EXHIBIT 15

Comment on "A Simple Test for the Extent of Voter Fraud with Absentee Ballots in the 2020 Presidential Election"*

Andrew C. Eggers[a], Haritz Garro[b], and Justin Grimmer[c]

[a]Political Science. University of Chicago
[b]Democracy and Polarization Lab. Stanford University
[c]Democracy and Polarization Lab, Political Science, and Hoover Institution. Stanford University

January 8, 2021

Abstract

In a recent paper,[1] John Lott Jr. claims to find evidence of anti-Trump fraud in the 2020 U.S. presidential election. We show that Lott's analysis is fundamentally flawed and provides no evidence of unusual patterns (let alone fraud) in either voting or turnout. Lott's analysis of voting patterns in Fulton County, GA, and Allegheny County, PA, uses an unusual estimation strategy that suffers from a subtle but fundamental flaw: his conclusions about fraud in those two counties are entirely dependent on the arbitrary order in which pairs of precincts in *other* counties are entered in the dataset. Using a more appropriate specification, we find nothing unusual about voting patterns in these two counties. Lott (2020) also claims that turnout unusually increased in counties where Republicans have made accusations of fraud; we show that Lott (2020)'s test is seriously biased and that turnout rates in these counties are consistent with broader patterns in contested states. In short, Lott's (2020) analysis provides no evidence of anything distinctive or suspicious about voting or turnout in the 2020 election.

*We thank John Lott for sharing the precinct-level data on the same day we made the request.

[1]John Lott Jr., "A Simple Test for the Extent of Voter Fraud with Absentee Ballots in the 2020 Presidential Election". https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3756988.

1 Introduction

We reexamine the evidence for voter fraud presented in "A Simple Test for the Extent of Voter Fraud with Absentee Ballots in the 2020 Presidential Election" (hereafter Lott (2020)). Lott (2020) claims that a comparison of adjacent election precincts in Georgia and Pennsylvania supports the Trump campaign's allegations that the 2020 presidential election was "stolen" through fraud. In Lott (2020)'s abstract, he estimates that fraud in Fulton County contributed 11,350 votes to Biden (which would account for nearly all of Joe Biden's margin of victory in Georgia) and fraud in Allegheny County contributed about 55,270 votes to Biden's victory in Pennsylvania (which would account for around 2/3 of Biden's margin in Pennsylvania). Lott (2020) also claims to detect unusually large turnout increases in a set of counties where Republicans have made post-election accusations of malfeasance, interpreting this as evidence of fraud that could account for "up to 289,000 excess votes." If true, these claims would cast serious doubts on the integrity of the 2020 election. The paper has already received widespread attention.[2]

In this comment, we show that Lott's claims are entirely baseless. Our reanalysis of Lott (2020)'s data shows that Lott's claims about absentee voting in GA and PA depend on an entirely arbitrary decision about how counties are entered in the dataset: the conclusion is reversed when an alternative and equally justified data entry rule is used. When we replace Lott's unusual specification with a more standard approach that does not depend on arbitrary coding rules, we find absolutely no evidence for fraud in either Fulton County or Allegheny County. We also find that, once simple allowances are made for differences in turnout trends across states, there is nothing unusual about the turnout rate in the counties that Republicans have targeted with fraud claims.

In short, even if we accept Lott's premise that small differences in Trump's share of the absentee vote between adjacent precincts or small amounts of unexplained turnout in a set of counties targeted for post-election appeals constitute evidence of fraud (which we do not), we find that Lott (2020)'s analysis provides no evidence of fraud whatsoever in the 2020 presidential election.

2 Lott's (2020) Precinct-Level Analysis Depends Entirely on an Arbitrary Coding Rule

Lott (2020) first seeks to estimate the effect of the absentee ballot counting procedure in two counties where fraud has been alleged by Trump and other Republicans: Fulton County, GA, and Allegheny County, PA. Lott (2020)'s approach assumes that Trump's share of the

[2] Peter Navarro, the outgoing Assistant to the President and Director of the Office of Trade and Manufacturing Policy, promoted the paper in a tweet on December 29 (https://twitter.com/RealPNavarro/status/1343979253659004928). The next day, Donald Trump also tweeted about the study (https://twitter.com/realDonaldTrump/status/1344173684983017473).

absentee vote in a precinct is related to Trump's share of the *in-person* vote in the precinct and voter demographics. Lott (2020) recognizes, however, that a difference in Trump's share of the absentee vote across neighboring counties, even controlling for Trump's share of the in-person vote and demographics, is not necessarily evidence of fraud. There may be other factors that vary across counties that could produce such differences.

To eliminate some of these alternative explanations for differences in Trump's absentee support between "suspect" counties and neighboring counties, Lott (2020) focuses on precincts that lie along county borders. Specifically, he forms pairs of precincts that lie along a boundary separating a suspect county (i.e. one where Republicans have alleged that fraud took place) and an adjacent county where Trump won a majority of the vote and no fraud allegations have been made.[3] Lott (2020) also forms pairs of precincts that lie along the boundary between two of these Republican counties, which serve as a kind of control group for the other pairs. Lott (2020) then conducts his analysis using within-pair *differences* in each variable: he regresses the difference in Trump's share of the absentee vote between the two precincts on the difference in Trump's share of the in-person vote between the two precincts and an indicator for whether the pair contains a precinct in a suspect county.[4] That is, his basic regression equation is

$$\left(\text{Absentee}_i - \text{Absentee}_j\right) = \beta_0 + \beta_1\left(\text{InPerson}_i - \text{InPerson}_j\right) + \delta\text{SuspectCounty}_i + u_{ij},$$

where Absentee_i is Trump's share of the absentee vote in precinct i, InPerson_i is Trump's share of the in-person vote in precinct i, SuspectCounty_i indicates whether precinct i is located in a "suspect" county, and i and j are adjacent precincts that Lott assigns to a pair. Thus β_0 measures the within-pair difference in Trump's share of the absentee vote among pairs that don't involve a suspect county (adjusting for the within-pair difference in Trump's in-person share), and the key coefficient is δ, which compares the adjusted difference in Trump's share of the absentee vote within pairs involving the suspect county against the corresponding adjusted difference within pairs not involving the suspect county. The underlying logic seems to be that fraud is the likely explanation if there is a bigger drop in Trump's share of the absentee vote when we cross from, for example, Coweta County to Fulton County than when we cross from Coweta County to Carroll County, two Republican counties where no fraud has been alleged.

Even if we stipulate that focusing on adjacent precincts eliminates all between-county differences in true absentee support for Trump (conditional on Trump's in-person support),[5] Lott (2020)'s design suffers from a fatal flaw. As noted, Lott (2020)'s design measures a

[3] Lott (2020) provides no justification for not comparing Fulton and Allegheny counties (or others where fraud was alleged) with surrounding counties carried by Biden. By ruling out these comparisons, Lott severely restricts his sample size and likely excludes the most similar comparisons.

[4] In some specifications he also includes differences in various race-and-gender groups between the two precincts.

[5] This is doubtful. For example, Trump won just 9.6% of the in-person vote in a precinct in Fulton County (FA01B) that is adjacent to a precinct in Coweta County where Trump won 78% of the in-person vote (Fischer Road). It seems unlikely that precincts that differ so markedly in voting outcomes would be similar in e.g. voters' propensity to vote in person vs. absentee conditional on their vote choice.

difference between two differences: is the drop in Trump's share of the absentee vote larger when we cross the Fulton County border into Republican counties than when we cross the border of one Republican county into another Republican county? The problem arises in measuring the second drop: there is no clear rule for determining the order of the difference. For example, should we record the change in Trump's absentee vote share as we move from Carroll to Coweta, or as we move from Coweta to Carroll? Neither county is "suspect", so either approach could be justified. Lott (2020, footnote 13) chooses one rule (subtracting east from west and north from south) but the opposite rule or indeed any rule would be equally justified. This arbitrariness is a symptom of the underlying lack of compelling logic behind this aspect of the design: there is no clear reason to benchmark the difference in voting patterns across the key county boundary against the corresponding difference across another boundary.[6]

As it turns out, Lott (2020)'s evidence for fraud in Fulton County, GA, and Allegheny County, PA, relies entirely on this arbitrary coding rule: if a different but equally valid rule is used we reach the opposite conclusion from Lott (2020). Figure 1 illustrates the point for Fulton County. In both panels, each red dot corresponds to a pair of precincts lying on opposite sides of the Fulton County boundary; each blue dot corresponds to a pair of precincts lying on opposite sides of the boundary between two nearby Republican counties. The vertical axis shows the difference in Trump's share of the absentee vote within the precinct pair; the horizontal axis shows the difference in Trump's share of the in-person vote within the precinct pair.

