An Evaluation of Fraud Claims from the 2020 Trump Election Contests*

Justin Grimmer[a] and Abhinav Ramaswamy[b]

[a]Democracy and Polarization Lab, Political Science, and Hoover Institution. Stanford University. 616 Jane Stanford Way, Stanford CA 94305. jgrimmer@stanford.edu.
[b] Democracy and Polarization Lab, Political Science. Stanford University

January 16, 2024

Abstract

Even years after the 2020 election, Donald Trump continues to claim that fraudulent and illegal votes cost him the 2020 election. In this paper we provide the most comprehensive assessment of his empirical claims to date. All of the claims we evaluate fail to provide evidence of fraud or illegal voting. Trump's claims of fraud or illegality are riddled with errors, hampered by misunderstandings about how to analyze official voter records, and filled with confusion about basic statistical techniques and concepts. Often, the claims are based on the casual impressions of what happens in a "normal" election based on little more than intuitions. Worse yet, several claims are simply misstated by Trump's legal team or Trump. As a result, sometimes the public claims do not even match the weak evidence in Trump's legal challenges. This paper provides a resource for assessing many of the most prevalent claims made about the 2020 election and a guide to anticipating potential objections in future elections.

1 Introduction

After the 2020 election, Donald Trump, his campaign, and political allies engaged in a sustained effort to challenge Joe Biden's victory. This effort culminated with the events of

*For comments and data we thank Benjamin Ginsberg, Marc Meredith, Al Schmidt, Jonathan Mummolo, Ryan Germany, Eitan Hersh, Andrew Hall, Sean Westwood, Michael Herron, Matthew Tyler, Andrew Eggers, Duncan Carling, Samuel Beckerman, Christina Wang, Chris Rickert, Daniel Hopkins, Daniel Thompson, Apoorva Lal, and Max Hailperin. We thank Jasmine Sky Nguen for dedicated research assistance. Grimmer discloses that he served as an expert witness in the State Bar of California vs John Eastman disciplinary hearings as an expert for the State Bar.

January 6th, when Trump supporters stormed the US capitol to protest the official counting of ballots from the electoral college. But Donald Trump didn't stop objecting to the election just because he left the White House.

Even after officially losing the election and vacating the White House, Trump continues to object to the election results. A common refrain from Trump's objections is that many of his election cases were never actually heard in court, because the plaintiffs lacked standing. According to Trump and his allies' objections, if the cases had been heard, they would have revealed that Trump won the election. For example, in August 2023 Trump announced, and then canceled, a news conference that would detail "A Large, Complex, Detailed but Irrefutable REPORT on the Presidential Election Fraud which took place in Georgia." Then in January 2023 Trump released an unsigned 32 page memo supposedly detailing evidence of fraud and illegality in the 2020 election (Anonymous, 2023) .

In this paper we perform an extensive examination of claims made in Trump's legal challengers and in subsequent public statements about the 2020 election. We focus specifically on two sorts of claims. One set of empirical claims were made using registration files, voter history files, absentee voter files, and merges with other administrative data sets: such as data from the post office, individuals who are deceased, or lists of felons or other supposedly ineligible voters. A second set are statistical claims made about aggregate voting patterns that identify supposedly anomalous patterns consistent with voter fraud.

We examine 38 empirical claims, constituting the most complete assessment of empirical voter fraud claims from the 2020 election to date. The claims were alleged in courts throughout the country, including Arizona, Georgia, Michigan, Nevada, Pennsylvania, and Wisconsin. In each instance, we find that these claims fail to provide any evidence of fraud, illegality, or even an abnormality. One reason that the claims fail is that they are not based on facts (Eggers et al., 2021). We find that many claims misstate turnout histories, incorrectly characterize candidate's performance, or are just simply mistaken. A second reason that claims fail is that they are based on an inappropriate or incorrect application of statistical

models.

Regardless of the reason why, every claim we analyze fails to provide evidence of illegality or fraud. But what is even more remarkable is that none of the claims presented are remotely convincing. We document that the supposed evidence of fraud that Trump relies upon is riddled with basic statistical misunderstandings and errors, confusion about how to use voter files or absentee voter history to analyze turnout and registration, and invented statistical techniques based on the impressions of what happens in a "normal" election from "experts" who never previously analyzed election data and provide no argument to justify their procedures. At no point did Trump or his allies present even remotely plausible evidence of consequential fraud or illegality.

This paper is part of a much broader literature evaluating claims made about the 2020 election and voter fraud claims more generally. Our effort to provide a more comprehensive assessment of empirical voter fraud claims follows Danforth et al. (2022) who provide an expansive and impressive comprehensive view of claims made in Trump litigation. As we document in this paper, many of the claims made in Trump's initial court filings were refuted in expert reports that we discuss in this paper (Stewart, 2020a; Rodden, 2020; Rodden and Marble, 2020; Ansolabehere, 2020; Mayer, 2020; King, 2020; Herron, 2020). There have also been other academic papers that examine voter fraud claims from the 2020 election. Eggers et al. (2021) examine several broadly disseminated claims of "anomalous" election results and show these claims are either simply false, or technically true but not evidence of fraud. Grimmer et al. (2023) examine more expansive claims made after the 2020 election that allege every election is being manipulated. Grofman and Cervas (2023) examines the logical structure of voter fraud claims and show that they fail as a matter of basic logic. And Herron (2023) shows that there is no evidence Biden performed better in counties with Dominion machines.

We organize our assessment of the claims made about the 2020 election based on the specific states the claims target and the type of allegation made. That said, many of the

claims analyzed in one state were also made in other states. We first focus our analysis on nationwide claims, or claims involving several states' results and voting systems. We then provide an extensive analysis of Georgia elections, because Georgia is the location of perhaps the most aggressive objections to the 2020 election. We then analyze claims in Pennsylvania, Michigan, and Wisconsin. But the claims we analyze also cover many of the objections raised in Arizona and Nevada.

2 Nationwide Claims About Voting Patterns

The claims made that cover several states' elections are particularly troublesome, because they indict a large number of election systems across the United States. Yet, we find that the claims are baseless. Across all of the allegations, we show that they lack basic logical coherence, rest upon confusion about statistical terms, and ultimately have no evidentiary basis. In short, the expansive or nationwide claims about the 2020 election are based upon incredibly weak and often nonexistent evidence.

2.1 Contrast Analyses and Supposedly Anomalous Vote Returns

Skeptics of the 2020 election have claimed that the results were anomalous. Some of the claims were based on mere disbelief that Joe Biden could win the election. For example, Donald Trump tweeted on November 26th that "Just saw the vote tabulations. There is NO WAY Biden got 80,000,000 votes!!! This was a 100% RIGGED ELECTION" (Trump, 2020c). On November 30th, Trump continued to doubt Biden's popularity arguing that "We have some big things happening in our various litigations on the Election Hoax. Everybody knows it was Rigged. They know Biden didn't get more votes from the Black community than Obama, & certainly didn't get 80,000,000 votes. Look what happened in Detroit, Philadelphia, plus!" (Trump, 2020e). And in a speech in Georgia on December 5th, Trump exclaimed that

> I got more votes than any sitting president in history. 11 million more votes than we got in 2016. And we thought that if we could get 68 million, 67 million that would be the end. All of our great, brilliant geniuses said you'd win if you get 67 or 68. It's over. We got 74 million-plus and they're trying to convince us that we lost. We didn't lose. They found a lot of ballots to be nice about it (Trump, 2020a).

But other claims about anomalous results are based on differences from prior elections. Trump expressed one version of this skepticism in his December 5th speech in Georgia, when he said that

> President Obama beat Biden all over the country, except in some of the swing states where Biden beat him badly. How does that work? And they say it's statistically impossible. He beat crooked Hillary. Think of this. He beat crooked Hillary in the swing states, but she beat him everywhere else. (Trump, 2020a).

Trump never explains who declared this to be statistically impossible or who told him about what vote total targets to hit to guarantee victory. Nor does he point to concrete evidence that Joe Biden outperformed Barack Obama and Hillary Clinton in only the swing states or a report that makes that claim. One reason for the lack of citation might be that this claim is completely false. Compared to Barack Obama in 2012, Joe Biden in 2020 received a higher share of vote in 22 of 51 states plus DC. Compared to Hillary Clinton in 2016, Joe Biden received a higher share of the vote in 43 of 51 states plus DC.

But other analysts have taken up this basic idea—detecting anomalous shifts in votes—and claimed to find anomalies in the 2020 election. For example, in a series of reports authored after the 2020 election, Young and Blehar (2021) argue that they have a statistical technique to identify anomalous state or county level election results.[1] Young and Blehar's (2021) contrast analysis analyzes whether the vote counts in the 2020 election were similar

[1]The contrast analysis was also included in an expert report filed in *King v Whitmer* in Michigan.

to the the vote counts in the 2016 election. To do this, Young and Blehar (2021) examine state or county-level changes in support for a particular candidate. Given the primary focus on counties, we derive their test at the county level first.

In a state i and county c Young and Blehar (2021) computed the difference in the *number* of votes Joe Biden received in 2020 (Biden_{ic}), and the number of votes Hillary Clinton received in that county in 2016 (Clinton_{ic}): Biden_{ic} - Clinton_{ic}. They also computed the difference in the number of votes that Donald Trump received in 2020 ($\text{Trump}_{ic,2020}$) and the number of votes he received in 2016 ($\text{Trump}_{ic,2016}$): $\text{Trump}_{ic,2020}$ - $\text{Trump}_{ic,2016}$. They then call the "contrast" in county c in state i the difference of those differences,

$$\text{Contrast}_{ic} \quad = \quad (\text{Biden}_{ic} - \text{Clinton}_{ic}) - \left(\text{Trump}_{ic,2020} - \text{Trump}_{ic,2016}\right)$$

Young and Blehar (2021) then visually inspect the contrasts of counties within a state. They order counties according to the size of the contrast. They declare some counties as anomalous based on contrasts they view as large, despite offering no formal definition of "large".

Young and Blehar (2021) offer no systematic justification for their use of contrasts to detect supposedly anomalous election results. Young and Blehar (2021) do not cite a single study that analyzes the use of contrasts to detect election anomalies. They also fail to even cite a common understanding of elections that justifies the use of contrasts for assessing where anomalous election results take place. And while Young and Blehar (2021) argue that "If there is fraud, a contrast and distribution analysis will likely point to where it happened" (Young and Blehar, 2021, 4), they offer no assessment of their method's performance when applied to actual election results and provide no evidence or calculation to support this conclusion.

There is no reason to expect that the contrast analysis will reliably detect anomalous

election results. Instead, contrast analyses merely detect that larger counties have bigger changes in the number of votes. This is because the contrast analysis focuses on vote count. This implies that the potential size of a contrast depends entirely on the size of the county. One simple way to see the limitation of a contrast analysis is to note that uniform swings in candidate vote share in a state will create the appearance of a larger anomaly in the larger counties, merely because there are more voters in those larger counties. For example, suppose that there is a state with 99 1,000 voter counties and one 1,000,000 voter county and that there is a 5 percentage point shift in support from Trump in 2016 to Biden in 2020. If this occurs in the 1,000 person county, this will result in a 100 person contrast. In the 1,000,000 voter county, however, this will result in a 100,000 contrast. So, despite all counties experiencing the same shift in support for candidates, the larger counties appear more anomalous merely because there are more voters.

In fact, the maximum size of a contrast is determined mechanically by the number of voters in the county. This maximum possible contrast occurs when all voters in a county change their support from one party's candidate to the other. When this occurs, the contrast will be twice the number of voters. Necessarily, any other shifts must be below the maximum contrast. But this means that it is effectively impossible to discover an "anomaly" in small counties. Because the size of anomalies in a county is determined in large part by the number of voters in the county, it is not surprising that supposedly anonymous contrasts are found in every nearly every state in every presidential election over the past 60 years. In Figure 1 we show that in every presidential election in Georgia, the contrast analysis method detects supposedly "anomalous" election results.

Consider the top-left plot first. This plot recreates the contrast analysis found in Young and Blehar (2021) for the state of Georgia. This plot computes the county-level contrasts comparing the 2020 and 2016 presidential election. The vertical axis displays the contrast. Counties are ordered by the size of the contrast, and their rank in this ordering is displayed on the horizontal axis. The red dots in the plot represent the four largest counties in Georgia

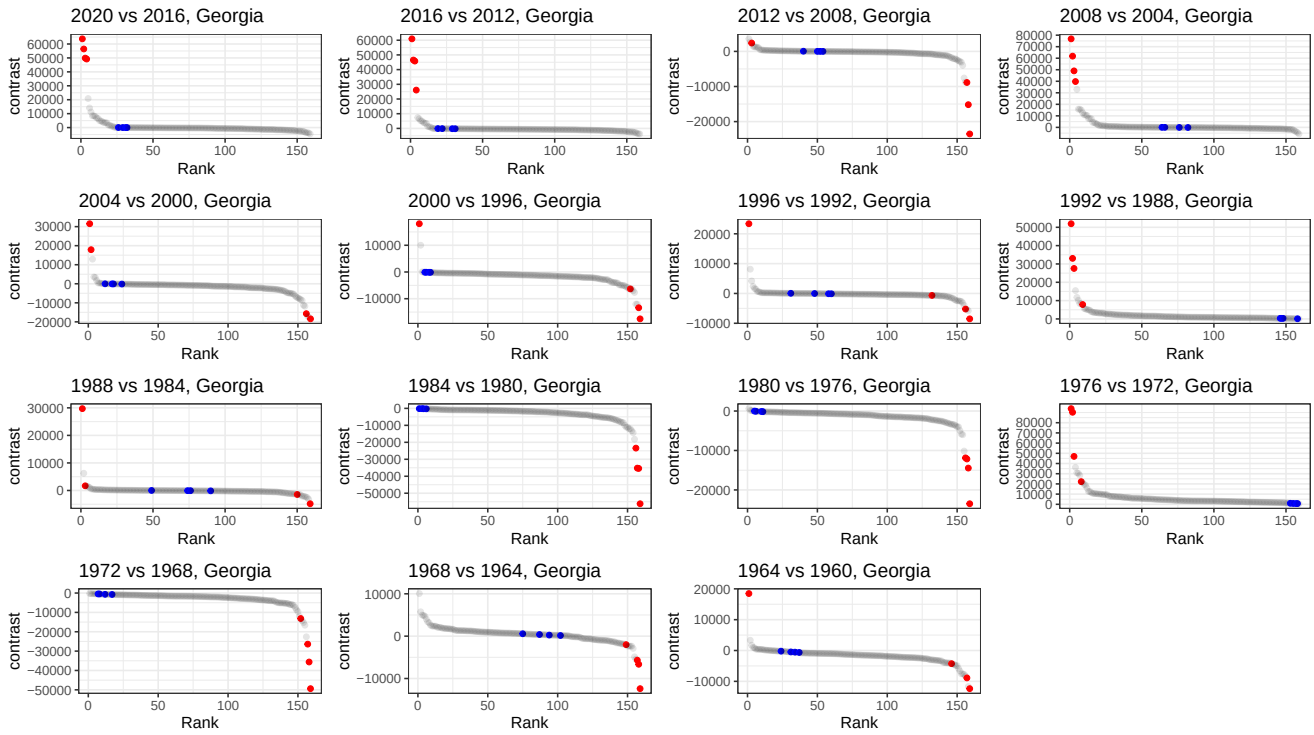

Figure 1: Examining the county-level contrasts for all Georgia presidential elections since 1960 reveals that supposedly anomalous contrasts occur in every election.

in 2020: Fulton, DeKalb, Cobb, and Gwinnett. The blue dots represent the four smallest counties in Georgia in 2020: Clay, Webster, Quitman, Taliaferro. Young and Blehar (2021) interpret the large contrast in the largest counties as suggestive evidence that there were illegal or fraudulent ballots cast in those counties.

Yet, looking at the other plots in Figure 1 reveals that Fulton, DeKalb, Cobb, and Gwinnett counties are almost always identified as "anomalous". When comparing 2016 to 2012 we find that the four largest counties also have a supposedly large positive anomalous deviation when comparing 2016 to 2012 (the top row second column from the left plot). When comparing 2012 to 2008 three of the four counties have large negative "anomalous" deviations. In fact, in every presidential election since 1964 these four counties produced supposedly anomalous deviations from the prior election.

Similarly, the four smallest counties are never anomalous. Comparing 2020 to 2016 we

find very small deviations in these counties. That makes sense, because the counties are much smaller than the largest counties in Georgia, so the smallest counties largest possible contrast is much smaller than the largest counties. And as the plots show, in every election since 1964 the small counties necessarily have small contrasts.

The plots in Figure 1 show that the contrast analysis is a useless tool for diagnosing surprising or anomalous results. Young and Blehar's (2021) contrast analysis methodology merely uncovers the unsurprising fact that larger counties have larger deviations in their vote counts. And smaller counties have smaller deviations.[2]

Related Claims About Anomalous Election Results There have been other, very similar, claims that the results of the 2020 election were somehow different than the results of prior elections. These other claims fail for nearly identical reasons as the contrast analysis.

In the bill of complaint for the *Texas V Pennsylvania* lawsuit, the authors claim that an expert report from Charles Cicchetti demonstrated that there was a "one in quadrillion" chance of Joe Biden winning the 2020 election (Paxton, 2020b). But as Eggers et al. (2021) demonstrate, Cicchetti never calculated the probability Biden would win. Instead, Cicchetti calculated the probability that the 2020 and 2016 elections had the same underlying rates of support for the Democratic and Republican candidates *and* the same number of voters in both elections. It is not at all surprising that candidates' vote shares change across elections or that there will be different numbers of voters. And, at any rate, this test has nothing to do with the probability of Biden winning. In fact, in Cicchetti (2022) he explains that the authors of the bill of complaint in *Texas V Pennsylvania* misinterpreted his expert report.

In a case filed by Trump in Arizona, *Bowyer v Ducey*, Seth Keshel filed an expert report

[2]Young and Blehar (2021) also seem to be unaware that there are changes in voter turnout in different elections. The contrast analysis report exerts considerable effort to explain changes in vote count based on turnout. But of course, the number of votes candidates receive can be the result of changes in the population of a state. But they can also be the result of persuading individuals to vote for a different candidate, or the result of candidates increasing (or decreasing) interest among individuals who hadn't voted in the prior election.

where he claimed to find evidence that the 2020 election results were anomalous relative to prior election results and changes in party registration (Keshel, 2020). To support this conclusion, Keshel demonstrates that the changes in Arizona deviate from his predictive models.

But deviations from predictive models are simply not evidence that fraud or illegal voting occurred. It is much more plausible the reason Keshel's model fails to make accurate predictions is that he did a poor job constructing the model (King, 2020). Or, it is also plausible that the 2020 election deviated from the patterns in prior elections, but for completely legal reasons. Keshel offers no evidence that the deviations he identifies are the result of illegal or fraudulent activity.

2.2 Joe Biden Performed Better than Hillary Clinton in Urban Areas Throughout the Country

Skeptics of the 2020 election claim that Joe Biden had suspicious performance in urban areas throughout the country. For example, some conservative-leaning media organizations ran articles claiming that Biden generally performed worse than Clinton in urban areas, but that Biden outperformed Clinton in urban areas found in swing states (Vespa, 2020). Those reports, however, had to be corrected as the outlets were alerted that their analysis was simply incorrect.

