7
Rhode Island	59.4	38.6	20.8	NA	6.4
New York	60.8	37.7	23.1	4.1	6.0
Missouri	41.4	56.8	-15.4	NA	5.9
Florida	47.9	51.2	-3.3	2.9	5.8
Alabama	36.6	62.0	-25.4	10.9	5.4
Indiana	41.0	57.0	-16.0	NA	5.2
Michigan	50.6	47.8	2.8	6.4	5.2
Arkansas	34.8	62.4	-27.6	NA	4.9
Washington	58.0	38.8	19.2	-2.5	4.6
New Jersey	57.1	41.3	15.8	NA	4.5
South Carolina	43.4	55.1	-11.7	5.5	4.2
Connecticut	59.2	39.2	20.0	NA	4.1
Texas	46.5	52.0	-5.5	1.9	4.0
Nevada	50.1	47.7	2.4	3.8	3.7
Pennsylvania	50.0	48.8	1.2	1.8	3.5
Oregon	56.5	40.4	16.1	1.4	3.3
New Hampshire	52.8	45.5	7.3	2.8	3.3
Maine	52.8	44.1	8.7	4.8	3.1
North Carolina	48.6	49.9	-1.3	2.2	3.0
Vermont	66.1	30.7	35.4	NA	2.7
Virginia	54.1	44.0	10.1	1.1	2.4
Arizona	49.4	49.1	0.3	5.3	2.3
Mississippi	41.0	57.5	-16.5	NA	2.2
New Mexico	54.3	43.5	10.8	NA	2.2
Minnesota	52.4	45.3	7.1	0.3	2.0
Kansas	41.3	56.4	-15.1	NA	1.9
Alaska	42.8	52.8	-10.0	NA	1.6
Massachusetts	65.6	32.1	33.5	NA	1.4
Nebraska	39.3	58.5	-19.2	NA	1.2
Illinois	57.5	40.6	16.9	NA	1.0
California	63.5	34.3	29.2	-4.2	0.8
Georgia	49.5	49.3	0.2	2.6	0.7
Louisiana	39.9	58.5	-18.6	NA	0.0
DC	92.1	5.4	86.7	NA	-0.8
Colorado	55.4	41.9	13.5	-0.2	-1.8
Margland	65.4	32.2	33.2	NA	-1.8
Popular Vote	51.3	46.8	4.5	3.9	3.6

As can be seen, *all* the states targeted by the "Stop The Steal" challenges actually exhibit some degree of *red* shift from both exit (EP) and tracking (right-hand column) polls: Wisconsin 9.8% and 7.6%, respectively; Michigan 6.4% and 5.2%; Arizona 5.3% and 2.3%; Nevada 3.8% and 3.7%; Pennsylvania 1.8% and 3.5%; New Hampshire 2.8% and 3.3%; and Georgia 2.6% and 0.7%.

Thus there were *no* baselines indicating Trump victories in these critical states and *no* conventional forensic basis for suspecting he was robbed. In fact – again, barring systemic polling failure – the signal was that Biden's margins were *trimmed* as voters moved from poll to ballot.

Indeed, viewed as a function of national popular vote, that trimming was rather egregious. The aggregate disparity in national popular vote is shown here:

E2020 National Popular Vote Disparities

	Biden%	Trump%	Margin%	Margin In Votes	Disparity
Official Vote	51.3	46.8	4.5	7,060,140	NA
Aggregate Tracking Polls	53.4	45.3	8.1	12.6 Million*	5.5 Million*
Exit Poll (National Sample)	53.2	44.8	8.4	13.1 Million*	6.0 Million*

* Numbers rounded to significant figures

Trump Approval

Election Day Aggregate (538)	44.6%
Exit Poll (National Sample)	46.0%

It may come as a surprise to anyone keeping score at home who has been caught up in the furor over the "legitimacy" of Biden's victory that statistical forensics point in the opposite direction – that the magnitude of Trump's defeat was significantly mitigated in moving from *all polling* (with unadjusted exit polls and tracking

polls all but congruent) to official counts. Biden's 8.1% margin in aggregate national tracking polls translated to a win by some *12.6 million* votes; Biden's 8.4% margin in the unadjusted exit poll national sample translated to a *13 million-vote* win. It should be apparent that such margins would have been *catastrophic* from the standpoint of the Republican Party and its leadership and operatives: margin-of-defeat mitigation was effectively imperative. **The disparities with the official votecount margin of 7,060,140 were 5.5 million and 6 million votes, respectively. Please keep those last two large numbers in mind for subsequent reference.**

For anyone who suggests that it was all the polls that were "off," I have included, above, a key validation of both the tracking poll and exit poll samples: Trump approval, a highly stable measure, very strongly correlated with respondents' candidate preference (i.e., votes). In the tracking poll aggregate, Trump's approval stood at 44.6%; in the unadjusted national exit poll it was 46.0%. Clearly neither sample was bedeviled by Trump-supporter under-participation or -representation (i.e., selection or response bias). Both approval numbers are well *above* stable trends and indicate *at worst* fair samples – slanted, if anything, a few percentage points *against* Biden.

I will return to the question of what might have happened to the between 5 and 6 million Biden votes that were measured by what appear to be robust and fully validated polling samples but did not make it into the official results of E2020.

The Senate

After popping the cork in January to celebrate their bare 50+VP Senate majority, the Democrats forthwith had to face the reality that in many respects Mitch McConnell and the GOP's power in the chamber was effectively undiminished. With little hope of mustering even the bare partisan majority needed to nix, in whole or part, the filibuster, both the Biden agenda and a record to run on in E2022 are in great jeopardy – not to mention federal voting rights legislation essential to counter state-level GOP "reforms" aimed at further suppressing the vote. And this is *with* the aid and comfort of the two razor-thin Georgia runoff victories, outcomes attributable by most accounts to the perverse insistence of Donald Trump on injecting himself into those elections and attempting to make them referenda on his "Stop The Steal" agenda.

Comparative Performance By Office in Select Venues E2020

VENUE	DemVote	GOPVote	Total Vote (2-Party)	D-Margin (2-Party)	Dem%	GOP%	D-Margin (2-Party)
MAINE							
Total Vote for President	435072	360737	795809	74335	54.7%	45.3%	9.3%
Total Vote for House	468978	340236	809214	128742	58.0%	42.0%	15.9%
Total Vote for Senate	347223	417645	764868	-70422	45.4%	54.6%	-9.2%
House/Senate Delta	121755	-77409		199164	12.6%	-12.6%	25.1%
Pres/Senate Delta	87849	-56908		144757	9.3%	-9.3%	18.5%
House/Pres Delta	33906	-20501		54407	3.3%	-3.3%	6.6%
WISCONSIN							
Total Vote for President	1630673	1610065	3240738	20608	50.3%	49.7%	0.6%
Total Vote for House	1566671	1661399	3228070	-94728	48.5%	51.5%	-2.9%
House/Pres Delta	-64002	51334		-115336	-1.8%	1.8%	-3.6%
NEW MEXICO							
Total Vote for President	501614	401894	903508	99720	55.5%	44.5%	11.0%
Total Vote for House	495781	407786	903567	87995	54.9%	45.1%	9.7%
Total Vote for Senate	474483	418483	892966	56000	53.1%	46.9%	6.3%
House/Senate Delta	21298	-10697		31995	1.7%	-1.7%	3.5%
Pres/Senate Delta	27131	-16589		43720	2.4%	-2.4%	4.8%
House/Pres Delta	-5833	5892		-11725	-0.6%	0.6%	-1.3%

Heading into E2020, the Democrats were given a fair chance of winning an outright 51+ Senate majority. Although Democratic candidates held narrow tracking-poll leads on November 1 in Iowa and North Carolina, both of which wound up in the GOP column, by far the biggest surprise was Maine, an ES&S state where Democratic candidate Sarah Gideon went from a steady 2- to 3-point lead to an

8.6% defeat by incumbent Susan Collins (who was given McConnell's blessing to vote No on the polarizing Barrett SCOTUS nomination, though any boost from that vote was not reflected in a significant shift in tracking polling). Mainers I have consulted are divided about what caused Gideon's collapse – and thus about whether the results should be viewed with skepticism. It was striking how dramatically Gideon under-performed relative to her Democratic ticket-mates up *and* down the ballot. The table above at left shows the deltas (differences in margins for House, presidential, and Senate contests) within each state; I have included Maine at top along with a couple of other states to illustrate the more ordinary deltas (green highlight) found elsewhere.

