**Fallacies in Statistically-Based Claims about Massive Election Fraud in 2020:**

**A Compendium\***

Bernard Grofman

University of California, Irvine

[bgrofman@uci.edu](mailto:bgrofman@uci.edu)

https://orcid.org/0000-0002-2801-3351

Jonathan Cervas

Carnegie Mellon University

[cervas@cmu.edu](mailto:cervas@cmu.edu)

https://orcid.org/0000-0001-9686-6308

\*This research was partially supported by the Peltason Chair of Democracy Studies, University of California, Irvine

**ABSTRACT**

The amount that has been uttered/written about the assertions of massive electoral fraud in the 2020 presidential election is immense. Here, we examine claims of fraud based on aggregate election data. We limit ourselves to claims in which the data itself is essentially undisputed, but where the implications of that data for the presence of massive fraud have been largely or entirely misinterpreted through invalid statistical or logical reasoning. Our goal is not to provide new insights, since virtually all the points we make have been made by others, but rather to put together in one place a compendium of recent glaring misuses of statistics that we believe will be useful both as a teaching tool and as a warning about sloppy use of statistics. We deliberately discuss the fallacies in a non-technical way so as to make our critiques broadly accessible to non-specialist audiences.

I. Introduction

Years after the election, tens of millions of voters, including a clear majority of Republicans and a non-trivial number of independents and Democrats believe that there was massive fraud in 2020. Almost all of those who believe in massive fraud also believe that this fraud denied former President Trump his legitimate election victory (Gardner, 2021; Solender 2020 ).[[1]](#footnote-1) For anyone who is interested in the uses of statistics and statistical reasoning in political debate the claim of massive election fraud in the 2020 presidential election is a fertile ground for exploration. And considering the enormous stakes for the maintenance of democracy, misinformation surrounding election fraud has broad implications.

There are many reasons why voters hold to their belief about massive fraud in 2020.Despite the implausibility of a multi-state conspiracy of such magnitude, limited to several background states and in particular counties which delivered overwhelmingly to President Biden, the volume and variety of such claims make them like a hydra-headed monster almost impossible to successfully rebut.[[2]](#footnote-2) The fact that many of these claims about massive fraud in 2020 (including most of those we discuss in this essay) are plausible on their face, even though fallacious, make it harder to refute any individual claim sufficently to satisfy those who are in disbelief. Exploring why voters believe what they do is not the purpose of this essay.[[3]](#footnote-3) Similarly, we do not wish to argue what state or federal courts should or should not have decided about election law. And our list of statistical fallacies committed by those who alleged massive fraud in the 2020 election in no ways claims to be an exhaustive one.[[4]](#footnote-4)Indeed, to the contrary,we are not going to discuss the vast bulk of claims made about fraud. Our concern here is a narrowly focused one. We deal solely with claims about fraud that are grounded, at least in part, on indisputable facts about statistical features of the 2020 presidential election.

Our goal is not to provide new insights about these statistical fallacies, since the particular fallacies we give have already been rebutted by others. Rather, our goal is put together in one place a useful compendium of many of the most glaring recent misuses of statistical reasoning as applied to understanding elections,and to do so in a way that is readily accessible to non-technical readers.[[5]](#footnote-5)We believe strongly that a discussion of statistical fallacies based on real-world examples should be part of any statistics or public policy curriculum. Moreover, the salience of the topic and misinformation availalble has not subsided over time. Rather, politicians continue to push claims, some candidates even preemptively claiming fraud, nearly a year before a single vote is cast (Corasaniti et al, 2022), and a major political science journal has acceped a peer-reviewered scholarly article with claims of statistical evidence (Lott, forthcoming, *Public Choice*; though, Eggers and Grimmer (n.d) have swiftly and effectively shown flaws in Lott’s methods) for massive fraud in Georgia and Pennsylvania.

We begin our inventory of fallacies with (a) arithmetic fallacies of a simple sort, such as drawing conclusions from unweighted averages where use of weighted average was required, cherry-picking the data to emphasize only those facts that lead to the desired conclusion, and confusing percentages and percentage point changes Then we discuss (b) improper use of statistical significance, and then turn to (c) inaccurate probabilistic reasoning, such as improperly using as an indicia of fraud having voters with the same name and date of birth. Then we discuss (d) syllogistic arguments based on cross-election statistical comparisons that are either fallacious in form, or that have at least one premise that is indubitably false, and thus which give rise either to invalid or unfounded conclusions. Finally, we briefly consider similar types of statistical errors in (e) syllogistic arguments based on within-election comparisons. We believe we have a typology that includes the most common errors involving analysis of aggregate election data.

II. Fallacious Statistically Based Claims about Massive Fraud in the 2020 Presidential Election

A. Arithmetic Fallacies

Failing to weight units

Failing to recognize that large changes in one direction in low population subsets or low population demographic subgroups can be compensated for by small changes in the other direction in high population subsets or demographic units is a common mistake. There were various instances of this types of error in discussions of the 2020 election. We first focus on the most common instance.

It was observed that Trump won more counties in 2020 than he did in 2016, with the implication being that he must have done better in 2020 than in 2016 in terms of the popular vote.**[[6]](#footnote-6)** But, of course, that is nonsense, since he could have done better in the remaining counties. These counties, though fewer in number, had more voters in them. Indeed, Biden received over three million votes in Los Angeles County, alone. In fact, Biden netted an additional 609,000 more votes in 2020 than Clinton did in 2016, just in this one county! Out of the over 3,000 counties in the United States, the top 150 contained half of the total votes casted. Biden won 125 of those 150 (83.3%).

Errors in using unweighted averages when it is appropriate to use a weighted average can be plotted in several ways. In **Figure 1A**, we plot the Democratic vote share of each county as a histogram. There’s a strong skew left, indicating that Trump did well in a far larger number of counties. However, in **Figure 1B**, we properly weight each county to match the number of voters in each county. Now, there is a distribution that has a more normal curve, slightly favoring Biden (which makes sense, since he won the popular vote by 7 million votes). The same error can be made when looking at a choropleth map of election results by county (**Figure 2A**). As Chief Justice Earl Warren famously quipped in *Reynolds v. Sims* (1964), “legislators represent people, not trees or acres”. Trump certainly won more acres, but he did not win more voters. f you look at map of U.S. counties showing those won by President Trump in red and those by Hillary Clinton (or Joe Biden) in blue, you will see a sea of red and only a relative handful of pockets of blues. But those pockets (mostly big cities) have lots of voters in them. When counties are instead represented on a map in proportion to their votes, such a with a cartogram or bubble map (**Figure 2B**) -- a map that has been resized so the units’ area is equal to its population weight – it become clear that the election was close, but certainly not an overwhelmingly “red” county shown by a county-level election map.[[7]](#footnote-7) However, even this weighted maps fail to identify the winner of the election, since counties are shown as either won by Trump or won by Biden, not the magnitude of victory.[[8]](#footnote-8) I

<<Figure 1 about here>>

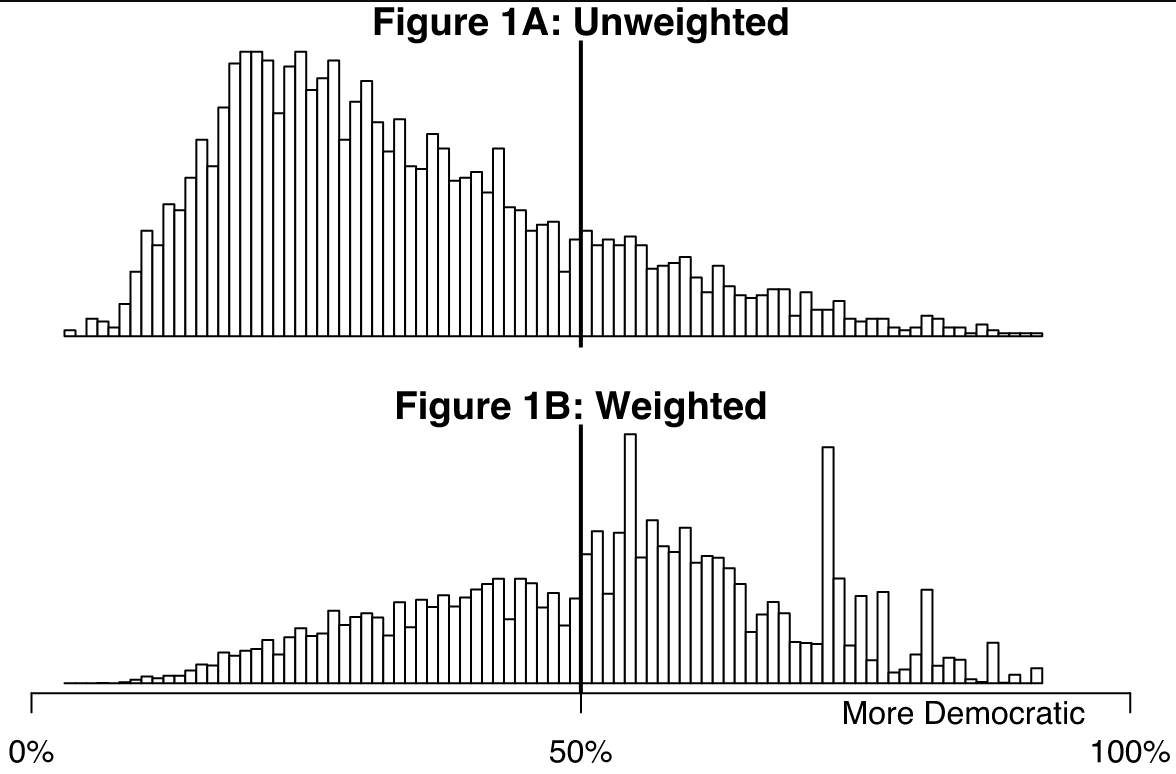
2020 Presidential Election Results, by county

