LOSING VOTES BY MAIL

*Charles Stewart III**

INTRODUCTION

The 2000 election was a wake-up call for America, demonstrating the vulnerability of the democratic process to breakdowns of voting technology, election law, and election administration. It shamed states and the federal government into action, yielding, in its most expansive (and expensive) manifestation, the Help America Vote Act (HAVA) of 2002.[1] HAVA contained many provisions; the one that most concretely addressed the Florida recount controversy required states to phase out mechanical lever machines and punch card voting. Hundreds of millions of federal dollars were authorized to underwrite this requirement.

The implementation of HAVA funds yielded equipment upgrades that, in turn, led to the recovery of at least a million votes in the 2004

* Kenan Sahin Distinguished Professor of Political Science, MIT, Cambridge, Massachusetts 02139. Email: cstewart@mit.edu

1. 42 U.S.C. §§ 15301–15545 (2006).

and 2008 presidential elections—votes that would have otherwise been lost because of the decrepitude of punch card and mechanical lever voting machines. HAVA increased the likelihood that a voter who wakes up on Election Day intending to vote, and then does everything required of him to cast a ballot, will have his vote counted as intended.

HAVA solved one set of problems, reducing "lost votes" due to voting technology failures, but it failed to address others. In particular, HAVA has been less effective in strengthening all of the ties that bind a citizen's desire to vote to the successful completion of the act. Technological failures are only one reason why votes are lost. HAVA addressed other reasons, such as registration problems and poor polling place practices; however, with the exception of the requirement that states maintain centralized voter registration lists, HAVA only addressed these reasons indirectly.

In retrospect, the biggest shortcoming of HAVA may have been its virtual lack of attention to voting by mail. As legislators respond to calls to make voting more convenient, and public officials respond to demands to make elections less costly, voting by mail is becoming more prevalent.[2] Yet despite the increasing prevalence of vote-by-mail, there has been virtual silence on the question of whether it causes more lost votes, compared to the in-person modes it is replacing.

Compared to in-person voting, either in traditional precincts on Election Day or in early-voting centers, vote-by-mail is highly decentralized. It relies on millions of people who are unschooled in election law and out of the sight of election administrators to perform a series of clerical tasks they otherwise rarely encounter. The chain-of-custody of ballots is less exacting. There are fewer opportunities to correct mistakes or clarify how to mark the ballot. Finally, the technological safeguards mandated by HAVA to protect against unintended over- and under-votes do not exist.

The trend toward vote-by-mail raises the question about whether the gains of HAVA, which have cured many of the ills of in-person

2. *See, e.g.*, Paul Gronke & Eva Galanes-Rosenbaum, *The Growth of Early and Non-Precinct Place Balloting: When, Why, and Prospects for the Future*, *in* AMERICA VOTES! A GUIDE TO ELECTION LAW AND VOTING RIGHTS 261, 268–72 (Ben Griffith ed., 2008) [hereinafter Gronke & Galanes-Rosenbaum, *The Growth of Early and Non-Precinct Place Balloting*]; Paul Gronke, *Early Voting Reforms and American Elections*, 17 WM. & MARY BILL RTS. J. 423, 423–24 (2008) [hereinafter Gronke, *Early Voting Reforms*]; Paul Gronke et al., *Convenience Voting*, 11 ANN. REV. POL. SCI. 437, 438 (2008) [hereinafter Gronke, *Convenience Voting*].

voting, may be undercut by the short-comings of this alternative voting method.

The answer to this question is mixed and preliminary. It is *mixed* because the best evidence suggests that the pipeline that moves mail ballots between voters and election officials is very leaky. On the other hand, the rise of voting-by-mail has not caused a precipitous rise in the residual vote rate, despite the lack of technological safeguards against over- and under-voting. The answer is *preliminary* because the quality of the best evidence we have is highly variable and reliant on reports from state election officials, who have fifty different ways of defining and gathering data about mail-in ballots. Some of the evidence also relies on the recalled memories of voters, who may have psychic incentives to blame others (i.e., election administrators) for their failures to vote.

The larger purpose of this paper is not to argue that voting methods that rely on the mail, whether they are mail-in absentee ballots or Oregon's statewide vote-by-mail system, *do* or *must* result in an inordinate number of lost votes.[3] Rather, this paper aims to show that we should be monitoring the lost-votes problem in the context of voting by mail, and that the current state of post-election data gathering is insufficient to identify where the biggest problems with vote-by-mail exist.

The remainder of this article proceeds as follows. First, I frame the problem of lost votes by introducing the notion of a "voting pipeline," which is inspired by the 2001 report of the Caltech/MIT Voting Technology Project (VTP), which articulated a holistic perspective of the lost-votes problem in the context of voting in-person on Election Day.[4] Second, I apply that metaphor to the vote-by-mail system and demonstrate that the voting pipeline in this context has many more weak points. Third, having framed the issue, I discuss the rise of vote-by-mail over the past four decades and identify the regions of the country in which the practice has become more prevalent. Fourth, relying on data from a unique public opinion survey, the Survey of the

3. The reader should be alerted to the fact that I generally treat vote-by-mail and absentee voting as synonymous. When the distinction is important, they are distinguished in the text. Almost all absentee voting is by mail, but some absentee ballots are cast in-person in election departments. In states that conduct all voting by mail, such as Oregon, it is technically a misnomer to refer to this as absentee voting. Overall, though, the variety in nomenclature is a set of distinctions without a difference, as far as addressing the first order set of theoretical and empirical issues is concerned.

4. CALTECH/MIT VOTING TECHNOLOGY PROJECT, VOTING: WHAT IS/WHAT COULD BE (2001), http://vote.caltech.edu/drupal/files/report/voting_what_is_what_could_be.pdf.

Performance of American Elections (SPAE) and the Election Assistance Commission's 2008 Election Administration and Voting Survey (EAVS),[5] I show that the number of lost votes through the vote-by-mail system in 2008 may have been as large as 7.6 million, or approximately one in five individuals who attempted to vote by mail. These votes were lost largely because of problems in the distribution system of mail-in ballots. There is little evidence that the vote-by-mail system is prone to excessive residual vote rates, despite the lack of feedback mechanisms that are supposed to alert in-precinct voters that they have over- or under-voted their ballots.

I conclude by considering some objections to the analysis I provide, in order to suggest an agenda for addressing the lost-vote problem along the vote-by-mail path. Three points are emphasized. First, addressing the lost-vote problem in the vote-by-mail context depends critically on improving the data-gathering and -analysis capacity in the domain of election administration. Second, progress will not be made in addressing the problems identified in this paper without more careful attention to the normative position that voting-by-mail occupies in American elections. Third, the empirical investigation of problems associated with voting-by-mail will be assisted by making sharper distinctions between situations in which voters are *allowed* to vote by mail, as opposed to *required* to vote by mail.

I.
THE VOTING PIPELINE AND LOST VOTES, 2000 TO 2008

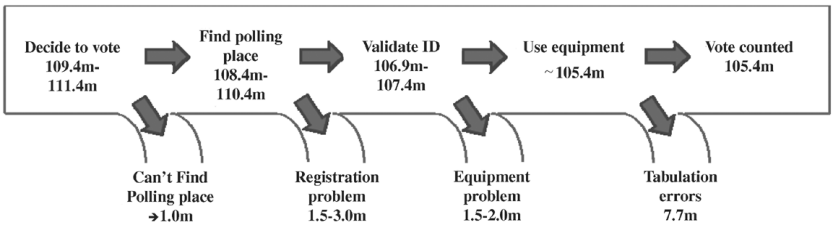
In their 2001 report, *Voting: What Is/What Could Be?*, the Caltech/MIT Voting Technology Project (VTP) argued that we will significantly underestimate the size of the lost vote problem if we focus only on voter confusion (illustrated by the butterfly ballot) and equipment malfunctions (illustrated by hanging chad) and fail to grasp the process of voting more holistically.[6] Figure 1 illustrates this line of thinking, using the metaphor of a pipeline.

---

5. R. MICHAEL ALVAREZ ET AL., CALTECH/MIT VOTING TECHNOLOGY PROJECT, 2008 SURVEY OF THE PERFORMANCE OF AMERICAN ELECTIONS (2008), http://vote.caltech.edu/drupal/files/report/Final%20report20090218.pdf [hereinafter SPAE]; U.S. ELECTION ASSISTANCE COMM'N, 2008 ELECTIONS ADMINISTRATION AND VOTING SURVEY (2009), http://www.eac.gov/assets/1/Documents/2008%20Election%20Administration%20and%20Voting%20Survey%20EAVS%20Report.pdf [hereinafter 2008 EAVS].

6. CALTECH/MIT VOTING TECHNOLOGY PROJECT, *supra* note 4, at 8–9 (explaining that votes were lost for a number of reasons besides problems with voting technology).