The delta numbers in Maine are quite egregious, the state's ticket-splitting reputation notwithstanding. Without deeper drilldown, culminating in an examination of ballots and/or ballot images (cf. Maricopa County, Arizona), they remain an anomaly that may or may not be signaling interference. One footnote is that the exit poll in Maine, which had Collins the victor by 1.6%, relied on a mere 1,119 respondents, while other highly competitive states ranged from a low of 1,639 respondents (Arizona) to a high of 4,734 (Texas). As the size of a jurisdiction's population has little impact on the sample size needed for a given accuracy, and as Maine was highly competitive and significant, the exceptionally low sample size (cf. Solid-R Alabama at 998) is hard to explain from a methodological standpoint. Also of note is the fact that Collins was credited with a majority of the vote (51.2% of the *total* vote, of which 6.6% went to minor candidates and would have been redistributed had Collins not topped 50%), which meant that she avoided having state-mandated ranked-choice voting (RCV) come into play for this election. As the RCV process would have entailed a separate, secondary mode of counting, any manipulator of the first count would have been strongly motivated to avoid it by exceeding the 50% threshold.

The current Senate struggle should be viewed in light of the durable structural imbalance of the upper chamber embedded in the Constitution as applied to modern America. As of 2020, half the U.S. population was represented by just 17 senators, the other half by 83. The gross over-representation of low-population rural states (e.g., Wyoming has Senate representation equal to that of California, while having less than 1.5% of the larger state's population; this imbalance also serves to weight the Electoral College) tilts the Senate red, a skew that has become fully evident only with the complete erosion and transformation of what was once the Democratic "Solid South." Under the current alignment, a bare Senate majority is generally a *ceiling* for the Democrats, and it would take abysmal governance and great unpopularity by the GOP to enable any enduring Democratic breakthrough.

U.S. Midterm Election Turnout: Total Votes Cast for House of Representatives 1962 – 2018

Election Year	Total Votes Cast (Turnout)*	Change From Previous Midterm	Percent Change From Previous Midterm	Bar Chart of Election-to-Election Turnout Change
2018	**114,016,831**	**35,204,062**	**44.7%**	**44.7**
2014	78,812,769	-7,972,188	-9.2%	-9.2
2010	86,784,957	5,809,420	7.2%	7.2
2006	80,975,537	6,268,985	8.4%	8.4
2002	74,706,552	8,101,750	12.2%	12.2
1998	66,604,802	-3,888,846	-5.5%	-5.5
1994	70,493,648	8,138,795	13.1%	13.1
1990	62,354,853	2,596,456	4.3%	4.3
1986	59,758,397	-4,122,447	-6.5%	-6.5
1982	63,880,844	9,296,922	17.0%	17.0
1978	54,583,922	2,270,465	4.3%	4.3
1974	52,313,457	-1,945,428	-3.6%	-3.6
1970	54,258,885	1,356,910	2.6%	2.6
1966	52,901,975	1,659,787	3.2%	3.2
1962	51,242,188	With admission of Hawaii and Alaska, U.S. becomes 50 states.		
Average 1962 - 2014			**3.7%**	
Abs. Val. Avg. 1962 -2014			**7.5%**	

In fact, the Trump presidency presented just such a scenario and opportunity. Consider the explosion in midterm turnout for pre-COVID E2018 (see table at left), in which Trump was effectively "on the ballot" in a plethora of proxy elections. *Remarkable*, and yet the Democrats (who were defending two more competitive Senate incumbencies than was the GOP) *lost* a net of two seats to the GOP. In E2014, with congressional approval at a dismal 8% (with the GOP holding the House majority and enough senators to filibuster anything of significance proposed by Obama

and the Democrats), the GOP returned 220 out of 222 House incumbents seeking re-election and *picked up a net of nine seats* in the Senate.

There is a long history of "shocking" GOP victories in Senate races, dating back at least to Cleland/Chambliss (GA) in 2002, the year that HAVA passed and Georgia lost no time in deploying unverifiable, paperless touchscreen voting machines (DREs). Anyone looking at America's two-decade political veer to the right should study *all* the thumbs the GOP has sitting on the electoral scales – structural, administrative, and most likely electronic.

State Legislatures

Heading into E2020, the GOP had a major overall advantage in control of state legislatures, which had largely survived the supposed "blue tsunami" of E2018. Of the 99 chambers (Nebraska is unicameral), the GOP controlled 59 to the Democrats' 39 (one chamber had a power-sharing arrangement) and held 21 "trifectas" (control of both chambers and the executive) to the Democrats' 15. To a significant extent, this breakdown is a function of there being more "red" than "blue" states, though of course that begs the question of what makes a state red if not these very numbers – the post-HAVA, computerized voting era trend has been strong reddening at the state level.

With E2020 essentially a referendum on Trump, and with forecasts of expanded turnout, the Democrats were amped up about their prospects of cutting into the GOP state-level advantage. With the anticipated Biden coattails and high turnout, several states, including such long-time GOP bastions as Texas, seemed to be in play for the Democrats. Instead, the GOP *widened* its advantage in chambers to 61-37 and *added* two trifectas (Montana and New Hampshire), yet another manifestation of Biden's bizarrely negative coattails.

Gerrymandering is, of course, an obvious factor in state legislative control, with districts, to an even greater extent than US House districts, being carved by the majorities to protect their incumbents and hold on power. But the E2020 results at the state level were nonetheless unanticipated and, to many observers, shocking. Although this foundational, or infrastructural, level of American politics receives minimal media attention and is off the radar for all but the political caste and its strategists, it is of profound national importance, as can be seen in the current wave of legislation restricting voting rights and handing control of electoral administrations (and outcomes) to highly partisan legislatures in key states where the GOP has control. As with the US House, relatively few state legislative seats are actually competitive and in play – and even these contests are rarely polled. So there are effectively no baselines available to get a forensic handle on votecounts and outcomes – which makes these contests, sprinkled around the country but bearing outsized national significance, the lowest hanging fruit from the standpoint of operatives who would interfere with the vote casting and/or counting processes. These key contests can be manipulated with virtually no risk of investigation or detection and for enormous political reward.

A Key Indicator

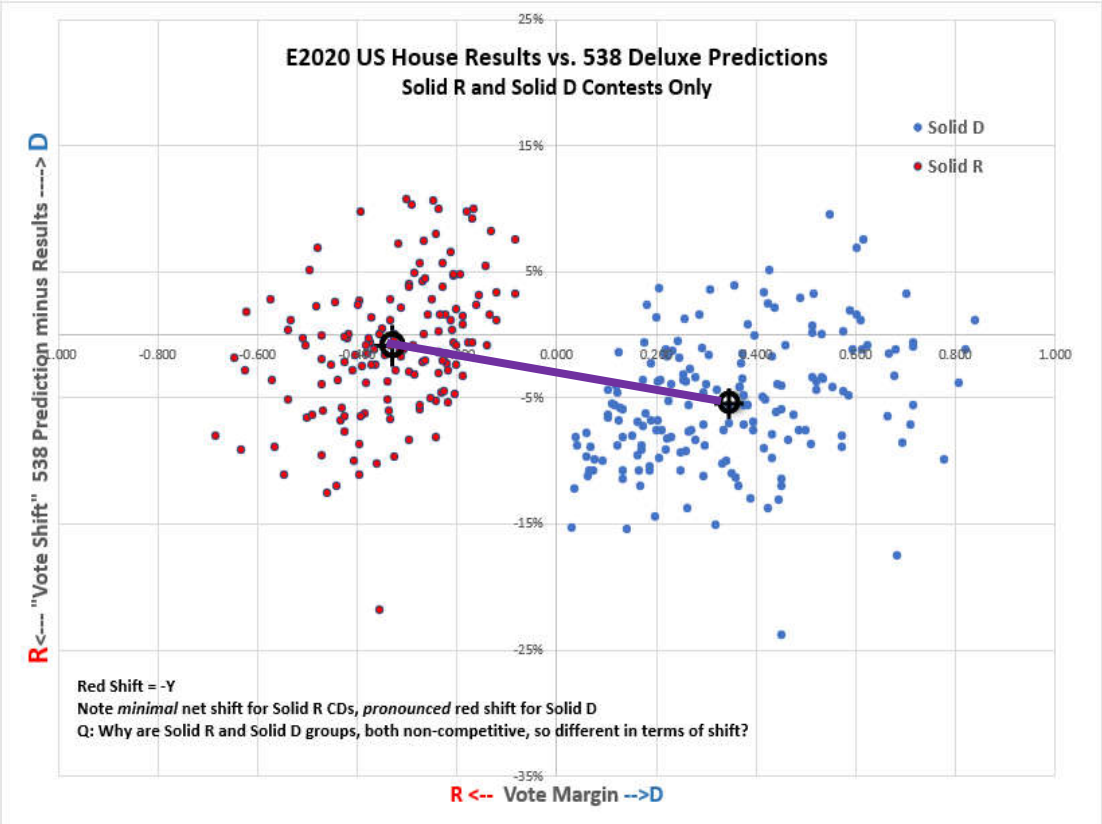
A national election generates a vast amount of data, ranging from the counts themselves (down to the precinct level) to a plethora of tracking polls, exit polls, turnout statistics, and of course archival data from prior elections. At varying degrees of granularity, signals emerge from this great numerical geyser that may, viewed through the appropriate lens, indicate improbable and anomalous patterns worthy of further investigation. In E2020, we saw unexpected, perplexing results at every level – US House, US Senate, presidential, and state legislative – as detailed above. And we saw that the shifts relative to baselines were virtually all in the same direction: blue to red. For the

GOP to win 27 of 27 US House races rated tossups on the basis of tracking polls, for example, is, from a probabilistic standpoint, beyond bizarre – and indicative of either a massive systemic polling failure or targeted manipulation of the contests themselves. But, because these baselines are poll-based, whether tracking or exit or both, without additional evidence analysts cannot get beyond that either/or assessment. And the default position regarding U.S. elections has long been that when the polls do not match the votecounts, however egregious the disparity, it is the polls that will be deemed to be "off." Trump's all-out and continuing assault on democracy and the legitimacy of our elections has only strengthened that resolve and the blanket refusal to consider the possibility that the fault may lie not in the polling but in the counting.