A more coherent version of the allegation comes from the litigation around Texas lawsuit in early December at the Supreme Court. In a reply brief, Texas' expert, Charles Cicchetti, compared Hillary Clinton's performance in urban counties to Joe Biden's performance (Paxton, 2020a). Cicchetti writes that "In 2016 limited to the major-party candidates Clinton had 51% of all the votes nationally compared to Trump's 49%. She won in counties with large urban areas with a much larger percent than Trump. Her share of the votes in counties where the Top-50 cities are located was 67.8% compared to Trump's 32.2%" (Paxton, 2020a, 155a). Cicchetti goes on to claim that "Biden did not do as well as relative to Clinton in

2020 the counties [sic]. He had a lead in the Top-50 cities of 66.4% compared to Trump's 33.6% ignoring third part [sic] shares. The difference in Democrat vote percentage between 2020 and 2016 is a loss of 1.4% from Biden compared to Clinton" (Paxton, 2020a, 155a-156a). Cicchetti then compares Clinton and Biden's performance for large urban counties in Georgia, Pennsylvania, Wisconsin, and Michigan and finds that in four of the five largest urban counties, Biden outperformed Clinton. Based on this data, Cicchetti argues that "Coincidences are possible, but relying on them is questionable. Analyzing data and stubborn facts are a better way to determine what we observe. The facts show that while Biden had more ballots in urban areas than Clinton, she outperformed him in terms of the percentage of the vote she received" (Paxton, 2020a, 157a). Cicchetti concludes that "it is worthwhile to understand that in the battleground states comprised of the four defendants that in four of the five urban areas Biden had higher percentages than Clinton. This result is surprising given the Trump national gains in support from Hispanics and African Americans. This contrary observation supports further investigation to determine what happened and why. This would not be a worthless outcome" (Paxton, 2020a, 158a).

Cicchetti does not explicitly identify the fifty counties that used in his analysis and he has never posted replication code for his analysis. To analyze his claims about Biden's performance in urban counties, we use data from the US Census Bureau to identify the fifty counties with the largest population and voting returns from the 2020 and 2016 elections. Using the election data, we then contrasted Biden's share of the two-party vote in 2020 to Clinton's share of the two-party vote in 2016 in Table 1.

Table 1 shows that Biden actually outperformed Clinton in forty of the fifty urban counties. The first row shows that the largest increase occurred in Salt Lake County, Utah where Biden in 2020 obtained 11.5 percentage points more of the two-party vote share than Clinton in 2016. There was also a large increase in Collin County, Texas where Biden secured 8.1 percentage points more of the two-party vote than Clinton in 2016. And in Hennepin County, Minnesota Biden obtained 7.3 percentage points more of the two-party vote share

County	Biden Share - Clinton Share	Biden Share	Clinton Share
Salt Lake,UT	0.115	0.530	0.415
Collin,TX	0.081	0.470	0.389
Hennepin,MN	0.073	0.705	0.631
Tarrant,TX	0.062	0.493	0.431
Middlesex,MA	0.062	0.715	0.653
St Louis,MO	0.060	0.612	0.552
Travis,TX	0.059	0.716	0.658
Fairfax,VA	0.055	0.699	0.644
Maricopa,AZ	0.053	0.501	0.448
Pima,AZ	0.051	0.584	0.533
King,WA	0.051	0.750	0.698
Oakland,MI	0.050	0.562	0.513
Franklin,OH	0.049	0.647	0.598
Fulton,GA	0.049	0.726	0.677
Wake,NC	0.049	0.623	0.574
Suffolk,NY	0.048	0.494	0.446
Mecklenburg,NC	0.044	0.667	0.623
Dallas,TX	0.043	0.651	0.608
Contra Costa,CA	0.041	0.716	0.675
Duval,FL	0.040	0.511	0.471
Bexar,TX	0.040	0.582	0.542
San Diego,CA	0.039	0.602	0.563
Montgomery,MD	0.039	0.786	0.747
Fresno,CA	0.037	0.529	0.492
Allegheny,PA	0.035	0.594	0.559
Sacramento,CA	0.034	0.614	0.580
Riverside,CA	0.033	0.530	0.497
Nassau,NY	0.028	0.541	0.513
Westchester,NY	0.027	0.676	0.649
Orange,CA	0.025	0.535	0.509
San Bernardino,CA	0.021	0.542	0.521
Harris,TX	0.020	0.560	0.540
Wayne,MI	0.020	0.683	0.664
Alameda,CA	0.017	0.798	0.781
Hillsborough,FL	0.017	0.527	0.510
Clark,NV	0.012	0.537	0.524
Orange,FL	0.011	0.609	0.598
Honolulu,HI	0.010	0.625	0.615
Cuyahoga,OH	0.009	0.664	0.654
Cook,IL	0.003	0.742	0.739
Santa Clara,CA	-0.001	0.726	0.727
New York,NY	-0.001	0.864	0.866
Palm Beach,FL	-0.003	0.560	0.562
Los Angeles,CA	-0.007	0.710	0.718
Philadelphia,PA	-0.011	0.812	0.823
Broward,FL	-0.016	0.645	0.661
Kings,NY	-0.027	0.768	0.795
Queens,NY	-0.033	0.720	0.754
Bronx,NY	-0.052	0.833	0.885
Miami-Dade,FL	-0.099	0.533	0.632

Table 1: Comparing the vote share in the fifty largest counties in the US reveals that Biden improved over Clinton's two-party vote share in forty out of fifty counties.

than Clinton.

All of these increases are larger than the counties in swing states that Cicchetti identified as supposedly anomalous. For example, Fulton County, Georgia the official results show Biden winning 4.9 percentage points more of the two-party vote share than Clinton. In Wayne County, Michigan Biden received 2 percentage points more of the two-party vote share than Clinton, in Allegheny County, Pennsylvania Biden received 3.5 percentage points more of the two party vote than Clinton, and in Philadelphia County, Pennsylvania Biden saw a 1.1 percentage point decline in vote share compared to Clinton. Cicchetti includes Milwaukee County, Wisconsin even though it is not a top fifty county in population size. There, Biden obtained 3.6 percentage points more of the two-party vote share than Clinton. Averaged across the counties, we find that Biden's average vote share was higher than Clinton's vote share in 2016 in the urban counties. Averaged across the counties in 2020 Biden received 64.8% of the two-party vote share, while Clinton received 64.6% of the two-party vote share across counties.

Cicchetti's conclusion that Biden received a smaller share of the two-party vote share than Clinton is based on tallying all the votes each candidate received across the fifty largest counties. Then Cicchetti computed the share of votes each candidate received in the corresponding election. When we perform that calculation, we find that Biden received 64.9% of the two-party vote in 2020, while Clinton received 65.7% of the two-party vote. This disparity is driven entirely by the differences found in counties in Florida, New York, and California. As Table 1 shows, Biden consistently did better in urban areas throughout the country.

Cicchetti never explains why this calculation is relevant to his claim that the 2020 election was anomalous or that the smaller share of votes for Biden overall because of Florida, New York, and California are indicative of Biden's performance otherwise being surprising. Further, Biden actually secured a much larger total vote margin across urban counties in 2020 (13,099,708 more votes than Trump across urban counties) than Clinton in 2016 (11,179,626

more votes than Trump across urban counties).

It is simply false that Biden performed worse than Clinton in urban areas. And it is not clear what implication that would have for identifying an anomalous result.

2.3 Allegedly Surprising Vote Additions through "Vote Spikes"

One of the first claims of fraud in the 2020 election was that there were inexplicable large numbers of votes that were allocated to Joe Biden. The morning after the election President Trump tweeted that "Last night I was leading, often solidly, in many key States, in almost all instances Democrat run & controlled. Then, one by one, they started to magically disappear as surprise ballot dumps were counted. VERY STRANGE, and the pollsters got it completely & historically wrong!" (Trump, 2020f). Trump made a similar claim in tweet from November 19th, 2020 when he linked to a chart of the vote counts from Wisconsin and claimed that "Look at this in Wisconsin! A day AFTER the election, Biden receives a dump of 143,379 votes at 3:42AM, when they learned he was losing badly. This is unbelievable!" (Trump, 2020d). In a speech in Georgia on December 5th, Donald Trump continued to make allegations about supposedly erroneous large increases in votes for Joe Biden. Trump declared that "the alleged Biden margin of victory in several states is entirely accounted for by extraordinarily large midnight vote dumps. You saw them going up to the sky, all extremely skewed to Biden. Like, Biden would get 10,000 votes and Trump would get three in a Trump area. Oh. For, example, at 6:31 AM in Michigan, reported 141,238 votes for Joe Biden and 5,900 votes for Donald Trump. 96% for Biden, 4% for Trump in an area where I should be doing quite well" (Trump, 2020a). Donald Trump then repeated his allegation about Wisconsin vote tallies, claiming that "A similar vote dump occurred in Wisconsin at 3:42 AM. This is 3:42 in the morning. In fact, when I went to bed that night, everybody was calling, "Congratulations. You're up 600,000 votes in Pennsylvania" (Trump, 2020a).

These claims misunderstand basic facts about how votes are tallied in American elections and where candidate support is located. The votes Trump referenced in Wisconsin were

merely ballots from Milwaukee being reported, a densely populated city where Democrats have historically performed well. The votes Trump referenced from Michigan correspond to votes from Wayne County, where Biden was expected to receive many more votes than Trump (Team, 2020). In fact, the geography of partisan support in US elections leads to an expectation that there will be some large increases in Democratic candidate's vote totals as urban centers report their ballots. This is because support for Democratic candidates tends to be concentrated in a small number of large population counties, while support for Republican candidates tends to be concentrated in a large number of low population counties. When the larger counties report their vote totals, it will create the illusion of a "spike" in support, merely reflecting the large number of ballots from the county that tend to support the Democratic candidates at higher rates.

All of the claims appear to be based on election counting data from Edison, a media company that reports unofficial vote tallies as they are made available. Edison collects both official and unofficial election data from states and reports them to networks, who use the information in their election night reporting of the voting counting process.

Using the information from the Edison data, Trump alleged that there was something surprising, illegal, or fraudulent about large tranches of ballots being reported late in the night. While many of the claims merely involved observing the presence of the spikes, other analysts have produced reports that claim to provide statistical evidence that the vote tallies in the 2020 election were anomalous. These efforts were collected in the "2020 Presidential Election Startling Vote Spikes", which alleged to uncover evidence that the large changes in the vote count tended to favor Biden (Quinnell et al., 2021).[3] Quinnell et al. (2021) argue that "On average, we expect Trump/Biden jumps to be of the same order of magnitude for each candidate, especially for close races. Wild differences ('jumps') in magnitudes, and especially ones that favor a particular candidate, are signs of potential anomalies" (Quinnell et al., 2021, 3). Quinnell et al. (2021) provide no citation or argument to explain why they

[3] An earlier version of this report was included in an expert report filed in *King v Whitmer*.

expect close races to also have close vote updates. Nor do Quinnell et al. (2021) explain why this expectation would occur given the well known geographic distribution of support for political parties.

Quinnell et al. (2021) instead perform a null hypothesis test to assess whether particular updates were surprising. Specifically, Quinnell et al. (2021) count the change in Biden and Trump votes and then divide by the number of seconds since the last update, which they call the "the 'rate' of vote counts added" (Quinnell et al., 2021, 3). Quinnell et al. (2021) then calculate the average "rate" of vote counts added and take the standard deviation of those counts.[4]

Quinnell et al. (2021) then use the calculated rates to perform a null hypothesis test based on the assumption that the rate of vote counts added follows a normal distribution. Using this assumption Quinnell et al. (2021) calculate a z-statistic for each update by subtracting the average vote count added over the updates and diving by the standard deviation. Quinnell et al. (2021) then call z-statistics greater than 10 "statistically impossible." They also claim that the probability of vote jumps in several state is extremely low. For example, in Florida they claim that the probability of one jump is "1 in 10^{23}" and another jump is "1 in 10^{25}" (Quinnell et al., 2021, 5).

Quinnell et al. (2021) calculated probabilities depend *entirely* on the assumption that in a free and fair election election night vote count updates will follow a normal distribution. If that assumption fails—if the updates follow some other distribution—all of their calculated probabilities and claims about statistical impossibility are completely meaningless. Yet, Quinnell et al. (2021) provide no evidence, discussion, citation, or proof that vote counts

[4]Quinnell et al. (2021) fail to explain why they think it is appropriate to divide by the number of seconds since the last update in the state to calculate a "rate." An obvious objection is that counting in a state occurs in parallel with many small localities counting their votes at the same time. Because of this, it is not accurate to allege that all the ballots reported in an update had been counted since the last update. Nor would it even be accurate to make the same claim within a county or counting center. This is because ballots can be counted in parallel even within the same location. If we perform the same ill-advised null hypothesis test Quinnell et al. (2021) perform using the vote counts, rather than the vote count rates, we find fewer anomalies according to Quinnell et al's (2021) own standard. But as we explain below, even this test is inappropriate and meaningless for diagnosing fraud.

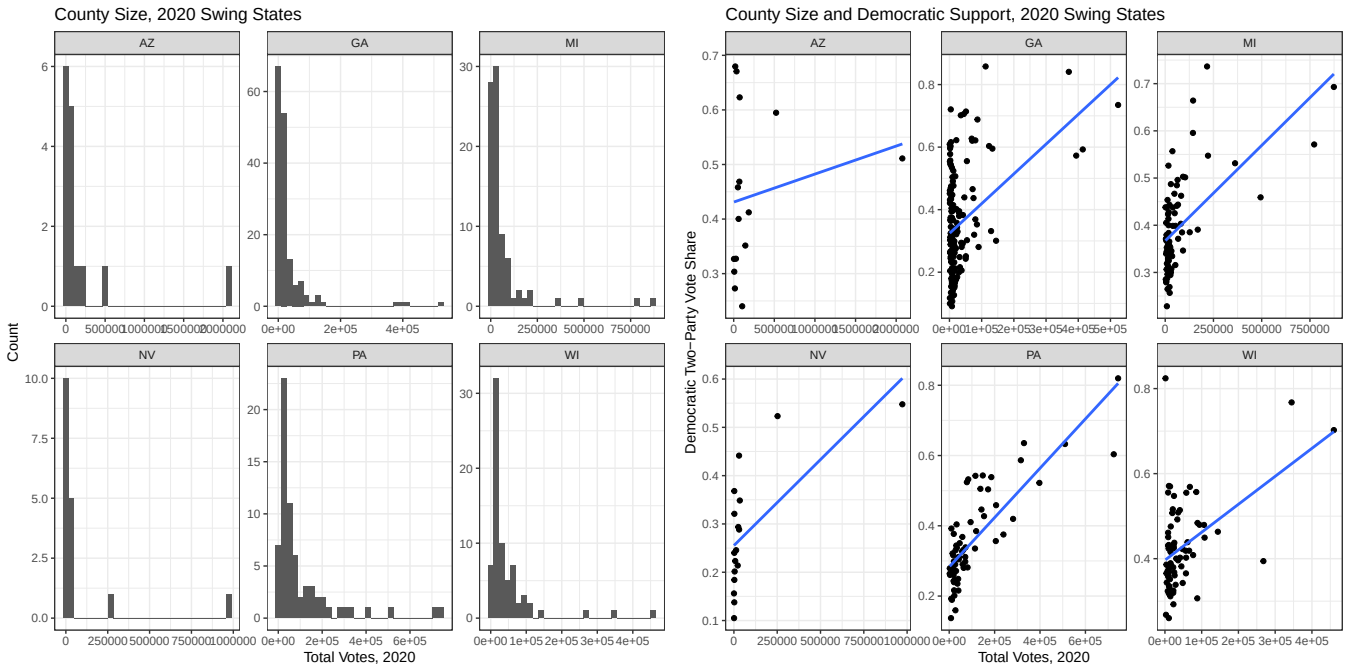

Figure 2: Distribution of County Population Size and Support for Biden in the 2020 election 6 swing states

being added will follow a normal distribution in the absence of illegality or fraud. Rather, they merely make the assumption and apply it.

But a basic understanding of how votes are counted in US elections and the geography of political support would easily show that a normal distribution is inappropriate. First, in every state—but especially the swing states in 2020—there are a small number of large population counties and a large number of small population counties. The left-hand plot of Figure 2 shows a histogram of the total number of votes reported across six swing states in the 2020 election—Arizona, Georgia, Michigan, Nevada, Pennsylvania, and Wisconsin. The histogram shows that most counties in the swing states have a relatively small number of votes cast in 2020. But in the swing states there also is a small number of large population counties, which results in the histogram having a long tail to the right.

The right hand plot shows that counties with more total votes tended to allocate a larger share of their votes to Joe Biden. The horizontal axis of Figure 2 is the total number of votes in the county, while the vertical axis is Joe Biden's share of the two-party vote. The

blue line is a linear regression line, or a line of best fit. In all six swing states the regression line is upward sloping. This indicates that voters in counties with more total votes allocated a bigger share of their votes to Joe Biden than smaller counties.

The combined plots in Figure 2 demonstrates that we would not expect updates of vote counts to follow a normal distribution. This is because updates from large counties will tend to report many more votes than updates from smaller counties and those votes will tend to be cast for Joe Biden. If the large counties report the same share of their votes as the small counties when they provide an update, the result will be vote totals that are much larger than the updates from other counties and where most votes are cast for Biden. This will create long right-hand tail, unlike a normal distribution. And this occurs without any fraud, illegality, or other "suspicious" manipulation of votes.

2.4 Claims About Addresses, Movers, and Other Allegedly Illegal Activity

In several states Trump and his allies claimed that lax regulation of voter files resulted in illegally cast ballots. For example, in a speech given the day before the Georgia Senate runoff, Trump declared that in the 2020 presidential election "At least 15,000 ballots were cast by individuals who moved out of the state prior to the November 3rd election, or maybe they moved back in. I don't know. I mean, I can't tell. They moved out, ah, let's go back. Usually takes a little time, right? We moved out. Let's go back, darling" (Trump, 2021a). This is a reference to claims made in Georgia, where expert reports supposedly demonstrated that voters had moved and still cast votes. Similarly, on December 16th Trump tweeted that "Chris Krebs was totally excoriated and proven wrong at the Senate Hearing on the Fraudulent 2020 Election. Massive FRAUD took place with machines, people voting from out of state, illegals, dead people, no signatures—and so much more!" (Trump, 2020b). This is a reference to an expert report from Jesse Kamzol in Nevada, which was summarized in testimony by the attorney in the case, Jesse Binnal. Binnall's testimony was

based on a five-page expert report submitted by Jesse Kamzol in the *Law v Whitmer* case in Nevada. These claims about voters with invalid addresses were made across the country, as expert reports by Matt Braynard from *Texas v. Pennsylvania* and *Stevenson v. Ducey* alleged there were thousands of such ineligible voters in Arizona, Pennsylvania, Wisconsin and Michigan (Braynard, 2020b,c).

All of the claims made in these reports misunderstand the basics of election law in each state or incorrectly label individuals as ineligible. For example, in Georgia, Davis (2020a) and Braynard (2020a) claim to show evidence that individuals cast votes who also files national change of address forms with the United States Postal Service. Yet, Grimmer et al. (2021) explain that filing a national change of address form does not cancel a voter's registration. In both Georgia and Nevada voters have to establish residences in new states before their prior registration would be canceled. There are many reasons why voters could file a change of address and still retain their registration—including visiting family, using a vacation home, or traveling out of state for work. In fact, during Jesse Kamzol's deposition in *Law v Whitmer* he sometimes admitted he didn't know if a voter he labeled as suspicious actually violated Nevada law by voting. Other times he conceded that there are legitimate reasons individuals could have moved and maintained eligibility. Davis (2020a), who asserts that some Georgia voters moved across counties without reregistering, makes the same concession in his expert report.

Kamzol (2020) also alleged that individuals cast votes in Nevada from vacant or non-existing residential addresses. Kamzol argued that "I reviewed the Washoe County and Clark County voter database files and found 8,111 voters who voted in Nevada but whose registered addresses are physically non-existent, as in, these are vacant lots, apartment numbers that do not exist, and house address numbers that do not exist" (Kamzol, 2020, 1). Similar claims were made in Geels' (2020) and Braynard's (2020) expert reports in Georgia. In both instances, the expert reports fail to understand the basics of state election law. For example, Kamzol's report relies upon the "CASS" system from the United States Postal Service

which provides an incomplete listing of residential addresses (Cegavske and Wlaschin, 2021). Further, the Nevada Secretary of State's office performed a search of supposedly vacant addresses from Kamzol's report and found that all are actual addresses. Similarly, Braynard's (2020) report that supposedly identified illegal registrations was shown in a hearing to actually be apartment buildings. Geels (2020) report does appear to identify a small number of individuals who registered at PO Boxes. However, this number is no where near the number of votes Trump lost by in Georgia, and in any case, Georgia law is explicit that challenges to registration have to come *before* the election, not after the vote has already been cast (Georgia Code 21-2-230 (2022)). So even if these claims were accurate (which they aren't), the Trump campaign should have surfaced them before the election.