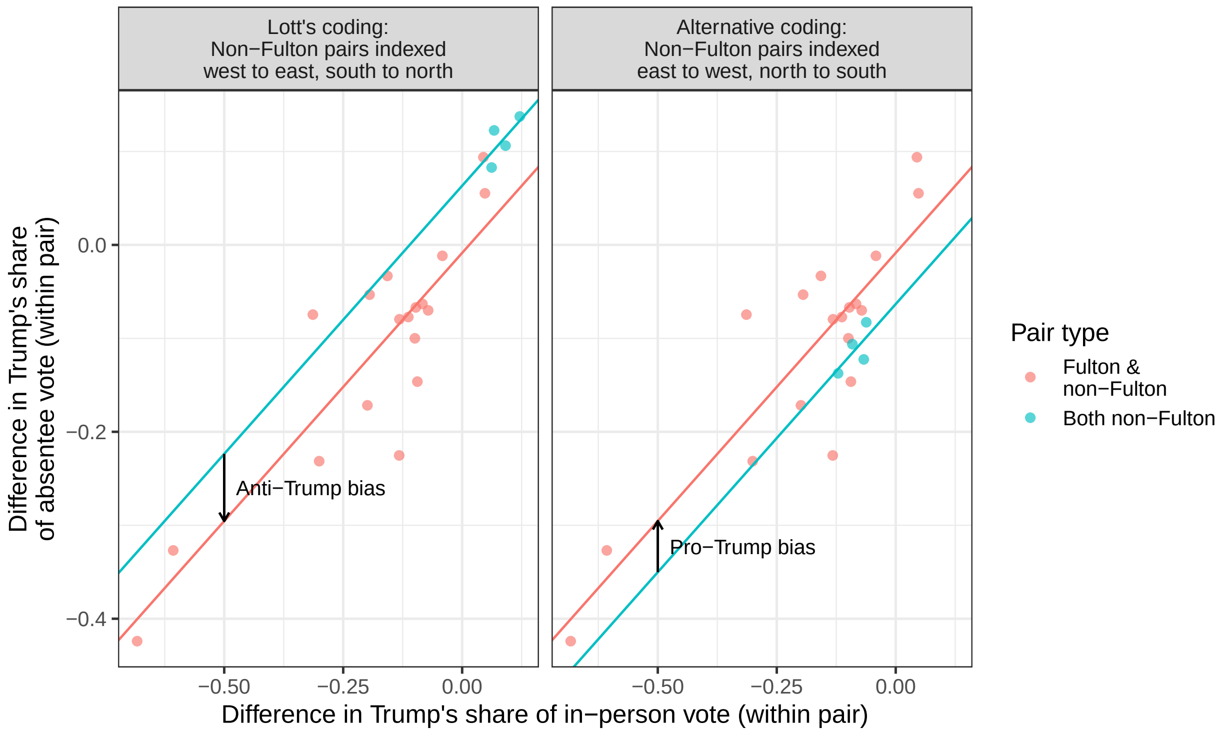
The left panel of Figure 1 shows the analysis using Lott (2020)'s coding: for pairs including a Fulton County precinct, the Trump share for the non-Fulton County precinct is subtracted from the Trump share for the Fulton County precinct; for pairs not including a Fulton County precinct, Lott (2020) uses the arbitrary rule noted above. This coding results in what Lott interprets as evidence for anti-Trump bias in Fulton County. Conditional on the difference in Trump's in-person vote share within a precinct pair, the difference in Trump's absentee vote share is lower in precinct pairs involving Fulton County than in other precinct pairs.

In the right panel of Figure 1 we show that the conclusion is reversed when we reverse Lott's arbitrary coding rule: instead of subtracting east from west and north from south in computing differences for non-Fulton precinct pairs, we subtract west from east and south from north. The scatterplot looks identical to the left panel except that the four blue dots (representing non-Fulton precinct pairs) are reflected through the origin. This small change reverses the conclusion, however: by Lott (2020)'s logic we now have evidence of pro-Trump bias in Fulton County.

Table 3 (Appendix) reports coefficient estimates and standard errors for both sets of

[6] One could imagine a better design that compared the *magnitude* (i.e. absolute value) of differences across suspect boundaries and other boundaries. In this case the ordering of precinct pairs would not matter. This is not Lott's design.

Figure 1: Evidence for fraud in Fulton County, GA, is reversed if arbitrary coding rule is reversed



analysis depicted in Figure 1. The evidence of pro-Trump fraud with the alternative coding rule has a similar absolute t-statistic (t = 1.67) as Lott's evidence of anti-Trump fraud with the original coding rule (t = 1.89).

The Pennsylvania results also depend on Lott's arbitrary coding rule, as we show in the same manner in Figure 2 and Table 4 (Appendix). Lott (2020) concludes from his analysis that anti-Trump fraud took place in Allegheny County, but if we apply a different but equally valid coding rule we find (by the same logic) stronger evidence for *pro-Trump* fraud in Allegheny County: the positive coefficient we obtain with the alternative coding rule is both larger in magnitude and more significant than the negative coefficient Lott reports.

We can further highlight the dependence of Lott's results on arbitrary coding decisions by exploring the universe of possible fraud estimates that Lott could have reported with equally justified alternative coding rules. In Figure 3 we show that, among the possible rules that could be used, any alternative rule would have produced weaker apparent evidence for anti-Trump fraud in Fulton County and almost any rule would have produced weaker evidence for anti-Trump fraud in Allegheny County.[7] In the Fulton County analysis, there are four non-Fulton precinct pairs and thus $2^4 = 16$ possible rules for computing differences within non-Fulton pairs. The left panel of Figure 3 shows the histogram of the key coefficient across

[7]In personal communication Lott said the ordering of precincts followed a rule in a prior AER paper. We believe that is Bronars and Lott (1998).

Figure 2: Evidence for fraud in Allegheny County, PA, is reversed if arbitrary coding rule is reversed