With all that in mind, I combed the E2020 data dump for an indicator of some sort that was not poll-based. In examining non-competitive contests for the US House, I found one.

The scatterplot below includes only those E2020 US House contests that were rated either Solid-Republican (red dots) or Solid-Democratic (blue dots) by FiveThirtyEight.com; altogether there are 324 contests; the 74 competitive and 37 uncontested US House races have been excluded. For each of these contests, FiveThirtyEight generated a predicted result: a winner and both a voteshare and a win percentage (likelihood of winning) for each candidate. ***However, because these contests were all seen as non-competitive, virtually none of them were polled.*** The predictions, therefore, were based almost entirely on *other* factors, such as voter registration, prior results, prior candidate performance, campaign expenditures, etc. – none of which involved sampling of respondents. This is, as will be seen, of great significance.

In the scatterplot, the x-axis shows the margin of victory, Democratic being positive (+x); the y-axis shows the disparity between predicted and actual results, with "red shift" (results better than predicted for GOP candidate) being downward (-y). The crossed circles represent mean x,y values for each group, Solid-R and Solid-D.



The first thing to note is that no dots cross the center line (y-axis): blue are all +x and red are all -x, which means that all 324 predictions *correctly* identified the winners. The next is that the distribution is clearly (and oddly) bi-modal with respect to y – that is, the blue and red clusters are visibly distinct in their y-values or degree of red shift. The Solid-R mean is just barely below the x-axis, a y-value of -0.9%, a very minimal red shift; the Solid-D mean, on the other hand, is well below the x-axis, a y-value of -5.7%, a major red shift.

We naturally asked what might account for such a distinct pattern. Why were the predictions so much worse in places with high concentrations of Democratic voters (and votes) than in Republican strongholds? When I first showed this pattern to one-time pollster (CBS) and long-time polling expert David Moore, his first thought was a particular type of sampling bias in which GOP/Trump voters would be more comfortable responding to pollsters in Republican strongholds where they were in the majority than in Democratic strongholds where they were a dwarfed minority – call it the "shy Trump voter away from MAGA-land" hypothesis. There is a working theory out there that Trump/GOP voters in general are more likely to be "shy" and refuse to respond to polls (which they associate with the despised, liberal, "fake-news" media) than are their Democratic counterparts – a theory often trotted out to explain the major red-shift disparities of E2020 (and prior elections) as a massive *polling* failure. It certainly is no great stretch to extend that theory to take into account the politically friendly or hostile environments in which such voters find themselves.

But recall that, with very rare exceptions, the predictions that generated this scatterplot were *not based on polls*. There could be virtually no "shy" Trump/GOP respondents *because there were virtually no respondents*. The bi-modal distribution above consists of a large number of data points and the statistical significance is very high. We are not surprised by the relatively wide dispersion of the y-values in both clusters – we expect a fair amount of dispersion or "noise" with non-poll-based predictions. But we *would* expect roughly the same *mean* y-value (or red/blue shift) for both clusters – there should be no correlation between partisanship and prediction/result shift. A line connecting the crossed circles in the plot above should be *horizontal*, not a slant.

So something clearly was happening in Blue America that was not happening in Red America. Were these particular contests targeted for manipulation? Of course that makes no sense: these were blowout races and, as we saw, every winner was correctly predicted. No rational rigger would target any of these contests. How, then, to read this riddle?

Our working hypothesis is that these contests resided on ballots in which <u>other</u> contests may well have been competitive and of national significance and therefore attractive targets for interference of one form or another.

Many of the blue dots above represent gerrymandered urban CDs in presidential or senatorial battleground states; some even contain within them or overlap with competitive *state* legislative districts. These Solid-D U.S. House races were not targeted but it appears that at least several million *whole ballots* on which they resided were either not successfully cast, went uncounted, or were mistabulated (i.e., flipped). The "safe" U.S. House contests were, in other words, collateral and politically insignificant damage, but a strong *signal* of process disenfranchisement of predominantly Democratic voters living in Democratic strongholds.

Many of these would be the very voters of color, residing in primarily urban and easily identified zip codes, who have been the targets of both open and, we suspect, also less visible GOP suppression tactics. They might be seen waiting on hours-long lines, many leaving, by necessity or in discouragement, before voting; they might be

dropped from the voting rolls in sweeping, targeted county- or state-wide purges; their mail-in ballots might be delivered late or not at all; their signatures might be rejected as "not matching." The fact is there are *many* ways to tamp down on the casting and counting of whole ballots – especially in venues where the votes to be suppressed are heavily clustered and readily identified by the intersection of demography and geography: the same "Big Data" used for ruthless, precision gerrymandering easily flags precincts and zip codes for suppression. As Willy Sutton put it when asked why he robbed banks: "Because that's where they keep the money." The scatterplot above strongly suggests that "banks" holding lots of Democratic ballots were "hit" in E2020.

This signal of large-scale interference with the casting and/or counting processes jibes with the investigative reporting of Greg Palast into voter suppression, specifically the systematic suppression of the votes of traditionally Democratic constituencies. Throughout the computerized voting era, Palast has uncovered illegal purges of voter rolls as well as disqualification and spoliation of ballots in numbers large enough to produce the extra 5 percent red shift in the Democratic areas on the SolidR/SolidD scatterplot above. In addition, the surge in mail-in ballots spurred by the pandemic and the tweaking of voting protocols to accommodate it made the U.S. Postal Service a potential choke point for vote casting, whether through delayed delivery of requested ballots to voters, delayed delivery of completed ballots to counties, or the loss or destruction of ballots in USPS custody. It remains unclear whether and to what extent suspected schemes by Trump's postmaster general, Louis DeJoy, came to fruition.

What is clear is the signal flashed by the scatterplot presented above: a significant red shift of votecounts relative to non-poll-based predictive baselines, collectively impacting venues with a high concentration of Democratic voters – millions of votes, nearly enough to account for the 6 million-vote poll-votecount red shift in the presidential race, and likely impacting competitive Senate contests and state legislative contests as well. If the same level of interference is extrapolated to the competitive U.S. House contests, it would of course provide an explanation for the Republican 27-for-27 table-run of "toss-up" races and flipping of seven Democratic leaners.