2.5 Survey-Based Claims of Unrequested or Uncounted Absentee Ballots

In several of Trump's challenges to election results his lawyers attached a two page expert report from William Briggs, who claims to be the "statistician to the stars" and holds a Ph.D. in Cornell University.[5] Briggs' expert report uses survey data provided to him by Matt Braynard. According to the report, Briggs analyzes survey results for individuals who, official state data reports, did not return their absentee ballots. Briggs' survey then attempts to identify two kinds of errors in that official data. First, Briggs examines the share of individuals who report requesting a ballot. Briggs calls individuals who report not requesting a ballot an error. Second, among those who say they did request a ballot, Briggs examines the share of respondents who report returning their ballot. Because the sample frame was limited to only those who records show did not return their ballot, Briggs calls this second response a second source of error.

Briggs' survey methodology is deeply flawed and is a poor basis for challenging absentee

[5]The affidavit was submitted, at least, as part of *Pearson v Kemp* in Georgia, *Bowyer v. Ducey* in Arizona, and the data was also submitted as part of *King v Whitmer* in Michigan

voting procedures. Ansolabehere (2020) provides an extensive critique of Briggs' methodology. Perhaps the most important limitation of Briggs' analysis is that he fails to consider that survey respondents might imperfectly recall their own behavior. This undermines the entire premise of Briggs' analysis because an "error" could either be the result of illegal voting or simply that the respondent fails to recall their voting appropriately. In fact, there is good evidence that voters generally struggle to recall how they voted in elections, casting doubt on the ability to use surveys to uncover supposedly fraudulent votes. Shino et al. (2022) use extensive survey data to show that voters alter reports of how they plan to vote and how they actually voted based on which candidate they supported in the 2020 election.

Another limitation Ansolabehere (2020) identifies is that Briggs fails to recognize that in many states, some voters can apply to be placed on a permanent absentee voter list. For example, in Georgia, voters over 65 or self-declared disabled can apply for an absentee ballot once during an election cycle and then receive a ballot for the remainder of the election cycle. As a result, a Georgia voter could apply for a ballot in the primary, then receive a ballot for the general election without remembering that they had applied in the primary.

Ansolabahere (2020) diagnoses a series of other fatal flaws in Briggs' analysis that undermine his claims. For example, Ansolabahere (2020) shows that the questions used to screen respondents allowed for individuals other than the targeted individual to respond to the survey. Further, Ansolabehere (2020) shows that the response rates were low enough to raise concerns about the validity of the survey. And finally, Ansolabehere (2020) shows that ambiguously worded questions in Briggs' survey undermines the validity of the responses.

3 Georgia Claims

In many ways Georgia has been the primary battleground for voter fraud claims. This is true, in part, because Georgia has been extremely closely contested. Georgia's 2020 presidential election had a razor thin margin, with Joe Biden winning the election by a mere 11,779 votes.

Immediately after the 2020 general election, there was a runoff election for both Senate seats that would determine which party would hold a majority in the Senate. To contest the results of the presidential election the Trump campaign legal team and private citizens filed lawsuits in state and federal court to challenge the results of the election. These claims were dismissed in court (Danforth et al., 2022) and have subsequently become the basis for Trump and his allies' indictment for their alleged roles in challenging the results of the 2020 election.

Here, we perform an extensive analysis of claims made in the Georgia legal cases. We divide the claims into empirical claims made about the number of illegal votes cast in the Georgia elections and statistical claims that the aggregate election results were anomalous and inconsistent with a free and fair election.

In all instances we find no support for alleged claims of fraud or illegality. Many of the claims are not evidence of fraud because they are simply not true. For example, many claims confuse signature rejection rates or are based on incorrect analyses of administrative data. Other claims are not evidence of fraud because supposedly anomalous patterns are not anomalous at all. For example, we show that several claims made about "surprising" returns are merely based on fundamental misunderstandings about election data.

3.1 Claims Made About Illegal Votes in Georgia

3.1.1 Signature Matching in Georgia

The 2020 general election in Georgia saw a much higher rate of mail-in absentee voting than in previous elections. This surge in mail-in voting is unsurprising given public concerns about the COVID-19 pandemic. After the election, Trump and his allies argued that these ballots were not treated appropriately. In a rally before the January 5th, 2021 Senate runoff election in Georgia, Trump claimed that "Across all of the key swing states, there were impossibly low rejection rates, you saw that, for mail-in ballots, drastically less than the

historic norm. In Georgia 0.5% of the mail-in ballots were rejected in 2020 compared to 5.77%. That's a difference of 11 times more. It's hundreds of thousands of votes" (Trump, 2021a). He attributed this decline to a change in how signatures on absentee ballots were verified, arguing that "Democrats, state Supreme Court judges and Democrat Secretary of State effectively abolished the signature verification process right here" (Trump, 2021a). Trump alleged that this occurred because of a supposed consent decree signed between the Secretary of State's office and Stacey Abrams, whose organization sued the Secretary of State to modify the signature verification process before the 2020 election (Trump, 2021a). In reality, the Secretary of State's office merely issued non-binding guidance on best practices for evaluating signatures. No actual rule changes occurred.

Trump allies testified to similar claims in election hearings. For example, at a Georgia Senate hearing John Eastman claimed that the "settlement agreement that was entered into with your Secretary of State and certain Democrat committee challengers that effectively altered the signature verification process for your absentee ballots in ways that are contrary to state law" (Eastman, 2020a).

These claims also appear in the Trump campaign's legal challenges. In *Trump v Raffensperger*, Trump's legal team argued that "The explosion in the number of absentee ballots received, counted, and included in the tabulations for the Contested Election, with the simultaneous precipitous drop in the percentage of absentee ballots rejected, demonstrates there was little or no proper review and confirmation of the eligibility and identity of absentee voters during the Contested Election." Claims about signature rejections were found across many different lawsuits in both Georgia state court, federal court, and at the Supreme Court—*Trump v Raffensperger*, *Trump v Kemp*, *Texas v Pennsylvania*, *Pearson v Kemp*, and *Wood v Raffensperger*.

The claims also surfaced as part of Trump's efforts to pressure Georgia election officials to alter the election outcomes. Trump claimed in his phone call with Secretary of State Brad Raffensperger that "We think that if you check the signatures – a real check of the

signatures going back in Fulton County you'll find at least a couple of hundred thousand of forged signatures of people who have been forged" (Staff, 2021).

Overall Rejection Rates Declined, But Not Because Of A Change in the Signature Rejection Rate Trump, his allies, and his litigation all focused their claims on the overall rejection rate of mail-in absentee ballots. But this is an error if their goal was to make a claim about changes in the rate ballots were rejected for signature reasons because there are many reasons that mail-in absentee ballots can be rejected. And even though the overall rejection rate did decline, the rate at which ballots were rejected for incorrect signatures did not.

Trump's public claims were based on expert reports that were filed with Trump's litigation contests. These claims about mail-in absentee voter rate rejection rates were based, largely, on analyses of the Georgia Absentee Voter History file.[6] The Georgia Secretary of State's Office compiles information on every mail-in ballot application from across the state. This data then records whether the application is accepted, whether the ballot is returned, and whether the ballot is accepted or rejected once it is returned.

Using this data, several expert reports calculated rejection rates for recent elections. For example, in affidavits submitted in Georgia for *Trump v Raffensperger*, Bryan Geels calculated the rejection rates in the 2016 general election, 2018 general election, 2020 primary election, and the 2020 general election in Georgia using the absentee voter history file. In *Texas v. Pennsylvania*, an attached expert report from Charles Cicchetti used a combination of data from the Election Administration and Voting Survey (EAVS) and the absentee voter history file.

We obtained the absentee voter history file from Georgia's secretary of state's office to calculate the overall rejection rate for mail-in absentee ballots and the rate ballots were rejected for various reasons. The first row of Table 2 calculates the overall rejection rate for

[6] Cicchetti's claims in *Texas v PA* are based on the EAC's survey data from the state of Georgia.

mail-in absentee ballots for the 2016, 2018, and 2020 general elections and the 2020 primary election. We focus on this set of elections because these were the elections analyzed in Geels (2020) and includes all the elections analyzed in Cichetti's expert report.

The top of Table 2 shows that the overall rejection rate clearly declined in 2020 relative to recent prior rejections. In 2016, 2.90% of mail-in absentee ballots were rejected and in 2018, 3.46% of mail-in absentee ballots were rejected. This number declined first to 0.97% in the 2020 presidential preference primary and then 0.34% in the 2020 General Election. Focusing on the steepest decline, from 2018 to 2020 the overall mail-in absentee ballot rejection went down 3.12 percentage points.

But even though the overall rejection rate declined, the rate at which ballots were rejected for signature reasons remained essentially unchanged. In the second row of Table 2, we have calculated the share of ballots that were rejected for a signature related reason–either because the signature is mismatched or the signature is missing.[7] In 2016, this was 0.28% of all returned mail-in absentee ballots and in 2018 0.2% of all mail-in absentee ballots. In the 2020 general election the signature-related rejection rate declined, but only slightly to 0.15% of all returned mail-in absentee ballots. Relative to 2018 this was only a 0.05 percentage point decline, or only a small fraction—1.6%— of the total decline in mail-in ballot rejections.

A second way to examine the rate of signature related rejections is to focus explicitly on rejections due to a signature mismatch, rather than any signature issue, is mentioned. This will necessarily be a subset of the rejection reasons in the second row of Table 2.[8] When

[7] To make this determination we use the "STATUS_REASON" field in the Georgia Absentee voter history file. In 2020 there were a small number of predetermined reasons that official could select for their reason for rejecting an absentee ballot. In 2018 and 2016 officials had the option to type in their reason. We hand classified the reasons and our replication code provides details on this coding process. The vast majority of coding decisions are clear. In instances where more than one reason was offered and a signature reason was one reason offered, we classified the reason as signature based reason. This biases our findings towards identifying a decline from 2016 and 2018 to 2020.

[8] In 2016 the reasons for rejection included are: "SIGNATURE/NAME DOES NOT MATCH","SIGNATURE/DOB NOT A MATCH", "SIGNATURE NOT A MATCH", "SIGNATURE MATCH AND INCORRECT DOB", "Signature Match", "SIGNATURE DOES NOT MATCH", "SIGNATURE DID NOT MATCH", "SIGNATURE AND DATE OF BIRTH DOES NOT MATCH", "SIG NOT A MATCH, DOB NOT A MATCH", "SIG NOT A MATCH", "OATH & SIGNATURE NOT A MATCH FOR OATH", "CURRENT DATE FOR DOB / SIGNATURE NOT A MATCH". In 2018 the reasons are: "CURRENT YEAR AS YEAR OF BIRTH, SIG MATCH", "SENT SIG MISMATCH LETTER, DIDNT

	2016 General	2018 General	2020 Primary	2020 General
Overall	2.90%	3.46%	0.97%	0.34%
Signature	0.28%	0.20%	0.27%	0.15%
Deadline	1.38%	1.57%	0.70%	0.18%
Oath	1.06%	1.50%	0%	0%
ID/Ineligible	0.18%	0.05%	0.01%	0.01%
Other	0.01%	0.14%	0%	0%

Table 2: The overall rate of ballot rejections went down, but not because of a change in signature matching.

explicitly focusing on the signature rejections because of a stated match issue, we reach a similar conclusion: we cannot explain the decline in the mail-in absentee rejection rate based on the signature rejection rate. In 2016, we find that 0.21% of mail-in ballots were rejected for a signature mismatch. In 2018, 0.09% of mail-in ballots were rejected for a signature mismatch issue, and in the 2020 primary election, 0.08% of mail-in ballots were rejected for a signature match issue. This declined further in the 2020 general election to 0.05% of mail-in ballot rejected for a signature mismatch.

Signature-related rejections cannot explain why mail-in ballot rejections declined in Georgia in the 2020 election. Instead, the decline in rejections was due to fewer ballots arriving after the deadline and fewer ballots rejected for issues with the "oath" envelope. In Georgia, mail-in absentee ballots have to arrive before polls close on Election Day. In 2016, 1.38% of all returned mail-in ballots arrived after this deadline. This number rose to 1.57% in 2018. But in the 2020 general election, only 0.18% of all returned ballots arrived after the deadline. The decline in ballots rejected for arriving after the deadline accounts for 45.4% of the total decline in ballot rejections from 2018 to 2020.

In short, there were fewer ballots rejected in 2020 because mail-in absentee voters returned their ballots much earlier than in prior years. This is evident in Figure 3. This figure shows

FOLLOW", "SIG AND YOB NOT A MATCH", "SIG NON MATCH, ADDITIONAL ID NOT", "SIG NON MATCH; ADDITIONAL ID NOT", "SIG NOT A MATCH", "SIGNATURE DOES NOT MATCH", "Signature Match", "SIGNATURE MISMATCH", "YOB AND SIG NOT A MATCH". In both the 2020 primary and general election, we examine "Invalid Signature" status reasons.)

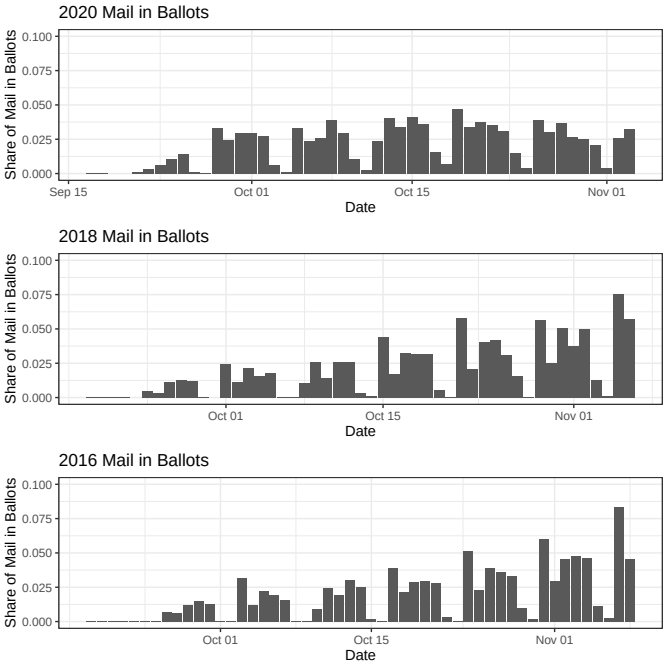

Figure 3: Mail-In ballots were returned earlier in the 2020 election.

the share of mail-in absentee ballots returned on each day before the general elections in 2020 (top facet), 2018 (middle facet), and 2016 (bottom facet). This figure shows that in 2020, a much larger share of ballots were returned earlier than in prior years. The median ballot in 2020 was returned 18 days before the election, but in 2018 and in 2016 the median ballot was returned only 13 days before the election. In 2016 and 2018, the day with the most ballots returned was the day before the election. In 2020, the day with the most ballots returned was October 19th.

A second reason for the decline is a change in Georgia election law. In 2019, Georgia passed HB 316. One result of this reform is that the information voters had to provide when returning their ballot was greatly simplified. This is referred to as the "Oath" envelope. Consequently, voters could no longer have their ballots rejected for failing to provide some information (such as their address) when returning their ballot. In 2018, 1.5% of all returned ballots were rejected for reasons related to this oath envelope. In 2020, this was reduced to 0%. This 1.5 percentage point decline accounts for 47.9% of the total decline in ballot

rejections.

HB 316 also provided voters with the opportunity to "cure" ballots when voters made a mistake—either because of a signature mismatch or because they failed to provide required information. An expert report from Ken Mayer in *Pearson v. Kemp* estimates that 403 ballots were cured during the election. The primary reason to cure a ballot would be a signature related issue (Mayer, 2020). If those 403 ballots were all due to a signature related issue, then the signature rejection rate in the 2020 general election would be 0.18% of all mail-in ballots, nearly identical to the rejection rate in the 2018 general election.[9] While there was a court case in late 2018 that enabled a small amount of ballot curing in 2018, there is no indication that there was ballot curing on the level that was found in 2020.[10]

3.1.2 Underage Voters in Georgia

Another startling claim from Trump's team was that voters younger than 18 cast votes in the 2020 election. In President Trump's speech delivered on the morning of January 6th, 2021 "Sixty-six thousand votes, each one of these is far more than we need. Sixty-six thousand votes in Georgia were cast by individuals under the legal voting age" (Trump, 2021b). He made a similar claim at the January 4th rally for Senate Republican candidates in Georgia, arguing that "66,000 votes in Georgia were cast by people under the legal voting age" (Trump, 2021a).

The claim comes from one of Trump's challenges to the election results in Georgia, *Trump v Raffensperger*. A press release from Trump's campaign issued on December 4th, 2020 alleged that "Data experts also provided sworn testimony in the lawsuit identifying thousands of illegal votes: 2,560 felons; 66,247 underage voters" (Trump, 2020g).[11] A similar claim is

[9] Using a separate methodology, another estimate of the number of ballots cured in the 2020 general election is 345 ballots (Swift and Gomen, 2021). To obtain this number, we took the total number of rejections for signature reasons discussed in this paper and then subtracted the final set of rejections. The result is 345 ballots.

[10] See: https://www.aclu.org/news/voting-rights/court-blocks-georgia-rejecting-ballots-over

[11] As we explain below, the actual complaint labels these voters as underage registrants, not underage votes.

repeated by State Senator William Ligon in a report published the previous day on December 3rd, 2020 (Ligon, 2020). Summarizing testimony from John Eastman, a constitutional scholar who was hired to assist with Trump's post-election legal challenges, Ligon stated that Eastman "reiterated failures such as counting the votes of approximately 66,000 underage individuals" (Ligon, 2020). The specific claim made about underage voters casting votes appears to be a reference to John Eastman's written testimony provided to the committee where he claimed that "66,000 underage individuals to cast ballots, which were then counted" (Eastman, 2020b).

But no expert report filed in *Trump v Raffensperger* claims to have evidence of underage voters. Instead, the expert report filed with *Trump v Raffensperger* claims to have evidence that individuals were underage when they registered to vote. This distinction is clear in John Eastman's verbal testimony to William Ligon's committee on December 3rd. At the hearing, John Eastman verbally testified that "the number of underaged individuals who were allowed to register contrary to statutory authority amounts allegedly up to 66,000 people" (Eastman, 2020a). This is a key difference—what Eastman is alleging in his testimony is that people were underage *at the time of registration*, rather than *underage at the time they cast a vote.*

Eastman's verbal testimony (but not his written testimony) aligns with the evidence presented in Bryan Geels' first report, which was attached in the complaint filed in *Trump v Raffensperger*. In this first expert report, filed on December 1st, 2020 , Geels opines "that there are 66,247 individuals whom the State's database identifies as having cast a ballot whose records indicate that they were registered to vote prior to their 17th birthday" (Geels, 2020). Later in the report Geels describes this group as "Voters Who Were Underage When Registered" (Geels, 2020). Geels then provides pseudo code to describe the match that was performed, which clearly shows that he was attempting to calculate the difference between what Geels thought was the date of registration and the registrants' birth year (Geels, 2020).

But even when interpreted correctly, Geels' claim is simply false. Charles Stewart, in a reply expert brief, first identified an issue with Geels' calculation (Stewart, 2020a). In his

expert report he explained that he was unable to verify Geels' claim, finding a dramatically different number than Geels (Stewart, 2020a). After Geels' initial affidavit, an updated report was attached to a second filing of *Trump v Raffensperger*. In this updated report filed on January 7th, 2021, Geels reported "Category (17): Underage Registrations. At least *2,047* individuals voted in the Contested Election who, according to the State's records, were registered to vote prior to their 17th birthday, below the minimum age permitted under the Election Code." (Geels, 2021). This is more than a 64,000 downward revision in the estimate. And there was no explicit reference to this revision—even though the 66,000 underage voter claim had been repeated by the Trump, the campaign's lawyers, John Eastman in his testimony in Georgia, and across public efforts to persuade state legislators to rescind electoral college votes from Georgia. In the January 7th report, there is no reference to the 66,000 figure found in the first report, no independent reference to underage voters, and no attempt to address Charles Stewart's criticism of the initial analysis.