Figure 2A - Choropleth Plot, 2020 Presidential Election by county

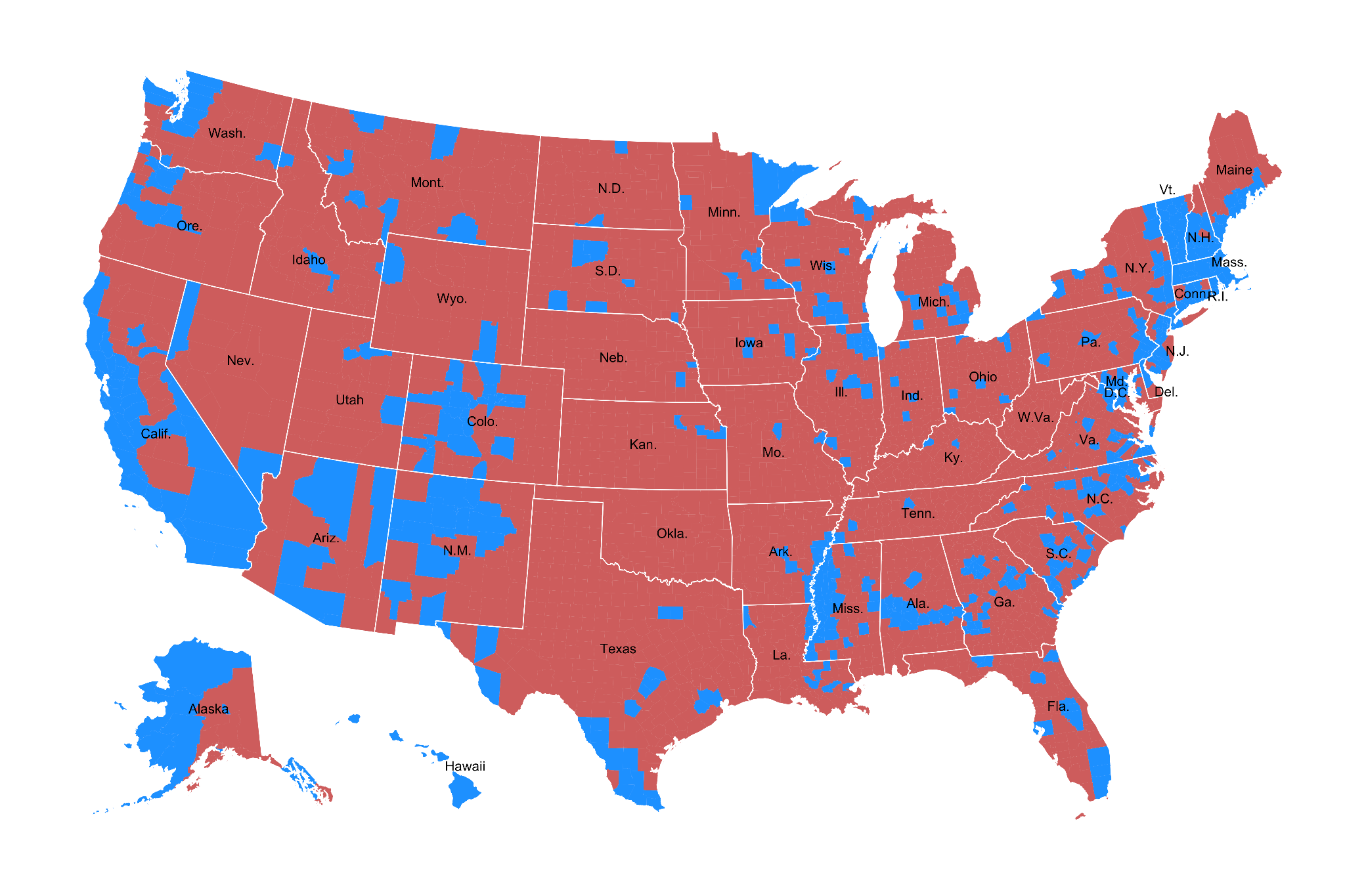
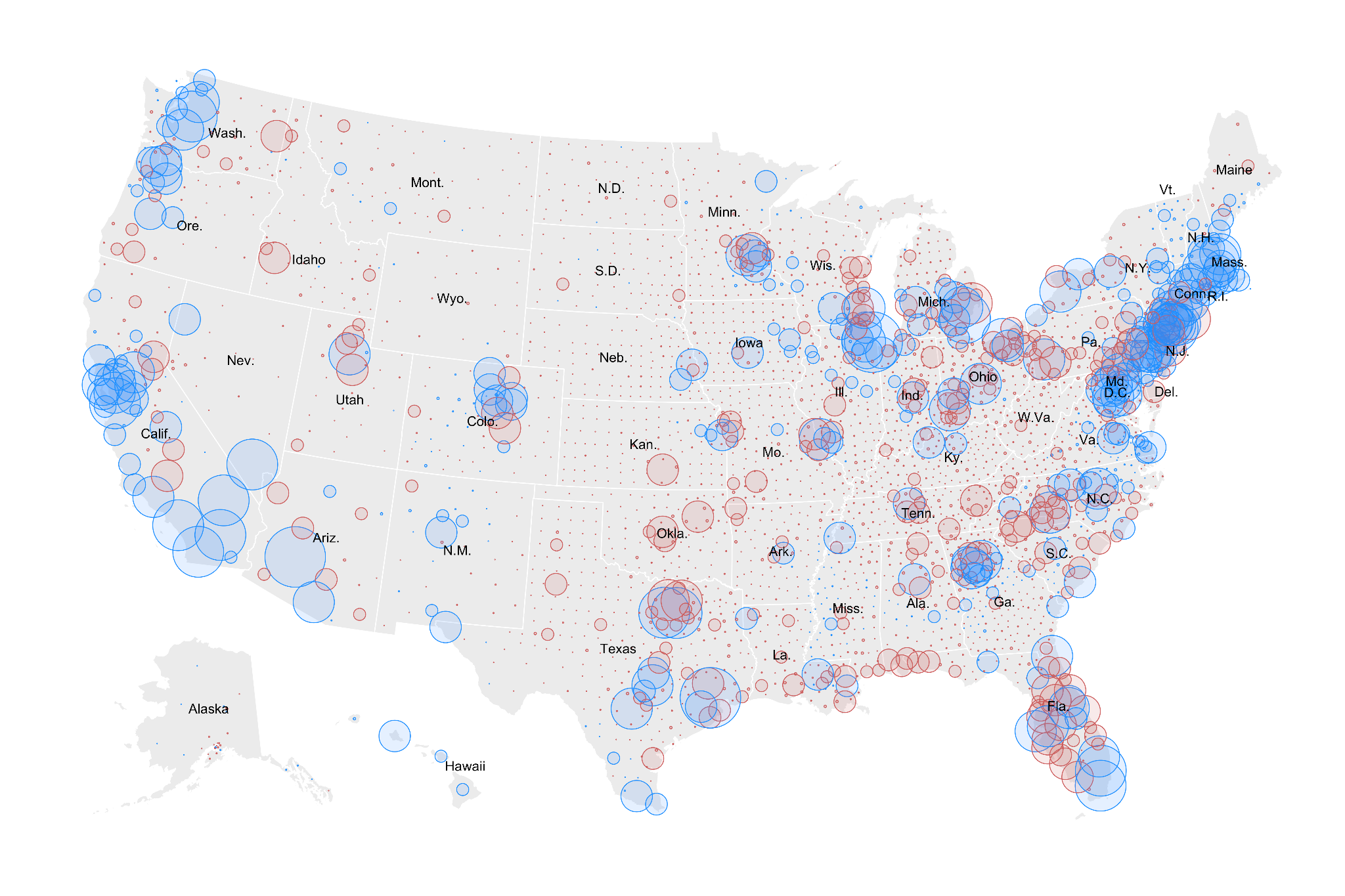


Figure 2B - Bubble Plot, 2020 Presidential Election by county



Note: GIS from NYT; County-level presidential results from MIT Election Data and Science Lab, 2018, "County Presidential Election Returns 2000-2020",<https://doi.org/10.7910/DVN/VOQCHQ>. For an even better visual display of county level election returns using force-directed graphs, see NYT: https://www.nytimes.com/interactive/2020/11/03/us/elections/results-president.html

Failure to weight units also lead some commentors to focus on Trump’s gains among minority voters. Trump did show a substantial percentage point gain among minorities as compared to 2016. He won 12% of the Black vote in 2020, compared to 8% in 2016. Likewise, he won 32% of the Latino vote in 2020, an increase of 4-percentage points from 2016. He even won 1-percentage point more of the White vote in 2020 compared to 2016.[[9]](#footnote-9)It might seem that these facts imply that Biden must have done worse in popular vote share than Clinton did in 2016. But there is one critical data point that suggests otherwise. While Trump did 1-percentage point better among White voters, Biden did 4-percentage points better among this same demographic! In 2020, 67% of the electorate identified as White in the exit polls, so increasing the margin among that category of voters will swamp decreases among other demographics..The actual combination of percentage point losses among minority voters and percentage point gains among non-minority voters, gave rise to an election in which Biden won81,268,924 (51.3%) of the popular vote (i.e., total vote) in 2020 compared to Clinton’s vote in 2016 of65,853,514 (48.2%).[[10]](#footnote-10) Moreover, the increased total turnout further increased Biden’s margin of victory over Trump. Trump lost to Clinton by 2,868,686 million votes in 2016; he lost to Biden by 7,052,770 million votes.**[[11]](#footnote-11)** Moreoever, what really matters is *where* these changes occur. Since presidential elections are actually 51 separate state-wide (or at least determined by states in the electoral college), winning many more votes in states where the plurality is already secured creates “wasted” votes, since they do not contrbute to winning any new electors. Taken all together, it is clear that Biden was the benefactor of changes from 2016 that propelled him to the presidency.