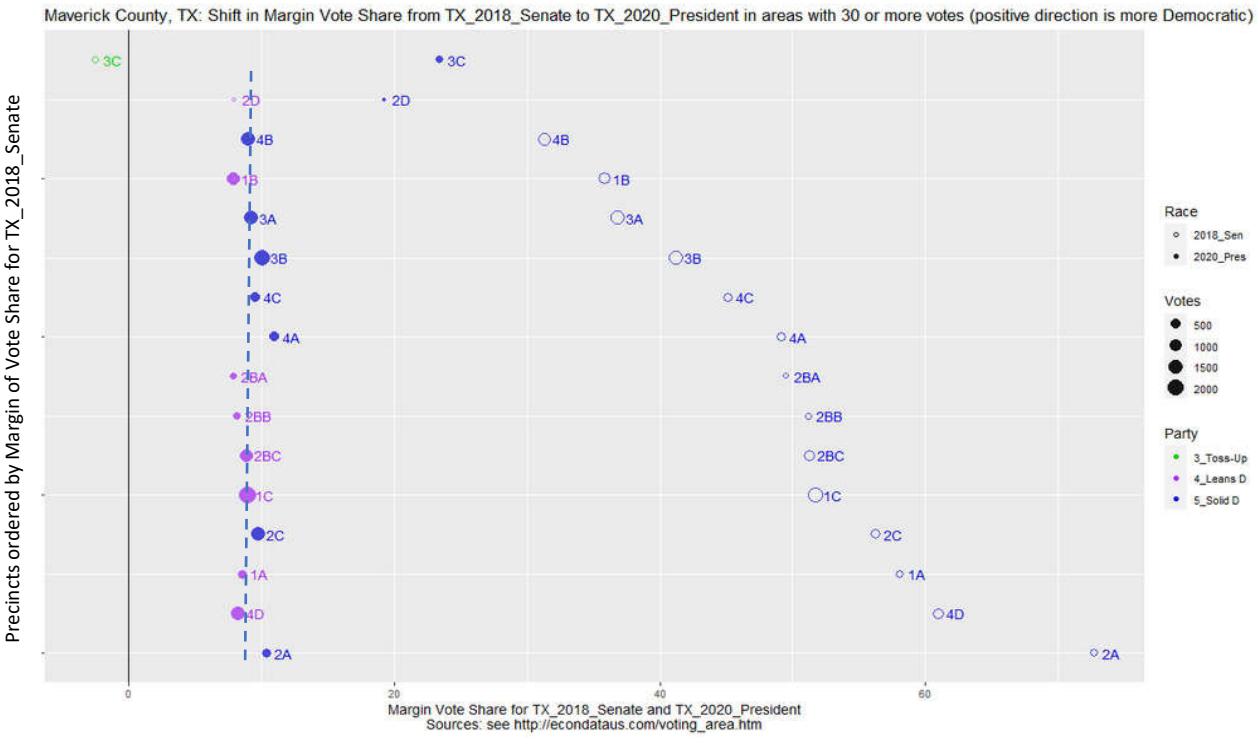
In Microcosm: Texas Red Flags

Our examination to this point has consisted primarily of a bird's-eye view of E2020's forensic "lean" – the "red shift" from predictions and expectations to results at the level of national significance. If in fact red thumbs tilted the electoral scales in a way that showed up in the national vote for president and table-run of U.S. House races, we should be able to see odd patterns and anomalies reflected when we drill down to the precinct level. Let's accordingly examine a few examples to illustrate how such patterns might present, a preliminary guide to some of the potential red flags to look for in the county- and precinct-level data available for public inspection.

Maverick County, Texas

Maverick is a sparsely populated county bordering on the Rio Grande in southwest Texas. It is serviced by ES&S and uses hand-marked paper ballots, optically scanned, for its able voters; these ballots should be preserved for 22 months post-election by federal law and subject to FOIA examination under Texas law (assuming a public records exemption has not been carved out). Maverick is included here to illustrate a precinct-level shift pattern that is not readily explained by organic factors (all graphs in this section are the product of www.econdataus.com; for additional data specific to Maverick County, please see https://econdataus.com/voting_area.htm#maverick).

The plot below shows the individual precinct results for the 2018 U.S. Senate election (the diagonal subset of hollow dots to the right) and the 2020 presidential election (the nearly vertical set of solid dots to the left), *as they stood on June 2, 2021, six months after Texas certified the 2020 results,* archived in *The New York Times* Election 2020 database. Democratic margin of victory is plotted on the x-axis (+x); Maverick is a strongly Democratic-leaning county. The downward direction indicates increasing delta-x, or shift from E2018 to E2020 (only two outlier precincts, 2D and 3C, exhibited a reverse shift; 2D is an extremely small precinct):

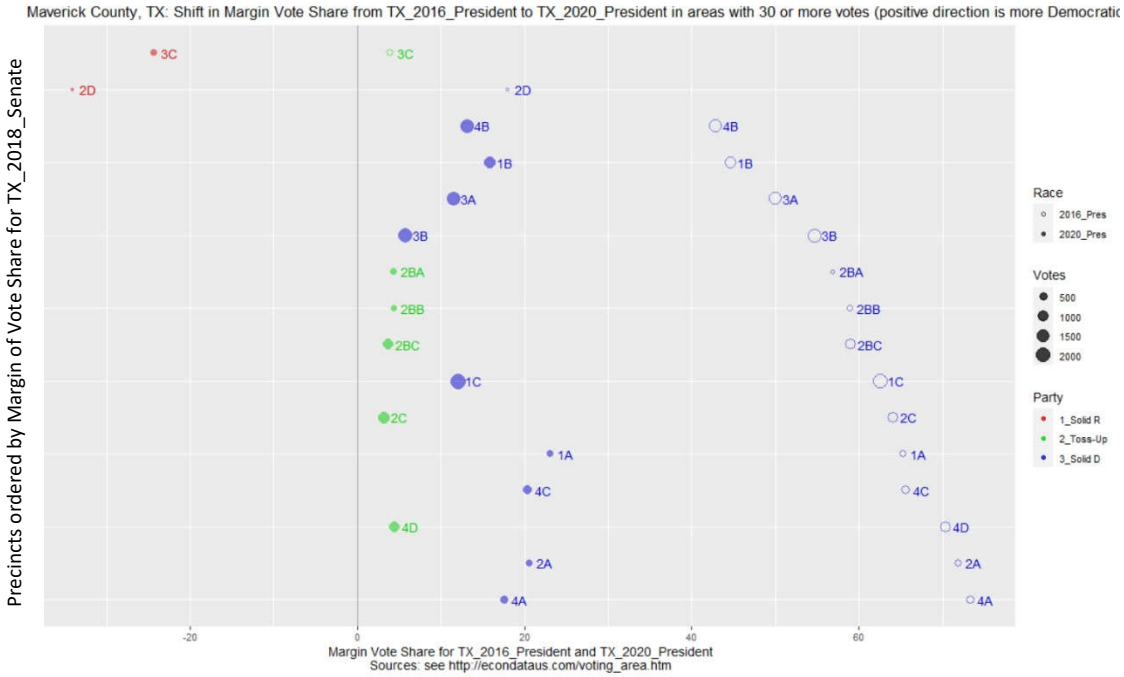


What is remarkable about this chart is that the E2020 result appears to bear no relation to the E2018 result. That is, no matter how large the 2018 Democratic margin for a particular precinct (except 3C and tiny 2D), the E2020 margin appears to be "fixed" within a percentage point or two of 10%. So, for example, a precinct (4B) that went 66%D/34%R (32% margin) in 2018, wound up 55%D/45%R (10% margin) in 2020; while a precinct (2A) that went 87%D/13%R (74% margin) in 2018, *also* wound up 55%D/45%R (10% margin) in 2020. And this "regression" to 55%D/45%R (the near-vertical line to the left) occurred in every precinct except the two outliers noted.

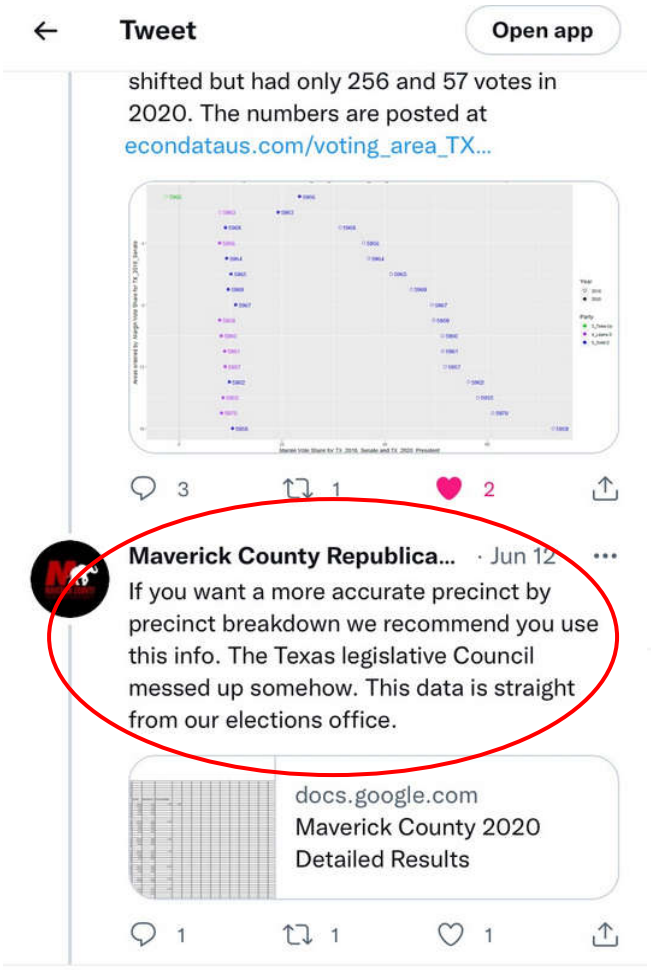
The question raised by this distribution is what would cause 14 of the 16 precincts in Maverick County, regardless of how blue in 2018, to vote essentially the same – i.e., as a bloc – in 2020? If there had been a county-wide attrition of Democratic votes (caused by some organic factor such as the jobs impact of border-wall construction or a spectacularly successful GOP GOTV operation), one would expect something like a consistent or proportional loss in each precinct. That would result in a translation of the 2018 diagonal distribution to the left – but it would *remain diagonal*. The vertical line that represents the actual 2020 vote (as tabulated) is far more difficult to explain organically. In fact, it has the look of a programmed mistabulation – as if a kind of "quota" had been established for Maverick, such that this solidly blue county would not merely bleed Democratic votes but would come in at a narrow, but still plausible, county margin of 10%, 55%D/45%R (the official county result was 54.3%D/44.8%R).

