Electoral Studies xxx (2017) 1–20

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election☆

David Cottrell [a], Michael C. Herron [b, *], Sean J. Westwood [b]

[a] Program in Quantitative Social Science, Dartmouth College, 6108 Silsby Hall, Hanover, NH 03755, United States
[b] Department of Government, Dartmouth College, 6108 Silsby Hall, Hanover, NH 03755-3547, United States

ARTICLE INFO

Article history:
Received 20 March 2017
Received in revised form
23 August 2017
Accepted 1 September 2017
Available online xxx

ABSTRACT

As Republican candidate for president and later 45th President of the United States, Donald Trump has claimed repeatedly and vociferously that the 2016 General Election was tainted by massive voter fraud. Here we use aggregate election statistics to study Trump's claims and focus on non-citizen populations across the country, state-specific allegations directed at California, New Hampshire, and Virginia, and the timing of election results. Consistent with existing literature, we do not uncover any evidence supportive of Trump's assertions about systematic voter fraud in 2016. Our results imply neither that there was no fraud at all in the 2016 General Election nor that this election's administration was error-free. They do strongly suggest, however, that the expansive voter fraud concerns espoused by Donald Trump and those allied with him are not grounded in any observable features of the 2016 election.

In addition to winning the Electoral College in a landslide, I won the popular vote if you deduct the millions of people who voted illegally — Donald Trump, November 27, 2016[1]

1. Introduction

Regular and fair elections are the keystones of democratic governance (Lipset, 1959; Katz, 1997). These mechanisms translate voter preferences and opinions into elected officials, who ultimately make policy. Electoral fraud distorts the relationship between constituents and representatives, and for this reason alone the threat of voter fraud is inherently serious. Moreover, elections perceived as unfair can decrease electoral legitimacy (Norris, 2014), reduce governmental credibility (Magaloni, 2010), and undermine perceptions of voter efficacy (Elklit and Reynolds, 2002).

Insofar as it was repeatedly tarred by allegations of widespread voter fraud, the 2016 American General Election exemplifies these concerns. Despite a dearth of evidence that fraudulently cast ballots play an important role in American elections (e.g., Levitt, 2007; Minnite, 2010; Goel et al., 2016), as the Republican nominee for president Donald Trump claimed that he was at risk of losing the presidential contest to Democratic rival Hillary Clinton because of systematic voter fraud. Later as president-elect, Trump asserted that Clinton had received "millions" of improper votes, and he blamed his loss of the popular vote on illegal activity. And finally, as the 45th President of the United States, Trump asserted that voting in New Hampshire was tainted by fraud and that, in the absence of illegal Massachusetts voters, Trump would have won the Granite State's four electoral votes and then-United States Senator Kelly Ayotte, who lost a close election to former New Hampshire governor Maggie Hassan, would have been reelected.

Trump's expansive claims merit attention because of the role that elections play in democratic politics and on account of Trump's status—45th President of the United States. Moreover, assertions of voter fraud are a significant source of political division and conflict in American politics (Hasen, 2012; Bentele and O'Brien, 2013; Hicks

☆ The authors released an early version of this article on the Internet on November 29, 2016. We thank Hollye Swinehart and Morgan Waterman for research assistance; Rich Houseal of Research Services, Church of the Nazarene Global Ministry Center, for religious affiliation data; Dwayne Desaulniers of the Associated Press for assistance with 2016 election returns; three anonymous referees; and, seminar participants at the Freie Universität Berlin and the Hertie School of Governance for helpful comments. This academic research project is not affiliated with any of the candidates in the 2016 General Election, and none of the authors received external funding for the work described here. Herron thanks the Hertie School of Governance in Berlin for support while he was writing this article.

* Corresponding author.
E-mail addresses: david.cottrell@dartmouth.edu (D. Cottrell), michael.c.herron@dartmouth.edu (M.C. Herron), sean.j.westwood@dartmouth.edu (S.J. Westwood).

[1] https://twitter:com/realDonaldTrump/status/802972944532209664 (accessed December 8, 2016).

https://doi.org/10.1016/j.electstud.2017.09.002
0261-3794/

et al., 2014), and they are believed by a non-trivial segment of the voting population (Ansolabehere and Persily, 2008; Stewart et al., 2016). Lastly, simply because there was little voter fraud prior to November, 2016, does not imply perforce that Trump's claims are necessarily vacuous; it is always possible that 2016 was the first year in which systematic voter fraud was a meaningful factor in a presidential contest. These points motivated us in mid-2016 to develop an election fraud research project premised on the question, what could we academics say about election fraud in the aftermath of the then-upcoming presidential election? Our concern as of the summer of 2016 was that Trump might suffer a close loss in his bid for the presidency and react by leveling widespread accusations of voter fraud that, in principle, could explain his defeat at the polls.

Given the tenor of the Clinton-Trump presidential contest at the time of the Republican and Democratic party conventions, we anticipated post-election fraud allegations that centered on illegal voters, in particular non-citizens. To prepare ourselves to scrutinize such allegations, we assembled a county-level dataset that included historical election returns, demographics, and economic indicators. We also contracted with the Associated Press so that we would be able to access their national database on county presidential election returns. Our plan was to begin work on fraud allegations on Election Day evening (November 8, 2016), and we were prepared for an intense post-election week or two.

Since its inception, our research project has evolved in reaction to two developments. First, Trump did not lose the 2016 presidential election; this relieved us of the pressure to investigate fraud allegations made in the aftermath of a close Trump loss. Second, and seemingly in spite of his victory, Trump continued to invoke the specter of widespread voter fraud. This latter development has spurred on our project, the result of which is this article.

We would like to draw particular attention to our use of the term, "widespread," in the sense of what we are calling allegations of widespread voter fraud. Donald Trump, as candidate and then later as president, has not anchored his voter fraud claims on the likelihood of a person, here or there, voting illegally.[2] Rather, Trump and key supporters have spoken literally of "millions" of illegal votes, as our introductory quote makes clear. With this as context, our research project, an attempt to introduce scientific rigor into a debate largely dominated by bombastic claims, is not aimed at ferreting out what one might argue are more minor instances of voter fraud. While all instances of voter fraud are troubling, not all frauds are pivotal and not all frauds are systematic and widespread. Our research focuses solely on the possibility of massive and systematic fraud because fraud of this type in principle had the potential to be pivotal to the 2016 presidential election and because this is precisely the type of fraud against which Trump and his supporters, both before and after November, 2016, have regularly inveighed.

One can think of the analysis that follows as the proverbial canary, one that is an appropriate yet far-from-final step on the path of testing for voter fraud in the 2016 General Election. Detailed, individual-level audits, conducted on random samples of voters across jurisdictions spanning the United States, might be the ideal method to test for instances of voter fraud. However, in the absence of such audits, our analysis of aggregate county voting represents a valuable start. As will be clear shortly, we leverage variation in election outcomes across thousands of counties and connect that variation to a litany of explanatory variables, including counts of non-citizens provided by the American Community Survey. In the absence of a very expensive—and possibly unfeasible—audit of voter lists in jurisdictions across the United States, we believe that our aggregate analysis provides a significant advance in testing claims of voter fraud.

One could argue that an alternative method for testing voter fraud allegations would be to leverage a large-scale survey that questions respondents about, say, citizenship status and voting history. Such a survey would have the benefit of assessing the eligibility of voters individually as opposed to in the aggregate. However, unlike an audit, a survey in this vein would depend on the accuracy of the information volunteered by its respondents. This dependence is exemplified by Richman et al. (2014), who use the Cooperative Congressional Election Study to analyze the voting behavior of self-identified non-citizens; drawing on survey data, they estimate that 1.2 million non-citizens voted in the 2008 General Election. Ansolabehere et al. (2015) show, however, that this estimate reflects respondent data errors. Our use of aggregate data in conjunction with a corresponding lack of dependence on surveys allow us to avoid the sort of response problems that confound Richman, Chattha and Earnest.

We consider three allegations of voter fraud in the 2016 General Election: participation across the United States by non-citizens who supported Hillary Clinton in her presidential bid; concerns about voting in three states, California, New Hampshire, and Virginia, with particular attention to the possibility that Massachusetts voters tampered *en masse* with the United States Senate election in New Hampshire; and, finally, a conspiracy of election officials who attempted to "rig" the presidential election against Trump. The voter fraud accusations that we examine here span both national (non-citizen voting) and state-specific (e.g., New Hampshire), and all are associated with Donald Trump.

Briefly, we find little evidence consistent with widespread and systematic fraud fomented by non-citizens. Our analysis of returns in California, New Hampshire, and Virginia likewise turns up no evidence of problems in the vein raised by Donald Trump. And, our closer look at New Hampshire also yields nothing concrete. Lastly, and keeping in mind that the concern about a "rigged" election is ambiguous—we operationalized this idea by considering patterns in the way that election returns were released starting on the evening of November 8, 2016—we find no suspicious patterns in result timing.

Our results do not imply that there was no fraud at all in the 2016 presidential contest; indeed, we already know that the rate of fraud in the 2016 presidential election was not literally zero.[3] Nor do our results imply that the administration of the 2016 General Election was error-free. Nonetheless, they do strongly suggest that Trump's voter fraud allegations are not grounded in any observable features of the 2016 presidential election.

This article proceeds as follows. In the next section, we provide additional details on the motivation for and development of the research project whose results are described here. We then consider the aforementioned three sources of voter fraud, and we present results on them in sequence. Our final section concludes with suggestions for future research and how the academic community might want to consider studying voter fraud in upcoming elections.

[2] One of the more prominent, post-2016 election fraud situations in the United States involves a Mexican native who entered the country as a child. This case is documented in "Illegal Voting Gets Texas Woman 8 Years in Prison, and Certain Deportation," *The New York Times*, February 10, 2017, available at https://www.nytimes.com/2017/02/10/us/illegal-voting-gets-texas-woman-8-years-in-prison-and-certain-deportation.html (accessed June 15, 2017).

[3] See fn. 2.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

2. Studying voter fraud: motivation and project development

Allegations of voter fraud in the 2016 General Election from the Trump camp were alarmingly common. Beyond those already mentioned, Donald Trump and officials allied with him have asserted that records of deceased individuals are regularly used in the commission of voter fraud and insinuated that some urban areas, specifically Chicago, Philadelphia, and St. Louis, are hotbeds of fraudulent voting.[4]

As Goel et al. (2016) summarize, there are three general classes of voter fraud: impersonation (a voter casts a ballot while claiming to be someone else), double-voting (an individual votes more than once), and ineligible voting (an individual who is not supposed to have access to the franchise in a particular location casts a ballot). A voter casting a ballot out of her jurisdiction is an example of the third type of fraud; this form of fraud would also characterize a citizen ex-felon, who has lost the right to vote due to a state law restricting ex-felon voting rights (Manza and Uggen, 2006), improperly casting a ballot.

Efforts to uncover evidence of widespread voter fraud in American elections have come up empty. Surveys like Levitt (2007) and Minnite (2010) find only a small number of cases of verified voter fraud. Focusing on double-voting and using a database of 129 million individuals records, Goel et al. (2016) conclude that the maximum double-voting rate is approximate 0.02 percent and that "many, if not all, of [such] double votes could be a result of measurement error in turnout records" (p. 30). Christensen and Schultz (2014) conclude that, "if [voter fraud] occurs, [it] is an isolated and rare occurrence in modern U.S. elections" (p. 313).

This literature notwithstanding, it is challenging to prove that rampant voter fraud does not exist. Indeed, proving a negative is in general a strenuous task. Even though existing findings on voter fraud are clear in the sense of not uncovering evidence of widespread conspiracies, one could always argue that scholars are simply looking in the wrong places. Providing proof of American citizenship, for example, is not needed when registering to vote in federal elections; instead, to register an individual needs only to swear under penalty of perjury that he or she is a citizen.[5] Hence, it is theoretically possible for non-citizens to vote in federal elections in the United States and in so doing risk potential jail time and even deportation. Given the likelihood of being pivotal in an election, it is hard to believe that any non-citizen would want to behave in this way. Still, it remains the case that non-citizen voting is possible.

We know of no systematic audit of voter rolls that checks the citizenship status of individuals who voted. And it is not obvious how such an audit would proceed, give that there is no national identification system in the United States. Not all voters possess documents—like passports and birth certificates—that are definitive in the sense of proving citizenship. Thus, what one might call direct evidence of fraud might be difficult to come by, even in the presence of actual non-citizen voting.

While we lack the ability to verify the citizenship statuses of individual voters, we can nonetheless attempt to leverage the best data available to test national claims about widespread voter fraud. With this in mind, our research design and concomitant data are indirect insofar as our design seeks to identify downstream consequences of fraud as opposed to identifying directly which illegal voters cast ballots in the 2016 General Election.

If, as we wrote earlier, Trump were to have lost a close election in November, 2016, and responded with accusations of fraud, there would be little time for election officials and scholars to respond to what might be serious accusations of impropriety. It is hard to imagine that a truly comprehensive study of voter fraud in any presidential election could be carefully executed prior to deadlines imposed by institutions such as the Electoral College. States were required, for example, to have resolved issues surrounding appointment of their Electors no later than December 13, 2016, six days prior to Elector meetings on December 19, 2016.

Given what seemed like a plausible scenario as of the summer of 2016, we designed our fraud project to be national in scope and, most importantly, not conditioned on a particular post-election fraud claim. In addition, our approach needed to be feasible given data typically available in the aftermath of a presidential election. We did not want to rely on comparisons between election returns and pre-election polls or exit polls because comparisons like this are inevitably confounded with questions about sampling frames and representativeness.

In light of these exigencies, in the summer and early fall of 2016 we assembled an extensive county-level data set on historical election returns and other county features. Our plan was, starting on the evening of November 8, 2016, and as returns began to trickle in, to seek a rationalization of the presidential election using tools typically applied by academics studying presidential elections. We would then ask, do we observe across the country large and systematic deviations from our rationalization in a way that is redolent of the forms of voter fraud that Trump was regularly citing during his then-presidential campaign? Did we observe post-election, for example, evidence at the county level that the presence of certain classes of non-citizens was associated with increased support for Clinton? If we were to observe nothing like this, we would be more confident—albeit not completely confident—that voter fraud in the fashion envisioned by Trump did not play a pivotal role in the 2016 presidential election.

