RECONSIDERING LOST VOTES BY MAIL

Charles Stewart III

Department of Political Science, Massachusetts Institute of Technology

cstewart@mit.edu

July 25, 2020

Abstract

A "lost vote" occurs when a voter does all that is asked of her, and yet her vote is uncounted in the final tally. Estimating the magnitude of lost votes in American presidential elections has followed the work of the Caltech/MIT Voting Technology Project, which initially estimated the magnitude of lost votes in the 2000 presidential election, due to failures of voter registration, polling-place management, and voting technologies, to be between 4 and 6 million out of 107 million cast that year. Because of data and conceptual limitations, lost vote estimates have tended to focus on in-person voting, ignoring lost votes due to mail voting. This paper revisits the one previous effort to estimate lost votes, by considering data available from the 2016 presidential election. Conceptually, the paper highlights how differing mail-ballot legal regimes produce lost mail votes in different ways, and at different rates, on account of differing laws, regulations, and practices. Empirically, the paper draws on administrative records and surveys to provide an estimated number of lost mail votes in 2016. That estimate works out to approximately 1.4 million votes in 2016—4.0% of mail ballots cast and 1.0% of all ballots. These estimates are relevant in light of efforts to expand mail balloting in the 2020 presidential election. States that will see the greatest growth in mail ballots tended to have higher lost vote rates than those with vote-by-mail systems. This implies that a doubling or tripling of the number of mail ballots in 2020 will result in a disproportionate growth in the number of lost votes due to mail ballots.

RECONSIDERING LOST VOTES BY MAIL

Charles Stewart III

Department of Political Science, Massachusetts Institute of Technology

cstewart@mit.edu

July 25, 2020

The COVID-19 pandemic has brought renewed attention to voting by mail in the United States. The pandemic reached critical proportions in early March 2020, right as the presidential primary season was beginning. Starting in mid-March, states began delaying their spring primaries to give election officials more time to plan responses to the pandemic. In many cases, this response included taking affirmative action to encourage voters to cast their ballots by mail.

Encouraging greater use of mail balloting is attractive from a public health perspective, because it promises to "de-densify" in-person polling places. However, the increased attention to mail balloting has set off a political backlash on some fronts and a less strident "not so fast" on others. The political backlash is related to President Donald R. Trump's tweets, charging that mail balloting promises that the 2020 election will be "the most RIGGED Election in our nations [sic] history" in 2020,[1] resulting in the "END OF OUR GREAT REPUBLICAN PARTY [sic]."[2] This backlash is also evident in public opinion polls, which place overwhelming majorities of Democrats in favor of expanded vote by mail in 2020, while robust Republican majorities oppose.[3] A less strident opposition against the mail-ballot push has come from some citizen and

---

[1] Donald R. Trump, Twitter post, Jun 22, 2020 08:45:23 AM, http://twitter.com/realdonaldtrump.
[2] Donald R. Trump, Twitter post, May 28, 2020 08:00:39 PM, http://twitter.com/realdonaldtrump.
[3] Quinnipiac Poll, "Biden Ahead Of Trump By 8 Points In Presidential Race, Quinnipiac University National Poll Finds; Nearly 6 In 10 Support Vote By Mail In November," June 18, 2020, https://poll.qu.edu/national/release-detail?ReleaseID=3664.

civil rights groups, who claim that mail ballots risk disfranchising minority voters (Garrison 2020.)

Despite the political ruckus that has been kicked up by President Trump's aggressive opposition to expanded vote-by-mail options, it appears clear that the number of mail ballots cast in the November 2020 election will vastly exceed past levels. This has certainly been true during the presidential primary season. In an examination of the election records of sixteen states that held 2020 presidential primaries after April 1, the fraction of ballots cast by mail averaged 80 percent of turnout, compared to 12 percent among the same states in 2016.[4]

These historic mail-ballot usage levels may not be repeated in November, because in many states, they are the result of one-time emergency measures implemented only for the primaries. Yet, even if only half of Americans vote by mail in November 2020, the shift will result in an unprecedented volume of mail ballots. This shift will have significant consequences for the management of the election, which is done by local governments that are struggling from the fiscal and governing challenges produced by the pandemic. A voting system that is managed around an expectation that 5% of ballots will arrive by mail is both qualitatively and quantitatively different from one in which the mail-ballot volume is greater by an order of magnitude. Estimates of what it would cost to shift all states to a 100% vote-by-mail (VBM) system come up with price tags of at least one billion dollars (Brennan Center 2020). (For context, the direct costs for running the presidential election nationwide are probably in the $2 billion range.)

---

[4] These data were gathered under the direction of the author and are available upon request. The states in question are Wisconsin, Ohio, Nebraska, Idaho, Indiana, Iowa, Maryland, Montana, New Mexico, Pennsylvania, Rhode Island, South Dakota, Georgia, Nevada, North Dakota, and the District of Columbia. They were chosen using the following criteria: Presidential primary after April 1; presidential primary in both 2016 and 2020; not a vote-by-mail state under normal circumstances; and reliable data available for both 2016 and 2020, preferably from administrative data.

Evidence has accumulated during the unfolding presidential primary system, in which widespread reports have documented deficiencies such as undelivered and uncounted mail ballots, underscores the fact that a VBM system at any scale requires attention to a multitude of details that are likely to need years to perfect.[5]

Nonetheless, the United States finds itself facing an expansion of mail balloting for the 2020 election. Despite the clear public health imperative that mail balloting be increased in the 2020 primary season and general election, the expansion of mail balloting comes with risks. Those risks relate to the probability that a voter who decides to cast his or her ballot by mail will have that ballot uncounted once the votes have been tallied. To be sure, these risks are small, and should not be sensationalized. But, they are greater than zero; they should be the part of any risk/reward calculus that informs policy decisions about how much effort should be placed into encouraging greater use of voting by mail, compared to devoting the same effort to making in-person polling places available and safe.

This paper examines the risks of voting by mail through the lens of "lost votes," focusing mainly on the 2016 presidential election as a case study. The concept of lost votes was originally developed by the Caltech/MIT Voting Technology Project (VTP) following the 2000 election, to help quantify the degree to which deficiencies in election administration—ranging from inaccurate voter registration files to malfunctioning voting machines—deprived voters of their voice.

The lost-votes concept was developed with in-person voting in mind, and can be explained through the following thought experiment: Suppose a voter awoke on Election Day,

---

[5] An insightful report on the challenges of a rapid pivot to widespread mail balloting after a tradition of discouraging absentee balloting was produced by Wisconsin Election Commission staff following its April 7 primary. See Wisconsin Elections Commission (2020).

intending to vote for her candidate of choice. If she attempts the actions necessary to effect that choice—goes to the polling place, authenticates her identity, marks her ballot, and deposits it for tabulation—how likely is it that a vote for her candidate will be registered in the official tally? Administrative errors at each of the stages just delineated that are not the fault of the voter could result in her vote being "lost." For instance, her name could incorrectly be missing from the registration roll, or a poorly calibrated tabulator could fail to register her mark on the ballot. After an analysis of data from the 2000 election, the VTP concluded that between 4 and 6 million votes had been lost, out of 107 million cast (Caltech/MIT Voting Technology Project 2001; United States Elections Project 2011).

The VTP's report omitted an estimate of lost votes that resulted from mail ballots. At the time, there was virtually no nationally available data to even attempt an answer. In the next few years, the data landscape improved, primarily through the inauguration of the Election Administration and Voting Survey (EAVS) by the U.S. Election Assistance Commission, an agency that was created by the 2003 Help America Vote Act. This allowed Stewart (2010) to analyze the question of votes lost through mail balloting, by relying on administrative data contained in the EAVS, supplemented with national surveys. That study concluded that a significant fraction of mail ballots, as many as 21% of all requests, were "lost" in 2008.[6] In numbers, that was 7.6 million votes, an eye-catching sum amidst an electorate of 132.6 million in 2008 (United States Elections Project 2012).

---

[6] Ardent vote-by-mail opponents have recently discovered this article, touting the 21% "lost votes" number, as if this was intended as an estimate of the number of ballots that literally were lost in the mail, rather than as an estimate of lost votes as defined here. Authors who have claimed that Stewart (2010) found that 21% of mail ballots in 2008 were "lost," i.e., were mislaid in postal facilities and never found, have either not read the article, or have willfully misrepresented it.

It is clear that this estimate needs to be revised and reconsidered, both for data and conceptual reasons. On the data side, the EAVS has continued to improve since 2008. Much of Stewart (2010) was given over to imputation of missing data. That is no longer necessary. Conceptually, as the use of mail balloting has increased in the U.S., it has proceeded along different administrative tracks that require us to consider carefully what various administrative statistics actually imply if we are to take the underlying idea of lost votes seriously.

To illustrate this latter point, consider two states, A and B, which each mail out one million ballots to voters and receive back 900,000 for counting. State A is a "vote-at-home state," and sends ballots by mail to every registered voter.[7] State B is a "no-excuse absentee ballot state," and only mails ballots to voters who request them.

Both states have a 10% non-return rate of mailed ballots, but the implications of this statistic from a lost-votes perspective are different for each. In State B, the fact that 100,000 voters had taken the initiative to ask for a ballot and had not returned it clearly raises the possibility that these are lost votes as previously defined, although the mechanisms still need to be specified. How many ballots had never been received by the voters? How many had been received by the voter too close to Election Day to return it on time? How many voters changed their minds and decided to vote in person instead?

The mechanisms in State A include the ones suggested in State B, but there is a third that is likely dominant: abstention. A percentage of registered voters fail to turn out to vote in every state in every election. In states that do not mail ballots to all voters, their failure to vote is not considered a lost vote. The fact that such a voter might have a ballot mailed to him by the state and left sitting on the kitchen table does not make it a lost vote. If abstention are the primary

---

[7] In 2016, these states were Colorado, Oregon, and Washington. In 2020, California, Hawaii, and Utah have been added to the list.

driver of the non-return of mail ballots in VBM states, then the non-return rate is uninformative about lost votes in those states.

The rest of this paper proceeds as follows. First, I lay out the administrative context of voting by mail in the United States. I then turn my attention to measurement issues related to estimating the number of lost votes due to mail balloting. I conclude by estimating that in 2016, 1.6 million votes were lost due in the mail-ballot process, and discussing how the expansion of mail balloting in 2020 will expand that number, both in terms of raw votes, and as a percentage of mail ballots cast.