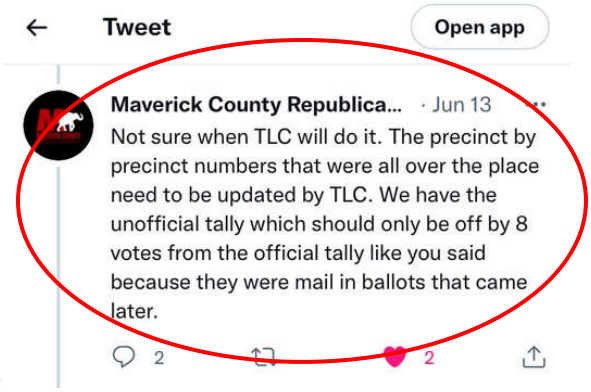
There are several known ways to effectively pre-program electoral outcomes, on both touchscreen (DRE) and optical scan computers; those mechanics are beyond the scope of this analysis. One would assume that a careful manipulator would avoid the vertical line to the left, and the strange pattern seen above, as a potential calling card or fingerprint of manipulation. But that is also to assume that: 1) A precinct-level forensic drilldown will be undertaken; 2) Its odd result will be interpreted as a red flag; 3) That red flag will be taken seriously, prompting further investigation; and 4) Such investigation will be successful in obtaining the hard evidence (in this case, the hand-marked ballots, sorted by precinct and chain of custody maintained) needed to confirm or dismiss the suspicion. That is a lot of assuming, so from a manipulator's standpoint, such care is unlikely to seem essential.

But the Maverick story does not end there. Subsequent to the screencapturing and posting of the E2020 results giving rise to the plot presented above, the Texas Secretary of State published revised numbers, redistributing the Maverick votes among its precincts without significantly changing the county totals. With the votes redistributed, the above chart now looked like this:



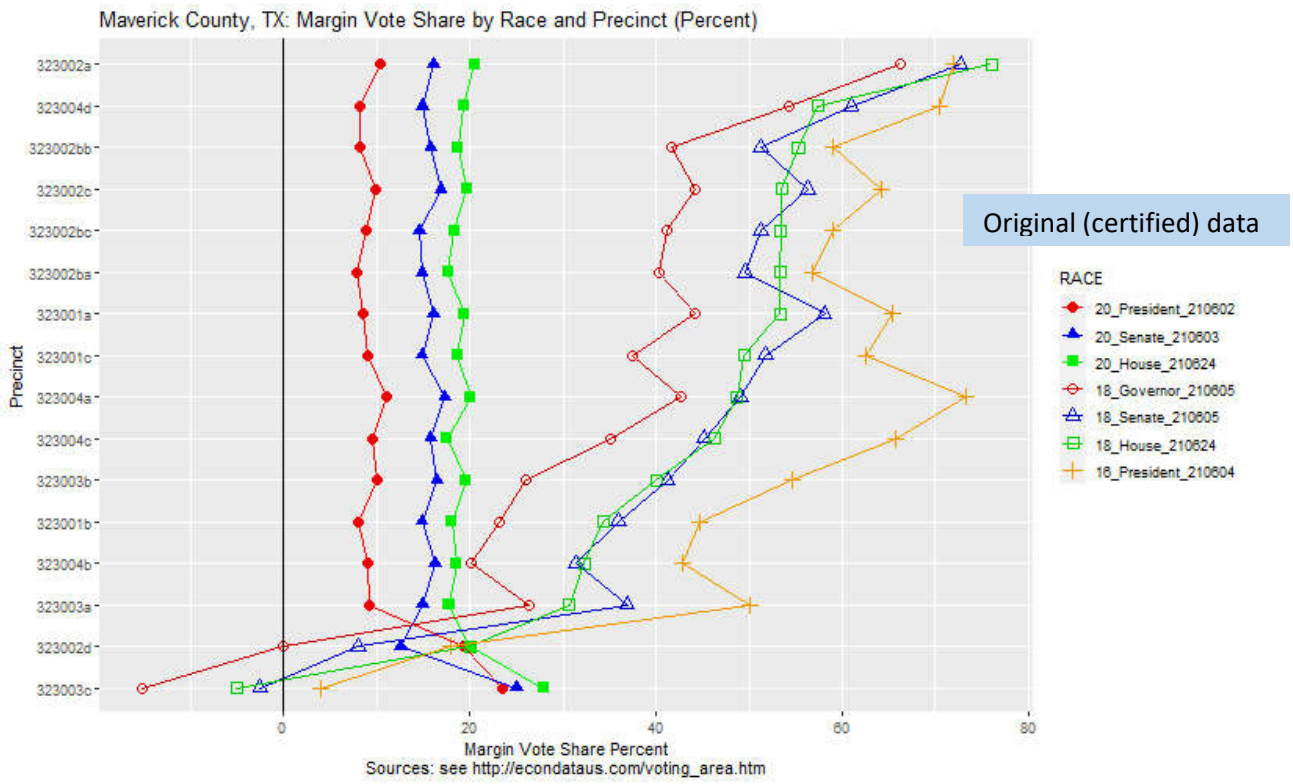
As can be seen, the near-linear distribution of the E2020 presidential votes on the left now varies widely from the vertical (as we would expect it to) and has no particular shape overall, although a half dozen of the precincts do remain more or less vertically aligned. So some *six months* after the election was certified, the suspect E2020 pattern that we posted publicly was "repaired" and no longer stood out like a numerical sore thumb. After our posting on Twitter of the original chart, showing the votes as originally distributed, the Maverick County Republican Party responded with these cryptic posts, fingering the "Texas legislative Council" (TLC):



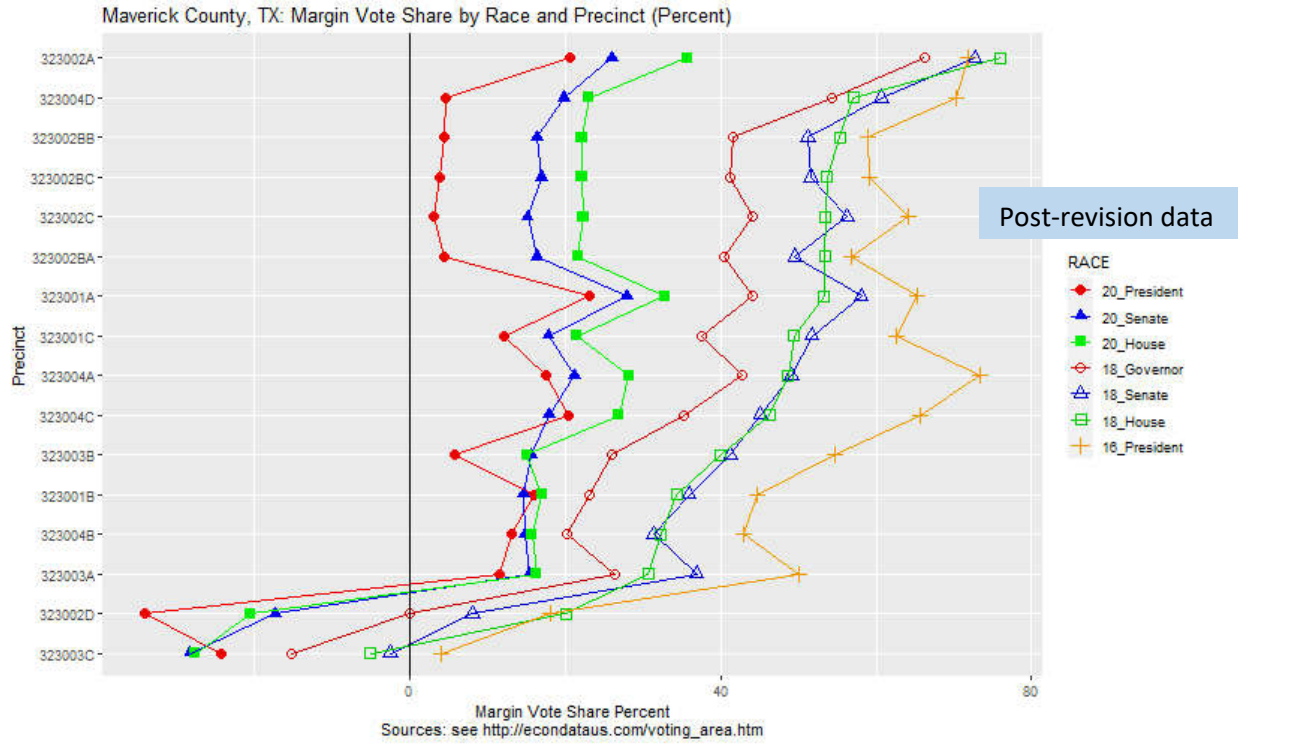


The timing of this "corrective" action should raise additional eyebrows. The TLC, whose task it is to provide and model data used in the redistricting process, did recently respond with an explanation that raised as many questions as it answered. It is not clear why the TLC was relied upon at any point to make public – such that it was incorporated into official results as reported by the *Times* – precinct-level vote tallies; or why, if these tallies were, in effect, place-holder "dummies," they were presented and published without any asterisk indicating such was the case.