As we noted in the introduction, our plan was interrupted by two events. First, Trump did not lose the 2016 presidential election. Second, and despite his Electoral College victory, Trump has continued to maintain that the election was affected by widespread fraud and in so doing has leveled some specific claims. With these

[4] Allegations about non-citizen voting are described in "Donald Trump's got a new theory: Illegal immigrants are crossing the border to vote," *The Washington Post*, October 7, 2016, available at https://www.washingtonpost.com/news/the-fix/wp/2016/10/07/donald-trumps-got-a-new-theory-illegal-immigrants-are-crossing-the-border-to-vote (accessed June 15, 2017); about a rigged election at "Trump labels Clinton 'the devil' and suggests election will be rigged," *The Guardian*, August 2, 2016, available at https://www.theguardian.com/us-news/2016/aug/02/donald-trump-calls-hillary-clinton-the-devil-and-suggests-election-will-be-rigged (accessed June 15, 2017) and "Trump: Clinton should be in jail, the election is rigged," *USA Today*, October 15, 2016, available at http://www.usatoday.com/story/news/politics/elections/2016/2016/10/15/donald-trump-maine-new-hampshire/92143964 (accessed June 15, 2017); about California, New Hampshire, and Virginia at "California Official Says Trumps Claim of Voter Fraud Is Absurd," *The New York Times*, November 28, 2016, available at https://www.nytimes.com/2016/11/28/us/donald-trump-voter-fraud-california.html (accessed June 15, 2017); and, specifically about New Hampshire at "Trump to Dems: 'Pocahontas is now the face of your party'," *CNN*, February 11, 2017, available at http://edition.cnn.com/2017/02/10/politics/donald-trump-elizabeth-warren-voter-fraud/index.html (accessed June 15, 2017) and "Trump aide repeats debunked voter fraud claim, offers no new evidence," *CNN*, February 14, 2016, available at http://edition.cnn.com/2017/02/12/politics/stephen-miller-trump-voter-fraud (accessed June 15, 2017). Regarding deceased voters, see "Campaign 2016 updates: Donald Trump and Hillary Clinton prepare for final debate," *Los Angeles Times*, October 16, 2016, available at http://www.latimes.com/nation/politics/trailguide/la-na-trailguide-updates-pence-and-trump-disagree-over-whether-1476646363-htmlstory.html (accessed June 15, 2017). Finally, for allegations about Chicago, Philadelphia, and St. Louis, see "Trump tells supporters to watch polls in Chicago, St. Louis and Philadelphia on Election Day," *Chicago Tribune*, October 18, 2016, available at http://www.chicagotribune.com/news/nationworld/politics/ct-trump-voter-fraud-chicago-st-louis-philadelphia-20161018-story.html (accessed June 15, 2017).

[5] The National Voter Registration Act of 1993 required that states without same-day registration accept a uniform federal voter registration form that allows voters to register for federal elections. That form does not require documentary proof of citizenship. Recent attempts to require such documentation in Arizona and Kansas have been struck down by the Courts in *Arizona v. Inter Tribal Council* (2013) and in *Kobach* et al. *v. The United States Election Assistance Commission* (2015), respectfully.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

two developments in mind, neither of which we expected as of the summer of 2016, our research design reflects both continuity and change. With respect to the former, below we report results from our exercise of looking nationwide for evidence of voter fraud associated with non-citizens and a "rigged" election. Conducting research in this vein was our plan since mid-2016, and it would be inappropriate to abrogate a research project on voter fraud because the anticipated loser of an election turned out to be the winner. And with respect to the latter, given the specificity of some of Donald Trump's post-election claims about voter fraud, we decided to focus attention on these claims to check if there is evidence consistent with them. Overall, our research design represents a compromise between adhering to a pre-election plan and reacting to events that we could not have anticipated. We will return to this compromise in the conclusion, when we consider future efforts aimed at studying widespread voter fraud.

3. Results

Our results are in three sections. First, we offer a county-level analysis that addresses Donald Trump's claims about non-citizen voter fraud; allegations about non-citizens were promulgated pre-election, and we highlight the possibility of non-citizen voting in California, New Hampshire, and Virginia, three states which were mentioned explicitly by Trump post-election. Second, we continue our focus on states by analyzing a specific, post-election claim about New Hampshire. And third, we consider the timing of election results, and this reflects pre-election claims about a rigged election. What follows draws on many sources of data, all of which are listed in the appendix.

3.1. Non-citizen voting

Here we address the possibility of non-citizen voting in the 2016 presidential election. Counties are the smallest jurisdictions in the United States for which presidential election returns are tabulated nationally, and our non-citizen voting analysis is thus conducted at the county level. Among New England states, election results are tabulated by town, which in principle could push us in the direction of a town-level analysis (towns are in general smaller than counties). However, outside of election returns, the other variables we use in our non-citizen analysis are not nationally available below the level of county.

Counties are aggregate units that range in size from hundreds of residents to hundreds of thousands. This said, in an ideal world a study of non-citizen voting in the 2016 General Election might take the form of an individual-level audit where lists voters in this election are compared against lists of American citizens. We already mentioned this in the introduction and noted as well that we know of no publicly-available list of American citizens. This point notwithstanding, an individual-level study of non-citizen voting would avoid possible ecological fallacies that our aggregate analysis risks (e.g., Kramer, 1983). However, ballot secrecy makes it impossible to model vote choice at anything but the aggregate level. Accordingly, the majority of work on election outcomes relies on aggregate election data (e.g., Brians and Grofman 2001; Wand et al., 2001; Beber et al. 2012).

We drop counties not present in the Associated Press election returns database, meaning, the state of Alaska and Kalawao County, Hawaii. We also drop Oglala Lakota County, South Dakota, which experienced changes to its Census coding after 2010 and two Virginia localities, the city of Bedford and Bedford County, the latter of which incorporated the former prior to the 2016 General election. Of the 3142 counties and county equivalents in the United States, our non-citizen analysis covers 3,111, equivalent to a coverage rate of approximately 99 percent.[6] To keep our language as straightforward as possible, henceforth we refer to counties and county equivalents simply as counties.

3.1.1. Modeling approach

Our non-citizen voting analysis is based on a series of linear regression models. These models consider differences between a 2016 election variable—i.e., Hillary Clinton's vote share—and a corresponding variable from a previous election—i.e., Barack Obama's vote share from the 2012 General Election. The reason we model differences, as opposed to levels, in our study of non-citizen voting in the 2016 presidential election is because, intuitively, the former represent features of the election that require explanation. Whether Clinton did well in a particular county in the United States may be noteworthy (if, in contrast, Obama did badly there) or not (Obama did well there, too). If we want to understand whether widespread fraud affected the 2016 presidential contest, we argue that we should study the difference between Clinton's and Obama's vote share, not simply Clinton's alone.

Formally, our regression study of non-citizen voting in 2016 is based on three dependent variables. These variables are as follows:

- **Difference in Democratic vote share**: the difference between Clinton's share of the two-party vote in 2016 and Democratic presidential candidate Barack Obama's share of the two-party vote in 2012.
- **Difference in Democratic turnout**: the difference between the number of votes received by Clinton in 2016 and the number of votes received by Democratic presidential candidate Barack Obama in 2012, divided by citizen voting age population.
- **Difference in Republican turnout**: the difference between the number of votes received by Trump in 2016 and the number of votes received by Republican presidential candidate Mitt Romney in 2012, divided by citizen voting age population.

Of the three dependent variables above, most important is the Clinton-Obama vote share difference. Per Donald Trump, non-citizens planned to turn out in November, 2016, and cast votes for Clinton. As such, the first place we should look for evidence of widespread non-citizen voter fraud is in an elevated Clinton vote share, relative to Obama, in counties that contain disproportionately many non-citizens, *ceteris paribus*.

In terms of our second dependent variable, Trump's hypothesis of non-citizen voter fraud posits that, on account of ineligible voters casting ballots, there should be a surge (normalized by county size) in Democratic turnout in heavily non-citizens areas in the United States. If this were to have happened, then Clinton should have received more votes than Obama in counties with disproportionately many non-citizens, *ceteris paribus*.

We estimate a third regression with Trump-Romney turnout differences (total Trump votes from 2016 minus total Romney votes in 2012, normalized by county size) on its left hand side because this regression leads to a placebo test. Trump's theory of non-citizen voters is that they supported Hillary Clinton, not simply that non-citizens turned out to vote. With this in mind, a natural placebo test is one that involves the presence of non-citizens and Trump-Romney turnout differences. If we find evidence that counties with many non-citizens had disproportionately many Trump votes compared to Romney votes, *ceteris paribus*, that would

[6] On some census county boundary changes, see "Geography Note: U.S. - Changes to counties and county-equivalent entities," *Moody's Analytics*, July 15, 2015, available at https://www.economy.com/support/blog/buffet.aspx?did=50094FC4-C32C-4CCA-862A-264BC890E13B (accessed June 22, 2017).

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

indeed be odd. However, it would not look like the sort of non-citizen voter fraud that Trump predicted.

Our placebo test is useful because of potential misspecification problems in our regression analyses. These analyses are national in scope, which we believe is crucial, and purport to model election returns (vote share differences and turnout differences) in counties that differ from one another in size, partisanship, and economic conditions. Moreover, counties may differ in the extent to which their voters are politically mobilized, and this is very hard, if not impossible, to measure over the entire United States. Key political units in the country—United States Congressional Districts and state legislative districts, for example—do not coincide with county boundaries. This leads to a well-known problem in the study of American election administration: the geographies (counties) with the best data availability on socioeconomic variables are not units (for example, Congressional Districts) with particularly good political data (Chen, 2017).

Each presidential election is unique in some fashion, and it is hard to imagine that we can capture all of the important dynamics of the 2016 presidential contest with a county-level model. Our placebo test provides leverage over this problem.

Our regression models of vote share and turnout differences are grounded in the maintained hypothesis that the 2012 General Election was not affected by extensive and systematic voter fraud. This is consistent with literature already cited, none of which identifies fraud as a major problem in American presidential contests. Our use of differences also implies that we are not explicitly interested in understanding, say, where Clinton received more or fewer votes in 2016. This is an important matter for scholars of American electoral politics, but it is not our focus. Rather, we want to know where Clinton received more or fewer votes than one might have otherwise expected, based on how Obama performed in 2012, and whether these so-called extra votes were cast in locations with many non-citizens.

In terms of independent variables that ostensibly explain county-level vote share and turnout differences, our regressions contain control measures that draw on existing literature on American presidential elections. These variable touch on the role of the economy and retrospective evaluations in presidential voting (e.g. Bartels and Zaller, 2001; Fair, 2002; Cho and Gimpel 2009), the role of race (e.g., Tate, 1991; Abrajano and Alvarez 2010; Stout, 2015), and moral issues, which can be closely tied to religion (e.g., Hillygus and Shields 2005). We include as well in our regressions a measure of foreign born citizens insofar as immigration was a prominent feature of the 2016 presidential campaign. Finally, our regression models include state fixed effects, which are intended to proxy for across-state variability in areas like election administration and generic partisanship. To the extent that United State Senate races affected turnout and vote choices in the 2016 presidential rate, our state fixed effects should pick these up.

Beyond these control variables, our three regressions include the fraction of a county that is composed of non-citizens. This is the key variable in our study of non-citizen voting, and the tests that follow below turn on whether various non-citizen coefficient estimates are different than zero. That is to say, we want to know if the presence of non-citizens in counties is associated with unusual election vote share or turnout differences, even in the presence of variables that are routinely used to study aggregate vote choice in American presidential elections.

Our regression models have differences on their left hand sides (e.g., Clinton vote share from 2016 minus Obama vote share from 2012) yet include variables which purport to explain presidential vote share. Our logic here is as follows. To the extent that economic conditions are correlated with presidential vote shares, such conditions will also explain differences in candidates. Donald Trump's populist economic message in 2016 was not identical to Mitt Romney's in 2012, this despite the fact that both Trump and Romney are Republicans. When a county-level economic variable is correlated with Clinton-Obama vote share differences, then, it thus likely follows that this variable reflects the extent to which Trump's and Romney's economic messages (and, implicitly, Clinton's and Obama's messages) resonated with different slices of the American electorate.

3.1.2. County changes between 2012 and 2016

To fix intuition, Fig. 1 displays raw differences between Clinton's 2016 vote share and Obama's 2012 vote share. In the pictured map, darker (lighter) counties are those where Clinton lost (gained) vote share relative to Obama in 2012.

Fig. 1 shows that Trump was very successful in the upper Midwest, where Democrats lost significant support in 2016: Minnesota is quite dark, for example, as are most of Wisconsin and most of Iowa. In addition, there is a band of Trump support (that is, anti-Clinton support) in New England that starts in Maine and heads southwest, maintaining some distance from the east coast of the United States. Moreover, Clinton did poorly compared to Obama in parts of the West, notably in Nevada but not in Utah, where it appears Romney (as opposed to Obama) had particular support from the state's large Mormon community.

Perhaps the most notable pattern displayed by the map in Fig. 1 is the difference between urban and rural locations, the former of which saw gains in Democratic support. In the Midwest, for example, Chicago is lightly shaded, as are Detroit, Milwaukee, and Minneapolis-St. Paul. In contrast, less-urban counties in the periphery shifted in a pro-Republican, anti-Clinton direction. While Fig. 1 seems to suggest that Trump managed to gain an advantage in 2016 by securing the votes of the electorate living outside of major city centers, this sort of a map tends to over-represent conditions in non-urban counties, which are geographically expansive and hence relatively visible but at the same time lightly populated.

Fig. 2 contains a map of the total number of votes received by Clinton minus the number of votes received by Obama, normalized by county citizen voting age population. The figure shows that Clinton-Obama turnout differences varied in similar locations as the previously shown Clinton-Obama vote share differences. This is particularly true in regions of the United States with large Mormon, immigrant, and minority populations, for example, in Utah, California, Texas, and Chicago. Fig. 2 suggests that voters in 2016 were mobilized, as one might expect, at different levels by county.

