Simple tests for the extent of vote fraud with absentee and provisional ballots in the 2020 US presidential election

John R. Lott, Jr., Ph.D.
Crime Prevention Research Center

Abstract

This study reports three tests measuring vote fraud in the 2020 US presidential election, although they provide inconsistent evidence. To isolate the impact of a county's vote-counting process and potential fraud on candidates' vote margins, I first compare voting precincts in a county with alleged fraud to adjacent precincts in neighboring counties with no allegations of fraud. I compute the differences in President Trump's vote shares on absentee ballots in those adjacent precincts, controlling for the differences in his vote shares on ballots cast in person. I also control for registered voters' demographics and compare data for the 2016 and 2020 presidential elections. When I examine Georgia and Pennsylvania separately, weak evidence of vote fraud on absentee ballots is found. However, combining the samples produces significant results and implies at least 10,000 additional votes for Biden in Pennsylvania's Allegheny and Georgia's Fulton counties. I then apply the same method to provisional ballots in Allegheny County, where, contrary to state law, voters were allowed to correct alleged defects in absentee ballots by submitting provisional ballots on Election Day. My analysis finds that such permission contributed to a statistically significant additional 6,700 votes for Biden. Finally, vote fraud can show up as artificially larger voter turnouts, higher rates of filling out absentee ballots for people who hadn't voted, dead people voting, ineligible people voting, or payments for votes. The estimates for Arizona, Georgia, Michigan, Nevada, Pennsylvania, and Wisconsin combined indicate an average of 255,000 excess votes for Biden.

1. Introduction

Brennan and Buchanan (1984) compared voting to cheering at a sporting event – one shouts in support of one's team without any demonstrable impact on the outcome. Perhaps we can push their comparison even further. Sometimes fans (as well as voters) are convinced that the referees make the wrong call.

This paper offers a unique method of testing the existence of vote fraud by comparing adjacent precincts – ones in counties with alleged vote fraud to ones where no fraud is alleged. The method applies to any state with fraud allegations in at least part of the state and to precinct-level voting data on absentee and in-person voting. I then apply the test to allegations that some heavily Democratic counties adopted different rules to process provisional ballots. (Voters receive provisional ballots when questions arise over their eligibility to vote.) Finally, I report an additional test by looking at voter turnout rates.

In 2005, the bipartisan Commission on Federal Election Reform, chaired by former President Jimmy Carter and former Secretary of State James Baker III, concluded that "Absentee ballots remain the largest source of potential voter fraud." Intimidation and vote-buying also were key concerns of the commission: "Citizens who vote at home, at nursing homes, at the workplace, or in church are more susceptible to pressure, overt and subtle, or intimidation. Vote-buying schemes are far more difficult to detect when citizens vote by mail" (Carter and Baker 2005, p. 46).

Worries about vote-buying have a long history in the United States. They helped drive the move to the secret ballot, which US states adopted between 1888 and 1950. Secret ballots made it harder for vote buyers to monitor the candidates whom voters actually supported. Vote-buying had been pervasive, with voter turnout falling by about 8% to 12% after states adopted the secret ballot (Lott and Kenny 1999). Absentee ballots reverse that trend, making it easier to buy votes. The Carter-Baker report identifies additional instances of absentee ballot fraud in modern times; other recent examples can be found (Lott 2020a; Crime Prevention Research Center 2020).

Lyndon Johnson was both a victim and perpetrator of fraud in his 1941 and 1948 Senate races (Caro 1982, pp. 736-738; Caro 1991, Chapter 13). The fraud allegedly involved vote-buying, stuffing ballot boxes, and destroying votes. Vote fraud in Chicago until the 1980s is well documented (Bauer 2018). The city's Democratic political machine added dead people to voting rolls and then had people vote in their names. Precinct captains also would vote on behalf of individuals who didn't show up at the polls, and they would "help" senior citizens in nursing homes cast absentee ballots. The nursing homes cooperated with Democratic party activists because they wanted good relationships with the city.

Much of the work on vote fraud in the United States involves audits that look at double voting or dead people voting (Goel et al. 2020; Wu et al. 2020; Viebeck 2020). Other studies count the number of convictions for vote fraud (Heritage Foundation 2022). But both approaches are of

very limited value. Many types of vote fraud are hard to detect.[1] It is difficult to catch someone who fills out and mails in absentee or mail-in ballots for others. Providing provisional ballots for individuals who are told that they have already voted often doesn't even avoid the problem of voting an absentee ballot in someone else's name. Less than half of people vote in most elections, so only a small portion of fraudulent votes can be detected, replaced, and corrected. Finally, vote-buying also is tough to detect because buyers and sellers of votes both gain from the transaction.

It's not just Republicans who are concerned about absentee ballots. Indeed, almost all European countries have adopted stricter anti-fraud voting rules than the United States (Lott 2020a, 2021c). For example, 35 of the 47 countries in Europe ban absentee voting entirely for citizens living there.[2] Another ten countries allow it, but require voters to show up in person and present photo IDs to pick up their ballots. Six of those ten countries limit the practice to people in the military or a hospital, and they require third-party verification. Another 16 European countries ban absentee ballots for voters living abroad on Election Day. Similar requirements are imposed in other nations.

Developed countries, with few exceptions, did not adopt emergency voting measures during the coronavirus pandemic. Poland allowed mail-in ballots as a one-time option for everyone during 2020, as did Bavaria for a run-off election and two cities in Russia (Flis and Kaminski 2020; Lott 2020b). France adopted more limited exceptions, temporarily allowing sick or at-risk individuals to vote absentee.

Many countries have learned the hard way about what happens when mail-in ballots aren't secured. They have discovered how hard it is to detect vote-buying.

France banned mail-in voting in 1975 because of massive fraud in Corsica -- voters cast multiple votes with stolen or purchased postal ballots. Mail-in ballots also were cast for dead people (Briquet 2016; *World Today News* 2020; *New York Times* 1973).

---

[1] Other countries, including the United Kingdom and France, have learned about the difficulty of uncovering vote fraud firsthand. The United Kingdom's Electoral Commission provides examples of vote fraud for in-person voting (United Kingdom's Electoral Commission 2017, 2018). As the Electoral Commission describes it: "Later in the day the same voter attended again and sought to vote again, this time in his own name. Due to certain physical characteristics of the voter (he was very tall and wore distinctive clothing) and the vigilance of the presiding officer he was suspected of having already voted earlier and formally challenged." In another UK case from 2017, police caught a person voting multiple times only because he bragged openly about it on Twitter.

[2] Eight European countries allow proxy voting as a substitute for absentee ballots, permitting voters to designate others to vote in their names (Lott 2021c). But proxy voting is very strictly regulated by requiring photo IDs and signed request forms. In Poland, the proxy voter must have a power of attorney and obtain special permission from the municipal mayor. In France, voters must visit a municipality's office in person with proof of identity and a reason for absence (for example, a certified letter from an employer or hospital) and then nominate a proxy. Proxy voting avoids the problem of absentee ballots being unsecured. Proxy voting requires that the proxy vote be cast in person in a voting booth. Switzerland is the only European country adopting a comparatively liberal proxy voting policy, requiring only a signature match.

The United Kingdom, which allows postal voting, has uncovered notable mail-in ballot fraud cases. Prior to recent photo ID requirements, six Labor Party councilors in Birmingham won office after what the judge described as a "massive, systematic and organised" postal voting fraud campaign (Britten and Jones 2005). The fraud apparently occurred with the full knowledge and cooperation of local Labour Party officials. "Widespread theft" of postal votes (as many as 40,000 ballots) was documented in areas with large Muslim populations because Labour's organizers were worried that the Iraq war would induce such voters to oppose the incumbent government.

In 1991, Mexico mandated voter photo IDs with thumbprints and banned absentee ballots. The then-governing Institutional Revolutionary Party (PRI) had for many years engaged in fraud and intimidation for mail-in ballots to win elections (Lott 2006). Only in 2006 were absentee ballots again allowed, and then only for voters living abroad who requested them at least six months in advance (McKinley 2005).

Unsecured absentee ballots create the potential of either introducing fraudulent ballots or destroying legitimate ones. Some safeguards, such as matching signatures, can minimize the problems, but even signature requirements are not as secure as requiring voters to present government-issued photo IDs. Nor do the safeguards prevent ballots from being destroyed.

One of the 2020 election controversies was that states like Georgia, Nevada, Pennsylvania, and Wisconsin did not match signatures on the outer mail-in ballot envelopes to voters' official registration records (Navarro 2020). In other states, particularly Pennsylvania, accusations were raised of accepting absentee ballots not enclosed in outer envelopes displaying the voter's signature and that signatures on those envelopes weren't checked (Navarro 2020, pp. 13-14).