Geels updated this analysis a third time, revising the number of alleged underage registrants to be even lower. In a report filed for John Eastman's disbarment hearing in California, Geels issued an updated estimate (Geels, 2023). In this report, Geels claimed there were "778 violations" once he used the `REGISTRATION_DATE` field in the voter file (Geels, 2023). Even if we accept this estimate as accurate (and we shouldn't, as we show below), this implies that Geels' initial estimate was more than a 8,300% overestimate.

But even if this last claim from Geels' claim were true, the way he wrote his query caused him to identify individuals who were substantially older than 18 during the 2020 election. This is because Geels' analysis looks across all Georgia registered individuals and calculates the difference between their date of registration and date of birth. This means that individuals obviously old enough to vote were labeled as "underage."

To evaluate Geels' methodology, we use a version of the Georgia voter file from December 2020, close to the date used in Geels' expert report. Consider one example in the data set. One voter was born in 1928 and, according to Geels' calculation was 15 years old when they

were added to the voter file in 1943. Despite Geels considering this person "underage" at the time of voting, they were 92 years old in the 2020 election. In fact, in our replication of Geels' procedure, we identify 2,116 individuals based on the code outlined in Geels' report whom he would classify as having been underage at the time of registration. Among these individuals, the average age in 2020 was *42 years old.*

When we execute Geels' query, we find no evidence of underage voting or underage registration in Georgia in 2020. Based solely on the voter file and the voter history file, there are three individuals who were supposedly younger than 18 at the time of the election. But further investigation of all three voters reveals that they were all over 18 years old during the 2020 election.[12]

Similarly, we find no evidence of systematic underage registration in Georgia. Geels' analysis uses the "DATE ADDED" file, which is not necessarily the date that a voter was registered to vote in the State of Georgia. Once we analyze individuals' registration date, we find that 1,063 individuals have registration dates that are 17 years or less after their birth year. Once we eliminate individuals registered after 2002—when identification was required to comply with HAVA—another 147 individuals were eliminated as having been an underage registrant.

Among the remaining 906 individuals, there is clear evidence of clerical errors. Almost half (440) have the same date added and birth year in the data set—an indication of an error made when adding the voter to the voter file. In fact, 87% of the supposedly underage voters were 14 or less according to Geels' query.

In summary, there is no evidence of systematic underage voting or registration. There is some evidence of a small number of clerical errors in the Georgia voter file. This is not evidence of fraud or illegal voting. Instead, it represents the challenges of maintaining a voter file across a heterogeneous state.

[12]One voter has a birth date of 2020, but was added to the voter file in 1995. A second voter has a birth date in 2018, but was added to the voter file and registered in the same year, 2018. We were able to identify a public record of this individual and they were born in the 1940's. A final voter is listed with a birth year of 2003, but public records (based on an arrest) indicate that they were born in 2002.

3.1.3 Early Application for Absentee Ballots in Georgia

Another claim about mail-in votes in Georgia is that many ballots were mailed earlier than allowed. The origin of this claim is Brian Geels' expert report in *Trump v Raffensperger*. There he says that "305,701 individuals have records indicating that they applied for absentee ballots more than 180 days prior to the general election (i.e., prior to May 6, 2020) exceeding the statutory maximum according to state law" (Geels, 2020, paragraph 13). Geels' reference to the Georgia code, however, fails to mention that Georgia state law explicitly allows for individuals over 65 and individuals who are self-identified as disabled to apply once for a mail-in absentee ballot.[13] To implement this portion of the law ,the mail-in ballot applications in 2016, 2018, and 2020 allowed applicants to declare themselves over 65 or disabled and therefore automatically receive an absentee ballot for the general election after the primary election.

Almost all of the individuals who applied early for an absentee ballot in Georgia was over 65 (Stewart, 2020a). Using the absentee voter history file from the 2020 election for the state of Georgia, we find that among individuals whose applications were filed more than 180 days before the election, 92.8% were over 65 years. The remaining 7.2% of individuals who applied early are easily explained as individuals who declared themselves to have a disability and therefore eligible under the provisions of the Georgia law.

This claim simply fails to understand Georgia law and is therefore in error. Rather than evidence of a conspiracy to swing the election or an illegal attempt to avoid Georgia law, the applications alleged to be outside of the statutory window actually comply with Georgia election law.

[13] 21-2-318 1. G. "Any elector meeting criteria of advanced age or disability specified by rule or regulation of the State Election Board or any elector who is entitled to vote by absentee ballot under the federal Uniformed and Overseas Citizens Absentee Voting Act, 42 U.S.C. Section 1973 ff, et seq., as amended, may request in writing on one application a ballot for a presidential preference primary held pursuant to Article 5 of this chapter and for a primary as well as for any runoffs resulting therefrom and for the election for which such primary shall nominate candidates as well as any runoffs resulting there from.

3.1.4 Early Distribution of Absentee Ballots in Georgia

The Geels report also claimed that some ballots were distributed earlier than it was legally allowed for ballots to be sent out. Geels claims that 2,664 mail-in absentee votes "were issued before they were eligible to be sent out on September 15, 2020" (Geels, 2020). We examined these claims using the 2020 absentee voter history file for the state of Georgia. Grady County in Georgia accounts for 40% of these ballots, with almost all of those ballots recorded as issued on September 8th, 2020, a week earlier than allowed. There are similar apparent clerical anomalies in other counties with early ballots. For example, in Fulton County, where 17% of the supposedly early ballots are located, nearly all of the ballots were supposedly issued on August 19th or August 20th, 2020. And every individual who supposedly had their ballots issued on these dates applied the day before, right as mail-in absentee ballot applications in Georgia opened for the general election. The most likely explanation is that a clerk incorrectly recorded the issued date for these ballots. Hart County is the county with the next most ballots issued. There, for 330 applicants, the application date was erroneously set equal to the issued date, making it seem like applications were made before ballots could be distributed. We last consider Fayette County, the last county that supposedly issued more than 100 ballots earlier than allowed under Georgia law. Fayette County, like Fulton, erroneously recorded ballots as being issued during the start of the absentee application process.

Regardless of when the ballots were issued—whether it was a clerical error in the absentee voter file or an error in when the ballots were distributed—none of the ballots in any of the counties are recorded as returned before September 20th, after the official dates ballots could be released. As a result, it is hard to imagine how the early release of ballots–if they actually occurred–advantaged any candidate.

3.1.5 Claims Made About "Suspicious" Registration in Georgia

A different claim made about the Georgia registration file is that there are a large number of individuals with registration records that Trump's team classified as "suspicious". In Bryan Geels' first expert report he claims that there are "6,635 individuals whom the State's database identifies as having cast a ballot whose records indicate that they were added to the voter registration file subsequent to the 2016 general election (since they weren't listed in the 2016 Voter Registration File)" (Geels, 2020, paragraph 25). But Geels revised this claim in the expert report he submitted as an update in *Trump v Raffensperger* filed on January 7th, 2021. In this updated report, Geels claims that "86,880 individuals who voted in the Contested Election were listed as having been added to the voter registration file prior to 2016, but were not listed in the 2016 Voter Registration File" (Geels, 2021). In Geels' updated expert report he provides no explanation for the 1,209% increase in his estimate of the number of voters in this category.[14]

We examined Geels' report, along with copies of the Georgia voter file and the 2020 voter history file and found that the increase in the number of suspicious registrations occurred because Geels changed the criteria for determining if a voter had a "suspicious" registration. We were able to make this assessment because Geels attached to his updated report a pdf print out of the voter records of all 86,880 individuals he classified as having a suspicious registration record, along with specific information about the queries he ran. We digitized the records in the exhibit and successfully recovered the 86,880 individuals that Geels reports in his updated report.

Using these digitized records, we find that Geels' number of "suspicious" registrations dramatically increased because he changed how he ran his query, resulting in many individuals with registration records that were not suspicious at all. Initially, Geels' first expert report used the "REGISTRATION DATE" field to identify voters he thought should be in

[14]It is the case that the first report only examined individuals who voted mail-in absentee and the second report examined all voters, but we show in our replication code that this cannot even come close to explaining the difference between the two reported numbers.

the 2016 voter file. But then in the second report, Geels uses the "DATE ADDED" field in the voter registration file to determine when an individual was registered to vote. He uses this field because of a "discussion with an employee in the Georgia Secretary of State office, the Date Added is the right field to be using for the registration date, because it will show the date the voter was added to the list" (Geels, 2020, Paragraph 47).

We find that 6,584 individuals in Geels' data set cast a vote by absentee ballot and have a registration date prior to November 8th, 2016, according to the "REGISTRATION DATE" field. If we include individuals who voted in person but continue to restrict the sample to individuals with a registration date prior to November 8th, 2016, we find that Geels would identify 8,785 voters. Only when Geels alters the query to focus on the "Date Added" field, rather than the registration date, does he find the 86,880 voters whom he claims have suspicious registrations.

We inquired with the Secretary of State's office about the registration file and learned that there are many reasons why voters might have a "DATE ADDED" field that is earlier than 2016 but not appear in the 2016 voter file. In fact, any time a voter cancels their registration and subsequently registers, they will have a registration date later than their date added to the voter file. One reason this can occur is if a voter leaves Georgia, cancels their registration, and then returns to the state and re-registers. Further, if a voter is convicted of a felony, their registration is canceled and they are removed from the registration list. But after completing their sentence, the voter is then able to have a valid registration again. When this occurs the registration date will be updated, but the "DATE ADDED" field will remain unchanged.

We also see clear evidence that these are not "phantom" individuals added for the 2020 election. Rather, they are voters who are restoring previously canceled registrations. To make this assessment we used a version of the Georgia voter file from January 2019. Among the 8,785 voters who Geels' identifies as having suspicious registrations 4,743 were in the January 2019 voter file–54.0% of individuals. And even if we focus on the entire 86,880

voters, we find that 45.5% of the identified voters appear in the January 2019 voter file.

In short, there is nothing suspicious about these registrations. Geels altered the query to inflate the number of suspicious registrations in his updated affidavit.

3.1.6 Claims Made About Dead Voters in Georgia

Another claim made in Georgia and many other states is that ballots were fraudulently cast in the name of deceased voters. For example, at a rally in Georgia on January 4th, 2021 Donald Trump declared that "We were up 10,315 ballots were cast by individuals whose name and date of birth matches a Georgia resident who died in 2020 prior to the election. Then your wacky secretary of state said two people, two people" (Trump, 2021a). This corresponds with testimony from Trump's legal team. For example, Ray Smith, Trump's attorney in *Trump v Raffensperger*, testified there were "10,315 *or more* individuals" who voted but "were deceased at the time of the election" (Smith, 2020). And later, John Eastman declared on January 6th in his speech at the Elipse that "we know dead people voted" (Eastman, 2021).

The claims about dead voters in Georgia come from Brian Geels' expert report in *Trump v Raffensperger*. Geels claims "that there were as many as 10,315 individuals who cast ballots in the November 3,2020 election that were accepted and counted but who deceased prior to Election Day. 8,718 of these individuals are recorded as having perished prior to the date the Sate records accepting their ballot" (Geels, 2020, Paragraph 28).

To calculate this number, Geels matched a list of voters who cast a mail-in absentee ballot in the 2020 election with a list of deceased individuals he obtained from the state of Georgia. To match across the two lists, Geels matched on First Name, Last Name, and year of birth.

This matching procedure, however, will lead to a large number of false positive matches, causing Geels to claim voters are dead when they are still alive. Geels recognizes that his procedure will lead to false positives, arguing that "Because the Voter Registration file only

contains the Birth Year for each registered voter, a more exact match cannot be made and there may indeed be false positives" (Geels, 2020, paragraph 50). Indeed, because many first name, last name, and birth year combinations in Georgia are common, we should expect a large number of false positives.

To assess how often false positives will occur, we examined the share of registered voters in the December 2020 Georgia voter file who share a first name, last name, and birth year with at least one other registered individual. We find that about 5.2% of all first name, last name, and birth year combinations in the December 2020 Georgia Voter File are not unique. The lack of information provided by Geels suggests that the vast majority of his matches are false positives.

Despite public pronouncements that dead voters cast ballots in Georgia, communication among Trump's advocates revealed they held doubts about the accuracy of their estimates of the number of dead voters in Georgia. In a January 4, 2021 email released as part of the Congressional January 6th committee, an unnamed expert reported serious problems with estimates of the number of dead voters in Georgia. In an email addressed to "Mayor" (presumably Rudy Giuliani) the redacted author explains that that there was little evidence of deceased voters illegally cast a ballot in the 2020 election. The redacted author of the email explains that "many of the dead voters on the GA list sent their vote in before they passed. I don't think makes a particularly strong case, and I think it is possible that Chairman Graham [Senator Lindsey Graham] will push back on that. Almost all of them died in October 2020, a few in November 2020, and a few in September 2020, I think this makes the case for unfortunate timing – many sent their ballots before they passed- rather than nefarious activity. Am raising this just so that everyone is aware of what the data actually says" (Redacted, 2021). The email confirms what Geels already conceded in his expert report: there was little credible evidence of many dead voters casting ballots in the 2020 election.

On January 6th, Brad Raffensperger, Georgia's Secretary of State, explained in a letter

to Congress that there was no evidence dead voters cast ballots en masse in the 2020 election. Raffensperger cited the claim from the Geels report that "10,315 dead people voted" and then went on to explain that "Our office has discovered 2 potential dead voters and both instances are under investigation" (Raffensperger, 2021).

Despite the public claims made from the Secretary of State, at least one election integrity group in Georgia continues to claim that dead people cast ballots in the 2020 election. But this group's claims are not based on any publicly verifiable analysis and instead claim to use data that this group cannot possibly have. In a letter allegedly refuting statements from Georgia's Secretary of State, Voter GA, an election integrity group in Georgia, responded to Raffensperger's claims before Congress, arguing that "The actual evidence shows that at least 873 people received credit for voting in the November 2020 election although they had died in 2020 prior to Election Day. This is based on the voter history file from November 28, 2020 and a deceased individuals file from the State of Georgia's Public Health Department. These files were matched on First Name, Last Name, Date of Birth and Address" (Favorito and Coovert, 2022). It is completely unclear what data and analysis this claim is based on. The refutation from Voter GA links to a version of Geels' expert report: the updated report attached to the revised complaint in *Trump v Raffensperger* filed in January 2021. But Geels' expert report only matches from the voter file to the deceased individuals list using First Name, Last Name, and birth year—a set of criteria less stringent than a match based on First Name, Last Name, Date of Birth and Address. But it isn't even clear how Voter GA could perform this match. Date of birth is not available in the public voter file. And it isn't clear what address is available in the deceased individual list.

We have further reason to doubt that Voter GA actually identified illegal votes from deceased voters, because the numbers they cite correspond to the number of ballots cast from individuals who were alive and then died after casting their ballot. We can make this assessment using the Georgia canceled voter file, the Georgia voter history file, and the absentee voter history file. The canceled voter file contains information about when

individuals were removed from the voter registration file, why they were removed, and a voter's unique registration number. This enables us to exactly identify voters removed from the roles for being deceased and then to assess whether they returned their mail-in absentee ballot before being removed from the roles.[15]

Using the canceled voter file, we find no evidence of widespread voting from already dead voters. In fact, as the email to Giuliani made clear, almost every instance of a "dead" voter was someone who legitimately cast their ballot before dying. Using the canceled voter file and the absentee voter history file, we find that there are 871 people who voted in the 2020 election and whose registration was canceled before November 15th, 2020 because they died. This number is strikingly similar to the claimed number of dead voters in the Voter GA refutation letter. First, we find that 77 individuals who voted in the 2020 general election and then had their registration canceled for being deceased cast their ballot in person—either on Election Day or early in person. Among the remaining 794 individuals who cast their ballot using mail-in absentee, 787 of those individuals had their ballot returned before they were removed for being deceased. This means 99.1% of the supposedly dead voters are easily identified as legitimately cast ballots. This group of individuals returned their ballots, on average, 23 days before being removed from the voter file for being deceased.

This leaves 7 remaining voters who were removed on the day their ballot was accepted or had their ballot accepted after they were removed from the voter file–constituting approximately 0.0001% of all ballots cast in the 2020 election. One voter had her ballot accepted the same day she was removed and, according to public obituaries, died approximately five days before her ballot was received by the state and therefore could have plausibly voted her ballot before dying. Of the remaining 6 individuals, it appears plausible, given the timing, that two individuals could have been alive when they cast their ballots and then died. This leaves approximately four individuals with ballots cast after they were decreased with a large enough gap to warrant further investigation.

[15]To obtain the canceled voter file we filed a Freedom of Information Act (FOIA) request with the Georgia Secretary of State's office.

Our estimate of approximately four coming from voters after they were deceased is in line with other estimates of dead voters, for example (Wu et al., 2020) find similarly low levels of ballots from dead voters. And regardless of the explanation for these four potential dead voters, this is hardly the 10,315 illegal dead votes Trump and his allies claimed illegally cast ballots in the 2020 election.

3.1.7 Claims Made About Felony Voters in Georgia

Trump and his advocates also claimed that voters currently serving felony sentences were allowed to cast votes in Georgia. In his speech in Georgia on January 4th, 2021 Trump claimed that "2,506 ballots were cast by individuals whose name and date of birth matches an incarcerated felon in a Georgia prison" (Trump, 2021a). Similarly, a press release announcing the *Trump v. Raffensperger* litigation from the Trump campaign pronounced that "Data experts also provided sworn testimony in the lawsuit identifying thousands of illegal votes: 2,560 felons" (Trump, 2020g). In testimony to a Georgia state legislative committee Ray Smith declared that 2,506 felons were allowed to vote in the 2020 general election (Smith, 2020).

This claim also comes from Brian Geels' expert report in *Trump v Raffensperger*. Geels reports that "there could have been up to 2,560 individuals who cast a ballot that were accepted and counted but who were inmates" (Geels, 2020). Geels goes on to explain that "2,560 represents the number of individuals in the Active Inmate File whose First Name, Last Name, and Birth Year matched exactly to the Voter History file" (Geels, 2020). But Geels did warn that, "similar to the analysis in #17 for deceased individuals, only the Birth Year is available in the State's voter record files. This analysis therefore may also contain false positives due to the imperfect nature of the match."

Geels does not provide an estimate of the false positive rate, nor have we been able to find the list of individuals he identified as felons who voted in the 2020 election. Instead, to study the rate of false positives we investigate the number of duplicate matches that

we would expect if we merged the Georgia voter file using first name, last name, and an individual's birth year. We find that false positives will occur at an extremely high rate.

To obtain a baseline rate of false positives from the Georgia voter file, we simulated a matching exercise where we know the correct answer. Specifically, we first created a synthetic set of names to "match" by randomly selecting a subset of Georgia registered voters. We then proceed as if this is a list of voters we wanted to match to the voter file. Using our sampled subset, we then assess our rates of false positives when we attempt to match our sampled subset to the original voter file. To determine the size of the group of voters we match to the Georgia voter file, we use the actual number of felons in Georgia during the 2020 election. According to Reform Georgia, an advocacy group for the restoration of felon rights, 266,317 Georgians were ineligible to vote in 2018 due to a felony conviction.[16] So we examine the rate of false positive matches if we were to randomly select 266,317 voters to match the remaining voter file to match against. We also examine the rate of false positives if we focus on the individuals the state of Georgia was actively incarcerating at the time of the election: 54,806 individuals.

Regardless of the estimate of the size of the felon population used to form a match, we find large numbers of false positives if we match on First Name, Last Name, and Birth Year alone. For example, if we examine the entire population of individuals serving a felony sentence we find approximately 32,000 false positives when matching against the Georgia voter file using first name, last name, and birth year. If we examine those actively incarcerated we find approximately 6,750 false positives. This high rate of false positives isn't surprising, because 5.2% of all individuals who voted in the 2020 general election have a non-unique first name, last name, birth year combination. In other words, over 5% of the voters in Georgia can contribute to a false positive when matching against records.

In short, the procedure Bryan Geels used to identify potential felon voters was guaranteed to generate false positives at a high rate. So the evidence he presented is not indicative fraud

[16] Statistics on the felon population is available here: https://www.senate.ga.gov/committees/Documents/ReportonFelonyD

or illegal voting.

