EXHIBIT 3

Caltech/MIT Voting Technology Project, Summary Report, Election Auditing, Key Issues and Perspectives (2018)

ELECTION AUDITING

KEY ISSUES AND PERSPECTIVES

SUMMARY REPORT

MIT ELECTION DATA + SCIENCE LAB

Caltech/MIT Voting Technology Project

ELECTION AUDIT SUMMIT
DECEMBER 7-8, 2018

ACKNOWLEDGMENTS

The coordinators of this summary report extend their deepest appreciation to each of the panelists who participated in the Election Audit Summit, without whom this report could not have taken form, and to all those involved in making the event and this report possible, especially Claire DeSoi, Daniel Guth, and Kathryn Treder. They also gratefully acknowledge the many attendees of that summit, who continue to advance the conversation on election administration and auditing in the United States.

DISCLAIMER

The findings, interpretations, and conclustions expressed in this report do not necessarily represent the views or opinions of the California Institute of Technology or the Massachusetts Institute of Technology.

PREFACE

On December 7 and 8, 2018, The Caltech/MIT Voting Technology Project (VTP) hosted the Multidisciplinary Conference on Election Auditing, or "Election Audit Summit," for short, at the Massachusetts Institute of Technology in Cambridge, Massachusetts. The conference was organized by a small group of academics and practitioners from across the United States:

- R. Michael Alvarez (Caltech)
- Jennifer Morrell (Democracy Fund, Election Validation Project)
- Ronald Rivest (MIT)
- Philip Stark (UC Berkeley)
- Charles Stewart III (MIT)

Inspired by the groundswell of interest in risk-limiting audits and other rigorous methods to ensure that elections are properly administered, the conference assembled an eclectic mix of academics, election officials, and members of the public to explore these issues. The essays in this report briefly summarize many of the presentations made at the Audit Summit, while the first chapter ties together the themes of the Summit into one package.

A permanent record of the conference, including video of all the sessions, exists online at https://electionlab.mit.edu/election-audit-summit.


TURN OVER FOR PAGE

CONTENTS

INTRODUCTION: THOUGHTS FROM THE ELECTION AUDIT SUMMIT

CHARLES STEWART III
Massachusetts Institute of Technology

R. MICHAEL ALVAREZ
California Institute of Technology

For nearly twenty years Americans have been faced with questions about the integrity of their country's elections. Challenges to election integrity arise for a variety of reasons, ranging from bad luck, to mistakes, to malicious behavior. The possibility that something might happen in the conduct of an election that might place the correctness of its conclusions at risk have led many to ask the question:

> *"How do we know that the election outcomes announced by election officials are correct?"*

Ultimately, the only way to answer a question like this is to rely on procedures that independently review the outcomes of elections, to detect and correct material mistakes that are discovered. In other words, elections need to be audited.

But how?

The broad topic of auditing elections was the subject of the Election Audit Summit, a public conference held at the Massachusetts Institute of Technology in December 2018. This report presents a summary of the viewpoints presented at that conference. This introductory chapter frames the issues that brought the conference together and presents some summary thoughts about how the practice of auditing can be more thoroughly incorporated into the practice of administering elections in the United States.

From the outset, it should be said that the purpose of the Summit was not solely to share ideas about auditing. A second purpose was to help build bridges between academic researchers and practitioners in the elections field—communities that have of-

ten been at loggerheads over the need for, and proper scope of, election auditing.

We believe the conference was a success, both on the intellectual and community-building fronts. Evidence of that success—at least on the intellectual side—is contained within the covers of this report.

The short papers that follow contain summaries of nearly all the presentations made at the conference. These papers range across a variety of topics, including theoretical and practical issues related to post-election tabulation audits, audits of non-tabulation processes, changes needed in the legal and business environments to accommodate the greater implementation of election audits, and applications of audits to settings outside the United States.

Readers of this report who want more than what is contained in this report are invited to visit the conference website, where slides from the presentations and videos of all panels are located. The URL for that website is https://electionlab.mit.edu/election-audit-summit.

The rest of this chapter provides an introduction to the issues addressed at the conference and in the rest of this report. It is organized around eight questions:

- What are audits for?
- Why do we need audits?
- What do we want to audit?
- Who should do audits?
- Why should people believe the results of post-election audits?
- How often are audits needed?
- How do we get states and counties to implement election audits?
- What can Americans learn about auditing from other countries?

WHAT ARE AUDITS FOR?

Election audits are intended to accomplish two things. The first is to ensure that the election was properly conducted, that election technologies performed as expected, and that the correct winners were declared. The second is to convince the public of the first thing. ***Convincing the public*** that the election was properly conducted and that the correct winners were declared is a core activity of establishing legitimacy in a democracy.

Of course, whether audits actually instill confidence is an empirical question. There is scant research into whether post-election audits in the United States actually serve this legitimating purpose. And, indeed, as Emily Beaulieu's presentation and essay in this report demonstrate, there are cases in overseas elections where the process of election scrutiny has undermined public confidence in those elections.

Still, the purpose of the Summit was to help the nation move ahead in applying higher quality control standards to the conduct of elections. As the presentations and the discussion made abundantly clear, it is insufficient simply to develop fine-tuned and scientifically justified modes of auditing. It is also necessary to develop communications plans, so that the public understands the purpose and processes behind these audits, that the results of any auditing are available to stakeholders and the public, and that the conduct of audits becomes part of the pride a community has in conducting clean elections.

WHY DO WE NEED AUDITS?

Solid evidence can be adduced that elections in the United States have become, on the whole, better-run since the 2000 presidential election highlighted serious shortcomings in vote tabulation, ballot design, voter registration, mail-ballot administration, poling-place operations, and recount laws. At the same time, the shortcomings identified in 2000 have only been incompletely addressed, as new challenges—such as cybersecurity threats and aging voting equipment—have emerged.

One result is that although Americans remain as confident that their own ballots are counted as intended as they were in the early 2000s, their confidence in the vote-count nationwide has fallen steadily since then.[1] Following the 2018 election, approximately 40% of respondents to a post-election academic poll stated that people breaking into election computer systems and voting equipment was either a "major problem" or "a problem."[2] Furthermore, although most Americans are confident that the voting equipment they use is hard to hack, recent criticism of electronic voting equipment has led to a decline in support for those systems.[3]

Moving beyond confidence in the vote and the voting equipment, recent stories emanating from the 2017 and 2018 elections show what happens when attention to all the details that make up an election are not carefully attended to. For instance, in a 2017 state legislative election that determined which party would control the Virginia House of Delegates, it was discovered that 26 voters in that district had been incorrectly assigned to vote in that district.[4] The 2018 U.S. Senate race in Florida may have been determined by poor ballot layout in parts of Broward County.[5] In 2018 there were a number of reported SNAFUS with respect to California's new "motor voter" registration process, in particular regarding how the state's Department of Motor Vehicles was collecting and processing voter registration and re-registration requests, and reports that the registration system itself may have been the target of hackers.[6]

Events like these illustrate why it is important for states and localities to engage in comprehensive programs of auditing and quality assurance for every aspect of election management. Election margins are as close these days as they have been in American history; with partisan polarization, small electoral margins can produce huge policy swings. A lot is riding on getting all the details right, and on communicating that to stakeholders and voters.

[1] Michael W. Sances and Charles Stewart III. "Partisanship and confidence in the vote count: Evidence from US national elections since 2000." Electoral Studies 40 (2015): 176–188; Betsy Sinclair, Steven S. Smith, and Patrick D. Tucker. "'It's Largely a Rigged System:' Voter Confidence and the Winner Effect in 2016," Political Research Quarterly 71, no. 4 (2018): 854–868.

[2] These findings are based on responses to the MIT module of the 2018 Cooperative Congressional Election Study (CCES).

[3] Charles Stewart III and Dunham, James, "Attitudes toward Voting Technology, 2012–2018." Paper presented at the annual meeting of the Midwest Political Science Association, April 4–7, 2019. Available at SSRN: https://ssrn.com/abstract=3363708 or http://dx.doi.org/10.2139/ssrn.3363708.

[4] Laura Vozzella and Ted Mellnik, "Va. election officials assigned 26 voters to the wrong district. It might've cost Democrats a pivotal race." *Washington Post,* May 13, 2018.

[5] Larry Barszewski , Lois K. Solomon , Rafael Olmeda and Skyler Swisher, "Broward recount appears to confirm thousands skipped voting in hotly contested Senate race," *South Florida Sun Sentinel,* Nov. 16, 2018.

[6] John Myers, "Hackers attacked California DMV voter registration system marred by bugs, glitches." *Los Angeles Times*, April 9, 2019.

WHAT DO WE WANT TO AUDIT? (EXPANDING THE CONCEPT OF AN AUDIT)

Formal audits of vote tabulations have been occurring in the United States ever since California mandated post-election audits in the 1960s. By the 2018 election, roughly 30 states required some form of post-election tabulation audit. The typical form of these audits is to require a hand recount of the ballots in a fixed percentage of precincts, usually 1%. These percentages vary considerably across states and, of course, there is still a substantial minority of states with no requirement for post-election tabulation audits at all.

A lot has changed in elections since the 1960s, and this is reflected in advances in the practice and theory of election auditing. Dylan Lynch's contribution to this report discusses the state of auditing requirements across the states, as of late 2018.[7]

Among the audit topics covered in the conference and in this report are the following:

Risk-limiting audits (RLA). An RLA is a post-election tabulation audit in which a random sample of voted ballots is manually examined for evidence that the originally reported outcome of the election is correct.[8] The RLA examines an increasing number of ballots until there is sufficiently strong evidence that looking at all ballots would show that the originally reported outcome is correct. In the limiting case, which is likely to be rarely encountered, all ballots must be examined, as in a recount. If the originally reported outcome is in fact incorrect, there is a pre-specified minimum chance that the audit will correct the result. The correction is made by performing a full manual tally. As its name suggests, an RLA limits the risk of certifying a contest with the wrong winner.

Much of this report concerns RLAs. Philip Stark's essay on "RLAs and Evidence-Based Elections" provides a grounding in the general topic of RLAs.[9] Essays by Neal McBurnett and Hillary Rudy, that arise out of the experience with RLAs in Colorado, provide insightful comments from experienced practitioners.[10] The essay by Jay Bagga and Bryan Byers provides insights into RLA pilots conducted in Indiana — a state that currently has no statewide post-election audit requirement of any sort.[11]

Auditing who gets which ballot. Ballots in the United States are the longest in the world, at least when measured by the number of offices and questions (referenda and initiatives) that appear on the ballots. This is only partly because of federalism and the need to elect officials at three levels of government, federal, state, and local. It is also because state and local governments put offices on the ballot that in other countries would be appointed by the governing authorities. Members of the U.S. House of Representatives and many state and local officials are elected in districts. These

[7] Lynch's presentation at the Summit may be found at 48:08 of the conference video: https://youtu.be/t-cYEVOKWxc?t=2888.

[8] Here, *correct* means that an accurate manual tabulation of all validly cast ballots would give the same winner(s). RLAs can correct tabulation errors, but assume the paper trail is trustworthy; establishing this would take the form of a ***compliance audit***.