While such questions received the most attention in those swing states, they weren't limited solely to them. In Missoula County, Montana, which held an entirely mail-in ballot election in November 2020, an audit found that 4,592 out of the 72,491 mail-in ballots lacked envelopes—representing 6.33% of all votes (Lott 2021b). Absent an officially printed envelope containing registration information, the voter's signature, and a postmark indicating that a ballot had been cast on time, election officials cannot verify a vote's legitimacy. It is a violation of Montana's election laws to count such votes, yet such votes were counted.

Measures to prevent vote fraud are important for ensuring accurate election results and encouraging voter participation in US elections (Lott 2006). For example, after Mexico instituted strict anti-fraud provisions in 1991, voter turnout increased substantially, rising from 59% in the three presidential elections before the reforms to 68% in the three elections afterwards. That increase occurred despite massive stuffing of ballot boxes with fraudulent ballots and registration lists filled with non-existent people (Rossie 1988; Uhlig 1991).[3]

---

[3] It is only in recent years that reports of vote-buying allegations have become more common in Mexico (Associated Press 2005; Associated Press 2018; Linthicum 2018).

Courts frequently have rejected Republican challenges to the 2020 presidential vote, citing the lack of evidence of enough fraud to alter the outcome in a particular state. Republicans sometimes argued that since their observers couldn't watch the vote counts, they couldn't provide such evidence without investigations backed by subpoena power. Still, while some courts agreed that irregularities had occurred in 2020, they weren't willing to grant discovery unless Republicans first presented enough evidence of fraud that could overturn the election. Republicans thus faced a kind of Catch 22.

The following sections provide evidence on the existence of vote fraud in the 2020 US presidential election. First, precinct-level estimates for Georgia and Pennsylvania are inconsistent in revealing vote fraud with absentee ballots, although the evidence of broken rules is more robust for provisional ballots. The results show evidence of vote fraud with absentee ballots when the data from both states are combined. The precinct-level approach can be applied to other states. I then look at all swing states at the county level to ask if counties wherein fraud was alleged experienced higher voter turnout rates.

2. Georgia

The Trump campaign and its supporters focused primarily on what they claimed were voting irregularities in Fulton County, which includes Atlanta.[4] In Georgia's certified ballot count, former Vice President Joe Biden led President Trump by 12,670 votes (BBC 2020). Biden won Fulton County by a margin of 243,904 votes and he out-polled Trump on the country's absentee ballots by 86,309 votes.[5]

Part of the controversy surrounding Fulton County's absentee ballots arises from a burst water pipe at a vote counting facility that resulted in the removal of poll watchers. According to the Chair of the Georgia Republican Party, David J. Shafer, "counting of ballots took place in secret after Republican Party observers were dismissed because they were advised that the tabulation center was shutting down for the night" (Letter dated November 10, 2020, from Doug Collins and David Shafer to Georgia Secretary of State Brad Raffensperger, p. 3).

If election workers processed absentee ballots when Republican observers were not present, can statistical evidence of bias be found in those absentee ballot counts? While in-person voting took place at the precinct level, absentee votes were counted at one common facility at the county level. If the type of fraud that Mr. Shafer worried about actually occurred, it would have affected only Fulton County's absentee ballots.

To examine that possibility, I gathered precinct-level data for Fulton County and four Republican-majority counties that border it wherein no allegations of fraud were raised: Carroll,

---

[4] Donald J. Trump and David J. Shafer v. Brad Raffensperger et al., Fulton County Superior Court, December 4, 2020 (https://cdn.donaldjtrump.com/public-files/press_assets/verified-petition-to-contest-georgia-election.pdf).

[5] November 2020 differed from previous elections in that respect. For example, in 2012, while Obama captured 64% of the total vote in Fulton County, he won barely a majority of the absentee vote (50.89%) (data from Clark Bensen at Polidata).

Cherokee, Coweta, and Forsyth.[6] Precincts adjacent to one another on opposite sides of a county's border are relatively small areas and should be similar demographically. The idea is a simple one: compare Trump's share of absentee ballots in adjacent precincts on opposite sides of a county's border. The comparison in Georgia is between precincts in Fulton and the four other just-named counties as well as between precincts in those four counties that are adjacent to one another. Comparing a county in which fraud was claimed to ones without it is simpler than comparing counties where hard-to-specify, varying degrees of fraud may have materialized.

The model is given as:

$$((A/TA)_i^0 - (A/TA)_i^1) = \text{ß}((P/TP)_i^0 - (P/TP)_i^1) + \delta\ D*X_i + \mathbf{\varepsilon_I},$$

where A = absentee ballots for Trump, TA = total absentee ballots for both candidates, P = in-person votes for Trump, and TP = total in-person votes. The superscripts 0 and 1 indicate adjacent precincts in neighboring counties and the subscript i represents the set of precincts being compared. The binary variable D = 1 if one of the adjacent precincts is in Fulton County (in that case Fulton County is superscript 0); D = 0 otherwise. The error term is $\mathbf{\varepsilon_I}$.

The null hypothesis is that δ = 0, meaning that precinct pairs in which one is a Fulton County precinct are no different from other pairs. The alternative hypothesis is that δ < 0, implying that precinct pairs in which one is the Fulton County precinct undercounted Trump's absentee ballots.

Matching precincts across county lines was a very time-consuming process. While each county (or, in the case of Pennsylvania, each township) had its own precinct maps, comparisons across the political boundaries had to be done by hand. The goal is to compare the precincts of Fulton County that are most similar to precincts in nearby counties with no fraud allegations. In that way, we can isolate the impact of Fulton county's vote-counting process (including potential fraud).

In 2020, Fulton County contains 384 precincts and Cherokee County had 42 precincts. In one case, Fulton County precinct ML02A matches up with four different precincts in Cherokee County (Mountain Road 28, Avery 3, Union Hill 38, and a small portion of Freehome 18).[7]

Any difference in Trump's share of the absentee ballots would not have been caused by the general shift to absentee voting among Democrats, because Democrats encouraged their voters to vote by absentee ballot at both the national and state level. The absentee voting rate shouldn't differ substantially between two precincts next to each other and similar in terms of

---

[6] Corrected data were not available for Fayette County but including those observations resulted in no changes in the statistical significances reported in either Tables 1 or 2.

[7] The other counties are matched west to east and south to north. For a related discussion, see Bronars and Lott 1998.

their in-person voting support and demographics. After all, Democrats cared about winning the state, not a county, or even a precinct.

In precincts with alleged fraud, one would expect Trump's proportion of absentee votes to be depressed, even when such precincts returned in-person Trump vote shares similar to surrounding counties. It would be suspicious if the shift happened only among absentee ballots on opposite sides of a county line.

One challenge when entering a standard intercept term in the model specified above can be the direction of difference between the precincts. If I estimate a typical regression equation with an intercept term, either subtracting control precinct A from B or B from A, the order chosen flips the signs for a given data point. But a way can be found for ensuring the same result no matter the ordering of the between-precinct comparisons: exclude the intercept.[8] The results are robust and unaffected when I drop the intercept term from the regressions.

What is the problem with an intercept for the control precincts? It implies that two identical adjacent precincts have a predictable difference in voting behavior. A linear regression without an intercept is robust because it is symmetric around the origin and thus unaffected by the data points surrounding it.

I ran the test on data from both 2016 and 2020 elections. No accusations of fraud involving absentee ballots were voiced in 2016, so one should expect the absentee ballot percentage for Trump in Fulton County's precincts to behave no differently than the adjacent precincts in Carroll, Cherokee, Coweta, and Forsyth counties. In 2016, the average difference between Trump's share of the absentee ballots and his share of the in-person votes in the adjacent precincts was only 0.75 percentage points.

In Tables 1 and 2, if the estimate of the "Difference in Trump's percentage of the two-candidate in-person vote" between the two adjacent precincts equals one, it means that the differences in the percentage of the in-person vote that Trump received in the adjacent precincts would perfectly track the difference in the absentee ballots. In the estimate for 2016, the coefficient is 0.88, and the 95% confidence interval ranges from 0.684 to 1.076, so it is not statistically different from one.[9] But for the 2020 data, Trump's share of in-person votes did not align as closely with the differences in absentee ballots, which can be seen in the smaller coefficient of the control variable for Trump's share of in-person votes. Indeed, the coefficient for 2020 (0.6059) has a 95% confidence interval of 0.3851 to 0.8267, so it is statistically significantly

---

[8] The geographical ordering convention in (Bronars and Lott 1998) was applied, but by excluding the intercept, that ordering doesn't matter for the results reported herein.

[9] The source for the 2016 precinct borderlines was obtained here: http://rynerohla.com/index.html/election-maps/2016-south-atlantic-republican-primaries-by-precinct/

different from one.[10] However, Trump's shares of absentee ballots cast in Fulton and other counties are not statistically different in 2016 or 2020.

Place Tables 1 and 2 here

Compared to 2016, the in-person vote shares in 2020 explain only about two-thirds of that year's variation in absentee ballot voting in adjacent precincts.