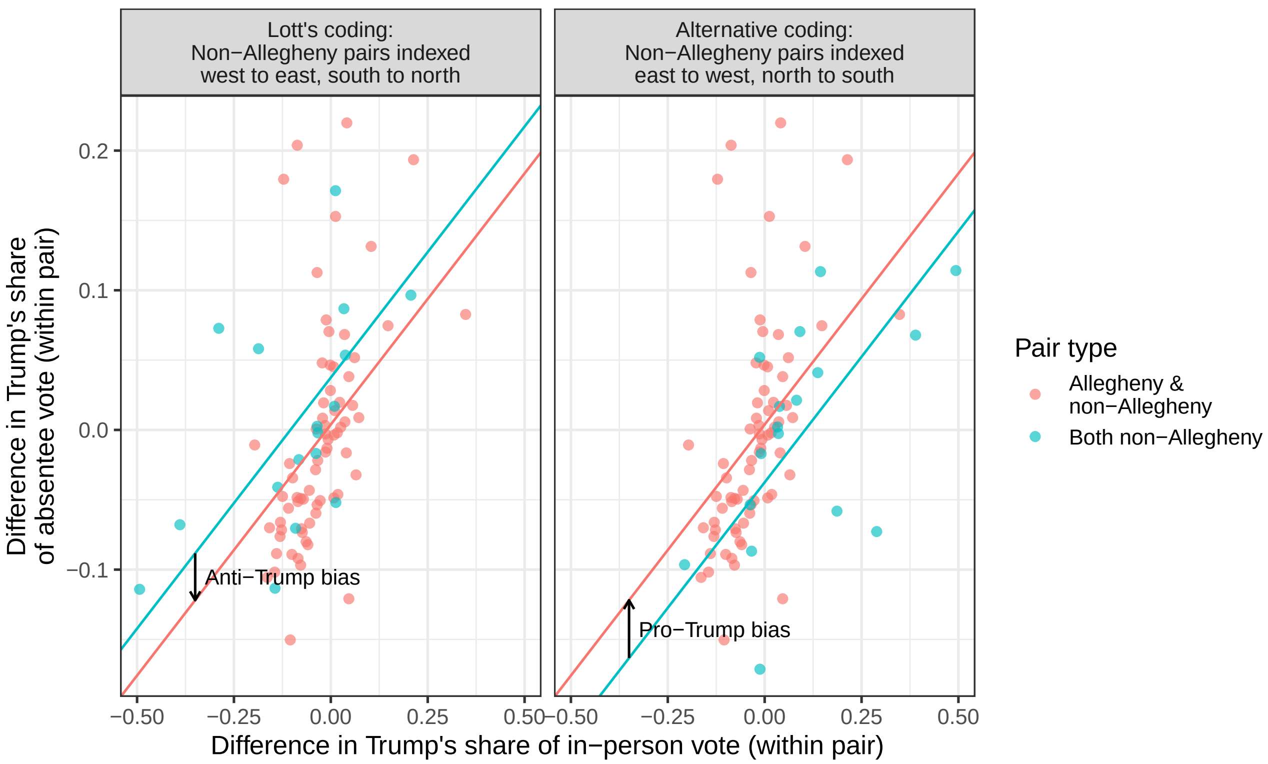
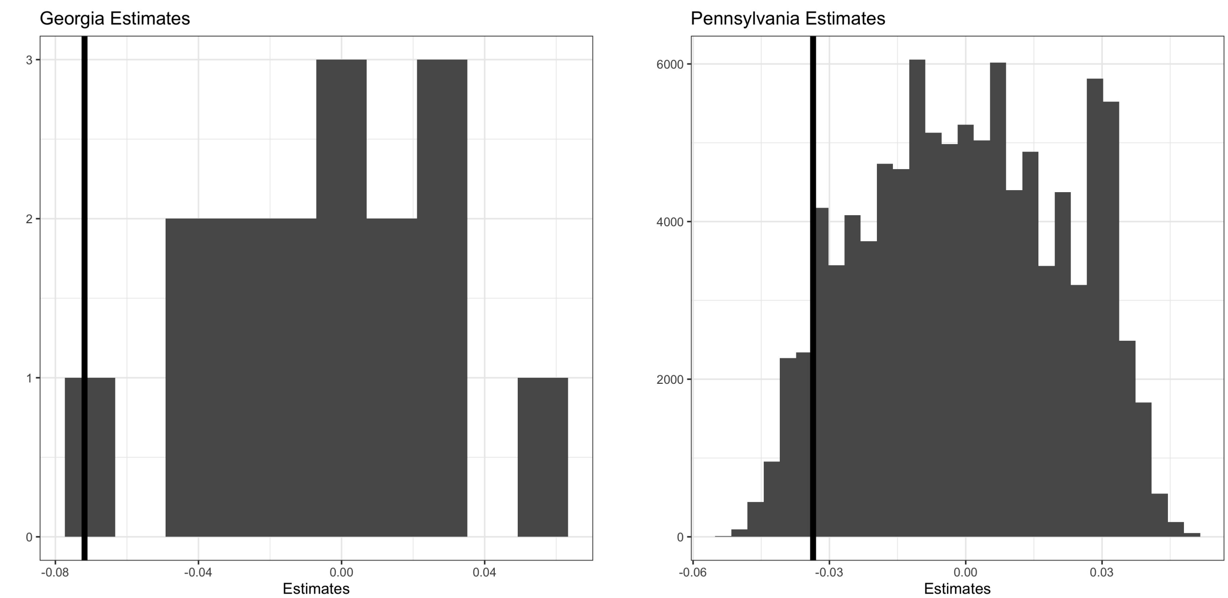


Figure 3: Evidence for fraud in Georgia and Pennsylvania depends on arbitrary coding rules; Lott's estimates are outliers in the distribution of estimates



these sixteen possible rules, with a vertical line highlighting the estimate for the rule Lott used. Among the sixteen possible rules, Lott's rule produces the strongest apparent evidence of anti-Trump fraud; six possible rules produce apparent evidence of pro-Trump fraud. In the Pennsylvania analysis we have seventeen non-implicated precinct pairs, allowing for over 130,000 possible coding rules. The right panel of Figure 3 shows the distribution of estimates for a random sample (with replacement) of 100,000 of these rules,[8] with the actual estimate again shown with a vertical line. The distribution is centered around zero, with roughly as many rules producing apparent evidence of pro-Trump and anti-Trump fraud; Lott's rule again happens to produce among the strongest apparent evidence of anti-Trump fraud.

Although the issue we highlight was not obvious to us on first reading Lott's study, it is an example of a known problem that crops up in research studying pairs of observations, or "dyads." When there is a clear distinction between members of dyads, such as aggressor/victim or source/destination, it can be sensible to address unobserved differences across dyads by studying within-dyad differences as Lott does. When no such distinction exists for some or all dyads (as in Lott's case), it becomes arbitrary how to define within-dyad differences. In such cases, "there is no consistent, non-arbitrary way to order the two members" of a dyad (Olsen and Kenny, 2006) and, as pointed out in Wheeler, Updegraff and Umaña-Taylor (2018), dyads whose members cannot logically be classified in a meaningful way "cannot be easily analyzed with the difference approach", i.e. the approach that Lott

[8] To explore the space of changes to the difference order, we first sample the number of difference orders to change from a Uniform(1, 16). Once this number is obtained, we then randomly sample the specific units that will have the difference order changed. This explores the space, but does not provide a sampling distribution that gives an equal probability to each rearrangement, because our sampling method is biased towards either too few or too many rearrangements.

(2020) uses.[9]

3 A More Standard Estimation Strategy Produces No Evidence of Fraud in Absentee Voting

Although Lott's specification problematically depends on arbitrary coding decisions, Lott's basic strategy of examining differences in voting patterns across a county boundary has some merit. Such differences in voting patterns could of course be explained by differences in voter behavior rather than fraud (particularly because county boundaries determine school districts and other policy outcomes, and some precincts along county boundaries are rather large geographically), but focusing on precincts along the county border does seem likely to reduce the role of these differences.[10]

To more effectively achieve Lott's objective of comparing voting patterns across county boundaries, we reanalyze Lott's data using a more standard specification that does not suffer from the problems highlighted in the previous section. Rather than using within-pair differences as Lott does, we employ a simple fixed effects model. The regression equation can be written as

$$\text{Absentee}_i = \beta_1 \text{InPerson}_i + \delta \text{SuspectCounty}_i + \sum_{k=1}^{K} \alpha_k I(\text{pair}_i = k) + \epsilon_i \qquad (1)$$

where Absentee_i and InPerson_i denote Trump's share of the absentee and in-person vote (respectively) in precinct i, SuspectCounty_i indicates whether precinct i is located in a "suspect" county (Fulton or Allegheny, depending on the state being analyzed), and each precinct is identified with one of K precinct pairs indexed by k, with α_k indicating the fixed effect for pair k. The regression thus asks whether Fulton or Allegheny county precincts have lower absentee support for Trump than would be expected controlling for their in-person support for Trump and any factors (observable or unobservable) that are common to paired precincts. Precinct pairs that do not involve a suspect county contribute to estimating the coefficient β_1 but do not otherwise contribute to the estimation of the key coefficient δ. Crucially, no arbitrary coding decisions are necessary.