<p align="center">The Growth of Mail Balloting</p>

American election administration has seen an expansion of "convenience voting" over the past quarter century. As Figure 1 illustrates, nearly 90% of votes were cast in-person on Election Day in 1996, nearly 3% were cast early in-person, and almost 8% were cast by mail. By the 2016 presidential election, the share of votes cast on Election Day had fallen to 60%, with roughly equal proportions voting early in-person and by mail.[8]

<p align="center">[Figure 1 about here]</p>

This trend toward greater convenience voting has not been equally distributed across the nation. Over the past twenty years, states have taken three general policy paths that have determined how much voting occurs before Election Day, and how those votes are distributed between in-person early voting and mail voting. The first path is to maintain the status quo, so that the lion's share of voting still occurs on Election Day. Among states that have abandoned the status quo, two additional paths are possible. The first has encouraged greater voting in-

---

[8] The saw tooth pattern for in-person early voting in Figure 1 is likely due to decisions by national presidential campaigns to emphasize mobilizing their supporters via early voting more than state candidates in non-presidential general elections.

person before Election Day (usually called "early voting" or "early in-person voting"). The second has encouraged greater use of mail ballots. (Florida is alone among the states in pursuing both paths with roughly equal zeal.)

As a result of policy choices, the distribution of voters across the three major voting modes—Election Day, early in-person, and mail/absentee—varies across the states. This is illustrated in Figure 2, which uses a ternary diagram to describe how voters distributed themselves across these modes in 1996 and twenty years later, in 2016. For each of these diagrams, states that appear at the very top of the triangle saw almost all their votes cast on Election Day. As states migrate down and to the left, their voters have abandoned Election Day voting in favor of mail balloting. As they migrate down and to the right, they have abandoned Election-Day voting in favor of in-person early voting.

[Figure 2 about here]

Consistent with the overall national trend, in 1996 the vast majority of states had traditional Election-Day voting systems, with a small percentage voting absentee—thus the preponderance of states that hug the left-hand leg of the triangle toward the top. There were some exceptions. Notably, Oregon and Washington were already moving toward the VBM systems they have today; Arizona and California were encouraging vote-by-mail through the use of permanent absentee lists. On the other hand, Tennessee and Texas were beginning to experiment with the creation of vote centers, at which voters could cast ballots in person, either before or on Election Day.

Twenty years later, there was much greater variability across states. Roughly a dozen still maintained laws that made it difficult to vote before Election Day. This means that over the past two decades, the vast majority of states have actively expanded the options voters have to

cast ballots. In 2016, a majority of voters from seven states cast their ballots by mail (all states from the west); in another seven states, a majority cast their ballots in-person before Election Day (all states from the Sunbelt). The remaining states expanded voting options through a mix of mail and early voting, but Election-Day voting still predominated.

Focusing on mail balloting, states can be categorized into four types which can be arrayed from most- to least-restrictive.[9] Most restrictive are states that allow mail balloting only if the voter presents an excuse that is identified by law. The two most common excuses presented by voters are being out of town on Election Day or being disabled and unable to go to a polling place. Less restrictive are states that require voters to apply for a mail ballot, but do not require an excuse to accompany the application. Less restrictive still are states that maintain a permanent absentee ballot list, allowing voters to automatically receive a mail ballot in the future even without applying for it. The least restrictive states maintain vote-by-mail systems, wherein all voters—usually those on the "active" registry[10]—receive a ballot by mail by virtue of being registered. Table 1 lists states by mail-ballot regime in 2016, and notes those states that have changed their regime, either permanently or temporarily, for the 2020 election.

[Table 1 about here]

Each of these categories is crude. The policy details chosen by states can further influence the ability of voters to actually vote by mail. Consider North Carolina, which is a no-

---

[9] This classification of states draws on the work of the National Conference of State Legislatures. See NCSL, Voting Outside the Polling Place: Absentee, All-Mail and other Voting at Home Options, https://www.ncsl.org/research/elections-and-campaigns/absentee-and-early-voting.aspx, 10 July 2020.

[10] The 1994 Help America Vote Act encouraged states to distinguish between two types of registrants on their voter rolls. New registrants are identified as "active." As a general matter, the voter's status remains active unless the voter fails to vote in two federal elections, at which time a confirmation process attempts to verify that the voter is still alive and living at his/her address on file. If the confirmation process is unsuccessful, the voter can be moved to "inactive" status, which is one step away from removing the voter from the rolls. Inactive voters usually cannot be removed from the voter rolls until they have failed to vote—or otherwise communicate with the local elections office—for two additional federal elections.

excuse absentee ballot state. It requires absentee voters to fill out their ballot in the presence of either a notary or two witnesses, who must also sign the ballot.[11] No doubt because of this hurdle, and the fact that North Carolina makes it relatively easy to vote early in person, only about 3 percent of North Carolina voters have cast ballots by mail in recent federal elections.

Despite the policy details that lie below these gross mail-ballot policy categories, the choice of mail-ballot policy regime does influence the fraction of voters who choose to vote by mail. This is illustrated through a simple regression reported in Table 2, where the dependent variable is the fraction of voters who reported casting a ballot by mail in the 2016 election, and the independent variables are two sets of dummy variables. The first set describes the policy regime of the state in 2016. ("Excuse required" is the omitted category.) The second records whether the state required notary or witness verification of returned ballots. The constant records the average mail-ballot usage in 2016 among the excuse-required states, or 8.0%. Voters in no-excuse absentee states voted by mail at a 12.6% rate (.080 + 0.046), whereas half the voters in states with a permanent absentee list voted by mail, and 87% of voters in vote-by-mail states voted remotely.[12] Adding a requirement that the returned mail ballot be notarized or verified by a witness reduced the usage of mail ballots by 4.6 percentage points—virtually the same amount that mail-ballot usage is increased when states drop the excuse requirement for absentee ballots.[13]

---

[11] This requirement was reduced back to one witness for the 2020 election. In addition, the state legislature passed a law providing for other ways in which absentee balloting became more accessible to voters. See Associated Press, North Carolina Senate passes bill to make mail-in absentee voting easier during fall 2020 election, https://www.wxii12.com/article/north-carolina-election-senate-bill-mail-in-absentee-voting/32830528, 11 June 2020.

[12] The percentage of voters in vote-by-mail states is less than 100 because all vote-by-mail states allow voters to cast ballots in person, if they wish. Colorado is the most expansive in this regard, opening in-person vote centers throughout each of each counties in the weeks before Election Day, and on Election Day itself.

[13] The analysis is Table 2 is a simple cross-sectional regression, and thus no causal claims can be made. However, in a pooled time series analysis from 1996 to 2016, where the observations are the state-year, the results relating to mail-ballot regime reported in Table 2 are substantively unchanged. Unfortunately, I do not have a time series of

[Table 2 about here]

The fact that policy choice influences the usage of mail ballots so strongly has implications for how we understand the expansion of mail balloting over the past two decades, and ultimately how we frame the question of lost votes via mail balloting. Although the relative usage of mail ballots grew nearly three-fold from 1996 to 2016, the lion's share of that growth has been in the six states that were either vote-by-mail in 2016 or had permanent absentee lists. In 1996, these six states accounted for only half (51%) of mail ballots cast in the U.S. By 2016, they accounted for nearly two-thirds (64%). In contrast, the remaining states that continued to require some sort of absentee ballot application to vote by mail—whether or not an excuse was required—actually saw their contribution to mail balloting decline, from 48% of all mail ballots in 1996 to 27% in 2016.

Thus, the consequences of mail balloting for lost votes prior to 2020 have been largely regional, concentrated increasingly on the western states that have loosened restrictions on mail balloting the most. Lost votes by mail have much greater consequences for the conduct of elections in these states than those in the rest of the country. The great variability in how mail-ballot regimes are implemented requires us to be mindful of those differences as we tote up how many lost votes there might be, and where they are.

Available Data

On what basis can we measure lost votes in the context of mail balloting? To begin answering this question, we must first distinguish between data that would be necessary in an ideal research design and the data that are actually available to do the estimation.

---

policies relating to witness and notary requirements, which precludes me from casting the Table 2 analysis in a more causal-inference-friendly framework, such as fixed effects regression.

The calculation of lost votes in mail elections focuses on summing up six quantities:

1. Requests for mail ballots that are not received by the election authorities.

2. Absentee ballot requests that are unfulfilled by election authorities.

3. Absentee ballots transmitted by election authorities that are not received by the voter.

4. Absentee ballots returned by the voter but not received by the election authorities.

5. Returned absentee ballots that are rejected by election authorities.

6. Tabulated mail ballots that fail to record the choice(s) made by the voter, i.e., residual votes.

In Table 3, I have associated each of these sources of lost votes with ideal data we would like to have in order to quantify their magnitudes, contrasting each with the actual, available data. For the remainder of this section, I discuss the challenges of obtaining the ideal data and the nature of the actual data.

[Table 3 about here]

Ideal Data

I will start with the data that would most directly help calculate these quantities. Elections are managed by local governments in the United States—counties in most of the country, municipalities in New England and much of the upper Midwest. Therefore, this discussion will primarily center on data that might be produced and reported by local election offices.

Votes lost by the postal service. The first, third, and fourth quantities of interest—requests for mail ballots not received by the election authorities, absentee ballots transmitted by election authorities at are not received by the voter, and absentee ballots returned by the voter but not received by election authorities—all involve the performance of the U.S. Postal Service. The

ideal data to estimate these quantities would be based on administrative audits that gauge the performance of the USPS as it delivered election mail.

The USPS is already engaged in an extensive performance review of its delivery record (USPS 2020a). The quantification of lost mail votes because of the non-performance of the Postal Service would ideally be built off the methodology behind these studies. Barring that, Postal Service statistics that measure the rate of postal non-delivery could be used as a proxy for problems potential mail voters face because of poor postal service performance.

The current performance assessment system combines sampling of ingested mail pieces with information gathered by the postal service's mail-handling equipment. Using this information, it is possible to estimate the time it takes different classes of mail to travel from sender to receiver. The current system, which was instituted in 2019, replaced an earlier one, which essentially relied on sending test items from randomly selected locations to other locations, and then calculating service times based on the average travel time of these items from point-to-point.

Most ballots are sent from election offices to voters using marketing ("bulk") mail. The current USPS standard is to deliver marketing mail within a period of 3-to-18 days, depending on the distance between the origin and destination of the piece. The USPS (2020a) reports that in FY 2016, 97.4% of marketing mail arrived within one day of the standard, 98.8% within two days, and 99.3% within three. For a mail ballot sent locally, this means that 99.3% would arrive within six days. There are no statistics about mail that never arrived.