Cherry-picking the data

A more primitive form of failing to weight the data properly is cherry-picking the data to emphasize only those facts that lead to the desired conclusion. In employing only the partial story, a claimant can appear tohave been honest while suppressing pertinent information that otherwise would prove their claims either false, or incomplete.**[[12]](#footnote-12)** Because the data being cited is accurate, cherry-picking can prove a persuasive tool. One example of cherry-picking the data about the 2020 election is to highlight the situations where Biden did worse than Clinton, to suggest that it was impossible for Biden to have done better against Trump than did the previous Democratic candidate and, since Trump beat Clinton, absent fraud, Trump must also have beaten Biden.

For example, it is true that, in terms of percentage voting for the Democratic candidate, racialminority support for Biden was (marginally) lower than for Hillary Clinton. However, support among categories of white voters, namely the college educated and those living in suburbs, was higher for Biden than for Hillary Clinton. Similarly, there were some urban areas where support for Biden was lower than for Hillary Clinton but there were also urban areas where support for Biden was higher than for Hillary Clinton. It is, of course, the combination of all the subgroup patterns of voting and the pattern of their geographic location as it affects Electoral College outcomes that matters. Looking only at some subsets of voters, or only some geographic areas, e.g., only rural counties, is misleading and can lead to ridiculous claims that the candidate who received more votes did not actually receive more votes. Moreover, as we show below, decreases (increases) in the *percentage* of votes from given group or area can be offset by increases (decreases) in the *number* of raw votes in terms of impact on the difference in raw votes between two candidates.

Changes in vote percentage vs. changes in vote totals

President Trump’s percentage of support among minority voters increased in 2020 from what it had been in 2016. It might seem obviousthat, whatever happened overall, President Trump should therefore have done better in 2020 than in 2016 among this group of voters in terms of reducing the gap between his total votes from this group and those received by his opponent. But that need not be true! We need always to take into account raw vote totals. **In particular, total turnout rose in 2020 as compared to 2016 among both minority and non-minority voters -- but the increase was greater among non-minority voters, so that the proportion of the electorate that was minority fell in 2020 from what it had been in 2016. JONATHAN IS THIS CORRECT?**

**Below we offer a very simple example in which President’s Trump percentage of the minority vote increased, and yet his overall disadvantage relative to his Democratic opponent actually increased among minorities in terms of total votes received. To understand what is happening we need to look at changes in turnout. The toy example below matches many of the main qualitative features of the 2016 and 2020 elections, but we simplify by keeping turnout and vote share among non-minority voters fixed, whereas in reality Trump did less well among non-minority voters in 2020 than he did in 2016.** [[13]](#footnote-13)

**Assume Trump in 2016 received 30% of the votes from minority voters in 2016 and Clinton 70%. In 2020, let us posit an increase in the minority turnout of say 5%, Imagine further that, in 2020, President Trump received 32% of the minority votes to Biden’s 68%. Thus, in 2020, Trump would have received a higher share of the minority vote than he did in 2016; President’s Trump’s vote tally from minorities rises two percentage points. However, as can also be readily calculated, among minority voters, the net vote margin for the Democratic candidate *increases* by 1.4 percent (=.68\*1.05 - .7\*1) in the second election as compared to the first. In other words, in this toy example, once we take changes in turnout into account, despite the increase in President’s Trump’s percentage of the minority vote from 2016 to 2020, he actually did worse in 2020 than in 2016 in terms of the gap between his total votes from minorities and those received by his Democratic opponent.**

**More generally, if V was the previous vote percentage for the Democratic candidate among the members of the group, and we normalize, T, the percentage turnout among the group’s members in the previous election, to be 1, while we set d as the percentage point decline in Democratic support from that group in the subsequent election and t as the percentage point increase (or decrease) in turnout among that group’s members in the subsequent election, than the Democratic gain or loss from the group’s voting between the two elections is**

**(V - d)\*(T + t) – VT = Vt -d(1 + t)**

**When t (V- d) > d, this value is positive. The belief that change in percentages tell us everything we need to know about the popular vote is an easy mistake to make, and to further confuse us, changes from previous years may sometimes be given as percentage point changes, and sometimes as percentage changes. Of course, if d is negative, i.e., there was a gain in Democratic vote share in the group, and if there is also a rise in turnout in that group, then the raw votes for Democrats from the group must be higher in the latter year.**

**JONATHAN, CAN YOU REPLACE THE TOY EXAMPLE WITH REAL DATA IF THE REAL DATA GIVES RISE TO THE SAME SEEMING PARADOX and also can you please doublecheck that I did the algebra correctly?**

B. Misinterpreting Statistical Significance

In statistics, warning about misinterpreting statistical significance are legion. One point frequently made is that statistical significance can only be interpreted in the context of the specific null hypothesis being tested. Another is a reminder that p values are very strongly linked to sample sizes. A third point is a warning not to confuse statistical significance with substantive importance. A further note of caution is that high statistical significance in regressions does not imply a causal relationship between variables.

Our first example of misuse of statistical significance fall into the second and third categories. Our second example involves misinterpretation of causality. Both are found in the expert witness testimony of Dr. Charles Cicchetti in the lawsuit brought by Texas Attorney General Ken Paxton challenging election results in Georgia, Michigan, Pennsylvania, and Wisconsin (*Texas v. Pennsylvania*, 592 U.S. \_\_\_, 2020).**[[14]](#footnote-14)** Dr. Charles Cicchetti’s calculations were picked up and repeated in the blogosphere.

First, in comparisons between what was found in 2020 and what was found in 2016 , Dr. Cicchetti noted that he could demonstrate beyond any possibility of error that the vote share distribution for Joe Biden in 2020 differed from that of Hillary Clinton to a statistically improbable degree. He is certainly right about that fact.

But what that shows about election fraud is — exactly nothing! What it shows is that the vote-share distribution for Donald Trump in 2016 was not the same as in 2020. But we could have picked ANY two adjacent presidential election years and showed that the vote shares of the Democratic candidates in those two elections at the level of states or counties or precincts differed from one another. With a large enough sample size, such comparisons are virtually certain to generate statistically significant differences.

Second, in a comparison of partisan voting patterns in ballots in 2020 among ballots that are tallied early in the counting process and those that are tallied later, Dr. Cicchetti pointed out that, in some states, President Trump’s share of the vote declined relative to those first reported as polls closed on election night as more ballots were tabulated, while in other states he found the reverse pattern. He found the difference between the early vote-share for Trump and later vote-share for share to be statistically significant beyond any reasonable doubt, though of course, the directionality of the difference was not uniform. And about those facts and the statistical significance of the observed differences he is again quite correct. But what that shows about election fraud is — once again, exactly nothing! Indeed, for those familiar with elections it shows a pattern that was predicted in advance (Foley and Stewart 2020).

**While once it was believed that mail in ballots are disproportionately Republican, for the last two decades or so the evidence has been clear that mail-in-ballots are more Democratic leaning than is the case for ballots cast in-person. In some states in 2020, mail-in ballots were tallied as they came in; thus, they were the first ballot results to be reported; in other states, the tallying of mail-in ballots did not begin after the polls were closed and they were often the last to be tallied. Thus we would expect that there would be differences across states in early vote versus later vote partisan shares that would be in part predictable by when mail-in ballots were tallied, and this could account for why, in some states, the early vote was more Democratic than the later vote while the reverse was true in other states.** In fact in **most** **JONATHAN, CAN WE BE MORE PRECISE?** states where the late-tallied ballots were disproportionately mail-in ballots, they were disproportionately more Democratic than ballots reported earlier. In states where the late-tallied ballots were disproportionately in person ballots, they were disproportionately Republican.

But there is also another reason why late ballot tallies and early ballot tallies might be expected to be different in their partisan outcomes. Ballots tallies come in from counties and then get tallied statewide. When ballots get reported depends upon the efficiency of the county officials doing the tallying. But it is often easier to complete in-person voting tallies in small rural counties where there are few polling stations than in counties with large cities. But rural counties differ from urban counties in their partisan propensities. So, it is essentially guaranteed that there will be differences in partisan vote share as county reports come in at different times on election night.

In sum, when the null hypothesis is a silly one, there is nothing at all surprising about it being rejected at a high level of statistical significance, especially when sample size is high. Indeed, what would be surprising is that such silly null hypotheses would NOT be rejected.

C. Meretricious Probabilistic Reasoning

Voters with the Same Name and Date and Year of Birth

**It has been noted that Joe Frazier voted in Pittsburgh in 2020?? even though he had been dead for a while. This fact was taken as evidence of fraud (CITE NEEDED). But of course, not all Joe Fraziers are the famous (and now dead) boxer. Similarly, instances where individuals with the same name and same birthday and/or same birth year were found on voter rolls have been taken as evidence of fraud. But the likelihoodof occurrence of voters appearing on voting lists who have identical names and even identical birth dates and even identical year of birth is dramatically underestimated by virtually everybody, except for those who have truly understood the famous “Birthday Paradox.” The Birthday Paradox in its basic form is about how the number of people in a group affects the probability that some pair of them have the same birthday.**

**When first confronted with the question of how the number of people in a group affects the probability that some pair of them have the same birthday, students in statistics classes (or at least those in the undergraduate statistics classes taught by one of the present authors) think that, since any given birthday has a probability of only 1/365 (or 366 if you take into account leap years), it would take around 183 people before you would get a probability of 50% that two people in the group had the same birthday. The Birthday Paradox is called a paradox because, actually, it only takes a group of size 23 for the probability of two people in the group to have the same birthday to exceed 50%.**

**The probability that two people in a group of size n do NOT share the same birthday can be written as 1 \*(1- 1/365) \*(1- 2/365)\* (1- 3/365)\*… \*(1- n/365). This product goes down much faster than one might think, and thus the probability that at least two people in the group share the same birthday, which is one minus this product, goes up much faster than one might think.** [[15]](#footnote-15) **Indeed, with only 75 people the probability of a birthday match rises to 99.95%. In a similar way one can calculate the expected proportion of a group of size n who share a birthday with at least one other person in the group.**