Two additional plots – showing contests for other offices both in 2020 and previous years – provide a more comprehensive picture of the initial and redistributed precinct-level data:



Original (certified) data

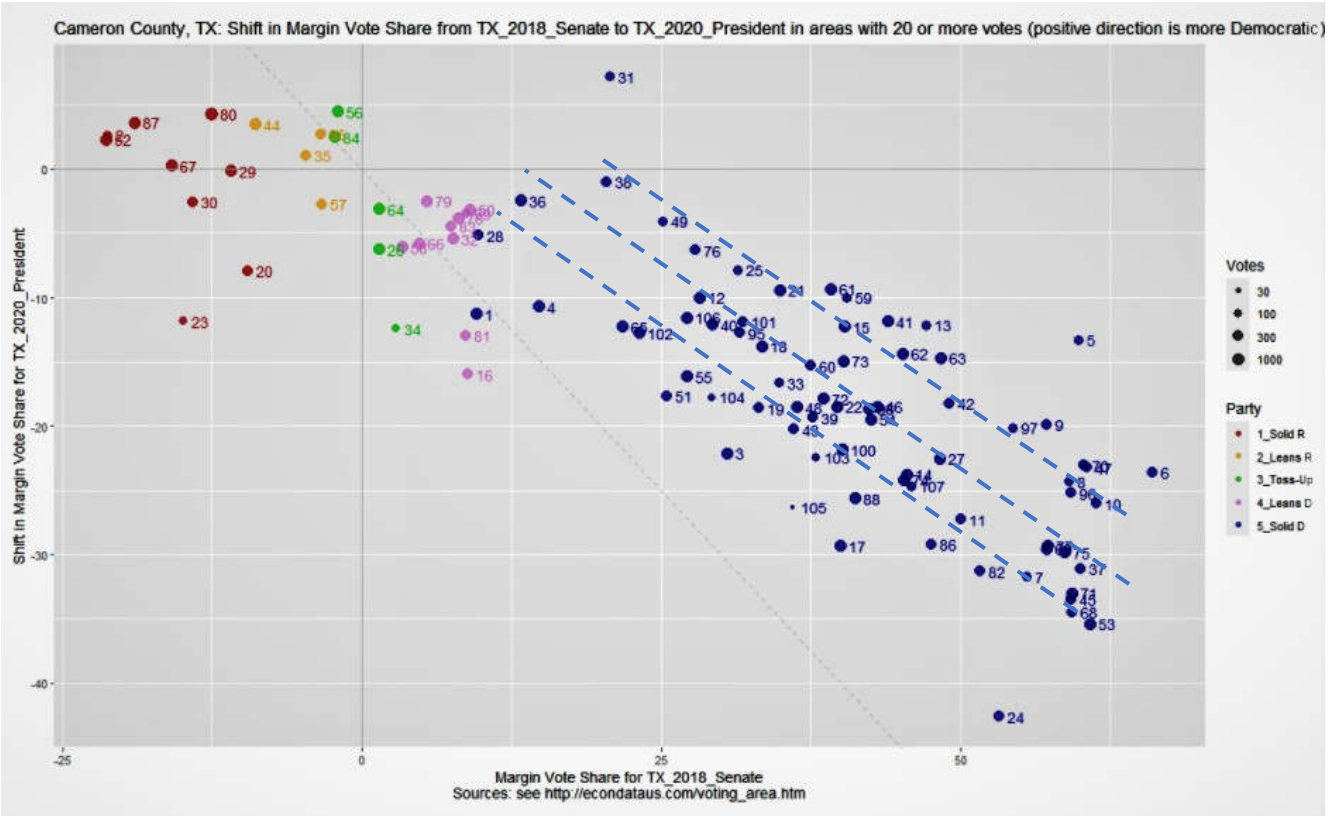


Post-revision data

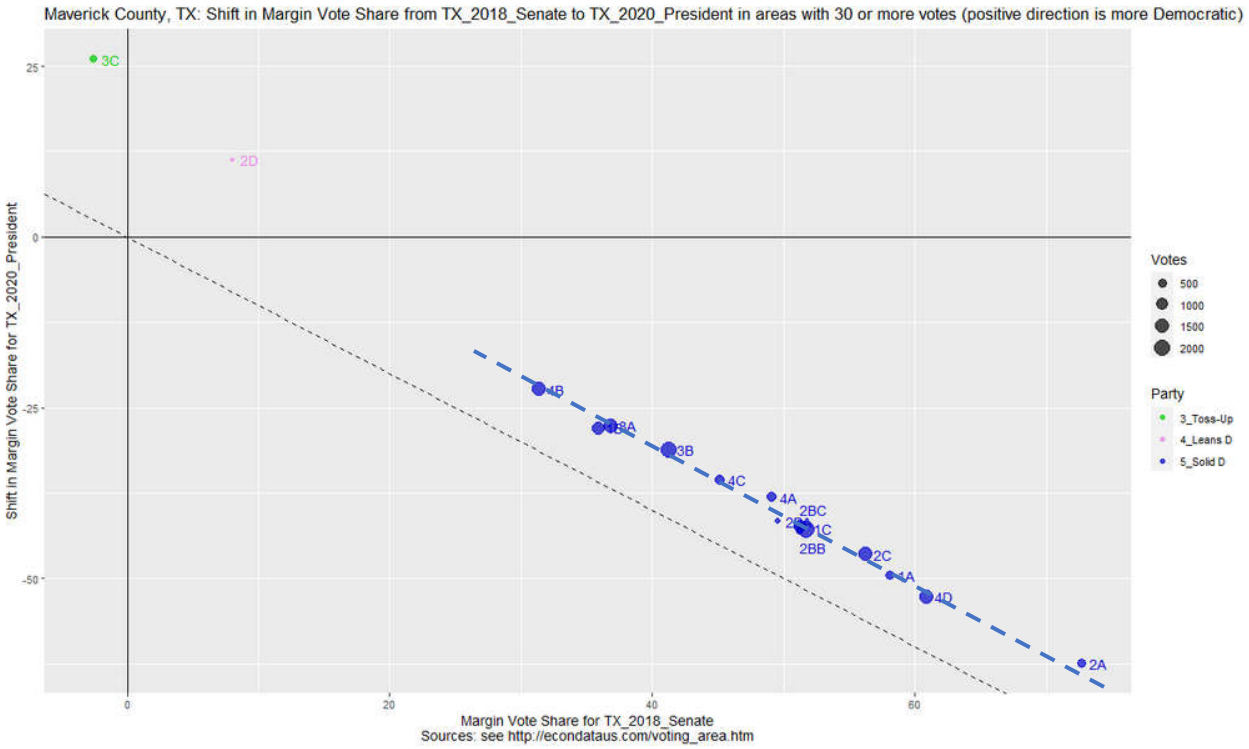
We are all too aware that various putative red flags presented as of great import in the course of Trump's "Stop The Steal" campaign turned out to have simple and benign explanations. Elections – like most large, multi-layered counting tasks – are inherently "noisy" and not every anomaly is indicative of fraud. Maverick is, as noted, a tiny county – shifting *all* its votes would barely make a dent in a statewide election (though it might have an impact in a U.S. House election and certainly could have an impact on a competitive state legislative election). We are viewing it here as a potential iceberg tip – indicative perhaps of a broader pattern with heavier impact. It warrants clear explanation: the ballots, if preserved, should serve to replace speculation with knowledge. There are 253 other counties in Texas and over 3,400 counties nationwide; given the resources, similar precinct-level drilldowns and comparisons could be undertaken elsewhere and the results provide guidance for where the search for "hard" evidence should be focused.