3.1.3. Regression results

Table 1 contains results of estimating three regressions involving vote share and turnout differences. Our regressions are estimated with least squares and are weighted by the total two-party vote in 2016 (difference in vote share regression) and the total number of Clinton votes in 2016 (the two turnout regressions).

Before turning to the non-citizen variable in our regression models, we summarize Table 1's control variable results. The Clinton-Obama vote share of the table column shows that, *ceteris paribus*, counties with disproportionately high unemployment, low income, many men, many uneducated whites, fewer Mormons, fewer Jews, and fewer Muslims had less support for Clinton in 2016 than they did for Obama in 2012 (that is, more support for Trump in 2016 than they did for Romney in 2012). The Clinton-Obama turnout column in Table 1 is roughly similar. And, the Trump-Romney turnout column implies that, in those counties where Trump under-performed Romney, voters moved to Clinton and not to third party candidates. It is also the case that, *ceteris paribus*, counties with high unemployment rates, large groups of unemployed white males and large Hispanic populations showed greater

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

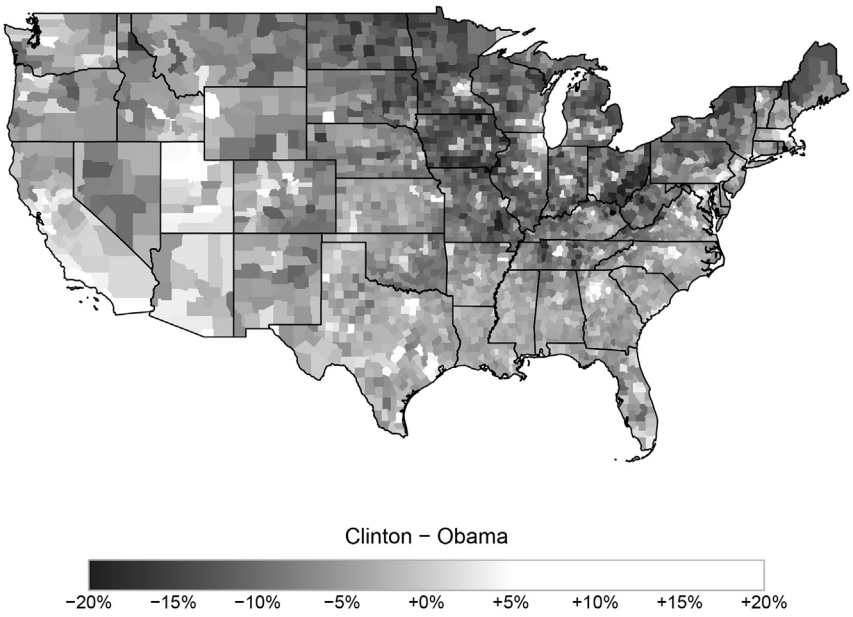


Fig. 1. Map of Clinton-Obama differences in vote share.

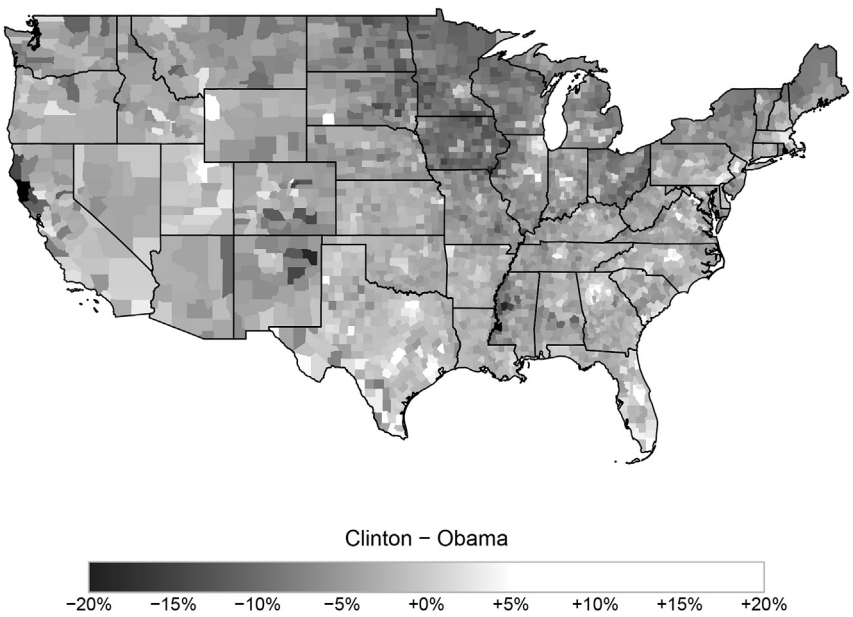


Fig. 2. Map of Clinton-Obama differences in turnout.

support for Trump than Romney.

It is worth emphasizing that the "Foreign born citizen" row in the table includes United States citizens only; the more such citizens, the greater the number of Clinton 2016 votes compared to Obama 2012 votes, *ceteris paribus*.[7]

With respect to our key voter fraud measure, the percentage of a county that is composed of non-citizens, the Clinton-Obama vote share column of Table 1 contains no evidence that Clinton's vote share in 2016 was elevated disproportionately in counties with many non-citizens, *ceteris paribus*. In particular, the non-citizen estimate in the Clinton-Obama vote share difference column is not statistically significant at conventional confidence levels, and this is not consistent with Trump's allegations of non-citizen voter fraud.

However, the middle column of Table 1 has a different result; in particular, it shows that counties with many non-citizens had disproportionately many Clinton votes, *ceteris paribus*. The Trump-Romney turnout column (recall, this is our placebo test) of the table has no such finding. This combination raises concerns that are in

[7] The racial percentages that appear in Table 1 are calculated using the citizen voting age population; religion affiliation variables are calculated by taking raw number of adherents from the 2010 U.S. Religion Census and dividing by county population; and, education variables are calculated based on the over 25 year old population.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

Table 1
Regression analyses of election differences.

	Dependent variable:		
	Clinton-Obama vote share	Clinton-Obama turnout	Trump-Romney turnout
	(1)	(2)	(3)
% Unemployed	−16.671***	4.279	15.853***
	(3.612)	(4.028)	(3.589)
Log median household income	1.942***	4.054***	1.503***
	(0.296)	(0.348)	(0.310)
% Employed in manufacturing	0.126	−3.418***	−4.147***
	(0.972)	(1.145)	(1.020)
% Urban	0.011***	0.003	−0.013***
	(0.002)	(0.003)	(0.002)
% Male	−11.383***	11.323***	−5.748*
	(3.466)	(3.778)	(3.367)
% White	38.849***	32.393***	−31.153***
	(8.933)	(10.069)	(8.972)
% Black	11.969	14.555	−22.112**
	(9.042)	(10.192)	(9.082)
% Hispanic	19.794**	41.598***	−22.300**
	(9.626)	(10.840)	(9.660)
% Asian	−2.928	10.393	−0.547
	(9.422)	(10.683)	(9.519)
% No college degree	13.509	21.734*	−14.200
	(10.187)	(11.434)	(10.188)
% White, no college degree	−53.210***	−38.964***	41.757***
	(10.354)	(11.628)	(10.361)
% Black, no college degree	−8.518	−19.712*	18.717
	(10.426)	(11.707)	(10.432)
% Hispanic, no college degree	−17.638	−42.566***	23.373**
	(11.122)	(12.502)	(11.140)
% Asian, no college degree	3.964	−1.024	6.604
	(11.429)	(12.998)	(11.582)
% Mormon	8.132***	6.502***	−21.403***
	(1.598)	(1.753)	(1.562)
% Evangelical Christian	6.065***	3.570***	−5.081***
	(0.515)	(0.607)	(0.541)
% Jewish	−8.682***	−7.206**	11.915***
	(3.069)	(3.637)	(3.241)
% Muslim	10.406**	17.912***	−2.751
	(4.257)	(5.109)	(4.553)
% Foreign born citizen	−0.046	−1.778	0.841
	(2.777)	(3.216)	(2.865)
% Non-Citizen	4.012	13.773***	−1.290
	(2.805)	(3.187)	(2.840)
Observations	3111	3111	3111
R^2	0.896	0.729	0.802
Adjusted R^2	0.894	0.723	0.797

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

line with Trump's theories of voter fraud. The concerns are certainly puzzling insofar as Table 1 highlights evidence of an elevated number of votes for Clinton yet no effect on Clinton vote share. *Prima facie*, this does not present like the sort of pro-Clinton voter fraud that Trump alleged. Still, this oddity is troubling enough for it to warrant additional analysis. We return to it shortly after considering particular allegations of voter fraud in California, New Hampshire, and Virginia.

3.2. Non-citizen voting in California, New Hampshire, and Virginia

If it were true that voter fraud among non-citizens occurred in California, New Hampshire, and Virginia, as claimed post-election by Donald Trump, we should expect that, one, Clinton outperformed Obama in these states (via accumulated extra voters) and, two, that Clinton gained vote share in these states. We look for the possibility that Clinton attracted non-citizen voters by including in our three regressions interaction terms for the three aforementioned states and non-citizen percentage. The results of our interaction-augmented model are in Tables 7, 8, and 9 in the appendix.

Contrary to Trump's claims, we find that, in California, New Hampshire, and Virginia counties with disproportionately many non-citizens, Clinton actually attracted *fewer* voters and *lost* vote share relative to Obama (coefficients for relevant interaction terms are negative and statistically significant at conventional confidence levels). We find no such effects in our Trump-Romney placebo.

To further probe the relationship between non-citizen voters and the 2016 election, we add to our regressions interaction terms to check if counties with larger numbers of non-citizens favored Clinton in battleground states (where non-citizen votes would be more likely to be pivotal) and in counties that border Mexico (where the theoretical supply of non-citizens is large). These interactions are both statistically significant and negative in our turnout model and in our vote-share model, implying that Clinton received fewer votes and less vote share in these locations. This is

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

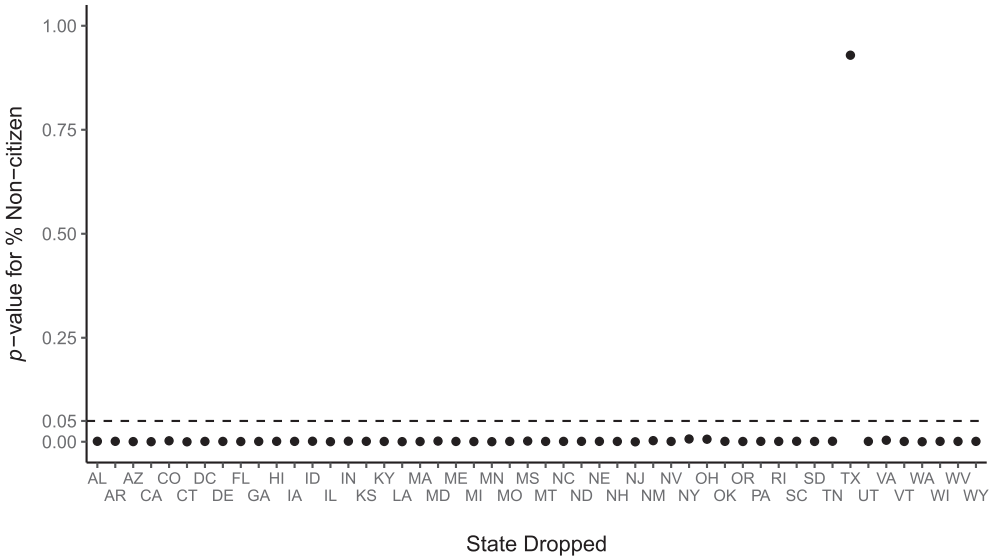


Fig. 3. Effect of Texas on the significance of percentage non-citizen.

inconsistent with the sort of non-citizen voter fraud alleged by Donald Trump.[8]

3.2.1. Non-citizen voting in Texas

While Table 1 shows no evidence of a relationship between the fraction of non-citizens in a county and Democratic improvements in vote share, we have already noted that they do indicate a relationship between the fraction of non-citizens in a county and improvements (Clinton votes minus Obama votes) in the total number of Democratic votes. Barring Donald Trump's fraud allegations, we would not have expected to find such a thing, and this significant relationship is therefore a cause for concern.

There are two potential explanations for the finding. The significant relationship between non-citizens and Clinton turnout, relative to Obama, could be an indication that many non-citizens indeed voted in 2016 (although they apparently did not all vote for Clinton—Table 1 does not show any non-citizen effects on Clinton's vote share). Or, the result could reflect excessive and legitimate pro-Democratic turnout in counties with many non-citizens.

As a robustness check on our puzzling finding, we test whether the relationship between non-citizens and Clinton-Obama turnout differences is sensitive to the inclusion of a particular state. If non-citizen voting were rampant and widespread in the 2016 presidential election, then removing a single state from our analysis should not have an effect on the significance of our results. Thus, we repeatedly estimated our Clinton-Obama vote share regression model from Table 1, each time removing a single state, and we record the *p*-value associated with the effect of non-citizens on Clinton-Obama vote share differences. This gives us 49 total regressions (recall that Alaska is not part of our analysis).

The non-citizen *p*-values from these 49 models are plotted in Fig. 3; the horizontal axis in the figure indicates the state that was removed from a given regression. It is evident that, for every removed state other than Texas, there is a significant effect of non-citizens on Clinton-Obama turnout differences in 2016. However, when Texas is removed, the significance of this relationship goes away. We conclude from this exercise that the effect in Table 1 of non-citizens on Clinton-Obama turnout differences is driven almost entirely by Texas, a state with 254 counties (approx. 8% of all counties) and 13.6% percent of the non-citizens in the United States.

Fig. 3 does not rule out the possibility that non-citizen voter fraud boosted Clinton's vote share in 2016, but it does imply that any further look at this matter needs to consider Texas in particular. With this in mind, Table 2 contains results of three regression models, all of which are restricted to Texas counties only. The variables in this table are identical to the variables in our earlier, national regression.