**The difficulty in appreciating how the increasing number of possible pairs that could share a birthday (and other attributes) increases in a non-linear way with increasing n is very relevant to the claim made in 2020 (and earlier) that finding examples of people with the same name and same date and same year of birth on the voting rolls was *proof* of “double-voting” fraud.**[[16]](#footnote-16) **While, for any given name, it is obviously harder to find someone with both the same birthday and the same year of birth as it is merely to find people with the same name and same birthday, the logic of figuring out the probability of such a happening is basically the same as the simple product formula given above, but now the divisor is no longer 365 since we need to take into account the years in which a person might have been born, though we may reasonably assume that everyone who votes was born at least 18 years ago.** **Moreover, as a further complication, we must take into account the degree of heterogeneity in the distribution of names.** [[17]](#footnote-17) **But if we take name, birthday, and birth year as mutually independent factors,** [[18]](#footnote-18) **then we can simply multiply probabilities.** [[19]](#footnote-19) **And we can further simplify by assuming a uniform distribution across the first two factors and assess the likelihood of two randomly chosen individuals bearing the same name from the name distribution in empirical data.**[[20]](#footnote-20)

**Of course, multiplying probabilities for three different factors gives us low probability values, but not as low as one might think. For example, if a randomly chosen person has a 0.000074 percent chance of sharing both a first and a last name with the next randomly chosen person, which was the estimate from the McDonald and Levitt (2008: p. 119, fn. 26)) study,**[[21]](#footnote-21) **a total electorate of only 21,071 is enough to bring the probability of finding two people with the same name and the same birthday and the same birth year above 50%. And an n of 57, 314 brings that probability to 99.5%. Moreover, as McDonald and Levitt (2008) demonstrate with their detailed analyses of New Jersey electoral rolls, there are other reasons why we see individuals shown with identical data identifiers, including flaws in the data such as the same individual simply being entered twice, or individuals with missing data birth data (assigned a particular missing data code) being treated as having identical records.**

Benford’s Law

**There have been many attempts to use statistical tools to detect election fraud. Some of these involve looking at suspicious data, e.g., vote tallies that disproportionately end in 0’s or 5’s since, in “inventing” data, there can be a human tendency to call these numbers to mind more often that would be expected from a purely uniform distribution** **(*Economist*, 2021). While it makes no sense to look at first digits of election returns since these will be obviously contingent on the mean size of the units,**[[22]](#footnote-22) **in investigating fraud the frequency of digits other than the last or first digit has also been investigated. *Benford’s Law* , which is an hypothesized frequency distribution of kth digits, is an example of one tool that has been used to assess prima facie evidence of voter fraud, with the political scientist, Walter Mebane, having applied this tool to elections in multiple countries (CITES NEEDED) In the 2020 presidential election analyses based on Benford’s Law were provided in videos and tweets as evidence that, in various locales, elections had been rigged.**[[23]](#footnote-23) **We will make no attempt to repeat the logic that leads to *Benford’s Law* (see the discussion of the supposed Law in Wikipedia and references there in**[[24]](#footnote-24)**); we simply note that almost all of those who have investigated it empirically are dubious about its application to elections.**[[25]](#footnote-25)

**According to Mebane (2011), *Benford’s Law* implies that the mean value of the second digit of a distribution of votes at the precinct or other unit level should equal 4.187. But a little reflection will show that the frequency of even second digits will be linked both to the size of units and their variance.**

**Imagine, for example, that the mean Democratic votes cast in a set of precinct is 1650 voters, with a variance of only 200. The second digit will have disproportionately many more 4’s, 5’s, 6’s, 7’s and 8’s than it has other numbers. If we change either the mean or the standard deviation of that distribution, then the expected mean value of the second digit will also change. Thus, there cannot be some fixed value that is the expectation of the mean value of the 2nd digit in all distributions. And the same point applies whichever digit we focus on, or whether we look at all of them, since scale of units will affect which digits are most likely to deviate from the Benford’s Law expectation.**

**But perhaps even more importantly, it is quite possible for Benford’s Law to work well for some candidates and badly for others, since different candidates will have different means and variances.**

**The next two examples of meretricious probabilistic reasoning are more about human psychology than they are about statistics, *per se*.**

Tip of the iceberg fallacy

**One fallacy of probabilistic inference is to infer from the indubitable fact of some electoral wrongdoing**[[26]](#footnote-26) **in most states that electoral fraud is massive in both magnitude and geographic spread. The careful studies of fraud, such as the one of the 2020 election undertaken by the Ohio Secretary of State, reveals fraud at a miniscule level of fraud: 27 cases out of 6 million or so ballots cast in Ohio.**[[27]](#footnote-27) **However, finding such as those of the Ohio Secretary of State, rather than being taken as evidence that the fraud level was trivial, can instead be interpreted as supporting evidence for a high probability of massive fraud.**[[28]](#footnote-28) **The implicit and quite wrong-headed** **probabilistic argument is that any examples of fraud that are found should be presumed to be “only the tip of the iceberg**.”**[[29]](#footnote-29)**

Straw man fallacy

**On the other hand, the claim that there was no massive fraud can be rhetorically equated to the claim that there was no fraud, and the latter claim rebutted as a straw man, as if successfully rebutting the claim of no fraud demonstrated the truth of the claim of massive fraud.**

Now let us turn to a set of other claims that have as their general form: “The only way these election results could have happened is if there was massive fraud.”

D. Logically Valid Arguments with False Statistical Premises involving Historical Election Results Comparisons

We begin with an example of a logically invalid argument.

Spoiled ballots

An example of invalid argument based on comparisons of past and present election results is based on the empirically accurate observation that the spoiled ballot rate of mail-in ballots in 2020 was much lower (in some states remarkably lower) than in 2016. This fact is taken to be evidence of mail-ballot fraud by Trump supporters. Presumably the theory is that the lowered spoilage rate came about either because more ballots were illegally cast in 2020 than in 2016 and that those stealing unfilled in mail ballots and submitting them knew how to properly fill out the forms, or that ballot workers in polling stations where the mail ballots were likely to be Democratic ones deliberately allowed improperly filled out mail ballots to pass muster in a way that they had not in 2016.

We can write the argument as

|  |
| --- |
| If there is ballot fraud involving mail-in ballots (A),  then the spoilage rate among mail-in ballots will be lower than in the past (B).  The spoilage rate among mail-in ballots was lower than in the past (B)  Therefore, there was ballot fraud involving mail-in ballots (A). |

Here we have the fallacy of *affirming the consequent*.

There are good reasons why ballot spoilage was lower in 2020 than in 2016 that have nothing to do with fraud, namely much greater effort on the part of election administrators to inform voters of what they needed to do to cast a valid ballot. For instance, popular late night comedy Stephen Colbert created a rather sophisticated website aimed at informing those in all 50 states about the specifics for casting a ballot in each of those states.**[[30]](#footnote-30)** His “Better Know a Ballot” also aired many times in the months before the elections on his highly rated “The Late Show.” Ads developed by the states themselves aired on television channels and as ads on streaming services.**[[31]](#footnote-31)** The Democratic National Committee also spent millions of dollars on television ads with information about returning mail-in ballots.**[[32]](#footnote-32)** There was also ample coverage in newspapers about properly filing out and mailing a ballot so that it would not be rejected.**[[33]](#footnote-33)**

Moreover, in some states, there were greater efforts to ensure that those who submitted a mailed-in ballot with an envelope which had some correctable error that would prevent the ballot inside the still unopened envelope from being counted were informed of the error and given the opportunity to correct it. Eighteen states allowed voters to “cure” their ballots if there is a discrepancy.**[[34]](#footnote-34)** These states are disproportionately Democratic; Trump won just 5 of the 18. But in our federal system, absent genuine issues of potential voting rights violations, states can and do differ in the details of their election administration.

Now we turn to claims about election fraud that fall into the category of valid arguments with false premises.

Presidential coattails

One claim about election fraud was based on the observation that winning presidential candidates have coattails that aid members of their party in the House of Representatives to gain seats. The Democrats lost 12 seats in the House of Representatives in 2020, which violates this expectation, and so the claim goes that the implication is that there must have been massive multi-state fraud. The structure of this argument is:

|  |
| --- |
| If a presidential candidate wins election (A),  then there will be a gain in the number of members of his party in the U.S. House of Representative (B).  There was no gain for the Democrats in the House in 2020 (not B)  Therefore, Biden must have lost the election (not A) |

This is a valid argument. It is an example of *denying the consequent*. However, the premise on which is built, that presidential coattails are inevitable, is false.

By coattails, we are referring to an increase in the number of members in the U.S. House of Representatives that share the incoming president’s party (Campbell 1986). Actually, negative coattails are not uncommon, and in contemporary politics, have becomemorelikely.**[[35]](#footnote-35)** Since 1868, there have been thirteen elections where a president has had negative coattails (including 2016 and 2020). Negative coattails are more likely when (a) elections are close in popular vote (b) there is substantial partisan bias against the party of the presidential winner in the House, (c) a substantial portion of the votes for the winning presidential candidate are wasted in states that are won by large margins, and (d) there are many more seats won by the previous president in the election four years earlier which have a sitting House member of the other party than there are seats lost by the previous president in the elections four years earlier which have a sitting House member of his party. All four of these features are found in 2020.

Biden’s share of the major party vote was only 52.27%; the estimated partisan bias in 2020 in the House of Representatives in 2020 was 2.7%.**[[36]](#footnote-36)** Congressional districts have become far less competitive in recent elections, leaving fewer chances for a president to provide coattails large enough to flip seats (Engstrom 2020). If we eliminate the states that gave the widest raw margin to Biden (California and New York and Massachusetts) from the calculations, Trump has a majority of the vote in the remaining states – hence, we would not expect to see Biden coattails in those remaining states.**[[37]](#footnote-37)** Democratic gains in the House in the 2018 midterm were significant, and turnout was a level not seen before universal adult franchise (Jacobson 2019). Moreover, up through 2016 there is a time trend of decreasing presidential coattails which, when projected onto 2020, would create an expectation of a negative coattail in the 2020 election.**[[38]](#footnote-38)** But perhaps most importantly, there were 35 House constituencies carried by Trump in 2016 but with a Democratic House member elected in 2018,**[[39]](#footnote-39)** and only 5 House constituencies lost by Trump in 2016 but with a Republican House member elected in 2018.**[[40]](#footnote-40)** Thus, Democrats in 2020 had many more vulnerable House seats than did the Republicans.