3.2 Claims Made About Statistical Evidence of Voter Fraud in Georgia

In addition to making specific allegations of voter fraud, such as underage, deceased, or felon individuals voting, Trump and his allies also claimed that statistical patterns imply the existence of systemic voter fraud. The statistical claims cover a wide variety of voting results, allegations, and theories of how the fraud was conducted. Yet, they tend to follow the same pattern. There is some assertion made about what would happen in a normal election, perhaps based on results from prior elections. Then, those alleging the fraud suggest the results of the 2020 election are statistically surprising or improbable given their assertions about what is expected, and they conclude that the election results must be fraudulent.

Such claims are found not only in public speeches, but also in expert witness reports from the court cases challenging the Georgia election results, *Pearson v. Kemp* and *Wood v. Raffensperger*. The reports have the appearance of science, as they analyze big data sets and use statistical language like predictive modeling and probability distributions. But upon inspection, the analyses are clearly flawed. They misinterpret data sources, rest upon unjustified statistical assumptions, and employ dubious methods which render their conclusions implausible and often useless. The arguments lack basic statistical rigor and are, at times, incomprehensible. They also demonstrate how insufficient domain knowledge about election data and voting patterns can lead to spurious conclusions. Many conclusions appear to be based on the intuitions the scholars have what statistical patterns look like in a "normal" election. Yet, none of the authors of these reports have any prior experience analyzing election data. So, it is unclear what the basis for their claims about what a typical election would look like were.

3.2.1 Claims Made About Anomalous Vote Distributions

Dr. Eric Quinnell's report in *Pearson v. Kemp* makes the startling claim that 32,347 Biden votes in Fulton County are statistically "anomalous". To make this conclusion, Quinnell claims there is something anomalous about the distribution of Biden's vote totals. Quinnell suggests that we "often expect our data to be close to a normal distribution" (Quinnell, 2020), and thus, if voting distributions deviate from the normal curve, this "can indicate an event that is statistically unlikely"(Quinnell, 2020, paragraph 6). Upon observing that the distribution of Biden's vote gains in Atlanta precincts, calculated as the difference between Biden and Clinton's vote totals, is more right-skewed with a fatter right-tail than the distribution of Trump's vote gains from 2016 to 2020, Quinnell concludes that Biden's performance is statistically anomalous.

Building on this observation, Quinnell then estimates the specific number of suspicious votes using a simple procedure he appears to have invented for his expert report. He adjusts Biden's vote total in each Fulton County precinct to fit Clinton's vote share in the same precinct and then sums the differences between Biden's actual vote totals and these "predicted values". This produces 32,347 statistically anomalous votes in Fulton County.

There is no basis for Quinnell's invented procedure, nor his conclusion that 32,347 ballots were anomalous in Fulton County. There is no statistical or social scientific reason why changes in vote totals across elections at the precinct, district, or county level should be normally distributed.[17]

Simple analyses and basic logic explain why Biden's distribution is more right-skewed than Trump's, and why it has a fatter right-tail. Biden outperformed Clinton's vote share

[17]There are a multitude of other issues with Quinnell's analysis. For example, as Jonathan Rodden documents in his expert report (Rodden, 2020), Quinnell's approach ignores the fact that precinct boundaries often change across elections. In Fulton County, 82 out of the 342 precincts in 2016 had their boundaries redrawn before 2020 (Rodden, 2020, paragraph 17). Analyzing the precinct-level differences in vote totals across elections is thus meaningless, as you cannot make "apples-to-apples" comparisons in over 20% of precincts. Quinnell does acknowledge that a few precincts were redrawn, but he misses the vast majority of them and does not adjust his methods or conclusion in any way. He simply compares Biden's and Clinton's precinct-level performances, despite many of these precincts no longer being comprised of the same voters. Reliable inferences, let alone evidence of fraud, cannot be drawn from such an analysis.

in the densely-populated and growing suburban counties of Georgia, while Trump primarily gained in Georgia's sparsely-populated and shrinking rural counties (Rodden, 2020, 14). This pattern holds at the precinct-level in Fulton County as well. Biden's largest gains came in the suburban precincts, which experienced the greatest growth in registered voters between 2016 and 2020, while Trump's vote totals largely remained stagnant in these precincts (Rodden, 2020, 22). If a candidate's vote share increases in populous and/or growing areas, this produces greater changes in raw votes than a candidate whose vote share increases in small and/or shrinking areas. This trend thus adds greater mass to the far-right of Biden's distribution than Trump's, as a subset of precincts experience outsize Biden vote gains. It also clarifies why Quinnell's assumption that Clinton's and Biden's precinct-level vote shares should be equal, which leads to him identifying over 32,000 anomalous votes, is unjustified. Clearly, vote shares fluctuate based on the groups a campaign mobilizes (in this case, suburban voters) and their spatial distribution. There is no logical link between a particular vote change distribution and fraud. The supposedly "anomalous" distribution that Quinnell identifies is perfectly explainable by geographic voting patterns.

3.2.2 Claims About Dominion Voting Machines

Another statistical claim made in expert reports from Trump's challenges in Georgia was that Joe Biden performed particularly well in locations where Dominion voting machines were used. Perhaps the clearest articulation of this claim is from Russell Ramsland's expert report in *Wood v. Raffensperger* and *Pearson v Kemp*, where he argues that Dominion and Hart voting machines contributed to Biden's success (Ramsland, 2020). Ramsland regresses 2016 US county votes on unspecified "census variables" and uses this model to predict Biden's "expected" vote share. He finds that Biden outperforms the model in 78% of the counties which use Dominion or Hart voting machines, and argues this is statistically indicative that "something strange is occurring", because "in normal circumstances any candidate should perform above expectations roughly 50% of the time" (Ramsland, 2020, 11). Ramsland

computes that Biden performed 5-6% above expectation in such counties, which, in Georgia, leads to at least 123,725 "statistically invalid" (Ramsland, 2020, 10) votes for Biden that "must be disregarded when tabulating the election results" (Ramsland, 2020, 18).

Ramsland makes erroneous statistical assumptions and arguments that, once addressed, nullify his claims about illegal voting. First, Ramsland is implicitly assuming his model will produce unbiased predictions of vote totals in a free and fair election (Angrist and Pischke, 2009). Yet, Ramsland cites no evidence that this claim is true. In fact, he doesn't even demonstrate it is true within his training data, let alone for the 2020 election. It is well known that across elections things change. Support for candidates might swing or one campaign might be more effective at mobilizing sympathetic voters. This swing will cause vote totals to shift, which make predictions from prior elections inaccurate. So on a basic level, Ramsland simply fails to establish that his method is valid.

Second, and most important, Ramsland mistakes correlation for causation. Counties using Dominion or Hart machines in 2020 differ from those that did not. Most notably, they were much more likely to vote for Clinton in 2016. For example, over 40% of the counties Clinton won in 2016 election used Dominion voting machines in the 2020 election, while just 25% of the counties Trump won in 2016 used Dominion machines in 2020 (Rodden, 2020). To identify causal evidence that Dominion/Hart machines increased Biden's county-level vote share, one must ensure that Dominion/Hart counties do not differ from other counties in all the other factors that affect Biden's vote share. Ramsland makes no attempt to correct for such confounding factors. That Biden's vote shares were on average higher in Dominion/Hart counties is not evidence of machine-related fraud. In this case, it is simply driven by the fact that these counties were on average more Democratic in the first place.

Eggers et al. (2021) demonstrate that more appropriate statistical models fail to find any relationship between Biden's performance and the type of machine used in the county. Eggers et al. (2021) use simple statistical strategies like fixed effects to adjust for unchanging differences across counties. After including this information, Eggers et al. (2021) fail to find

any effect of Dominion or Hart machines on Joe Biden's support.

3.2.3 Claims Made About the Timing of Absentee Ballots Counted

Other reports claim there were anomalies in how absentee ballots were received in certain Georgia counties. For example, Quinnell and Young's report in *Wood v. Raffensperger* and *Pearson v Kemp* allegedly discovers anomalous patterns in how absentee ballots were returned in Fulton County (Quinnell and Young, 2020a). They claim to discover patterns in the data that "fail basic quality and sanity checks mathematically" and are only explainable by some exogenous "event or constraint... outside mere voting pattern[s]" (Quinnell and Young, 2020a, 33).

Quinnell and Young (2020) identify three specific patterns they allege are anomalous. First, they claim it is anomalous that approximately half of absentee ballots were reported in the first batch of ballots from Edison and approximately half were reported later. They claim this an "exceptional mathematical anomaly" (Quinnell and Young, 2020a, 27).

This claim, however, is not based on any actual analysis, reference to prior election data, or citation to prior work. Rather, Quinnell and Young (2020) merely have the impression that this is anomalous and, based on that impression, declare the result to be strange. There are many legitimate reasons that this pattern could occur in Edison's return data. But regardless, whether or not the pattern is anomalous is irrelevant, because Edison's batch updates do not correspond to the timing and cumulative totals of official vote tabulations (Rodden, 2020). For example, official records from the Georgia Secretary of State's office show that nearly 99% of total absentee ballots were received by November 3rd, but Edison's first update on November 4th contained only half of the total absentee ballots (Rodden, 2020, 5). As a result, it is impossible to attribute this pattern in Edison's data to anything fraudulent in official vote counts.

Second, Quinnell and Young (2020) claim to discover "unexplainable" partisan differences in the proportion of total absentee ballots reported in Edison's first timestamp (Quinnell

and Young, 2020a, 14). Quinnell and Young's analysis is based on differences in the share of ballots Trump and Biden had reported after the first Edison time stamp. Quinnell and Young (2020) find suspicious that Trump had many precincts with 100% of his total absentee counted in Edison's first timestamp, but Biden had none above 71%. Quinnell and Young (2020) also find it odd that the share of ballots counted for Biden doesn't follow a normal distribution.

Quinnell and Young (2020) continue to make the same error: they use their own impressions of what is likely in a fair election rather than conducting rigorous analysis. Quinnell and Young's supposedly anomalous patterns are, in fact, consistent with precincts randomly sampling ballots to count. The reason there are many more precincts with 0 or 100% Trump ballots counted is that there are many fewer Trump mail in absentee ballots to count in Fulton County precincts (Rodden and Marble, 2020). Because there are many fewer Trump ballots cast, it is much more likely that either none or all of the Trump ballots in a precinct would be counted. The intuition here is simple. Suppose we flip a coin with a 50% chance of obtaining a heads and a 50% chance of tails. If we flip the coin once, obviously we will either get all heads or all tails. If we flip the coin twice, we have a 25% chance of obtaining two tails and a 25% chance of two heads. So, 50% of the time we will have all heads or all tails. As the number of flips increase, the probability of obtaining all heads or tails will decline. After three flips the probability of obtaining all heads or tails declines to 25%. If we include ten flips the probability of all heads or all tails declines to $\frac{1}{2^9}$. As the number of coin flips increase, the probability of getting all heads or tails will continue to drop.

This same basic idea is one explanation for why there are precincts where Trump had 100% of his ballots counted, but Biden had no precincts with all of his ballots counted. There are 82 precincts where Trump received 5 or fewer votes from mail in ballots, but only 15 precincts where Biden received 5 or fewer votes from mail in ballots. So if a coin flip determines whether a ballot is counted, Trump has a higher chance than Biden of receiving all heads (e.g. all ballots counted) or all tails (no ballots counted) in many precincts, because

Trump has fewer coin flips (i.e. fewer ballots to count). A simple simulation where each ballot in each precinct has a 50% chance of being counted shows that Trump has all his ballots counted in approximately 3.2% of precincts, while Biden has all his ballots counted in only 0.2% of precincts. This means that, on average, we expect less than one precinct to have all its Biden ballots counted.

Relatedly, Quinnell and Young also appear concerned that the distribution of shares of mail in ballots counted do not follow a normal distribution. Quinnell and Young argue that the pattern they observe "represent some kind of major external constraint or non-linear event that is corrupting the fidelity of the data" (Quinnell and Young, 2020a, 32). They go on to assert that "such a signature would certainly block any attempts of any device going to production until a root cause could be found to describe the phenomenon and prove it either true, safe, or discarded" (Quinnell and Young, 2020a).

Yet, Quinnell and Young never explain what device they are referring to, what production they envision occurring, or what a "linear event" (as opposed to a "non-linear" event) would mean for election data. Nor do they explain why any of the patterns they observe are indicative of illegal or fraudulent voting. In fact, it appears their assertions are based on a very basic statistical misunderstanding. In a reply expert report, Jonathan Rodden and Will Marble explain there is no reason to expect the shares of votes counted to follow a normal distribution (Rodden and Marble, 2020). In a response to this critique, Quinnell and Young assert assuming normality is the " 'standard method' in all of sciences that uses comparisons against the 'standard normal' until proven otherwise. Note they are named 'standards' because they are standards" (Quinnell and Young, 2020c, 3).

This is *not* why the standard normal distribution is called the standard normal distribution. As any high school statistics student can tell you, the standard normal distribution is a "standardized" normal distribution. That is, once a normally distributed random variable is subtracted by its expected value and divided by its variance, it follows a *standard* normal distribution. There is no literature in statistics or political science where the normal

distribution is the default. Rather, distributional assumptions are derived from principles or data. This basic misunderstanding undermines nearly all of Quinnell and Young's claims and reveals a profound ignorance of basic statistics.

Third, Quinnell and Young (2020) are concerned that precinct's cumulative tallies at each timestamp are correlated with each other, as it suggests that absentee votes of all precincts are "centralized and coordinated" (Quinnell and Young, 2020a, 26). Yet, Rodden and Marble (2020) explain there are numerous reasons why there would be correlated shares of ballots counted across precincts. Certainly, Quinnell and Young (2020) offer no evidence that correlation in precinct counts imply the presence of fraud. Yet again, Quinnell and Young (2020) use their own poorly informed impressions of what should happen when votes are counted as evidence of fraud or illegality.

In summary, Quinnell and Young's (2020) report is incoherent as a statistical argument. They appear to not grasp basic statistical arguments. They base most of their tests on intuition rather than data, argument, or well documented common understanding. And the concepts they do apply appear to be imported from other fields without even a preliminary attempt to justify their application.

3.2.4 John Eastman's False Allegations of Machine Fraud During the Senate Runoff Election

Speaking to Trump supporters outside the Capitol Building with Rudy Giuliani by his side just hours before the January 6th riot, John Eastman claimed there were fraudulent ballot dumps in the Georgia Senate run-off elections between Jon Ossoff and David Perdue and Raphael Warnock and Kelly Loeffler. He argued that ballots were placed in a "secret folder" of voting machines, and at the close of polls, when the machines "know how many [votes] they need" as well as who has yet to vote, enough fake ballots from yet-to-vote voters were conjured to "barely get over the finish line" (Eastman, 2021).

His evidence for this fraud comes from his recollection of how votes were counted on

election night. At Trump's January 6, 2021 rally on the ellipse, John Eastman explained that when he was watching election returns he saw that "when it got to 99% of the vote total, and then it stopped. The percentage stopped, but the votes didn't stop" (Eastman, 2021). He went on to explain that "the data shows that the denominator, how many ballots remain to be counted...That number started moving up. That means they were unloading the ballots from that secret folder, matching them to the unvoted voter, and voila. We have enough votes to barely get over the finish line" (Eastman, 2021).

Eastman's allegations and supposed evidence are demonstrably false and fail to understand the basics of how election night reporting of votes works. Most importantly, Eastman appears to misunderstand what is reported on election night. News organizations, like Edison media or the Associated Press, provide an *estimated* number of votes cast. As the name suggests, this is not an official number of the number of ballots cast nor is it even an estimate from official sources. Rather, media organizations "start by estimating the overall expected turnout in a state based on a variety of factors, such as turnout in recent elections and details on early voting. This is done as a comparison to the turnout in a past election" (Staff, 2022) along with information about ballots that have already been returned and estimates about the number of votes cast and returned on Election Day. But the media organizations caution that "our reporting of expected vote is an estimate. It can and will fluctuate as the vote tabulation takes place and we learn more about how many people have actually cast a ballot" (Staff, 2022). So if media organizations underestimated Election Day turnout, then the estimated number of ballots cast increase while the ballots are counted after polls close.

A second reason Eastman's allegations fail is that he is simply wrong about what happened: Eastman does not accurately describe what happened on election night with estimated vote totals. To make this determination we examined updates to the vote counts and estimated votes cast posted by the New York Times for the Raphael Warnock and Kelly Loeffler runoff election. The data contains the reported vote counts for Loeffler and Warnock at particular times in the evening, along with the estimated total number of votes cast at

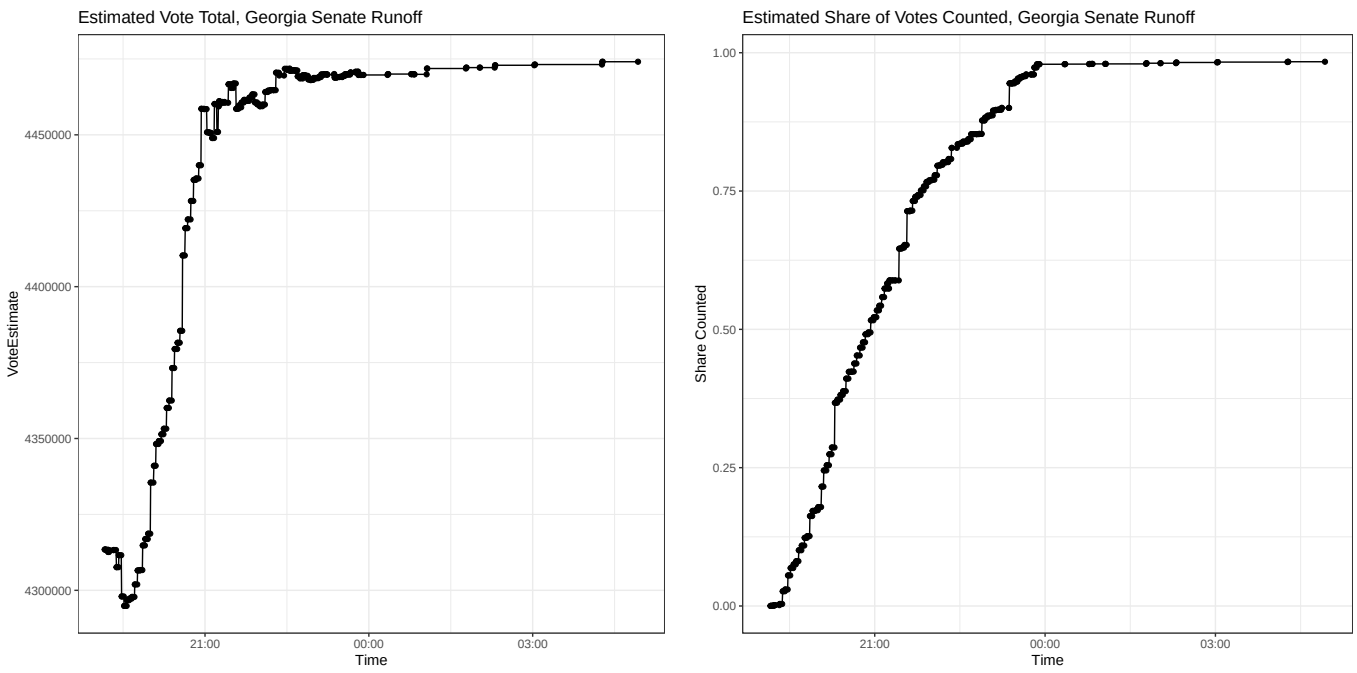

Figure 4: Examining how votes were actually reported on election night directly contradicts Eastman's supposed evidence of fraudulent votes in the Georgia runoff election.

each time stamp.

Using this data, we find that there is an increase in the estimated votes cast. In the left-hand plot of Figure 4 we show the estimated number of votes cast as the ballots arrived. After the polls closed the estimated vote total was 4,313,444. This total initially decreased to 4,294,871 and then quickly increased. By approximately 1030 pm Eastern Time the maximum estimated number of votes counted occurred, 4,474,071. And then subsequently, from 1030pm Eastern Time until 455 am on January 6, the estimated number of votes changed only slightly. Therefore, the left-hand plot of Figure 4 shows that the biggest changes in the estimated vote count occurred from 730pm to 1030pm.