The results in Table 2 are not consistent with non-citizen voter fraud in Texas in the vein of Donald Trump's allegations. Namely, in this state we find, as in our national analysis, no effect of non-citizens on Clinton-Obama vote share differences. While we do find an effect of non-citizens on Clinton-Obama turnout differences, we find a similar effect on Trump-Romney turnout differences. What we have in Texas, then, is either evidence of non-citizen voter fraud in favor of both Clinton *and* Trump or unmodeled political mobilization in Texas counties with many non-citizens. Trump voiced concerns about fraud directed at him, as opposed to for him, and as such we interpret Table 2 as inconsistent with Trump's allegations.

Continuing with Texas, we now break down non-citizen percentages into three race/ethnic categories, Hispanic, Asian and white, and these categories appear in regression results reported in

[8] We experimented with including a measure of county voter registration in our regression models, in particular registration divided by total county population, albeit not in North Dakota, a state which does not have voter registration at all; such a measure might in principle proxy for local engagement or political mobilization. Our registration figures combine data from Leip Atlas with data that we downloaded for Maryland and calculated directly for Florida (by counting the number of lines for each county in the January, 2017, voter extract file). Sources are specified in the appendix. The lack of patterns in regression that we discussed previously do not change when our voter registration measure is included. This said, heterogeneity across states in how voter registration is handled complicates the interpretation of the fraction of a county that is registered. Some states distinguish between registered voters who are "active" and those who are registered yet inactive, and other states have no such distinction; procedures for removing voters from registration lists vary by state; some states allow pre-registration of individuals not yet of voting age; deadlines for registering to vote vary across states; and, as of this manuscript's writing 14 states allow same day voter registration. On this point see the report titled "SAME DAY VOTER REGISTRATION," published on January 11, 2017, by the National Conference of State Legislatures, available at http://www.ncsl.org/research/elections-and-campaigns/same-day-registration.aspx (accessed June 14, 2017). On account of this heterogeneity, we do not believe that the fraction of voters registered by county can be interpreted similarly across states, and for this reason we do not include voter registration in our main regression models.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

Table 2
Regression analyses of election differences, Texas counties only.

	Dependent variable:		
	Clinton-Obama vote share	Clinton-Obama turnout	Trump-Romney turnout
	(1)	(2)	(3)
% Unemployed	−31.940***	−13.795	19.601
	(9.080)	(9.313)	(15.941)
Log median household income	1.571**	0.987	2.847**
	(0.695)	(0.744)	(1.273)
% Employed in manufacturing	10.587***	13.025***	−8.268
	(3.023)	(3.234)	(5.536)
% Urban	0.029***	0.005	−0.028**
	(0.006)	(0.007)	(0.012)
% Male	12.629**	15.190***	−4.383
	(5.771)	(5.681)	(9.724)
% White	−37.908	−82.339**	−1.485
	(34.058)	(35.433)	(60.653)
% Black	−51.270	−96.594***	8.726
	(34.817)	(36.190)	(61.948)
% Hispanic	−37.532	−56.656	0.684
	(34.051)	(35.386)	(60.572)
% Asian	−102.843***	−72.923**	16.197
	(34.410)	(35.787)	(61.259)
% No college degree	−115.701***	−94.871**	94.496
	(43.162)	(45.121)	(77.235)
% White, no college degree	95.176**	79.346*	−76.815
	(43.705)	(45.703)	(78.232)
% Black, no college degree	110.834**	92.427**	−107.377
	(44.056)	(46.088)	(78.890)
% Hispanic, no college degree	94.805**	49.688	−88.594
	(43.427)	(45.379)	(77.677)
% Asian, no college degree	148.083***	70.240	−112.207
	(47.711)	(50.109)	(85.773)
% Mormon	5.183	16.005	51.917
	(18.473)	(20.168)	(34.522)
% Evangelical Christian	−2.309**	−5.129***	−8.421***
	(1.006)	(1.105)	(1.892)
% Jewish	703.223***	415.293***	−323.105***
	(65.709)	(70.491)	(120.662)
% Muslim	−15.108*	−43.459***	11.217
	(8.999)	(9.578)	(16.394)
% Foreign born citizen	64.867***	10.070	−18.359
	(8.044)	(8.338)	(14.272)
% Non-citizen	−3.932	20.515***	12.116**
	(3.028)	(3.103)	(5.312)
Observations	254	254	254
R^2	0.942	0.901	0.642
Adjusted R^2	0.938	0.893	0.611

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

Table 3. The leftmost three columns in this table are identical to our previous regression columns with the exception that non-citizen fraction has three categories. What we see in these columns is that the fraction of non-citizen Hispanics is associated with decreased Clinton-Obama vote share differences, increased Clinton-Obama turnout differences, and increased Trump-Romney turnout differences. This is consistent with excessive political mobilization in Texas counties with many non-citizens and is not evidence of non-citizen voter fraud as Trump characterized it.

The rightmost two columns of Table 3 constitute a final analysis of Texas, an analysis motivated as follows. One limitation of our results on turnout differences is that county population in the United States has not been static between 2012 and 2016. If counties with many new non-citizens in 2016 are also counties with many new citizens in 2016, we might expect to see more votes for Clinton and Trump in 2016, *ceteris paribus*, simply because there are more people in said counties. We do not have good measurements of the dynamics of the non-citizen population between 2012 and 2016, but we do know registered vote totals in Texas counties in 2012 and in 2016. Thus, we calculate modified Clinton-Obama and Trump-Romney turnout differences by differencing, for the former, the number of Clinton votes in 2016 for each county votes divided by the number of registered voters in 2016 with the number of Obama votes in 2012 for each county votes divided by the number of registered voters in 2012. We carry out a similar calculation for Trump-Romney turnout differences, and corresponding regression results are in the rightmost columns of Table 3. These columns show very similar, albeit perhaps slightly attenuated, results to the results using 2016 population data to compute the dependent measures.

In sum, we have followed up what looked to be a suspicious finding regarding the presence of non-citizens and the difference between total votes received by Clinton in 2016 and Obama in 2012. This finding led us to Texas, and a closer look at Texas does not turn up evidence of non-citizen voter fraud consistent with Donald Trump's allegations. Rather, what we see in Texas looks like political mobilization in counties with many non-citizens, mobilization that acted in both a pro-Democratic and pro-Republican direction. We suspect that within-county and within-city dynamics in Texas are responsible for our Texas results and that a detailed study of this state and its population dynamics would be necessary to understand fully why counties with non-citizens had more Clinton voters

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

Table 3
Regression analyses of election differences, Texas counties only.

	Dependent variable:				
	Clinton-Obama vote share	Clinton-Obama turnout	Trump-Romney turnout	Clinton-Obama registered turnout	Trump-Romney registered turnout
	(1)	(2)	(3)	(4)	(5)
% Unemployed	−27.230***	−16.504*	20.143	−13.432	4.440
	(8.603)	(9.158)	(16.137)	(8.535)	(15.752)
Log median household income	1.479**	0.310	2.853**	−0.321	−2.880**
	(0.713)	(0.759)	(1.337)	(0.722)	(1.333)
% Employed in manufacturing	10.491***	14.140***	−8.791	10.471***	−2.613
	(2.989)	(3.182)	(5.607)	(3.027)	(5.588)
% Urban areas	0.026***	0.005	−0.027**	0.015**	−0.029**
	(0.006)	(0.007)	(0.012)	(0.006)	(0.012)
% Urban clusters	0.008	−0.004	−0.036**	0.0004	−0.031**
	(0.008)	(0.009)	(0.015)	(0.008)	(0.015)
% Male	12.602**	18.535***	−2.543	17.430***	11.264
	(5.330)	(5.674)	(9.998)	(5.755)	(10.622)
% White	−29.178	−68.435	2.137	5.242	15.171
	(32.787)	(34.900)	(61.498)	(33.410)	(61.664)
% Black	−42.241	−85.092**	13.906	−12.433	16.173
	(33.363)	(35.513)	(62.579)	(33.974)	(62.705)
% Hispanic	−34.526	−46.037	1.749	7.220	20.039
	(32.662)	(34.767)	(61.265)	(33.261)	(61.389)
% Asian	−89.418**	−32.258	19.101	23.062	102.231
	(34.810)	(37.054)	(65.293)	(35.325)	(65.198)
% No college degree	−98.902**	−67.302	102.554	−2.294	166.106**
	(42.045)	(44.755)	(78.865)	(42.788)	(78.972)
% White, no college degree	79.268	49.355	−83.837	−8.531	−144.475
	(42.705)	(45.458)	(80.103)	(43.455)	(80.203)
% Black, no college degree	93.243**	64.215	−117.295	3.322	−162.123**
	(42.938)	(45.706)	(80.540)	(43.590)	(80.452)
% Hispanic, no college degree	86.010**	23.735	−92.947	−11.622	−155.233
	(42.245)	(44.968)	(79.240)	(42.968)	(79.305)
% Asian, no college degree	115.135**	38.968	−125.444	−46.409	−218.089**
	(46.948)	(49.974)	(88.062)	(47.743)	(88.118)
% Mormon	13.136	23.041	50.800	17.661	26.620
	(18.574)	(19.771)	(34.839)	(18.507)	(34.158)
% Evangelical Christian	−2.176**	−5.152***	−8.522***	−3.214***	−6.183***
	(1.015)	(1.080)	(1.904)	(1.019)	(1.881)
% Jewish	651.918***	397.705***	−350.347***	239.070***	−134.226
	(66.589)	(70.881)	(124.902)	(66.580)	(122.885)
% Muslim	−18.104**	−38.543***	8.951	−15.601	10.234
	(8.948)	(9.525)	(16.784)	(9.108)	(16.810)
% Foreign born citizen	66.202***	6.035	−17.488	27.039***	−39.276***
	(7.675)	(8.169)	(14.396)	(7.647)	(14.114)
% Non-citizen Hispanic	−7.137**	21.015***	11.708**	6.574**	12.366**
	(2.969)	(3.161)	(5.570)	(2.950)	(5.446)
% Non-citizen Asian	1.656	−38.339	28.401	−12.599	−9.837
	(19.388)	(20.637)	(36.365)	(19.477)	(35.948)
% Non-citizen White	25.473	24.457	−7.481	11.866	−64.543
	(19.937)	(21.222)	(37.396)	(19.966)	(36.850)
Observations	254	254	254	254	254
R^2	0.944	0.908	0.644	0.831	0.726
Adjusted R^2	0.939	0.898	0.609	0.814	0.699

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

and more Trump voters in 2016 than expected. Nonetheless, our goal was to ask whether our Texas results look like voter fraud in the way that Trump envisioned it, and we conclude that they do not.

3.2.2. Alternative baselines

Our tests thus far have offered little evidence to suggest that non-citizen voting increased Clinton's electoral performance in 2016 relative to Obama's in 2012. Perhaps, however, 2012 was not the appropriate baseline for our previous analysis. This could be the case because 2012 General Election results may not be comparable to 2016 results on account of idiosyncrasies. The dynamics of United States Senate and House races were different in 2012, and this could have affected political mobilization and thus presidential vote shares in 2012. While differences between 2012 and 2016 may be addressed by intercepts in our regression models, we nonetheless address the use of 2012 as a baseline now by providing an alternative baseline.

Rather than comparing Clinton's performance in 2016 to Obama's performance in 2012, here we use Obama's performance in 2008. This allows us to check whether our original results hold across elections. By using 2008 as a baseline, we risk confounding our results with the Great Recession, but this sort of a tradeoff is unavoidable. We repeat our previous regression analysis, now using a different baseline for the dependent variable. That is, we simply swap Obama's 2012 vote share and turnout with Obama's 2008 vote share and turnout, respectively, and results are in Table 10 in the appendix.

This table shows that our original results on the lack of evidence for non-citizen voter fraud are not sensitive to the use of 2012 as a

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

baseline. Namely, the fraction of a county that is non-citizen is still not significantly associated with Clinton's 2016 outperformance of Obama in terms of vote share. However, like before, the fraction of non-citizens remains significantly associated with Clinton's outperformance of Obama in terms of total votes. If the state of Texas is dropped from the 2008 baseline analysis, this finding goes away, as it did before.

We repeated our alternative baseline analysis using an average of 2008 and 2012 election returns to form a baseline. Results are in Table 11 in the appendix, and this table shows that the conclusions from the 2008–2012 average baseline are not qualitatively different than our original results.

3.3. The New Hampshire busing hypothesis

Allegations about voting in New Hampshire returned in mid-February, 2017. According to journalistic accounts, at a meeting with United States Senators, "[Donald Trump] claimed that he and [former United States Senator Kelly] Ayotte both would have been victorious in the Granite State if not for the 'thousands' of people who were 'brought in on buses' from neighboring Massachusetts to 'illegally' vote in New Hampshire."[9] This accusation—henceforth called the busing hypothesis—was detailed by presidential adviser Stephen Miller, who said, "I can tell you that this issue of busing voters into New Hampshire is widely known by anyone who's worked in New Hampshire politics. It's very real. It's very serious."[10] Post-winter 2017, neither Donald Trump nor Stephen Miller has offered additional evidence about New Hampshire and the busing hypothesis in particular.

There are three research approaches that one might take in light of this hypothesis. One, search for visual or photographic evidence of buses ferrying Massachusetts residents to New Hampshire voting locations on November 8, 2016. Two, conduct an audit of voter checklists maintained by New Hampshire towns; (a set of randomly selected) individuals on these lists could be queried to see whether they are domiciled in New Hampshire, as required by state law, or are actually residents of Massachusetts. Three, look for evidence of unusual patterns in election figures like turnout rates; oddities in these variables could in principle highlight unusual and possibly nefarious voting in the Granite State.