Mail ballots are returned as single-piece first-class mail, as are applications for absentee ballots. Current service standards call for local mail to be delivered within 2 days and mail to other locations within 5 days. Because this paper is most concerned about the risks faced by

voters who vote by mail but could plausibly vote in person, I focus here on the 2-day standard. In FY 2016, 98.5% of "2-day mail" arrived within three days, 99.3% within four days, and 99.6% within five days. Again, there are no statistics about mail that never arrived.

If we assume that marketing and first-class mail that does not arrive within three days of the service standard never arrives, then a voter who mails in a request for an absentee ballot has a 98.5% probability of eventually getting a mail ballot back to the election office, assuming a total allowance of 16 business days for all the mail to complete their voyages.

Through the use of the intelligent barcode program, it would certainly be possible for a local election jurisdiction to track whether an absentee ballot sent by the election office reached its intended target, and whether it arrived within postal service standards. State and local election offices have begun using the intelligent barcode system to track absentee ballots, even providing web-based systems that allow voters to track ballots themselves.[14] However, at the present time, most jurisdictions do not formally track the journey of mail ballots from the main office, to the voter, and back. Those that do, do not publish performance statistics based on these systems.

One can imagine a program of tracking all mail ballots in their round-trip journey. The same cannot be easily said of the first step of the process, requesting a ballot, given the fact that requests for mail ballots originate from outside the election office. Short of surveilling all registered voters, it is hard to imagine using administrative data to determine how many voters requested absentee ballots in the first place.[15]

---

[14] The two best-known systems are Ballot Trace, developed by Denver Elections Department, and Ballot Scout, developed by Democracy Works. See Visram (2020) for a description of these systems.

[15] A further complication is that absentee ballots are increasingly being requested through electronic means, such as Internet portals. In 2020, at least a dozen states appear to have statewide portals to request mail ballots. (It is unclear how many had them in 2016, but it was undoubtedly many fewer.) Even in states without a centralized portal, some local jurisdictions run request websites of their own. In these cases, an auditing process of mail-ballot requests that focused solely on requests that are mailed in would miss electronic requests. Finally, websites have emerged that claim to facilitate the mail-ballot application process for all states, whether or not they have a

Votes lost by the election office. The second quantity in Table 3, the number of mail-ballot requests that are unfulfilled by election authorities, could be measured in one of two ways. The simplest, though least precise, would be agnostic about why the request was unfulfilled—for instance, because the application gave the wrong address for the voter, or because the request got lost in the paper shuffle—and record the number of requests received and the number honored. The second would be more granular, asking jurisdictions to record why requested ballots applications were denied, counting as lost votes only those cases that were not the fault of the voter.

The U.S. Election Assistance Commission's Election Administration and Voting Survey, which will be described below, does not ask local jurisdictions how many requests they received for mail ballots, although it does ask for the number of mail ballots transmitted. Therefore, the discrepancy between received requests and resulting transmitted ballots cannot be measured on a national basis.

From my knowledge of administrative records in state and local election administration, few jurisdictions disseminate the number of requests. The most notable exception is North Carolina, whose Board of Elections releases a detailed data file documenting the fate of each absentee ballot request. In the file for 2016, 1,791 absentees are recorded as having been requested but not sent out, out of a total of 137,234 mail requests. Further investigation shows that all but 402 of these were part of a chain of requests involving a series of spoiled ballots, or ballots rejected for technical administrative reasons.

---

centralized portal. One such organization is vote.org. Although such organizations can, and do, track whether the voter successfully completed the ballot request, they do not share that data publicly.

If all of these 402 unaccounted-for requests do in fact represent applications "lost in the office," then it suggests that in North Carolina, at least, lost votes due to this reason is around 0.3% of requests. Of course, further investigation of these 402 cases could very well reveal that they merely *appear* not to have been sent out, or that a legally legitimate reason prohibited them from being sent. On the other hand, requests that were received but not logged into the ballot tracking system would not show up at all.

The case of North Carolina underscores the general lack of solid data to quantify the number of absentee ballot requests that die in local election department offices due to clerical errors. The number is clearly "small," but how small is anyone's guess. As with the question of votes lost by the postal service, it is clear that alternative methods are needed to estimate the number of votes lost because of clerical errors.

Rejection of returned ballots. The fifth quantity in Table 3, the number of ballots that are returned for counting, but are then rejected, is the item where the ideal data source comes the closest to what is actually available. The desired data in this case would account for the number of voted ballots received by local election offices and the number accepted by the office for counting.

Although some states disseminate a precise accounting to the public of the status of returned absentee ballots to the public—North Carolina, for instance, is the best example—most do not. However, the U.S. Election Assistance Commission's Election Administration and Voting Survey (EAVS) requests local election jurisdictions to report the number of mail ballots returned by voters for counting, and the number that were subsequently counted.

The difference between these two quantities can be used as a proxy for rejected mail ballots, with one important caveat, which pertains to late-arriving ballots. All states have a

deadline for the receipt of mail ballots. For most, it is on Election Day; for the rest, the deadline ranges up to two weeks after Election Day, so long as the ballot is postmarked by Election Day.[16] There is some disagreement between the academic and election administration communities about whether these late-arriving ballots are properly termed "rejected." That distinction need not detain us here, since regardless of the terminology used, a returned ballot that is uncounted, for whatever reason, can be considered a lost vote.[17]

Lost votes at tabulation. The sixth and final quantity is the number of tabulated ballots that fail to record the choice made by the voters, because of equipment failure or voter confusion. This problem was at the core of the 2000 recount controversy in Florida that led to the court case *Bush v. Gore.* The two most infamous ballot defects—"hanging chad" and the "butterfly ballot," were instances of mismarking and the failure of tabulation equipment. With hanging chad, which interfered with the tabulation of punch-card ballots, voters could believe they had marked their choice for a candidate correctly, only to have the vote not counted because the chad was still attached to the voting card, blocking the electro-mechanical sensors that counted ballots. In the case of the butterfly ballot, not only did the ballot design encourage voters intending to vote for presidential candidate Al Gore to mark the ballot for Patrick Buchanan instead, it also produced a tendency for voters to double-vote, thus causing an overvote (Agresti and Presnell 2003; Herron and Sekhon 2003; Mebane 2004).

---

[16] National Conference of State Legislatures, "VOPP: Table 11: Receipt and Postmark Deadlines for Absentee Ballots," https://www.ncsl.org/research/elections-and-campaigns/vopp-table-11-receipt-and-postmark-deadlines-for-absentee-ballots.aspx, 15 June 2020.

[17] An important exception can be made for those few ballots that are rejected because it is clear that the voter did not certify the return envelope. An example would be a family member voting the ballot, or the ballot being intercepted fraudulently and the voter's signature forged. Databases that track prosecutions for voter fraud have found miniscule numbers of prosecutions for this type of forgery. See Heritage Foundation, A Sampling of Recent Election Fraud Cases from Across the United States, https://www.heritage.org/voterfraud.

The 2003 Help America Vote Act contained provisions that were intended to warn voters that their ballots may have contained inadvertent errors. In addition to banning punch card ballots in federal elections—and thus eliminating hanging chad from the election lexicon—HAVA also required ballot scanners to provide feedback to voters when their ballots were under- and overvoted.

Research since 2000 has shown that these provisions in HAVA reduced lost votes due to tabulation and voter errors, thus saving millions of votes from being lost (Ansolabehere and Stewart 2005; Stewart 2006; Stewart 2011). However, when votes are cast by mail, there are no fail-safe features that warn voters if they have inadvertently over- or undervoted. The empirical question for us is how to quantify an excess of over- and undervotes (collectively termed "residual votes") caused by voters casting mail, rather than an in-person, ballots.

Because of the secret ballot, it is impossible to measure directly the number of ballots that may have been *inadvertently* under- or overvoted, nor is it possible to directly distinguish an inadvertent undervote from a conscious abstention (Stewart et al 2020). Therefore, residual vote rates that are attributable to voting technologies or voting modes have generally been estimated using statistical techniques, taking advantage of "natural experiments," such as the replacement of one voting technology type with another (Ansolabehere and Stewart 2005; Stewart 2006, 2014) or the diffusion of absentee voting to replace in-person voting (Alvarez, Beckett, and Stewart 2013).

The best data to use in conducting such analysis are election returns. Because most states fail to distinguish between over- and undervotes in their election return reports, the residual vote rate is generally measured indirectly, by subtracting the number of votes cast and recorded in an election—typically at the "top of the ballot," such as for president—from the number of voters

who cast a ballot in the election, i.e., turnout. The limiting factor here is turnout data, which most states now report at least at the county level.

The subject of this paper is lost votes <u>by mail</u>, so there is an added requirement here. In order to estimate the excess in residual votes in ballots cast by mail versus ballots cast in person most directly, it is necessary to separate out the total number of ballots cast by mail (whether or not they contained votes that were counted) from the total number of ballots cast in person. Very few jurisdictions do this, however. As a consequence, even under the best of circumstances, calculating the excess residual vote rate among mail ballots will be incomplete, given current data-reporting practices.

*   *   *

The second column of Table 3 identifies the ideal data sources that one would want to rely on in an investigation of the number of lost votes. As this discussion suggests, with the exception of the rejection rate of returned mail ballots, and perhaps residual votes, there is currently no practical way to conduct a direct lost-votes analysis of mail ballots that gives due attention to each major way in which a vote could be lost. Therefore, for the most part, we must consider alternative measurement strategies that would allow us to gain a sense of the magnitude of the mail-ballot lost-votes problem.

 Actual Data

As a consequence of data limitations, we must rely on existing sources to fill in the gaps. These sources are not as precisely focused as we would prefer, nor do they always measure the quantities of interest directly. The data sources that are available include public opinion surveys that ask voters about their experience with the election administration process. These include the Survey of the Performance of American Elections (SPAE) and certain modules of the

Cooperative Election Study (CES, formerly the Cooperative Congressional Election Study), both of which ask voters who attempted to vote by mail whether they received the requested ballot. Another survey, the Voter and Registration Supplement of the Current Population Survey (CPS), which is conducted by the U.S. Census Bureau immediately following each federal election, asks non-voters why they failed to vote, with one response category being "Registration problems (i.e. didn't receive absentee ballot, not registered in current location)."

Despite being the only data currently available on a national basis to address the front-end of the mail-ballot process, these surveys have acknowledged shortcomings. The first is that the survey items, at best, elide problems related to the ballot request not being received, the ballot request not being fulfilled, and the mail ballot not being mailed to the voter. The second is that the relevant question in the CPS combines two reasons for nonvoting, by giving as examples of registration problems both "didn't receive absentee ballot" and "not registered in current location." The CES and SPAE battery of reasons for not voting separates out not receiving a mail ballot from registration problems. But, other question wording quirks in these surveys, also limit their usefulness.