I control for demographic variables to help account for any differences that might still exist. Georgia collects precinct-level information on registered voters' racial and gender demographics by precinct. Table 3 accounts for the differences in the adjacent precincts based on detailed demographic information instead of the change in the in-person difference in Trump's share of the votes. It supplies information on the differences between the precincts in the population percentages of Black males, Black females, Hispanic males, Hispanic females, Asian males, and Asian females. Table 4 enters those variables as well as the "Difference in Trump's percentage of the two-candidate in-person vote." Thus, the table adopts three ways of accounting for differences in Trump's share of the absentee ballot vote: the geographic closeness of relatively small areas (voting precincts), differences in Trump's share of the in-person vote, and differences in the demographics of registered voters.

Place Table 3 here

The results are not consistently significant in answering whether Trump's absentee votes uniformly were lower in precincts bordering Fulton County than in the precincts just across the street in neighboring counties. The estimates for the Fulton County effect range from 0.3% to 11.5%, but only one of those estimates is statistically significant. In nearly all of the estimates, the race and gender demographics variables are not statistically significant, although that is not surprising given how highly correlated those variables are. That makes it difficult to interpret individual coefficients on the demographic variables. However, they are statistically significant as a group in Table 3 (a joint F-test for the demographic variables generates a value of 4.17, which is statistically significant at the 5% level).

That result indicates that the demographic values are worth including and, moreover, that Table 4 is the preferred model. But all models agree that Trump's absentee ballot share was depressed in Fulton County precincts.

Place Table 4 here

The empirical results imply an unusual drop-off in Trump's share of the absentee ballots for Fulton County, which ranges from zero to 11.53 percentage points. Given that 145,267 absentee ballots were cast for Trump and Biden in Fulton County, the largest estimate of 11.53

---

[10] The average difference between Trump's share of the absentee ballots and his share of the in-person votes in the adjacent precincts was -6.04 percentage points.

percentage points equals approximately 16,749 votes, which is 32% more than Biden's margin of victory over Trump.

The average of the various estimates (7.81%) implies an unusual drop in Trump's share of the absentee ballots in Fulton County alone. It amounts to 11,350 votes or 90% of Biden's vote lead in Georgia.

DeKalb County also drew charges of vote-counting irregularities, but it no adjacent Republican counties are available to test those claims. With 128,007 absentee ballots cast for the two major-party candidates in DeKalb, fraud also could have been significant in determining the election outcome.

While some critics of the research methods employed herein argue that errors correlate across precincts within a county,[11] it isn't a concern here. The estimates look at the difference between adjacent precincts across county lines. But, again, there's no apparent reason why Democrats would treat absentee ballots differently in two adjacent precincts having similar political and demographic makeups. Still, I reran the estimates in the first four tables clustering standard errors by county. That change actually makes some of the estimates more statistically significant.[12]

Given that I am measuring how the vote Biden-Trump vote gap between adjacent precincts is changing over time, the results are equally consistent with vote fraud in Republican counties and the destruction of Democrat absentee ballots. But no allegations of such fraud have been made for Republican counties.

If Biden was helped by fraud in Fulton County's in-person voting, the estimates reported herein would underestimate the amount of fraud associated with absentee ballots. Not just in Georgia, but also in Nevada, Pennsylvania, and Wisconsin, critics alleged that large numbers of in-person voters were not legally registered (Navarro 2020). In Fulton County, Georgia, 2,423 voters were not listed on the State's records as registered, and 2,560 felons voted even though they had not completed their sentences.[13,14]

3. Pennsylvania

---

[11] Austan Goolsbee in a Tweet regarding the present research project wrote: "you don't seem to be clustering the standard errors at the county level and instead treating every precinct as though it is independent" (https://twitter.com/Austan_Goolsbee/status/1344361588535521280).

[12] Rerunning the estimates by clustering standard errors by county pairing doesn't fundamentally alter the results. The coefficient for county fraud remains statistically insignificant in Table 1, with a p-value of 0.035%. The t-statistic for Table 2 becomes 0.31 (probability for a two-tailed t-test = 0.79), for Table 3, it is 4.08 (probability for a two-tailed t-test = 0.055), and for Table 4 it is 0.50 (probability for a two-tailed t-test = 0.664). A one-tailed t-test is actually the more appropriate statistical test for the various results that we are showing, but reporting those wouldn't really change which results we note are statistically significant.

[13] The Superior Court of Fulton County, State of Georgia, Trump v. Raffensperger, December 4, 2020. https://www.democracydocket.com/wp-content/uploads/sites/45/2020/12/Trump-v.-Raffensperger.pdf

[14] Survey and other evidence indicate that felons vote almost exclusively for Democrats (Lott 2007b, pp. 182-184).

The Trump campaign focused its allegations of voting irregularities in Allegheny County (which includes Pittsburgh) and Philadelphia County (which matches the boundaries with Philadelphia), although it also raised concerns about the counties surrounding Philadelphia. In Pennsylvania's initial ballot count, former Vice President Joe Biden led President Trump by 81,361 votes. Biden won Allegheny and Philadelphia Counties by margins of 146,706 and 471,305 votes, respectively, and led the absentee vote by margins of 206,505 and 310,553 votes. Unusually large numbers of provisional votes were cast in those same counties, with Biden leading the tallies by 1,489 and 9,045, respectively.

Many concerns arose about possible vote fraud in both counties. Republican poll watchers complained that they were too far away from the ballots to observe the process closely (Li and Ramey 2020; Payne 2020). Trump's lawyers said that voters in Pittsburgh and Philadelphia submitting invalid mail-in/absentee ballots were notified and allowed to correct the defects by casting provisional ballots on Election Day. In contrast, election officials in Republican-leaning counties followed Pennsylvania's election laws more strictly and did not notify voters about similarly defective mail-in/absentee ballots.[15] Complaints arose when voters were allowed to cast provisional votes because they supposedly had requested a mail-in ballot already, even though some voters claimed to have not done so.[16] That process raises concerns that someone other than the registered voter may have voted by submitting an absentee ballot in that person's name.

While affidavits attested to such potential problems, an open question has been whether the issues were widespread enough to alter the 2020 election's outcome.

To explore that question, I adopted the same approach with precinct-level data as I did for Georgia. I collected voting data from adjacent precincts in Allegheny County and the four Republican counties that border it: Beaver, Butler, Washington, and Westmoreland. My comparisons are between Allegheny and those four adjacent counties, and also between the four counties themselves when they are adjacent to one another. However, unlike Georgia, I could obtain the breakdowns of absentee and provisional voting only for Allegheny County in 2020. The estimates reported below thus apply to that single year, In addition, while large-scale fraud allegedly occurred in Philadelphia County, no Republican counties are adjacent to it, thereby limiting the analysis further.

The precincts in the relevant Pennsylvania counties encompass very small, homogenous areas. For example, Allegheny County contains 1,323 precincts – on average, a different precinct every half mile. The more rural, less populous counties also contain large numbers of precincts: Westmoreland 307, Washington 180, Beaver 128, and Butler 111.[17]

---

[15] Rudy Giuliani, Trump Campaign News Conference on Legal Challenges, C-SPAN, November 19, 2020 (https://www.c-span.org/video/?478246-1/trump-campaign-alleges-voter-fraud-states-plans-lawsuits).

[16] See the complaint filed in Trump v Boockvar et al. in the United States District Court for the Middle District of Pennsylvania (p. 48).

[17] https://www.butlercountypa.gov/DocumentCenter/View/1982/Precincts-List?bidId=

The results reported in Table 5 show that in 2020, Trump's percentage of absentee votes was smaller in border precincts within Allegheny County than in the precincts just across the street in neighboring counties. Trump's share was just 0.25 percentage points lower on the Allegheny County side, a difference that is not statistically significant.

To the extent that any fraud occurred with in-person voting, the estimates here will underestimate the prevalence of fraud in absentee ballots.[18]

Place Table 5 here

Because of the concerns expressed about the provisional ballots relied on to solve problems with improperly voted absentee ballots in Allegheny and Philadelphia counties, I compared the previously described adjacent precincts to examine the possibility of provisional ballot fraud.

Table 6 is the same as Table 5, except that it applies to provisional, rather than to absentee votes. The estimate implies a 3.6 percentage point reduction in support for Trump in precincts adjacent to Allegheny County than in Allegheny County itself, but the result is not statistically significant. One obvious reason is the small number of observations: 53 of the 87 precincts have no provisional ballots for Trump and, since one cannot divide by zero, those vote share differences are not defined.[19]

Place Table 6 here

An alternative approach that avoids the loss of those observations is to look at the rates at which provisional ballots were issued. In that case, a clear difference exists. Approximately 1.5% of the votes in border precincts on the Allegheny County side were cast by provisional ballots, which is 3.2 times the 0.48% rate in the surrounding counties' adjacent precincts. That difference is statistically significant at below the 0.1% level on a two-tailed t-test.[20]

Table 7 compares across adjacent precincts the percentages of Biden's votes that were on provisional ballots, after accounting for the same difference for Trump. The share of Biden's votes from provisional ballots is about 0.65 percentage points higher in Allegheny County than in the adjacent precincts, which amounts to roughly 2,800 more votes for Biden. If the same pattern held for Philadelphia, that would mean another 3,925 votes. Adding together the estimates for the two counties, our test implies approximately 6,700 extra ballots for Biden. That total represents about 8.3% of Biden's vote margin in Pennsylvania.