[9] Stark's presentation may be found at 26:20 of the conference video: https://youtu.be/t-cYEVOKWxc?t=1580. Ronald Rivest's presentation addressed new developments in the area of RLAs: https://youtu.be/kY5siXsgWUI?t=116.

[10] The panel on the Colorado experience may be found on the conference video here: https://youtu.be/1cbElHGePrA.

[11] The panel on "Looking beyond Colorado" may be found on the conference video here: https://youtu.be/r4jX6CVeBpk.

districts overlap one another in haphazard ways. It is usually the case that the unique combination of offices and questions that appear on the ballot in one precinct are different from the neighboring precinct. To ensure that voters vote on the correct matters—that is, are given the correct ballot—requires careful attention to detail among state and local officials.

The Summit presentation of Michael McDonald, from the University of Florida, powerfully made the case that that voters are oftentimes given the wrong ballot, because the legal definitions of precincts do not always align with geography.[12] These problems can be caused by a number of reasons. For instance, states that define districts using Census Bureau geography often don't account for the fact that the definitions of this geography can change between decennial censuses. Or, addresses may be improperly geo-coded. McDonald makes a strong argument for periodic auditing of the assignment of voters to districts, and that states and localities do a better job at collecting data on district boundaries.

Auditing ballot design. Many of the people who are now academic leaders in the study of voting technology got their start because of the poor ballot design in Florida during the 2000 presidential election. As Whitney Quesenbery, of the Center for Civic Design, points out in her contribution to this report, history has shown that poorly designed ballots, including hand-marked and –verified ballots, can mislead voters.[13] Organizations like the Center for Civic Design have made election officials aware of ballot-design best practices,[14] and these best practices have been disseminated by the EAC.[15] But, even well-intended ballots may hold unanticipated problems, which raises the importance of pre-testing ballots on humans. It also suggests an opening for vendors and civic tech groups to create applications to help test ballot designs against these practices.

Auditing everything else. The essay by R. Michael Alvarez in this report contains the most direct expression of the need to "audit everything."[16] Noting the importance of procedures such as logic-and-accuracy tests and post-election audits, Alvarez writes that, nonetheless:

> *a "logic and accuracy" test of voting equipment used for in-person ballot marking on Election Day or in a vote center doesn't shed any light on the integrity of a jurisdiction's voting-by-mail process, nor does a post-election ballot audit help us determine the integrity or accuracy of a jurisdiction's voter registration process and databases. For a more complete assessment of the integrity of an election in a state or county, we need different and more comprehensive methodologies that can evaluate the performance of the entire election jurisdiction's "eco-system."*

[12] McDonald's presentation may be found at 43:23 of the conference video: https://youtu.be/eFks-xHZH5o?t=2603.

[13] Quesenbery's presentation may be found at 1:06:21 of the conference video: https://youtu.be/eFks-xHZH5o?t=3981.

[14] https://civicdesign.org/fieldguides/

[15] https://www.eac.gov/election-officials/designing-polling-place-materials/.

[16] Alvarez's presentation may be found at 25:59 of the conference video: https://youtu.be/eFks-xHZH5o?t=1559.

In the spirit of this quote, Alvarez discusses a comprehensive assessment project he and his students at Caltech undertook during the 2018 primaries and general election in Orange County, California to assess a variety of election procedures, ranging from mail-ballot transmission to voter registration accuracy, to the monitoring of social media.

WHO SHOULD DO AUDITS?

As audits become regarded as more of a central feature to election administration, an important question emerges: who does the auditing? Presently, post-election audits are typically conducted by the authorities who conduct the elections, with the State of Connecticut being a notable exception.[17] However, it is conceivable (some might even say advisable) that post-election auditing be done by independent third parties.

Taking the lead from the world of finance, there would seem to be advantages to establishing independent election audit boards. Related to this point, Bill Kresse, a CPA who teaches auditing and financial forensics, made the point at the last Summit panel that all states have financial auditors who could supply an instant and willing army of individuals who would be at home in the world of ballot-level audits.[18]

Unfortunately, most jurisdictions seem unwilling to go the next step to establish completely independent auditing procedures, but that does not mean these jurisdictions are unresponsive to the need to "guard the guardians." For instance, as was noted in the Summit panel that reviewed Colorado's experience implementing post-elect auditing, even though election workers were the ones who provided the person-power to audit the results, they did not know which ballots would be reviewed until the audit began. (Furthermore, all stages of the post-election audit were viewed by the public.)

Taking a step away from the formal auditing process, the Summit raised the issue of the public reviewing the results of the election and effectively crowdsourcing an audit. The case of North Carolina's 9th congressional district in 2018 is close to an example of this. In that election, Republican Mark Harris initially appeared to beat Democrat Dan McCready by 905 votes. However, stories quickly emerged alleging absentee-ballot irregularities in Bladen County that were orchestrated by Republican political operative McRae Dowless. After a hearing by the North Carolina State Board of Elections (NCSBOE), the board failed to certify the election because of the irregularities and called a new election.

A lot went into the charges of irregularities and the investigation that ensued. Certainly, one factor that helped the charges gain traction is the fact that the NCSBOE maintains one of the most complete election data sites in the country that include detailed data files that document the request, distribution, return, and resolution of every mail ballot requested in the state. This allowed state investigators, journalists, academics, and citizen enthusiasts to search the record on their own, not only to confirm what officials were finding, but to examine whether there were instances of "District Nine behavior." The 9th CD episode illustrates the importance of making administrative

[17] The Center for Voting Technology Research (VoTeR) at the University of Connecticut. Audit reports are contained at this Web site: https://voter.engr.uconn.edu/voter/audits/.

[18] Kresse's presentation may be found at 41:44 of the conference video: https://youtu.be/LLNX-0eJ9JmU?t=2504.

data from elections available to the public in a usable format, and also illustrates that wrongdoing can sometimes be detected outside of formal post-election auditing programs.

WHY SHOULD PEOPLE BELIEVE THE RESULTS OF POST-ELECTION AUDITS?

A criticism made of states that have no requirements for post-election auditing is that they provide no way for the public to be assured that election outcomes are the correct ones, other than accept election officials when they say, "trust us." The promise of more sophisticated techniques, such as risk-limiting audits, is that they not only require a strict adherence to chain-of-custody and auditing protocols, but they can provide a mathematically rigorous away to quantify how confident we should be that election results are correct.

There are problems with both sets of claims made in the preceding paragraph, of course. As to the criticism that the lack of a formal program of post-election auditing leaves candidates and the public simply to trust election administrators, it can be said that even states that do not require audits have practices that allow for independent observation of polling places, vote counting, tabulation, and canvassing. All states allow close elections to be re-counted, and the results of recounts no doubt inform the public about the quality of vote counting overall.

As a result, voters trust election returns even in the absence of auditing. In the 2016 Survey of the Performance of American Elections, for instance, 91% of respondents from states that required no post-election audits at all stated that they were very confident or somewhat confident that their votes were counted as cast. This contrasts with 90% of respondents from states that required post-election audits. Even without formal audits, voters already express a high level of confidence that votes are counted accurately.

As to the promise that most sophisticated, mathematically rigorous techniques will convince candidates and the public of the veracity of election returns, one only need remember the notoriously poor level of "numeracy" that besets the American public. Even among the numerically sophisticated, understanding how risk-limiting audits work requires a level of statistical knowledge few people possess. As a result, adopting risk-limiting audits risks asking the public to shift blind trust from election officials to statisticians, which, in this age of skepticism about elite expertise, would seem to be a non-starter.[19]

The answer to this conundrum lies in the middle. Even trustworthy individuals make mistakes, and at the very least, rigorous auditing regimes can protect against those mistakes. Beyond this minimalist justification for pursuing better auditing methods, we should remember that some of the most critical electoral crises in recent memory have occurred due to problems that were flying below the radar, unnoticed by the public. The fact that the public at large does not appear to be overly alarmed at the quality of vote-counting does not mean that quality controls are currently adequate.

At the same time, proponents of more sophisticated measures, such as risk-limiting audits, have work to do in explaining how their procedures work and why the public should trust them. At the Summit, the par-

[19] The Summit presentation of William Kresse, cited above, provides further insights into the need to make RLAs "judge-friendly" and "media-friendly."

ticipants took part in an hour-long simulation of a ballot-polling RLA. As the exercise proceeded, it was clear that many of the participants failed to grasp the instructions and got lost in the process. This was a palpable sense to many in the room that moving RLAs from being the preferred method of auditing among the in-the-know experts to being widely accepted among regular citizens still has a long way to go.

HOW OFTEN ARE AUDITS NEEDED?

An important and overlooked issue in the movement toward more and more sophisticated post-election tabulation audits is the question of which elections to audit, and how frequently to audit them. As William Kresse noted in the final panel of the Summit, financial audits do not always cover the same material, nor at the same level of detail every time.

Is there something to be learned in the election auditing realm? Certainly, returns for high-visibility offices, such as U.S. president and state governors, should be subjected to risk-limiting audits every time. But, should every school board race or state legislative seat be equally scrutinized every time? This is where the American "long ballot" raises practical issues regarding post-election tabulation audits. As states become comfortable with risk-limiting audits and anticipate expanding them down the ballot, an important topic to consider is which down-ballot races should be audited, at what frequency, and chosen based on what process?

Furthermore, as already noted, tabulation is not the only election administration detail that should be subjected to auditing and other quality control procedures. How often should an audit of district assignments, of the sort discussed by Michael McDonald, be conducted, for instance?

HOW DO WE GET STATES AND COUNTIES TO IMPLEMENT ELECTION AUDITS? (THE HERE-TO-THERE PROBLEM)

Expanding the prevalence of auditing is a goal shared by a wide variety of election reformers and election administrators. The auditing culture has certainly expanded over the past decade. In 2008, fewer than half of the states, 23, required any sort of post-election tabulation audit. By 2016, that number had grown to 34, plus the District of Columbia.[20]

Of course, with only 34 states currently requiring any sort of post-election tabulation audit, and only three states requiring RLAs, there is still a long way to go before RLAs become ubiquitous.

It is clear that the expansion of election auditing will most likely be a state-by-state affair. Recent legislation introduced in both the House and Senate would mandate that all federal elections include post-election auditing. However, the legislation has stalled, over White House opposition and conflict over states' rights issues. At the same time, concern over cyber threats has caused states without auditing requirements to consider them, and for states with those requirements to investigate strengthening them.

Three of the chapters in this report, by Neal McBurnett, Hilary Rudy, and Jay Bagga and Bryan Byers, provide insights into how RLAs might be expanded, based on observations from one state that has already implemented them (Colorado) and from another state that is exploring the issue (Indiana).

[20] https://elections.mit.edu/#indicatorProfile-PEAR.

Colorado's experience lays out one blueprint for how RLAs might be rolled out on a statewide basis. Colorado, which first implemented statewide RLAs for the election in 2017 (which included local, municipal, and special district elections), has been the pioneer in the field. One factor that aided Colorado's embrace of RLAs is that it was integrated into a transformation of the voting model altogether, to a "vote-at-home" system, where ballots are mailed to all residents and they are then returned either by mail or at official locations.