The fourth source of lost votes, the failure of a returned ballot to reach the election authority, has no obvious second-best measurement strategy. Lacking administrative audits of marked ballots put into the mail-stream that do not reach the local election office, and lacking universal use of systems that allow voters to track their ballots back to the local election office, this is a major data hole.

As already discussed, the fifth source of lost votes by mail, the non-counting of mail ballots that are received for counting, is in the best shape. The EAC's EAVS survey records how

many ballots were returned for counting in each jurisdiction of the U.S. and, of these, how many were accepted for counting.

Finally, the last source of lost votes by mail, the excess of residual votes compared to in-person voting, cannot be measured directly, because most states do not allow us to calculate residual vote rates by voting mode. Therefore, we must rely on indirect methods in most cases. These methods include performing the preferred calculations for the states that do allow a calculation of residual vote rates by mode. In addition, borrowing from Alvarez, Beckett, and Stewart (2013), it may be possible to estimate a nationwide rate of residual votes via mail ballots in a fixed-effects regression framework, leveraging off the within-state variation in mail-ballot usage over time.

Tracing "Lost" Votes by Mail in 2016

In this section, we finally reach the estimation of lost votes by mail in 2016. Because the quantity that is best measured with existing data is the number of ballots received for counting that are left uncounted, we start there. I then turn my attention to the other sources of lost votes.

The non-counting of mail ballots.

We start with ballots that are left uncounted after they have been returned. The data source here is the EAVS.[18] The two variables that are needed here have very little missing data—0.13% in the case of the number of ballots returned and 0.36% for the number of ballots counted.[19] In the few cases where data were missing, we imputed missing values, so that we could get a more

---

[18] The particular items are C1b (domestic mail ballots returned for counting) and C4a (ballots counted).
[19] These percentages are calculated by weighting each local jurisdiction by the number of registered voters. Only three states had any missing data for these two variables.

complete view of lost votes due to non-counting.[20] However, this imputation does not change any of the substantive results.

Overall, I estimate that of 33.4 million mail ballots returned for counting in 2016, 377,000 were left uncounted. This produces a nationwide rate of uncounted ballots of 1.14%. (This rate was 0.33% as a percentage of all votes cast in the election.[21])   (See Table 4.)

[Table 4 about here]

As discussed above, states can be arrayed along a continuum of mail-ballot access, from states that allow only absentee voting for a limited number of reasons to states that affirmatively mail ballots to all registered voters.[22] Because the laws and regulations that define these mail-ballot regimes reflect long-standing attitudes states about access to the ballot box, we should expect that rejection rates should be correlated with the level of access afforded to voters to mail ballots.

The data in Table 4 affirm this expectation for the most part. Among states that allowed mail balloting only for a limited set of reasons, nearly two percent (1.80%) of returned absentee ballots were left uncounted. At the other end of the spectrum, among the three vote-by-mail states in 2016, less than one percent (0.92%) of returned ballots were uncounted. Voters in states

---

[20] Because the amount of missing data was very small, we employed a very simple imputation method to "plus-up" the number of returned and counted ballots. If a state had missing data on one of these variables, we simply divided the total number of ballots accounted for by the non-missing data by the fraction of counties from which the data were not missing. (Counties were weighted by the relative fraction of registered votes in that county, as a percentage of the entire states.) For instance, counties accounting for 90.4% of Arkansas's registered voters reported how many ballots were returned for counting, summing to 27,525 ballots. To estimate the total number of returned ballots, we divided 27,525 by (1-.904), resulting in an imputed estimate of 30,433.

[21] Turnout at the state level and nationwide is taken from the United States Election Project, www.electionproject.org.

[22] To be precise, ballots tend to be mailed to all registered voters on the active registry. "Inactive voters," who have been tagged for eventual removal from the rolls, due to evidence that they may have moved or died, are still considered to be registered voters, although vote-by-mail states generally do not automatically mail ballots to them.

with permanent absentee lists were actually a little less likely to see their ballots uncounted (0.79%), whereas voters in no-excuse absentee states saw an overall no-counting rate of 1.33%.

Of course, there is another way to think about the relative magnitude of these no-counting rates. For instance, although 1.80% of mail voters in excuse-required states saw their mail ballots go uncounted, only 10% of voters cast ballots by mail in these states. Therefore, the uncounted mail ballots in the excuse-only states only accounted for 0.18% of <u>all</u> ballots cast. On the other hand, 97% of voters in the vote-by-mail states returned their ballots by mail. Thus, the uncounted mail ballots in these states accounted for 0.89% of all ballots cast. Therefore, the issue of uncounted mail ballots in vote-by-mail states has greater overall consequences for election outcomes in vote-by-mail states, even though the uncounted rates are relatively low, compared to the excuse-required and no-excuse states.

This analysis is purely correlational, and does not address what is likely to happen to the rate of uncounted votes if a state experiences an exogenous shock in the number of mail ballots without fundamentally changing the laws, regulations, and practices governing mail ballots, such as is likely to happen in the November 2020 election. Anecdotal reports from the primaries suggest that these non-counting rates might be quite high.

For instance, nearly ten percent of mail ballots returned in New Jersey's local elections in May 2020—which were conducted entirely by mail—were uncounted (O'Dea 2020). In the 2016 presidential election, New Jersey, which at the time was a "no-excuse" state, had a mail-ballot non-counted rate of 3.1%, the ninth-highest in the nation.

Experiences such as New Jersey's suggest that the rapid expansion of mail ballots due to exigencies such as the COVID-19 pandemic will produce higher non-counting rates among states that have tended not to facilitate the use of mail ballots. This effect, plus the simple increase in

the number of mail ballots, will make the failure to count mail ballots more consequential for election outcomes in states throughout the nation.

The non-return of mail ballots by voters

As discussed in the previous section, in an ideal case, we would have systematic and comprehensive measures about the two major processes that precede the decision by a local election office to accept a mail ballot for counting: (1) whether the absentee ballot transmitted by election authorities was received by the voter and (2) whether the returned absentee ballot was received by the election office.

Left unsaid before, but clearly relevant to this chain of events, is the intermediate step of the voter marking the ballot and returning it for counting. In other words, it is possible that a ballot mailed by the local election office was not received back for counting because the voter decided not to return the ballot—either to abstain or to vote in person instead. In recent analysis of the non-return of mail ballots in the 2020 presidential primaries, approximately 10% of voters who did not return a requested mail ballot subsequently turned out to vote in person (Read 2020; Curiel and Dagonel 2020).

Leaving aside for the moment this issue of voters deciding not to return a mail ballot after requesting one, and focusing on postal issues alone, the lack of systematic administrative audits of the mail-ballot mail stream makes the ideal measurement strategy infeasible. The closest we can come to addressing the round-trip voyage of mail ballots is to analyze the number of ballots transmitted by local election offices to voters and the number of those ballots returned for counting. Both of these quantities are reported in the EAVS.

As reported in Table 5, 41.7 million mail ballots were transmitted to voters in 2016;[23] 33.5 million were returned for counting, leaving 8.2 million unreturned. This is a 19.7% nationwide non-return rate. However, as the data in Table 5 also suggest, a state's mail-ballot legal regime has a strong influence on the non-return rate. Non-return rates were 6.5% among excuse-required states and 13.8% among no-excuse states. Among states with a permanent absentee list or vote-by-mail laws, the non-return rates were 26.8% and 21.8%, respectively. Therefore, we need to examine non-return dynamics in states with different mail-ballot regimes more closely.

[Table 5 about here]

The non-return of mail ballots in no-excuse and excuse-required states. What do we make of these differing rates? Let us start with the states that require a contemporaneous[24] absentee ballot application, either with or without an excuse. Among these states, the non-return rates are a bit of a puzzle. Why were voters who requested absentee ballots in these states unable to complete the voting transaction? Did the ballots not make it to the voters? Were the returned ballots lost in the mail? Did the requesters have a change of heart?

The discussion in the previous section about performance reports released by the USPS suggests that the primary explanation for the non-return of mail ballots is not that the ballots got lost in the mail. Using USPS performance reports, we would expect for the non-return of ballots

---

[23] It is likely that some of these reported ballot transmissions are double-counted, because the EAVS survey item asks about the number of ballots transmitted, not the number of voters to whom ballots are transmitted. Some voters receive more than one ballot, mainly because they were lost, mis-addressed, or spoiled. How many is unknown. However, based on the absentee ballot request files for the 2016 election published by North Carolina, we can get a rough sense about how many this would be. In North Carolina, 231,782 individuals requested a mail ballot, 240,641ballots were transmitted, with 3.6% of requesters receiving more than one ballot.

[24] This adjective is not literally true, since many states now allow voters to sign up to receive ballots for several elections over a limited period of time, such as an election cycle or a year. This is different from the permanent absentee list, since the practice in the more traditional states resets the list of absentee ballot recipients to zero periodically.

that were due to the USPS itself would amount to approximately one percent of transmitted ballots.[25]

Evidence from North Carolina, which publishes a detailed public file that accounts for all mail-ballot requests, provides some insight into this issue. The North Carolina absentee ballot file from the 2016 election records 240,641 ballots transmitted and 42,606 (17.7%) that were unreturned.[26] These unreturned ballots had been sent to 40,934 individuals. The difference in the number of ballots transmitted and the number of individuals is due to the fact that some people had more than one ballot sent to them, because the original ballot was lost or spoiled.

Investigation of the North Carolina voter history file further reveals that of the 40,934 individuals who requested a mail ballot but did not return it, 19,880 (48.5%) ended up voting. Of these, 4,475 (22.5%) are recorded as having voted by mail in the voter file — a clear discrepancy between two datasets that should, by design, be consistent. Beyond this discrepancy, however, most of the non-returners who cast a ballot did so during early voting (11,954, or 60.1%), while a smaller number (3,451, or 17.4%) voted on Election Day.

Whether a non-returner eventually voted in person was curvilinear with respect to time. In particular, the largest number of ballots that were later not returned in North Carolina— 5,471—were sent out on September 9, the first date when general election mail ballots were

---

[25] This is calculated as follows. The probability of a ballot transmitted to a voter getting to the voter within three days of the marketing-mail service standard is 99.6%. The probability of a ballot mailed back to the election office arriving within three days of the first-class service standard is 99.3%. Therefore, the probability of a successful completion of this circuit is $0.996 \times 0.993 = 0.989$. The probability of this circuit not being completed is therefore $(1 - 0.989) = 0.011$.