---

[18] Republicans argue that in-person vote fraud is a problem in Pennsylvania. One instance is the 2020 conviction in of a Philadelphia Judge of Elections charged with election fraud for allegedly stuffing ballot boxes on behalf of Democratic candidates in three different races (Meyer 2020). The former president's lawyer, Rudy Giuliani, also claimed that people from New Jersey voted illegally in Philadelphia (Sobey 2020).

[19] I also applied the same method to Georgia data, but with so few places there allowing provisional-ballot voting, only 12 observations remained, none of which were located in Fulton County.

[20] The provisional-ballot rate is slightly higher for the entire county: 1.98%.

Place Table 7 here

As a control, I tried running this comparison for Georgia. No one claims that Fulton County used provisional ballots to correct absentee ballots, so one would not expect a statistically significant result for that state. Indeed, those results were statistically insignificant, with a t-statistic of only 0.27.

Finally, I reproduced the estimates from Tables 5, 6, and 7A with information taken from Polidata on the racial demographics of voting-age populations in the relevant Pennsylvania precincts (see Table 8). While information on voters' gender wasn't available, data from the 2010 Census was available on the differences between the precincts in the voting-age population percentages of Blacks, Hispanics, and Asians. The results are similar to those reported before, although the estimates corresponding to Table 5 become statistically significant at the 10% level on a one-tailed t-test; the results corresponding to Table 7A remain significant, implying a swing to Biden of about 8,500 votes.

Place Table 8 here

Overall, Sections 2 and 3 provide evidence of vote fraud in 2020, but it often isn't statistically significant. The evidence regarding provisional ballots is the strongest.

4. Combining Data for Georgia and Pennsylvania

It made sense to examine Georgia and Pennsylvania separately because some information was available only for one state but not the other. Georgia, for example, published more detailed demographic information at the precinct level than Pennsylvania did. In addition, the allegations regarding provisional ballots were specific to Pennsylvania. But it is possible to combine some of the data for the two states in our examination of absentee ballots. One reason for doing so is to ask whether the failure of some estimates to reach statistical significance arises from the small numbers of observations for each state.

Interestingly, the results reported in Table 9 show statistically significant evidence of large-scale problems with absentee ballots. The estimates presented in column 1 correspond to those reported in Tables 2 and 5, while the estimates in column 2 correspond to the first part of Table 8. I couldn't reproduce Table 3's specification because racial data by gender is not available for Pennsylvania. The one difference from the earlier analyses method is that I now enter fixed effects for the two states.[21]

Place Table 9 here

---

[21] Clustering by county reduces the absolute t-statistic for a two-tailed t-test in column 1 of Table 9 from 2.73 to 1.83 and column 2 from 2.89 to 1.96.

The estimates in Table 9 imply a drop-off of between 3.5 and 3.9 percentage points in Trump's share of absentee ballots for the precincts in the counties where fraud was alleged. For Allegheny and Fulton counties, the models imply that Biden received an extra 10,000 to 11,300 votes. Applying those estimates to Philadelphia County adds another 10,800 to 12,000 more Biden votes, for a total of 20,800 to 23,300.

5. Voter turnout rates

I have accounted for three differences between voting precincts: geographically contiguous areas, in-person vote shares for Trump to account for other political differences, and demographic variables. But it's possible that other differences across county lines may explain variations in how absentee ballots were counted in 2020. It isn't readily apparent what those differences would be because the push for absentee ballots by Democrats appears to have been a statewide and national effort. If once compares two adjacent precincts that have the same demographics and same support for Trump, it would seem that Democrats would undertake the same efforts in both precincts to get absentee votes for Biden. Still, even if the absentee-ballot factor is independent of other voting irregularities, another qualitatively different test might help make alternative explanations less plausible.

Vote fraud can elevate official voter turnout statistics. Fraud can take many forms: higher rates of filling out absentee ballots for people who hadn't voted, dead people voting, ineligible people voting, or even bribes to encourage legally registered people to vote. However, the official turnout rate also will be reduced to the extent that votes for opposing candidates are lost, destroyed, or replaced with ballots filled out for the other candidate.

Republican plaintiffs in the United States District Court for the District of Arizona claimed that up to 94,975 voters returned absentee ballots that were marked as unreturned and, hence, not counted. Peter Navarro's (2020, p. 9) election report characterizes those lost or destroyed ballots as "consistent with allegations of Trump ballot destruction."

No one claims that Republicans systematically encouraged people to submit affidavits about vote fraud, specifically in places where Democratic voter turnout had been heaviest.

Other scholars have looked at voter turnout in Russia as a proxy for fraud (Mebane and Kalinin 2009). They examined high turnout rates and strange patterns in them, such as percentages being reported in round numbers. Research on Bolivia studied unusual reporting of late-counted votes (Idrobo et al. 2020).

Numerous contributions to the literature ask why voters vote. More competitive ("close") political contests generate more voter participation and larger campaign expenditures (Bursztyn et al. 2018; Matsusaka 1993; Lott 2000). Turnout also varies between the two major US political parties based on whether they control other offices (Jung et al. 1994). Voting regulations regarding mail-in ballots, absentee voting, voter IDs, same-day voter registration, and ballot harvesting all can impact voter turnout rates (Alvarez et al. 2008; Lott 2007a).

To examine voter turnout, we start with county-level election participation in six states wherein Trump advisor Peter Navarro (2020) claims voting irregularities occurred in 2020: Arizona, Georgia, Michigan, Nevada, Pennsylvania, and Wisconsin.

For example, a legal challenge in Georgia's Fulton County Superior Court by State Republican Chairman David Shafer and President Donald Trump revealed hundreds of thousands of possible extra votes: 40,279 people who had moved within the state without re-registering; 4,926 voters who registered in another state after they had registered in Georgia; 305,701 people who, according to state records, applied for absentee ballots past the application deadline; 66,247 ineligible voters under 17 years of age; 2,560 felons; 8,718 who remained on the voting rolls after they were dead; and 2,423 who were not listed on the state's voter rolls.[22]

In Nevada, more than 42,000 people voted more than once.[23] In testimony before the Senate Hearing on Election Security and Administration, Jesse Banal, lead counsel for the Trump Campaign, compiled the list by reviewing voter registrations and finding the same name, address, and birthdate for some registered voters. In some cases, two registrants might have the same last name, same birthdate, and same address, but one is "William" and the other "Bill" – they, in fact, are the same person. More than 1,500 dead people allegedly voted. Another 19,000 votes didn't live in the state (excluding military personnel or students). In excess of 1,000 voters listed non-existent addresses.

Similarly, 28,395 people allegedly voted without identification in Madison and Milwaukee, Wisconsin. Republican lawyers claimed that 200,000 absentee ballots did not submit the proper signatures.[24] Payments to Native Americans to vote allegedly were "orchestrated by the Biden campaign . . . [with] Visa gift cards, jewelry, and other 'swag'" (Navarro 2020, p. 8; Bedard 2020).

Another reason for higher turnouts could be a much lesser absentee ballot rejection rate. For example, Ballotpedia notes that in the 2016 general election, Georgia rejected 6.42% of absentee ballots, but that rate was only 0.60% in 2020 – a difference of about 76,971 votes.[25] Other swing states also saw drops in rejection rates, although they were much smaller than Georgia's. Pennsylvania's rejection rate went from 0.95% in 2016 to 0.28% in 2020, a difference

[22] Donald J. Trump and David J. Shafer v Brad Raffensperger et al., Fulton County Superior Court, December 4, 2020 (https://cdn.donaldjtrump.com/public-files/press_assets/verified-petition-to-contest-georgia-election.pdf).
[23] Senate Hearing on Election Security and Administration, December 16, 2020 (https://www.c-span.org/video/?507292-1/senate-hearing-election-security-administration).
[24] Ibid.
[25] Election results, 2020: Analysis of rejected ballots, Ballotpedia, December 23, 2020 (https://ballotpedia.org/Election_results,_2020:_Analysis_of_rejected_ballots). The number of absentee ballots cast (1,322,529) is from the Georgia Secretary of State's website (https://sos.ga.gov/index.php/elections/number_of_absentee_ballots_rejected_for_signature_issues_in_the_2020_election_increased_350_from_2018).

of 17,361 votes.[26,27] Nevada's rejection rate dropped by 0.6 percentage points, a difference of 4,143 votes. Michigan is the only other swing state for which Ballotpedia measures rejected absentee ballots; its rate essentially was unchanged from 2016 to 2020, falling from 0.49% to 0.46%.