In transitioning to the new system, Colorado was able to integrate the purchase of new voting equipment into the new auditing regime. With the vote-at-home model relying on the central counting of ballots, the record-keeping load on administrators was made manageable. The wholesale changeover to a new voting model also provided an opportunity to engage a variety of stakeholders into rethinking the election workflow, not just to facilitate RLAs, but also to improve administration overall.

Colorado still has challenges to surmount before the RLA path is completely smooth. Colorado has learned that implementing RLAs is software-intensive, and that the software doesn't write itself. It is still considering how to expand auditing beyond the top-of-the-ticket races. Much work still needs to be done.

(On the issue of software for RLAs, this is yet another example of how the implementation of a common data format for election returns, cast-vote records, and the like is needed to implemented critical reforms in election administration.)

Because of the enormous heterogeneity in terms of size, scope, and timing of elections in the United States, there is unlikely to be a one-size-fits-all auditing system for the entire nation, or even for local jurisdictions within states. The Summit heard examples of pilot projects in Colorado, New Jersey, California, Rhode Island, and Indiana that seemed to be successful in giving state and local officials information about how RLAs might be adapted to their own settings, and getting them comfortable with the ideas overall.

Although the purpose of conducting rigorous election audits is to assure the public that an election was conducted accurately, as well as to provide convincing evidence to losers that they in fact lost, the critical stakeholder in determining whether a state mandates audits, and whether those audits are rigorous, is local election officials.

Local election officials bear most of the administrative burden of implementing election audits, especially post-election tabulation audits. The typical local election office is small and runs on a tight budget. Anything that increases work without an obvious benefit to local officials will be met with howls of opposition from these local officials who, by the nature of their job, have the ears of those officials.

Bringing local officials on board to advocate for rigorous post-election tabulation audits requires more than simply explaining how they are done and why they are important. Showing how they are done, through the pilots mentioned earlier, seems to be one mechanism for opening up local officials to the feasibility of audits.

To the degree that explanation is important, one factor seems to trump all others: Under most circumstances, once the requisite systems are in place, RLAs require

less work after the election than do traditional fixed-percentage audits. With most elections decided by comfortable margins, RLAs will often require only the examination of a few hundred ballots in most cases.

Even when elections are close, the number of ballots examined under RLAs will likely be less than the number examined under more traditional methods.

WHAT CAN AMERICANS LEARN ABOUT AUDITING FROM OTHER COUNTRIES?

The attention paid to post-election auditing in the United States has tended to focus entirely on American elections, despite the fact that assessing the veracity of elections has long been a major issue in the administration of elections in other countries, as well as an important subject of scholarship. Observation of elections by international observers, such as the Organization for Security and Cooperation in Europe (OSCE), has been regarded as an important element in reducing corruption in countries that are considered problematic.

One challenge that auditing has to face in the developing world is that of sovereignty. Developing nations, trying to come out under centuries of colonial control, are keen to develop their own election apparatuses. This puts a premium on countries doing their own in-house audits. At the same time, lack of capacity often leads these countries to rely on international experts to supply statistical expertise.

Political scientists have piggy-backed their research on top of these efforts, to develop rigorous techniques to document how election-observation regimes can reduce corruption.[21] Many of these efforts can be grouped under the heading of "election forensics."[22]

In both her presentation to the Summit and her essay in this report, Emily Beaulieu, a leading scholar of international election observation and corruption, offered both optimistic and cautionary observations about election auditing internationally.[23]

The 2010 election in Haiti is one success story, where scrutiny of precincts with above-average (and in some cases, above-100%) turnout overturned the results of the preliminary election ended up with a result in which the original third-place finisher in the preliminary was allowed to go into the final round, ultimately winning. On the other hand, recent experience in elections in Afghanistan, Honduras, and Kosovo illustrate how audits alone can be insufficient to ensure that clear evidence of elections being stolen by fraud will result in new elections being demands, or consequences being felt for the perpetrators.

Whether these comments apply directly to the American case can be questioned. However, one point made by Beaulieu does seem applicable: Using audits to detect and correct election fraud will be more effect if citizens already have trust in elections. If they do not, then the results of audits will

[21] R. Michael Alvarez, Thad E. Hall, and Susan D. Hyde, eds. *Election fraud: detecting and deterring electoral manipulation*, Washington: Brookings, 2009.

[22] Allen Hicken and Walter R. Mebane Jr., "A guide to election forensics," USAID Research and Innovation Grants Working Paper Series, July 28, 2017, https://pdf.usaid.gov/pdf_docs/PA00MXR7.pdf.

[23] Beaulieu's presentation may be found at 40:55 of the conference video: https://youtu.be/kY5siXs-gWUI?t=2455.

become just another source of conflict over which competing political factions compete. At its worst, audits have the potential to deepen suspicious and cause a decline in voter confidence.

The techniques discuss by Walter Mebane have been applied to both American and non-American elections.[24] Unlike the techniques based on election observation, Mebane's methods primarily rest on the analysis of aggregate election data, matching that data against comparison statistics, such as turnout data, previous election results, and demographic data.

Mebane's presentation and essay return us to the point that all methods of auditing do not have to rest on an examination of individual ballots, as proposed by RLAs. Certainly, ballot-based audit methods are statistically superior to other methods, but ballots are not always available. In those cases, less powerful methods may be powerful enough to convince the public, local election authorities, and/or the international community that something was amiss in a nation's election.

NEXT STEPS AND MOVING FORWARD

The conference was a success, especially as it brought election officials, academics, and other stakeholders in election auditing together for two days of productive conversation and interaction. In a number of cases, conversations between academics and election officials, begun at the conference, have sparked subsequent conversations and perhaps even eventual collaborations.

We would like to see more collaborations between election officials and academics on election auditing, and to that end, we will start by proposing that conveningsl like this conference be held more regularly.

> *There is a strong and pressing need to continue to build trust and communication between election officials and academic researchers, in particularly when it comes to election auditing.*

There is a growing interest among academics in different areas of the election auditing process, and facilitating that interest by keeping academics in contact with election officials is important. Many election officials are interested in post-election ballot audits and comprehensive election auditing, but lack the time and statistical expertise to implement election audits on their own, so giving them the opportunity to connect with academics who might help them is important.

There is also a need to continue to facilitate the scientific study of election administration and technology, in particular as it relates to election auditing. The academics interested in election auditing have made significant progress in recent years developing auditing techniques and tools to perform different types of election audits. However, the research initiatives often exist within academic disciplines, and there is a need for more interdisciplinary communication about election auditing. So we also believe that there should be periodic workshops and conferences for the academics interested in studying election audits and integrity, which will help grown and strengthen scientific knowledge of auditing practices and methodologies.

[24] Mebane's presentation may be found at 14:23 of the conference video: https://youtu.be/kY5siXs-gWUI?t=863.

CURRENT STATUS OF POST-ELECTION AUDITING AND RECOUNT PRACTICES

RISK-LIMITING AUDITS AND EVIDENCE-BASED ELECTIONS

PHILIP B. STARK
University of California, Berkeley

No way of counting votes is perfect. Every system—manual or electronic—can make mistakes. Electronic systems are particularly vulnerable to misconfiguration, bugs, hacking, data loss, etc.

If there is a trustworthy paper record of voter intent, reported outcomes can be checked against that paper trail by suitable audits. But an audit is no better than the paper trail is relies on. If there is no paper trail, there is no way to verify whether the reported results are correct. If the paper trail is not voter-verifiable (e.g., the paper record produced by some ballot-marking devices), an audit cannot verify who won. If the paper trail is not trustworthy, the audited outcome is not trustworthy.

The key elements for ensuring reported election outcomes are trustworthy can be summarized with *"5 Cs"*:

- ***Create*** durable, trustworthy record of voter intent. Hand-marked paper ballots are best for voters who have the dexterity and visual acuity to use them; ballot-marking devices (BMDs) are helpful for voters with disabilities that make it difficult or impossible to mark a ballot by hand.[1]
- ***Care*** for the paper record. The chain of custody should be verifiable; there should be two-person custody rules, ballot accounting, good seal protocols, etc.
- ***Compliance*** audit. Auditors need to establish whether paper trail is trustworthy, through ballot accounting, checking against pollbooks and voter registration databases, reviewing chain of custody logs, video security surveillance, checking eligibility determina-

[1] See, e.g., https://www.stat.berkeley.edu/~stark/Preprints/bmd19.pdf

tions, checking signature verification, etc.
- » ***Check*** reported outcome against the paper (using a risk-limiting audit).
- » ***Correct*** the reported outcome if it is wrong (by conducting a full hand count).

A risk-limiting audit (RLA) is any procedure such that:

> *If an accurate full hand count of the paper would find different winners than were reported, the procedure has a known minimum chance of requiring a full hand count.*

The ***risk limit*** of a RLA is the largest possible chance that, if the reported outcome is wrong, the audit won't correct it. Here, ***outcome*** means the electoral outcome: the winner or winners, not the exact vote tallies.[2]

Many state audit laws go into great detail to specify how many ballots (or precincts) to audit. That focus is misplaced, in my opinion: the starting sample size is not important. What matters is when you *stop* auditing.

A RLA does not stop auditing until and unless there is strong statistical evidence that a full hand count would simply confirm the reported outcome—that it would be a waste of time. If it does not find strong evidence that the reported outcome is correct, a RLA progresses to a full hand count to set the record straight. If the outcome is wrong but the paper trail is trustworthy, a RLA has a known minimum chance of correcting the outcome. RLAs do not involve assumptions about voter preferences, nor about how or why errors might occur.[3]

Risk-limiting audits can be used with a broad variety of approaches to drawing random samples of ballots or groups of ballots, allowing audits to be tailored to the logistics and equipment of individual jurisdictions. The ***sampling unit*** can be a group of ballots or an individual ballot. The sample can be ***stratified*** or ***unstratified***. The sampling units can be drawn with equal probability, or with different probabilities (for instance, sampling with probability proportional to an error bound is useful when the sampling unit is a group of ballots). The sample can be drawn with replacement, without replacement, by Bernoulli sampling, by Poisson sampling, or many other methods.

Once the sample is collected, there are two main approaches to analyzing the data to determine whether the audit can stop. ***Poll-***

[2] In general, it is impossible to get the tallies right to the last vote without a full, accurate hand count. But getting the electoral outcome right seems like the minimal acceptable standard. If we do not audit enough to determine with high confidence who won, we are not auditing enough.

[3] Bayesian audits are not, in general, risk-limiting audits. Bayesian audits assume voter preferences are random, with a known distribution. They answer the question, "if the current election had been selected at random from a particular hypothetical population of elections, then, given the audit data, what is the probability that the current election is one of those hypothetical elections for which the reported result is correct?" The "upset probability" for a Bayesian audit is in general much smaller than the risk that a Bayesian audit will not correct the outcome if the outcome is wrong. There are examples where the "upset probability" is 5 percent, but there is a 55 percent chance that the Bayesian audit will not correct a wrong outcome.

RLAs and Bayesian audits both require a trustworthy paper trail, random sampling, etc. The biggest operational difference between them is the rule for deciding whether the audit can stop—but they answer very different questions. In particular, a Bayesian audit might not have a large chance of correcting the outcome if the outcome is wrong.

ing audits use the audit data directly. They are like exit polls, but instead of asking voters how they voted, they get that information directly from the ballots. Unlike voters, ballots have to reply (and have to reply honestly). The only information a polling audit needs from the voting system is the reported winner(s).