Cameron County, Texas

Here is a slightly different way of plotting shift from E2018 to E2020, more suitable for larger counties with a greater number of precincts. Cameron County, at Texas's southern tip, is, like Maverick, a blue county serviced by ES&S. The chart shown here plots the E2018 U.S. Senate Democratic margin (x-axis) against the shift in margin from that contest to the E2020 presidential contest (y-axis) – an x/dx (x against change-in-x) scatterplot of precincts (all the blue dots in the bottom right quadrant are Democratic precincts that became less Democratic in 2020).



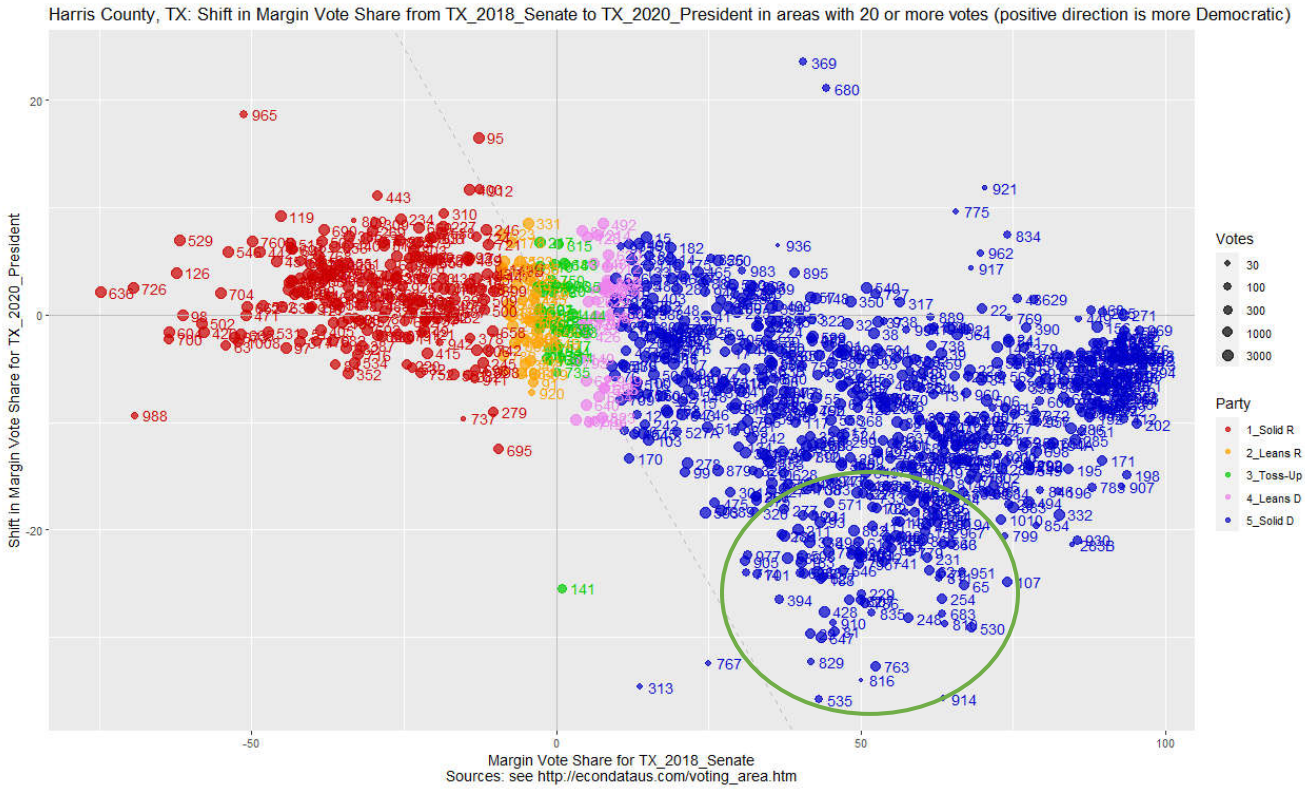
Here is the equivalent chart from Maverick, using the original, certified data:



As can be seen, the near-linear x/dx distribution in Maverick equates to the first Maverick chart shown above, where the E2020 distribution is near-vertical and seemingly independent of the diagonal E2018 distribution. When assessing Cameron, therefore, we recognize that the closer to linear the distribution of precincts is, the more it would translate to the suspect Maverick vertical pattern. We do in fact see a relatively low overall dispersion among the Solid Democratic precincts in Cameron but perhaps more significant are the near-parallel lines within that overall distribution: each of these lines would translate to a "Maverick" distribution (with E2020 near vertical). As with Maverick, it is only possible to speculate what, if anything, might be afoot, whether manipulation of targeted sets of precincts might be occurring. But, as with Maverick, this distribution may serve as a pointer or "scout film" for further investigation.