   The report of the USPS's Inspector General that investigated postal service problems related to the conduct of the 2020 Wisconsin primary confirms the essence of this conclusion (USPS 2020b). That report documented that hundreds of mail ballots were undelivered in the primary. If our expectation is that ballots transmitted to voters would be successfully delivered 99.6% of the time, then of 1,282,097 ballots sent to voters in the primary, we would expect 5,128 not to be delivered (or at least delivered on time).

[26] If we examine this from the perspective of underlined requesters, rather than underlined requests, 15.2% of people who requested a mail ballot did not end up returning a mail ballot, even if they requested more than one.

distributed. The voting rate among these requesters, 39.9%, was significantly below that of the others. Because North Carolina gives absentee voters the option of requesting absentee ballots for a calendar year if they are ill or disabled, it is perhaps unsurprising that these requesters would eventually end up voting at a lower rate, since the intervening time between the initial request and the mailing of the ballots, the voters could have moved, died, or lost interest in voting. In addition, the voting rate of non-returners begins to decline sharply within two weeks of the general election. Given USPS performance standards, we can conclude that these requesters probably received their ballot too late to get it back to the voting authorities in time to count.

If this excursion into the North Carolina absentee ballot file is generalizable to other states that require voters to request ballots, it suggests that perhaps half of the unreturned mail ballots in presidential elections cannot be considered a "lost vote," even though they might be considered an inconvenience to the voter. If half of those who request a mail ballot and do not return it eventually vote in person, leaving the remainder as potential lost votes, then the upper bound of lost votes due to postal problems and administrative problems in election offices is half the first two percentages in Table 5, that is, 3.3% in excuse-required states and 6.9% in no-excuse states.[27]

The non-return of mail ballots in permanent absentee and vote-by-mail states. The non-return of mail ballots in states that have permanent absentee lists and vote-by-mail system are qualitatively different from unreturned ballots following a contemporaneous request. In permanent absentee ballot states, the penalty for a voter on the list who chooses to vote in person in a particular

---

[27] Left for future research is the question of why so many people appear to request an absentee ballot and then eventually fail to vote at all.

election is extracted in inconvenience. When the voter appears at an in-person polling location, the voter registry will note that the voter has a mail-ballot outstanding. If the voter does not surrender the unmarked ballot, she will typically be required to vote a provisional ballot, to guard against double voting. Filling out the provisional ballot form is a minor inconvenience the voter runs the risk that the provisional ballots could be rejected even if she had not already returned a mail ballot.

As reasonable as the explanation in the previous paragraph may be, there is no simple way to confirm it within the confines of this paper. For starters, the states that had a permanent absentee ballot list in 2016 all established their lists in the mid-2000s, before the availability of data from the EAVS to allow us to analyze this issue with any precision over time.[28] Furthermore, none of the states with permanent absentee lists in 2016 appear to make the archived list, along with the voter history file,[29] available to the public.

In addition, the EAVS does not gather information about the number of provisional ballots cast because a voter was on the permanent absentee list.[30] Detailed analysis using data from state voter files could in principle be performed, but that analysis will have to await future research. Until that analysis is done, the best we can do is note that California, which was the first state to establish a permanent absentee list, had a provisional ballot usage rate of 8.9% of all ballots cast in 2016, the highest in the nation, and over eight times the national average.

---

[28] California began its permanent absentee list in 1984, well before the EAVS project began. Both Arizona and Montana began between the 2004 and 2008 elections, which in theory could allow us to use the EAVS to analyze the relationship between unreturned mail ballots and cast provisional ballots. However, close analysis of the quality of the EAVS data conducted by the Pew Charitable Trusts suggests that data prior to 2008 is often unreliable.

[29] A voter history file is a database that records who voted in which elections and by which mode they voted. All states that are under the jurisdiction of the National Voter Registration Act must maintain some sort of voter history file. However, the conditions under which they make those files available to the public, including academic researchers, varies tremendously.

[30] Indeed, the EAVS does not gather any information about why provisional ballots are cast.

Finally, there are the vote-by-mail states. By sending a ballot to all registered voters, there ends up being an observational equivalence between a voter who failed to vote because of a problem with the postal service and a voter who just decided to sit out the election.

This is best illustrated by the experience of Colorado, which instituted vote-by-mail in 2013, having used a permanent absentee list for a number of years before that. Figure 3 shows the number of active registered voters (solid blue line), turnout (dashed blue line), ballots mailed (solid red line), and ballots returned (dashed red line). The implementation of vote-by-mail is evident by the surge in ballots mailed out between the federal elections of 2012 and 2014, from 2.1 million to 3.0 million. Prior to 2013, the number of unreturned ballots averaged 250,000 in the two presidential years (out of an average 1.9 million transmitted, or 13.3%) and 355,000 in the one off-year election (out of 1.6 transmitted, or 22.0%). These non-return rates are similar to other permanent-list states. After 2013, the number of ballots transmitted rose to over 3 million, despite only marginal changes in turnout rates across the next three elections. This caused the number of unreturned ballots to quadruple, to over 1 million, in the two midterm years and to double, to 756,000, in 2016.

[Figure 3 about here]

Thus, although some non-academic studies have attempted to tie the small turnout increases in Colorado in 2016 and 2018 to the institution of the vote-by-mail system, the primary effect of the system seems to have been to increase dramatically the number of ballots mailed to registered voters who have no interest in, or ability to, vote.[31]

---

[31] There is another complication to Figure 3: the institution of Election-Day Registration (EDR). The time series showing the number of registered voters in Colorado shows a distinct upward climb between 2012 and 2014, which coincides with the implementation of EDR. The expansion of the number of registered voters occurred at a faster pace than the expansion of turnout. Thus, a second-order effect of EDR in Colorado may have been to increase the number of non-returned ballots by increasing the number of voters abstaining.

The purpose of this excursion into Colorado's recent mail-balloting history is not to litigate the question of whether VBM increases turnout, but rather, to note that it is obvious that the lion's share of unreturned mail ballots in VBM states are due to the disinterest of voters in voting. If anything, it is likely that postal-service issues are less of a problem in VBM states, because of those states' higher level of engagement with local postal authorities, and owing to the fact that VBM states have pioneered in using ballot-tracking systems. There is no evidence that postal service on-time performance and delivery rates are lower in VBM states than elsewhere.[32]

Estimating non-delivery of requests using survey research. An alternative way to measure whether mail votes are lost because requested ballots are not delivered is through survey research. Since 2008, the Survey of the Performance of American Elections has asked a national sample of registered voters whether they voted in the most recent presidential election; if they answer, "no," they are asked about which of a number of reasons was a "major" or "minor" factor in the absence of their vote. One of those reasons is "I did not receive my ballot in the mail, or it arrived too late for me to vote." Another is "I requested, but did not receive, an absentee ballot." In 2016, 84.0% of non-voting respondents stated that the first reason was not a factor, 5.9% stated it was a minor factor, and 10.2% said it was a major factor. As to requesting but not receiving an absentee ballot, 80.1% said it was not a factor, 9.6% said it was a minor factor, and 10.2% said it was a major factor.

---

[32] USPS performance statistics by region show very little inter-region variation—certainly not enough to suggest that postal-service problems are a major factor distinguishing states according to their mail-ballot systems. See, for instance, USPS (2020).

There is a statistically significant correlation between those who said they did not receive a mail ballot and those who said they requested an absentee ballot that did not come.[33] Therefore, I combine answers to these two items and divide the sample into two groups: those for whom neither mail-related problem was a reason for non-voting and those for whom at least one of the problems was. Considering the sample as a whole, 8.0% of the reported non-voters stated that not receiving a mail ballot was a factor contributing to their non-voting. When I examine the relationship between the mail-ballot regime and the probability that a respondent chose one of these factors for not voting, the relationship fails to reach statistical significance.[34] Thus, from the perspective of voters, it appears that non-voting due to problems in receiving mail ballots is not associated with the type of law governing the use of mail ballots.

Summary. In trying to estimate the degree to which problems with the mail channel of transmitting absentee ballots contribute to lost votes, it is clear that administrative statistics can be misleading. The non-return rate of mail ballots transmitted is clearly inflated by abstention in VBM states and in states with permanent absentee ballot lists. Because there is no good evidence to suggest that the performance of the postal service is significantly better in one set of states than another, it seems reasonable to assign a uniform "postal service failure rate" across all states.

Lost votes due to voter confusion and machine malfunction

Finally, votes might be lost because of voter confusion or the failure of the tabulation device. As noted previously, voters who use mail ballots do not benefit from feedback required by the Help

---

[33] Goodman and Kruskal's gamma measuring the degree of association between responses to the two items is .91, $p < .018$.

[34] The chi-squared statistic for the contingency table describing the relationship between the mail-ballot regime and choosing one of these two responses as a factor for non-voting is 2.6, d.f. = 3, $p = .45$.

America Vote Act to voters who over- and undervote a ballot. If voters who cast ballots by mail are otherwise similar to those who cast ballots in person, then it is reasonable to expect that mail-ballot voters will experience higher residual vote rates than in-person voters.

By how much is difficult to pin down precisely. Estimation is empirically difficult because few states report the basic statistics necessary to calculate the residual vote rate separately for each voting mode. (These statistics would be the number of people who voted using each mode and the number of legal voted counted on each mode. The difference is the residual vote by mode.) The only state that has regularly and meticulously reported this data for a long time is Florida, which in 2001 mandated the Secretary of State to produce a report of the total number of over- and undervotes for the top-of-the-ballot races after each biennial election.[35]

Figure 4 displays this data for the 2004 – 2016 presidential elections. In each year the vote-by-mail mode has outpaced the other two in-person modes in overvotes, undervotes, and therefore, in residual votes. The average difference between the vote-by-mail residual vote rate and that of the combined rate of the two in-person modes is 0.36 percentage points. To put this in perspective, if we apply this difference to the 2.7 million mail ballots cast in Florida in the 2016 presidential election, this works out to nearly ten thousand votes—a small number in a sea of a total of 9.6 million votes cast in the state, but a large number when compared to the 537-vote margin by which George W. Bush beat Albert Gore in Florida in 2000.