Comparison audits use the audit data together with detailed information exported from the voting system. They compare how the equipment tabulated randomly selected ballots with how humans would tabulate the same ballots. Comparison audits are like checking someone's reported travel expenses: First, add up the reported expenses to check the math. Second, spot check the reported expenses against the underlying paper receipts to make sure the expenses were reported accurately.

Similarly, a comparison audit starts with data exported from the voting system: vote subtotals for individual ballots or groups of ballots. First, auditors check that the reported subtotals add up to give the overall reported results. Second, auditors draw ballots or groups at random and manually check whether the reported subtotals were correct. If the audit finds convincing evidence that the tabulation was accurate enough that the reported winner must have won, the audit can stop.

Any jurisdiction that uses paper ballots (and keeps track of the ballots) can perform a ***ballot-polling risk limiting audit***; no special voting equipment is needed. However, the efficiency of the audit—measured by the number of ballots that must be inspected before the audit can stop—does depend on the capabilities of the voting system. If the voting system can report how it interpreted individual ballots (i.e., if it can report a cast-vote record for each ballot) in such a way that the corresponding physical ballot can be identified and retrieved for manual inspection, then a ballot-level comparison audit is possible. When the reported electoral outcome is correct, ballot-level comparison RLAs generally require inspecting far fewer ballots than ballot-polling RLAs, especially when the margin is small.

COMPREHENSIVE ELECTION PERFORMANCE AUDITING

R. MICHAEL ALVAREZ
California Institute of Technology

Over the last two decades, the technology and administration of American elections have become hot topics in public discourse. No longer is the conduct of elections a matter of discussion among a small group of academics, nor is it a relatively obscure area of state and local public administration. Discussions about the integrity of recent elections have dominated headlines and been central topics of debate in the 2016 presidential and 2018 midterm elections. From allegations of cyber-attacks on election administration and database systems in recent years, to debates about election malfeasance in some states, there is more discussion of election security and integrity than ever before.

Given the public focus on the integrity of elections, the question that continues to arise is how does the public know that an election has been conducted with a high degree of integrity? How can we be sure that there weren't successful attempts to hack voter registration databases, to stuff ballot boxes, or to impersonate vote-by-mail voters? Furthermore, how can we confirm that proper procedures were followed in all vote centers and polling places, that voter rights were maintained, and that in the end, all ballots were counted as intended? Confirming the integrity of an election is no simple matter.

In the past, many election jurisdictions used certain forms of auditing approaches to attempt confirmation that aspects of their election process and voting systems were functioning as expected. For example, in many jurisdictions, pre-election "logic and accuracy" tests have been conducted on sampled voting machines to ensure that they record votes as they should, and in some states, certain types of post-elec-

tion ballot audits are used to provide some assurance that ballot recording and tabulation may have functioned as expected in an election. States and counties continue to improve and innovate with respect to these practices, for example, by testing and implementing "risk-limiting audits." These newer forms of post-election auditing can provide statistical confirmation that ballots were tabulated correctly.

These types of auditing procedures are important. But they only can help to assess the integrity of some aspects of election administration. For example, a ***"logic and accuracy"*** test of voting equipment used for in-person ballot marking on Election Day or in a vote center doesn't shed any light on the integrity of a jurisdiction's voting-by-mail process, nor does a post-election ballot audit help us determine the integrity or accuracy of a jurisdiction's voter registration process and databases.

> *For a more complete assessment of the integrity of an election in a state or county, we need different and more comprehensive methodologies that can evaluate the performance of the entire election jurisdiction's "eco-system."*

In addition, the analyses that serve as the justification for that assessment should be transparent and available to the public.

Working in collaboration with the Orange County Registrar of Voters (OCROV), Neal Kelley, and his team, our research group at Caltech pilot-tested an ambitious set of comprehensive election performance auditing methodologies in the 2018 primary and general elections in Orange County. In our pilot project, we wanted to develop and deploy auditing and performance measurement tools that would be both relevant and actionable for the OCROV, as well as timely and transparent for stakeholders and the public. We also sought, as much as possible, to focus on election performance data that were already being generated by OCROV ("trace data") or on data that we could produce and analyze independently of OCROV; this strategy would minimize the amount of time and resources that OCROV needed to devote to this pilot project in the course of a busy and complex election cycle, while also producing an independent evaluation of the administration of the 2018 primary and general elections in Orange County, California.

Orange County was chosen because it is an ideal location for a pilot project like this. First off, the OCROV and his team have an established record as innovators, and prior to our collaboration were already generating a great deal of data. Secondly, election administration in California is changing rapidly; for example, in 2020 Orange County will be moving away from the traditional neighborhood voting model towards universal vote-by-mail and voting centers; starting our pilot project in 2018 in Orange County provided an important baseline for longitudinal analysis of these changes and their potential implications for voter confidence. Third, Orange County is a large (approximately 1.5 million active registered voters) and diverse election jurisdiction in Southern California. Finally, in 2018 we expected to see many hotly contested elections, in particular for U.S. House seats in Orange County—helping us gauge which performance measures might be more relevant and important in competitive elections.

For both the 2018 primary and general elections in Orange County, we built and implemented a number of different performance methodologies:

1. mail ballot transmission and return tracker;
2. in-person observation studies of early and Election Day voting;
3. post-election precinct-level turnout and candidate forensics and anomaly detection analytics;
4. post-election voter surveys (general election);
5. voter registration auditing;
6. observation and study of OCROV's post-election risk limiting audits; and
7. social media monitoring.

Reports and summaries of these election performance methodologies for both election cycles are available on the project's website (https://monitoringtheelection.us).

Our 2018 Orange County comprehensive election performance auditing project has yielded a great deal of important analytical data, and a number of conclusions for our continued research on developing this approach for providing a data-driven evaluation of election integrity.

First, of the methodologies that we developed and deployed in 2018, we believe that the most useful for election officials is our voter registration auditing methodology: we developed an approach that flags anomalous changes in the voter registration data for further investigation. Second, producing timely and actionable performance measurement is crucial for both election administrators and the public; during the immediate post-election canvass period is when concerns about election integrity arise, and it is imperative for maintaining voter confidence that performance measures and analyses be up-to-date and available to the public in the days and weeks following Election Day. Third, some of our methodologies, like social media monitoring and turnout/candidate vote share forensics, have considerable promise as election performance tools, but they require continued research and further development. Finally, and most crucially, the 2018 election cycle in Orange County was quite competitive—our comprehensive election performance audit provided substantial data-driven evidence that these elections were administered with integrity, and that voters should be confident that their votes were tabulated as they intended.

Looking forward, our research group will continue to test, develop, and implement the methodologies used in 2018 in Orange County in future elections. Our immediate plan is to scale our comprehensive election performance auditing approach to cover Southern California in the 2020 presidential election cycle. By adding additional counties, we will build important variance in context that will give us greater opportunity to compare performance across jurisdictions within the same state. This will also move us closer to being able to deploy comprehensive election performance auditing for most American states (covering Orange and Los Angeles Counties in 2020, for example, would include approximately 7 million registered voters in our analysis, a greater population of registered voters than in all but the largest American states. And this will give us important longitudinal data for Orange County, which we can use to track election integrity over time.

ELECTION VALIDATION PROJECT

INCREASING TRUST IN ELECTIONS THROUGH AUDITS, STANDARDS, AND TESTING

JENNIFER MORRELL
Democracy Fund

This is a unique time in election administration. Never before have there been so many resources, tools, and information for election administrators. In fact, towards the end of my time as a local election official, I felt overwhelmed by all of the ideas and information pouring in. I could not keep up with reading it all, let alone implementing it.

A serious discussion about the role that election audit standards might play in validating our elections must also include a discussion about strategies for taking an idea as broad as audits and creating something that will actually be read and used by state and local election officials. One way to start is by clearing away the clutter.

Think about the strategy you would follow to teach someone a new skill, such as baking a loaf of bread. You would not start by asking them to read all of the literature ever written about bread making. You start by forming a basic understanding of the important principles. Next you demonstrate what tools are available, which ones are necessary and which ones are optional but might make the task easier, followed by providing a recipe or guidelines that include the necessary ingredients and instructions. Finally, you have someone with experience demonstrate how the process works and act as a mentor.

We can do the same for risk-limiting audits by providing a practical guide for state and local election officials that covers the following:

- Terms and definitions
- Policy considerations
- Voting equipment and technology
- Implementation considerations

Implementing RLAs becomes more likely when practical guidelines are coupled with templates for ballot organization and storage, pilot audits to provide hands on experience, auditing software, and pairing up states who are or plan to conduct risk-limiting audits to help mentor each other.

It is also important to recognize the gaps that need to be addressed before a concerted push is made for wide-spread implementation. With respect to post-election, risk-limiting audits, there is a need to create better communication about the process, a universal audit tool or software to assist with the audit, better ways to retrieve ballots (especially for ballots scanned at the polling location), ways to ensure the audit material does not compromise voter anonymity, and deciding if it is appropriate to create national standards. These are all areas where outside organizations can contribute possible solutions.

Not every state may be able to or have a desire to implement risk-limiting audits. Resistance to change is universal and there may be reasons to make a careful and gradual move towards risk-limiting audits. It does not need to be an immediate destination but can be viewed more as a path with steps leading to it. Some steps that will help ease the transition toward RLAs:

- Strong collaboration among state and local election officials
- Making the RLA terms and definitions a regular part of election vocabulary
- Creating documented voter intent guidelines
- Developing a well-crafted plan for ballot storage and organization
- Requiring precise ballot reconciliation
- Implementing dates and deadlines to accommodate time for a post-election audit prior to certification
- Basing the number of ballots selected for audit on the contest margins
- Using dice or similar method to randomly select the ballots, precincts, voting machines, etc. that will be audited
- Purchasing a voting system that produces a voter verifiable paper ballot and cast vote record

Most of the focus has been on robust, post-election audit of the vote tabulation equipment, such as risk-limiting audits. But it begs the question why only audit and test the voting equipment? Why not audit and validate other critical components of the election system? Auditing how votes are tabulated plays an important role in validating the outcome of an election. However, it is only one of several elements in the election system that needs to be examined.

A risk-limiting audit provides modest benefit if you cannot provide evidence of a solid chain of custody from the beginning of an election to the end, for both ballots and voting equipment.

As we start to think about incorporating audit principles into election administration, consider other critical components in the election system that can be audited:

- Voter registration databases
- Voter district and precinct assignments
- Security procedures (physical and cybersecurity)
- Pre-election testing of voting equipment (focused on paper ballots)
- Ballot reconciliation and chain of custody

- Ballot layout and design
- Resource planning and allocation (enough equipment, supplies, and people to meet demand)

State and local election offices increasingly employ or contract with a number of experts. This includes individuals with a professional background or expertise in law, communications, data analysis, project management, and cybersecurity. It may be time to consider including auditing and quality control professionals into that mix.

Audits have played a role in U.S. companies for many years. Many of those same auditing standards and definitions can be applied to elections. For instance, an audit can be defined as a systematic, independent and documented process for obtaining audit evidence and evaluating the evidence objectively to determine the extent to which the audit criteria are fulfilled.