FIGURE 1: THE VOTING PIPELINE AND LOST VOTES IN 2000



Source: Caltech/MIT Voting Technology Project, *Voting: What Is/What Could Be* (2001), http://vote.caltech.edu/drupal/files/report/voting_what_is_what_could_be.pdf.

In thinking about the voting pipeline, it is helpful to imagine a representative voter who wakes up on Election Day intending to vote for her favored candidate for President. For this voter, "success" consists of having her vote recorded as she intended when the final tally is complete. To achieve success, four steps in a stylized election administration system must be navigated successfully. First, the voter must locate the polling place, travel there, and get to the front of the check-in line.[7] Second, the voter must identify herself and have her identity verified.[8] Third, she must use the equipment, and it must function flawlessly.[9] Fourth, the vote must be counted accurately.[10]

Using the 2000 Voting and Registration Supplement (VRS) of the Current Population Survey, the VTP estimated that almost one million would-be voters had their efforts to vote thwarted by polling

7. *Id.*
8. *Id.*
9. *Id.*
10. This fourth step, the accurate counting of votes, was not included in the original VTP formulation. However, as scholars have continued to explore how to improve elections, the problem of making sure votes are counted accurately has become a great concern, especially among those who are worried about the functioning of "black box" electronic voting machines. *See, e.g.*, Donald P. Moynihan, *Building Secure Elections: E-Voting, Security, and Systems Theory*, 64 PUB. ADMIN. REV. 515, 518–20 (2004). *See generally* BEV HARRIS, BLACK BOX VOTING: BALLOT TAMPERING IN THE 21ST CENTURY (2004). A series of international vote-counting controversies in countries like Ukraine, Iran, and Afghanistan have also sensitized many to the reality of vote tampering after ballots have been cast. *See* MIKHAIL MYAGKOV ET AL., THE FORENSICS OF ELECTION FRAUD: RUSSIA AND UKRAINE 138 (2009); Walter R. Mebane, Jr., *Fraud in the 2009 Presidential Election in Iran?*, 23 CHANCE 6, 6 (2010); Scott Worden, *Afghanistan: An Election Gone Awry*, 21 J. DEMOCRACY 11, 18–19 (2010). There is currently no way to estimate the extent of this problem in the United States, but it is an important potential source of losses in the voting pipeline that deserves to be acknowledged. For efforts to quantify, or at least identify, tampering with vote totals, see generally MAYAGKOV ET AL., *supra*; Mebane, Jr., *supra*; Walter R. Mebane, Jr., *Election Forensics: The Second-Digit Benford's Law Test and Recent American Presidential Elections*, *in* ELECTION FRAUD: DETECTION, PREVENTION, AND CONSEQUENCES (R. Michael Alvarez et al. eds., 2008).

place practices (i.e., long lines), and that between 1.5 and 3.0 million would-be voters had registration problems that kept them from voting.[11] Relying on statistical analysis of residual votes from 1988 to 2000, they also estimated that between 1.5 and 2.0 million votes were lost because of machine-related problems.[12] The VTP study did not originally identify the problem of accurately counting votes as part of the pipeline, so it failed to estimate the number of lost votes here.[13]

Approximately 105.4 million votes were recorded in the 2000 presidential election.[14] Working backward along the pipeline, adding in the four to six million votes that were "lost" when a link of the voting chain broke in 2000, as just described, we can estimate that 109.4 to 111.4 million people "woke up on Election Day intending to vote."

We can update these estimates using subsequent analysis. Later research estimated that, compared to 2000, approximately one million fewer votes were lost in 2004 because of machine problems;[15] it is reasonable to conclude that an additional half million votes were recovered because of the upgrades to voting machines that occurred between 2004 and 2008. Fewer respondents mentioned registration problems as a reason for not voting in 2004 and 2008 versus 2000, although the percentage of non-voters who blamed polling place practices was virtually unchanged across the same time period.[16] There-

11. CALTECH/MIT VOTING TECHNOLOGY PROJECT, *supra* note 4, at 9.

12. *Id.*

13. In all likelihood, the number of votes lost due to tabulation errors is already included in the number of machine-related problems. An example is provided by the 2008 election in the Republic of Georgia, in which a car carrying ballots to the central vote-counting center in Tbilisi was involved in an automobile accident that resulted in nearly all the ballot papers being lost. *Ballot Papers Lost in Car Accident in Georgia*, RIA NOVISTI (June 1, 2008, 12:46 PM), http://en.rian.ru/world/20080106/95543298.html. If this had occurred in most American jurisdictions, the fact that voters had checked in at the polls would have established that they had voted, but the aggregate vote result would fail to record an actual vote. To the outside observer, a vote lost because paper ballots have been lost in a car accident before they have been counted is indistinguishable from a vote lost due to an under-voted ballot.

14. David Leip, *2000 Presidential General Election Results*, DAVE LEIP'S ATLAS OF U.S. PRESIDENTIAL ELECTIONS, http://uselectionatlas.org/RESULTS/national.php?year=2000&off=0&f=1 (last visited Oct. 19, 2010).

15. Charles Stewart III, *Residual Vote in the 2004 Election*, 5 ELECTION L.J. 158, 158 (2006).

16. *Compare* THOM FILE & SARAH CRISSEY, U.S. CENSUS BUREAU, P20-562, VOTING AND REGISTRATION IN THE ELECTION OF NOVEMBER 2008 14 (2010), http://www.census.gov/prod/2010pubs/p20-562.pdf [hereinafter 2008 VRS] (5.5% reporting "registration problems"), *and* KELLY HOLDER, U.S. CENSUS BUREAU, P20-556, VOTING AND REGISTRATION IN THE ELECTION OF NOVEMBER 2004 15 (2006), http://www.census.gov/prod/2006pubs/p20-556.pdf [hereinafter 2004 VRS] (6.6% reporting "registration problems"), *with* AMIE JAMIESON ET AL., U.S. CENSUS BUREAU, P20-

fore, if voters had faced the same conditions in 2000 as they did in 2008, there would have been one million fewer lost votes due to equipment problems and approximately half a million fewer lost votes due to registration problems.[17] This represents an overall reduction in lost votes in the range of 20–30%, most of which can be attributed to efforts associated with the implementation of HAVA.

II.
THE VOTE-BY-MAIL PIPELINE

One flaw with the description of the voting pipeline is that for many voters, the decision to vote is not made on Election Day. It may be made days or weeks ahead of time, when the voter decides whether to vote by mail (usually absentee) or wait until Election Day. If the voter decides to wait until Election Day, then the pipeline illustrated in Figure 1 still applies. However, if the voter decides to use the mail route, a different pipeline is involved.

Figure 2 provides a schematic voting pipeline that includes the possibility of vote-by-mail. The first decision is whether to vote in-person or by mail.[18] If the voter decides to vote in person, then the pipeline proceeds as before. If the voter decides to vote by mail, then the new pipeline is illustrated by the bottom track.

With the exception of voters in Oregon, Washington, and places with permanent absentee voting, the vote-by-mail process begins when the voter requests a ballot.[19] The first leak in the pipeline can

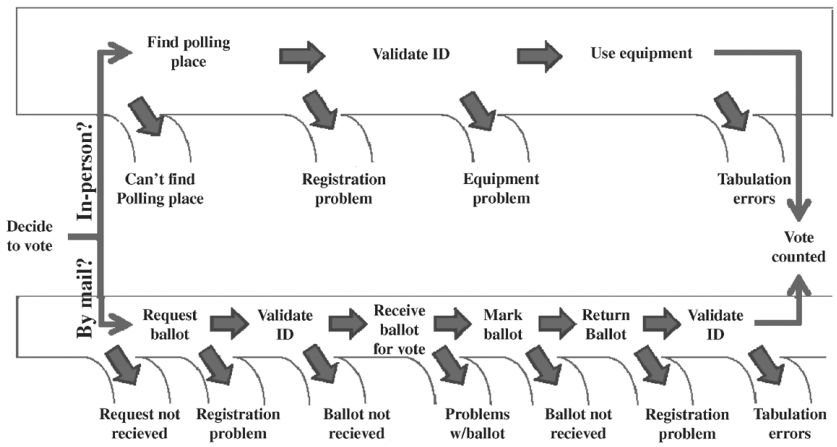
542, VOTING AND REGISTRATION IN THE ELECTION OF NOVEMBER 2000 10 (2002), http://www.census.gov/prod/2002pubs/p20-542.pdf [hereinafter 2000 VRS] (7.4% reporting "registration problems").