With respect to these three approaches, we know of no photographic evidence consistent with the busing hypothesis. The only compelling piece of visual evidence we have been able to locate on this matter is in fact a negative one. *The Boston Globe* included a story on February 10, 2017, which described how "[New Hampshire Secretary of State William Gardner] got a report that a busy precinct had a parking lot full of cars with Massachusetts license plates. When [Gardner] drove there [on November 8] to check it out, the report was correct. But the people who drove the cars were standing outside with campaign signs, not inside casting ballots."[11] In terms of auditing New Hampshire voters to ensure that they are all residents, the second approach to the busing hypothesis, the office of the New Hampshire Secretary of State mailed letters to "6033 individuals who signed domicile affidavits [when registering to vote] during the period from May 9, 2016 through December 31, 2016." Of these letters, 458 (approximately 7.6%) were returned by the United States Postal Service as "undeliverable."[12] We do not know of any follow-up efforts made to contact these individuals. Notwithstanding this Secretary of State effort, an audit of (randomly-selected) New Hampshire checklists would be an extensive and labor-intensive undertaking, one we believe would be best carried out by election administrators. Audits can be very useful exercises, but actually implementing one is beyond our scope as academics. The remaining research approach for the busing hypothesis is one involving actual election returns. Insofar as frauds are aimed at influencing these returns, it is natural to study them when looking for evidence of fraudulent voting. We turn to this now.

The 2016 General Election is not the first time that New Hampshire has been involved in accusations involving potential electoral irregularities. After the 2008 Democratic Presidential Primary, in which Hillary Clinton won a surprise victory over then-United States Senator Barack Obama, there were seemingly odd correlations between voting technology and Clinton returns: Clinton performed better in New Hampshire areas that used Accuvote optical scanning machines as opposed to hand-counted paper ballots. Herron et al. (2008) show that this correlation is an artifact of the relationship between local demographics and voting technology and that there is no evidence that Clinton's victory in the 2008 Primary was due to voting machine manipulation.

Key to the analysis to come is the following maintained hypothesis: the 2010 and 2012 elections in New Hampshire were not affected by voter fraud. We are comfortable with this hypothesis as we know of no credible allegations of widespread voter fraud in New Hampshire in 2010 and 2012. Our maintained hypothesis allows us to ground our analysis of 2016 election data with returns from 2010 and 2012. If one were to believe that elections in New Hampshire prior to the 2016 contest were riven by systematic voter fraud, our analysis would not be credible.

Kelly Ayotte's losing margin in November, 2016, was quite small—only 1,017 votes based on 353,632 votes cast for her and 354,649 for Maggie Hassan.[13] An aggregate data approach like ours is unlikely to be sufficiently sensitive to reject with 100% certainty the possibility that Ayotte's loss was driven by fraudulent votes. However, our approach is useful if one is concerned that an alleged fraud is widespread and systematic, which is appropriate given the allegation of "thousands" of Massachusetts residents voting in New Hampshire.

With this in mind, our analysis of the New Hampshire busing hypothesis is twofold. It reflects the fact that elections are characterized both by turnout and the way that voters choose candidates

[9] This quote is taken from "Trump brings up vote fraud again, this time in meeting with senators," *Politico*, February 10, 2017, available at http://www.politico.com/story/2017/02/trump-voter-fraud-senators-meeting-234909 (accessed June 15, 2017).

[10] This passage is drawn from an interview with Stephen Miller, the transcript of which can be found at "'This Week' transcript 2-12-17: Stephen Miller, Bob Ferguson, and Rep. Elijah Cummings," *ABC News*, February 12, 2017, available at http://abcnews.go.com/Politics/week-transcript-12-17-stephen-miller-bob-ferguson/story?id=45426805 (accessed June 15, 2017). See as well "Trump Adviser Repeats Baseless Claims Of Voter Fraud In New Hampshire," *NPR*, February 12, 2017, available at http://www.npr.org/sections/thetwo-way/2017/02/12/514837432/trump-adviser-repeats-baseless-claims-of-voter-fraud-in-new-hampshire (accessed June 15, 2017).

[11] Annie Linskey, Matt Viser, and James Pindell, "Trump makes groundless N.H. voter fraud claims," *The Boston Globe*, February 10, 2017, available at https://www.bostonglobe.com/news/nation/2017/02/10/trump-makes-groundless-voter-fraud-claims/fcnMJfLgOx0UAVhJeTS8TP/story.html (accessed June 13, 2017).

[12] This information is drawn from a February 15, 2017, letter from Anthony Stevens, Assistant Secretary of State of New Hampshire, to Brian W. Buonamano, Assistant Attorney General of New Hampshire. The authors thank David M. Scanlon, New Hampshire Deputy Secretary of State, for providing this letter, a copy of which is available from the authors.

[13] This margin is based on the summary of the New Hampshire United States Senate race on the website of the New Hampshire Secretary of State; see http://sos.nh.gov/2016USSGen.aspx?id=8589963690 (accessed March 17, 2017). According to the town level data issued by this same office, Kelly Ayotte received 353,627 votes and Maggie Hassan, 354,636. These numbers are very close, but not identical, to the summary numbers which appear in the body of the manuscript. Some press reports refer to a 743 vote margin between Ayotte and Hassan; see, for example, the aforementioned "Trump makes groundless N.H. voter fraud claims," noted in fn. 11.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

conditional on turnout. With respect to the former, if the busing hypothesis were true, then we argue that we should observe unusual turnout patterns in New Hampshire on Election Day. This leads us to compare absentee voting in New Hampshire—which by definition was beyond the scope of the busing hypothesis—with voting on Election Day. We consider voter turnout in New Hampshire overall and subsequently look within towns. With respect to voter choices conditional on turnout, we examine election returns in Kelly Ayotte's United States Senate race. We carry out this partisan analysis because the busing hypothesis states explicitly that then-Senator Ayotte was negatively affected by fraud. Our analysis of the Ayotte race includes a number of different elements, and it draws on historical election returns and the physical locations of towns, among other things. It also includes a placebo test that draws on the fact that New Hampshire shares a border with Vermont as well as Massachusetts.

3.3.1. Voter turnout in New Hampshire

New Hampshire tabulates election results by town (or city, but henceforth we use the term town to describe both cities and towns), and there are 259 such units in the state. Of these 259, the New Hampshire Secretary of State reports that 241 had a positive number of votes in the 2016 General Election. Some New Hampshire towns are divided into wards, and when necessary we aggregate wards to the town level.[14]

As initial leverage over the busing theory, we invoke the fact that this theory addresses Election Day turnout—what is called *regular* turnout in New Hampshire—but not absentee turnout. This type of variance can be useful, as illustrated by Wand et al.'s (2001) study of the effects of the butterfly ballot in Palm Beach County, Florida, in the 2000 presidential election. There are certainly many factors that can affect regular and absentee turnout in New Hampshire, but *prima facie* overall turnout trends across the state are not consistent with the busing theory. In particular, there were 68,013 absentee ballots cast in New Hampshire in 2012 and 75,303 in 2016; this constitutes an approximate 10.7 percent increase. Corresponding counts for regular ballots are 650,667 from 2012 and 680,529 from 2016, an approximate 4.6 percent increase. In other words, absentee voting in New Hampshire increased more in 2016 compared to 2012 than did regular voting. Insofar as the busing theory posits the existence of improper voting on Election Day, these two percentage changes are not consistent with it. This does not, of course, imply that the busing theory is incorrect, but it is a piece of evidence that is at odds with this theory.

We now disaggregate our regular and absentee ballot counts to the town level. The busing hypothesis leads us to expect to observe towns with surges in 2016 regular turnout, due to Massachusetts residents being bused in, compared to 2012 turnout. In contrast, we do not expect to see unusual turnout patterns in absentee balloting when we compare 2012 and 20116 turnout by town. This said, Fig. 4 plots the relationship between 2012 turnout and 2016 turnout, via both regular and absentee voting.

Fig. 4a shows that most New Hampshire towns had similar regular turnout numbers in the 2012 and 2016 presidential elections. Note that most points in the figure fall very close to the depicted 45-degree line; this is true for towns (grey dots) in the three southern counties in the New Hampshire that border Massachusetts as well as for towns (black dots) in non-southern counties. A similar relationship between 2012 and 2016 absentee voting can be seen in Fig. 4b. Indeed, a comparison of both panels in Fig. 4 suggests that turnout between 2012 and 2016 was relatively stable in New Hampshire, regardless of time of voting. Insofar as the busing theory highlights regular voting only, this pattern is notable.

With an eye on individual town turnout, 198 New Hampshire towns did have greater 2016 regular turnout than in 2012; however, over half of these—127, approximately 64 percent—also had greater 2016 absentee turnout compared to 2012 absentee turnout. This relationship is stronger among southern towns, where one might expect it to be weaker if the busing theory were correct. In particular, among southern towns, 78 had greater regular turnout in 2016 compared to 2012, and of these 56—approximately 72 percent—also had greater absentee turnout. Among non-southern towns, 120 had more regular votes in 2016 than in 2012, and of these approximately 59 percent also had more absentee ballots as well.

Our final comment about turnout in New Hampshire focuses on total votes cast in 2016 and census-based estimates of citizens per town. If the busing theory were correct, then we expect to see towns with more votes than voting-age citizens, and in fact there are five New Hampshire towns where this happened in 2016. These towns are as follows: Ellsworth (three more votes than citizens of voting age), Hebron (22 more votes), Millsfield (six more votes), Newfields (seven more votes), and Watterville Valley (62 more votes). *Prima facie* these five towns are indeed puzzling. The American Community Survey, however, publishes 90 percent margins of errors for their citizenship estimates, and we calculated associated confidence intervals for these five towns. For all five towns noted above, the total number of votes cast in 2016 is within a 90 percent confidence interval for total number of voting age citizens.

3.3.2. The Ayotte-Hassan United States Senate race

According to allegations put forth by Donald Trump, Kelly Ayotte would have won her bid for re-election to the United States Senate in the absence of Massachusetts residents being bused into New Hampshire to vote. Accordingly, New Hampshire towns where regular voting in 2016 surged compared to absentee voting should be of particular interest. With this as background, we turn to the New Hampshire United States Senate race.

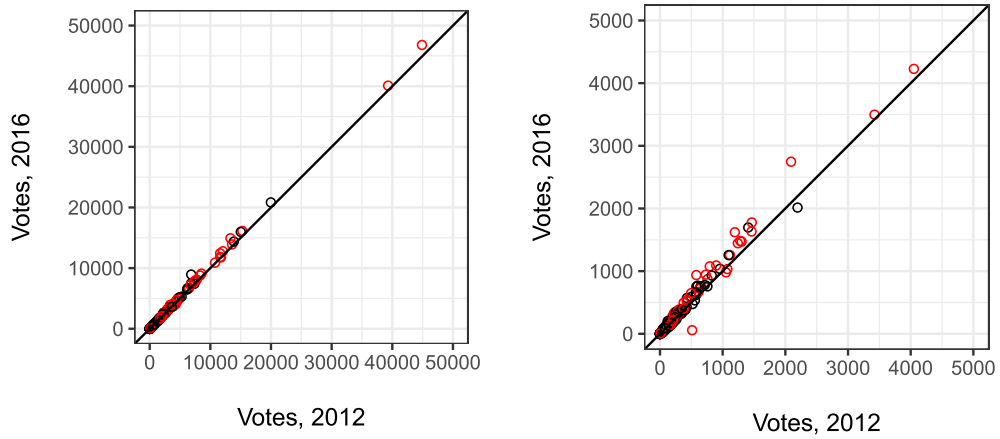
Fig. 5 contains two plots that speak to the busing theory and Kelly Ayotte's performance in her United States Senate contest. To fix intuition, Fig. 5a shows that, across New Hampshire towns, Ayotte's 2016 vote was systematically—apparently linearly—related to her 2010 vote. In 2010, Ayotte won a Senate seat decisively over Democrat Paul Hodes, and in 2016 she lost a very close election to former New Hampshire Government Maggie Hassan. Nonetheless, as the grey regression line in Fig. 5a shows, Ayotte's 2016 vote closely tracks her 2010 vote. This line has an R^2 value of approximately 0.99 and a slope of approximately 1.27. It is not surprising that Ayotte received more votes in 2016 than in 2010; the latter election was a midterm, as opposed to a general, election. Regardless, a systematic relationship between Ayotte's 2010 vote and her 2016 vote is not consistent with busing theory, which would suggest that Ayotte should have areas of substantial underperformance in 2016 compared to 2010.

New Hampshire does not tabulate ballots by voting time (regular versus absentee), and thus we cannot contrast Ayotte's 2016 performance on regular ballots with her 2016 absentee performance.[15] However, we can examine the relationship between, one, the difference in a town's increase in 2016 from 2010 in regular

[14] Of the 18 towns for which the Secretary of State reports zero total votes cast in 2016, 17 of them are in Coos County, the most northern and least populated county in the state.

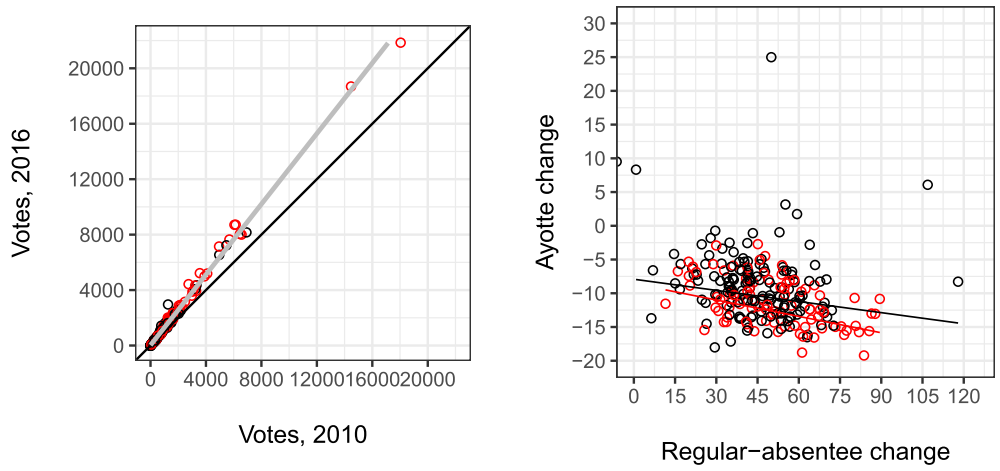
[15] We return to this point in the conclusion, but here it bears mentioning that an ability to distinguish between Election Day and absentee returns would have been be very useful for our efforts. Indeed, one can gain leverage against a fraud allegation aimed against a particular form of voting if returns from this form of voting can be distinguished from returns generated by other forms of voting.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002



(a) Regular turnout (b) Absentee turnout

Fig. 4. Voter turnout in New Hampshire, 2012 and 2016.
Note: each dot reflects a New Hampshire town where grey dots denote towns in the three counties, Cheshire, Hillsborough, and Rockingham, that share a border with Massachusetts.