According to Eastman's account of what happened on January 5th, 2021 the increase in the estimated vote count occurred when the share of the votes counted reached 99%. This is simply not true. At approximately 7:30 pm Eastern Time, when the estimated vote total began increasing only 5% of the votes had been counted at that point. By the time the estimated vote total was done increasing at 10:30 pm Eastern Time only 83.6% of the

estimated total votes cast had been counted. By the time the vote share was approaching a large share of ballots counted, the estimated vote total was changing very little.[18]

Two years after John Eastman's speech he revealed his source of information about voting machines and the supposed injection of fraudulent votes. During Eastman's disbarment hearing he testified that he learned about the supposed ability to hack voting machines the night before his speech on January 5th. Eastman reported that Joe Oltmann, a former startup CEO, showed John Eastman a single-page diagram that graphically depicted voting machines, external registration systems, and alleged that it was possible to inject false ballots. The only proof that John Eastman had before making the allegations was this diagram and, allegedly, a conversation with Eric Quinnell and Russell Ramsland.

Merely watching election returns and inferring fraud based on simplistic diagrams in bars is no way to infer election fraud. Further, Eastman fails to even recall what happened on Election Night correctly. Put simply, there is no proof for Eastman's assertions.

3.2.5 Timing of Absentee Ballots Does Not Indicate Fraud

In addition to arguing that Dominion voting machines contributed to Biden's success, Russell Ramsland's report in *Wood v. Raffensperger* and *Pearson v Kemp* suggests that the timing in which mail-in ballots were received is indicative of fraud. He claims that 96,600 were voted but have no return record, and 5,990 mail-in ballots were received either before, the same day, or within two days of being mailed to the voter. He thus concludes that "at least 96,600 votes were illegally counted in the Georgia general election" and that the remaining 5,990 suspicious ballots constitute an "impossible phenomenon occur[ing] throughout the counties

[18]Eastman's supposed evidence of voter fraud is not based on any assessment from research. We are unaware of any study that uses changes in the estimated vote total to detect fraudulent ballots. And it also fails logically. During Eastman's disbarment proceeding he claimed that he was uncovering evidence that the number of ballots to be counted and the number of ballots counted were increasing at the same rate. If the the number ballots to be counted and the number of ballots counted are increasing at the same rate, the share of ballots counted would have to be increasing. The primary issue appears to be that Eastman's arm chair observations on election night, even if they had been accurate, were based on numbers that are insufficiently precise to detect the relevant changes.

of Georgia" (Ramsland, 2020, 15).

Neither pattern that Ramsland identifies is actually indicative of fraud. The claim that 96,600 ballots were voted with no record of return is not indicative of fraud because it is based on a basic misunderstanding of what is contained in the Georgia absentee voter history file. Specifically, it appears that Ramsland incorrectly assumed that the "C" in the Georgia absentee voter history file referred to "counted" ballots when it actually refers to "canceled" ballots.[19] We know this is the case because we find that there are 96,600 ballots that canceled and have no date of being returned, a finding first reported in an expert report from Kenneth Mayer submitted who responded to Ramsland in *Pearson v Kemp*, a case where Ramsland's expert report was also used (Mayer, 2020).

Ramsland's claims about 5,990 suspicious ballots also fails to appropriately use the absentee voter history file. Ramsland claims that there are 42 ballots returned before they were mailed, 1,887 returned the same day, 1,786 returned the day after they were issued, and finally 2,275 ballots returned two days after being issued, for a total of 5,990 ballots Ramsland labels "suspicious".

This total, however, includes ballots that were never accepted to be tallied and instead were canceled. When we focus on accepted ballots only we find that there are 467 ballots that were issued and returned on the same day, 374 returned the day after they were issued, and 963 returned two days after it was issued. The small number of ballots returned on the same day could be part of a program where county election officials visit hospitals to ensure sick individuals can vote. In those instances, ballots are applied for, issued, and returned all on the same date. The other ballot returns time are actually quite feasible. For example, in the 2022 election the USPS took less than two days, on average, to deliver ballots to voters.[20] If a voter received their ballot and then immediately voted and returned the ballot, it could easily be recorded within two days after it was issued. Finally, the small number of ballots

[19] Ballots can be canceled for several reasons, including individuals changing how they want to vote, deciding to not vote at all, or because the ballot was lost or destroyed

[20] https://about.usps.com/what/government-services/election-mail/pdf/usps-2022-post-election-analysis.pdf

returned before they were issued are obviously clerical errors, rather than an indication of an underlying nefarious plan to steal an election.

3.2.6 Claims Made About Algorithmic Voter Fraud

Trump often suggests that Dominion voting machines algorithmically transferred his votes to Biden. For example, at a Georgia rally in December 2020 he claimed:

> "In one Michigan county using Dominion voting systems, nearly 6,000 votes were... wrongly switched from Trump to Biden....Numerous times we found glitches, and every single time, the glitch went a hundred percent to Biden and no percent to Trump. The same systems are used in 30 states."

Trump is referring to Antrim County, Michigan, the primary inspiration for such claims. The Antrim county clerk made a clerical error when setting up the election software, which briefly inflated Biden's unofficial vote totals due to the incorrect loading of a key for scanning ballots. The mistake was remedied by a rescanning of all county ballots and the consensus is that the error was not caused by the Dominion machine or ElectionSource software, but rather by human error (Halderman, 2021). This did not stop the Trump team from arguing that Dominion machines algorithmically gave Trump votes to Biden across the country. Related claims were filed in the Georgia court cases.

Ramsland claims that "rogue actors" from outside the US provoked Dominion voting machines to employ ranked choice voting (RCV). He argues that the algorithm may have contributed to Biden's success, as its use is "completely consistent with the mathematical advantage for Biden [in counties] using Dominion or Hart equipment" (Ramsland, 2020, 16). His evidence is that raw vote totals from Edison Research, reported in live-time, included non-integer values, which means the Dominion voting machines abandoned traditional vote counting for RCV.

Similar to the reports and claims discussed in Section 3.2.4, Ramsland misinterprets

the Edison data. Edison Research's live reporting produces *estimates* of vote totals, not actual values. Each timestamp includes an estimate of Biden's and Trump's vote shares. As Jonathan Rodden explained in an expert report responding to Ramsland, if you multiply them by the total number of votes, you are likely to get a non-integer estimate of the vote total (Rodden, 2020). Ramsland's argument therefore fails because RCV does not produce non-integer totals.

3.2.7 Alleged Evidence of Other Algorithmic Vote Transfers

Shiva Ayyadurai's report in *Pearson v. Kemp* also argues that votes were algorithmically transferred to Biden (Quinnell and Young, 2020b). His analysis focuses on DeKalb County, Georgia, and his evidence is summarized in a line plot. The x-axis is "number of cumulative votes in DeKalb County" and the y-axis is the "difference between Mr. Biden's votes... and what he should have received based on the ethnic distribution of DeKalb County" (Quinnell and Young, 2020b, 30). Ayyadurai says that "as we move from left to right on the x-axis, the percentage of white voters in each precinct increases" (Quinnell and Young, 2020b, 30). The line plot shows that Biden increasingly outperforms Ayyadurai's expected totals calculated from the county's "ethnic distribution" after a cumulative vote total of 150,000, which according to Ayyadurai, is when the "percentage of white voters increases beyond approximately ten-percent" (Quinnell and Young, 2020b, 30-31). He concludes this is "unequivocal evidence" (Quinnell and Young, 2020b, 28) that "as the percentage of white voters increases beyond approximately ten-percent...a mathematical algorithm comes into play, to transfer a weighted factor of total votes from President Trump to Mr. Biden in a very specific – un-natural [sic], machine-like manner" (Quinnell and Young, 2020b, 31). He calls this a "weighted race" algorithm and suggests it transferred 48,000 Trump votes to Biden in DeKalb County alone.

Several components of Ayyadurai's analysis are poorly defined and not based on any well established procedure for detecting illegal voting or fraud. Or even based on a logical

and well defined procedure invented in the report. Ayyadurai defines the horizontal axis as the "number of cumulative votes" without formally defining what he means. A reasonable interpretation is that this is the the vote totals as they were reported on election night by Edison. But if that is true, it is entirely unclear how the horizontal axis could possibly correspond to precinct-level ethnic distributions (Rodden, 2020). Ayyadurai also does not explain how he uses DeKalb's ethnic distribution to calculate Biden's expected vote totals, or why race is even relevant (Rodden, 2020). It is possible that Ayyadurai is assuming a strict partisan distribution for each ethnicity group, where whites and blacks overwhelmingly vote for Trump and Biden, respectively, and he is concerned that Biden outperforms such a model in DeKalb's whiter precincts.

This ethnic voting assumption is deeply unfounded in a racially diverse county like DeKalb. Even the majority white suburban precincts in DeKalb County are, on average, firmly Democratic (Rodden, 2020, 34). This is borne out in previous general elections, as well as in Biden's vote totals, where he received a vast majority of the votes in even the DeKalb precincts that are over 75 percent white (Rodden, 2020, 34). A strict ethnic voting model is thus implausible in such counties. White support for liberal candidates in diverse Democratic stronghold counties like DeKalb is typical across the US. That Biden performed well among such white voters in Georgia is not evidence of algorithmic fraud but expected. Similar to other case reports in Georgia, Ayyadurai's analysis reflects insufficient scientific rigor and knowledge about historical voting patterns, which clouds his inferences.

4 Pennsylvania Claims

The Trump campaign waged a multistate campaign to cahllenge the election results. We focus first on claims in Georgia because of the large amount of litigation there and because many of the claims made in Georgia were made in other states as well. That means, of course, that our refutations will also apply to the arguments made in other states. But in

other states there were also unique claims made about alleged illegality and fraud in the 2020 election. In this section we examine some of these distinct allegations in Pennsylvania.

One distinct feature of the election challenges in Pennsylvania is that a state legislature produced a report alleging evidence of illegal voting. After the 2020 election, Francis Ryan, a state legislator in Pennsylvania, wrote a letter to Scott Perry, a member of the House of Representatives from Pennsylvania, alleging that the 2020 general election "was fraught with inconsistencies, documented irregularities and improprieties associated with mail-in balloting, pre-canvassing, and canvassing" and the result was "that the reliability of the mail-in votes in the Commonwealth of Pennsylvania is impossible to rely upon" (Ryan, 2020). In fact, Ryan asserts that "The analysis below substantially confirms that the mail-in ballot process in the Commonwealth of Pennsylvania in the 2020 General Election was so defective that it is essential to declare the selection of presidential electors for the Commonwealth to be in dispute. The United States Congress is asked to declare the selection of presidential electors in this Commonwealth to be in dispute and to intervene in the selection of the electors for the Commonwealth of Pennsylvania for the 2020 General Election" (Ryan, 2020).

The Ryan Report was broadly reported across several media outlets and its content was used in *Texas v Pennsylvania.* The letter is brief, providing few details on how quantities were computed and, in some instances, sources of information. Nevertheless, we examined the empirical claims made in the report and evaluated those claims using relevant data from the 2020 Pennsylvania presidential election.

4.1 Mail-in and Absentee Voting on the Same Day in Pennsylvania

Perhaps the most consequential claim made in the Ryan Report is that mail-in and absentee ballots had sent dates and returned dates that were impossible, raising suspicions that the ballots may have been cast illegally. The Ryan Report makes five claims about mail-in ballots, though it provides few details about why these particular patterns are problematic and provides no details about how the analyses were conducted. We analyze all five

claims here and we find that each fundamentally misunderstands Pennsylvania election law or misrepresents what is found in the data.

The first claim in the Ryan Report is that there are "Ballots Mailed on or BEFORE 9-11-2020. That total is 27995". This claim appears to be the result of Representative Ryan misunderstanding the timeline for mail-in and absentee ballots in Pennsylvania. In 2019 Pennsylvania passed Act 77, which created no-excuse mail-in balloting in the state and modified the timeline at which counties could distribute ballots. According to Act 77 (with emphasis added):

> "(a) General rule.–Applications for mail-in ballots shall be received in the office of the county board of elections not earlier than 50 days before the primary or election, *except that if a county board of elections determines that it would be appropriate to the county board of elections' operational needs, any applications for mail-in ballots received more than 50 days before the primary or election may be processed before that time*"

Act 77 also states that once a ballot application is processed, the county board of elections "shall commence to deliver or mail official mail-in ballots as soon as a ballot is certified and the ballots are available."

As a result, state law enables county election officials to change the 50 day threshold that Representative Ryan appears to be referencing. The Secretary of Commonwealth of Pennsylvania confirmed this facet of the law in a letter addressing Representative Ryan's allegations. In a letter to Senators Ron Johnson and Gary Peters, Secretary Kathy Boockvar explained that Ryan's identified ballots were "neither unlawful nor unusual" (Boockvar, 2020). That's because "Pennsylvania counties are authorized to begin processing mail-in and absentee ballots more than 50 days prior to the election and are required to begin mailing ballots as soon as they are available" (Boockvar, 2020).

Consistent with counties exercising their discretion to process applications early, we find that the ballots Ryan identifies are concentrated in a small number of counties. To assess the

claims in the Ryan Report, we use the same reported source for data about the distribution of mail-in and absentee ballots in Pennsylvania. Following the Ryan Report, we downloaded the statewide mail ballot file that was uploaded to the Pennsylvania Open Data site. Using this data set, we nearly exactly replicate the claimed number of ballots that were sent on or before September 11, 2020: we find 28,718 ballots, compared to the 27,995 identified in the Ryan Report. Among these 28,718, we find that almost all are concentrated in just a few counties: 81.6% of the ballots are found in just five counties[21] and that 97.5% of the early ballots were distributed in just ten counties[22] These same ten counties represent only 11.9% of all mail-in or absentee ballots in Pennsylvania. Thus, there is clear evidence that the early ballots that Ryan identified merely were from counties who chose to distribute their ballots early, as allowed under Act 77.

The second claim in the Ryan Report is that there are "Ballots Mailed on November 1, 2 or 3. That total is 8163." These ballots are also neither suspicious nor unlawful. Individuals have several reasons for why they can obtain their mail-in ballot in the final days of the election. Using data downloaded from Pennsylvania's open data site, we find there were 9,603 ballots that were distributed on November 1st, 2nd, or 3rd. Among those ballots, 26.1% were "emergency" ballots: ballots that were issued after the close of the formal application period but when voters are still eligible to apply for and obtain a mail-in ballot in Pennsylvania.

The remaining 73.9% of the ballots reflect individuals who had their ballot issued by a county election official in the last days of the election. This is not surprising, because Act 77 created the ability for voters to cast their no-excuse mail-in ballots in one stop, just like Pennsylvania had allowed for several prior elections with its excused-based absentee voting system. Voters who had previously applied for a mail-in ballot but had not received it could go to a county election office, have their mail-in ballot printed, vote the ballot, and then return the ballot. These voters are merely ensuring they obtain their ballot before the election, as the law allows.

[21] Those five counties are Indiana, Carbon, Lawrence, Venago, and Schuykill.

[22] The additional five counties are Wayne, Butler, Lancaster, Cumberland, and Northampton.

The final three claims in the Ryan Report are related, so we consider them at the same time. The third claim in the Ryan Report is that there are "Ballots with NO MAILED date. That total is 9005." The fourth claim is that there are "Ballots Returned on or BEFORE the Mailed Date. That total is 58221" and the fifth claim is that there are "Ballots Returned one day after Mailed Date. That total is 51200."

Each of these claims appear to be the result of Representative Ryan being confused about how Act 77 altered the voting process and how the data on mail-in and absentee voting were recorded in the statewide mail ballot file. As mentioned above, Act 77 created the ability for Pennsylvania voters to do "one stop" mail-in or absentee voting. In that one stop, a voter can apply for a mail-in or absentee ballot, have the ballot printed, receive the ballot, vote the ballot, and then return the ballot to the county election official. While this process enables voters to complete the mail-in balloting process in one stop, it is legally equivalent to voting a mail-in absentee ballot. This is made clear on county websites. For example, Allegheny County describes this process as "over the counter" voting. When explaining the process, the Allegheny County election page explains that when performing one stop voting, "the process is the same as if a person applied for an absentee or mail-in ballot and had the ballot packet mailed to them."[23]

A second source of confusion is the reliance on a field called "MAILED DATE", which is a field in the statewide mail ballot request file loaded to the open Pennsylvania site. But the belief that this only refers to ballots that were placed in the mail is in error and, in fact, the date in this field references when a voter receives the ballot. We have several ways to know this. First, in Act 77 the requirement for this particular field is that it contains "The date on which the county board mails *or delivers* the mail-in ballot to the elector" (emphasis added). Second, the documentation on Pennsylvania's open data set makes clear that this field is actually called the "ballot sent date" field or the date that a "ballot label" is printed for a voter. Third, in a conversation with now Secretary of the Commonwealth Al Schmidt,

[23]https://www.alleghenycounty.us/Government/Elections/Voting/Over-the-Counter-Voting

who led Philadelphia County's election office in 2020, Schmidt confirmed that voters who receive their mail-in ballot in person would have that date logged as the Ballot Sent Date, which would then be the "Mailed Date" on the open data Pennsylvania site.

The Secretary of Commonwealth explained why the three claims are in error in a letter dated December 15th, 2020. For example, when explaining the number of voters who received their ballot and then returned the ballot the next day, the Secretary of Commonwealth explains that

> "Again, as Rep. Ryan should know, Act 77 authorized eligible Pennsylvania voters to vote early in person (by mail ballot) at their county election offices, and over 100,000 Pennsylvania voters availed themselves of this option. Most of these voters would be shown as having been approved and provided their ballot on the same date they cast it at their county election office. Far from an anomaly, the data Rep. Ryan cites is an obvious result of the legislation he himself supported." (Boockvar, 2020)

The letter went to explain that

> "Any dates listed incorrectly as cast before the mail date are likely routine data entry errors by the county election offices. DOS worked with county election officials to ensure that dates in the state's election registration system accurately reflected the date on which ballots were actually mailed. In fact, DOS implemented a utility in August to help counties manage this function to ensure that voters could access accurate information regarding when their ballots were mailed. Similarly, the allegations of some ballot entries lacking a mail date are caused either by a data entry error or omission, which occurred on occasion when counties printed ballot labels for voters who voted a mail-in ballot in person at their county election office." (Boockvar, 2020)

Our own analysis of the statewide mail ballot file confirms the Secretary of the Commonwealth's observations, while casting doubt on several of the reported numbers in the Ryan Report. First, we find only 186 ballots that were supposedly returned before they were sent out. Those 186 ballots represent 0.006% of all mail-in ballots in Pennsylvania. We also find 42,404 ballots that were sent and returned on the same date. Together, these 42,590 ballots are substantially fewer than the 58,221 reported in the Ryan Report. It is also many fewer ballots than reported in some initial post-election reports about fraudulent ballots. According to a November 10th, 2020 article in the *Epoch Times* "more than 23,000 have a return date earlier than the sent date" (Svab, 2020). Obviously, there is no evidence for this claim. And finally, we find 55,574 ballots were returned the day after the ballot was obtained. This number is higher than the number claimed in the Ryan Report and again, there is nothing suspicious about this empirical pattern.

After being confronted with evidence that this claim was false, a revised claim was made during the John Eastman disbarment proceedings. Rather than claim that all ballots provided clear evidence of illegal activity, the claim from Eastman's defense was revised to apply to only ballots applied for online. Specifically, the revised claim was that it was impossible for an individual to apply online for a ballot, obtain the ballot, and return it on the same day.

To investigate this claim, we used information available in the statewide mail ballot request file. We find evidence consistent with individuals completing their application and voting in one stop using offline applications and then obtaining their ballot and voting after previously applying online several days before for mail-in or absentee ballots. The vast majority of mail-in and absentee ballots that were obtained and voted on the same day were not applied for online. At least 81.4% (or 34,509 in total) of ballots received and voted on the same day were applied for offline. Of these 34,509 ballots, 73.7% were applied for on the same day or the day before they were voted. Online ballots have a very different pattern. Only 4.6% of the online ballots that were received and voted on the same day also had an

application from the day before. This represents only 60 total ballots. The remaining ballots indicate that online applicants who received and voted their ballots on the same day had applied for those ballots previously.

In summary, we find no evidence to support the claims of anomalous votes in the Ryan Report. Instead, we find clear evidence that support a misunderstanding of basic facts about how Pennsylvania administered its elections.