Bellwether counties

An argument of a similar form is that Biden lost most of the counties that had been bellwether counties, and therefore, since bellwether counties predict presidential elections, Biden must really have lost the election.

Again, we can write this argument as

|  |
| --- |
| If a presidential candidate wins the election (A)  then they can be expected to carry almost all the bellwether counties (B),  Biden lost almost all of the bellwether counties (not B)  Therefore, Biden must have lost the election (not A) |

This, too, is actually a valid argument -- another example of *denying the consequent*. However, even though “not B” is empirically accurate, the premise on which the argument is built, namely that bellwethers predict elections, is false.

Many decades ago, the political scientist Edward Tufte (1974, chapter 3; 1975) wrote a devastating rebuttal to work on the power of bellwethers. Tufte showed that, over the period 1916-1968, there were no real state-level bellwethers and, most importantly, the U.S. counties identified as presidential bellwethers at time had no better track record at the next presidential election than the non-bellwether counties.**Hopkins (2017) showed the same result for much more recent data.** **Yet, belief in bellwether units of geography, more particularly in the existence of bellwether counties, refuses to die.**

**Grofman and Chen explain the predictive failures of bellwethers partly in terms of classic work of Deutsch and Madow (1961) which uses a simple binomial model to show how, by chance alone, in large groups, some individuals can appear to have repeated (predictive) success even though, for any given event, the probability of success of any actor is only .5, and we are dealing with independent trials where past success tells us exactly nothing about future success.**[[41]](#footnote-41) **But Grofman and Chen also elaborate on the conditional probability model that generates the likelihood of bellwethers by showing that as partisan polarization increases, but presidential politics nationally is competitive in the Electoral College so that sometimes we have Democrats winning and sometimes Republican winning (three each in the 21st century), the likelihood of bellwethers declines.**

**The intuition here -- which they examine at the level of counties -- is a very simple one: in order to be a bellwether county, a county must vote for the winner both when the winner is a Democrat and when the winner is a Republican, but *ceteris paribus*, the increasingly polarized patterns of voting have by now put virtually all counties firmly into one partisan camp or the other. In the 21st century , Grofman and Chen (2020) show that more than 70%+ of all counties vote consistently Republican, and another portion are more or less permanently Democrat. These counties make great bellwethers only when their party is always winning, but awful bellwethers otherwise.**[[42]](#footnote-42)

Other Cross-Election Comparisons

Other arguments with specious premises supposedly demonstrating that Biden could not have won in 2020 also make use of differences between the 2020 election and patterns found in previous elections. For example, Trump supporters such as Shurk (2020) noted that no incumbent who has won more than 75% of the primary vote has lost their election. Therefore, since Trump had won 94% of the primary vote, he must have won re-election. **[[43]](#footnote-43)** Shurk also observes that “no incumbent in over 100 years who has gained votes in his reelection bid has lost his quest for reelection.**[[44]](#footnote-44)** On the other hand, we could just as easily to claim that incumbents whose overall presidential approval rating was as low as that of President Trump must lose re-election.Unfortunately for this type of argument, there is no guarantee that what was true in the past will be true in the future. And, of course, these historical comparisons are based on small *n*, especially when we limit our comparisons to those where there is an incumbent running for re-election. **[[45]](#footnote-45)**

E. Logically Valid Arguments with False Statistical Premises Using Comparisons Based on Features or Components of the Same Presidential Election.

Matching design: within election split ticket voting versus straight ticket voting

**One matching comparison involves comparison of elections where there are two kinds of votes: straight ticket votes (where a single check casts a vote for all candidates of a given party for all offices) and ballots where this box is not checked. The claim by Ayyudurai (2020) is that in four counties in Michigan where the straight ticket vote option was available, the difference between the vote share of former President Trump in the straight ticket votes and in the split ticket votes should be unrelated to the level of success in President Trump in straight ticket votes unless there is voter fraud. Moreover, Ayyudurai asserts that a negative correlation indicates that the split ticket votes have been manipulated to reduce the apparent Trump share of the vote, i.e., that a negative correlation is a sign of voter fraud.**

**We can write this argument as**

If the difference between straight ticket vote share for the candidate of a given party and split ticket vote share for the candidate of that party is negatively correlated (A),

then there is vote fraud in favor of the other party (B).

The difference between straight ticket vote Republican share and split ticket Republican vote share is four Michigan counties is negatively correlated with straight ticket Republican vote share (A).

Therefore, there must have been vote fraud favoring the Democrats in those counties (B).

**This is, in principle, a valid argument. But the premise makes no sense. The notion that the two vote shares should be uncorrelated if there is nothing suspicious going on might appear on its face to be plausible, but some reflection reveals that it is completely misguided. Indeed depending upon whether we take the difference as** **split ticket vote share minus straight ticket vote share or as straight ticket vote share minus split ticket vote share) it should be inevitable to get a non-zero correlation. Below we work out a toy example.**

**If we take as an approximation that Republican vote propensities are the same among straight ticket voters as among split-ticket vote, then we are correlating r1s with r1s-r2(1-s), or with r2(1 - s) – r1s, where s is the share of the Republican voters who cast a straight ticket vote and ri is the Republican share of the vote cast by voters of the given type, with i =1 straight ticket voters and i=2, split ticket voters. Let us simplify further by positing that r1 = r2. Now, doing what Ayyudurai (2020) wishes us to require simply correlating r with either 2r-1 or 1-2r, thus guaranteeing either a positive or a negative correlation depending upon which way we do the difference. What there is NO reason to expect is a zero correlation.**

**But, if the argument above did not convince you that Ayyudurai’s (2020) syllogism is misguided, just think about the failure of symmetry. Under our simplifying assumptions, the exact same argument he makes for Republicans goes through for Democrat votes as well, so that his argument leads us, on the one hand, to the conclusion that there was fraud favoring Democrats (when we look at Republican vote shares) and on the other hand, to the contradictory conclusion that there was fraud favoring Republicans (when we look at Democratic vote shares).** [[46]](#footnote-46) **JONATHAN CAN YOU PLEASE VERIFY WHAT THE MICHIGAN EMPIRICS REALLY WORK FOR BOTH DEM AND REP VOTE SHARE. I THINK I HAVE THIS RIGHT BUT I AM NOT 100% SURE.**

Matching design: within-election comparisons of areas with and without fraud claims

**An apparently sophisticated attempt to prove election fraud via statistical analysis is Lott ( 2020), which is now forthcoming in a peer-reviewed journal (Lott, 2022). While Lott offers various forms of analysis, the most ingenious one, which appears to be quite plausible,**[1] **is a matched pairs analysis which compares the difference in support for Trump in adjacent precincts in pairs of counties in Georgia and pairs of counties in Pennsylvania, such that one of each pair is in a county which voted for Biden and one is in a county which voted for Trump.**[2] **Some precinct comparisons involved pairs of precincts which voted for the same party’s presidential candidates, and others are ones where the two precincts voted for different presidential candidates. The county which voted for Biden is treated as one in which fraud might be present; the country which voted for Trump is treated as a baseline in which no fraud would be expected. Lott compares the difference in absentee votes for Trump between adjacent pairs of precincts in the two types of counties with the difference in in-person votes for Trump in these same pairs of precincts in a regression with a dummy variable for precincts located in the county which voted for Biden. He claims that a positive sign on this dummy variable indicated the presence of fraud. He obtains positive but not statistically significant results for both Georgia precinct pairs and Pennsylvania precinct pairs and then combines the data for both states to find a positive and statistically significant result on his dummy variable for the pooled data. Lott (2020) also argues that adjacent Republican counties should not show this effect.**

**Eggers, Garro and Grimmer (2021) provide a variety of arguments, both theoretical and empirical, to show that the econometric designs in the Lott (2020) study are so flawed that its conclusions are completely vitiated. That study has the advantage of making use of the data provided by Lott for replication purposes**[3]**. Here we simply make two points about conceptual flaws in the Lott analyses.**

**On the one hand, it is a huge leap to believe that a positive sign on the dummy variable for precincts in pro-Biden counties in his analyses can be taken as evidence of fraud. There can be many reasons that have nothing to do with fraud for why Trump vote shares among mail-in voters as composed to in-person voter vary across precincts. For example, while we do expect the area on the border between the two paired counties are similar, in Pennsylvania where we are most familiar with the data, there are key differences, such as income levels, homeownership levels and tax rates between the two paired counties that can be expected to affect voting choices.**[4] **These differences are unlikely to be captured by control variables such as age and gender, and can lead to differences in mean levels of Trump voting that can have major consequences for how the expected shape of in-person versus mail-in ballots across the two counties. JONATHAN PLEASE ELABORATE**

**More particularly, implicit in the Lott regression model is that, in all precincts, both in counties where Trump wins and in counties where Trump loses, the share of the Trump vote that is cast in-person is independent (at least on average) of the level of Trump support in the precinct. In other words, Trump in person-vote share is k times Trump total vote share; while Trump mail-in vote share is (1 - k) times Trump total vote share, where 0 ≤ k ≤ 1. If this assumption holds, then we can write the in-person vote difference between the precincts in the pro-Trump county (County 1) and those in the pro-Biden county (County 2) as k\*(T1-T2), and we can write the in-person vote difference between the precincts in the pro-Trump county (County 1) and those in the pro-Biden county (County 2) as (1- k)\*(T1-T2). Hence we can form a regression with**

**Trump absentee vote share in County 1 minus Trump absentee vote share in County 2 as the y variable**

**and**

**Trump in-person vote share in County 1 minus Trump in-person vote share in County 2**

**as the x variable.**

**But, in accord with the homogeneity assumption above, this leads us to the regression**

**y -= [(1 -k)/k] x.**

**In this regression, because of the posited linearity, we should find an intercept of zero. Lott interprets a non-zero intercept as evidence of fraud but, as we have said ,that is a huge leap.**