[Figure 4 about here]

Other states are not so forthcoming with data. Therefore, efforts to estimate the additional residual vote rate due to voting by mail have taken the indirect route. Alvarez et al (2011) examined California counties from 1990 – 2010 in a fixed-effect regression framework to

---

[35] Section 101.595, Florida Statutes.

study the effect of changing voting technologies and the rise in vote-by-mail balloting on the residual vote rate. The point estimates from that study suggested that the additional residual vote rate of a vote-by-mail ballot had been 2.2 percentage points during this period—an order-of-magnitude greater than what is suggested by the Florida data just discussed (Alvarez et al 2011, Table 2). In another study analyzing state-level data, Stewart et al (2019, Table SM1) derived point estimates that suggested nationwide levels of additional residual votes through voting by mail of 0.6 percentage points in 2012 and 1.6 points in 2016.

None of these studies directly accounts for an important consideration in exploring the residual vote rate differences between mail and in-person voting modes: the voters who use the two modes are not randomly assigned. If demographic factors that are associated with higher residual vote rates are also correlated with voting by mail, then the finding of a mode effect on the residual vote rate would be due to non-random sample selection. There is some research suggesting that minority and low-education voters tend to contribute disproportionately to the residual vote count (See Buchler, Jarvis, and McNulty 2004). However, if anything, until recently, minority and low-education voters have been *less* likely to vote by mail. Therefore, it may be that the causal effect of switching from in-person to vote-by-mail balloting is actually *greater* than that reported here.

To conclude this subsection, Table 6 reports the residual vote rate, broken down by the mail-ballot regime in 2016. Consistent with Stewart et al (2020), states with the most liberal mail-ballot laws had the highest residual vote rates in 2016. States that required a contemporaneous application to in order to cast an absentee vote saw residual vote rates of 1.47% (excuse-required states) and 1.51% (no-excuse states). States with permanent absentee

lists had an average residual vote rate of 2.67%; the vote-by-mail states averaged a 3.42%

residual vote rate.[36]

[Table 6 about here]

Because Table 6 shows a one-year cross-sectional analysis, the differences cannot be

considered causal. However, the differences are cautionary, and open up the real possibility that

the added residual vote risk of voting by mail is in the ballpark of 1.5 %.

Toting Up Lost Votes in 2016

The ultimate goal of this paper has been to provide an updated estimate of the number of lost

votes by mail in the 2016 election, in an effort to understand the risks that are been taken on by

the expansion of vote-by-mail in 2020. In this section, I provide a series of estimates based on

the discussion in the previous two sections.

As those discussions should make clear, it is impossible to produce an estimate of lost

votes due to mail balloting with much precision. However, it is possible to provide back-of-the-

envelope estimates that can help orient an understanding of the risks involved in moving voters

who have been habituated to voting in person to voting by mail instead.

Table 7 provides an overview of this summary. The estimate starts with an assumption

that the probability of a ballot application not being received by the local election office and the

probability that the ballot returned by the voter for counting is equal to the percentage of first-

class mail that does not reach its destination within three days of the USPS service standard.

This percentage is 0.4%. Keeping with postal-service failures, this estimate also assumes that

the probability a ballot that is transmitted to a voter does not reach the voter is equal to the

---

[36] In an ANOVA analysis, the differences among these states were statistically significant using an F-test: $F = 10.89$, d.f. = 3, 43.

percentage of marketing-mail that does not reach its destination within three days of the service standard, or 0.7%.

The most speculative source of lost votes is a legitimate request not being fulfilled by the election office. In the interest of retaining this potential source of lost votes without over-weighting it, I have set it to 0.1%, in other words, assuming that 99.9% of good applications that reach the elections office actually result in a ballot going back to the voter.

Note that the first two sources of lost votes do not apply to VBM states, because no mail-ballot request is required. For that reason, the value of the first two sources is set to zero in VBM states. In addition, in recent years about 75% of voters in permanent-list states have also not had to request a ballot. Therefore, I have set the values of the first two sources of lost votes to 0.1% and 0%, respectively for these states.

The next quantity, the ballot-rejection rate, is taken from the average rates reported in the EAVS in 2016. The final quantity, the additional residual vote rate of mail ballots, is taken from a consolidation of the comparison reported in Table 6 and the previous statistical studies of the topic cited above. Here, I assume that someone casting a ballot by mail is 1.5 percentage points more likely to commit an over- or undervote than someone voting in person.

Summing all these percentages together, we have estimates ranging from 4.9% in excuse-required states to 3.5% in permanent-list and VBM states.

It is possible to construct a lower estimate, by reducing the source of lost mail ballots due to problems with the USPS even further. Recall that the estimate in Table 7 is based on the fraction of mail ballot pieces that do not arrive within three days of the service standard. If we are willing to believe that 99.9% of all mail reaches its destination eventually, then we can reduce these quantities down even further, to 0.1%. In that case, the lost-votes statistics become

3.0% in VBM states, 2.9% in permanent-list states, 3.5% in no-excuse states, and 4.0% in excuse-required states.

Finally, we can apply the lost-vote percentages to the number of mail ballots cast in each of the states to arrive at a nationwide lost-vote estimate on the mail-ballot side of the equation from 2016. That number is 1.4 million votes, which would be 4.0% of mail ballots attempted. It is also equal to 1.0% of all ballots cast in 2016.

This estimate is significantly less than the 21% figure suggested by Stewart (2010). One major difference between the two estimates is due to how non-returned ballots are treated in the previous study and the current one. In the 2010 study, transmitted mail ballots that were not returned from counting were treated as a "leak" in the mail-ballot pipeline, and added to the other "leaks" on an equal basis.[37] The 2010 study also took at face value, as a starting point, the claims of non-voters who blamed problems with receiving mail ballots for their not casting a ballot. The current study focuses on documented service statistics from the USPS to estimate the portion of requests and transmitted ballots that never are received. Using those statistics, the estimated lost-vote rate is much lower. Although both estimation approaches can be defended and criticized, upon reflection, the approach used in this paper seems closer to the reality of election administration as it is practiced in the United States.

These estimates are, therefore, more in proportion to the lost-vote rates first established for in-person voting by the VTP in 2001. Because of improvements in in-person voting, mostly wrought by HAVA, lost votes through in-person voting have fallen to the 2 – 3 million vote level. To this, we must add another million, due to mail ballots.

---

[37] Note that Stewart (2010, pp. 572 – 73) is hesitant to even call these "lost votes."

Epilogue: Applicability to 2020

The reason why this analysis might be of interest to a wider audience is that voters are being encouraged to vote by mail in 2020, rather than risk their health voting in person. If voters shift to voting by mail, what are the risks to their votes not being counted as a consequence? Of course, most voters who are being urged to shift to mail balloting are in states that have never experienced the volume of mail voting they will certainly encounter in November 2020. These administrative disruptions will inevitably increase the number of lost mail votes over the 2016 baseline, as is suggested by the response of states to increased demand for mail voting during the presidential primaries (Timm 2020). Nonetheless, there is value in knowing what the baseline is.

The estimates provided in this paper are a starting point for estimating the potential magnitude of the lost-votes problem in the 2020 election, assuming the fraction of votes cast by mail is much greater than it was in 2016. If we do nothing more than simply assume that the excuse-required and no-excuse states will see 60% of their ballots cast by mail in 2020, and we apply the lost-vote percentages in Table 7 to these inflated numbers of ballots, the estimate of lost votes soars from 1.4 million to 4.0 million. This would represent 2.9% of all ballots cast, compared to 1.0% in 2016.

Of course, with the rise of lost votes due to mail ballots, there will be a decline in lost votes due to in-person ballots. Even if the rise of mail-ballot lost votes outweighs the decline of in-person lost votes, the effect on the election outcome, per se, is likely to be relatively negligible. The more significant effect that the rise in lost votes due to mail ballots will produce will be to provide more fodder for legal disputes of the 2020 election, should it be close. Unlike in-person lost votes, most mail-ballot lost votes leave a paper record, in the form of a mail ballot that was uncounted. These uncounted mail ballots could become the focus of intense scrutiny,

should the presidential election turn on one or two states that are "too close to call." Thus, the greatest risk of an increase in mail ballots in 2020 may not befall individual voters as much as it affects any post-election controversy over whether the election was "rigged" or legitimate.

Citations

Agresti, Alan, and Brett Presnell. 2002. Misvotes, undervotes and overvotes: The 2000 presidential election in Florida. *Statistical Science* 17(4): 436 – 440.

Ansolabehere, Stephen, and Charles Stewart III. 2005. Residual votes attributable to technology. *Journal of Politics* 67(2): 365 – 389.

Brennan Center for Justice. 2020. *Estimated Costs of Covid-19 Election Resiliency Measures.* March 19, 2020 (updated April 18, 2020), https://www.brennancenter.org/our-work/research-reports/estimated-costs-covid-19-election-resiliency-measures.

Buchler, Justin, Andrew Jarvis, and John E. McNulty. 2004. Punch card technology and the racial gap in residual votes. *Perspectives on Politics* 2(3): 517 – 524.

Caltech/MIT Voting Technology Project. 2001. *Voting: What Is/What Could Be.* http://vote.caltech.edu/reports/1.

Curiel, John and Angelo Dagonel. 2020. *Wisconsin Election Analysis.* Report of the Stanford-MIT Healthy Elections Project. Forthcoming.

Garrison, Joey. 2020. Amid vote-by-mail push, civil rights groups say in-person voting still needed during pandemic. *USA Today* 20 April 2020, https://www.usatoday.com/story/news/politics/elections/2020/04/20/naacp-others-person-voting-still-needed-during-coronavirus-pandemic/5144452002/.

Herron, Michael C., and Jasjeet S. Sekhon. 2003. Overvoting and representation: An examination of overvoted presidential ballots in Broward and Miami-Dade Counties. *Electoral Studies* 22(1): 21 – 47.

Mebane, Walter R. 2004. The wrong man is president! Overvotes in the 2000 presidential election in Florida. *Perspectives on Politics* 2(3): 525 – 535.

O'Dea, Colleen. 2020. One in 10 ballots rejected in last month's vote-by-mail elections, 10 June 2020. *NJSpotlight,* https://www.njspotlight.com/2020/06/one-in-10-ballots-rejected-in-last-months-vote-by-mail-elections/

Read, Blair. 2020. *North Carolina Election Analysis.* Report of the Stanford-MIT Healthy Elections Project. Forthcoming.

Stewart, Charles, III. 2006. Residual vote in the 2004 election. *Election Law Journal* 5(2): 158 – 169.

Stewart, Charles, III. 2010. Losing votes by mail. *New York University Journal of Legislation and Public Policy* 2010(3): 57a – 602.

Stewart, Charles, III. 2011. Voting technologies. *Annual Review of Political Science* 14: 353 – 378.

Stewart, Charles, III, 2014. The performance of election machines and the decline of residual votes in the United States. In *The Measure of American Elections,* Barry C. Burden and Charles Stewart III, eds. New York: Cambridge.