17. In 2000, the percentage citing "registration problems" in the VRS as a reason for not voting was 7.4%. 2000 VRS, *supra* note 16. That number fell to 6.6% in 2004 and 5.5% in 2008. 2004 VRS, *supra* note 16; 2008 VRS, *supra* note 16. These numbers differ from those reported in the final VRS reports, but were obtained directly from the Census Bureau's public use data. The datasets are available at the Census Bureau's DataFerrett website. *DataFerrett*, U.S. CENSUS BUREAU, http://dataferrett.census.gov/index.html (last visited October 12, 2010). Those citing "inconvenient polling place or hours or lines too long" in the VRS amounted to 2.8% of non-voters in 2000 and 2.9% in 2004 and 2008. 2000 VRS, *supra* note 16; 2004 VRS, *supra* note 16; 2008 VRS, *supra* note 16. Below, I estimate that the total number of non-voting registered voters in 2008 was 57.6 million, compared to 40 million non-voters in 2000. Therefore, the *absolute number* of votes lost due to polling place practices and registration problems may have risen over the decade, but that is only because the number of non-voters has grown.

18. I assume that issues related to early in-person voting may be treated as simply a part of traditional in-precinct voting. This assumption may or may not be reasonable, but it does not affect the line of reasoning that follows.

19. According to the 2008 EAVS Report, fifteen states maintained permanent absentee ballot databases in 2008. 2008 EAVS, *supra* note 5, at 9. 37% of absentee

FIGURE 2: THE VOTE-BY-MAIL PIPELINE



occur if the request is never received. If it is received, then a leak can occur at the point of verifying the voter's identity and eligibility to receive a ballot. If there is a registration problem, this too represents a leak. If the ID is validated, then a ballot is sent to the voter to be cast. If the voter does not receive the mailed ballot, another pipeline leak occurs. Once the voter marks the ballot, it is returned—but it can get lost in the return mail, yet another potential leak. If the ballot is returned for counting, it must again undergo identification verification. If there is a registration error that did not manifest earlier, another vote leaks from the pipeline. The final step is the accurate tabulation of the returned ballot.

Before considering the likelihood of problems to emerge at each step, a few things leap out at us when we compare the top and bottom tracks of Figure 2. First, there are simply more ways to lose votes along the bottom than along the top. It can be argued that this is because the bottom track is described at a greater degree of granularity than the top, but this argument brings us to the second major point: the reason the bottom path provides more ways to lose votes is that many steps of the vote-by-mail process have to be accomplished twice. The voter sends two pieces of mail to the central elections office: (1) the request for the absentee and (2) the absentee ballot itself.[20] Only a

ballots transmitted in 2008 were off of these permanent lists. *Id.* at 38. If we include the transmission of mail ballots from Oregon and Washington in this calculation, the percentage of automatic transmissions approaches 50%.

20. Of course, the "first piece of mail" could actually be a phone call or a visit to the elections office, but the point remains that the voter has to navigate the elections office twice when it comes to mail-in ballots.

miniscule fraction of the mail is actually lost, but navigating mail channels involves more than simply surviving the U.S. Postal Service, including, for example, the handling and sorting process at both ends of a letter's journey.[21] We simply do not know how reliable the system is once a ballot has left the hands of a postal worker.

Below, I estimate how many votes are lost at every point along the pipeline. For the moment, it is sufficient to note that the *opportunities* to lose votes appear to be greater along the mail route than along the in-person route.

III.
The Rise of Vote-by-Mail since 1972

The logistics of requesting and delivering absentee ballots introduce more opportunities for lost votes. In order to estimate the potential magnitude of the problem, I start by reviewing the rise of voting-by-mail over the past four decades.

Figure 3 shows the percentage of ballots cast by mail in federal elections from 1972 to 2008. The estimates are provided by the VRS and are based on self-reporting in a national survey (the Census Bureau did not gather information about voting mode in 1988).[22] The prevalence of mail-in ballots has grown exponentially since 1972.[23] The pace of growth has quickened over the past decade, as more states have relaxed their "for cause" absentee laws, developed permanent absentee ballot databases, and mandated the use of the mails for an

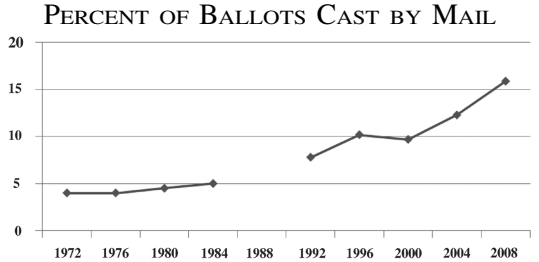
21. The United States Postal Service (USPS) does not release estimates of the number of letters that are never delivered. However, in its most recent statistical report, the USPS did report that the average piece of pre-sorted first-class mail was delivered in 2.3 days, with 99.9% of mail delivered within ten days. *See* U.S. Postal Quarterly Statistics Report: Postal Quarter III Fiscal Year 2010 81 (2010), http://www.usps.com/financials/_pdf/QSR_FY10QT3.pdf.

22. 2008 VRS, *supra* note 16, at 1. The EAVS collects reports from states about the actual number of ballots cast, including mode (in-person, civilian absentee, etc.). Although the response rate in 2008 was nearly 100%, in previous years, the survey suffered from significant non-response problems. *Compare* 2008 EAVS, *supra* note 5, at 16 (showing response rates for all questions above 70%, with most above 90%), *with* Kimball W. Brace & Michael P. McDonald, U.S. Election Assistance Comm'n, Final Report of the 2004 Election Day Survey 1-6 (2005), http://www.eac.gov/assets/1/Documents/2004%20Election%20Administration%20and%20Voting%20Survey%20Report%20EAVS%20with%20tables.zip ("Even with the follow-up review, many responses to the Election Day Survey are incomplete."), *and* U.S. Election Assistance Comm'n, The 2006 Election Administration and Voting Survey 27 (2007), http://www.eac.gov/assets/1/AssetManager/2006%20EAVS%20Report%20(All%20Chapters).pdf (showing response rates as low as 35.5% for some questions). Because we cannot use EAC data to describe the long-term trend, I rely on the Census Bureau self-reports, which are cross-validated by the EAC data for 2008.

23. *See infra* Figure 3.

increasing number of voters.[24] Approximately 16% of all ballots cast for president in 2008 were sent through the mails.[25]

FIGURE 3: THE RISE OF VOTING BY MAIL, 1972–2008



Source: Census Bureau, Current Population Survey, Voting and Registration Supplement, various years.

Figure 4 shows the percentage of ballots cast by mail in each state in 2008, estimated using the VRS.[26] Self-reported mail-in ballots ranged from 1.9% of respondents in West Virginia to 97.5% of respondents in Oregon. The laws and practices of the states have a significant influence over these absentee rates. In the twenty-seven states that allowed "no-excuse" absentee voting by mail in 2008,[27] 22% of ballots were cast by mail, compared to 6.0% of ballots in the twenty-one states and the District of Columbia that still require an excuse to vote absentee. Focusing only on the no-excuse states, the twelve with permanent absentee voting saw 39.8% of their ballots cast by mail, compared to an 11.2% mail rate in no-excuse states without permanent absentee voting.