We report the results of these analyses for Georgia in Table 1 below. In column 1 we simply regress Trump's share of the absentee vote on Trump's share of the in-person vote and a dummy for Fulton County; in column 2 we add precinct-pair fixed effects as in equation 1, essentially allowing the intercept to vary across Lott's precinct pairs; in column 3 we instead use county-pair fixed effects, with one intercept for Fulton-Coweta pairs, another for Carroll-Coweta pairs, etc. None of these specifications shows a substantively or statistically

[9] See also Chapter 2 in Kenny, Kashy and Cook (2006) for a rigorous overview of the problems with unordered or indistinguishable pairs in dyadic data.

[10] Even if we could find a difference in voting patterns between county A and county B that is so suspicious as to suggest fraud, we may not know which county conducted the fraud.

significant difference between Trump's share of the absentee vote in Fulton County precincts and other precincts.

Table 1: A Fixed Effects Specification Shows Nothing Suspicious in Fulton County, GA

	Dependent variable:		
	Trump Share Absentee		
	(1)	(2)	(3)
Trump Share, In-Person	0.760	0.606	0.654
	(0.049)	(0.077)	(0.056)
Fulton County	0.019	−0.003	0.006
	(0.019)	(0.020)	(0.018)
Observations	44	44	44
Precinct-Pair Fixed Effects		✓	
County-Pair Fixed Effects			✓

Table 2 shows the same analysis for Pennsylvania in the same manner. Again, none of the specifications shows a substantively or statistically significant difference between Trump's share of the absentee vote in Allegheny County precincts and other precincts.

In short, when we reanalyze Lott (2020)'s data with a more sensible fixed effects specification, we find no evidence of differences in voting patterns between precincts in Fulton County or Allegheny County and adjacent precincts in Republican-leaning counties. If such differences existed they would hardly be convincing evidence of fraud, given possible differences between precincts located in different counties that are served by different school systems. But we find no such differences, undermining the basis for Lott (2020)'s claims.[11]

4 No Evidence of Distinctively High Turnout in "Suspicious" Counties

Lott (2020) provides a second analysis that he claims demonstrates evidence for voter fraud. First, Lott argues that fraud can increase turnout through a variety of mechanisms. He then claims to show that turnout rates increased more in 2020 in a set of counties where Republicans have alleged fraud. Lott argues that there was an "unexplained increase in voter

[11] In the Appendix we also replicate and extend Lott's analysis of provisional ballots in Pennsylvania. As with his analysis of absentee voting, his conclusions about provisional ballots depend on the arbitrary coding of non-Allegheny precinct pairs (Figures 8 and 9) and fixed effects estimation shows no difference in Biden's share of the provisional vote in Allegheny precincts and other precincts (Tables 5 and 6).

Table 2: A Fixed Effects Specification Shows Nothing Suspicious in Allegheny County, PA

	Dependent variable:		
	Trump Share, Absentee		
	(1)	(2)	(3)
Trump Share, In-Person	0.511	0.307	0.442
	(0.042)	(0.066)	(0.048)
Allegheny County	0.003	0.003	0.006
	(0.008)	(0.009)	(0.009)
Observations	174	174	174
Precinct-Pair Fixed Effects		✓	
County-Pair Fixed Effects			✓

turnout" in the key counties of between 1.26 and 2.42 percent, which Lott says is equivalent to 150,000 to 289,000 votes in those states. Lott concludes that this is evidence consistent with fraud.

While the first half of Lott's (2020) paper focuses on narrow comparisons across county boundaries, this section engages in analysis that spans hundreds of counties across nine states. Specifically, Lott checks whether turnout in the 2020 election was higher than would be expected (given previous turnout, political leaning, and local demographics) in counties where, according to Republican lawsuits, fraud may have taken place. Lott identifies 19 counties across six swing states where Republicans have alleged that fraud took place.[12] Lott (2020) compares turnout in these counties to turnout in other counties in the same six states plus counties in three other swing states (Florida, Ohio, and North Carolina). He argues that, if turnout is higher in these counties than would be expected given covariates, it would be evidence of fraud.

Before digging deeper into Lott (2020)'s turnout analysis, we emphasize that we question the premise of Lott (2020)'s analysis; that is, we do not believe that even a robust finding of slightly higher than expected turnout in a set of counties Republicans targeted in post-election lawsuits would constitute convincing evidence of electoral fraud. The differences Lott claims to have found are small (1-2 percentage points), and in the absence of fraud turnout is not perfectly explained by the covariates Lott (2020) uses: a particularly energetic local mobilization campaign (on either side) or an especially effective down-ballot candidate could affect turnout by these amounts. Perhaps more to the point, Lott (2020)

[12]Lott identifies the following "suspicious" counties—Georgia: Fulton, Dekalb; Pennsylvania: Allegheny, Centre, Chester, Delaware, Montgomery, Northampton, Philadelphia; Arizona: Apache, Coconino, Maricopa, Navajo; Michigan: Wayne; Nevada: Clark, Washoe; Wisconsin: Dane.

looks for unexplained turnout in places Republicans chose to target in post-election lawsuits. We do not know how Republicans chose which counties to target, but it seems plausible that they targeted counties based on district characteristics that are related to turnout (but not modeled by Lott (2020)) or even based on observed results (including turnout). This creates a thorny selection: was fraud the cause of high turnout, or was high turnout the cause of allegations of fraud? Highly anomalous turnout figures could provide evidence of a problem, but a percentage point or two of unexplained turnout has other more plausible explanations and could not on its own establish fraud.

Nevertheless, given the possible implications of such a serious claim, we investigate the issue to see if Lott (2020) has shown a genuinely unexplained anomaly in the counties where Republicans have alleged that fraud took place. We assembled an original dataset that would allow us to assess Lott (2020)'s claims beyond his chosen set of states, if necessary. We use turnout rates for the county citizen voting-age population. For total votes, we use Dave Leip's county-level vote results for 2020 and 2016. For the number of voting-aged citizens we use the five-year ACS from 2019 and 2015.[13]

If we visually examine how turnout in 2020 compared to turnout in 2016 for counties in the six states where Lott alleged fraud, we find that there is nothing remarkable about the turnout rate in the suspicious counties. In Figure 4 we plot turnout in 2020 against turnout in 2016 for counties in the six states with counties that Lott codes as having alleged fraud; we do this separately by state, with counties where fraud was alleged colored red and a linear regression line superimposed.[14] On a simple visual inspection, there is nothing puzzling about 2020 turnout in the highlighted counties. In fact, turnout seems to have been lower on average in these counties than in other counties in the same state, conditional on prior turnout. In light of this observation, Lott (2020)'s finding is puzzling: why would he conclude that turnout is suspiciously high in these counties, given the information in this figure?

The answer is that Lott's conclusions are driven by the inclusion of states that have lower turnout increases and no "suspicious" counties—namely Florida, North Carolina, and Ohio. Figure 5 shows that, conditional on turnout in 2016, turnout in Florida, Ohio, and North Carolina was lower than turnout in the six states that contain a suspicious county in Lott's analysis. This is relevant because Lott (2020)'s analysis compares changes in turnout in suspicious counties with changes in turnout in all other counties, so these smaller increases in turnout rates across states will be conflated with the suspicious county indicator in his analysis. The smaller the turnout increase in these three "non-suspect" states, the more turnout in the suspect counties will appear to be suspiciously high, even if the changes in turnout in these suspect counties are unremarkable relative to the changes in turnout in other counties in their own state.

[13]This follows best practice from Michael McDonald http://www.electproject.org/home/voter-turnout/faq/congress. We provide our county-average turnout rates by state in the appendix. We note that our estimates of turnout are lower than Lott (2020)'s average turnout rates, but closer to official statistics.