4.2 Anomalies in the Number of Absentee Ballots

The Ryan Report also made a vague allegation about the number of absentee ballots requested in Pennsylvania. Specifically, the Ryan Report alleged that "additionally, in a data file received on November 4, 2020 the Commonwealth's PA Open Data site reported over 3.1 million mail in ballots sent out. The CSV from the state on November 4 depicts 3.1 million mail in ballots sent out but on November 2, the information was provided that only 2.7 million mail ballots had been sent out. This discrepancy of approximately 400,000 ballots from November 2 to November 4 has not been explained" (Ryan, 2020). This claim was quoted in the *Texas V Pennsylvania* lawsuit and has been cited as evidence of fraudulent ballots in the Pennsylvania election.

This claim is simply false. And it is easy to demonstrate that it is false. Perhaps this is not surprising given that the claim is extremely vague. While clear that the 3.1 million request number comes from the "PA Open Data Site", it is unclear what "information was provided" to reach the 2.7 million number. We are aware of no clarification from Representative Ryan or anyone else to explain where this information came from. So, it is unclear what information was made available to contradict the statewide mail ballot file.

It is straightforward to show this claim is false. To examine the number of mail in and absentee ballots that had been applied for on each date, we obtained the daily releases of the Pennsylvania statewide mail ballot file. Using this data set, we examined the total number of requested ballots from each daily snapshot. In Figure 5 we present the cumulative number

of requested ballots (vertical axis) against the day the statewide mail ballot file data was distributed (horizontal axis). The points represent the cumulative number of requests as of that day and the black line connects the reported number of requests.

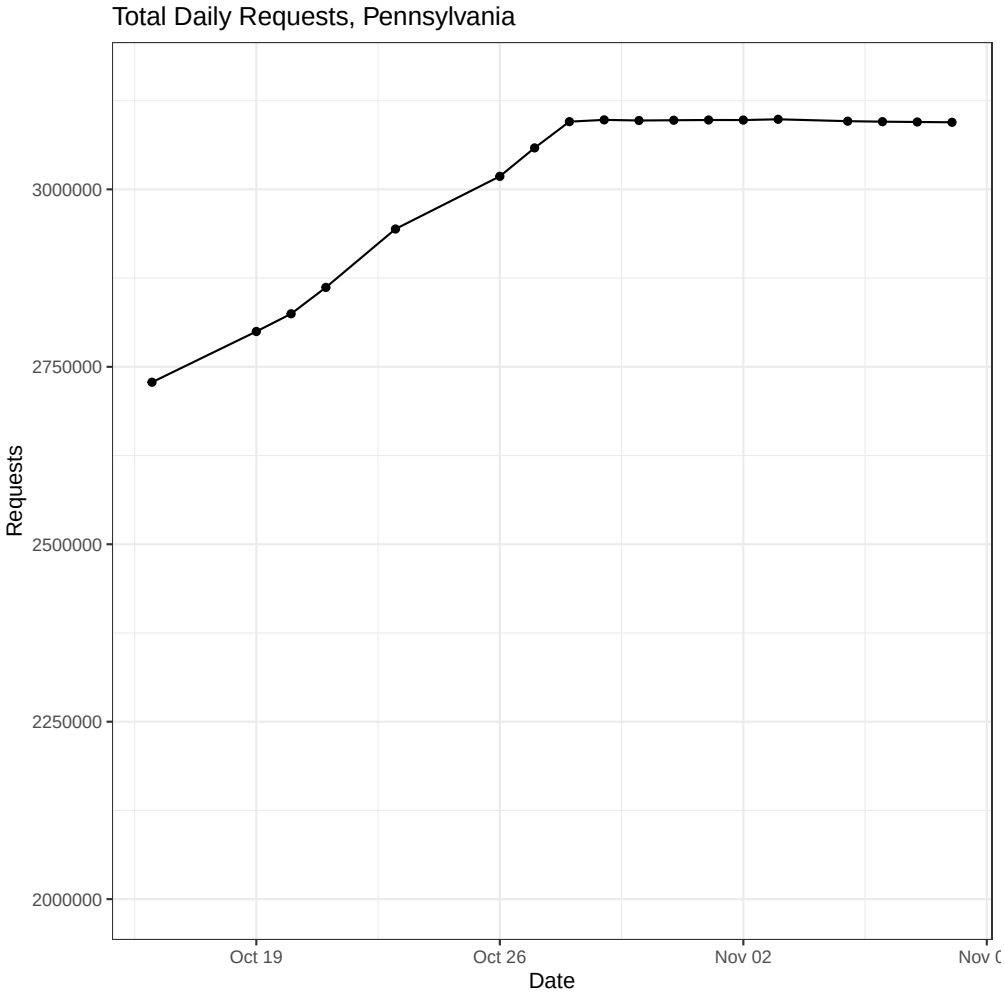

Figure 5: Cumulative Number of Requests by Day in Pennsylvania Before the 2020 Election

Figure 5 shows that the claim in the Ryan Report is simply false. There is no evidence that only 2.7 million ballots were requested as of November 2nd. In fact, there were more than 2.7 million ballots requested for every day shown on this plot. In general, the number of requested ballots grew steadily until applications for mail-in and absentee ballots closed. After the official close of applications, there were a small number of "emergency" requests made, slightly increasing the number of requested ballots to 3,098,701 on Election Day,

November 3rd, 2020. Then there is a small decline in the number of requests after Election Day as some voters cancel their ballots and vote in person.

Not only is it easy to refute this claim with a simple examination of a data set, information was widely available before the 2020 election that 3.1 million mail-in ballots had been requested in Pennsylvania. For example, the Associated Press distributed an article on October 27th with the headline "More than 3M in Pennsylvania apply for mail-in ballots" (Levy and Dale, 2020). On October 30th, the *Pennyslvania Capital Star* ran an article providing a "Map: How Pa. counties plan to count 3.1 million mail-in ballots" (Staff, 2020). Further, Secretary Boockvar held a press conference on October 30th, declaring that "More than 2.2 million mail or absentee ballots have already been returned" in the 2020 general election (Murray, 2020). She explained that this total "accounts for 73% of ballots requested" (Murray, 2020). Doing simple math shows that this is more than 3 million ballots. And finally, academics had well documented the number of ballots that were requested in Pennsylvania. Charles Stewart explained that 3.1 million ballots had been requested in Pennsylvania as of October 27th, 2020 (Stewart, 2020b).

In summary, this claim is vague, based on an unnamed data source, and easily disproven.

5 Michigan Claims

Like Georgia and Pennsylvania, the Trump campaign contested the results of the election. Many of the claims made in *King v Whitmer* repeat claims made in Georgia or in Pennsylvania.

But we analyze three Michigan specific claims that do not appear to be made in other states. First, we analyze a claim from *Texas v Pennsyvlania* that alleges a large number of mail in ballots were fraudulently cast in Wayne County. In fact, we show that the analysis is based on a basic error. And the claim is obviously false because it alleges every mail in ballot in Detroit was illegally cast. Second, we analyze supposed evidence of "algorithmic"

manipulation in Michigan elections. We show that the claim is obviously false. Even if we stipulate to the author's test, it shows the opposite of what they claim. But there is no reason to use their test to detect fraud or illegality. That's because they provide no evidence it is an effective tool for identifying fraud, illegal voting, or manipulation of electoral totals. Third, in Anonymous (2023) the anonymous author alleges that there are not enough voters in the voter file to explain Michigan's turnout. We explain that this is a common error made when analyzing voter file data.

5.1 False Claims About Absentee Ballots in Detroit

We first focus on one particular claim made in the *Texas v Pennsylvania* lawsuit. In particular, the suit claims that there were "more than 173,000 ballots in the Wayne County, MI center that cannot be tied to a registered voter" (Paxton, 2020b, 5-6). The specific allegation made in the Bill of Complaint was that "the Wayne County Statement of Votes Report lists 174,384 absentee ballots out of 566,694 absentee ballots tabulated (about 30.8%) as counted without a registration number for precincts in the City of Detroit" (Paxton, 2020b). The complaint goes on to allege that "The extra ballots cast most likely resulted from the phenomenon of Wayne County election workers running the same ballots through a tabulator multiple times, with Republican poll watchers obstructed or denied access, and election officials ignoring poll watchers' challenges, as documented by numerous declarations" (Paxton, 2020b).

The source for the allegation is an expert declaration from Charles Cicchetti. In the declaration Cicchetti explains that "I analyzed the absentee ballot data for Wayne County, Michigan, at the precinct level and I found that 174,384 absentee ballots out of 566,694 absentee ballots tabulated (about 30.8%) were counted without a registration number for precincts in the City of Detroit starting with Absentee Vote County Board [sic] 1 (ACVB 1) through (ACVB 134)." Cicchetti's declaration provides no more details about this allegation. Instead, he attempts to infer the margin for Joe Biden in the ballots he alleges were fraud-

ulently cast. He then attempted to connect the allegation to issues around the "balancing" of precincts in Michigan or whether the number of voters who arrived at the polls matches the number of ballots.

It is notable what's missing from Cicchetti's declaration. At no point does he identify a single specific ballot that was scanned without a registration number. But that does not stop him from alleging that nearly 1 in 3 absentee ballots in Wayne County were cast fraudulently.

Cicchetti's allegation is the result of profoundly misunderstanding how Wayne County, and Detroit in particular, provided precinct level election results for the 2020 presidential election. In order to count absentee ballots election administrators in the City of Detroit partially centralized the counting of absentee ballots. The City of Detroit consolidated its 503 precincts into 134 Absentee Vote Counting Boards (AVCB). When reporting the results for the 503 precints, the City of Detroit reported the number of registered voters at the precinct level and the number of ballots cast on Election Day. But for every precinct, the number of absentee votes cast is zero. Separately, when reporting results for absentee ballots, Detroit administrators reported the number of absentee ballots for a particular AVCB. So, for example, in "City of Detroit, AVCB 134" 356 absentee ballots were tallied. But all ACVBs in Detroit report zero registered voters, because voters are registered at the constituent precincts.

There is nothing in the reporting of the precinct level results in Wayne County that implies that ballots counted at AVCBs were counted without first verifying a voter's registration number. That conclusion is merely an unsupported inference from Cicchetti's expert declaration.

In fact, Cicchetti's declaration and the corresponding complaint for *Texas v Pennsylvania* implies that *every absentee ballot in the City of Detroit was fraudulently cast.* It appears that Cicchetti's allegation is based on merely adding together all the absentee ballots cast in the City of Detroit. When we use the precinct records from Wayne County we find that 174,384 absentee ballots were recorded as counted at Detroit Absentee Vote Counting Boards. This

also corresponds to the number of absentee ballots the Michigan Secretary of State reports from the City of Detroit in a post-election audit (Benson, 2021, 20).

This same post-election audit also showed that concerns about precincts being "out of balance", an issue raised in Anonymous (2023), could not possibly affect the election result. According to the post-election audit there were a total of 14 absentee vote counting boards out of balance after the votes were canvassed. These 14 boards were out of balance by 21 total votes, well below Joe Biden's 154,188 margin of victory in the state(Benson, 2021, 20).

The allegations of fraudulently cast ballots in the City of Detroit are obviously false and implausible. There are many eligible voters in the City of Detroit and setting aside any allegations, many of them obviously would vote using an absentee ballot. Instead, the allegations are based on rudimentary errors in reading data files with insufficient work done to assess whether the claims are true.

5.2 "Parallel Snakes" Analyses and Alleged Evidence of Algorithmic Voting

An expert report filed in the *King v Whitmer* lawsuit, Thomas Davis claimed to uncover "Irrational MI Absentee Ballot Findings" (Davis, 2020b). Davis claims to provide "very strong evidence that the absentee voting counts in some counties in Michigan have likely been manipulated by a computer algorithm" and that his analysis uncovered evidence that "the tabulating equipment in infected precincts has been programmed to shift a percentage of absentee votes from Trump to Biden" (Davis, 2020b). And based on his analysis he concluded that "If no other plausible explanation can be made for these unexpected findings, it appears that this computer software was installed sometime after the 2016 Presidential election" (Davis, 2020b).

Davis' (2020) conclusions about alleged voter fraud is based on visual inspections of graphs that he creates for three counties in Michigan: Ingham, Oakland, and Macomb. Davis (2020) first presents an analysis of Ingham County in the 2016 election. While he does

not explicitly describe his procedure, after careful inspection of the report we determined the figures were created using the following steps. First, for an unspecified subset of precincts in Ingham County in 2016 Davis (2020) calculated the share of Hillary Clinton's total votes in each precinct that were cast as an absentee ballots. Davis (2020) then plotted this share against an arbitrary precinct number and connected the points with a blue line. Davis (2020) fails to explain why the arbitrary precinct ordering he used was appropriate. Second, Davis (2020) performed the same calculation for Donald Trump, first calculating the share of Trump's total votes that were from absentee ballots in each precinct. Then, Davis (2020) plotted Trump's share against the same arbitrary precinct numbers, but now connected the points with a red line.

Davis (2020) then provides his visual analysis of this figure. He encourages the reader to "Note the irrequalities that occur: some precincts are higher for R, some higher for D ...**neither the red line nor blue line has a discernible pattern.** *This is what a normal result looks like!*" (Emphasis in Original). Davis offers no formal statistical test or calculation to reach a conclusion that the patterns across Ingham County precincts is "normal" or otherwise indicative of patterns consistent with an election free of irregularities.

Davis (2020) then creates the analogous plot for the 2020 election in Ingham County, but uses a different number of precincts for an undisclosed reason. In each precinct Davis now calculates the share of Biden's total votes that were cast using an absentee ballot and plots them against the arbitrary precinct number, connecting those points with a blue line. And Davis calculates the share of Trump's total votes in 2020 that were cast using an absentee ballot and then plots them against the arbitrary precinct number, connecting those points with a red line.

Using these calculations, Davis (2020) reports that his visual inspection leads him to the conclusion that something anomalous happened in 2020. Davis (2020) asserts that examining this plot he finds that "Except for one outlier, the percentage of Democratic absentee voters exceeds the percentage of Republican absentee voters in every precinct.

Even more remarkable (and unbelievable): these two independent variables appear to track one another" (Davis, 2020b). He goes on to assert that "There is no apparent legitimate explanation for the two absentee lines to be tracking each other like that – other than it being due to a computer algorithm (software program)." Again, Davis (2020) provides no statistical test or computes no particular number to reach the conclusion that there is something amiss in the results in 2020.

Davis (2020) never justifies his plots as a test of fraud, illegality, or algorithmic manipulation. He provides no derivation, no citation, no analysis, and he does not even point to a common understanding that would support his conclusion. Put simply, there is no reason to believe that the plots Davis (2020) created could possibly offer any information about illegal voting in Michigan or anywhere else.

Setting all that aside, even if we agreed with Davis' premise that the plots were useful for detecting fraud, the actual evidence in the plots directly contradicts his claims. Davis provides no explicit statistic to support his conclusion that in 2020 "these two independent variables appear to track one another." [24] Further, Davis provides no replication code nor does he provide sufficient detail to replicate the plots on our own. Because he is so vague, to test whether the relationship in 2020 was stronger than in 2016, we used a freely available software to manually extract the points from Davis' (2020) plots about Ingham County. Using this data, we then computed formal statistics to measure the relationship between the share of each candidates' total ballots cast using absentee ballots in the 2016 and 2020 presidential elections.

In direct contradiction to Davis' claims, we find that there was a stronger relationship between Clinton and Trump's share of ballots cast using absentee ballots in 2016 than Biden and Trump's share of ballots cast using absentee ballots in 2020. In 2016 the correlation between Clinton and Trump's share of votes from absentee ballots was 0.86, while in 2020 the

[24] Davis also appears to misunderstand basic statistical terminology. Just because two variables measure two different things–the share of Biden and Trump's votes that were cast using an absentee ballot across precincts—it does not mean that the variables are independent. Statistical independence generally refers to the idea that knowledge of one variable provides no information about the value of the other variable.

correlation between Biden and Trump's share of votes from absentee ballots was only 0.66. We reach a similar conclusion if we use other measures of association. If instead we calculate the root mean squared error across the Democratic and Republican candidate's share of votes from absentee ballots we find it higher in 2020 (27.7) than in 2016 (4.9)—indicating a stronger relationship in 2016 than in 2020. Or, if we calculate the average prediction error, relative to Trump's share of votes from absentee ballots we find that it is higher in 2020 (0.63) than in 2016 (0.30), again indicating that the relationship was actually stronger in 2016 than in 2020.

Davis' (2020) strong conclusions are based on vague or unstated criteria and, ultimately, an incorrect assessment of his own data. When we analyze his data set formally we find no evidence that there is more evidence the variables "track" each other in 2020 than in 2016. In fact, we find the opposite: a stronger relationship in 2016 than in 2020. But regardless of what this analysis finds, there is no reason to believe that a relationship between how votes are cast across precincts is an indication of fraud. After all common characteristics of precinct residents–like their age, income, education, or racial identification—could explain why they cast their votes in particular ways.

5.3 Voter Files and Reconciling with Official Data

In the unsigned memo released early in 2024, the anonymous author claimed that "In data obtained from the Secretary of State's office on nearly a monthly basis since the Election, the most voters ever recorded in the Qualified Voter File was 5,511,303 voters in April 2021. This means Michigan's own election records 80 showed 68,014 more votes than voters" (Anonymous, 2023, 27). While the author never explicitly explained why this is problematic, a common claim among election skeptics is that the failure of official turnout histories to align with official participation numbers is evidence that "phantom voters" participated in the election.

This belief, however, is actually just a misunderstanding about the purpose of voter files.

Voter files are used to provide a snapshot of the current citizens eligible to vote. This means that individuals who have become ineligible, died, or otherwise canceled their registration will be removed from the voter file and therefore not available to reconcile with official numbers. In other words the failure of the voter file to reconcile with official numbers merely reflects the churn expected in voter files.

6 Wisconsin Claims

One of the most prominent sources for post-election claims of illegality and fraud came from Michael Gableman, a former Wisconsin state Supreme Court justice. Robin Vos hired Gableman as special counsel to investigate the 2020 election. On March 1st, 2022 Gableman produced the "second interim report" from his office (Gableman, 2022). This report contained a series of allegations about illegality and fraudulent voting in the 2020 election, along with extended allegations of improper influence from the Center for Tech and Civic Life. In this section we investigate Gableman's specific allegations and find little support for his allegations.

Yet again, his claims fail under basic scrutiny. He either misrepresented the evidence that was available or used improper evidence to suppoort his claims. When we use the appropriate evidence, we find no support for his claims.

6.1 CTCL Grants in Wisconsin

Throughout Gableman's report he alleges that the Center for Tech and Civic Life (CTCL) exerted inappropriate influence over the administration of the 2020 election. Gableman goes on at length about the relationship between the five cities that received the most funding from the CTCL, which Gableman designates as the Zuckerberg 5, and CTCL. Gableman cites a study from the Wisconsin Institute for Law and Liberty that alleges money was inappropriately targeted at Democratic cities and caused a disproportionate increase in turnout

among Democrats (Flanders and an Kyle Koenen, 2021).

Gableman's conclusions and the conclusions from Flanders and an Kyle Koenen (2021), are based on poorly specified models that suffer from obvious shortcomings when estimating the effect of CTCL grants on election outcomes. Consider first where the grants were allocated in Wisconsin. Gableman compares the dollars distributed per voter in the five large cities in Wisconsin to (apparently) the next five largest cities in Wisconsin. This is an inappropriate method to infer whether the cities receiving CTCL grants were likely to cast votes for the Democratic party candidate in the 2020 presidential election. This is inappropriate, in part, because there are thousands of potential cities in Wisconsin that could be targeted.[25]

We use data from Lal and Thompson (2023) to examine where CTCL grants were targeted in the state of Wisconsin. We find little evidence that grants were targeted at municipalities in Wisconsin likely to support Democrats. To assess a municipality's likely vote in 2020, we examined the share of the vote in 2016 allocated to Hillary Clinton. While imperfect, this is likely the information available to anyone allocating grants before the 2020 election. We find that municipalities that did and did not receive grants voted in nearly identical ways in 2016. In 2016, Hillary Clinton received 38.9% of the vote in municipalities that received a CTCL grant and received 38.0% of the vote among municipalities that did not receive a grant. This difference is neither statistically nor substantively meaningful.