---

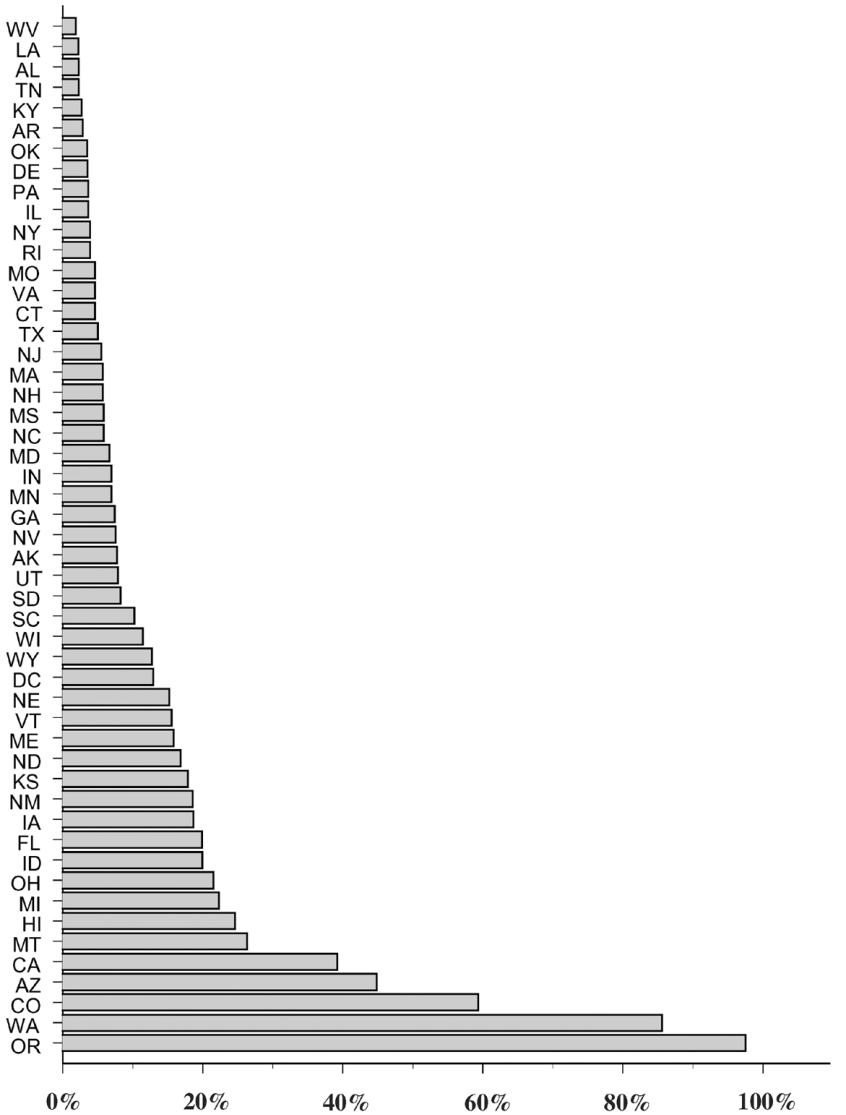
24. *Id.*

25. U.S. ELECTION ASSISTANCE COMM'N, 2008 ELECTION ADMINISTRATION AND VOTING SURVEY DATASET (2009), http://www.eac.gov/assets/1/Documents/2008%20EAVS%20XLS.zip [hereinafter EAVS DATASET]. Based on the 2008 EAVS dataset, it appears that 4% of the mail-in ballots were from overseas, with the rest being either traditional absentee ballots or by-mail ballots from jurisdictions that mandate it.

26. Inconsistencies across states in reporting statistics about absentee ballots preclude our use of the 2008 EAVS dataset to make this estimate. Forty-four states provided data about absentee voting that appear to be usable in the EAVS dataset. Focusing on these forty-four states, the cross-state correlation between the EAC data, which are based on ballot counts from election officials, and the Census Bureau data, which are based on self-reports from voters, is .97 (weighting each state by total turnout). Overall, the Census Bureau estimate among these forty-four states is (on aggregate) about 2.4 percentage points lower than the EAC count.

27. Oregon and Washington are excluded from this analysis. For a summary of absentee ballot laws see generally *Pre-Election Day and Absentee Voting by Mail Rules*, ELECTION REFORM INFORMATION PROJECT (Oct. 22, 2007), http://www.pewcenteronthestates.org/uploadedFiles/e%20and%20a%20voting%20laws.pdf.

FIGURE 4: STATE VOTE-BY-MAIL IN 2008



Source: Census Bureau, Current Population Survey, Voting and Registration Supplement, 2008.

Although mail is currently the minority mode for voting, it is growing, and growing faster in some places than others. In nine states, over 20% of voters use the mails. Because of concerns about the costs and logistical headaches associated with in-precinct voting, election officials in many parts of the country feel compelled to respond to citizen demands for greater convenience by expanding vote-by-mail

options.[28] However, for the reasons discussed in Part II, this trend raises questions about whether the voting pipeline is becoming more and more fragile for voters in these states.

IV.
An Estimate of Lost Votes in the 2008 Vote-by-Mail System

Estimating the strength of the vote-by-mail pipeline requires knowledge of how many registered voters attempted to use the mail to vote in 2008 and how many were successful. Estimating the strength of the vote-by-mail system requires us to have solid figures pertaining to all the possible sources of vote loss that were identified in Figure 2. Although election data for the purposes of diagnosing problems with the election system are better than they used to be, they are unfortunately still a work-in-progress. Therefore, this section takes a first stab at measuring lost votes along the voting-by-mail channel, but these estimates should be considered illustrative, not definitive.

I rely on two data sources. The first is the EAC's 2008 Election Administration and Voting Survey (EAVS).[29] The second is a national survey conducted by a team of researchers who were associated with the Caltech/MIT Voting Technology Project and funded by the Pew Charitable Trusts, the 2008 Survey of the Performance of American Elections (SPAE).[30] The EAVS data give us insights into most of the internal processes involving mail ballots. The SPAE data are useful for estimating how many individuals actually requested an absentee ballot.

EAVS was conducted by sending a survey to election officials in each state, who were asked to provide a large amount of data about election administration in each county for the November 2008 election.[31] Data gathered included the total number of ballots cast, the number of ballots cast by domestic absentee and overseas (UO-

28. *See, e.g.*, Gronke, *Early Voting Reforms*, *supra* note 2, at 427–28; Gronke, *Convenience Voting*, *supra* note 2, at 438; Gronke & Galanes-Rosenbaum, *The Growth of Early and Non-Precinct Place Balloting*, *supra* note 2, at 270–71.

29. EAVS Dataset, *supra* note 25.

30. SPAE, *supra* note 5. The 2008 Survey of the Performance of American Elections was generously funded by a grant from the Pew Center of the States, through their Make Voting Work initiative, along with the JEHT Foundation and the American Association of Retired Persons (AARP). The analysis presented here is solely the responsibility of the author.

31. 2008 EAVS, *supra* note 5. For New England, the survey asked that the data be broken down by municipality, not county. *Id.* at 18.

CAVA)[32] ballots, and the number of precincts in the county, among other facts.[33] Although the response rate for the 2008 survey was significantly improved over the 2004 and 2006 versions of this survey, the dataset is still incomplete. For instance, three states (Alabama, Massachusetts, and South Carolina) did not provide a breakdown of ballots by type.[34] Furthermore, in six jurisdictions (Connecticut, D.C., Hawaii, Indiana, Nebraska, and Texas) the sum of all votes cast in the different categories exceeded the total number of ballots cast in the jurisdiction.[35]

Although the EAVS data have limitations, they are the best data available for assessing most of the details of election administration. Problems with missing and inconsistent data in the domestic absentee portion of the survey are minimal, so any conclusions we draw should be robust with reference to decisions about how to impute missing data. With these cautions in mind about the quality of the data, I estimate that 27.9 million ballots were cast by mail in 2008. This estimate is based on the EAVS data, including data that must be imputed because of missing values.[36] The number of ballots transmitted to voters and returned to be counted can be estimated using the EAVS dataset.

A. Estimating the Number of Mail-In Ballots Transmitted to Voters

EAVS contains a variable for the number of domestic absentee ballots transmitted to voters, which is used as the estimate for forty of the states.[37] Two states reported precisely zero absentee ballots transmitted, Alabama and Washington.[38] To create an estimate for Alabama, I multiplied the Census Bureau's estimate of the percentage of ballots cast via mail times the number of ballots counted for President. Washington, which conducted all balloting by mail in every county except one, did not report the number of mail ballots that were trans-

32. These ballots are regulated by the Uniformed and Overseas Citizens Absentee Voting Act (UOCAVA), 42 U.S.C. §§ 1973ff to ff-6 (2006).

33. 2008 EAVS, *supra* note 5.

34. 2008 EAVS Dataset, *supra* note 25. Alabama provided a report for the number of domestic absentee ballots cast but otherwise did not provide an estimate of the number of ballots cast in precincts.

35. *Id.*

36. This estimate exceeds the EAVS report of 25.6 million absentee ballots counted in 2008, for several reasons. The major reason for the deviation from the raw EAVS report is that I have added the mail-in ballots cast in Oregon and Washington, which were not counted as absentee votes. I also imputed missing data for some counties and states that did not report the necessary statistics to the EAC. These imputations are less critical for the estimates that follow.

37. EAVS Dataset, *supra* note 25, variable C1a.

38. EAVS Dataset, *supra* note 25.

mitted to voters in the thirty-eight counties with mail ballots.[39] Because Washington mailed ballots to all registered voters, except those in Pierce County, my estimate of the number of transmitted ballots is equal to the number of registered voters in the state, excluding Pierce County.

Oregon, which conducts all its elections by mail, did not report the number of transmitted ballots, but did report a small number (19,782) of ballots transmitted to voters through a separate absentee procedure.[40] To estimate the number of transmitted mail ballots in Oregon, I used the number of registered voters in that state, as well. Oregon, like Washington, automatically mails ballots to all registered voters, at their registration address.[41]

In the remaining eight states, at least one county failed to report the number of mail ballots transmitted to voters, resulting in an under-reporting in each of these states.[42] Therefore, I imputed the number of transmitted ballots by multiplying the Census Bureau estimate of the percentage of ballots cast through the mails for the entire state by the number of ballots counted in the county with the missing data. This produces an under-estimate of the number of transmitted ballots in these states, but the error is likely to be small, because the missing counties are few and tend to have small voting populations.

Accounting for these corrections, I estimate that 31.6 million ballots were transmitted from election officials to voters via the mail in 2008.