[14]The regression line is drawn based on the non-suspect counties.

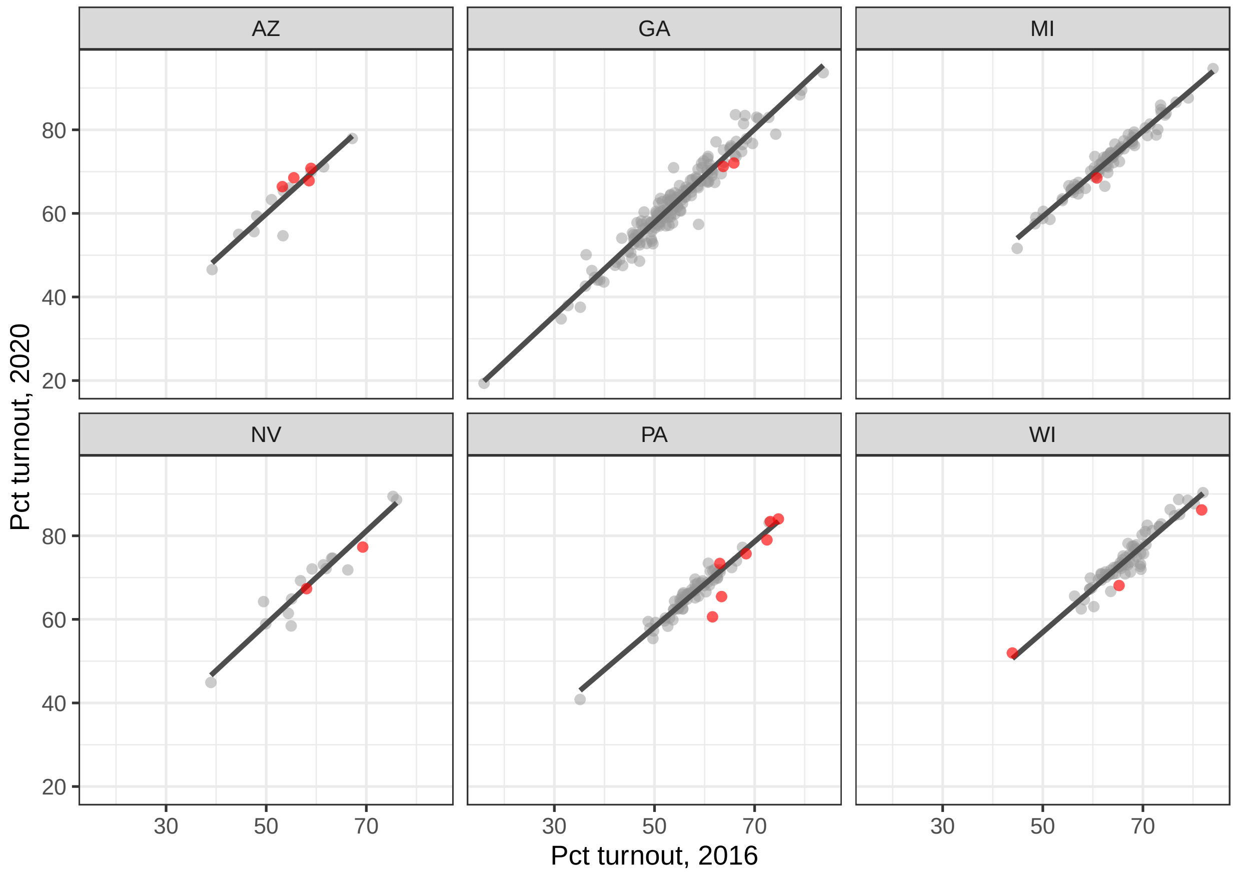


Figure 4: "Suspicious" counties (in red) are not remarkable relative to other counties in their state

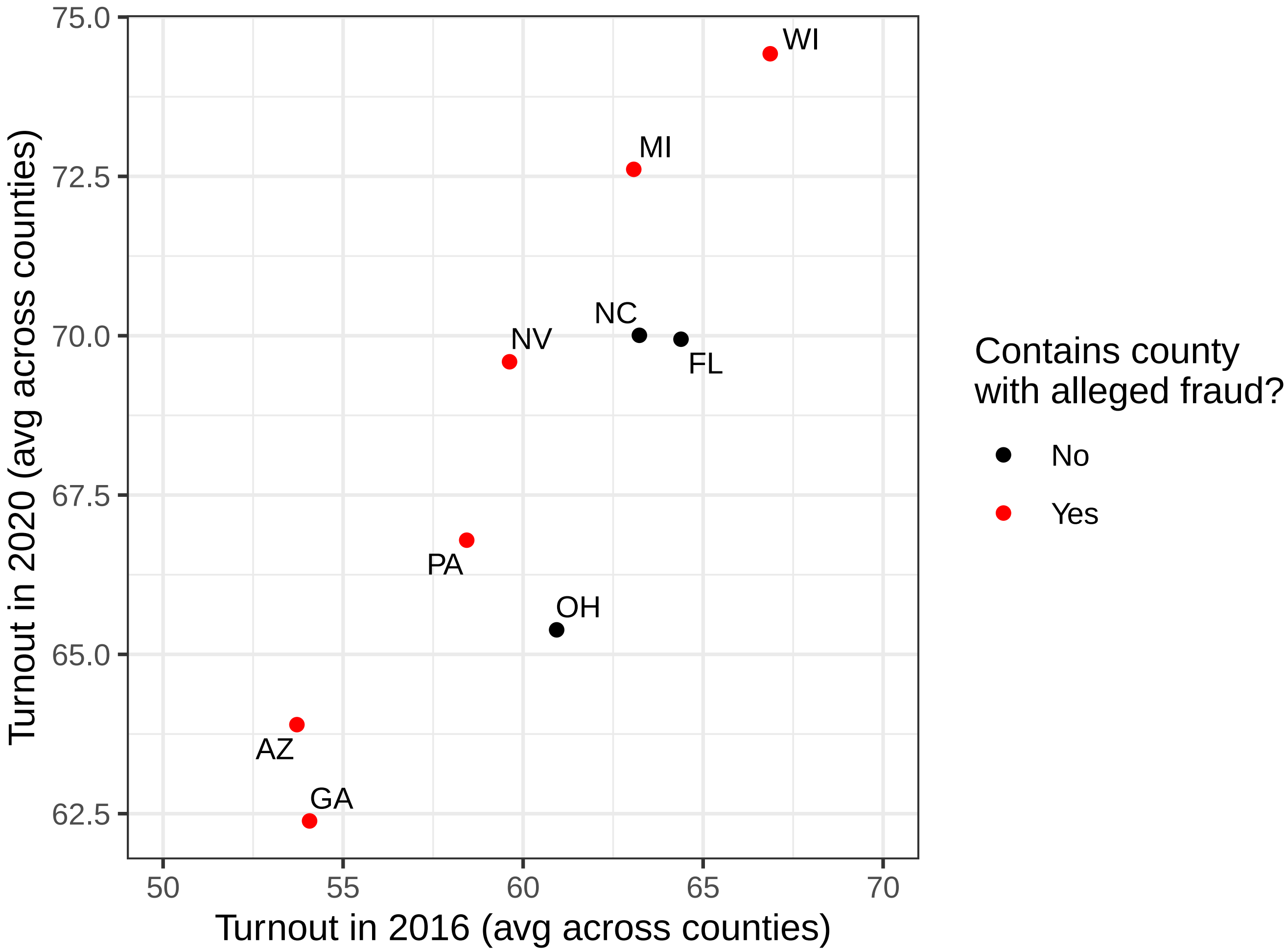


Figure 5: Swing states without suspicious counties had smaller average turnout increases, which drives Lott's (2020) results