Stewart, Charles, III, R. Michael Alvarez, Stephen S. Pettigrew, and Cameron Wimpy. 2020. Abstention, protest, and residual votes in the 2016 election." *Social Science Quarterly* 101(2): 925 – 939.

Timm, Jane C. 2020. States reject tens of thousands of mail ballots in this year's primaries, setting off alarm bells for November. *NBC News*, 18 July 2020, https://www.nbcnews.com/politics/2020-election/states-reject-tens-thousands-mail-ballots-year-s-primaries-setting-n1233833.

United States Elections Project. 2011. *2000 November General Election Turnout Rates,* http://www.electproject.org/2000g.

_____. 2012. *2008 November General Election Turnout Rates,* http://www.electproject.org/2008g.

United States Postal Service. 2020a. *Quarterly Performance for Single-Piece First-Class Mail, Service Variance, Mailpieces Delivered Between 01/01/2020 and 03/31/2020.* https://about.usps.com/what/performance/service-performance/fy2020-q2-single-piece-first-class-mail-service-variance.html.

United States Postal Service, Office of Inspector General. 2020b. *Management Alert: Timeliness of Ballot Mail in the Milwaukee Processing & Distribution Center Service Area.* 7 July 2020. https://www.uspsoig.gov/sites/default/files/document-library-files/2020/20-235-R20.pdf.

Visram, Talib. 2020. "'Track your ballot like a package': How technology will smooth the way for November's mail-in ballot surge." *Fast Company* 7 May 2020, https://www.fastcompany.com/90501588/track-your-ballot-like-a-package-how-technology-will-smooth-the-way-for-novembers-mail-in-ballot-surge.

Wisconsin Elections Commission. 2020. *April 7, 2020 Absentee Voting Report.* 15 May 2020, https://elections.wi.gov/sites/elections.wi.gov/files/2020-05/May%2020%2C%202020.Final_.pdf.

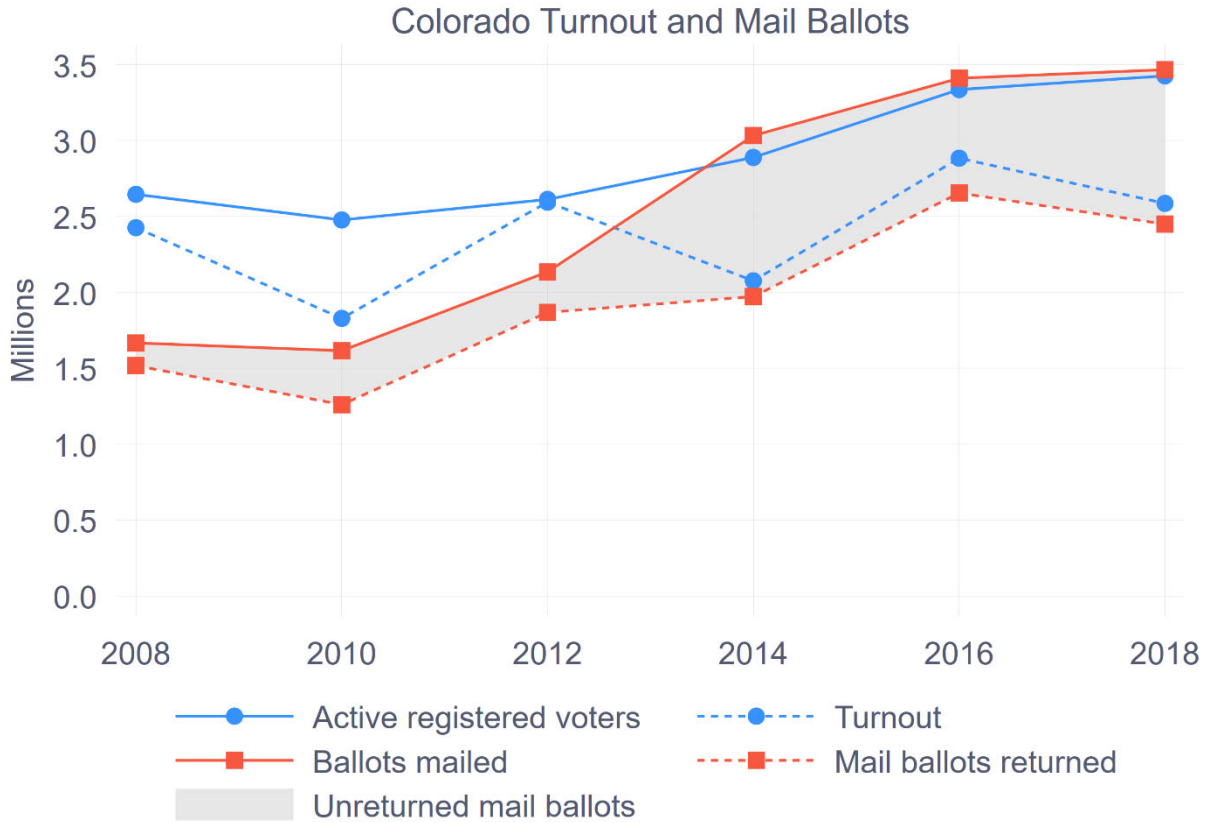Figure 1. Fraction of ballots cast by different voting modes, 1996 – 2018.
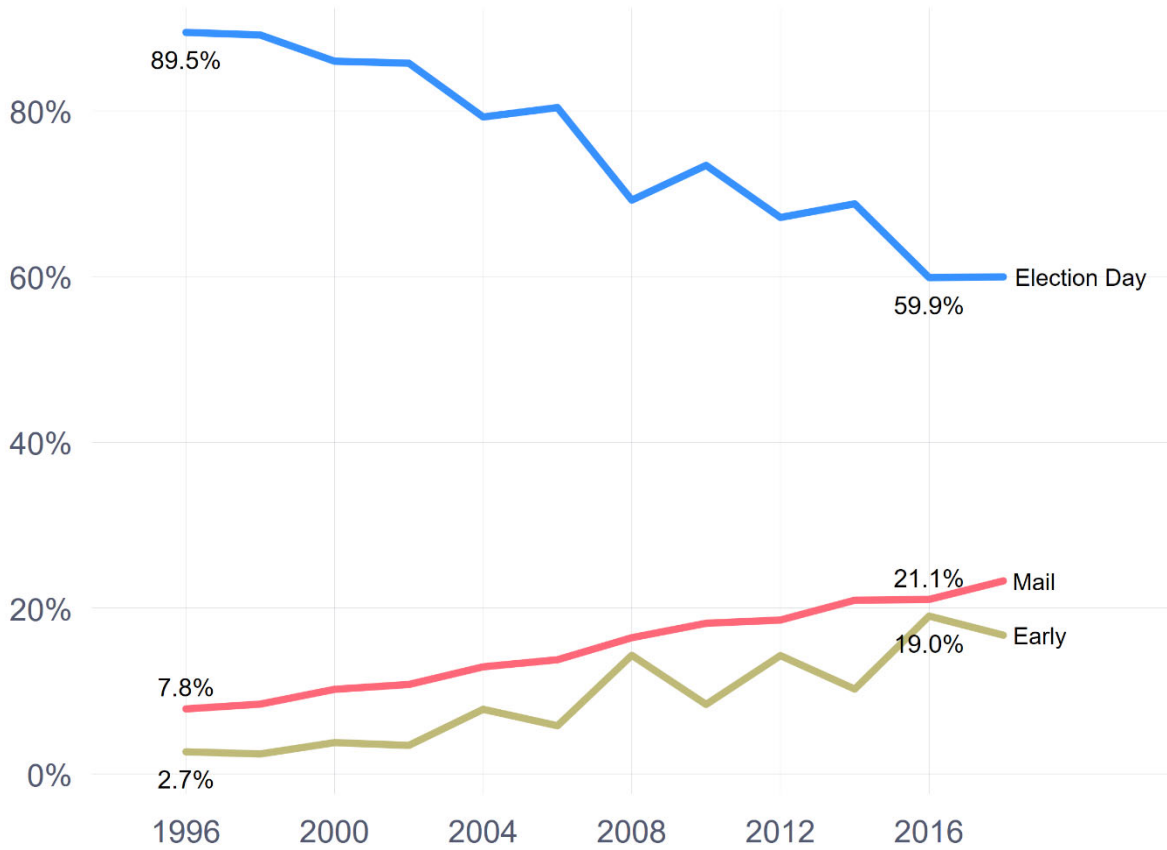
Data source: U.S. Census Bureau, Current Population Survey, Voting and Registration Supplement, various years.

Figure 2. Distribution of voting modes among voters in the fifty states, 1996 and 2016.
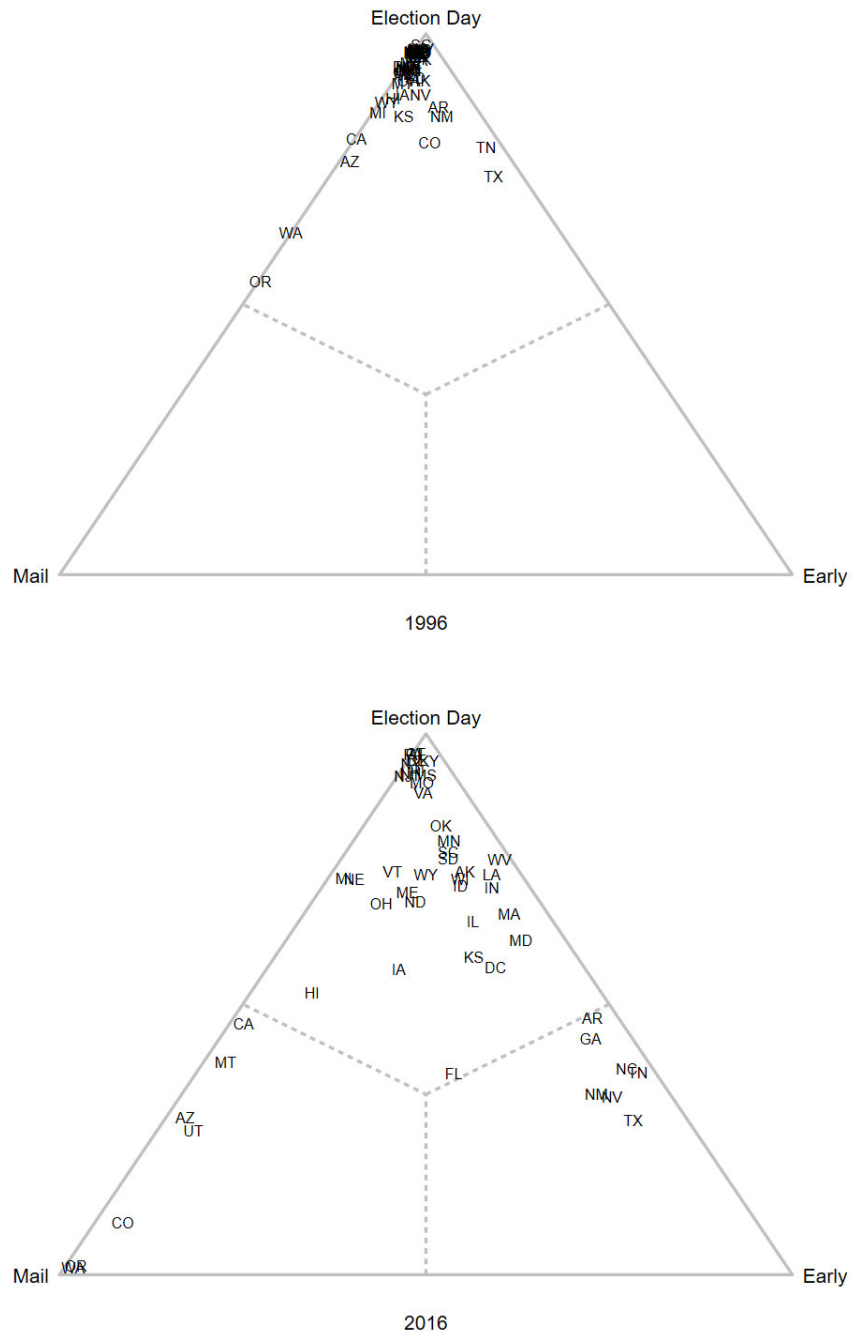
Data source: U.S. Census Bureau, Current Population Survey, Voting and Registration Supplement, 1996 and 2016.