(a) Votes cast for Ayotte, 2010 and 2016 (b) Voting method and Ayotte changes

Fig. 5. Kelly Ayotte United States Senate race.
Note: each dot reflects a New Hampshire town where grey dots denote towns in the three counties, Cheshire, Hillsborough, and Rockingham, that share a border with Massachusetts. The left panel contains a grey linear regression line that summarizes all towns, and the right panel contains two lines that summarize southern (grey) and non-southern (black) towns. The two right panel regression lines are weighted by total votes cast in 2016.

voting compared to absentee voting and, two, the difference in Ayotte support in 2016 and her support in 2010. This said, for each town we perform two calculations. First, we derive the rate of increase (or decrease) in regular ballots cast between 2016 and 2010 and we do the same for absentee voting; then we take the difference in these two rates, regular minus absentee. When a town has a large (small) difference, then said town had a relatively large increase (decrease) in regular voting in 2016 compared to 2010 absentee voting. When such a difference is zero, then the town had the same change (2010–2016) in regular voting compared to absentee voting. Second, we derive Ayotte's fraction of support in 2016 (her vote divided by total votes cast in her United States Senate race) and her corresponding fraction of support in 2010. We subtract these rates, 2016 minus 2010, and our two calculations described here are plotted in Fig. 5b.

What we see in this figure is, if anything, a negative relationship between change in voting method and Ayotte support. That is to say, the more a town had an increase in regular voting, 2010 to 2016, compared to absentee voting, the less of an increase in Ayotte support in 2016 compared to 2010. This relationship would not be particularly remarkable outside of the busing hypothesis; indeed, individuals who vote absentee versus election day can be different, and a correlation between voting method and political preferences should not be considered inherently troubling. In the context of the busing theory, though, we would have expected the opposite relationship in Fig. 5b compared to what is pictured there.

Our plots above shed light on the busing hypotheses (they are not consistent with it) but are limited for the follow two reasons: they do not incorporate sampling variance, and they are conceivably confounded by town-specific features that might be correlated with political partisanship and local support for Kelly Ayotte. To illustrate this latter concern in particular, consider Fig. 6, which describes New Hampshire town distances to the Massachusetts and Vermont borders and 2016 Ayotte support.[16]

Fig. 6a shows that support for Ayotte is high in New Hampshire

[16] A town's distance to a state border is defined as the shortest distance from anywhere on the border to the town's boundary.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

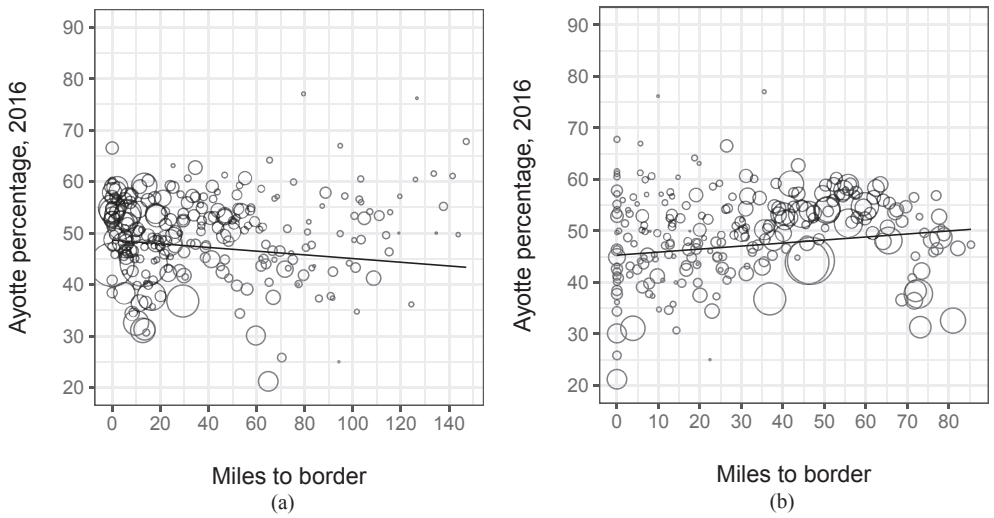


Fig. 6. Kelly Ayotte support and distance to Massachusetts and Vermont borders.
Note: each dot reflects a New Hampshire town. Small dots denote towns with up to 20,000 votes in the 2016 General Election, medium-sized dots towns with between 20,000 and 40,000 dates, and large dots towns with at least 40,000 votes. Regression lines are weighed by total votes cast in the United States Senate race.

towns that are close to the Massachusetts border (inconsistent with the busing hypothesis) yet low in towns close to the state's Vermont border (consistent with an alternative Vermont busing hypothesis). Of course the distribution of political preferences over United States Senate candidates may not be uniform across New Hampshire, and with this in mind our final look at the busing hypothesis considers a regression model that seeks to explain Ayotte support in 2016 using standard political indicators, town distance to the Massachusetts border (included because of the busing hypothesis) and distance to the Vermont border (included as a placebo). Table 4 contains regression results, which are weighted by total Senate votes cast in 2016.[17]

Our regression results are as follows. Regarding standard political variables, Ayotte share in 2010 is a strong predictor of Ayotte share in 2016; we previously saw evidence of this in Fig. 5a. In addition, the fraction of a town's voter check list that is composed of Democratic (Republican) voters is a strong predictor of whether the town is anti- (pro-) Ayotte in 2016. None of these results is surprising.[18]

In terms of variables associated with the busing hypothesis, we observe the following based on Table 4: *ceteris paribus*, towns with more election day voters were disproportionately pro-Ayotte, and towns far from the Massachusetts border were disproportionately anti-Ayotte. These two findings run counter to the busing hypothesis. Finally, our use of distance to the Vermont border as a placebo works as expected: there is no evidence that towns close to the Vermont border were different than towns far from it, *ceteris paribus*.

In summary, we have offered several perspectives on the 2016 General Election in New Hampshire, perspectives motivated by Donald Trump's allegation that Massachusetts voters were bused into the former state to cast ballots on November 8, in particular against Republican United State Senator Kelly Ayotte. The statistical evidence we have offered, based on town-level election returns from 2010, 2012, and 2016, offers no support for the so-called Massachusetts busing theory and in some cases runs counter to it. In addition, to the best of our knowledge, there exists at the time of this article's writing no visual evidence of buses on November 8, 2016, disgorging Massachusetts passengers in the vicinity of New Hampshire towns. This lack of evidence is notable in conjunction with our statistical results.

Table 4
Regression analysis of Kelly Ayotte vote share in 2016.

	Dependent variable:
	Ayotte share
Ayotte share 2010	0.396*** (0.031)
Fraction regular ballots 2016	0.106 (0.041)
Fraction Republican checklist	0.453*** (0.038)
Fraction Democratic checklist	−0.252*** (0.029)
Logarithm, Massachusetts distance	−0.005*** (0.001)
Logarithm, Vermont distance	0.002 (0.002)
County: Carroll	0.011 (0.008)
County: Cheshire	0.003 (0.009)
County: Coos	0.049*** (0.011)
County: Grafton	0.026*** (0.009)
County: Hillsborough	−0.005 (0.006)
County: Merrimack	0.002 (0.006)
County: Rockingham	−0.010 (0.006)
County: Strafford	0.005 (0.007)
County: Sullivan	0.034*** (0.010)
Constant	0.081* (0.044)
Observations	234
R^2	0.960
Adjusted R^2	0.958

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Belknap County omitted as base category.

3.4. A rigged election?

During the 2016 presidential campaign, Donald Trump regularly charged that the 2016 presidential election was going to be "rigged." The precise meaning of this characterization is difficult to pin down, but one interpretation of it is that Trump expected election officials across the United States to be biased against him.

We start by assuming that Trump's narrative of a rigged 2016 presidential election is valid. Insofar as most American elections are administered at the county level, this narrative impugns county

[17] One is added to town distances to the Massachusetts and Vermont borders because bordering towns have a distance of zero.

[18] Elections in New Hampshire towns are administered by town clerks. Clerk positions are elected locally but are non-partisan. We confirmed with David M. Scanlon, New Hampshire Deputy Secretary of State, that New Hampshire municipalities do not have partisan elections for local offices (email received on June 12, 2017). Because of this, we do not control for clerk partisanship in Table 4's regression.

election officials in particular. If these officials acted nefariously in a pro-Clinton way, then they might have tampered with the release of election returns. This might mean, for instance, that results from early-reporting counties were different than results from late-reporting counties, *ceteris paribus*, or that county results dramatically changed at the last minute. We recognize that timing-based fraud captures only one component of a potentially rigged vote, and we anticipate that future research will develop additional perspectives on this issue.

In the aftermath of November 8, 2016, "stringers" working for the Associated Press (AP) collected county-level election returns across the United States. When a stringer reported presidential results, the AP updated its nationwide database. In some cases, a county's report of its presidential returns covered all results for the county; for counties like this, the AP database includes only a single county report. In other cases, a county reported initial results based on a small set of precincts, later updated results (often many times) based on more precincts, and finally produced a full set of results based on all precincts in the county.[19] For a county that followed this pattern, the AP presidential election database includes multiple reports.

We associate with each county in the United States the number of minutes between the county's first report of election returns and its last. Overall, these reporting durations are roughly trimodal, and this is depicted in Fig. 7. As the histogram shows, some counties reported all of their results at once (283 counties in total) or nearly all at once, and the reporting durations for these counties is literally zero or close to this value. Other counties took a few hours before their results were finalized. And, in a few cases, some counties took several days and in some cases, a few weeks.

Suppose that anti-Trump county election officials deliberately tried to manipulate vote totals when it became apparent that Clinton was at risk of losing the 2016 presidential election. If this were to have happened, then we would have observed the following. First, later reporting counties would have been disproportionately pro-Clinton; and, these counties would have been pivotal counties in pivotal states. With this in mind, consider the cluster of reporting times around 10,000 min (almost a week). These counties represent Utah, easily won by Trump, and Washington, easily won by Clinton. Among late-reporting Utah counties, Trump won by 153,543 votes; among late-reporting Washington counties, Trump lost by 481,404 votes. In both states, late-reporting counties were not pivotal and thus had no effect on the Electoral College.

The cluster of counties reporting results far after the election (around 28,000 min, or approximately 19.4 days, after polls closed) includes just three states, California, Michigan and Wisconsin. Results in California are historically slow as the state allows for voters to mail in ballots up to the day of the election and also allows for provisional ballots that must be adjudicated. The outcome in Michigan appeared relatively clear for weeks, but the state endeavored to avoid reaching a re-count threshold and conducted an exhaustive tabulation before certifying. Wisconsin was extremely close and ended up completing a recount, which added 162 votes to Trump. In all cases, there is no evidence of results suddenly changing during extended tabulation periods. California

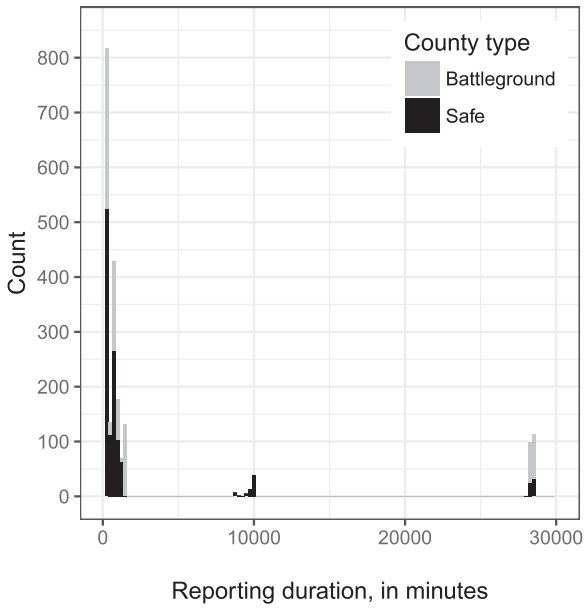


Fig. 7. Distribution of reporting time durations.

supported the Democratic presidential candidate, Clinton, as is typical of the state in recent years. Although Michigan results were not certified until November 28, 2016, the long delay in this state is not compatible with anti-Trump voter fraud: Trump won Michigan and the results from this state were not pivotal in the Electoral College. The situation with Wisconsin is similar.[20]

To explore potentially unexpected county-level changes that took place during vote tabulation processes, we consider changes in the presidential leader reported by a county as a function of percentage of precincts counted. A county can have zero changes if, for example, Trump were ahead of Clinton in every report made by the county. On the other hand, if Trump were ahead on the first report, Clinton in the second, and Trump in the third, then the county had two so-called leader flips. If, say, a subset of county election officials were rigging the election in favor of Clinton, then we might expect to see county vote reports suddenly changing, as corrupted election officials made their presence known. We note that leader flips often occur naturally as returns come in from different types of precincts, urban versus rural for example.

Of 15,796 Associated Press observations from 3111 counties in our data, we find only 251 instances where the winner of a county flipped from Trump to Clinton or from Clinton to Trump. These flips are summarized in Table 5. In the change from 75% to 100% of precincts reporting, 84 flips occurred—24 in battleground and 60 in safe states—representing a net gain of 41,388 votes for Trump (33,165 votes in battleground states). In battleground states, 12 flips were from Clinton to Trump. This is the opposite of what we would expect if the election were rigged against Trump.

Restricting attention to the 24 flips in battleground states, Table 6 shows that ten of these occurred in states with a small margin of victory. In two states—Florida and Pennsylvania—flips actually increased Trump's lead. In Arizona, Michigan, New Hampshire, Virginia, and Wisconsin, flips decreased Trump's margin but never by more than 6833 votes. In all of these cases, changes in votes associated with flips for Trump were not sufficient to sway a state-level election.

Perhaps the last opportunity to rig the 2016 election outcome occurred when counties reported final vote totals. Looking at just the last two reports from each county (in counties with more than two reports), there were flips in just three counties—Hillsborough

[19] We use the word "precinct" here generically to mean a local voting jurisdiction.