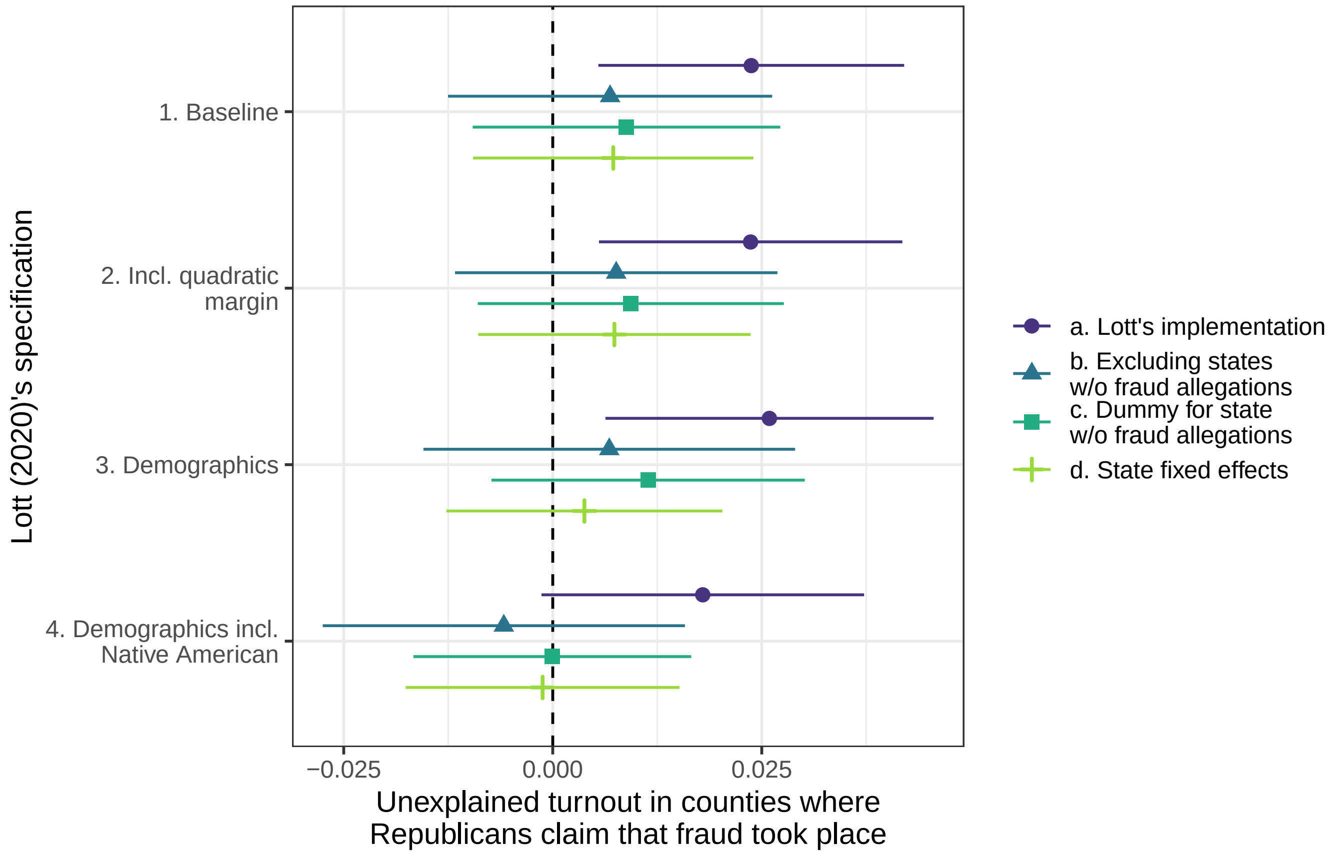


Figure 6: Lott's (2020) estimates of suspicious county differences in turnout are zero and null once we address state-level differences.

Figure 6 shows that once we address the level differences across states, Lott's (2020) estimates of the turnout differences in suspicious counties go to zero and become null. We examine all four of Lott's (2020) models (organized on the vertical axis) and present the estimate of the average difference in turnout rates for suspicious counties. The circle/purple estimates of suspicious county turnout depict the estimates using the four specifications for which Lott (2020) presents results in his Table 10. The triangle/dark-green estimates depict our estimates when we exclude Florida, Ohio, and North Carolina - three states in which no fraud was alleged. Across models, the difference in suspicious counties is close to zero and—in the case of model 4—the estimate is negative. The square/light-green estimates are from a model where we merely include an indicator for a state that has suspicious counties. Again, this reduces the estimate to null. Finally, the last plus/lime-green estimate includes state-level fixed effects. Across models, this gives a close to zero and null difference for suspicious counties. Thus, simply by focusing only on states where at least one county had alleged fraud (i.e. swing states that Biden won) or allowing that state-wide turnout trends may differ across states or groups of states, we are able to explain what Lott (2020) claimed was unexplained turnout in counties where Republicans had claimed fraud.

To highlight the deficiency of Lott's approach, we undertake a falsification test. To reiterate, the fundamental problem with Lott's analysis is that it compares "suspect" counties in states that experienced large turnout increases against a pooled control group comprising of non-suspect counties in states that experienced large turnout increases and counties in states that experienced smaller turnout increases. Given this flaw, we should find similar evidence of fraud if we replace Lott's coding of "suspect" counties with a random set of counties in the same states. To investigate this, we repeatedly draw a random set of counties from the states where Republicans alleged fraud, designate these counties (counterfactually) as "suspect", and conduct the same analyses reported in Figure 6.[15] If Lott (2020)'s design is valid, the coefficient on "suspect county" should be significant in about 5% of random draws. We expected otherwise: by including states with lower turnout increases in the control group (without including state fixed effects or otherwise accounting for cross-state turnout differences), Lott (2020)'s analysis builds in a bias toward finding "inexplicably" high turnout increases in counties where Republicans have alleged fraud.

Figure 7 shows the distribution of t-statistics across 1000 random reshufflings. The top row shows Lott (2020)'s specifications: the estimate from the true coding of suspect counties is statistically significant in each specification (as shown by the vertical red line at or above 2), but this t-statistic is actually typical of the distribution of t-statistics across random reshufflings (shown in the histogram). Across Lott (2020)'s specifications, the proportion of random reshufflings that produce a significant "effect" (the false discovery rate, or type I error, shown by the dark region of the histograms) is between .6 and .75. In fact, the t-statistic is larger on average when we randomly select counties than when we use the counties in which Republicans actually alleged fraud (according to Lott (2020))).

The next three rows of Figure 7 show the same exercise conducted for the alternative specifications we used in Figure 6 above. False discovery rates are near .05, suggesting that adjusting for differences in turnout across states renders Lott (2020)'s tests statistically valid.

5 Conclusion

Using precinct-level analysis of absentee voting in Georgia and Pennsylvania as well as county-level analysis of turnout, Lott (2020) claims to provide statistical evidence for voter fraud sufficient to explain Trump's defeat. After scrutinizing Lott (2020)'s analysis we conclude that this claim is false. Lott (2020)'s precinct-level findings in Georgia and Pennsylvania are reversed if we alter an entirely arbitrary coding rule, and we find no evidence of differences in voting behavior across county boundaries in those states using a more standard and appropriate estimation technique. Lott (2020)'s analysis of voter turnout collapses when we make simple adjustments to his specifications: what he claims is inexplicably high voter turnout is easily explained by differences in turnout trends across states. Thus, even if we accepted the questionable premise that minor differences in voting behavior or slightly elevated turnout rates constituted convincing evidence of fraud (we do not), we find that

[15] In a state where n counties had allegations of fraud, we randomly draw n counties to be the pseudo-suspect counties.