**But any non-uniformity in the posited linkages can produce a non-zero intercept. Below we demonstrate this non-uniformity for precincts in XX county and YY county in Pennsylvania, two of the counties analyzed by Lott. JONATHAN PLEASE PROVIDE A GRAPH OR SOME OTHER KIND OF DATA IF YOU CAN Indeed, what we find is that the higher the Republican vote share in the precinct, the higher the proportion of Republicans who cast in-person votes. We interpret this as a result of the more highly Republican areas having been more effectively cautioned by party leaders against the ue of mail-in ballots. JONATHAN is this vlaim empirically CORRECT??!**

**On the other hand, even if we reject the various rebuttals of the appropriateness of Lott’s matched pairs design, and take as given that the sign of the dummy variable tells us something about fraud, the evidence that the dummy variable he uses as an indicator of fraud is non-zero is not that strong. As noted earlier, Lott obtains statistically significant results for the dummy variable only for data pooled across two very different states. But that pooling can , in principle create a fallacy akin to *Simpson’s paradox*. If the two states differ, say, in their partisan propensities in the chosen counties, then an effect (e.g., a positive sign on the dummy variable) that is actually due to differences between the states can be attributed to within-county differences in each state.**[5]

[1] **Indeed, we regard Lott’s analysis as sufficiently plausible , on its face, to explain why, despite being either completely wrong or at best much overstated, it could nonetheless pass peer review.**

[2] **Lott (202) also offers a similar kind of argument about why turnout differences in adjacent pro-Biden versus pro-Trump counties could also be indicators of fraud. But that argument also is fundamentally flawed because it assigns to fraud differences across units that have other very plausible non-fraudulent explanations. See Eggers, Garro and Grimmer (2021)**

[3] **One point Eggers, Garro and Grimmer (2021) make is that the results we find for Republican vs. Republican county comparisons depend upon the order in which we select the units; another point they make is that the exactly which precincts are paired can reverse the sign of the results and that Lott’s specification is one of those that results in a higher level of statistical significance for the dummy variable. JONATHAN PLEASE THINK ABOUT. IS THIS CORRECT?**

[4]  **One of the present authors grew up in Pennsylvania in one of the precincts in Lott’s data set and is familiar not just with the county in which that precinct is located but also familiar with the adjacent county whose precincts are used for comparison purposes.**

[5] **Grofman (1989) looks at an analogous issue. Pooling data across nations he finds that minimum-winning cabinet coalitions are considerably more durable, on average, than non-minimum-winning cabinet coalitions, i.e., it takes longer until the coalition is dissolved. And yet, within each country, he finds that minimum winning coalitions are no more durable than non-minimum winning coalitions. Rather, he finds that countries with more durable coalitions also have a higher than average proportion of minimal winning coalitions. ADD TO REFERENCES Grofman, Bernard. 1989. The comparative analysis of coalition formation and duration: Distinguishing between‑country and within‑country effects. British Journal of Political Science, 19:291‑302**

III. Conclusions

To paraphrase Jeremy Bentham, claims of massive fraud based on aggregate level statistical features of the 2020 election are not just nonsense, but “nonsense on stilts.” While it is impossible to address all the misleading claims and specious arguments made in the blogosphere or by President Trump himself, we believe the compendium of statistical fallacies given above can be useful to those interested in the misuse of statistics. And, as noted earlier, this essay is deliberately written in a non-technical way to be comprehensible even to beginning students in statistics.

References

Ayyadurai, S. “Dr. Shiva live: MIT PhD analysis of Michigan votes reveals unfortunate truth of US voting systems” (video recording, 2020). . <https://www.youtube.com/watch?v=Ztu5Y5obWPk> Accessed February 12, 2022.

Bentham, Jeremy. 2002. Rights, Representation, and Reform: Nonsense upon Stilts and Other Writings on the French Revolution. In *The Collected Works of Jeremy Bentham* (edited by  Schofield, P., C. Pease-Watkin,  and C., Blamires).  Oxford University Press, 317–401.

Cervas, Jonathan R, and Bernard Grofman. 2019. Are Presidential Inversions Inevitable? Comparing Eight Counterfactual Rules for Electing the US President. *Social Sciences Quarterly* 100(4): 1322-1335.

Deckert, Joseph; Mikhail Myagkov, and Peter C. Ordeshook. 2011*.*["Benford's Law and the Detection of Election Fraud"](https://www.cambridge.org/core/journals/political-analysis/article/benfords-law-and-the-detection-of-election-fraud/3B1D64E822371C461AF3C61CE91AAF6D)*.* *Political Analysis*  19 (3): 245–268*.*[*doi*](https://en.wikipedia.org/wiki/Doi_(identifier))*:*[*10.1093/pan/mpr014*](https://doi.org/10.1093%2Fpan%2Fmpr014)*.*[*ISSN*](https://en.wikipedia.org/wiki/ISSN_(identifier))[*1047-1987*](https://www.worldcat.org/issn/1047-1987)

Deutsch, K. W., and W. G. Madow. 1961. “A Note on the Appearance of Wisdom in Large Bureaucratic Organizations.” *Behavioral Science* 6 (January): 72–78.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. Harper & Row.

*Economist*, The. 2021. “Russian Elections Once Again had a Suspiciously Neat Result.”[*Russian elections once again had a suspiciously neat result | The Economist*](https://www.economist.com/graphic-detail/2021/10/11/russian-elections-once-again-had-a-suspiciously-neat-result)

Eggers, Andrew C., Haritz Garro and Justin Grimmer. 2021. “Comment on ‘A Simple Test for the Extent of Voter Fraud with Absentee Ballots in the 2020 Presidential Election’." *Proceedings of the National Academy of Science*. 118(45). [10.1073/pnas.2103619118](https://dx.doi.org/10.1073%2Fpnas.2103619118)

Engstrom, Erik J. 2020. “Congressional Elections: Electoral Structure and Political Representation.” In *New Directions in American Politics*, eds. J L Carson and M S Lynch. Taylor & Francis. <https://books.google.com/books?id=-CnZDwAAQBAJ>.

Fiorina, Morris P. 2017. *Unstable Majorities*. Stanford: Hoover Institution Press.

Foley, E. B. 2013. “A Big Blue Shift: Measuring an Asymmetrically Increasing Margin of Litigation The Voting Wars: Elections and the Law from Registration to Inauguration.” *Journal of Law & Politics*, *28*(4), 501–546. <https://heinonline.org/HOL/P?h=hein.journals/jlp28&i=517>

Foley, E. B., and Stewart III, C. 2020. “Explaining the Blue Shift in Election Canvassing.” <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3547734>

Fryer, Roland G Jr. and Steven D. Levitt. 2004. “The Causes and Consequences  
of Distinctively Black Names.” *Quarterly Journal of Economics* 119 (3): 767-805.

Gardner, Amy. 2021. “In Ramp-Up to 2022 Midterms Republican Candidates Center Pitches on Trump’s False Election Claims.” *Washington Post*, July 5, 2021.

Gelman, Andrew. 2021. “Alex Jones and the Fallacy of the One-sided Bet.” <https://statmodeling.stat.columbia.edu/2021/08/21/alex-jones-and-the-fallacy-of-the-one-sided-bet/>

Golbeck, Jen. 2020. Benford’s Law Does not Prove Fraud in the 2020 U.. Presidential Election.” https://twitter.com/jengolbeck/status/1325166129321537536

Grofman, Bernard and Haotian Chen. 2022 forthcoming. “Understanding the Factors that Affect the Incidence of Bellwether Counties: A Conditional Probability Model.” *Political Research Quarterly.*

Grofman, Bernard and Jennifer Garcia. 2014. “Using Spanish Surname to Estimate Hispanic Voting Population in Voting Rights Litigation: A Model of Context Effects Using Bayes Theorem.” *Election Law Journal* 13(3): 375-393

Grofman, Bernard and Jennifer Garcia 2015. “Using Spanish Surname Ratios to Estimate the Proportion Hispanic in California Cities via Bayes Theorem.“ *Social Sciences Quarterly* 85: 1511-1527.

Holman, Mirya R and J. Celeste Lay, 2019. “They See Dead People (Voting): Correcting Misperceptions about Voter Fraud in the 2016 U.S. Presidential Election.” *Journal of Political Marketing* 18 (Issues 1-2).

Hopkins, David A. 2017. *Red Fighting Blue: How Geography and Electoral Rules Polarize American Politics*. Cambridge University Press.

Jacobson, Gary C. 2019. “Extreme Referendum: Donald Trump and the 2018 Midterm Elections.” *Political Science Quarterly* 134(1): 9–38. https://onlinelibrary.wiley.com/doi/abs/10.1002/polq.12866 (January 14, 2020).

Kabir, N. 2020. “The fraud of Dr. Shiva Ayyadurai: Oakland county, Michigan.” [The Fraud Of Dr. Shiva Ayyadurai: Oakland County, Michigan | by Naim Kabir | Medium](https://naim-kabir.medium.com/the-fraud-of-dr-shiva-ayyadurai-oakland-county-michigan-1bc51bcebf1b) Accessed February 12, 2022.

Lee, Frances E. 2016. *Insecure Majorities: Congress and the Perpetual Campaign*. Chicago: University of Chicago Press.

Lott, John R. 2021. “A Simple Test for the Extent of Voter Fraud with Absentee Ballots in the 2020 Presidential Election: Georgia and Pennsylvania Data," Unpublished Manuscript. https://papers.ssrn.com/sol3/papers.cfm?abstract id=3756988.

Mebane, Walter M. 2011. Comment on “Benford’s Law and the Detection of Election Fraud.”

*Political Analysis* 19: 269–272 .

Mebane, Walter M. 2020. “Inappropriate Applications of Benford’s Law Regularities to Some Data from the 2020 Presidential Election in the United States” [inapB.pdf (umich.edu)](http://www-personal.umich.edu/~wmebane/inapB.pdf) November 10, 2020. Accessed February 12, 2022.