We can also apply standard classifications for the type of audit being performed. A ***product audit*** is the examination of a particular product (such as a voting system) to evaluate whether it conforms to requirements and performance standards. A ***process audit*** evaluates an operation against predetermined instructions or standards and asks the questions: *Did the operation conform to the standards? Are the instructions for the operation effective?* The audit typically examines resources (equipment, materials, people), environment, and methods (procedures and instructions).

A ***system audit*** verifies that applicable elements of the system are appropriate and effective and have been developed, documented, and implemented in accordance with specified requirements. For example, an election system audit would determine if the election system conforms to state and federal policies and requirements.

When an election audit is implemented the standard phases of an audit are also applicable and include:

1. preparation;
2. conducting the audit;
3. reporting and feedback; and
4. closure.

Most of the workload that will fall to local election officials is in the preparation stage.

Audits can be both internal and external. In an internal, or ***first-party audit***, an organization measures its strengths and weaknesses against its own procedures or against external standards. In an external, or ***third-party audit***, the audit is performed by an independent audit organization and free of any conflict of interest.

We can also apply the benefits of conducting an audit to elections:

- Detects voting system errors
- Provides accountability to voters
- Deters fraudulent activity
- Limits the risk of certifying incorrect outcome
- Assures votes were counted & reported accurately
- Provides feedback for process improvement

We are at our best when we face complex challenges together.

We no longer have the luxury of working in our individual silos. The solutions to successfully planning for, conducting, and auditing an election will come from a diversity of professional backgrounds collaborating on research and exploring new ideas. This includes thinking about how technology and solutions already employed in other sectors can be used to improve election administration. We need each other—election officials, technology experts, academics, policy makers and election advocacy groups—all working together to build public trust in elections.

POST-ELECTION AUDITS: THE STATE OF THE STATES

DYLAN LYNCH
National Conference of State Legislatures

Much of what we do at the National Conference of State Legislatures is group and categorize state policies to provide a national overview. This allows state legislators and others to easily compare policies across the country. For post-election audits we have organized state policies into three general categories.

AUDIT TYPES ACROSS THE COUNTRY

1. Traditional. Traditional audits look at a fixed percentage of voting districts or voting machines and compare the paper record to the results produced by the voting system. Regardless of the closeness of a race, they will always count the same number of ballots. Although the way these states conduct audits is similar, differences do exist. For example:

- *What is counted:* In Alaska, a randomly selected precinct in each house race is selected and 5 percent of ballots cast in the district are audited. In Nevada, the audit looks at 2 percent or 3 percent of voting machines, depending on county size.
- *Who conducts the audits:* In New Mexico, an independent auditor in hired by the secretary of state (SOS) and oversees the audit. In Pennsylvania, the local boards of elections do it. And in Vermont, the SOS conducts the audit.
- *When is the audit conducted:* In Florida, the audit is conducted following the certification of election results while Illinois conducts theirs before the canvassing of ballots.

2. Risk-limiting audits (RLAs). By their simplest definition, RLAs are "statically based audit techniques." Washington's Revised Code §29A.60.185(c) further defines and explains a risk-limiting audit as:

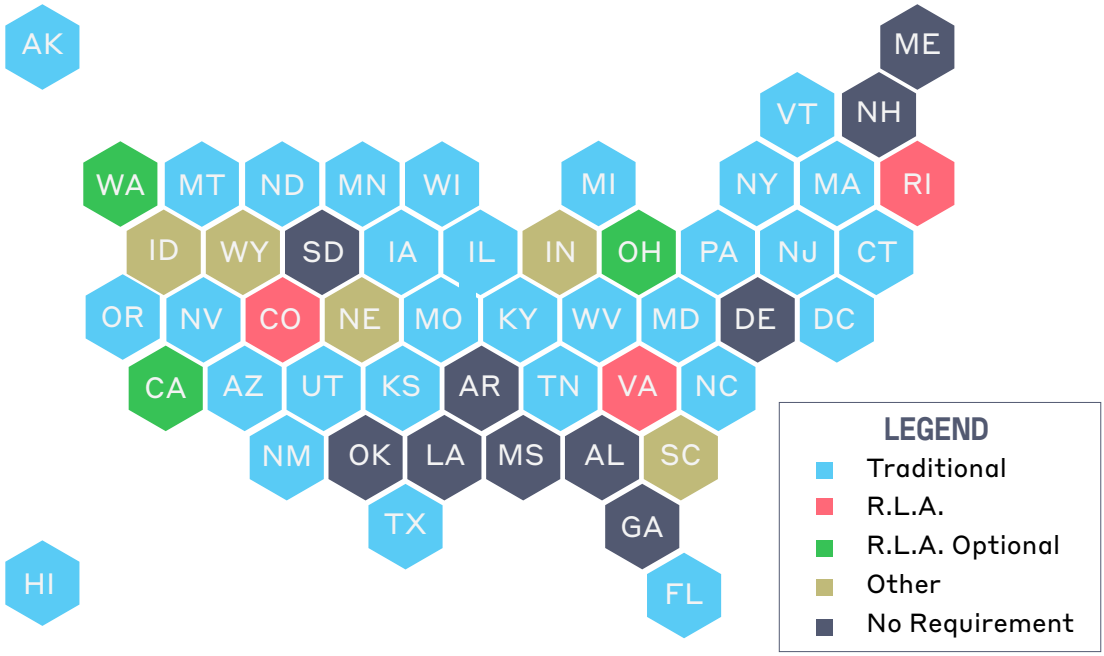
c) A risk-limiting audit. A "risk-limiting audit" means an audit protocol that makes use of statistical principles and methods and is designed to limit the risk of certifying an incorrect election outcome. The secretary of state shall:

(i) Set the risk limit. A "risk limit" means the largest statistical probability that an incorrect reported tabulation outcome is not detected in a risk-limiting audit;

(ii) Randomly select for audit at least one statewide contest, and for each county at least one ballot contest other than the selected statewide contest. The county auditor shall randomly select a ballot contest for audit if in any particular election there is no statewide contest; and

(iii) Establish procedures for implementation of risk-limiting audits, including random selection of the audit sample, determination of audit size, and procedures for a comparison risk-limiting audit and ballot polling risk-limiting audit as defined in (c)(iii)(A) and (B) of this subsection.

Only three states have enacted RLAs in statute and only Colorado has fully implemented a statewide RLA. However, many other pilot programs, generally done at the local level, have been completed across the country. In addition, three states are currently in the process of phasing-in RLAs (California), or allowing local jurisdictions the option of conducting a RLA, (Ohio and



Washington). More details about these states can be found below.

3. "Other." This, again, can vary. In some states, like Idaho, an audit is triggered only when a recount is required. Some states have a procedural audit, which is not necessarily an audit of ballots, but is instead an audit of the processes and procedures used in the election.

As the map on the previous page shows, traditional audits are by far the most common route states take, with a smattering of states that use RLAs and "other" types of audits. However, a few interesting things are going on in some states. New Mexico and Oregon are categorized as traditional audits, but really enact a tiered system.

In a tiered system, the margin of victory of a race dictates how many ballots are audited. The closer the race, the more ballots that get audited. Meanwhile, in September 2018, California passed Assembly Bill 2125, which stipulated that in lieu of the traditional audit, beginning with the spring primary of 2020, a county can choose to conduct a risk limiting audit. In addition, Ohio and Washington have "optional" R.L.A.s. In Ohio, a 2017 directive from the secretary of state recommended RLAs be used by counties but did not mandate them. In Washington, county auditors can choose among three post-election audits methods, with an RLA being one of the three.

Lastly, there was a decent amount of auditing action taken in Michigan in 2018. During the 2018 general election, the state ran an RLA pilot program in three cities. In addition, on the ballot in that election was Ballot Measure 18-3, a proposed state constitutional amendment that would have established many election policies, including a post-election audit, as rights in the state constitution. The measure passed, and the legislature enacted a traditional post-election policy in late December 2018.

YEAR	STATES WITH BILLS	BILLS INTRODUCED	BILLS ENACTED
2011	10	16	2
2012	9	14	2
2013	10	14	0
2014	11	21	4
2015	10	18	0
2016	10	22	2
2017	16	23	4
2018	21	48	10
TOTAL		**176**	**24**

LEGISLATION

Moving on to state legislation, post-election audits have been a topic throughout the states for many years. Starting in 2011, there was a consistent number of bills until an increase in 2016.

In total from 2011 to 2018, 176 post-election audit bills were introduced, with only 24 being enacted.

TOPICS FOR CONSIDERATION

Risk-limiting audits have certainly gained traction not only in legislatures, but also in the media. Yet, there are some things that may need consideration or discussion if state legislatures want to enact RLAs.

The Good:

- ***Implementable.*** We know that RLAs are possible. Colorado put in a lot of the legwork, but if other states look to implement RLAs, the process could become more streamlined and efficient.
- ***Integrity, Security and Confidence.*** RLAs provide integrity, security and confidence to the outcome of an election. Confidence in the democratic system is vital to the maintenance of our system governance.
- ***Cost/Time Savings.*** RLAs, as a system, could provide cost and time savings compared to traditional audits

The Bad:

- ***Not Everyone is Ready.*** Not every state is ready for RLAs. Colorado did put in a lot of work and time and money into their effort. Not every state can do so at this time.
- ***Technology Considerations.*** Many states may not have the equipment or technology necessary to efficiently conduct an RLA.
- ***Initial Time/Cost.*** Because this is somewhat unchartered territory, states may need to put a lot of initial work into creating a system that works with their structure and adheres to their laws.

The Ugly:

- ***Legislation for the Sake of Legislation.*** Many would agree that legislation should be based on informed decision making. Passing legislation just to have it in the books can lead to legislative and administrative issues that could ultimately sink the policy if a state is not ready.
- ***Complexity.*** If you asked the average voter if they wanted to learn about "statistically based audit techniques," how do you imagine that going? Because the point of RLAs is to provide confidence and security, that requires everyone having a basic understanding of the how and why. What is probably most detrimental to RLAs is the public believing RLAs operate like "magic".

CONCLUSIONS

Post-election audits aren't going away. In fact, audits are increasingly spreading into other areas of election administration, such as voter registration logs, voting equipment and other election procedures. It is possible that elections are becoming more audit centric. Still, we are fond of calling states "the laboratories of democracy" and as such, states will continue to make policy decisions that work for and suit them best.

NEW STATISTICAL APPROACHES TO AUDITING VOTE TABULATION AND RECOUNTS

SAMPLING WITH *K*-CUT, AND BAYESIAN AUDITS

RONALD L. RIVEST
Massachusetts Institute of Technology

INTRODUCTION

We present two new tools for the election auditor's toolbox that may provide increased efficiency, or additional flexibility in complicated situations.

Post-election statistical tabulation audits proceed by sampling cast paper ballots at random, and then figuring out what the sampled ballots tell you about the correctness of the reported election outcomes.

We propose a new procedure, ***sampling with k-cuts***, for drawing random samples of cast paper ballots for such statistical post-election audits. This procedure eliminates the need to count down to a specified pseudo-random position in a stack of ballots, by performing instead a sequence of ***k "cuts"*** (like cutting a deck of cards) and then taking the top ballot. Sampling with *k*-cuts works well with ballot-polling audits, but doesn't work at all for ballot-comparison audits (which need to find a ballot with a specified imprinted ballot ID).