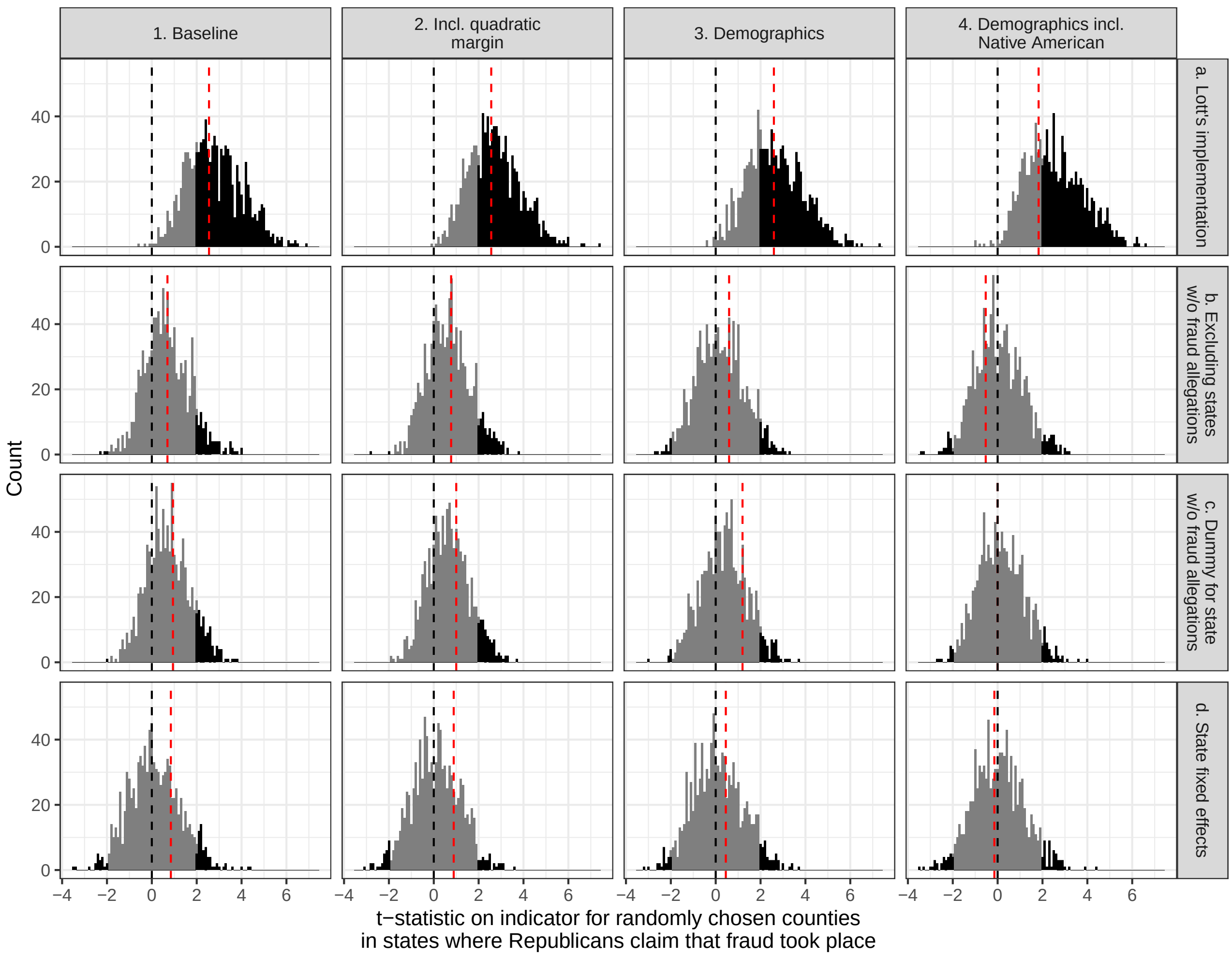


Figure 7: If "suspicious" counties were chosen at random rather than identified from Republican allegations as in Lott (2020), Lott (2020)'s test would usually find evidence of "fraud"; our improved specifications would not

Lott (2020)'s analysis provides no such evidence.

Like other claims of fraud following the 2020 election, Lott (2020)'s assertions have the potential to undermine belief in the integrity of American elections. Unlike most of these other claims, Lott's analysis has the appearance of careful social scientific research and cannot easily be dismissed as obviously illogical or mere hearsay. Indeed, it is because Lott (2020) shares several characteristics with rigorous social scientific research that we considered it especially important to investigate these claims more deeply.

Observers concerned about the integrity of the 2020 election can be reassured that Lott (2020)'s claims of election fraud have no basis in fact. We hope that our analysis helps undo some of the damage that has already been done by these and other unfounded claims of election fraud.

References

[] **Kenny, David A., Deborah A. Kashy and William L. Cook**, "Dyadic Data Analysis," *Guilford Press*, 2006, New York, NY.

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

[] **Olsen, Joseph A. and David A. Kenny**, "Structural Equation Modeling with Interchangeable Dyads," *Psychological Methods*, 2006, Vol. 11, 2, pp. 127-141.

[] **Wheeler, Lorey A., Kimberly A. Updegraff and Adriana J. Umaña-Taylor**, "A Dyadic Data Analytic Primer: An Illustration with Mexican-origin Couples," *Journal of Latina/o Psychology*, 2018, Vol. 6, 4, pp. 276-290.

Appendix

Table 3: Lott's Conclusions Are Reversed if the Arbitrary Ordering of Precinct Differences is Reversed (Georgia)

	Dependent variable:	
	Difference, Trump Absentee (Lott (2020), Table 2)	
	(1)	(2)
Difference, Trump In-Person Vote	0.574	0.574
	(0.073)	(0.073)
Fulton County	−0.072	0.055
	(0.038)	(0.033)
Observations	22	22
Reverse Coding		✓

Table 4: Lott's Conclusions Are Reversed if the Arbitrary Ordering of Precinct Differences is Reversed (Pennsylvania)

	Dependent variable:	
	Difference, Trump Absentee (Lott (2020), Table 5)	
	(1)	(2)
Difference, Trump In-Person Vote	0.359	0.359
	(0.069)	(0.069)
Allegheny County	−0.034	0.041
	(0.019)	(0.020)
Observations	87	87
Reverse Coding		✓

Table 5: Pennsylvania Provisional Ballot Results

	Dependent variable:			
	Difference, Trump Provisional (Lott (2020), Table 6)	Trump Provisional Vote		
	(1)	(2)	(3)	(4)
Difference, Trump In-Person Vote	1.038			
	(0.558)			
Trump, In-Person Vote		0.729	1.055	0.690
		(0.222)	(0.552)	(0.257)
Allegheny County	−0.125	−0.004	−0.036	−0.047
	(0.141)	(0.036)	(0.044)	(0.048)
Observations	34	120	120	120
Precinct-Pair Fixed Effects			✓	
County-Pair Fixed Effects				✓

Table 6: Pennsylvania Provisional Ballot Results, Total Ballots

	Dependent variable:			
	Difference, Biden Share of Votes From Provisional Ballots (Lott (2020), Table 7a)	Biden Share of Votes From Provisional Ballots		
	(1)	(2)	(3)	(4)
Difference, Share of Trump Vote from Provisional Ballots	0.364			
	(0.105)			
Share of Trump Vote from Provisional Ballots		0.371	0.385	0.342
		(0.078)	(0.103)	(0.082)
Allegheny County	0.010	0.007	0.007	0.007
	(0.004)	(0.002)	(0.002)	(0.002)
Observations	87	174	174	174
Precinct-Pair Fixed Effects			✓	
County-Pair Fixed Effects				✓

Figure 8: Distribution of Estimates for Alternative Precinct Differencing Orders, Pennsylvania Provisional Ballots