Parker, M. “Do these scatter plots reveal fraudulent vote-switching in Michigan?” (video recording, 2020). <https://www.youtube.com/watch?v=aokNwKx7gM8>. Accessed 12 February 2022

Shurk, J.B., “5 More Ways Joe Biden Magically Outperformed Election Norms”, *The Federalist*, November 23, 2020, <https://thefederalist.com/2020/11/23/5-more-ways-joe-biden-magically-outperformed-election-norms/>.

Tufte, Edward R. 1974. *Data Analysis for Politics and Policy*. Prentice-Hall.

Zorn, Eric. January 5, 2017. “Polls Reveal Sobering Extent of Nation’s Fact Crisis.” *Chicago Tribune*, http://www.chicagotribune.com/news/opinion/zorn/ct-polling-ignorance-facts-trump-zorn-perspec-0106-md-20170105-column.html

1. This is hardly surprising when viewed in the the context of Gabrial Lenz’s work on how the public change its policy views to match the politicians they support (Lenz, 2012). [↑](#footnote-ref-1)
2. There have been multiple supposed indicators of election fraud introduced, including personal affidavits alleging fraud in particular precincts, to videos allegedly showing direct evidence of vote tampering, to videos showing how easy it supposedly would have been to manipulate the record of votes produced by voting machines, to claims about a conspiracy by a particular voting machine vendor (for a devastating rebuttal of that particular claim see Eggers, Garro and Grimmer, 2021), to assertions (e.g., Ayyudurai, 2020) that more voters voted than were on the jurisdiction’s electoral roll (subsequently rebutted by showing that the list of eligible voters used for the comparison was incomplete; <https://www.bbc.com/news/election-us-2020-55016029>), to the kind of statistical evidence we consider in this article. [↑](#footnote-ref-2)
3. A number of authors have tried to understand why belief in massive fraud in persists (see e.g., Holman and Lay, 2019; Berlinski et al, 2021), with most arguing that the key factor is a siloization of communication channels. Political psychologist could inform this further by tying these beliefs to well-known cognitive biases like confirmation and desirability biases. The fraud claims are echoed as indisputable by a multiplicity of sources that particular voters trust. Those same sources denigrate the reliability of the mainstream media who refute the fraud claim and insist that the mainstream media are simply partisan mouthpieces for the Democrats. Also, we have the mesmerizing power of repetition. The claim of massive fraud in 2020 is stated again and again in conservative media sources and by former President Trump and his allies. And in some of these repetitions, the way the claim is stated suggests that denying massive fraud in 2020 is analogous to denying the reality of the Holocaust. Moreover, in the four years before the election, the claim that the 2020 election would be stolen by the Democrats was also repeatedly asserted by President Trump and his supporters, thus preparing the way for the post-election claims of fraud. But it was also impossible not to be constantly reminded about the fraud claims regardless of which media sources one followed. Virtually every claim made in the conservative press alleging fraud was given considerable space in the mainstream media, followed by a rebuttal that often went on longer than the statement of the claim. And there were follow-up pieces in the mainstream press repeating multiple claims of fraud and their rebuttals. Also (in the *Washington Post*) President Trump had previously been given tens of thousands of Pinocchio points based on other kinds of claims found untruthful, with points added every time the claim was repeated (Kessler and Fox 2021). Here, relevant in psychological terms, is the aphorism that “the lady doth protest too much,” in which repeated assertions that a given allegation is untrue become taken as support for a belief that it probably is true. [↑](#footnote-ref-3)
4. The Internet is rife with claims about fraud in the 2020 election. [↑](#footnote-ref-4)
5. We would call particular attention to the excellent survey and analysis of 2020 presidential election fallacies by Eggers, Garro and Grimmer (2021). We began writing our own essay before we were familiar with this article in its on-line form, but we have learned a lot from it and cite to it herein. Their detailed rebuttal of fallacies is more technical than ours, and we cover some fallacies that they do not.There are also various journalistic surveys of election fraud claims (Corasaniti et al, 2020; Feldman, 2020; Alba and Frenkel, 2021), but these primarily include claims that do not fall within the scope of this essay. As far as we are aware, none present the statistical claims have been discussed within a unified framework. [↑](#footnote-ref-5)
6. See Swenson, Ali, “Winning more counties doesn’t translate to an election win for Trump,” December 21, 2021, *Associated Press,* <https://apnews.com/article/fact-checking-afs:Content:9848943909>. [↑](#footnote-ref-6)
7. For a nice explanation of displaying political graphics of different kinds, see Bliss, Laura & Marie Patino, “How to Spot Misleading Election Maps,” November 3, 2020, *Bloomberg CityLab*, <https://www.bloomberg.com/news/articles/2020-11-03/a-complete-guide-to-misleading-election-maps>. See also “Cartographic Views of the 2020 US Presidential Election, November 27, 2020, *World Mapper*, https://worldmapper.org/us-presidential-election-2020/. [↑](#footnote-ref-7)
8. To best illustrate this point, it is imporant for the reader to understand that the county with the most Republican votes anywhere in the USA was Los Angeles County, in California. Trump received 1,145,530 votes, which was dwarfed by Biden’s 3,028,885. In fact, Trump’s 8 of the 10 highest county vote totals come in states won by Biden.. [↑](#footnote-ref-8)
9. Data from exit polls: <https://www.cnn.com/election/2016/results/exit-polls>; https://www.cnn.com/election/2020/exit-polls/president/national-results [↑](#footnote-ref-9)
10. ​​Clinton won a plurality of the vote in 2016, not a majority. Several third-party candidates combined had vote totals surpassing the margin between Trump and Clinton.

    <https://www.fec.gov/resources/cms-content/documents/2020presgeresults.pdf>

    https://www.fec.gov/resources/cms-content/documents/federalelections2016.pdf [↑](#footnote-ref-10)
11. There other types of statistical confusions related to the Electoral College that we simply mention in passing because they are not directly linked to issues of fraud. Many voters have trouble understanding how the Electoral College works. It can most simply be thought of an example of weighted voting, where the weights are the seats each states is allocated, namely the sum of the number of representatives the state has in the U.S. House plus two, the number of U.S. Senators from the state. A common fallacy is to believe that a reversal of the popular vote outcome in the Electoral College occurs simply because of the two-seat “federal bonus” favoring small states. But Donald Trump would have won the Electoral College in 2016 even without the two seat bonus. **The vote in an Electoral College of 438 members in 2016 would have been JONATHAN FILL IN XX.x% as compared to the actual EC percentage of JONATHAN FILL IN pp.p%. ; while Biden’s percentage an Electoral College of 438 members in 2020 would only one have gone up x.x percentage points from yy.y to zz.x JONATHAN FILL IN .**