B. Estimating the Number of Mail Ballots Returned for Counting

The number of ballots that were returned to election officials for counting is captured in a single EAVS question, which was answered in full by thirty-eight states.[43]

For Alabama, Oregon, and Washington—states that did not report the number of ballots returned for counting—I assumed the return rate was the same as the return rate for all states that *did* report the number of returned ballots, which was 90.8%.[44] I then multiplied this percentage by the number of transmitted ballots that were imputed for these three states to establish the estimate of the number of returned ballots.

---

39. *Id.*
40. *Id.*
41. OR. REV. STAT. § 254.465 (2009).
42. EAVS Dataset, *supra* note 25.
43. *Id.*, variable C1b.
44. *Id.*

Connecticut reported more ballots returned than transmitted.[45] Therefore, I set the number of returned ballots equal to the number transmitted.

The nine remaining states were missing data from at least one county.[46] For these nine states, I calculated the return rate in the counties that had reported a full set of data. Then, for each state, I took the state's overall return rate (for the counties with complete data) and multiplied it by the number of transmitted ballots reported for any county with missing data.

After making all the described imputations, I estimate that 28.7 million ballots were returned for counting in 2008.

C. Estimating the Number of Ballots Counted

The next step is to estimate the number of ballots counted. The beginning point for this estimation is the EAVS question that records the number of domestic absentee ballots that were counted.[47]

Forty-two states reported complete data for the relevant question,[48] leaving values for nine states to be imputed. The values for Alabama and Washington were imputed by first calculating the overall rate of returned ballots counted for the states with a full set of data. This rate was 97.4%. I then multiplied this percentage by the imputed number of returned ballots for these two states to produce the estimates. The number of counted mail ballots for Oregon was set to match the reported turnout for the state.

The remaining states had data missing for one or more counties. For each of these states, I calculated the "counting rate" within the state, using the counties that had the requisite data. I then multiplied the state counting rate by the number of returned ballots in the county to fill in the missing values for these counties.

Using this method, I estimate that 27.9 million mail ballots were counted nationwide.

D. Estimating the Number of Requested Ballots

Finally, I turn to the number of people who actually requested an absentee ballot. This is not an issue probed in the EAVS, but it *is* an issue examined by the 2008 Study of the Performance of American

45. *Id.*
46. *Id.*
47. *Id.*, variable C4a.
48. *Id.*

Elections (SPAE).[49] The SPAE was a nationwide post-election survey of 10,000 registered voters in November 2008 that focused on a series of election administration issues, from the perspective of voters.[50] The survey contained questions that can help to quantify the number of initial mail ballot requests, particularly the number of unfulfilled requests. For the foregoing analysis, all results have been weighted so that the sample reflects a national cross-section of registered voters.

The SPAE asked registered voters whether they voted in the 2008 November general election.[51] It then asked voters which mode they used to vote—in-person on Election Day, in-person before Election Day, or by mail.[52] Registered voters who reported that they did not vote were then asked several follow-up questions. One such question was the following:

> Sometimes when voters can't get to the polls on Election Day, they vote using an absentee ballot. Please indicate which of the following statements most closely describes why you did not vote absentee in the November 2008 General Election.[53]

Among the response categories was the answer, "I requested an absentee ballot, but it never came;" 6.8% of non-voters chose this response.[54] If we use the best estimate of the number of registered non-voters in 2008, 57.6 million,[55] then the estimated number of requests for absentee ballots that went unfulfilled is 6.8% of 57.6 million, which is 3.9 million.

Above, I estimated that 31.6 million mail ballots were transmitted from election offices to voters. If we add the 3.9 million estimate of

---

49. SPAE, *supra* note 5, at 26.

50. *Id.* at *i*. The SPAE also included a parallel telephone survey that was conducted in ten states, to cross-validate the results of the Internet survey. *Id.* at 3. The telephone responses have been omitted from this analysis.

51. *Id.* at 162 (question Q1).

52. *Id.* at 163 (question Q5).

53. *Id.* (question Q3).

54. *Id.* The response categories were "I had no interest in voting in this election" (24.0% of responses); "It was too late to request an absentee ballot once I thought about it" (11.1%); "I requested an absentee ballot, but it never came" (6.8%); "I wouldn't have been allowed to vote absentee according to my state's election law" (4.1%); "Requesting an absentee ballot requires too much effort" (2.2%); "I didn't know how to request an absentee ballot" (21.8%); and "Other" (9.8%). *Id.*

55. The 57.6 million figure is arrived at as follows. The 2008 EAVS Report estimates that 190.1 million Americans were registered to vote in 2008. 2008 EAVS, *supra* note 5, at 8. David Leip's *Atlas of U.S. Presidential Elections* reports that 131.5 million votes were cast for president in 2008. David Leip, *2008 Presidential General Election Results*, DAVE LEIP'S ATLAS OF U.S. PRESIDENTIAL ELECTIONS, http://uselectionatlas.org/RESULTS/national.php?year=2008&off=0&f=1 (last visited Oct. 19, 2010). Therefore, 190.1 million minus 132.5 million equals 57.6 million, the estimated number of registered non-voters in 2008.

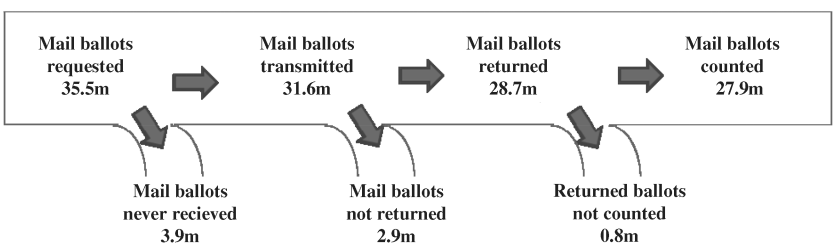
unfulfilled mail ballot requests to the number of estimated mail transmissions, we determine that 35.5 million Americans requested absentee ballots in 2008.

E. Summarizing the Calculations

Figure 5 provides a summary of these calculations. In words:

- 35.5 million requests were made for a mail ballot in 2008. Of these,
  - 3.9 million requests were unfulfilled, leaving
- 31.6 million mail-ballots transmitted to voters. Of these,
  - 2.9 million were not returned for counting, leaving
- 28.7 million mail-ballots returned for counting. Of these,
  - 0.8 million mail-ballots were not counted, leaving a total of
- 27.9 million absentee ballots counted in 2008.

FIGURE 5: LOST VOTES FROM MAIL-IN-BALLOTS



If we add together the estimated number of people whose ballot requests were unfulfilled, the number of ballots not returned for counting, and the returned ballots that were not counted, we determine that a total of 7.6 million votes left the mail-ballot pipeline at some point between requesting a ballot and the counting of the ballots. This amounts to 21% of all ballot requests.

What should we glean from these estimates? First, the 7.6 million figure (or the 21% rate) should not be taken as a firm measure of the number of potential mail voters whose votes were lost in 2008. If our goal is to calculate the number of lost votes through the mail-ballot route, these estimates are probably too high. Because non-voters are well known for rationalizing their failure to vote,[56] we must treat any

56. *See, e.g.*, Simon Jackman, *Correcting Surveys for Non-response and Measurement Error Using Auxiliary Information,* 18 ELECTORAL STUD. 7, 13–14 (1999); Brian D. Silver et al., *Who Overreports Voting?,* 80 AM. POL. SCI. REV. 613, 614 (1986).

survey-based estimate of the number of non-delivered absentee ballots with caution. Because voting is a socially desirable behavior, there are strong psychological pressures prompting non-voters to blame the actions of others, including election administrators, for their failure to vote.

However, even if the estimated number of unfulfilled requests for a mail ballot is off by an order of magnitude—that is, the correct estimate is closer to 390,000 than 3.9 million—the resulting lost vote estimate is still around 4.1 million, or a lost vote rate of 13%.

Turning our attention to the other estimates, the other "leaks" in the pipeline are not due entirely to errors beyond the voter's control. A ballot that is mailed out but not returned may reflect a voter's decision not to cast a vote after all, or a decision to go to the polls on Election Day instead.[57] Still, if we consider all the non-returned ballots as reflecting these voters' conscious decision to abstain, after they have gone through the effort to obtain an absentee ballot, it implies that the abstention rate among mail-in voters is 12%, an implausibly high rate.

Because many states do not record why absentee ballots are rejected, we cannot say with certainty why nearly one million returned mail-in ballots went uncounted.[58] Among states that do record why absentee ballots are rejected, the primary reason is the failure to return the ballot before the deadline.[59] A variety of reasons involving missing signatures are also common.[60]

Caveats noted, the number of potentially lost ballots through the vote-by-mail channel is significant, especially compared to estimates that focus on in-person voting. The 22% lost vote rate through the vote-by-mail channel is significantly larger than the overall rate of 4% estimated for all voters by the VTP in 2000. Even if future research with better data narrows the gap, there is no doubt that the magnitude of the phenomenon demands attention.