We also propose the use of ***Bayesian audits*** for determining whether to accept the reported election outcome or to continue the audit (by examining a larger sample). Bayesian audits are an alternative to "***risk-limiting audits***," and are of particular interest when no risk-limiting audit method is available or feasible.

SAMPLING WITH *K*-CUT

> *How can one pick a ballot "at random" from a given stack of ballots?*

The usual method is to generate a random ballot number (using cryptographic methods), and then to count down in the stack

to the ballot at that position. This method is tedious and error-prone when the stack is large.

Our proposed alternative, *k*-cut, works as follows to randomly sample a single ballot from a stack of ballots. (This procedure can be repeated to sample multiple ballots.)

- Pick a suitable small integer *k* (we suggest using *k* = 6).
- Perform *k* "cuts," where each cut removes a random fraction of ballots from the top of the stack and places those ballots at the bottom of the stack.
- Pick the ballot now at the top of the stack as your selected ballot.

Although each cut may individually be slightly non-uniform, repeating the operation *k* times smooths out the statistics to give acceptably uniform results.

A detailed analysis of the k-cut method appears in Mayuri Sridhar's Master's thesis,[1] and in Rivest and Sridhar, 2018.[2] Further research is underway to show how to improve (decrease) *k*, by making use of randomness "hints" when picking a cut size; decreasing *k* would provide further efficiency improvements.

The *k*-cut method has been used successfully in several pilot election audits (Indiana, Michigan, Rhode Island); going forward it is an attractive choice for use in actual (ballot-polling) audits.

[1] Mayuri Sridhar. Optimization for Election Tabulation Auditing. MIT EECS Master's Thesis. February 2019. https://mayuri95.github.io/main.pdf

[2] Mayuri Sridhar and Ronald L. Rivest. *k*-cut: A Simple Approximately-Uniform Method for Sampling Ballots in Post-Election Audits. Proceedings Financial Cryptography, February 2019, Fourth Workshop on Advances in Secure Voting. https://fc19.ifca.ai/voting/program.html

BAYESIAN TABULATION AUDITS

A ballot-level statistical post-election tabulation audit keeps drawing cast paper ballots and manually examining them, until it is determined that the sample drawn provides sufficient support for the reported outcome, or until all cast paper ballots have been examined.

There is more than one way to use statistical methods to define a "stopping rule" for the audit. "Risk-limiting audits" are one way; Bayesian audits are another (although there is some overlap).

A risk-limiting audit asks "What is the current risk if we stop the audit now?", and stops the audit if this risk is below a pre-defined ***risk limit***. Here risk is defined as the (conditional) probability that if the reported outcome is incorrect that the audit would accept the reported outcome as correct.

A Bayesian audit asks "What is the `upset probability'?"—the probability that examining all of the cast paper ballots would show the winner to be different than the reported outcome—and stops the audit if this upset probability is below a pre-defined ***upset probability limit***. Bayesian methods are used to define the upset probability as the posterior probability of an upset, given the sample and given a prior probability on ballots.

> *These definitions appear very close, but there are nonetheless significant differences.*

For one thing, risk may be viewed as a worst-case definition, while upset probabilities are more of an average-case definition. Given the adversarial nature of elections, a risk-limiting audit may in general be a more

appropriate choice than a Bayesian audit (and we recommend using risk-limiting audits whenever possible).

Also, risk and upset probabilities appear not to be on the same scale: a risk-limit of five percent may correspond (roughly) to an upset probability limit of one-half of one percent or so (ten times smaller). Determining the relationship between risk and upset probability is an active research area. Achieving risk below a certain risk-limit is not the same thing as achieving an upset probability below a certain upset probability limit. One can't naively switch back and forth between the two models; the definitions mean different things.

Nonetheless, risk-limiting audits and Bayesian audits are ***highly compatible***. Their high-level structure is identical: drawing increasingly large samples until a stopping rule says to stop. A Bayesian audit can easily "piggy-back" on a risk-limiting audit, using the same sample data, and computing upset probabilities while the risk-limiting audit is computing risk. This can provide additional comfort and confirmation that the reported outcome is likely to be correct.

How does one implement a Bayesian audit? The following outline sketches one approach (based on "Polya's Urn") for computing an upset probability:

1. Draw an initial random sample of the cast paper ballots; examine each sampled ballot manually to determine the voter's intent.
2. "Pretend" to examine the remaining ballots, but instead of drawing new ballots randomly to examine manually, look at randomly chosen ***previously examined*** ballots again (with probability proportional to the number of times each ballot has been previously examined).
3. Compute the winner of the set of all (really drawn and pretend-drawn) ballots.
4. Repeat steps 2–3 many times. The fraction of time that the reported winner loses is the "upset probability."

The Bayesian audit stopping rule says to stop the audit if the estimated upset probability is below the pre-defined upset probability limit.

The Bayesian method is quite simple. One nice feature is that it works at the ballot-level, and is independent of the voting method used. The same approach works for plurality, approval voting, instant-runoff voting, etc. All that is needed is a method to determine the winner (step three above) for a set of ballots, and one must have such a procedure anyway just to run an election!

It should be noted that Bayesian methods require the definition and use of a "***prior probability distribution***" giving the assumed likelihood of seeing any particular ballot prior to seeing any sample data. In this use of Bayesian methods for post-election audits, defining such a prior is much easier than for many other applications of Bayesian methods, since the only purpose of the prior here is to ensure that a priori all ballot choices are judged equally likely.

The prior is weighted to ensure that it "steps out of the way" when the sample data arrives. In the above sketch, a typical prior would be effected by including one extra ballot for each candidate in the sample as part of step one.

One may also easily extend Bayesian methods to handle ballot-level comparison audits, or various forms of stratified audits (where some strata are ballot-polling and some are ballot-comparison).

Details omitted here; see Rivest[3] for an expanded treatment, and see the original Rivest and Shen[4] paper for more variations.

Bayesian methods have been implemented and used in various pilot audits; typically as a "free add-on" to a risk-limiting audit. For example, in the December 2018 pilot audit of a proposition on the ballot in Rochester Hills, Michigan, Kellie Ottoboni and Philip Stark computed (for a sample of 76 ballots with 50 Yes votes and 26 No votes) a risk of 2.1%, while Mayuri Sridhar computed an upset probability of 0.3%.

Again, these numbers are not directly comparable, but both are significantly below their pre-defined limits, so both the risk-limiting audit and the concurrent Bayesian audit (on the same sample data) confirmed the reported election outcome.[5] Bayesian audits have been used in a number of other pilot audits as well.

In summary, Bayesian methods provide additional tools in the auditor's arsenal, and may in some cases (for complex voting methods where no risk-limiting audit method is known) be the only tools available. For typical plurality elections, Bayesian methods are probably best as a possible concurrent "second opinion" on the correctness of the reported election outcome.

[3] Ronald L. Rivest. Bayesian Tabulation Audits: Explained and Extended. arXiv https://arxiv.org/abs/1801.00528 (2018-01-02)

[4] Ronald L. Rivest and Emily Shen. A Bayesian Method for Auditing Elections. Proceedings 2012 EVT/WOTE Conference. https://www.usenix.org/conference/evtwote12/workshop-program/presentation/rivest

[5] Scott Borling, Tina Barton, Chris Swope, Virginia Vander Roest, Mayuri Sridhar, Kellie Ottoboni, Liz Howard. Eds. Ron Rivest, Jerome Lovato, Philip Stark. A Review of Robust Post-Election Audits: Various Methods of Risk-Limiting Audits and Bayesian Audits. Brennan Center for Justice and Verified Voting. 2019 (to appear).

ELECTION FORENSICS BEYOND AUDITS

WALTER R. MEBANE, JR
University of Michigan

INTRODUCTION

Election forensics[1] can be useful in at least two circumstances: when effective audits are not feasible or even not possible; when problems going beyond what an audit may detect are suspected of affecting an election. I define election forensics as:

> *the use of statistical methods to determine whether the results of an election accurately reflect the intentions of the electors.*

Most audits use statistical methods, but the range of methods considered as included in election forensics is wider. Election forensics methods may focus on trying to provide evidence that an election outcome is correct, but they may also—or instead—focus on suggesting why election returns are as they are, pointing out anomalies, revealing possible fraudulent manipulations or intimidations, explaining outcomes as due to routine strategic behavior or identifying areas that should be investigated further using more richly informed hands-on methods.

Election forensics were first developed to apply in cases where paper records of votes are not available, so the question was whether anything at all could be done to create evidence regarding an election's

[1] Mebane, Jr., Walter R. 2008. Election Forensics: The Second-digit Benford's Law Test and Recent American Presidential Elections. *In The Art and Science of Studying Election Fraud: Detection, Prevention, and Consequences*, ed. R. Michael Alvarez, Thad E. Hall and Susan D. Hyde. Washington, DC: Brookings Institution.

credibility in such circumstances.[2] By now many have contributed methodologies, and productive scientific controversies and research abound.[3] Election forensics are very far from perfect. Indeed, it's best to think of them as nascent. For example, methods based on the second significant digits of vote counts have been shown to be ambiguous: they respond both to normal political activities (strategic behavior, district imbalances, special mobilizations, coalitions) and to frauds.[4] Methods that examine the last digit of vote counts can be fooled if malefactors have sufficient control over the numbers.[5]

In my presentation I briefly reviewed four recent applications of election forensics analysis and discussed one extension being developed to incorporate new kinds of data. The applications refer to some kinds of statistics that are available via the Election Forensics Toolkit (available at http:// election-forensics.ddns.net:3838/EFT_USAID) and some others.

HONDURAS 2017

Polling station data from the 2017 Presidential Election in Honduras shows signs of frauds that may have affected enough votes to determine the election outcome. One Toolkit indicator ("P05") suggests that votes for the winning party may have been manipulated. Estimates from a likelihood finite mixture model[6] suggests fraudulent vote counts are present in about thirteen percent of polling stations, and that the overall number of fraudulent votes is greater than the difference in votes between the winner and the second-place candidate.

US 2016 WISCONSIN AND MICHIGAN

Briefly reviewing results reported more completely in Mebane and Bernhard,[7] I describe evidence that the voting technologies used in places that had the votes recounted in these states appear to have treated candidates Trump and Clinton symmetrically. In Wisconsin, there was a full recount and in Michigan there was a partial recount. In Wisconsin a variety of methods were used to recount the ballots, including both hand

[2] Wand, Jonathan, Kenneth Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Jr., Michael Herron and Henry E. Brady. 2001. "The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida." *American Political Science Review*

[3] e.g. Myagkov,. Ordeshook, and Shaikin in *The Forensics of Election Fraud: With Applications to Russia and Ukraine*; Levin, Cohn, Ordeshook, and Alvarez. in "Detecting Voter Fraud in an Electronic Voting Context: An Analysis of the Unlimited Reelection Vote in Venezuela;" Mebane, in"Fraud in the 2009 Presidential Election in Iran?;" Deckert, Myagkov, and Ordeshook in "Benford's Law and the Detection of Election Fraud;" Beber and Scacco in "What the Numbers Say: A Digit-Based Test for Election Fraud;" Klimek, Yegorov, Hanel; and Thurner in "Statistical Detection of Systematic Election Irregularities;" Mebane in "Election Forensics: Frauds Tests and Observation-level Frauds Probabilities;" and Ferrari, McAlister, and Mebane in "Developments in Positive Empirical Models of Election Frauds: Dimensions and Decisions."