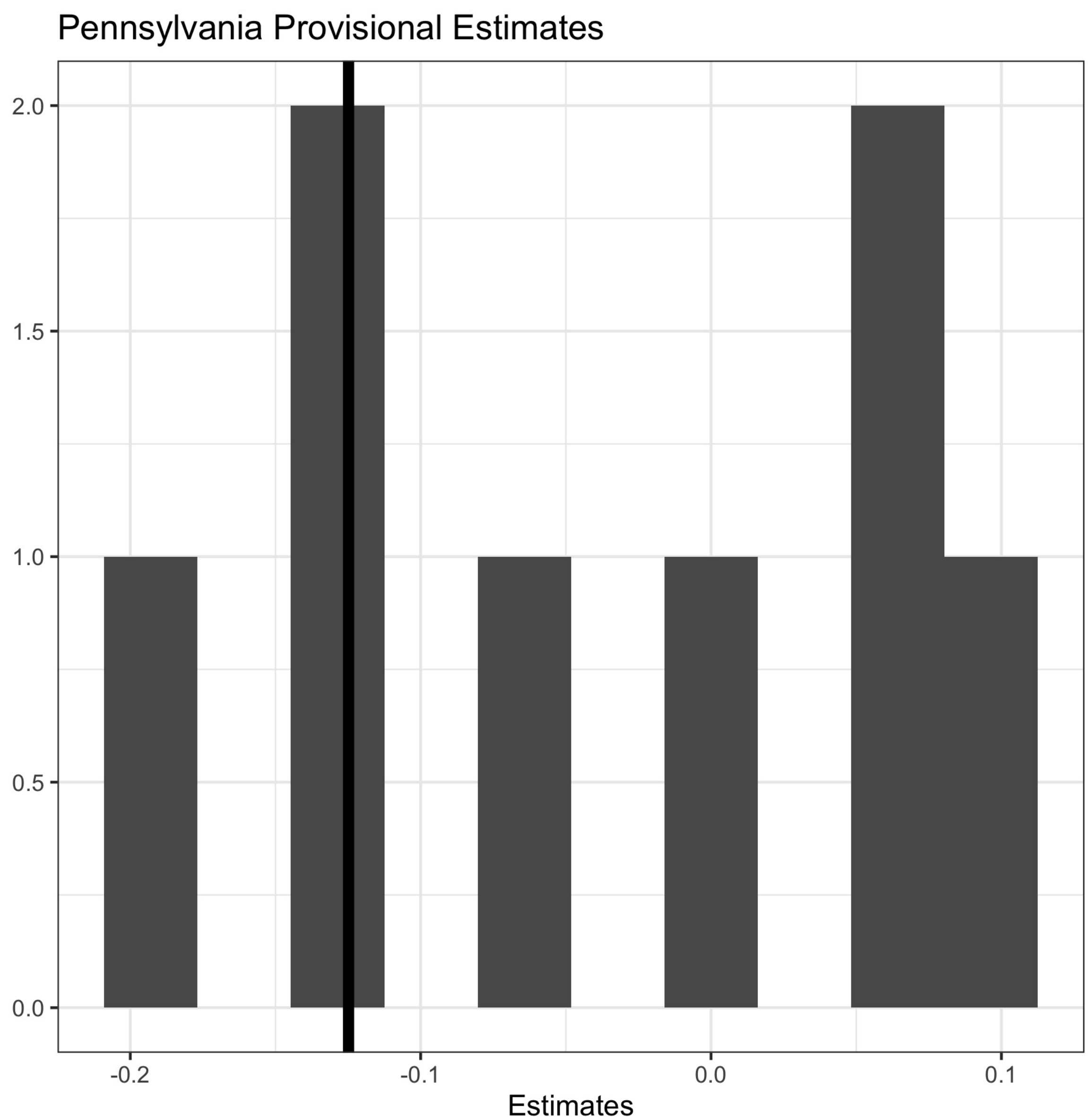
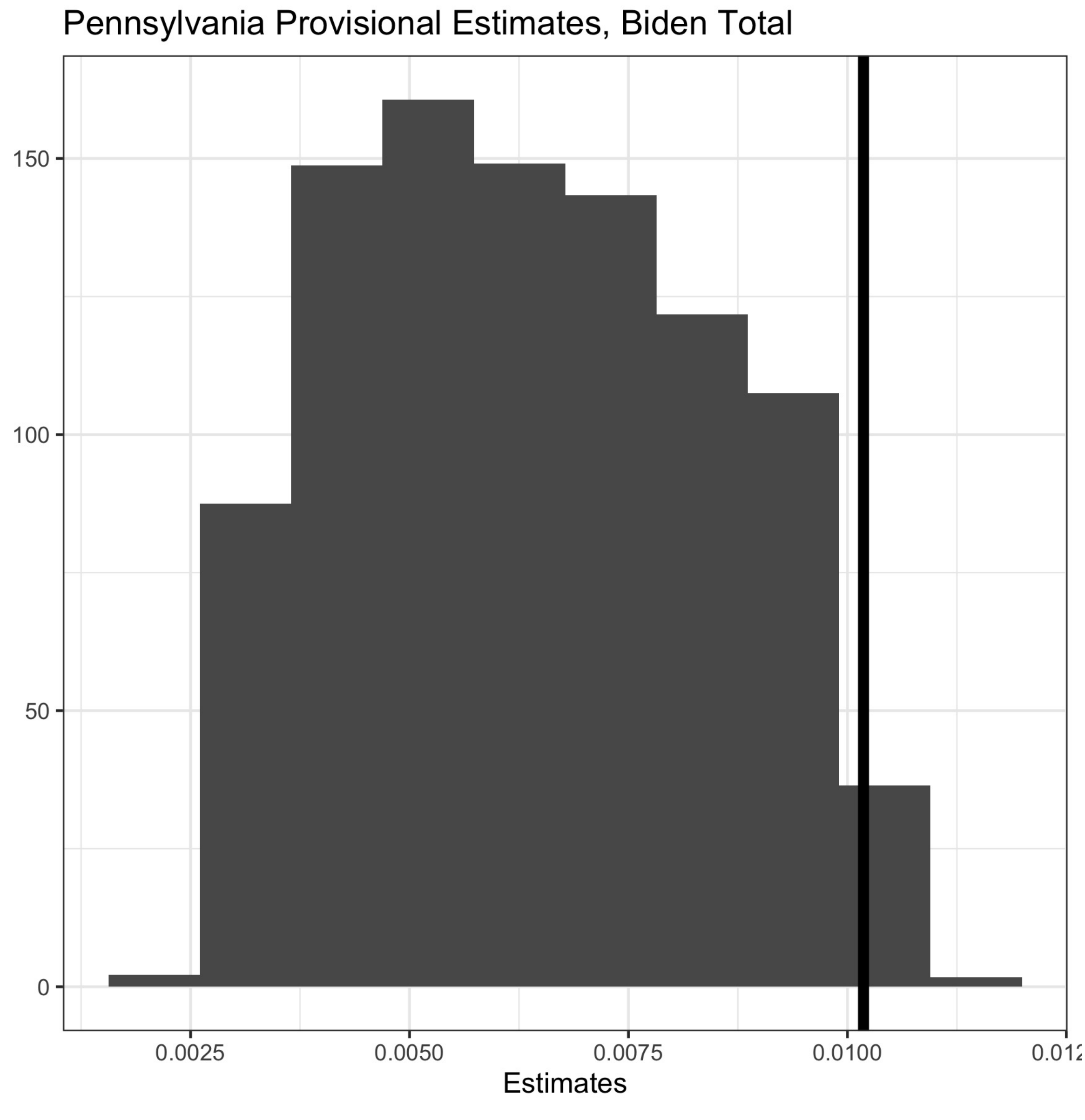


Figure 9: Distribution of Estimates for Alternative Precinct Differencing Orders, Share of Biden Ballots from Pennsylvania Provisional Ballots



6 Turnout Rate in States

Table 7: County-Average Turnout Rates

	Turnout 2016	Turnout 2020	Difference
States with "Suspect" Counties			
AZ	0.54	0.64	0.10
MI	0.63	0.73	0.10
NV	0.60	0.70	0.10
PA	0.58	0.67	0.09
GA	0.54	0.62	0.08
WI	0.67	0.74	0.07
States Without "Suspect" Counties			
NC	0.63	0.70	0.07
FL	0.64	0.70	0.06
OH	0.61	0.66	0.05