    Because the Electoral College is a state-based weighted voting game (except for the two states where seats are allocated at the congressional district level as well), once a unit is won by a candidate, the margin of victory is irrelevant for Electoral College seat share but is relevant in terms of the popular vote share. Thus, we can have a candidate winning some states by very close margins in such a fashion that there is a discrepancy between the popular vote winner and Electoral College winner. Of the four instances post-Civil War examples of such a discrepancy, two of the four, 1888 and 2016, would still have happened without the bonus. But, in 2000, without the bonus, Al Gore would have been elected president, and in 1876 there also would have been no inversion had the two seat bonus been eliminated. (Cervas and Grofman, 2019, Table 2, pp. 1328-29). Of course we must be careful in asserting counterfactuals. For example, had there been no two-seat bonus in 2000 Gore and Bush might have deployed campaign resources differently. [↑](#footnote-ref-11)
12. This is a strategy employed often by those who know that the full evidence is against them. [↑](#footnote-ref-12)
13. **Because in our toy example it is also posited that turnout arose among non-minority voters in this toy example the gap between President Trump and his Democratic opponent is 44 votes in 2020 as compared to 40 in 2016. The gap between the Democratic and the Republican candidate widens even among non-minority voters whose percentage share for the two candidates is fixed because turnout has increased, but now it is President Trump who is the gainer. So, in our toy example, Trump did better in the popular vote in 2020 than he did in 2016 since he lost 1 vote among minorities and gained 4 votes from non-minorities. In our toy example, President Trump’s percentage share of the total vote in 2016 was 49.5% (=(520+25)/1100) and 49.7% (=(572+27)/1205) in 2020. Of course, President Trump did NOT actually do better in percentage vote share terms in 2020 than he did in 2016! Our toy example, for simplicity of exposition, deliberately keeps the non-minority share of the vote for the two candidates fixed, since we wished to focus on what happens re minority voting. In reality, President Trump did less well among non-minority voters in 2020 than he did in 2016. Biden ran marginally better among non-minority voters in 2000 then Clinton had in 2016 largely due to a surge in support in the suburbs and among the more highly educated which compensated for losses in the rural areas and among the less well -educated. Jonathan IS what IS IN BLUE CORRECT?**  [↑](#footnote-ref-13)
14. As we noted earlier we are providing a compendium; this and other claims of Dr. Cicchetti have already been rebutted elsewhere (see e.g.,  <https://reason.com/volokh/2020/12/09/more-on-statistical-stupidity-at-scotus/>  and <https://statmodeling.stat.columbia.edu/2020/12/08/the-p-value-is-4-76x10%E2%88%92264-1-in-a-quadrillion/>, That report has been devastatingly critiqued in the expert witness report of Gary King in the same case. Claims about election fraud in *Texas v. Pennsylvania,* including Dr. Cicchetti’s report, can be found here: <https://www.supremecourt.gov/DocketPDF/22/22O155/163048/20201208132827887_TX-v-State-ExpedMot%202020-12-07%20FINAL.pdf> [↑](#footnote-ref-14)
15. **This product can be rewritten in terms of combination or permutations, which gives a somewhat more intuitive way to see how the probability changes with n. See** [**https://en.wikipedia.org/wiki/Birthday\_problem#Calculating\_the\_probability**](https://en.wikipedia.org/wiki/Birthday_problem#Calculating_the_probability) **A program that calculate the probability of the birthday paradox for any given n, is found at** [**https://www.statisticshowto.com/same-birthday-odds/**](https://www.statisticshowto.com/same-birthday-odds/) [↑](#footnote-ref-15)
16. **JONATHAN, CITES NEEDED FOR CLAIM OF doubLE VOTING IN 2020** [↑](#footnote-ref-16)
17. **Calculating this probability from the name distribution in the population requires calculating a self-weighted average based on name and surname. Names vary greatly in their frequency. The U.S. Bureau of the Census** **publishes a list of surname frequencies based on national values, but the distribution of names (and their degree of heterogeneity) will vary with the racial/ethnic composition of the political unit (Grofman and Garcia, 2014, 2015).**  [↑](#footnote-ref-17)
18. **Birth year and name are not fully independent of one another. Changing patterns of immigration affect the relative surname shares in different generations, and some rather dramatic changes over time in popularity of first names mean that first name probabilities are not independent of year of birth. Also first names and last names are far from independent, since both are linked to ethnicity. Indeed we could, in principle, estimate birth decade probabilities just by looking at the prevalence of first and last names of those born during the decade (cf. Grofman and Garcia, 2014, 2015).**  [↑](#footnote-ref-18)
19. **Note that the probability of finding such a match from among a set of n voters must not be confused with the probability of any particular name + birthday + birthyear combination being repeated. For obvious reasons the likelihood of a given combination being repeated will depend, *ceteris paribus*, on how common is the name.**  [↑](#footnote-ref-19)
20. **The most sophisticated study of double voting of which we are aware, McDonald and Levitt (2008), which investigates complaints of election fraud in New Jersey in 2004 that were based on apparent observance of thousands of instances of double voting, includes a comparison of the assumption of uniform distribution of year of birth and name distribution to actual data and show that assuming a uniform distribution over a 64 year interval and over the names in the data set tends to underestimate the prevalence of birth year matching by about 12% (=(487- 433)/433, see p. 119). The more there are birth bulges (as in the post-WWII baby boom), the more likely it is that two randomly chosen individuals will share the same year of birth, and some names are much more prevalent than others.**  [↑](#footnote-ref-20)
21. **For further compatibility with the McDonald and Levitt (2008) study, in the model presented below we also took our birth year time period as a 64 year span.**  [↑](#footnote-ref-21)
22. **Mebane (2020) notes that “It is widely understood that the first digits of precinct vote counts are not useful for trying to diagnose election frauds.”** **See also Golbeck (2020).** [↑](#footnote-ref-22)
23. **See e.g.,** “**Voter Fraud ! Joe Biden's Votes Violate Benford's Law.”** https://gnews.org/534248/ [↑](#footnote-ref-23)
24. [**https://en.wikipedia.org/wiki/Benford%27s\_law#cite\_note-37**](https://en.wikipedia.org/wiki/Benford%27s_law#cite_note-37) [↑](#footnote-ref-24)
25. **See e.g., Decker, Myagkov, and Ordeshook (2011). Even Mebane, who has been a repeated applicant of the Law as a fraud detector in elections in multiple countries, has reiterated that while its violation might be taken as suggesting an anomaly, a violation of the supposed Law does not prove fraud. Fraud would need to be directly investigated (Mebane, 2020).** [↑](#footnote-ref-25)
26. **No one who studies elections thinks that the level of election fraud in 2020 was zero. In an election with over 150 million voters, there will be some who voted who were not entitled to do so, some who voted twice, some whose ballots were lost, some whose mail ballots were stolen and cast by others, and some who voted while dead. And there will be some election officials who (at least initially) reported inaccurate results. But while fraud in any given state** **may have slain its tens, or perhaps even its hundreds in particular states, its level was not enough to change presidential election outcomes (see below) and it certainly did not involve millions of stolen votes**. [↑](#footnote-ref-26)
27. **McCade, Aaron. “Ohio Secretary of State Touts Security of Election Process After Referring 27 Fraud Cases.”** [**https://www.newsweek.com/ohio-secretary-state-touts-security-election-process-after-referring-27-fraud-cases-1675215**](https://www.newsweek.com/ohio-secretary-state-touts-security-election-process-after-referring-27-fraud-cases-1675215) **February 1, 2022.** [↑](#footnote-ref-27)
28. **There is also a belief in the conservative blogosphere that, if there is fraud, it will be fraud committed by Democrat officials and found among the kinds of voters (such as racial minorities) who are most likely to vote Democratic. But is it is amusing to note that one of the first documented examples of actual fraud, so-called “voting the graveyard,” was committed by a Republican who cast a (presumably pro-Trump) ballot on behalf of his recently dead wife** (**JONATHAN CITE NEEDED**) [↑](#footnote-ref-28)
29. **As Gelman (2021) points out, some people say things like “who’s to say” when they hear claims that are patently implausible but yet respond in a way suggesting that they believe that claim might be true. Here possibility is confounded with probability. A common folk aphorism also relates to this confusion, e.g., “Where there is smoke there is fire.”****Here the size of the fire remains unspecified. In the context of election fraud, confusing possibility with probability is likely to be more prevalent among those who see the world in conspiratorial terms. Gelman quotes one person who accepts claims such as Obama is a Muslim as saying: “We see what they want us to see, I mean anything could be anything.”** [↑](#footnote-ref-29)
30. “Better Know a Ballot,” Accessed January 29, 2021, *A Late Show with Stephen Colbert,* https://www.betterknowaballot.com/. [↑](#footnote-ref-30)
31. In Pennsylvania, this included pointing viewers to a website developed by the Department of State. Ads featured prominent actors and athletes from the state. <https://www.votespa.com/Pages/default.aspx>. [↑](#footnote-ref-31)
32. “DNC Launches New Digital Ads in PA Reaching Vote-By-Mail Voters: “How to Return Your Ballot!””, September 24, 2020, *Democratic National Convention*, https://democrats.org/news/dnc-launches-new-digital-ads-in-pa-reaching-vote-by-mail-voters-how-to-return-your-ballot/. [↑](#footnote-ref-32)
33. Lai, Jonathan. 2020. “How ‘Naked Ballots’ in Pennsylvania Could Cost Joe Biden the Election.” *The Philadelphia Inquirer*. [↑](#footnote-ref-33)
34. “States That Permit Voters to Correct Signature Discrepancies”, September 21, 2020, *National Conference of State Legislatures*, <https://www.ncsl.org/research/elections-and-campaigns/vopp-table-15-states-that-permit-voters-to-correct-signature-discrepancies.aspx>. [↑](#footnote-ref-34)
35. See Morris Fiorina (2016), *Unstable Majorities*, and Frances Lee (2016), *Insecure Majorities*. Modern elections are very likely to result in divided government, and control of any branch of government is often won or lost in the margins. Partisan bias, such as that introduced by malapportionment or gerrymandering, can also affect the ability to carry marginal House or Senate seats. [↑](#footnote-ref-35)
36. For more details on Biden’s overperformance compared to U.S. House Democratic candidates, see William A. Galston, “Why did House Democrats underperform compared to Joe Biden?,” December 21, 2020, *Brookings*, https://www.brookings.edu/blog/fixgov/2020/12/21/why-did-house-democrats-underperform-compared-to-joe-biden/. [↑](#footnote-ref-36)
37. House Democrats nationally underperformed their 2018 performance, which accounts for the eleven net seats gained by the Republicans. (As of this writing on February 7, 2021, one seats has yet to be called, though the Republican candidate leads by 12 votes). Relative to the 115th Congress (2016-2018), the 177th Congress (2021-2023) has 28 more Democrats. [↑](#footnote-ref-37)
38. Figure omitted for space reasons. [↑](#footnote-ref-38)
39. An increase of 22 from 2016. [↑](#footnote-ref-39)
40. A decrease by 19 from 2016. [↑](#footnote-ref-40)
41. Relatedly, Grofman and Chen (2022) provide a conditional probability model to further explain change over time in the ability to use past elections to predict future elections. [↑](#footnote-ref-41)
42. Indeed, Eggers, Garro and Grimmer ( 2021, Figure 2) show that only 2% of counties had a different party winner in 2020 than they did in 2016. [↑](#footnote-ref-42)
43. This claim links to Twitter user David Chapman (@davidchapman141), a self-proclaimed “Author & Historian”. This “thread” is filled with statistics purportedly showing how Biden is the historic underdog going into the 2020 election. , e.g., “Incumbents are 6/6 when facing re-election during civil unrest”. See more: <https://twitter.com/davidchapman141/status/1315440579485069314?s=20>

    Internal links to this claim on Twitter say that the first primary was in 1912 and that Trump had received a higher percentage of the primary vote than Eisenhower, Nixon, Clinton, and Obama. Moreover, only five incumbents have received at least 90% of their primary vote. [↑](#footnote-ref-43)
44. This claim fails to account for well-known political science theory that suggests that competitive elections increase voter turnout (Downs 1957), and the empirical fact that modern elections are increasingly competitive at the national level (Lee 2016) even though less competitive for other offices, such as the U.S. House (https://alarm-redist.github.io/fifty-states/) [↑](#footnote-ref-44)
45. We might also argue that neither Trump nor Hillary Clinton could possibly have been elected President in 2016 since both were so disliked. Of course, no matter how unpopular one candidate may be, the other person might be even more unpopular. But also, relative likeability, is only one of the factors affecting vote choice. [↑](#footnote-ref-45)
46. This rebuttal to Ayyudurai (2020) is presented in Eggers, Garro, and Grimmer (2020), acknowledging its previous elucidation by Kabir (2020) and Parker (2020), each of whom show empirical evidence of the party-independence of results. However the Eggers, Garro, and Grimmer main rebuttal to the Ayyudurai (2020) claim uses the logic of latent variable analysis by demonstrating how regression to the mean effects lead to negatively sloped regression lines in the situation posited by Ayyudurai. [↑](#footnote-ref-46)