[4] Mebane, Jr., Walter R. 2013a. "Using Vote Counts' Digits to Diagnose Strategies and Frauds: Russia;" and Mebane, Jr., Walter R. 2014. Can Votes Counts' Digits and Benford's Law Diagnose Elections? In *The Theory and Applications of Benford's Law*, ed. Steven J. Miller.

[5] Mebane, Jr., Walter R. 2013b. "Using Vote Counts' Digits to Diagnose Strategies and Frauds: Russia;" and Verena Mack and Lukas F. Stoetzer. 2019. "Election fraud, digit tests and how humans fabricate vote counts—An experimental approach." Electoral Studies 58(1):31–47.

[6] Mebane, Jr., Walter R. 2016. "Election Forensics: Frauds Tests and Observation-level Frauds Probabilities."

[7] Mebane, Jr., Walter R. and Matthew Bernhard. 2017. "Effects of Voting Technologies and Recount Methods on Votes in Wisconsin and Michigan;" and Mebane, Jr., Walter R. and Matthew Bernhard. 2018. "Voting Technologies, Recount Meth- ods and Votes in Wisconsin and Michigan in 2016.

and machine tabulations. In Michigan, all recounts were by hand. The diversity of voting and recount methods in the one case and incomplete coverage in the other makes it impossible say to anything with great confidence about whether voting machines were hacked in these states. Records from Wisconsin show that officials logged a variety of problems beyond "Voting machine error," and such errors were associated with some of the highest mean differences between the recounted and original vote count in Wisconsin wards.

KENYA 2017

Polling station data from the 2017 presidential election in Kenya produces extensive signs of frauds when analyzed using Toolkit methods.[8] This is the election that was annulled by the Kenyan Supreme Court. The most important challenge for election forensics analysis of Kenyan election data is that voting is very strongly polarized along ethnic lines. The strategic coordination around ethnicities creates patterns in vote count data that can look like frauds to statistical tests. The appearance of frauds is greatly reduced but not eliminated when the data are analyzed separately by county.

Unfortunately measures of ethnic composition are not available for Kenyan polling stations or other localities, but experts agree that many counties are ethnically much less heterogeneous than is the whole country. The analysis done separately by county still suggests there were many irregularities and thousands of fraudulent votes, possibly benefitting—in different counties—both leading candidates.

RUSSIA 2016

Polling station data from Russian national elections from 2000 through 2016 suggest that frauds in these election-type events have gotten worse over time.[9] A change appears to occur between the 2003 Duma and 2004 presidential elections: vote manipulations designed to "signal" that votes are being manipulated for United Russia are evident via the "P05" statistic from 2004 on. "P05" statistics suggests that turnout was being manipulated in all the elections. Estimates from the finite mixture model suggest the manipulations follow two basic patterns, with either substantial or very substantial proportions of votes being manufactured. While the "substantial" variant of the vote-manufacturing mechanism appears to have been in place during the 2016 presidential election, the proportion of polling stations in which frauds had effect is estimated to be the highest that year across all the elections. The number of fraudulent votes estimated to have been counted in that election dwarfs the other elections—the number is more than double the next highest estimated number of fraudulent votes.

TWITTER ELECTION OBSERVATORY

Based on 6.5 million keyword-filtered Tweets taken from the Streaming API, we used machine classification tools to extract 315,180 election incidents reported by 215,230 Twitter users during October 1–November 8, 2016: the initial implementation of a Twitter Election Observatory.[10]

An ***election incident*** is a report of a personal experience with situations such as lines

[8] Mebane, Jr., Walter R. 2017. "Anomalies and Frauds(?) in the Kenya 2017 Presidential Election."

[9] Kalinin, Kirill and Walter R. Mebane, Jr. 2016. "Worst Election Ever in Russia?"

[10] Mebane, Jr., Walter R., Patrick Wu, Logan Woods, Joseph Klaver, Alejandro Pineda and Blake Miller. 2018. "Observing Election Incidents in the United States via Twitter: Does Who Observes Matter?"

or wait-times of various lengths, success or difficulties voting, success or difficulties registering, participation in election-day, early or absentee voting, and more. At the time of the presentation, these incidents had been identified at more than 12,000 distinct locations in the continental United States (plus many locations abroad).

Using a measure of "presidential campaign partisan association" constructed using word embeddings derived from Twitter users' self-descriptions, we observe "communication silos" in which users tend to report their incidents to other users who have similar partisan associations. Users with differing partisan associations tend to report different kinds of incident—compared to "***clinton / hilary / hillaryclinton / strongertogether / democrat***" users, "***trump / donald / realdonaldtrump / maga / republican***" users are:

1. less likely to report unspecified line length incidents or long lines,
2. less likely to report unspecified polling place incidents, neutral polling place incidents or success voting,
3. less likely to report unspecified registration incidents or neutral registration incidents but more likely to report registration problems, and
4. less likely to report unspecified electoral system incidents or neutral electoral system incidents.

The reporting differences seem to reflect biases more than differences in real experiences. To implement the second version of the Observatory, we have used the Streaming API to collect about 65 million keyword-filtered original Tweets during October 1–November 6, 2018.

NEW STATISTICAL TECHNIQUES IN INTERNATIONAL PERSPECTIVE

EMILY BEAULIEU BACCHUS
University of Kentucky

Advances in the statistical techniques available for use in elections audits present exciting opportunities for those seeking to promote election integrity and strengthen the democratic process. At the same time, international experiences with elections provide several important points of consideration for election administrators, observers, and other actors providing technical assistance and support, even in the context of U.S. elections.

This presentation stressed the emphasis on a holistic definition of an election audit. Extending from this definitional orientation, the presentation emphasized: best practices for conducting audits, the critical importance of data availability, and the need for an appreciation of, and sensitivity to, the broader political context in which the audit takes place.

For international democracy promoters, an election audit is more than a recount.

While a recount can be employed to confirm results, an audit is a broader investigation (which may involve a recount) initiated in response to accusations of fraud, in order to verify the integrity of the election and establish whether election results should be considered legitimate. Thus, while both a recount and audit aim to determine whether election results are "correct," the charge of an audit extends beyond a narrow recount of ballots, to take a fuller picture of actions undertaken by key stakeholders and to evaluate the impact of those actions on election outcomes.

In the interest of accomplishing a more holistic audit, that assesses election integrity

rather than simply election outcomes, international practitioners have several ideas about best practices—all of which emphasize the need for election administrators to plan in advance for the possibility of an election audit.

First, any audit should ideally be undertaken by the country's own election management board (EMB). The emphasis on domestic election administrators performing an audit serves two purposes:

1. it prevents key stakeholders in the election (parties and candidates) from manipulating the election process (or being perceived as doing so).
2. It limits perceptions among stakeholders and the electorate of outside interference in the election process.

The main caveat to this advice is the recognition that some of the advanced statistical techniques being employed in election audits today may require expertise from third parties outside the country—such possibilities, however, should be planned for and made transparent well in advance of an actual audit.

Given that an election audit will include activities beyond a vote recount, the procedures and standards for the audit must be established well in advance. Audits that occur without such understandings in place are likely to do little to improve confidence in elections, or produce compelling evidence of election fraud.

In order to be legitimate, audit procedures should be constructed from a country's election law. These procedures must anticipate the kinds of issues that are likely to arise and trigger an audit, and make explicit plans to evaluate the election on the basis of those issues.

Finally, all the processes associated with the audit must be transparent and clearly communicated to key stakeholders (who should be able to observe, but not conduct, the audit). In particular, any advanced statistical methods that are employed in the course of the audit should be clearly communicated to stakeholders. Otherwise such evidence may undermine the credibility of the audit, which ultimately undermines the goal of evaluating the integrity of the election.

EXAMPLES:

- ***Venezuela Presidential Recall Audit (2004):*** The Carter Center worked with Venezuela's election commission, used simple statistical analyses, and addressed subsequent criticisms of the audit. This audit upheld the results of the election as valid.
- ***Haiti First-Round Presidential Audit (2010):*** The Organization for American States examined vote counts from outliers (high turnout, high vote margin) in a random sample of polling stations, in response to fraud accusations. Determined that 2nd and 3rd place in the first round should be reversed, affecting competition for the presidency.

While international democracy promoters stress these best practices for conducting an election audit, it is critical to acknowledge that a holistic audit requires data to be successful. The data made available to auditors should be extensive and include: voter registers, voter turnout, and election results data. Further, the turnout and results data should be disaggregated, ideally to the level of individual polling places. Finally,

data on the time that counts are taken, reported, and incorporated into higher levels of aggregation should also be collected and made available to auditors. If the country is using an electronic voting system, the codes used to transmit and aggregate results must also be made available to auditors.

This level of data availability requires first, that election management bodies establish procedures to collect the data and second, that election law permits such data to be made available to auditors in the event of an audit. In many countries today (e.g. Nigeria, Guyana, Cambodia) release of disaggregated results is not permitted once results have been aggregated and submitted to the central election management body. Further, issues and suspicions can be exacerbated when countries use electronic methods (or a mix of paper and electronic, as in Kenya) to aggregate election results. Finally, temporal data is critically important because demographic patterns can confound the meaning of irregularities that appear to be geographically distributed.

EXAMPLES:

- ***Kenya (2017):*** This election is an example of the limited conclusions that can be drawn when data are insufficient. Election results were aggregated both electronically and via paper ballot, but only electronic results were available for analysis. Furthermore, this is a country where demographic patterns make it difficult to draw conclusions from geographic irregularities, and time-stamped data would be useful.
- ***Honduras (2017):*** Because data was available at the polling location level, and by time of report and aggregation, this is a case where the Organization of American States was able to identify clear anomalies suggesting that there had, in fact, been fraud committed.

Photo credit: Annie Bolin

Finally, when undertaking an election audit, auditors and those overseeing the process must remain cognizant of the broader political context in which the election takes place. Again, if one thinks of an audit as analogous to a recount, in that it seeks simply to provide an accurate picture of what happened in the election, then the broader political context does not seem relevant.

Returning to the international definition of an audit, however, we know that these undertakings are motivated by accusations of fraud and electoral malfeasance. The kinds of elections where such accusations arise, are very often elections in political systems that are fraught with instability and the potential for conflict and/or democratic backsliding.

In such cases, those conducting the audit must way the value of providing accurate results against the value of maintaining peace and stability.

EXAMPLES:

- ***Afghanistan (2014):*** After accusations of fraud the United Nations supported an audit that included a full recount of all polling locations. The results of this audit, however, were never released publicly or to key stakeholders. Instead, the audit motivated a compromise between the winning and losing candidate for power sharing. This is an example where stability in the political systems was valued over electoral accuracy.
- ***Kosovo (2010):*** In this election an automatic, remedial audit was implemented where all tabulation sheets were checked against officially recorded results to ensure accuracy. Results of this audit were released publicly, showing that results from almost 50% of polling stations were inaccurate and should be invalidated. As a result, voter confidence in elections declined sharply, suggesting a case where accuracy was prioritized over enhancing confidence in democratic elections.