57. Los Angeles County, California alone reported that 33,078 absentee ballots were rejected because the voter also voted a provisional ballot. EAVS Dataset, *supra* note 25. These are probably voters who did not remember that they had already voted absentee, or who doubted their absentee ballot would be counted, for whatever reason.

58. *Id.*, Questions C5a–C5v. Of the 4,517 units that were asked to respond to the questionnaire, 947 reported no reasons for why absentee ballots were rejected. *Id.*

59. *Id.*, Question C5a.

60. *Id.*, Questions C5b (voter signature), C5c (witness signature), C5d (non-matching signature), C5e (election official signature).

V.
VOTE-BY-MAIL AND RESIDUAL VOTES

The analysis thus far has focused on the distribution channel of ballots as the major source of lost mail votes. Ballots can also be lost if the ballot-marking and -counting process fails. This is a problem shared with more traditional in-person voting, which is the context in which this problem has generally been studied.[61] Analogous problems in the mail ballot domain have gone almost entirely unstudied, even though there are reasons to believe that they may be more acute.

Research following the 2000 presidential election demonstrated that votes can be lost when voters are confused by the layout of the ballot or when the voting technology malfunctions.[62] Voter confusion and machine malfunction can lead to unintended over- and under-votes, both of which are treated as blank ballots for the purposes of counting votes.[63] The problem of over- and under-votes together is sometimes termed the "residual vote" problem.[64]

HAVA mandated that in-person voting technologies inform voters when they over- or under-vote their ballot, and that these technolo-

61. Stephen Ansolabehere & Charles Stewart III, *Residual Votes Attributable to Technology*, 67 J. POL. 365, 366 (2005); Stewart, *supra* note 15; Jonathan N. Wand et al., *The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida*, 95 AM. POL. SCI. REV. 793, 795–96 (2001); *see* Caltech/MIT Voting Technology Project, *supra* note 4, at 20–23; Douglas W. Jones, *Chad—From Waste Product to Headline*, http://www.cs.uiowa.edu/~jones/cards/chad.html (last visited Sept. 13, 2010).

62. *See* Ansolabehere & Stewart, *supra* note 61, at 366; Stewart, *supra* note 15, at 168; Wand, *supra* note 61, at 795–96; Caltech/MIT Voting Technology Project, *supra* note 4, at 27; Jones, *supra* note 61.

63. Caltech/MIT Voting Technology Project, *supra* note 4, at 23.

64. A "residual vote" is simply a ballot that is untallied because it is over- or under-voted. The residual vote *rate* is calculated by dividing the number of residual votes by the number of ballots presented for counting, or alternatively, the number of voters who check in at the polls. Residual vote analysis is usually done on the race at the top of the ballot, such as president, but in principle it can be applied to any race on a ballot, as the discussion of San Francisco below demonstrates. For research involving residual votes, see generally R. Michael Alvarez et al., *Studying Elections: Data Quality and Pitfalls in Measuring the Effects of Voting Technologies*, 33 POL'Y STUD. J. 15 (2005); Stephen Ansolabehere, *Voting Machines, Race, and Equal Protection*, 1 ELECTION L. J. 61 (2002); Ansolabehere & Stewart, *supra* note 61; Justin Buchler et al., *Punch Card Technology and the Racial Gap in Residual Votes*, 2 PERSP. ON POL. 517 (2004); Michael C. Herron & Jasjeet S. Sekhon, *Black Candidates and Black Voters: Assessing the Impact of Candidate Race on Uncounted Vote Rates*, 67 J. POL. 154 (2005); Jonathan I. Leib & Jason Dittmer, *Florida's Residual Votes, Voting Technology, and the 2000 Election*, 21 POL. GEOGRAPHY 91 (2002); Donald P. Moynihan & Carol L. Silva, *What Is the Future of Studying Elections? Making the Case for a New Approach*, 33 POL'Y STUD. J. 31 (2005); Stewart, *supra* note 15.

gies provide a way to correct such errors.[65] Mail-in ballots are exempt from this second chance feedback requirement, which raises the possibility that mail ballots will more likely contain unintentional residual votes than in-person ballots. Unfortunately, the research on this topic has been minimal, in large part because very few states break down residual vote rates by voting mode—in-person *vs.* absentee.

Furthermore, even if residual vote rates were reported for different voting modes, it is unclear what we would make of these comparisons, since voters do not randomly distribute themselves into the in-person and vote-by-mail categories.[66] In states where absentee balloting is primarily used for the convenience of a few voters, mail-in voters tend to be better educated and have higher incomes than in-person voters.[67] This population tends to produce relatively fewer residual votes.[68] The residual vote rate of mail-in ballots could therefore be lower than in-person ballots, simply for demographic reasons.

With this caveat in mind, there is no statistical evidence yet that a rise in vote-by-mail has led to an increase in the residual vote rate at the top of the ticket. The simplest way to test this relationship is to calculate the correlation across states of the changes in the vote-by-mail rate and the residual vote rate from 2000 to 2008. The correlation is an anemic -.06, which is no different than random chance.

A more sophisticated way to test whether the rise of vote-by-mail has led to an increase in the residual vote rate is to include a measure of the vote-by-mail rate in a larger regression model that tests for changes in the residual vote rate at a lower degree of aggregation (such as the county), controlling for factors such as changes in voting technology, turnout, and other local conditions. Prior research has done precisely this, and I adapt these previous estimation techniques

65. Help America Vote Act § 301, 42 U.S.C. § 15481 (2006).

66. Robert M. Stein, *Early Voting*, 62 PUB. OPINION Q. 57, 59 (1998); JOHN MARK HANSEN, EARLY VOTING, UNRESTRICTED ABSENTEE VOTING, AND VOTING BY MAIL 1 (2001), http://www.tcf.org/Publications/ElectionReform/NCFER/hansen_chap5_early.pdf.

67. For instance, in the SPAE, the average household income of in-person voters was $65,100, compared to $67,100 for in-person early voters and $72,300 for absentee and mail voters. SPAE, *supra* note 5, at 13–14. The average education for in-person voters was 13.7 years, compared to 13.9 for early voters and 14.1 years for mail voters. *Id.*

68. David C. Kimball & Martha Kropf, *Ballot Design and Unrecorded Votes on Paper-Based Ballots*, 69 PUB. OPINION Q. 508, 523 (2005); David C. Kimball & Martha Kropf, *Voting Technology, Ballot Measures, and Residual Votes*, 36 AM. POL. RES. 479, 492 (2008); Stephen Knack & Martha Kropf, *Roll-Off at the Top of the Ballot: Intentional Undervoting in American Presidential Elections*, 31 POL. & POL'Y 575, 584 (2003); Stephen Knack & Martha Kropf, *Voided Ballots in the 1996 Presidential Election: A County-Level Analysis*, 65 J. POL. 881, 889 (2003).

here.[69] Unfortunately, we do not have county-by-county vote-by-mail rates for all counties across the country, and so a comprehensive test is currently impossible to conduct. However, we can augment these models by including the state-level measure of the vote-by-mail rate for each election year as an additional control, and adding 2008 election returns to the analysis, in order to gain greater statistical leverage.

The appendix contains the details of the regression analysis.[70] The most important finding from a regression analysis of these data is that, in states that have seen the biggest increase in vote-by-mail rates over the past decade, the residual vote rates have fallen the most. The size of the effect predicts that, on average, a one standard deviation increase in the vote-by-mail rate in a state would reduce the residual vote rate of a county by 1.0% points. Therefore, the initial evidence shows that a rise in vote-by-mail has not led to a rash of new residual votes at the top of the ticket.

Another confounding factor in examining the residual vote rate of mail-in ballots is that some voters do not return all the ballot cards. This occurs infrequently when voters vote in person. San Francisco, California provides a cautionary illustration from 2008. Typical of California counties, the ballot in San Francisco in the November 2008 general election had to accommodate numerous races. Not only were there races for U.S. President, U.S. House, Superior Court, Bay Area Rapid Transit (BART) directors (for some precincts), County Board of Education (vote for four), County Board of Supervisors (rank-choice), and Community College Board (vote for four), but there were also twenty-two county-wide propositions and twelve state-wide propositions.[71] To accommodate all of these races, the optical scan ballot was distributed across four ballot cards, on which cards information was printed on both the front and back.[72] Card 1 contained the federal, state, and local offices, printed on the front and back. Card 2 contained the state propositions, with propositions 1A–5 printed on the front and 6–12 printed on the back. Card 3 contained the county measures, with measures A–J on the front and measures K–V on the back. Card 4 contained the rank-choice ballot for supervisors.

---

69. Ansolabehere & Stewart, *supra* note 61; Stewart, *supra* note 15.

70. *See infra* Appendix.