To test whether counties with alleged fraud had higher voter turnout rates, I examine the change in voter turnouts between the 2016 and 2020 general elections, by county, for the following swing states: Arizona, Florida, Georgia, Iowa, Michigan, Nevada, North Carolina, Ohio, Pennsylvania, and Wisconsin.[28] I focus on swing states because I expect turnout rates to be higher there, given that election campaigns tend to focus their resources there and, moreover, voters are more motivated to turn out. I attempt partially to account for other cross-state differences by comparing turnout rates in 2016 and 2020. But even within the sampled states, three were rated the closest "toss-ups" by *USA Today*: Arizona, North Carolina, and Pennsylvania. I also control for those three states separately.

The question is whether a larger increase in turnout rates is observed in the counties where voter fraud was alleged relative to other counties. I already have discussed counties where vote fraud was alleged in Georgia (Fulton and DeKalb) and Pennsylvania (Allegheny, Centre, Chester, Delaware, Montgomery, Northampton, and Philadelphia). Other counties were under critics' spotlights in Arizona (Apache, Coconino, Maricopa, and Navajo) (Davidson 2020; Smith 2020), Michigan (Wayne), Nevada (Clark and Washoe) (Bedard 2020), and Wisconsin (Dane, Menominee, and Milwaukee) (Bauer 2020).[29]

To examine differences in county turnout rates, I control for each county's turnout rate when Trump ran in 2016. I also look at how heavily Republican or Democrat the counties are, based on whether they voted for Trump or Biden. I classify those counties that Trump carried as Republican counties and Biden's as Democratic ones. Since the turnout change may differ for Democratic and Republican counties, I identify the counties supporting Trump and Biden with

[26] I obtained the number of absentee ballots cast in Pennsylvania for Biden and Trump from Pennsylvania's Secretary of State (https://www.electionreturns.pa.gov/General/SummaryResults?ElectionID=83&ElectionType=G&IsActive=1).

[27] While it isn't necessary for the results shown here, a higher turnout rate also could also show up in manufacturing false ballots. A possible example occurred in Atlanta, where, as noted, election officials ordered ballot-counting stopped because of a water leak (Chung 2020). The officials reportedly told observers that vote-counting would start up again the next morning. Then once poll watchers, observers, and the media left, the vote-counting continued with surveillance video catching large boxes of ballots being pulled from underneath a draped table. (Trump Campaign lawyers present video 'evidence' of ballot fraud, Senate Judiciary Subcommittee, December 4, 2020 https://www.youtube.com/watch?v=LJ0xDWhWUxk).

On the other hand, Fulton County Elections Director Richard Barron, a Democrat, claims that observers decided on their own to leave the building in Atlanta (Associated Press 2020). Similarly, Gabriel Sterling, Georgia's voting system implementation manager, says that Georgia Secretary of State's investigators were present even if political observers weren't on the spot (https://twitter.com/GabrielSterling/status/1334825233610633217?s=20).

[28] I exclude the two states regarded as longshot swing states (New Hampshire and Texas) (King 2020).

[29] See also https://web.archive.org/web/20201111220325/https://www.facebook.com/permalink.php?story_fbid=1539297286270372&id=573103029556474.

two separate variables. When Biden won a county, the values for the Republican variable are zero. Similarly, when Trump won, the values for the Democratic variable are zero. Elsewhere those variables equal Trump's share of the vote minus Biden's share. Since I don't have strong prior beliefs that the change in turnout is linear with respect to how partisan the county was, I try including the squares of the partisan county measures. Enthusiasm for the two major candidates might be increasing at either an increasing or a decreasing rate, but entering the squared terms does not affect the alleged fraud variable (see Table 10).

Place Table 10 here

The regression takes the form: (Turnout Rate$_i^{2020}$ - Turnout Rate$_i^{2016}$) = β_1(Vote Fraud County Dummy$_i$) + β_2(Vote Share in Republican Counties Trump Won)$_i$ + β_3(Vote Share in Democrat Counties Biden Won)$_i$ + β_3(Other Factors)$_i$ + α + ε_i, where α is the constant and $\boldsymbol{\varepsilon}$ is the error term. The null hypothesis is that $\beta_1 = 0$.

I rely on data from the US Census Bureau's 2019 American Community Survey for median household income, the percentage of the population that is female, different racial groups (Black, Hispanic, Asian, and two or more races), the highest level of education attained (high school graduate, some college, bachelor's degree, and graduate or professional degree), and by ten-year interval age groups (18 to 19, 20 to 29, 30 to 39, 40 to 49, 50 to 59, 60 to 69, 70 to 79, 80 and up). As voter turnout rates vary by gender, race, and age, all of the estimates account for those simple demographics.

Finally, the last regression in Table 11 includes state fixed effects. The state dummy variables may serve as a proxy for how states changed their election procedures between 2016 and 2020. However, any procedural changes implemented, such as mail-in ballots, would have affected all counties in a given state, making it more difficult to detect increases in fraud in the counties where fraud was alleged.

Place Table 11 here

The estimates displayed in Table 11 start from the simplest specification to one with more controls. They imply that the counties with alleged vote fraud had between 142,000 and 368,000 excess votes – in other words, a 1.1 to 3.1 percentage point increase in turnout. The county fraud variable's coefficient is statistically significant in every case, at least at the 8% level on a two-tailed t-test. All estimates cluster standard errors by state (see Moulton 1986, but a more skeptical approach is provided by Abadie 2017).

The first specification indicates that the more heavily Republican counties experienced higher voter turnout rates in 2020 than in 2016, although the effect is not statistically significant. More heavily Democratic counties show a slight drop in turnout, except for the counties wherein fraud was alleged, but the change is not statistically significant in the Democratic counties where fraud wasn't alleged. The F-test implies that Democratic and Republican counties did not behave differently in terms of voter turnout rates. The joint F-test for the gender, race, and age

variables is statistically significant at the 0.35% level. The second specification adds the other demographic information on income and education, but those independent variables have little impact on the results. The size and statistical significance of the fraud coefficient are reduced slightly. But the primary effect is to lower the statistical significance of the joint F-test for the demographic variables, so that it is now significant at the 8.9% level.

The next four estimates look at the percentages of Democratic or Republican voters in the counties and their squared values. The voter turnout rate increased at a decreasing rate in the most Republican counties and fell at an increasing rate in the most Democratic ones, but neither effect is statistically significant. The joint F-tests also indicate that the results are not statistically significant.

The fourth specification adds the changes in turnout between the 2012 and 2016 presidential elections to explain the change between 2016 and 2020, but that is not statistically significant and doesn't affect the other estimates. The fifth specification includes a dummy for the states rated "toss-ups" by *USA Today* on the morning of the November 3rd, 2020, election. That variable implies a three-percentage point larger increase in turnout in the swing states (Arizona, North Carolina, and Pennsylvania) and, moreover, the marginal effect is statistically significant at better than the 0.01% level. Including that variable cuts the size of the fraud effect in half, although it still significant at the 4% level or better on a two-tailed t-test.[30] Including state fixed effects cuts the size of the fraud variable further, but it remains statistically significant at the 8% level on a two-tailed t-test and implies an excess of 144,000 votes.

As one test of the sensitivity of sample selection, I re-estimated the models in Table 11 on just the two states that I examined earlier – Georgia and Pennsylvania – as well as the controlled-for swing states (Florida, North Carolina, and Ohio) (see Table 12). In all but one of the specifications, the results imply even larger increases in turnout in counties allegedly engaging in vote fraud; the marginal effects consistently are statistically significant at least at the 8% level on a two-tailed t-test. In the first five specifications, the estimates on the county-level fraud variable imply excess votes for Biden of between 2.21 and 4.34 percentage points, or about 70,000 to 79,000 votes. The total combined win margin for Biden in Georgia and Pennsylvania was 92,334.

Place Table 12 here

Again, the results reported above underestimate the extent of fraud if votes for opposing candidates are either lost, destroyed, or replaced with ballots filled out for the other candidate. It would also underestimate the amount of fraud to the extent that illegal votes are being cast in the comparison counties.

---

[30] I also ran a specification analogous to the fourth specification with census data but entering a dummy for the three states that were rated immediately before the election as the most likely toss-ups. The results remained stable, although the coefficient on the county fraud dummy is reduced from 0.015 (significant at the 3.8% level) to 0.012 (with a t-statistic of 2.53) – a value that is midway between those shown in specifications 4 and 6, implying about 143,000 excess Biden votes in the counties where fraud was alleged.