[20] On Michigan, see "Board of Canvassers certifies Trump victory in Michigan," *Detroit Free Press*, November 28, 2016, available at http://www.freep.com/story/news/politics/2016/11/28/trump-wins-michigan-board-canvassers/94547130/ (accessed June 15, 2017) and on Wisconsin, "Pennsylvania and Wisconsin End Election Recount Efforts," December 12, 2016, available at https://www.nytimes.com/2016/12/12/us/pennsylvania-and-wisconsin-end-election-recount-efforts.html (accessed June 15, 2017).

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

Table 5
Flips by precincts reporting.

Precincts reporting	Battleground flips	Safe flips
< = 25%	25	33
> 25% and < = 50%	23	33
> 50% and < = 75%	16	37
> 75%	24	60

County in New Hampshire, Dubuque County in Iowa, and Sauk County in Wisconsin—with these flips decreasing Donald Trump's margin by 856 votes. This number of votes is insufficient to have swayed the election in New Hampshire, Iowa, or Wisconsin (where the final margins were 2,736 votes, 147,314 votes and 22,748 votes, respectively).

4. Conclusion

We have presented here a number of analyses that speak to claims about widespread and systematic voter fraud in the 2016 United States General Election, claims promulgated by Donald Trump both before he became president and after his inauguration. These claims include allegations about non-citizen voter fraud, allegations specific to three states, and a national conspiracy of election officials. Our empirical results share a common theme: they are inconsistent with fraud allegations made by Trump. The results are, however, consistent with various state-level investigations conducted in the initial months of 2017, all of which have failed to find any evidence of widespread voter fraud in the 2016 General Election.[21]

The aftermath of the 2016 presidential election has witnessed the established of a presidential voter fraud commission, chaired by Vice President of the United States Mike Pence. Kansas Secretary of State Kris Kobach, known publicly for his advocacy of voter identification laws, is the vice chair of the commission.[22] While the future of this commission is hard to predict, the results we have offered here lead us to be skeptical of its grounding in claims about widespread vote fraud in the 2016 General Election.

We conclude with two broad comments intended to stimulate research in voter fraud, both within and beyond the United States. First, and with particular attention to our study of the possibility of widespread non-citizen voting, our reliance on aggregate data implies that our conclusions should be understood as exploratory. The research project we describe here is not an audit in which names on voter lists are checked against actual voters known with certainty to be citizens. Consequently, our analysis does not have any power over the question of whether a small number of non-citizens voted in the 2016 presidential contest. Instead, we have employed aggregate returns in a search for glaring irregularities that, were they to be discovered, might necessitate detailed exploration with refined, individual-level data. Audits have an advantage in that they are extremely detailed, but the time and expense required to verify addresses, identities, and citizenship statuses of millions of voters should not be underestimated.

Table 6
Net Trump vote gain in battleground flipped counties.

State	Unit	Flips	Net Trump gain
Arizona	Apache County	1	−6833
Florida	Pinellas County	1	12455
Iowa	Bremer County	1	2139
Iowa	Clinton County	1	1217
Iowa	Des Moines County	1	1325
Iowa	Dubuque County	1	2068
Iowa	Jefferson County	1	203
Iowa	Muscatine County	1	1906
Iowa	Warren County	1	4053
Michigan	Marquette County	1	−1473
Michigan	Muskegon County	1	−2229
New Hampshire	Hillsborough County	2	−382
Pennsylvania	Berks County	1	15855
Pennsylvania	Centre County	1	−1588
Pennsylvania	Erie County	1	4838
Pennsylvania	Northampton County	1	5543
Virginia	Montgomery County	1	−1747
Virginia	Staunton city	2	−83
Virginia	Suffolk city	1	−1679
Wisconsin	Douglas County	1	−1883
Wisconsin	Sauk County	2	−540

Note: counties and cities grouped by states, which are listed alphabetically.

Second, we believe that the sort of project described here should be a fixture of future American presidential elections and other important elections in the democratic world. It remains to be seen whether Trump's claims about voter fraud were idiosyncratic to his personality or whether the 2016 General Election is a harbinger of things to come. Either way, there are temporal and political pressures in the immediate aftermath of all important elections, and research projects aimed at ferreting out massive voter fraud should be initiated prior to voting day.

In thinking about future elections and fraud detection efforts, it is important to recognize that voter turnout and candidate (or party) vote shares are separate variables that merit attention. With respect to the former, a researcher interested in engaging claims about fraud should arm herself with historical turnout data at as low a level of aggregation as possible. Potentially troubling turnout spikes cannot be evaluated in the absence of historical context.

With respect to candidate vote shares, we suggest the following. Before an election occurs, a researcher interested in the sort of exploratory analysis of widespread fraud along the lines of our study of non-citizen voting should publicly pre-commit to a set of covariates that, when voting concludes, will be used to study an election outcome. In our case, these covariates appear in Table 1; we selected our socioeconomic variables based on literature that engages American presidential elections, and analyses of other contexts would need to pay heed to relevant scholarship prior to an election of interest. If such an approach were to yield a positive result for fraud, there would be two possible responses. One, the underlying model for studying the election is misspecified; or two, a more refined study of fraud, along the lines of an audit, is called for. Sorting between these explanations will depend on the idiosyncrasies of the election being scrutinized.

Cantú (2014) describes an alternative way to think about election fraud research, one that also could prove fruitful prior to elections. Briefly, Cantú studies the 2010 gubernatorial elections in Mexico and identifies fraud using the fact that voters in Mexico are assigned to local voting stations by name, i.e., via a device that is ostensibly uncorrelated with political preferences. Sorting devices like names may not necessarily commonplace, but the lesson implicit in Cantú is that researchers can learn a fair bit about problematic voting patterns in the presence of a useful sorting device. Boundaries between counties and states in the United States might

[21] For example, see "States find little evidence of voter fraud in months after elections," *The Hill*, March 6, 2017, available at http://thehill.com/homenews/state-watch/322522-states-find-little-evidence-of-voter-fraud-in-months-after-election, "Valid voter fraud complaints in California? Dozens, not millions," *CALmatters*, March 8, 2017, available at https://calmatters.org/articles/valid-voter-fraud-complaints-in-california-dozens-not-millions/, and "Ohio official finds 82 non-citizens had cast votes illegally," *The Toledo Blade*, February 28, 2017, available at http://www.toledoblade.com/Politics/2017/02/28/Ohio-official-finds-82-noncitizens-had-cast-votes-illegally.html (accessed June 15, 2017).

[22] See "Presidential Advisory Commission on Election Integrity," July 13, 2017, https://www.whitehouse.gov/blog/2017/07/13/presidential-advisory-commission-election-integrity (accessed August 23, 2017).

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

be valuable here. For example, if one believes that non-citizen voting is a serious matter, then a boundary between a state with strict voter identification laws and a state without such laws should distinguish areas that are subject to (illegal) turnout spikes and areas that are not. This could lead to a difference-in-difference design.

Finally, fraud detection efforts are for the most part dependent on official sources of data, and this raises the question of what types of data election officials should generate in the immediate aftermath of an election. Beyond the obvious (vote totals for candidates and overall turnout), a breakdown of election returns by type of voting (mail-in absentee, early, and Election Day) could prove very valuable for fraud research. The reason for this is clear, as we noted earlier in this manuscript: if an allegation of fraud targets Election Day voting, then comparisons of Election Day returns to absentee returns may in some cases be very useful. There are certainly privacy concerns associated with breaking down returns by voting time, and these concerns would have to be balanced against the importance of fraud detection efforts.

Another resource that would be useful is data on provisional ballots. Widespread non-citizen voting, if it exists, may trigger spikes in provisional ballots and other registration problems. If jurisdictions regularly reported counts of provisional ballots, and at low levels of aggregation, election researchers could use them to identify problematic areas. In light of allegations that Russian hackers may have accessed voter registration files across the United States in the run-up to the 2016 General Election, data on provisional ballots would be even more valuable. Suppose that a foreign entity were in fact to erase a large number of registration records from a county's list of registered voters prior to voting; this will probably have the observable implication of leading to a jump in provisional ballots, triggered when voters who records were compromised tried to vote.

We suggest as well that future research on election fraud track partial election returns as they are reported, as we have discussed in our results timing section above. Changes that occur overtime within polling locations or in other sub-county units might be useful as well. If it is true that elections are broadly rigged in the sense of that some politicians asserted prior to November, 2016, timing might be very valuable, particularly in the event of a close election where a small number of counties flipped results at the last minute.

We encourage as well the development of comprehensive data on local election officials. To the extent that one is concerned about voter fraud albeit not on a national scale, one might wonder about purely local efforts. Per the nature of federalism in the United States, there is significant variance across and within states as to how elections are handled locally. At the time of this article's writing, there is no comprehensive database on local election administration, the extent to which local officials are partisan, and in which direction. Election returns accompanied by local administrative data could prove very useful in fraud research.

Regardless of potential impediments to fraud research, many of which turn on limited data availability in the aftermath of an election, the role of fair elections as arbiters of the popular will implies that fraud analysis should be a regular part of electoral practices. We hope that the project described here constitutes a small part of what becomes a widespread and general effort.

Appendix

This appendix contains two sections. The first is a set of regression tables that are referenced in the text. The second is a list of data sources.

Additional regression tables

Table 7
Regression analyses of Clinton-Obama differences in vote share, with interaction variables.

	Dependent variable:		
	Clinton - Obama vote share		
	(1)	(2)	(3)
% Unemployed	−15.147***	−14.998***	−16.874***
	(3.660)	(3.663)	(3.609)
Log median household income	2.079***	1.958***	1.964***
	(0.300)	(0.296)	(0.296)
% Employed in manufacturing	−0.093	0.113	0.041
	(0.972)	(0.971)	(0.971)
% Urban	0.011***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)
% Male	−12.534***	−11.266***	−10.887***
	(3.474)	(3.463)	(3.468)
% White	42.137***	38.229***	36.186***
	(9.119)	(8.927)	(8.982)
% Black	14.945	11.243	9.296
	(9.196)	(9.038)	(9.091)
% Hispanic	25.242**	18.003*	17.187*
	(10.242)	(9.640)	(9.668)
% Asian	2.074	−0.371	−7.349
	(9.626)	(9.462)	(9.563)
% No college degree	17.869*	12.670	10.361
	(10.430)	(10.182)	(10.248)
% White, no college degree	−57.158***	−52.407***	−50.154***
	(10.572)	(10.348)	(10.410)
% Black, no college degree	−12.155	−7.747	−5.476
	(10.614)	(10.420)	(10.480)
% Hispanic, no college degree	−24.023**	−15.172	−14.854
	(11.787)	(11.150)	(11.162)
% Asian, no college degree	−1.419	1.591	6.779
	(11.593)	(11.453)	(11.468)
% Mormon	8.433***	8.039***	8.119***
	(1.597)	(1.597)	(1.597)
% Evangelical Christian	6.039***	6.071***	6.049***
	(0.514)	(0.514)	(0.514)
% Jewish	−8.156***	−8.857***	−10.241***
	(3.079)	(3.067)	(3.123)
% Muslim	9.756**	10.477**	10.150**
	(4.254)	(4.253)	(4.254)
% Foreign born citizen	−2.799	−0.102	2.256
	(3.134)	(2.774)	(2.910)
% Non-Citizen	6.655**	3.720	4.801*
	(3.124)	(2.804)	(2.818)
Borders Mexico	2.932***		
	(0.809)		
% Non-Citizen X borders Mexico	−14.995***		
	(5.023)		
NH-CA-VA		−0.042	
		(0.364)	
% Non-Citizen X NH-CA-VA		−5.605***	
		(2.112)	
Battleground state			−0.223
			(0.365)
% Non-Citizen X battleground state			−4.170***
			(1.590)
Observations	3111	3111	3111
R^2	0.896	0.896	0.896
Adjusted R^2	0.894	0.894	0.894

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

Table 8
Regression analyses of Clinton-Obama differences in turnout, with interaction variables.

	Dependent variable:		
	Clinton - Obama turnout		
	(1)	(2)	(3)
% Unemployed	4.759	12.000***	4.271
	(4.102)	(4.006)	(4.028)
Log median household income	4.082***	4.088***	4.058***
	(0.354)	(0.341)	(0.348)
% Employed in manufacturing	−3.310***	−3.250***	−3.425***
	(1.147)	(1.122)	(1.146)
% Urban	0.003	0.003	0.003
	(0.003)	(0.003)	(0.003)
% Male	11.524***	11.662***	11.388***
	(3.791)	(3.703)	(3.784)
% White	34.475***	30.496***	32.028***
	(10.301)	(9.868)	(10.134)
% Black	16.481	12.428	14.190
	(10.384)	(9.990)	(10.257)
% Hispanic	46.038***	33.323***	41.222***
	(11.628)	(10.649)	(10.905)
% Asian	11.528	23.912**	9.772
	(10.938)	(10.538)	(10.859)
% No college degree	24.026**	18.859*	21.292
	(11.737)	(11.208)	(11.518)
% White, no college degree	−41.317***	−36.368***	−38.542***
	(11.899)	(11.398)	(11.704)
% Black, no college degree	−21.920	−17.481	−19.293
	(11.939)	(11.474)	(11.781)
% Hispanic, no college degree	−47.379***	−30.935**	−42.157***
	(13.364)	(12.296)	(12.569)
% Asian, no college degree	−1.935	−14.094	−0.612
	(13.219)	(12.790)	(13.063)
% Mormon	6.423***	6.040	6.499
	(1.755)	(1.718)	(1.753)
% Evangelical Christian	3.539	3.632	3.569
	(0.608)	(0.595)	(0.607)
% Jewish	−6.872*	−7.626**	−7.440**
	(3.656)	(3.564)	(3.710)
% Muslim	17.808***	16.852***	17.862***
	(5.116)	(5.008)	(5.113)
% Foreign born citizen	−3.229	−1.525	−1.449
	(3.649)	(3.151)	(3.375)
% Non-Citizen	15.257***	11.781***	13.859***
	(3.624)	(3.128)	(3.199)
Borders Mexico	−0.509		
	(0.875)		
% Non-Citizen X borders Mexico	−0.725		
	(5.465)		
NH-CA-VA		−0.958**	
		(0.425)	
% Non-Citizen X NH-CA-VA		−26.634***	
		(2.368)	
Battleground state			−4.608***
			(0.444)
% Non-Citizen X battleground state			−0.613
			(1.914)
Observations	3111	3111	3111
R^2	0.729	0.740	0.729
Adjusted R^2	0.723	0.734	0.722

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

Table 9
Regression analyses of Trump-Romney differences in turnout, with interaction variables.