71. DEPARTMENT OF ELECTIONS, CITY AND COUNTY OF SAN FRANCISCO, NOVEMBER 4, 2008 STATEMENT OF VOTES, http://www.sfgov2.org/ftp/uploadedfiles/elections/ElectionsArchives/2008/november/SOV_081104.pdf [hereinafter SAN FRANCISCO 2008 STATEMENT OF VOTES].

72. CITY AND COUNTY OF SAN FRANCISCO, SAMPLE OFFICIAL BALLOT, NOVEMBER 4, 2008, BALLOT STYLE 39 (on file with author).

San Francisco is not unique in having a complicated ballot.[73] What makes it unique is that it publishes precinct-by-precinct statistics that record not only the number of votes for each candidate by election mode (in-person or absentee), but also publishes the number of cards that were physically returned for counting. Table 1 reports the basic statistics.

TABLE 1: RETURN RATES OF BALLOT CARDS, SAN FRANCISCO, CALIFORNIA, NOVEMBER 2008 GENERAL ELECTION

	In-person, not returned		Absentee, not returned		Total, not returned	
	N	Pct.	N	Pct.	N	Pct.
Card 1	—	—	—	—	—	—
Card 2	279	0.13%	4,697	2.6%	4,976	1.3%
Card 3	674	0.32%	5,943	3.3%	6,617	1.7%
N	209,527		178,585		388,122	
Card 4	1,727	1.4%	5,906	5.5%	7,633	3.3%
N	127,027		107,651		234,678	

Source: Department of Elections, City and County of San Francisco, Nov. 4, 2008 *Statement of Votes*, http://www.sfgov2.org/ftp/uploadedfiles/elections/ElectionsArchives/2008/november/SOV_081104.pdf.

Overall, 388,122 first cards were returned to be counted: 209,527 via the in-person route and 178,585 by mail. The San Francisco Elections Department published these numbers as the total turnout for the city and county, so they will be used as the denominators in calculating the non-return rate of the other cards.[74] A very small fraction of Cards 2 and 3 were not returned to be counted in the in-person precincts: 0.13% and 0.32%, respectively. Among the absentee ballots, a

73. For instance, in the November 4, 2008 general election, after Chicago residents voted in approximately two dozen races on the front of their ballots, they turned them over to face 71 judicial retention questions on the back of the ballot. CITY OF CHICAGO, SPECIMEN GENERAL ELECTION BALLOT, NOVEMBER 4, 2008, BALLOT STYLE 218, http://www.chicagoelections.com/nov2008ballot/es214.pdf. Harris County, Texas (Houston) voters faced forty countywide partisan elections for various judicial positions, in addition to nine statewide offices and a large number of special district elections, depending on one's residence. HARRIS COUNTY CLERK'S OFFICE, CUMULATIVE REPORT, GENERAL AND SPECIAL ELECTIONS, NOVEMBER 4, 2008, http://www.harrisvotes.net/HISTORY/110408/Cumulative/cumulative.pdf.

74. San Francisco lists these numbers as the "number of voters" in its turnout report and in its Statement of Votes. *See Historical Voter Turnout*, DEPARTMENT OF ELECTIONS, CITY AND COUNTY OF SAN FRANCISCO, http://www.sfgov2.org/index.aspx?page=1670 (last visited June 1, 2010); SAN FRANCISCO 2008 STATEMENT OF VOTES, *supra* note 71. It is likely that some (small) number of voters failed to return Card 1, while still returning one of the other voting cards. This number is unreported by the San Francisco Election Department, so I will follow their convention and treat the number of returned Card 1s as the number of voters.

much larger fraction, 2.6% and 3.3%, of Cards 2 and 3 were not returned. Calculating the non-return rates of Card 4 is more difficult, because races for the Board of Supervisors were not held in every supervisory district in 2008. Thus, the denominator is calculated by adding the number of Card 1's returned in the precincts in which Supervisor elections were held. In total, there were 234,678 voters in the precincts with supervisory races, 127,027 in-person and 107,651 by mail. The in-person non-return rate of Card 4 was 1.4%, compared to 5.5% for absentee ballots.

Stated another way, although absentee voters constituted only 46% of ballots cast in San Francisco in 2008, they accounted for between 77% and 94% of the ballot cards that were not returned for counting in 2008.

The failure to return cards to be counted affects the residual vote rate significantly, especially for "down-ballot" races that appear on every card but the first. This is illustrated in Table 2, which calculates the residual vote rate of each race that appeared first on each of the four ballot cards, along with the residual vote rate for Proposition 8 (gay marriage), which drew national attention. Residual vote rates are calculated using two denominators. The first is the total number of voters for that mode, that is, total turnout. The second is the total number of cards returned for that mode.

TABLE 2: RESIDUAL VOTE RATE DUE TO NON-RETURNED BALLOT CARDS, SAN FRANCISCO, CALIFORNIA, NOVEMBER 2008 GENERAL ELECTION

		In-person residual votes			Absentee residual votes		
	Top office on card	Number	As pct. of cards returned	As pct. of all voters	Number	As pct. of cards returned	As pct. of all voters
Card 1	President	2,075	1.0%	1.0%	2,241	1.3%	1.3%
Card 2	Prop. A1	13,760	6.6%	6.7%	9,261	5.3%	7.8%
	(Prop. 8)	5,671	2.7%	2.8%	4,438	2.6%	5.1%
Card 3	Meas. A	14,547	7.0%	7.3%	8,304	4.8%	8.0%
Card 4	County Bd.	18,561	14.8%	16.0%	8,520	8.4%	13.4%

Source: Department of Elections, City and County of San Francisco, Nov. 4, 2008 *Statement of Votes*, http://www.sfgov2.org/ftp/uploadedfiles/elections/ElectionsArchives/2008/november/SOV_081104.pdf.

The residual vote rate using all voters as the denominator is always higher than the rate using the number of returned cards. Because the number of non-returned cards is so much greater among absentee voters, the two rates diverge most significantly among these voters. Ironically enough, one of the races where the divergence may have

been the most significant was Proposition 8, which passed statewide, restricting marriage in California to opposite-sex couples.[75] Had absentee ballots not experienced the disproportionate "unreturned card" problem, San Francisco County would have likely contributed an additional 1,845 net votes against the proposition.[76] These additional votes would not have swung the result the other way,[77] but they are significant. A much smaller number of residual votes due to voter error determined the outcome of the 2000 presidential election in Florida,[78] so it is certainly possible that non-returned ballot cards could determine whether a ballot measure is adopted.

Not all states have ballots as complicated as California's, but the United States is known for its long ballots. San Francisco provides insights into how the rise of vote-by-mail may either exacerbate or ameliorate the lost-votes problem that emerges because of the length and complexity of American ballots.

CONCLUSION

If 20%, or even 10%, of voters who stood in line on Election Day were turned away, there would be national outrage. The estimates provided by this paper suggest that the equivalent may be happening among voters who seek to cast their ballots by mail, and yet there is nary a comment.

---

75. CALIFORNIA SECRETARY OF STATE, STATEMENT OF VOTE, NOVEMBER 4, 2008, GENERAL ELECTION 60–62 (2008), http://www.sos.ca.gov/elections/sov/2008_general/sov_complete.pdf.

76. The 1,845 figure is arrived at as follows. First, assume that the non-return rate for absentee Card 2 ballots in San Francisco had been the same as the in-person rate, 0.13%. That would have resulted in an additional 4,465 Card 2s being returned. Second, assume that support for Proposition 8 among these additional cards would have been distributed the same way as other absentee voters—28.1% yes, 63.4% no, 2.5% under-vote, and 0.05% over-vote. (The in-person percentages were less favorable to the Proposition—20.9% yes, 76.4% no, 2.7% under-vote, and 0.5% over-vote.) This yields 1,253 votes yes and 3,098 votes no, with 112 ballots under-voted and 2 ballots over-voted. 3,098 (no) minus 1,253 (yes) equals 1,845.

77. *See* CALIFORNIA SECRETARY OF STATE, *supra* note 75, at 13.

78. Walter R. Mebane, Jr., *The Wrong Man Is President! Overvotes in the 2000 Presidential Election in Florida*, 2 PERSP. ON POL. 525, 536 (2004) ("Bush's official, 537-vote margin of victory"); Wand et al., *supra* note 61, at 804 (noting studies of the residual votes showing margins of as low as 152 votes). As an aside, the residual vote rate in San Francisco was generally higher (regardless of the denominator) for absentee ballots than for in-person voting, *except* for the San Francisco County Supervisor races, in which the absentee residual vote rate was actually lower, even when we account for the fact that many more ballot cards were simply not returned. This provides some evidence in favor of the proposition that mail-in ballots may help voters navigate complicated situations, such as the new ranked-choice option in San Francisco.