6. Subsequent research

Eggers et al. (2021) respond to two of the three tests presented here. They object to a December 29th, 2020, version of the current paper where I entered an intercept term in analyzing the precinct-level data. As discussed above, those initial specifications were influenced by the ordering of precinct comparisons (whether one subtracts control precinct A from B or subtracts B from A). But the January 6th, 2021, revised version of the paper corrected that problem by removing the intercept term. Unfortunately, despite being immediately sent a copy of the revised study and not submitting their paper until the end of February 2021, Eggers et al. completely ignored the revised discussion for mail-in/absentee ballots. They have offered no response to the statistically significant results regarding provisional ballots.

Finally, my regressions in the January 6th version of my paper used the change in voter turnout rates between the 2016 and 2020 elections. As shown previously, even with many other determinants held constant, alleged fraud has a statistically significant and large effect on voter turnout. The authors ignored that information, despite having been made aware of it when I engaged in a public debate with one of the authors in competing mid-February 2021 op-eds in the *Washington Times* before they submitted their paper for review (Grimmer and Hall 2021; Lott 2021a).

7. Conclusion

The precinct-level estimates for Georgia and Pennsylvania provide some evidence of vote fraud in the 2020 US presidential election, but the sizes and statistical significances of the effects are not consistent. However, combining the Georgia and Pennsylvania samples implies additional mail-in/absentee votes for Biden. Allegheny and Fulton Counties alone account for at least 10,000 votes The evidence is much stronger for concluding that Democrats treated so-called provisional ballots differently than Republicans did.

The results reported herein might underestimate the extent of vote fraud because they assume that no fraud occurred with in-person voting. Yet, given that both results measure how the Biden-Trump vote gap between adjacent precincts is changing over time, they are equally consistent with more vote fraud in Republican counties than in Democratic counties. Nevertheless, no such misbehavior in Republican counties has been alleged either in terms of fraud or treating provisional ballots differently in 2020 than in previous elections.

The voter turnout rate data provide stronger evidence of significant excess Biden votes in Arizona, Georgia, Michigan, Nevada, Pennsylvania, and Wisconsin. The estimates imply that the counties wherein vote fraud was alleged returned between 142,000 and 368,000 excess Biden votes. While the findings reported here are dramatic, they may be underestimates because the voter turnout estimates do not account for ballots cast for the opposing candidate that are lost, destroyed, or replaced with ballots filled out for the other candidate. While it would involve a

lot more work, it would be possible to apply the precinct level tests to compare turnout rates in adjacent precincts across all the counties where fraud was alleged.

Acknowledgements

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When should you adjust standard errors for clustering? *MIT Working Paper* (October).

Alvarez, R. M., Bailey, D., & Katz, J. N. (2008). The effect of voter identification laws on turnout. *Social Science Research Network.*
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1084598

Associated Press. (2005). Mexican senate approves mail-in absentee ballots for Mexicans living abroad. *The Arizona Republic*, April 28.
http://www.azcentral.com/specials/special03/articles/0428mexicovote-ON.html

Associated Press. (2018). Observers say vote buying strong ahead of Mexico election. *Associated Press*, June 26. https://apnews.com/article/cd1f3c2f00584bdb896f54701dcb1555

Associated Press. (2020). Surveillance tape of vote counting breeding false fraud claims in Georgia. *HuffPost*, December 4. https://www.huffpost.com/entry/video-georgia-election-false-fraud_n_5fcac976c5b619bc4c330575?guccounter=1&guce_referrer=aHR0cHM6Ly9kdWNrZHVja2dvLmNvbS8&guce

Bauer, K. (2018). Chicago and rigged elections? The history is even crazier than you've heard. *Block Club Chicago*, October 24.

Bauer, S. (2020). Wisconsin issues recount order in 2 counties as Trump wanted. *Associated Press*, November 19. https://apnews.com/article/wisconsin-recount-2-counties-f408a7b43deb96e2ac7ff0b24a2f968a

BBC. (2020). US election 2020: Biden certified Georgia winner after hand recount. *BBC*, November 20. https://www.bbc.com/news/election-us-2020-55006188

Bedard, P. (2020). Pro-Biden effort offered Native Americans $25-$500 Visa gift cards and jewelry to vote. *Washington Examiner*, December 3.
https://www.washingtonexaminer.com/washington-secrets/pro-biden-effort-offered-native-americans-25-500-visa-gift-cards-jewelry-to-vote

Brennan, G., & Buchanan, J. (1984). Voter choice: Evaluating political alternatives. *American Behavioral Scientist, 28*(2), 185-201.

Briquet, J. L. (2016). Expatriate Corsicans and the Vote au Village: Mechanisms of control and expressions of sentiment (Nineteenth-Twentieth Centuries), *Revue française de science politique (English Edition), 66*(5), 43-63.

Britten, N., & Jones, G. (2005). Judge lambasts postal ballot rules as Labour 6 convicted of poll fraud. *The Telegraph (United Kingdom)*, April 5*.* https://www.telegraph.co.uk/news/uknews/1487144/Judge-lambasts-postal-ballot-rules-as-Labour-6-convicted-of-poll-fraud.html

Bronars, S. G., & Lott, J. R. (1998). Criminal deterrence, geographic spillovers, and the right to carry concealed handguns. *The American Economic Review, 88*(2), 475-479.

Bursztyn, L., Cantoni, D., Funk, P., & Yuchtman, N. (2018). Polls, the press, and political participation: The effects of anticipated election closeness on voter turnout. *University of Chicago Working Paper.*

Caro, R. A. (1982). *The Path to Power: The years of Lyndon Johnson*. New York: Alfred a Knopf Incorporated.

Caro, R. A. (1991). *Means of Ascent: The years of Lyndon Johnson*. New York: Alfred a Knopf Incorporated.

Carter, J., & Baker, J. (2005). Building Confidence in U.S. Elections: Report of the Commission on Federal Election Reform. *Center for Democracy and Election Management, American University*. https://www.legislationline.org/download/id/1472/file/3b50795b2d0374cbef5c29766256.pdf

Chung, F. (2020). 'Slow leak': Text messages cast doubt on Georgia officials' 'burst pipe' excuse for pause in counting. *News.com.au*, November 12*.* https://www.news.com.au/world/north-america/us-politics/slow-leak-text-messages-cast-doubt-on-georgia-officials-burst-pipe-excuse-for-pause-in-counting/news-story/19176f5113512210517c82debe684392

Crime Prevention Research Center. (2020). UPDATED: Examples of Vote Fraud for Mail-in Ballots. *Crime Prevention Research Center*, May 27. https://crimeresearch.org/2020/05/examples-of-vote-fraud-for-mail-in-ballots/

Davidson, J. (2020). In Nevada, a corrupt cash-for-votes scheme is hiding in plain sight. *The Federalist*, November 18. https://thefederalist.com/2020/11/18/in-nevada-a-corrupt-cash-for-votes-scheme-is-hiding-in-plain-sight/

Eggers, A. C., Garro, H., and Grimmer, J. (2021). No evidence for systematic voter fraud: A guide to statistical claims about the 2020 election. *Proceedings of the National Academy of Sciences of the United States of America, 118*(45).

Electoral Commission. (2017). Analysis of cases of alleged electoral fraud in the UK in 2016: Summary of data recorded by police forces. *UK Parliament* (March). https://www.electoralcommission.org.uk/sites/default/files/pdf_file/Fraud-allegations-data-report-2016.pdf

Electoral Commission. (2018). Analysis of cases of alleged electoral fraud in the UK in 2017: Summary of data recorded by police forces. *UK Parliament*. https://www.electoralcommission.org.uk/sites/default/files/pdf_file/Fraud-allegations-data-report-2017.pdf

Flis, J. & Kaminski, M. (2020). Multi-player electoral engineering and COVID-19 in the Polish presidential elections in 2020. *R&R Mind and Society.*

Goel, S., Meredith, M., Morse, M., Rothschild, D., & Shirani-Mehr, H. (2020). One person, one vote: Estimating the prevalence of double voting in US presidential elections. *American Political Science Review, 114*(2) 1-14.

Grimmer, J., & Hall, A. B. (2021). Dismantling Trump's election fraud claims. *Washington Times*, February 8. https://www.washingtontimes.com/news/2021/feb/8/dismantling-trumps-election-fraud-claims/

Heritage Foundation. (2022). A sampling of election fraud cases from across the country. Online database at Heritage.org/votefraud. *Heritage Foundation*. Last checked January 3, 2022.

Idrobo, N., Kronick, D., & Rodriguez, F. (2020). Do shifts in late-counted votes signal fraud? Evidence from Bolivia. *Social Science Research Network*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3621475

Jung, G. R., Kenny, L. W., & Lott, J. R. (1994). An explanation for why senators from the same state vote differently so frequently. *Journal of Public Economics, 54*(May), 65-96.