We also do not see any evidence that CTCL targeted larger grants at municipalities that were more supportive of Hillary Clinton in 2016. For each municipality in 2020, we calculated the number of dollars allocated per voting age resident, calculated using the estimated number of voting age residents from the state of Wisconsin. For municipalities that did not receive a grant, we calculated $0 per voter. We then regressed the CTCL grant money per voter on Hillary Clinton's vote share from 2016 in that municipality.

We find almost no relationship between the dollars allocated per voter and support for

[25] As a matter of basic logic, it is also the case that four of the five next largest municipalities in Wisconsin had a higher Democratic vote share in the 2016 presidential election than Green Bay, one of the cities in the "Zuckerberg 5".

Hillary Clinton in 2016. For every one percentage point increase in support for Hillary Clinton in 2016, on average CTCL allocated $0.006 *fewer* dollars per voter. In other words, even if we implausibly shifted Hillary Clinton's vote share from 0% to 100% there would only be an increase of $0.60 per voter in expenditure.

Lal and Thompson (2023) also found no evidence that CTCL caused the election administration process to favor Joe Biden or other Democratic candidates. Lal and Thompson (2023) analyze the effect of CTCL grants nationwide and in Wisconsin in particular. Rather than using poorly specified linear regression models as in Flanders and an Kyle Koenen (2021), Lal and Thompson (2023) use more sophisticated panel data methods that account for trends in Wisconsin municipalities. Using these more appropriate methods, they find essentially no effect of receiving a CTCL grant on either turnout or Democratic vote share. Lal and Thompson (2023) estimate that CTCL grants *decreased* turnout in Wisconsin and caused a modest 0.11 to 0.31 percenntage point increase in Democratic vote share. But Lal and Thompson (2023) are unable to reject the null hypothesis of no effect on Democratic vote share.

There is no evidence CTCL grants had any effect on the Wisconsin election and no evidence they were targeted at particularly Democratic locations.

6.2 Nursing Homes and Voting

Gableman also alleged that changes to the absentee voting procedures at nursing homes "result[ed] in many nursing homes' registered residents voting at 100% rates and many ineligible residents voting, despite a guardianship order or incapacity" (Gableman, 2022). Specifically, Gableman claims that "Rampant fraud and abuse occurred statewide at Wisconsin's nursing homes and other residential care facilities in relation to absentee voting at these facilities" (Gableman, 2022). The source of this fraud, according to Gableman, was that the Wisconsin Electoral Commission (WEC) altered how absentee voting occurred at nursing homes and other assisted living facilities in Wisconsin. According to Gableman (2022), in prior

elections, election officials would designate special voting deputies who would be dispatched to nursing homes to assist residents with casting ballots. With the pandemic, WEC altered those requirements and allowed assisted living employees to assist residents in casting their ballots.

Gableman (2022) alleged that this change in how absentee ballots were handled at Wisconsin nursing homes caused improbably high turnout rates at nursing homes. In a table in the report, Gableman claims that "vetted" nursing homes in Milwaukee, Racine, and Dane counties had 100% voter turnout, Kenosha county nursing homes had 97% turnout, and Brown county nursing homes had 95% turnout. Gableman provided few details on how he arrived at these turnout rates. Instead, he wrote that

> The Special Counsel intended to use a professional statistician in the nursing home setting. Using a controlled environment, the OSC could take a detailed sampling of nursing home abuse and voting irregularities to determine, statewide, the number of improperly cast ballots in residential care facilities. The OSC was not able to complete this task by the time this Report was due. Instead, the personnel conducting the nursing home investigation were also repurposed to assist in the drafting of this Report (Gableman, 2022, 12).

Gableman's report provides no details on how the quantities were calculated and provides no evidence that the votes cast from nursing home residents were cast using absentee voting data.

After examining the claims and appropriate data, we find that Gableman's reported turnout rates at nursing homes are based on profound errors and data misrepresentations. Ron Heuer, who performed the calculation while working for Gableman, reported in Rickert (2022) that "he used the state's voter database as of August 2021 to look at whether every registered voter at every nursing home in the five counties voted in the November 2020 election" (Rickert, 2022). Then, according to Rickert (2022), Heuer provided the information

used to arrive at the turnout rates to Mark Spreitzer a Democratic state legislator in Wisconsin who subsequently disclosed the information to Chris Rickert, a report for the Wisconsin State Journal.

We obtained the information from Chris Rickert and then reviewed the information. We find that the data does not support Gableman's claims about nursing home turnout. Consistent with the findings reported in Rickert (2022), Heuer regularly reports greater than 100% turnout—an obvious impossibility. One reason this could occur is that because Heuer used a voter roll from nearly a year after the election. As residents die they will be removed from the voter roll, but their ballot would remain recorded as being from the nursing home. And the result is over 100% turnout rates at those facilities.

We examined the actual turnout data from nursing homes in Wisconsin and found that turnout was far lower than reported in Gableman's report. We obtained the data from Chris Rickert, who also described the queries used to obtain the turnout data. Analyzing the data, we find that the average turnout across 29 nursing homes in Milwaukee County was 74.0%, with turnout ranging from 20% to 92.5% in the homes. In 2016 the average turnout in those same 29 nursing homes was 82.1%, with turnout ranging from 46.2% to 100%. In fact, according to this data there were no nursing homes in Milwaukee with 100% turnout in 2020, but two nursing homes had 100% turnout in 2016.

Examining other counties we also find no evidence of anomalous turnout at Wisconsin nursing homes. In Brown County we 80% of registered nursing home residents cast a ballot, with only one nursing home with perfect turnout. But at this nursing home there were only 8 registered residents, so it is not surprising that the nursing home recorded perfect turnout. In Kenosha County on average 72% of registered nursing home residents turned out to vote, while in Racine County 73% of registered nursing home residents cast a ballot. There were no nursing homes in either Kenosha or Racine County with 100% turnout.

In short, Gableman's claims about 100% turnout at nursing homes is not supported by the data. His report provides too few details to replicate the claims he made, but it appears

to be the result of applying the wrong data to analyze the question or simply incorrectly performing the analysis.

7 Fraud Claims and the 2020 Election

In this paper we have provided an extensive assessment of voter fraud claims from the 2020 election. Across all the distinct claims that we analyzed, we find no credible evidence of voter fraud. But most surprising is just how weak the evidence of illegal or fraudulent voting really is. The claims made in Trump's legal filings are based on error riddled analyses from "experts" who misunderstand basic facts about elections, basic statistical concepts, and fail to justify the statistical tests that are used. Not only was Trump's evidence insufficient to establish sufficient illegal or fraudulent voting to overturn the election. The supposed evidence is essentially useless and only reflects the profound misunderstandings of the individuals who provided the reports to the Trump campaign.

In the immediate future the 2020 election is likely to continue to be litigated. Trump and his allies face indictment in federal court, Georgia, and Trump will likely press his case on the campaign trail that the 2020 election was stolen from him. Yet, Trump simply lacks the evidence for these claims. Across numerous expert reports an academic papers, there simply is no evidence that fraud or illegal voting decided the 2020 election.

References

Angrist, Joshua D and Jörn-Steffen Pischke (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Anonymous (2023). Summary of election fraud in the 2020 presidential election in the swing states. Available at https://cdn.nucleusfiles.com/e0/e04e630c-63ff-4bdb-9652-e0be3598b5d4/summary20of20election20fraud20in20the20swing20states.pdf.

Ansolabehere, Stephen (2020). Expert report of stephen ansolabehere. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Benson, Jocelyn (2021, April). Audits of the november 3, 2020 general election. https://www.michigan.gov/-/media/Project/Websites/sos/30lawens/BOE$_2$020$_{P}ost_{E}lection_{A}udit_{R}eport_0$4$_2$1$_2$1.$pdf?rev$ = $a3c7ee8c06984864870c540a266177f2$.

Boockvar, Kathy (2020, December). Letter to chairman johnson and ranking member peters from secretary of commonwealth kathy boockvar.

Braynard, Matthew (2020a). Expert report of matt braynard. Trump v Raffensperger, No. 2020CV343255, Fulton County.

Braynard, Matthew (2020b). Expert report of matthew braynard. Stevenson v Ducey, No. CV2020-096490, Maricopa County.

Braynard, Matthew (2020c). Expert report of matthew braynard. State of Texas vs Commonwealth of Pennsylvania, State of Michigan, State of Wisconsin, and State of Georgia.

Cegavske, Barbara K. and Mark Wlaschin (2021). Re: Elections integrity violation reports. https://www.nvsos.gov/sos/home/showpublisheddocument?id=9428.

Cicchetti, Charles J. (2022, April). How the 2020 election was won. *Academia Letters* (5134).

Danforth, John , Benjamin Ginsberg, Thomas B. Griffith, David Hoppe, J. Michael Luttig, Michael W. McConnell, Theodore B. Olson, and Gordon H. Smith (2022). LOST, NOT STOLEN: The Conservative Case that Trump Lost and Biden Won the 2020 Presidential Election. `https://lostnotstolen.org//wp-content/uploads/2022/07/Lost-Not-Stolen-The-Conservative-Case-that-Trump-Lost-and-Biden-Won-the-2020-Presidential-Election-July-2022.pdf` (last accessed October 12, 2022).

Davis, Mark Alan (2020a). Expert report of mark alan davis. Trump v Raffensperger, No. 2020CV343255, Fulton County.

Davis, Thomas (2020b, November). Irrational mi absentee ballots findings, expert report of thomas davis submitted in *King v Whitmer*, 2:20-cv-13134, eastern district michigan. Michigan 2020 Voting Analysis Report.

Eastman, John (2020a, December). John eastman testimony at hearing to assess election improprieties and to evaluate the election process to ensure the integrity of georgia's voting system. https://www.youtube.com/watch?v=IHt6UEc_tQ8.

Eastman, John (2021, January). John eastman speech at march to save america. https://www.c-span.org/video/?c4953961/user-clip-john-eastman-january-6-rally.

Eastman, John C. (2020b, December). The constitutional authority of state legislatures to choose electors: Testimony of dr. john c. eastman.

Eggers, Andrew C. , Haritz Garro, and Justin Grimmer (2021). No evidence for systematic voter fraud: A guide to statistical claims about the 2020 election. *Proceedings of the National Academy of Sciences 118*(45), e2103619118.

Favorito, Garland and Bob Coovert (2022). Refutation of georgia secretary of state brad raffensperger's false election claims. https://voterga.org/wp-content/uploads/2022/02/VOTERGA-Raffensperger-Congress-Letter-Refutation.pdf.

Flanders, Will and Cori Petersen an Kyle Koenen (2021). Finger on the scale: Examining private funding of elections in wisconsin. *Wisconsin Institute for Law & Liberty*. https://will-law.org/wp-content/uploads/2021/06/WillLawFINGER-ON-THE-SCALE.6.pdf.

Gableman, Michael (2022). Office of the special counsel second interim investigative report on the apparatus procedures of the wisconsin elections system delivered to the wiscon-

sin state assembly on march 1, 2022. https://www.wpr.org/sites/default/files/osc-second-interim-report.pdf.

Geels, Bryan (2020). Expert report of bryan geels. Trump v Raffensperger, No. 2020CV343255, Fulton County.

Geels, Bryan (2021). Updated expert report of bryan geels. Trump v Raffensperger, No. 2020CV343255, Fulton County.

Geels, Bryan (2023). Registrations occurring prior to the age of 17, according to the georgia secretary of state's voter data. Submitted for John Charles Eastman Disbarment Proceeding.

Grimmer, Justin , Andrew Hall, and Daniel Thompson (2021). Evaluating look ahead america's "the georgia report" on illegal, out-of-state voting in the 2020 election.

Grimmer, Justin , Michael Herron, and Matthew Tyler (2023). Evaluating a new generation of expansive claims about vote manipulation. Available here: https://www.dropbox.com/s/92lp1gmqw2pei8m/Expansive.pdf?dl=0.

Grofman, Bernard and Jonathan Cervas (2023). Statistical fallacies in claims about 'massive and widespread fraud' in the 2020 presidential election: Examining claims based on aggregate election results. *Statistics and Public Policy*, 1–36.

Halderman, J. Alex (2021). Analysis of the antrim county, michigan november 2020 election incident. https://www.michigan.gov/sos/-/media/Project/Websites/sos/30lawens/Antrim.pdf?rev=fbfe881cdc0043a9bb80b783d1bb5fe9hash=ACE9

Herron, Michael (2020). Expert report of michael herron. Law v Whitmer, First Judicial District Court, Carson City, Nevada.

Herron, Michael (2023). Allegations made against dominion voting systems and the 2020 presidential election in wisconsin. *Election Law Journal: Rules, Politics, and Policy 22*(3), 247–267.

Kamzol, Jesse (2020). Expert report of jesse kamzol. Law v Whitmer, First Judicial District Court, Carson City, Nevada.

Keshel, Seth (2020). Expert report of seth keshel, in bowyer et al. v. ducey (governor) et al., us district court, district of arizona.

King, Gary (2020). Expert report of gary king. Bowyer v Ducey, No. 1:20-cv-02321, District Court of Arizona.

Lal, Apoorva and Daniel M Thompson (2023). Did private election administration funding advantage democrats in 2020? arXiv preprint arXiv:2310.05275.

Levy, Marc and Maryclaire Dale (2020, October). More than 3m in pennsylvania apply for mail-in ballots. https://apnews.com/article/election-2020-pennsylvania-elections-voting-2020-presidential-elections-b1fe12514c3a736ec33f84e4ce7e4161.

Ligon, William (2020, December). The chairman's report of the election law study subcommittee of the standing senate judiciary committee. http://www.senatorligon.com/THE_FINAL

Mayer, Kenneth (2020). Expert report of kenneth mayer. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Murray, Darryl C. (2020, October). 2.2 million mail or absentee ballots already returned in pa. https://whyy.org/articles/2-2-million-mail-or-absentee-ballots-already-returned-in-pa/.

Paxton, Ken (2020a). Charles cicchetti sur-rebuttal in state of texas vs commonwealth of pennsylvania, state of michigan, state of wisconsin, and state of georgia.

Paxton, Ken (2020b). State of texas vs commonwealth of pennsylvania, state of michigan, state of wisconsin, and state of georgia.

Quinnell, Eric (2020). Expert report of eric quinnell. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Quinnell, Eric et al. (2021). 2020 presidential election startling vote spikes. https://election-integrity.info/Vote$_S$*pikes*$_R$*eport.pdf*.

Quinnell, Eric and S. Stanley Young (2020a). Expert report of eric quinnell and s. stanley young. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Quinnell, Eric and S. Stanley Young (2020b). Export report of shiva ayyadurai. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Quinnell, Eric and S. Stanley Young (2020c). Sur-rebuttal of eric quinnell and s. stanley young. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Raffensperger, Brad (2021, January). Point by point refutation of false claims about georgia elections. https://sos.ga.gov/sites/default/files/2022-02/Letter_to_Congress_from_Secretary_Raffensperger_%281-6-21%29.pdf.

Ramsland, Russell (2020). Expert report of russell ramsland. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Redacted (2021). Exhibit s in *Thompson v Eastman* united states district court central district of california 8:22-cv-00099-doc-dfm. https://www.cacd.uscourts.gov/sites/default/files/Dkt%20350-20%2C%20Ex.%20S%20GA%20dead%20voter%20update$_0$*.pdf*.

Rickert, Chris (2022, April). Michael gableman's numbers on nursing home voting proven wrong again. *Wisconsin State Journal*. https://madison.com/news/local/govt-and-politics/elections/michael-gablemans-numbers-on-nursing-home-voting-proven-wrong-again/article_0b7d7a6d-5cf3-5068-83d6-7a4f7a6f3ca9.html.

Rodden, Jonathan (2020). Expert report of jonathan rodden. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Rodden, Jonathan and Will Marble (2020). Expert report of jonathan rodden and will marble. Pearson v Kemp, No. 1:20-cv-04809-TCB, Northern District of Georgia.

Ryan, Francis X. (2020). Letter, rep. francis x. ryan, pennsylvania house of representatives, to rep. scott perry, u.s. house of representatives (dec. 4, 2020).

Shino, Enrijeta , Daniel A Smith, and Laura Uribe (2022). Lying for trump? elite cue-taking and expressive responding on vote method. *Public Opinion Quarterly 86*(4), 837–861.

Smith, Ray (2020). Ray smith oral testimony before election law study subcommittee of the standing senate judiciary committee. https://www.youtube.com/watch?v=wAOmkZJFlr4.

Staff (2022). Counting the vote: Frequently asked questions. https://www.ap.org/about/our-role-in-elections/counting-the-vote.

Staff, CNN (2021, January). Read the full transcript and listen to trump's audio call with georgia secretary of state. https://www.cnn.com/2021/01/03/politics/trump-brad-raffensperger-phone-call-transcript/index.html.

Staff, Capital-Star (2020, October). Map: How pa. counties plan to count 3.1 million mail-in ballots. https://www.penncapital-star.com/election-2020/map-how-pa-counties-plan-to-count-3-1-million-mail-ballots/.

Stewart, Charles (2020a). Expert report of charles stewart. Trump v Raffensperger, No. 2020CV343255, Fulton County.

Stewart, Charles (2020b, October). What's pennsylvania to expect with early voting? https://electionupdates.caltech.edu/2020/10/27/whats-pennsylvania-to-expect-with-early-voting/.

Svab, Petr (2020, November). Tens of thousands of pennsylvania ballots returned earlier than sent date: Researcher.

Swift, Clint S. and Delaney Gomen (2021). Invalid signature rejections in georgia absentee voting: Comparing the 2020 general to the 2021 senate runoff. https://electionlab.mit.edu/sites/default/files/2021-07/swiftgomen$_g$*arejections.pdf*.

Team, Reuters Fact Check (2020, November). Fact check: Vote spikes in wisconsin, michigan and pennsylvania do not prove election fraud. https://www.reuters.com/article/uk-factcheck-wi-pa-mi-vote-spikes-idUSKBN27Q307/.

Trump, Donald (2020a). Donald trump georgia rally transcript before senate runoff elections december 5. https://www.rev.com/blog/transcripts/donald-trump-georgia-rally-transcript-before-senate-runoff-elections-december-5.

Trump, Donald (2020b). Trump tweet, december 16. https://www.presidency.ucsb.edu/documents/tweets-december-16-2020.

Trump, Donald (2020c). Trump tweet, november 16. https://www.presidency.ucsb.edu/documents/tweets-november-26-2020.

Trump, Donald (2020d). Trump tweet, november 19. https://www.presidency.ucsb.edu/documents/tweets-november-19-2020.

Trump, Donald (2020e). Trump tweet, november 30. https://www.presidency.ucsb.edu/documents/tweets-november-30-2020.

Trump, Donald (2020f). Trump tweet, november 4. https://www.presidency.ucsb.edu/documents/tweets-november-4-2020.

Trump, Donald (2021a). Donald trump rally speech transcript dalton, georgia: Senate runoff election, january 4th. https://www.rev.com/blog/transcripts/donald-trump-rally-speech-transcript-dalton-georgia-senate-runoff-election.

Trump, Donald (2021b). Read trump's jan. 6 speech, a key part of impeachment

trial. https://www.npr.org/2021/02/10/966396848/read-trumps-jan-6-speech-a-key-part-of-impeachment-trial.

Trump, Donald J. (2020g). Trump campaign press release - trump campaign files election contest in georgia. https://www.presidency.ucsb.edu/documents/trump-campaign-press-release-trump-campaign-files-election-contest-georgia.

Vespa, Matt (2020). Why a pollster is saying the 2020 results are ringing 'eerily true' to what joesph stalin said about elections.

Wu, Jennifer , Chenoa Yorgason, Hanna Folsz, Cassandra Handan-Nader, Andrew Myers, Tobias Nowacki, Daniel M Thompson, Jesse Yoder, and Andrew B Hall (2020). Are dead people voting by mail? evidence from washington state administrative records.

Young, Stanley and Ray Blehar (2021). 2020 presidential election contrast analysis. https://election-integrity.info/Contrast$_R$*eport.pdf*.