Three major objections may be lodged against the basic conclusions of this paper. First, it draws conclusions based on poor data. Second, it treats mail voters as innocent victims of election administration and the postal service. Third, even though the schematic view of voting-by-mail presented in Figure 2 is complicated, it is not complicated *enough.*

Each of these objections contains elements of truth, although they are ultimately irrelevant. The analysis is only as strong as the data supporting it. Close attention to the EAVS reveals that even the best source of nationally comparable data about election administration contains important gaps and inconsistencies. Yet while the amount of missing data in the EAVS might make us hesitant to trust its findings about the elections of 2004 and 2006, the 2008 study is substantially complete, at least as far as voting by mail is concerned. Instead, the problems are due to more vexing problems that are difficult to solve, such as inconsistent definitions across states. Similarly, the SPAE is based on self-reporting by voters who may feel pressured to give socially acceptable answers to questions about how they fulfilled their highest civic duty.

Still, even when we acknowledge the shortcomings of the data, it does not undermine efforts to quantify the general scope of the lost vote problem in the vote-by-mail context. Rather, the data shortcomings argue in favor of redoubling efforts to make the best data we have even better.

The question of who to blame for lost votes naturally arises in an analysis such as this. Are lost votes fundamentally the fault of voters? Election administrators? The Postal Service? Over-zealous campaign staff? The clichéd, but true, answer is "all of the above." Unfortunately, because election officials and voters often harbor contradictory expectations about the use of mail-in ballots, it is not always clear how to assign responsibility for lapses in the system. Therefore, it is not always normatively clear what would be the best way to limit lost votes, or how much positive effort election officials should exert to make sure that mail-in ballots are counted.

To take as an example the simplest of reasons for mail votes being lost, consider ballots that are rejected because of a missing signature on the ballot's return envelope. Is this a lost vote we should worry about? If the missing signature is due to the sloppiness or inattention of the voter, maybe not. We are often reminded that the reason for having an educated electorate is so that we can guard our rights ourselves. The mark of an educated voter is attention to details like signing absentee ballots. On the other hand, if the mistake results from

confusing instructions that were written in strict compliance with state law, maybe we ought to be concerned. Research has consistently shown that election materials often seem designed to confuse voters and produce mistakes.[79]

Where vote-by-mail is required, the question about where to assign responsibility shifts attention away from the voters. In these cases, it is hard to argue that voters must fully assume the risks in return for the added convenience of vote-by-mail; election officials must shoulder more of the responsibility to guard against lost votes. If they do not, election officials risk greater skepticism among the public.

There is evidence that such skepticism may already be present in the two states that require residents to vote by mail. In the SPAE, respondents were asked, "How confident are you that your vote in the General Election was counted as you intended?" Nationwide, 69% of respondents answered "very confident."[80] In Oregon, the percentage was 63%; in Washington, it was 52%. Interestingly, when asked whether they supported or opposed laws that mandate that all votes be cast by mail, 34% of Oregon voters opposed such laws, as did 47% of Washington voters.[81] The nationwide level of opposition was 84%. Thus, while most Oregonians and Washingtonians support mandatory vote-by-mail, significant minorities in each state do not.

Among respondents who opposed vote-by-mail in Oregon and Washington, 50% and 44%, respectively, answered that they were very confident that their ballots were counted as intended. In contrast, 62% and 61% of those in the two states who supported vote-by-mail said they were very confident. This relationship between confidence and attitudes about voting by mail did not exist in the rest of the nation. In other words, in Oregon and Washington, attitudes about mandatory vote-by-mail color voters' opinions about whether they trust the vote count, a pattern unobserved in the rest of the nation without mandatory vote-by-mail.[82]

It is possible that in the future, Oregon and Washington voters who oppose vote-by-mail may become less skeptical about votes being counted properly as they gain experience with the vote-by-mail

79. *See generally* MARCIA LAUSEN, DESIGN FOR DEMOCRACY: BALLOT AND ELECTION DESIGN (2007).

80. SPAE, *supra* note 5, at 29–32.

81. *Id*.

82. California has a mandatory vote-by-mail system that only affects a very small number of voters. Therefore, California is treated as a state without mandatory vote-by-mail.

system. However, their skepticism may persist. The point is that a significant minority of Oregon and Washington voters apparently are voting by mail against their wishes; election officials in these states have an obligation to these voters to take extra precautions to ensure that the vote-by-mail pipeline performs as well as the in-person pipeline.

Finally, although some may object because the pipeline metaphor illustrated in Figure 2 is overly complicated, others may argue it is not complicated enough. That is, they may observe that each state handles voting by mail differently, arguing that each state's processes, definitions, and normative standards must be taken into account. Each state has a unique pipeline. Others may note that there are fundamental differences between situations involving for-cause absentee voting (that is, the relatively small number of absentee ballots because the voter is sick, out of town, etc.), permanent absentee voting, and mandatory vote-by-mail. As suggested above, if a state requires all residents to vote by mail, as opposed to regarding vote-by-mail to be purely for the convenience of a limited number of voters, election officials could be considered to have a greater normative obligation to make sure all by-mail ballots are actually counted. Finally, the logistical success of the mail channel hinges more critically on the quality of the registration lists when voters are required to vote by mail or when states establish permanent absentee databases.

The first objection, about nationwide heterogeneity in voting by mail, implies that scientific analysis of the vote-by-mail phenomenon is impossible, a position fundamentally at odds with the underlying premises of this article. More challenging to this paper's analysis is the objection that there are *three* different types of voting by mail that must be analyzed separately. This challenge implies the need for more research. If this additional research helps to better quantify the quality of the voter registration lists that are used across the country, this would all be for the better.

One of the supreme ironies of the butterfly ballot problem in Palm Beach County, Florida in 2000 is that the county Supervisor of Elections deliberately chose this unusual ballot layout to solve a problem she faced. If the extremely long list of candidates on the ballot had been presented traditionally, in a single column, the necessary font size would have been so small that voters with limited eyesight could not have read the ballot. The butterfly ballot allowed the use of a larger, more readable font. The Supervisor was just trying to be helpful. The result was a disaster.

The drift toward more voting by mail is similarly benign. Busy voters demand convenience. Harried election officials demand greater control over the process. Cash-strapped local governments demand greater efficiencies. These are all valid reasons to encourage more voters to use the mail. But what are the hidden costs and trade-offs? As Palm Beach County demonstrated, significant dangers lurk when we make changes to how we vote before we understand all of the ramifications. As voting by mail seems destined to spread even further, it is incumbent upon us to understand better what this means in terms of lost votes.

APPENDIX

This appendix reports a fixed effects regression that estimates the degree to which the residual vote rate for president is related to the prevalence of vote-by-mail in a state. The model is adopted from research conducted by Ansolabehere and Stewart, who studied the relationship between voting technology and residual vote rates for the 1988 to 2004 elections.[83] Here, I start with a simple model that pools together observations from 2000, 2004, and 2008. The dependent variable is the residual vote rate in a county for a particular election year. The independent variables are (1) a series of dummy variables that indicate whether a county used a particular type of voting technology (punch card, paper, mechanical lever machine, optical scan, DRE, or mixed) in an election year, zero otherwise; (2) a dummy variable for each election year; and (3) the logarithm of turnout (i.e., votes cast).[84] A variable equal to the percentage of votes cast in the *state* to which the county belongs in the election year in question is added to this basic setup.[85]

The results of the regression are reported in Table A1.

The results that are relevant to this paper are contained in the first line of the table, which reports the coefficient associated with the percentage of mail ballots cast in the state. Substantively, the effect is relatively small. For instance, a state that increased its vote-by-mail percentage from 20% to 30% (from 0.20 to 0.30) would experience a drop in the residual vote rate of 0.12 percentage points in each county. However, because the same state rate is applied to each county in a state, this variable contains measurement error, since it is being used as an indicator of the vote-by-mail rate in each *county.* Therefore, it is likely that if we could have measured the vote-by-mail rate at the county level, the effect would have been even greater.

83. Ansolabehere & Stewart, *supra* note 61; Stewart, *supra* note 15.

84. The omitted categories for the dummy variables are optical scan voting machines and the year 2000.

85. The state is used, rather than the county, because we do not have nationwide estimates of the number of ballots cast by mail in each county.

TABLE A1: RESIDUAL VOTE RATE AND VOTES CAST BY MAIL, 2000–2008. (STANDARD ERRORS IN PARENTHESES)

Variable	Coefficient
% state ballots cast by mail	-0.012 (0.001)
Voting technology	
Optical scanning	Omitted category
Punch cards	0.0033 (0.0007)
Mechanical lever machine	-0.022 (0.001)
Paper	-0.017 (0.005)
DRE	-0.0042 (0.0007)
Election year	
2000	Omitted category
2004	-0.015 (0.0006)
2008	-0.017 (0.0008)
Log(turnout)	0.074 (0.002)
Intercept	-0.86 (0.03)
N	13,011
R^2	.89
Fixed effects (county)	
Number of categories	4,688
F-test	F(4687,8314) = 13,8 (p<.0001)