King, L. (2020). Battleground states: These are the 12 states that will determine the 2020 election. *USA Today*, November 3. https://www.usatoday.com/story/news/politics/elections/2020/11/03/key-states-watch-include-south-carolina-georgia-arizona-and-iowa-democrats-must-win-four-seats-take/6085491002/

Li, S. and Ramey, C. (2020). What are election observers? Role at crux of Trump lawsuits in Pennsylvania. *Wall Street Journal*, November 10. https://www.wsj.com/articles/what-are-election-observers-the-role-at-the-crux-of-trump-lawsuits-in-pennsylvania-11605053759

Linthicum, K. (2018). In Mexico, voters are wooed with gift cards, washing machines and cheap tortillas. *Los Angeles Times*, February 28. https://www.latimes.com/world/la-fg-mexico-tortilla-war-20180228-story.html

Lott, J. R. (2000). A simple explanation for why campaign expenditures are increasing: The government is getting bigger. *The Journal of Law and Economics, 43*(2), 359-393.

Lott, J. R. (2006). Evidence of voter fraud and the impact that regulations to reduce fraud have on voter participation rates. *Social Science Research Network* (August)*.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=925611

Lott, J. R. (2007a). Evidence of voter fraud and the impact that regulations to reduce fraud have on voter participation rates. *Social Science Research Network*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=925611

Lott, J. R. (2007b). *Freedomnomics*. Washington, D.C.: Regnery Publishing.

Lott, J. R. (2020a, April 10). Heed Jimmy Carter on the Danger of Mail-in Voting. *Wall Street Journal.*

Lott, J. R. (2020b). Why do most countries ban mail-in ballots?: They have seen massive vote fraud problems. *Social Science Research Network* (October). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3666259

Lott, J. R. (2021a). Taking another look at election fraud. *Washington Times*, February 15. https://www.washingtontimes.com/news/2021/feb/15/taking-another-look-at-election-fraud/

Lott, J. R. (2021b). A river of doubt runs through mail voting in Montana. *Real Clear Investigations*, March 24. https://www.realclearinvestigations.com/articles/2021/03/24/a_river_of_doubt_runs_through_mail_voting_in_big_sky_country_769321.html

Lott, J. R. (2021c). Why do most countries require photo voter I.D.s?: They have seen massive vote fraud problems. *Social Science Research Network* (May). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3849068

Lott, J. R., & Kenny, L. W. (1999). How dramatically did women's suffrage change the size and scope of government? *Journal of Political Economy, 107*(6, pt.1), 1163-1198.

Matsusaka, J. G. (1993). Election closeness and voter turnout: Evidence from California ballot propositions. *Public Choice, 76*, 313-334.

McKinley, J. C. (2005, June 29). Lawmakers in Mexico approve absentee voting for migrants. *New York Times*. https://www.nytimes.com/2005/06/29/world/americas/lawmakers-in-mexico-approve-absentee-voting-for-migrants.html

Mebane, W. R., & Kalinin, K. (2009). Comparative election fraud detection. *Social Science Research Network*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450078

Meyer, K. (2020). Philly judge of elections pleads guilty to election fraud, accepting bribes. *WHYY NPR*, May 21. https://whyy.org/articles/philly-judge-of-elections-pleads-guilty-to-election-fraud-accepting-bribes/

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics, 32*(3), 385-397.

Navarro, P. (2020). The immaculate deception: Six key dimensions of election irregularities. *The Navarro Report*, December 15*.*

New York Times. (1973). Corsicans of France are feeling the sting of publicity given to criminals. *New York Times*, January 7. https://www.nytimes.com/1973/01/07/archives/corsicans-of-france-are-feeling-the-sting-of-publicity-given-to.html

Payne, D. (2020). Pennsylvania poll watcher: 'We literally had no input and no ability to watch anything'. *Just the News*, November 9. https://justthenews.com/politics-policy/elections/pennsylvania-poll-watcher-we-literally-had-no-input-and-no-ability-watch

Rossie, C. (1988). 'Tacos' stuffed with fraudulent ballots with PM-Mexico-Elections, *Associated Press*, July 7. https://apnews.com/article/4be7794766c41f503e5a29f7d6e72034

Smith, A. V. (2020). How indigenous voters swung the 2020 election. *High Country News*, November 6. https://www.hcn.org/articles/indigenous-affairs-how-indigenous-voters-swung-the-2020-election

Sobey, R. (2020). Rudy Giuliani claims Trump campaign has found nationwide Democrat voter fraud conspiracy plot. *Boston Herald*, November 19. https://www.bostonherald.com/2020/11/19/rudy-giuliani-trump-campaign-has-found-nationwide-voter-fraud-conspiracy-plot/

World Today News. (2020). In Corsica, the tormented history of the vote by correspondence. *World Today News*, June 15. https://www.world-today-news.com/in-corsica-the-tormented-history-of-the-vote-by-correspondence/

Uhlig, M. (1991). Mexico overhauls electoral rolls. *New York Times*, January 23. https://www.nytimes.com/1991/01/23/world/mexico-overhauls-electoral-rolls.html

Viebeck, E. (2020). Minuscule number of potentially fraudulent ballots in states with universal mail voting undercuts Trump claims about election risks. *Washington Post*, June 8*.* https://www.washingtonpost.com/politics/minuscule-number-of-potentially-fraudulent-ballots-in-states-with-universal-mail-voting-undercuts-trump-claims-about-election-risks/2020/06/08/1e78aa26-a5c5-11ea-bb20-ebf0921f3bbd_story.html

Wu, J., Yorgason, C., Flosz, H., Handan-Nader, C., Myers, A., Nowacki, T., Thompson, D. M., Yoder, J., and Hall, A. B. (2020). Are dead people voting by mail? Evidence from Washington state administrative records. *Stanford University Democracy & Polarization Lab Working Paper.*

Table 1: 2016 Difference in Trump's share of the absentee ballot vote between adjacent precincts at the border of Fulton, Carroll, Cherokee, Coweta, and Forsyth counties		
Control variables	Coefficient	Level of statistical significance for a two-tailed t-test
Difference in Trump's percentage of the two-candidate in-person vote between two precincts	0.8800	0.000
Fulton County Effect	-0.0215	0.466
Number of Observations 45	F-statistic = 60.36 Level of significance = 0.0000	R-Squared = 0.7253

Table 2: 2020 Difference in Trump's share of the absentee ballot vote between adjacent precincts at the border of Fulton, Carroll, Cherokee, Coweta, and Forsyth counties

Control variables	Coefficient	Level of statistical significance for a two-tailed t-test
Difference in Trump's percentage of the two-candidate in-person vote between two precincts	0.6059	0.0000
Fulton County Effect	-0.00282	0.891
Number of Observations 22	F-statistic = 58.50 Level of significance = 0.0000	R-Squared = 0.8540

Table 3: 2020 Difference in Trump's share of the absentee ballot vote after adjusting for racial and gender demographics of registered voters

Control variables	Coefficient	Level of statistical significance for a two-tailed t-test
Fulton County Effect	-0.1153	0.011
Difference in the percent of voters who are Black males	1.6396	0.528
Difference in the percent of voters who are Black females	-1.8755	0.300
Difference in the percent of voters who are Hispanic males	-4.4266	0.196
Difference in the percent of voters who are Hispanic females	2.7631	0.394
Difference in the percent of voters who are Asian males	1.1089	0.534
Difference in the percent of voters who are Asian females	-2.3922	0.241
Number of Observations 22	F-statistic = 7.48 Level of significance = 0.0006	R-Squared = 0.7774

Table 4: 2020 Difference in Trump's share of the absentee ballot vote after adjusting for racial and gender demographics of registered voters and the difference in the in-person vote

Control variables	Coefficient	Level of statistical significance for a two-tailed t-test
Difference in Trump's percentage of the two-candidate in-person vote between two precincts	0.8846	0.0000
Fulton County Effect	-0.0225	0.568
Difference in the percent of voters who are Black males	-0.5052	0.768
Difference in the percent of voters who are Black females	0.8265	0.519
Difference in the percent of voters who are Hispanic males	-3.5121	0.116
Difference in the percent of voters who are Hispanic females	3.7800	0.082
Difference in the percent of voters who are Asian males	0.33894	0.767
Difference in the percent of voters who are Asian females	-0.9173	0.487
Number of Obs = 22	F-statistic = 18.98 Level of significance = 0.0000	R-Squared = 0.9156

Table 5: 2020 Difference in Trump's share of the absentee ballot vote between adjacent precincts at the border of Allegheny, Beaver, Butler, Washington, and Westmoreland counties		
Control variables	Coefficient	Level of statistical significance for a two-tailed t-test
Difference in Trump's percentage of the two-candidate in-person vote in the adjacent precincts	0.3068	0.0000
Allegheny County Effect	-0.0025	0.770
Number of Observations 87	F-statistic = 11.16 Level of significance = 0.0000	R-Squared = 0.2080