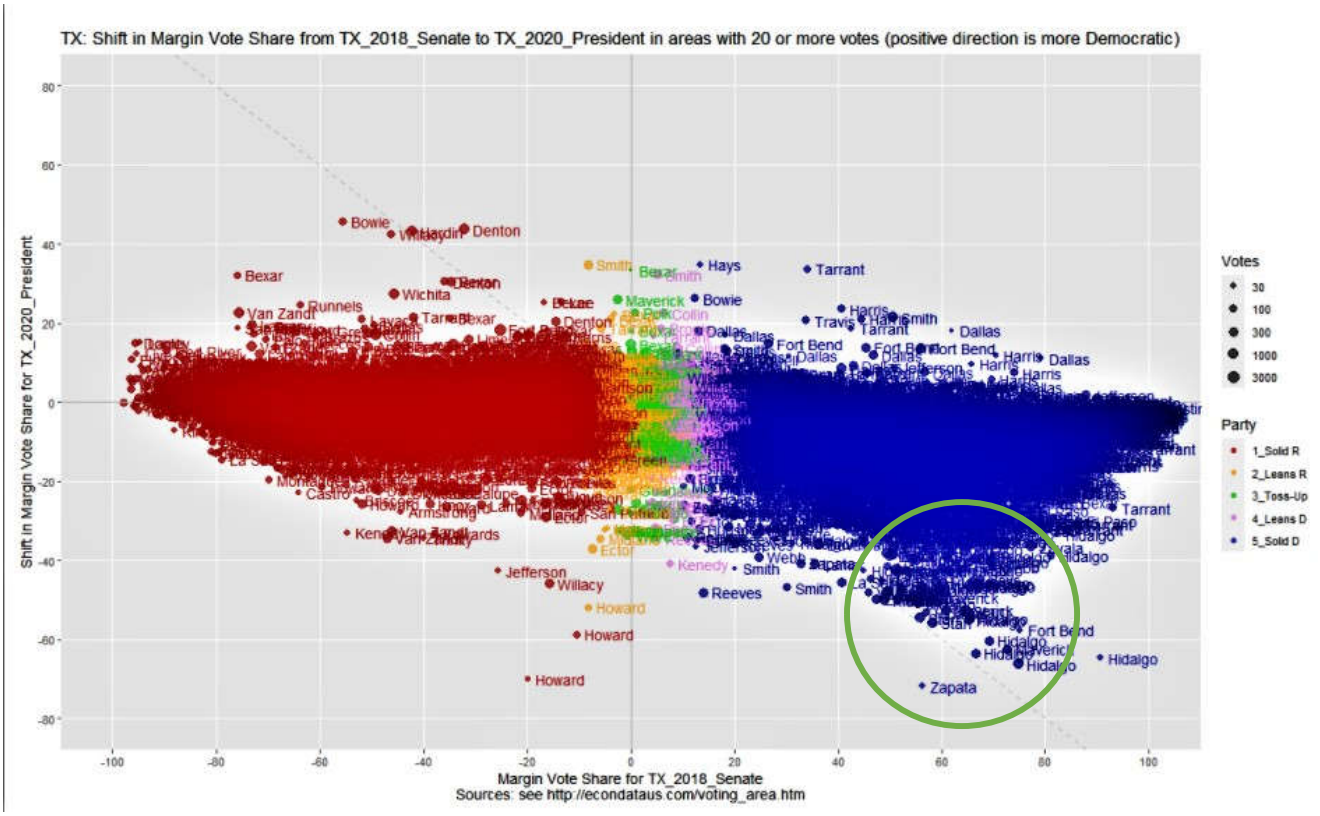
Harris County, Texas

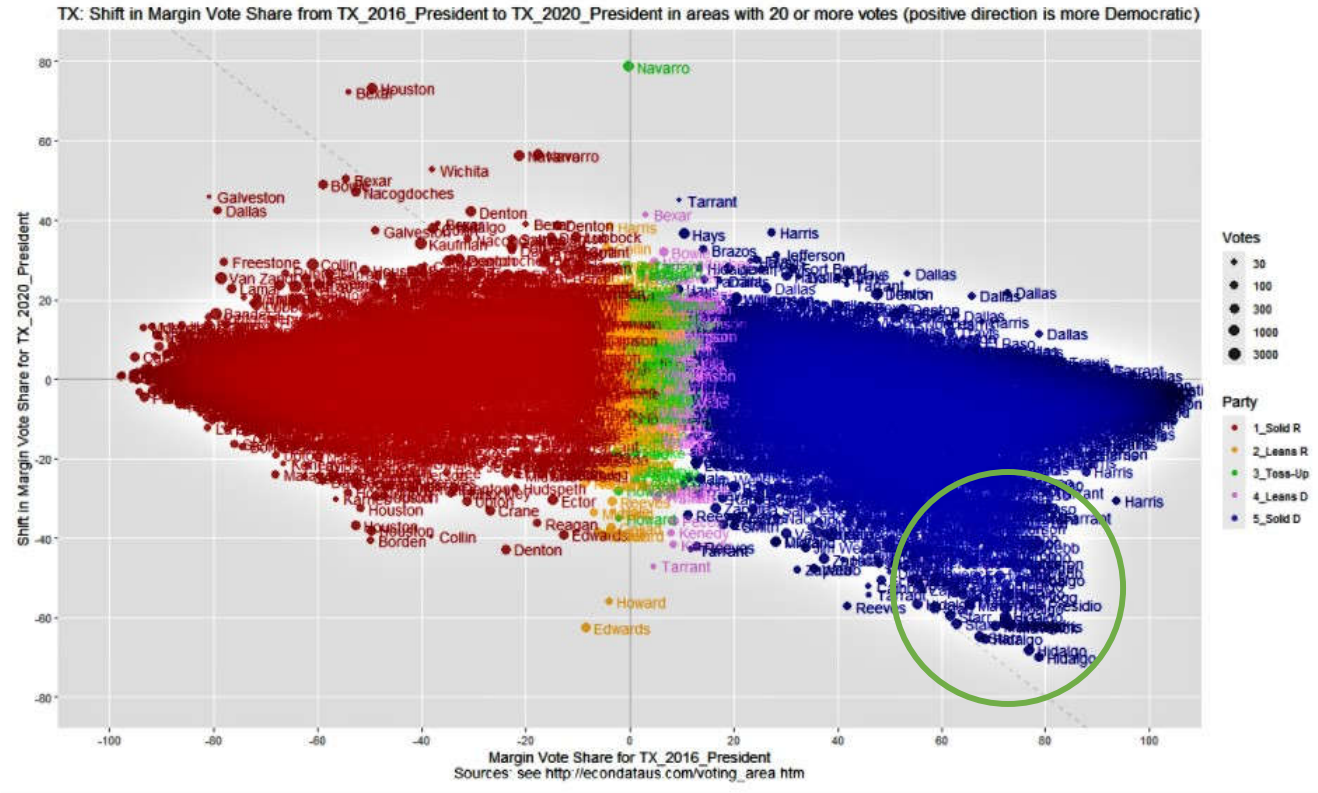
Greater Houston (Harris County, serviced by Hart Intercivic and using Ballot-Marking Devices for all in-person voters) presents a different pattern, but one that also begs explanation and one that we have seen replicated in certain other large red-state counties. The chart below presents, as did Cameron, an x/dx plot, E2018 Democratic margin on the x-axis and change in that margin in E2020 on the y-axis (shift from Democratic to Republican on all these graphs is downward, -y):



We have dubbed this the "Cuba" scatterplot for its distinctive shape: horizontal to the left (those are the Solid Republican and competitive precincts) and V-shaped to the right (the Solid Democratic precincts, in blue). To interpret, there is little or no net shift from E2018 (Senate) to E2020 (presidential) among the Republican and competitive precincts in Harris (to the extent there is any, its mean is above the x-axis; i.e., modestly toward the *Democrats* in E2020), but a strong shift in the Solid-D precincts (much more Republican in E2020). What is most noteworthy about this shift is that it all but spares the large cluster of *strongest* Democratic precincts (at extreme right of the plot) but takes a huge bite out of the vast swath of strong Democratic precincts clustered around x=50, or a 75%D/25%R margin in 2018. While it is possible that demographic/racial factors might account for this unexpected pattern, it is also possible that super-Solid Democratic precincts in the 95%D/5%R range would be regarded as off-limits for manipulation, as any substantial departure from historical patterns of near-100% Democratic voteshare would be immediately suspect, while the precincts toward the middle of the Solid-D block (in the 75%D/25%R neighborhood) would present a trove of suppressible and/or shiftable would-be Democratic votes without the heightened risk of detection associated with the near-100% Democratic precinct group.

Because the Harris County distribution pattern is one we found replicated in other large counties and in other states (for corresponding plots, see www.econdataus.com) , if it indeed turns out to be signaling skullduggery, the impact on election outcomes bearing national significance would be substantial, potentially seismic (it would also corroborate the signal from the SolidR/SolidD scatterplot we examined on page 10 above). On page 20 below are the statewide Texas scatterplots for first the E2018-E2020 shift and then the E2016-E2020 shift; in both we can observe echoes of the Harris County "Cuba" distribution. It may be useful to recall here Republican Texas Attorney General Ken Paxton's June 4, 2021 boast that Trump would have *lost* in Texas in 2020 if Paxton's office had not successfully *blocked* counties from mailing out applications for mail-in ballots to all registered voters. Trump took Texas by over 600,000 votes, so that is quite a boast – and quite a shift.


TX: Shift in Margin Vote Share from TX_2018_Senate to TX_2020_President in areas with 20 or more votes (positive direction is more Democratic)
Shift in Margin Vote Share for TX_2020_President
Margin Vote Share for TX_2018_Senate
Sources: see http://econdataus.com/voting_area.htm
Votes
30
100
300
1000
3000
Party
1_Solid R
2_Leans R
3_Toss-Up
4_Leans D
5_Solid D


TX: Shift in Margin Vote Share from TX_2016_President to TX_2020_President in areas with 20 or more votes (positive direction is more Democratic)
Shift in Margin Vote Share for TX_2020_President
Margin Vote Share for TX_2016_President
Sources: see http://econdataus.com/voting_area.htm
Votes
30
100
300
1000
3000
Party
1_Solid R
2_Leans R
3_Toss-Up
4_Leans D
5_Solid D

Comparable patterns have emerged from precinct-level surveys in Florida, Iowa, Maine, Arizona, South Carolina and North Carolina. All associated graphs are available for inspection at www.econdataus.com, produced by my colleague Reed Davis, whose programming work has greatly facilitated the capture and plotting of data and expedited the process of statistical forensics as applied to US elections.

Conclusion

The tables, graphs, and distributions presented here fall well short of proof that the 2020 election, or any other election, has been subject to covert manipulation. While there is little doubt that elections are the highest-stakes game, in which we may *expect* vulnerabilities and vectors to be exploited for partisan gain; while there is little doubt such vulnerabilities and vectors *exist*; and while there is little doubt that the statistical patterns presented here (and elsewhere) raise *serious questions* about the extent to which our elections, including the most recent, have been subject to such exploits, *proving* that such is the case will ultimately depend on collection and examination of "hard" evidence such as hand-marked ballots (and possibly digital ballot images), chain of custody preserved. That must be the next phase of any dispositive forensic undertaking.

We have witnessed (e.g., Arizona's Maricopa County "audit," which some have sought to template and export to other states) the dangers inherent in shoddy forensics, driven by little more than partisan zeal and the belief that defeat *cannot* be legitimate and *must* be the product of malfeasance. I do not condone such efforts and indeed regard them as dangerously destabilizing.

At the same time, to abandon all efforts to verify concealed, computerized votecounts – including efforts, such as that presented here with acknowledged limitations, guided by scrupulous adherence to the available data and objective analysis of the patterns that emerge from it – would be to foolishly confer absolute credibility on a vote recording and tabulation process that falls far short of having earned it.

Until eligible voters seeking to cast their ballots stop facing targeted impediments to so doing, and until the ballots cast are counted and/or audited publicly and observably, the need for the kind of forensic examination undertaken here will persist. We do not expect this work to result in reversal of any certified outcomes – this is, it should be clear, no "Stop The Steal" effort. But if this preliminary examination can guide deeper examinations and ultimately point the way to, and underline the urgency of, essential reforms in the security and transparency of our electoral processes, it will have succeeded in fulfilling its principal purpose.

This paper was originally published by WhoWhatWhy.com on August 23, 2021. It has been reformatted, expanded and revised to reflect updated information obtained post-publication.

Jonathan Simon is the former executive director of Election Defense Alliance (2006-16) and author of "CODE RED: Computerized Elections and the War on American Democracy." He may be contacted at 831-266-5111, verifiedvote2004@aol.com, or through his website, https://www.CodeRed2020.com.