Figure 3. Unreturned mail ballots in Colorado, 2008 – 2018.

Colorado Turnout and Mail Ballots

- Active registered voters
- Turnout
- Ballots mailed
- Mail ballots returned
- Unreturned mail ballots

Data source: U.S. Election Assistance Commission, Election Administration and Voting Survey, various years.

Figure 4. Residual votes in Florida, 2004 – 2016.
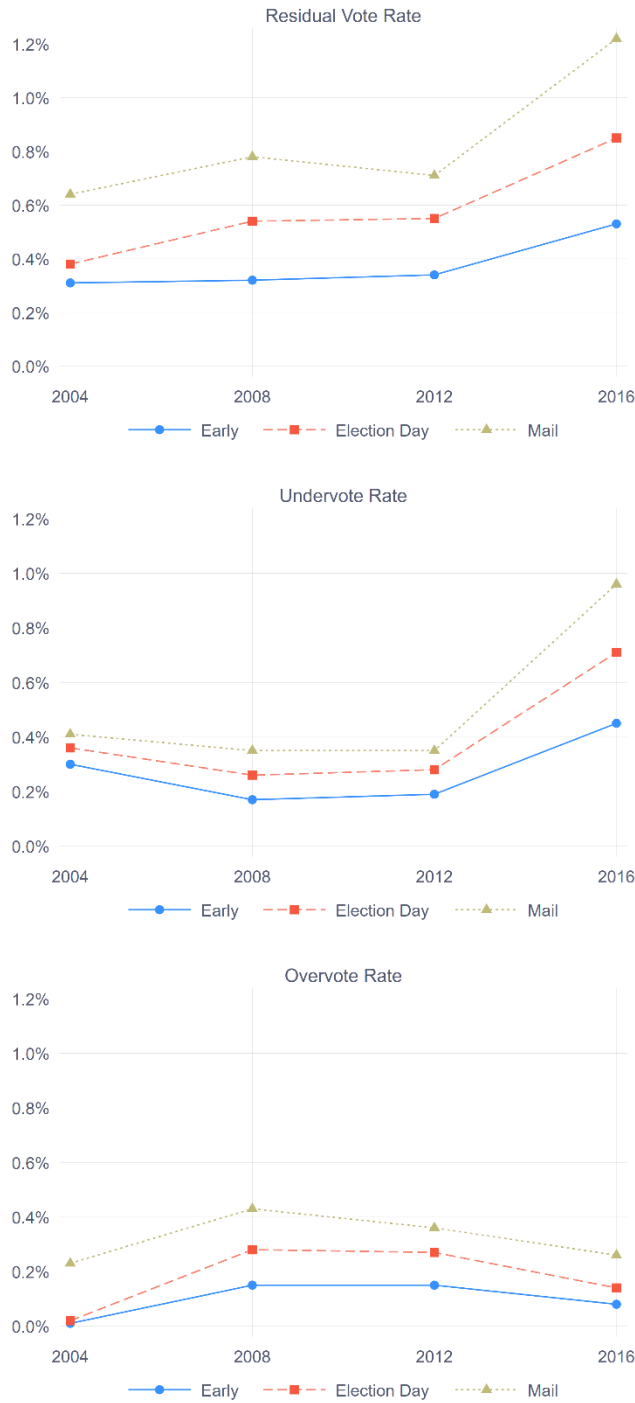
Data source: Florida Secretary of State.

Table 1. Vote-by-mail policy regime by state, 2016

Excuse required for absentee ballot	No excuse required for absentee ballot
• Alabama	• Alaska
• Arkansas	• Florida
• Connecticut	• Georgia
• Delaware	• Idaho
• Indiana	• Illinois
• Iowa	• Kansas
• Kentucky	• Maine
• Louisiana	• Maryland
• Massachusetts	• Minnesota
• Michigan[c]	• Nebraska
• Mississippi	• Nevada
• Missouri	• New Jersey
• New Hampshire	• New Mexico
• New York	• North Carolina
• Pennsylvania[c]	• North Dakota
• South Carolina	• Ohio
• Tennessee	• Oklahoma
• Texas	• Rhode Island
• Virginia	• South Dakota
• West Virginia	• Vermont
	• Wisconsin
	• Wyoming

Permanent absentee ballot list	Vote-by-mail
• Arizona	• Colorado
• California[a]	• Oregon
• District of Columbia	• Washington
• Hawaii[b]	
• Montana	
• Utah	

[a]California will mail ballots to all registered voters in the 2020 presidential election.
[b]Hawaii will be a vote-by-mail state in 2020.
[c]Michigan and Pennsylvania will be no-excuse absentee states in 2020.

Source: National Conference of State Legislatures, "Absentee and Early Voting," accessed through the Internet Archive, https://web.archive.org/web/20161108225513/https://www.ncsl.org/research/elections-and-campaigns/absentee-and-early-voting.aspx, 10 October 2016.

Table 2. Average percentage of voters who cast ballot by mail, 2016. (Ordinary least squares regression; robust standard errors.)

Variable	Coeff. (s.e.)
Mail ballot policy	
- Excuse required	Omitted category
- No excuse required	0.046
	(0.020)
- Permanent absentee list	0.46
	(0.06)
- Vote-by-mail	0.79
	(0.06)
ID or witness requirement	-0.046
	(0.016)
Constant	0.080
	(0.019)
N	50
R^2	.92
S.E.R.	0.075

Table 3. Causes of lost votes and sources of data.

Cause	Ideal sources	Available sources
1. Requests for mail ballots that are not received by the election authorities	Administrative audits	Public opinion surveys • Cooperative Congressional Election Study • Survey of the Performance of American Elections • CPS Voting and Registration Supplement
2. Absentee ballot requests that are unfulfilled by election authorities	Administrative records	
3. Absentee ballots transmitted by election authorities that are not received by the voter	Administrative audits	
4. Absentee ballots returned by the voter but not received by the election authorities	Administrative audits	None
5. Returned absentee ballots that are rejected by election authorities	Administrative records	Administrative surveys • Election Administration and Voting Survey
6. Tabulated mail ballots that fail to record the choices made by the voter, i.e., residual votes	Election returns	Election returns

Table 4. Lost votes due to uncounted domestic mail ballots, 2016.

Mail ballot legal regime	Ballots returned	Ballots Uncounted		
		Number	As pct. of ballots returned	As pct. of all ballots cast
Excuse required	6,070,962	109,309	1.80%	0.18%
No excuse required	7,521,342	99,866	1.33%	0.19%
Permanent absentee list	11,843,446	93,836	0.79%	0.48%
Vote-by-mail	8,039,705	73,842	0.92%	0.89%
Total	33,475,455	376,852	1.13%	0.33%

Source: U.S. Election Assistance Commission, Election Administration and Voting Survey, 2016.

Table 5. Unreturned mail ballots in the 2016 election

Mail ballot legal regime	Ballots mailed to voters	Unreturned Ballots		
		Number	As pct. of ballots mailed to voters	As pct. of all ballots cast
Excuse required	6,494,433	423,471	6.5%	0.72%
No excuse required	8,728,621	1,207,279	13.8%	2.34%
Permanent absentee list	16,184,284	4,340,838	26.8%	22.0%
Vote-by-mail	10,277,722	2,238,017	21.8%	27.0%
Total	41,684,050	8,209,605	19.7%	5.9%

Source: U.S. Election Assistance Commission, Election Administration and Voting Survey, 2016.

Table 6. Residual vote rate in the 2016 election

| | | Residual Votes | |
| | | | As pct. of |
Mail ballot legal regime	Turnout	Number	ballots cast
Excuse required	29,209,821	429,333	1.47%
No excuse required	63,796,464	961,134	1.51%
Permanent absentee list	18,300,443	448,441	2.67%
Vote-by-mail	8,274,104	283,334	3.42%
Total	119,580,862	2,162,242	1.81%

Note: The following states were excluded, because they did not report turnout levels: Mississippi, Missouri, Pennsylvania, and Texas.

Data source: Official state election returns

Table 7. Summary of lost-votes estimates

Location of lost votes	Mail ballot regime			
	Excuse required	No-excuse absentee	Permanent list	VBM
Ballot request not received by election office	0.4%	0.4%	0.1%	0.0%
Ballot request unfulfilled by election office	0.1%	0.1%	0.0%	0.0%
Transmitted ballot not received by voter	0.7%	0.7%	0.7%	0.7%
Returned ballot not received by election office	0.4%	0.4%	0.4%	0.4%
Ballot rejected by election office	1.8%	1.3%	0.8%	0.9%
Ballot fails to register a vote	1.5%	1.5%	1.5%	1.5%
Total	4.9%	4.4%	3.5%	3.5%