	Dependent variable:		
	Trump - Romey turnout		
	(1)	(2)	(3)
% Unemployed	16.285***	16.550***	15.732***
	(3.647)	(3.643)	(3.570)
Log median household income	1.516***	1.506***	1.558***
	(0.315)	(0.310)	(0.309)
% Employed in manufacturing	−3.916***	−4.132***	−4.260***
	(1.020)	(1.021)	(1.015)
% Urban	−0.013***	−0.013***	−0.012***
	(0.002)	(0.002)	(0.002)
% Male	−5.188	−5.718*	−4.723
	(3.371)	(3.367)	(3.354)
% White	−28.489***	−31.325***	−36.942***
	(9.158)	(8.973)	(8.982)
% Black	−19.632**	−22.304**	−27.900***
	(9.232)	(9.083)	(9.091)
% Hispanic	−16.301	−23.047**	−28.263***
	(10.338)	(9.682)	(9.665)
% Asian	0.161	0.673	−10.406
	(9.725)	(9.582)	(9.625)
% No college degree	−11.438	−14.459	−21.202**
	(10.435)	(10.190)	(10.208)
% White, no college degree	38.792***	41.991***	48.458***
	(10.579)	(10.363)	(10.373)
% Black, no college degree	15.917	18.918*	25.363**
	(10.615)	(10.433)	(10.442)
% Hispanic, no college degree	16.984	24.423**	29.857***
	(11.881)	(11.180)	(11.140)
% Asian, no college degree	6.364	5.425	13.131
	(11.753)	(11.629)	(11.578)
% Mormon	−21.600***	−21.444***	−21.463***
	(1.560)	(1.562)	(1.554)
% Evangelical Christian	−5.124***	−5.076***	−5.102***
	(0.540)	(0.541)	(0.538)
% Jewish	12.318***	11.877***	8.197**
	(3.251)	(3.241)	(3.288)
% Muslim	−2.774	−2.847	−3.554
	(4.548)	(4.553)	(4.531)
% Foreign born citizen	−0.883	0.863	6.051**
	(3.244)	(2.865)	(2.992)
% Non-Citizen	0.498	−1.470	0.070
	(3.222)	(2.845)	(2.835)
Borders Mexico	−1.436		
	(0.778)		
% Non-Citizen X borders Mexico	1.932		
	(4.859)		
NH-CA-VA		−2.133***	
		(0.386)	
% Non-Citizen X NH-CA-VA		−2.403	
		(2.153)	
Battleground state			−3.162***
			(0.394)
% Non-Citizen X battleground state			−9.730***
			(1.696)
Observations	3111	3111	3111
R^2	0.803	0.802	0.804
Adjusted R^2	0.798	0.797	0.800

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

Table 10
Regression analyses of vote share and turnout differences, 2008 baseline.

	Dependent variable:	
	Difference in vote share	Difference in turnout
	(1)	(2)
% Unemployed	−26.245***	−8.997*
	(4.250)	(5.384)
Log median household income	1.267	5.731
	(0.349)	(0.445)
% Employed in manufacturing	−0.932	−5.701
	(1.144)	(1.595)
% Urban	0.025	0.010
	(0.003)	(0.004)
% Male	−14.832	17.885
	(4.079)	(5.720)
% White	45.998	23.061*
	(10.511)	(11.982)
% Black	23.178**	5.631
	(10.639)	(12.082)
% Hispanic	39.073	41.723
	(11.326)	(12.955)
% Asian	4.206	−7.096
	(11.086)	(12.681)
% No college degree	33.586	20.990
	(11.986)	(13.503)
% White, no college degree	−70.975	−34.505**
	(12.183)	(13.769)
% Black, no college degree	−22.783*	−10.212
	(12.267)	(13.784)
% Hispanic, no college degree	−41.118	−43.379
	(13.086)	(14.838)
% Asian, no college degree	3.824	25.031
	(13.447)	(15.301)
% Mormon	5.985	12.816
	(1.881)	(3.274)
% Evangelical Christian	5.432	5.531
	(0.606)	(0.879)
% Jewish	−17.627	−13.830
	(3.611)	(4.122)
% Muslim	17.320	19.560
	(5.009)	(6.144)
% Foreign born citizen	1.907	−1.733
	(3.268)	(3.775)
% Non-citizen	4.941	19.160
	(3.300)	(3.866)
Observations	3111	3111
R^2	0.897	0.742
Adjusted R^2	0.894	0.736

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

Table 11
Regression analyses of vote share and turnout differences, 2008–2012 average baseline.

	Dependent variable:	
	Difference in vote share	Difference in turnout
	(1)	(2)
% Unemployed	−21.458***	−7.619
	(3.715)	(4.781)
Log median household income	1.604***	5.032***
	(0.305)	(0.395)
% Employed in manufacturing	−0.403	−5.350***
	(1.000)	(1.417)
% Urban	0.018***	0.007**
	(0.002)	(0.003)
% Male	−13.108***	17.051***
	(3.566)	(5.080)
% White	42.424***	25.097**
	(9.189)	(10.641)
% Black	17.574*	7.070
	(9.301)	(10.730)
% Hispanic	29.433***	39.186***
	(9.902)	(11.506)
% Asian	0.639	0.567
	(9.692)	(11.263)
% No college degree	23.548**	19.149
	(10.479)	(11.992)
% White, no college degree	−62.093***	−34.397***
	(10.651)	(12.228)
% Black, no college degree	−15.651	−11.317
	(10.725)	(12.242)
% Hispanic, no college degree	−29.378**	−39.994***
	(11.441)	(13.178)
% Asian, no college degree	3.894	13.823
	(11.756)	(13.589)
% Mormon	7.058***	11.667***
	(1.644)	(2.907)
% Evangelical Christian	5.749***	5.251***
	(0.530)	(0.781)
% Jewish	−13.155***	−10.773***
	(3.157)	(3.661)
% Muslim	13.863***	18.724***
	(4.379)	(5.457)
% Foreign born citizen	0.931	−1.747
	(2.857)	(3.353)
% Non-citizen	4.476	14.868***
	(2.885)	(3.433)
Observations	3111	3111
R^2	0.904	0.743
Adjusted R^2	0.902	0.737

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Intercepts and state fixed effects not displayed.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002

Data sources

The follow list describes data sources.

- 2010 United States decennial census. Data: Urbanization by county. Census data are available from https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml (accessed June 12, 2017).
- The 2010 U.S. Religion Census: Religious Congregations & Membership Study. Data: Religious affiliation by county.
- American Community Survey, 2010–2014, 5-year estimates. Data: Citizenship, gender, education, manufacturing, foreign born status, and racial breakdowns by county. This survey is available at https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2014/5-year.html (accessed June 12, 2017).
- The Associated Press, election services. Data: 2016 election returns by county. The service we used is described at https://www.ap.org/en-us/topics/politics/calling-election-winners (accessed June 12, 2017).
- Centers for Disease Control and Prevention. Data: Deaths by county. Death totals are summed by county deaths from 1999 though 2014 and then divided by the citizen voting age population. Data can be found at https://wonder.cdc.gov/ucd-icd10.html (accessed March 13, 2017).
- Dave Leip's Atlas of U.S. Presidential Elections. Data: 2012 election returns by county and 2016 voter registration figures. The website for the Atlas is http://uselectionatlas.org (accessed June 12, 2017).
- Local Area Unemployment Statistics, Bureau of Labor Statistics. Data: Unemployment data by county for 2015. These data can be found at https://www.bls.gov/lau (accessed June 12, 2017).
- Maryland State Board of Elections. Data: Maryland registration statistics for 2016. The website for these statistics is http://www.elections.state.md.us/pdf/vrar/2016_11.pdf (accessed June 14, 2017).
- New Hampshire Secretary of State, Elections Division. Data: New Hampshire election results and other election figures at the town level. The website for this office is http://sos.nh.gov/Elections.aspx (accessed February 28, 2017).
- Small Area Income and Poverty Estimates, United States Census Bureau. Data: household income by county. Data can be found at https://www.census.gov/did/www/saipe/index.html (accessed June 12, 2017).

References

Abrajano, Marisa, Alvarez, R. Michael, 2010. New Faces, New Voices: the Hispanic Electorate in America. Princeton University Press, Princeton, NJ.

Ansolabehere, Stephen, Persily, Nathaniel, 2008. Vote fraud in the eye of the beholder: the role of public opinion in the challenge to voter identification requirements. Harv. Law Rev. 121 (7), 1737–1774.

Ansolabehere, Stephen, Luks, Samantha, Schaffner, Brian F., 2015. The perils of cherry picking low frequency events in large sample surveys. Elect. Stud. 40, 409–410.

Bartels, Larry M., Zaller, John, 2001. Presidential vote models: a recount. PS Polit. Sci. Polit. 34 (1), 9–20.

Beber, Bernd, Scacco, Alexandra, Alvarez, R Michael, 2012. What the numbers say: a digit-based test for election fraud. Polit. Anal. 211–234.

Bentele, Keith G., O'Brien, Erin E., 2013. Jim Crow 2.0? Why states consider and adopt restrictive voter access policies. Perspect. Polit. 11 (4), 1088–1116.

Brians, Craig Leonard, Grofman, Bernard, 2001. Election day registration's effect on US voter turnout. Soc. Sci. Q. 82 (1), 170–183.

Cantú, Francisco, 2014. Identifying irregularities in Mexican local elections. Am. J. Polit. Sci. 58 (4), 936–951.

Chen, Jowei, September 6, 2017. The impact of political geography on Wisconsin redistricting: an analysis of wisconsins Act 43 assembly districting plan. Election Law J. 16 (3).

Cho, Wendy. K. Tam, Gimpel, James G., 2009. Presidential voting and the local variability of economic hardship. Forum 7. Article 1.

Christensen, Ray, Schultz, Thomas J., 2014. Identifying election fraud using orphan and low propensity voters. Am. Polit. Res. 42 (2), 311–337.

Elklit, J., Reynolds, A., 2002. The impact of election administration on the legitimacy of emerging democracies: a new comparative politics research agenda. Commonw. Comp. Polit. 40 (2), 86–119.

Fair, Ray C., 2002. Predicting Presidential Elections and Other Things. Stanford University Press, Stanford, CA.

Goel, Sharad, Meredith, Marc, Morse, Michael, Rothschild, David, 2016. One person, one vote: estimating the prevalence of double voting in U.S. Presidential elections. Working paper. Available at: https://www.dropbox.com/s/bbzgpeo1rh1s7dy/OnePersonOneVote.pdf?dl=0.

Hasen, Richard L., 2012. The Voting Wars: from Florida 2000 to the Next Election Meltdown. Yale University Press, New Haven, CT.

Herron, Michael C., Mebane Jr., Walter R., Wand, Jonathan N., 2008. Voting technology and the 2008 New Hampshire primary. William Mary Bill Rights J. 17 (2), 351–374.

Hicks, William D., McKee, Seth C., Sellers, Mitchell D., Smith, Daniel A., 2014. A principle or a strategy? Voter identification laws and partisan competition in the American States. Polit. Res. Q. 68 (1), 18–33.

Hillygus, D. Sunshine, Shields, Todd G., 2005. Moral issues and voter decision making in the 2004 presidential election. PS Polit. Sci. Polit. 38 (4), 201–209.

Katz, Richard S., 1997. Democracy and Elections. Oxford University Press on Demand.

Kramer, Gerald H., 1983. The ecological fallacy revisited: aggregate-versus individual-level findings on economics and elections, and sociotropic voting. Am. Polit. Sci. Rev. 77 (01), 92–111.

Levitt, Justin, 2007. The Truth about Voter Fraud. Report published by the Brennan Center for Justice at New York University School of Law.

Lipset, Seymour Martin, 1959. Some Social Requisites of Democracy: Economic Development and Political Legitimacy. Am. Polit. Sci. Rev. 53 (1), 69–105.

Magaloni, Beatriz, 2010. The game of electoral fraud and the ousting of authoritarian rule. Am. J. Polit. Sci. 54 (3), 751–765.

Manza, Jeff, Uggen, Christopher, 2006. Locked Out: Felon Disenfranchisement and American Democracy. Oxford University Press, New York.

Minnite, Lorraine C., 2010. The Myth of Voter Fraud. Cornell University, Ithaca, NY.

Norris, Pippa, 2014. Why Electoral Integrity Matters. Cambridge University Press.

Richman, Jesse T., Chattha, Gulshan A., Earnest, David C., 2014. Do non-citizens vote in US elections? Elect. Stud. 36, 149–157.

Stewart III, Charles, Ansolabehere, Stephen, Persily, Nathaniel, 2016. Revisiting public opinion on voter identication and voter fraud in an era of increasing partisan polarization. Stanf. Law Rev. 68 (6), 1455–1489.

Stout, Christopher T. (Ed.), 2015. Bringing Race Back in: Black Politicians, Deracialization, and Voting Behavior in the Age of Obama. The University of Virginia Press, Charlottesville, VA.

Tate, Katherine, 1991. Black political participation in the 1984 and 1988 presidential elections. Am. Polit. Sci. Rev. 85 (4), 1159–1176.

Wand, Jonathan N., Shotts, Kenneth W., Sekhon, Jasjeet S., Mebane Jr., Walter R., Herron, Michael C., Brady, Henry E., 2001. The butterfly did it: the aberrant vote for Buchanan in Palm Beach County, Florida. Am. Polit. Sci. Rev. 95 (4), 793–810.

Please cite this article in press as: Cottrell, D., et al., An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election, Electoral Studies (2017), https://doi.org/10.1016/j.electstud.2017.09.002