Table 6: 2020 Difference in Trump's share of the provisional ballots between adjacent precincts at the border of Allegheny, Beaver, Butler, Washington, and Westmoreland counties		
Control variables	Coefficient	Level of statistical significance for a two-tailed t-test
Difference in Trump's percentage of the two-candidate in-person vote in the adjacent precincts	1.0554	0.065
Allegheny County Effect	-0.0362	0.417
Number of Observations 34	F-statistic = 3.13 Level of significance = 0.0571	R-Squared = 0.1638

Table 7: 2020 Difference in the share of Biden's votes from provisional ballots in adjacent precincts

A) Examining Allegheny, Beaver, Butler, Washington, and Westmoreland counties

Control variables	Coefficient	Level of statistical significance for a two-tailed t-test
Difference in the share of Trump's votes from provisional ballots in the adjacent precincts	0.3855	0.000
Allegheny County Effect	0.0065	0.008
Number of Observations 87	F-statistic = 38.71 Level of significance = 0.0000	R-Squared = 0.4767

B) Examining Fulton, Carroll, Cherokee, Coweta, Fayette, and Forsyth counties

Difference in the share of Trump's votes from provisional ballots in the adjacent precincts	-0.1442	0.772
Fulton County Effect	0.0312	0.795
Number of Observations 22	F-statistic = 23.60 Level of significance = 0.0000	R-Squared = 0.7130

Regression Estimate	Coefficient on the Allegheny County Effect	Level of statistical significance for a two-tailed t-test	
			Table 8: Re-estimating Tables 5, 6, and 7A by including Census 2010 precinct demographic data on difference in the percent of the voting-age population who are Black, Hispanic, and Asian
Table 5	-0.0125	0.166	Number of Obs = 87 F-statistic = 6.61 Level of significance F-test = 0.0000 R-square = 0.2871
Table 6	-0.04196	0.376	Number of Obs = 34 F-statistic = 2.70 Level of significance F-test = 0.0400 R-square = 0.3180
Table 7A	0.0057	0.036	Number of Obs = 87 F-statistic = 15.13 Level of significance F-test = 0.0000 R-square = 0.4798

Table 9: Combining data for Georgia and Pennsylvania to examine the difference in the share of Biden's votes from provisional ballots in adjacent precincts, specifications analogous to those in Tables 2, 5, and 8

(the level of significance for a two-tailed t-test are in parentheses)

Endogenous variable: Difference in Trump's share of the absentee ballot vote between adjacent precincts		
Control variables	(1)	(2)
County where Fraud alleged	-.034725 (0.007)	-.03866 (0.005)
Difference in the share of Trump's votes from provisional ballots in the adjacent precincts	0.44896 (0.000)	0.51044 (0.000)
Difference in Percent of the Voting Age Population who are Black		0.14833 (0.228)
Difference in Percent of the Voting Age Population who are Hispanic		-.50884 (0.540)
Difference in Percent of the Voting Age Population who are Asian		-.58047 (0.027)
Georgia state dummy	0.04197 (0.001)	0.05289 (0.000)
State Fixed Effects	Yes	Yes
Number of Observations = 109	R2 = 0.5337 F-test level of significance = 0.0000	R2 = 0.5706 F-test level of significance = 0.0000

Table 10: Comparing voter turnout rates in 2020 swing states (Arizona, Florida, Georgia, Michigan, Nevada, North Carolina, Ohio, Pennsylvania, and Wisconsin) Observations = 668

Variable	Mean	Standard Deviation
Percent Voter Turnout in 2020 Election	0.7502149	0.0704998
Percent Voter Turnout in 2016 Election	0.6979785	0.0757554
Republican Counties (Trump's minus Biden's share of votes)	0.18628	0.21074
Republican Counties (Trump's minus Biden's share of votes squared)	0.0790	0.1228
Democrat Counties (Trump's minus Biden's share of votes)	-0.1369	0.200619
Democrats Counties (Trump's minus Biden's share of votes squared)	0.05894	0.10930
County where Fraud alleged	0.02844	0.1664

Table 11: Did counties accused of fraud have an unusual increase in voter turnout? (Arizona, Florida, Georgia, Michigan, Nevada, North Carolina, Ohio, Pennsylvania, Wisconsin)
(Clustering by state, the level of significance for a two-tailed t-test are in parentheses, results for percent in different age groups are not shown)

Control variables	(1)	(2)	(3)	(4)	(5)	(6)
Endogenous variable: The change in voter turnout rate between 2020 and 2016						
County where Fraud alleged	0.0308095 (0.013)	0.02964 (0.020)	0.02957 (0.018)	0.03008 (0.019)	0.01487 (0.038)	0.01136 (0.078)
Republican Counties (Trump's minus Biden's share of votes)	0.0135 (0.498)	0.01539 (0.342)	0.03472 (0.266)	0.03455 (0.286)	0.02408 (0.487)	-0.0006199 (0.980)
Republican Counties (Trump's minus Biden's share of votes squared)			-0.02377 (0.637)	-0.02123 (0.713)	-0.003800 (0.940)	-0.005518 (0.926)
Probability for Joint F-test for Republican Counties			0.2745	0.3066	0.1453	0.9804
Democrat Counties (Trump's minus Biden's share of votes)	0.008405 (0.593)	0.00755 (0.625)	-0.02767 (0.520)	-0.02604 (0.519)	-0.005774 (0.841)	-0.02089 (0.344)
Democrat Counties (Trump's minus Biden's share of votes squared)			-0.05635 (0.333)	-0.05457 (0.321)	-0.01698 (0.649)	-0.02209 (0.421)
Probability for F-test for how turnout rates vary differently between heavily Democratic and Republican counties	0.4537	0.6182				
Joint F-test for Democrat Counties			0.3527	0.3507	0.5930	0.5947
Percent Voter Turnout change between 2012 and 2016 Elections				0.01525 (0.808)		
Dummy for states rated as "Toss-ups" by USA Today the morning of Nov. 3 2020					0.029997 (0.000)	
Percent Female	0.2022785 (0.145)	0.17352 (0.179)	0.16992 (0.135)	0.17069 (0.136)	0.052754 (0.458)	0.0011253 (0.989)

Percent Black	0.011938 (0.443)	0.0016099 (0.949)	0.0013234 (0.957)	0.0044581 (0.827)	0.0069364 (0.820)	-0.037937 (0.073)
Percent Hispanic or Latino	-0.032705 (0.149)	-0.04526 (0.116)	-0.04666 (0.118)	-0.04713 (0.117)	-0.06818 (0.009)	-0.02579 (0.357)
Percent Asian	-0.23748 (0.326)	-0.2908 (0.143)	-0.2889 (0.132)	-0.2829 (0.104)	-0.1904 (0.162)	-0.1721 (0.060)
Percent two or more races	-3.81e-07 (0.168)	-3.71e-07 (0.197)	-3.62e-07 (0.217)	-3.66e-07 (0.219)	-2.40e-07 (0.314)	4.59e-08 (0.687)
Median household income		-2.27e-07 (0.449)	-2.23e-07 (0.479)	-2.12e-07 (0.466)	-1.75e-07 (0.493)	5.57e-08 (0.745)
Percent High School Graduate		-0.07682 (0.554)	-0.08142 (0.519)	-0.0819 (0.514)	-0.04632 (0.277)	0.02977 (0.535)
Percent Some College or Associate		-0.07325 (0.390)	-0.07852 (0.319)	-0.07579 (0.280)	-0.03529 (0.438)	0.03434 (0.605)
Percent Bachelor's Degree		0.03066 (0.699)	0.02924 (0.707)	0.02505 (0.722)	-0.01028 (0.857)	-0.02025 (0.672)
Percent Graduate or Professional		-0.06451 (0.634)	-0.07101 (0.587)	-0.06893 (0.589)	0.02340 (0.745)	0.05956 (0.459)
Probability for Joint F-test for Census Age Groups	0.0035	0.0890	0.1033	0.1752	0.0017	0.0009
State Fixed Effects	No	No	No	No	No	Yes
Constant	0.1139 (0.520)	-0.005957 (0.966)	0.00313 (0.981)	0.00426 (0.973)	0.1353 (0.054)	0.07914 (0.407)
Number of Observations = 668	R2 = 0.1074	R2 = 0.1183	R2 = 0.1210	R2 = 0.1216	R2 = 0.3040	R2 = 0.4214

Table 12: Focusing on voter turnout in Georgia and Pennsylvania. Using the specifications shown in Table 11, though not all results are reported. (Florida, Georgia, North Carolina, Ohio, Pennsylvania) (Clustering by state, the level of significance for a two-tailed t-test are in parentheses)						
Endogenous variable: The change in voter turnout rate between 2020 and 2016						
Control variables	(1)	(2)	(3)	(4)	(5)	(6)
County where Fraud alleged	0.0312286 (0.041)	0.040204 (0.038)	0.043622 (0.036)	0.04341 (0.039)	0.022067 (0.081)	0.0006146 (0.871)