In sum, the international election perspective cautions us that election audits must be holistic, and well-planned in advance, and that data must be fine-grained, prolific, and made available to auditors. Finally, auditors are advised to consider the broader political contexts and must be alert to the potential for goals such as accuracy to be in tension with other goals such as political stability and the promotion of voter confidence in elections and support for democracy.

WHAT COLORADO CAN TEACH US ABOUT POST-ELECTION AUDITS

RISK-LIMITING AUDITS: LESSONS LEARNED

NEAL MCBURNETT
Independent Consultant

In 2017, Colorado became the first state to regularly conduct risk-limiting audits (RLAs). Colorado's successes are grounded in 15 years of multi-partisan efforts to promote and pilot election auditing. Here are some of the lessons I've learned along the way.

For more background, and links to a wealth of material, I encourage you to explore The Colorado Risk-Limiting Audit Project (CORLA), available online.

> ***One of the clearest lessons is that pilot audits with input from people experienced with risk-limiting audits are enormously helpful and highly recommended. The whole community learns from pilot audits.***

BEST PRACTICE: BALLOT-LEVEL RLAS

Colorado has demonstrated that with good systems, processes and data, you can do ballot-level risk-limiting audits which limit the risk that tabulation errors or attacks result in getting the wrong outcome. This can be done at scale, in hundreds of contests, in dozens of counties, and across overlapping districts in a state.

They can also be done efficiently. Colorado audited less than ten thousand ballots statewide. Besides the fully risk-limiting audits, simultaneous "opportunistic" audtis can gather evidence on and report risk levels for all the rest of the contests.

Furthermore, Colorado's new statewide system is among the most cost effective and best for auditing: central-count scanners with BMDs available for accessibility.

RESOURCES AVAILABLE

These highly efficient ballot-level RLAs can be done with equipment from multiple vendors. In 2015, four vendors presented and piloted systems that could do ballot-level comparison RLAs in 2015. They were all central count scanning systems, from Dominion, Hart, ClearBallot and ES&S

Colorado funded the development of software to help manage the audits, which is now the open source ***ColoradoRLA***. This system continues to be enhanced, and can be used for free by any jurisdiction, with support available from multiple organizations.

IMPORTANCE OF RLAS

> *There is widespread, transpartisan consensus on the need for both paper ballots and audits.*

An early example was in 2003, when four local parties (Republican, Democrat, Libertarian and Green) supported a joint consensus in Boulder, Colorado. An excellent overview of the modern case is in the National Academies report from 2018: ***Securing the Vote: Protecting American Democracy***, from The National Academies Press.

While we've made huge steps forward, there is still much to do. Why is it taking so long to adopt robust audits?

- Elections are increasingly complicated
- You can't easily audit the data you've got
- You can't easily get the data you need

This underscores why it is critical to support and adopt the Common Data Standards work by the EAC / NIST VVSG-Interoperability task force.

COMMON DATA FORMATS

We need format standards! See a helpful overview presentation by John Wack: ***Overview of VVSG-Interoperability Common Data Formats*** (two presentations).

Common data formats are published or in-the-works for several use cases. Election Results reporting (SP 1500-100) is used in OH, NC, LA County. Other states are in progress. The Election Log Export CDF will soon be published as SP 1500-101. The Voter Records Interchange CDF is slated for review by VR vendors, to be published as SP 1500-102. I have seen initial use in OH and by OSET.

The Cast Vote Records CDF schema should be published soon as SP 1500-103. The ongoing development and documentation of election process business models and voting method descriptions is also very beneficial.

EVIDENCE PRESENTED AND CHECKED

Audits which are conducted by elections officials should also be highly accessible to the public, and the critical inputs to and results of the audit should be shared openly. Otherwise, audits may be convincing to officials, but leave losing candidates and the public without enough evidence to go on.

A document presenting details on what the public should have access to is available at ***Public RLA Oversight Protocol***, by Stephanie Singer and Neal McBurnett, 2017. Briefly, the elements it covers are: Chain of Custody, Tabulation, Manifest, Commitment, Random Selection, Ballot Card Retrieval, Ballot Interpretation and Data Entry, Ending the Random Selection and Examination of Ballot Cards, Hand Count, and Audit Conclusions Affect Outcomes.

COLORADO AUDIT RESULTS

The ColoradoRLA software includes an `rla_export` tool to provide necessary data for Oversight Protocol in csv/json formats.

`rla_report` software is in progress to interpret the exported data, confirm that the right ballots were selected, and check the risk level calculations, to help implement these oversight steps. This code will also be open source, and verifiers should be encouraged to check it and/or implement their own oversight processes and code.

In its recent audits, Colorado has shared more useful data on its audits, in more useful ways, than probably any other jurisdiction. Officials can be very proud of their results. Officials with access to all the audit data, including the Cast Vote Records (CVRs) etc., can be more confident in the outcomes of more contests than anywhere else in the country, and certainly more efficiently than anywhere else.

CONVINCING OTHERS OF ELECTION OUTCOMES

Unfortunately, while this is much more transparency than in the past, losing candidates and the public still encounter several crucial holes in the oversight protocol. Some summary data is not available yet, principally because due to an unusual confluence of challenging circumstances, the state is still wrestling with ballot anonymity issues which have limited the availability of the original CVRs to the public. That means the public can't check tally totals, and can't check ballot interpretations in real time, or sometimes at all.

We give kudos to the amazing ongoing accomplishments by both the state and the counties under very challenging circumstances, and look forward to resolving the various obstacles to full transparency.

A model for that sort of transparency has already been seen in the audits in Boulder CO in 2008, which, before the audit, successfully generated auditable data. In some cases that required merging small sets of ballots into larger sets, all to be audited together, in order to eliminate anonymity concerns. See Boulder County 2008 General Election Audit for the data and open-source software for those batch-comparison audits.

More detail on relevant challenges and good solutions is available at ***Preserving Anonymity of Cast Vote Records***, by Mark Lindeman, John McCarthy, Neal McBurnett, Harvie Branscomb, Ron Rivest, and Philip Stark, 2017-08-03.

DISCREPANCY INVESTIGATIONS

Detailed reporting on discrepancies in Colorado's audits is still in-progress. But it is evident that there are still some instances of errors in data entry. To avoid that, the software should inform the Audit Board that there was some sort of discrepancy right after it has been officially entered (and after preserving a record of that official entry). That would help with discrepancy investigations, provide much more useful and actionable quality control feedback, and enhance trust in the process on all sides.

REMAINING CHALLENGES

The software needs enhancements in reporting convenience and analysis. It should make it easy to view discrepancies, and risk levels for opportunistically audited contests. That is particularly challenging for the wide variety of districts, each involving samples taken in a variety of counties. The software should also automatically generate an "Audit Center" web site with full data for the public.

The software should be further modularized for use with external risk-level calculation modules, covering additional auditing methods like SUITE, Bayesian RLAs, etc.

We need new approaches to handle in-precinct/vote center scanners, which randomize ballots and/or CVRs. They complicate the process of matching paper ballots with CVRs.

We need upgraded support for batch-comparison audits, which yield risk reduction which is predictable, easy to plan for and easy to understand. We should also provide better support for ballot-polling audits, though they can be unpredictable and impractical for some of the most interesting contests with tight margins.

We should foster collaboration between clerks, privacy experts, and tool-smiths around preserving anonymity, especially for the complicated situation in Colorado. And we should audit more systems involved in elections: voter registration, signature verification, envelope sorting, ballot reconciliation etc.

TARGETED AUDITS

Often in any given election, public attention is focused on particular circumstances. Random selection of ballots to audit is essential for good risk reduction, but we should also be prepared to directly address specific concerns and unusual circumstances.

We should encourage candidates and the public to identify additional interesting ballots to target for auditing. They could be chosen based on analysis of the CVRs, based on mark density data, or even based on ballot images.

PUBLIC ENGAGEMENT IN VERIFICATION

Finally, we should promote more public participation in audits. We could print ballot tracking pages with QR codes, and provide an app that public observers could use to photograph ballots along with the tracking-sheet QR codes. That could assist the public in conducting their oversight, and facilitate sharing of a series of confidence-inducing tweets like *"I verified the votes on this ballot."*

ACKNOWLEDGEMENTS

I'm deeply grateful to Paul Tiger for being the first to encourage me to get involved in election verification back in 2002. A long list of colleagues since then have offered expertise, insights, enthusiasm, and comradery, including Joe Pezzillo, Paul Walmsley, Ron Rivest, Philip Stark, Harvie Branscomb, and Hillary Hall and her Boulder County team, The Election Verification Network has been invaluable in all of this work.

The Colorado legislators who helped us pass laws in 2005 and 2009 to require audits deserve much credit. I've been incredibly impressed at the dedication of the Colorado SOS staff and Clerks! The Free & Fair Team that took on the daunting challenge of signing up to write the initial ColoradoRLA software under incredible constraints of time and resources deserve my eternal gratitude. They went beyond the call of duty.

And the Democracy Works team that has continued to improve the software, enhancing the user interface and digging deep into the internals to re-work it for multi-county contests has been incredible also.

Updated versions of this narrative report will be available at http://bcn.boulder.co.us/~neal/elections/audit-summit.html

SUMMARY OF RLA IMPLEMENTATION: TRAINING AND CHANGE MANAGEMENT

HILARY RUDY
Office of the Colorado Secretary of State

RELATIONSHIPS AND COMMUNICATION

In our experience, one of the strongest factors to successfully deploying any major project or new system is building and maintaining strong working relationships between the state and counties, and with any outside vendors. The state leverages these relationships to gain buy-in and build support for the project, and to identify "cheerleaders" who will help test and pilot.

> *We've found that good communication is central to building trust and a strong relationship.*

A critical factor of success in project implementation is that counties must be comfortable calling with questions or training needs. It takes time and work to build trust, and face-to-face interactions are essential in the process. Our office goes out regionally to provide training and we attend the clerks' association conference. We also go out to visit the counties in their offices to understand their unique processes and specific challenges. Having established relationships helps our team plan resources and focus their energy on counties that may need more one-on-one training time.

We've found that it's critical for our team to be responsive when the counties call in with questions or need one-on-one training time. The team reaches out regularly to the counties to ensure they feel a level of comfort calling in. We layer the communication to counties and focus on ensuring that our messages are clear and effective. The voting systems manager sends regular emails to the counties listing upcoming deadlines and critical information. We include the same information in the weekly newsletter that we send to counties. And leading up to

the mock risk-limiting audit before the 2018 general election, we also highlighted in the weekly newsletter one piece of functionality in the audit system that had changed. We conduct a weekly county support call during the election period, and we include the upcoming deadlines and other critical information in that weekly agenda.

Before the first statewide audit, the team developed clear, comprehensive written documentation about the process and technology. This documentation also included how to use the ancillary systems, such as how to hash a document or access the SFTP site. In addition to the state's step-by-step technical documentation, we worked with the counties to develop a county playbook outlining process best practices for small and medium counties. We work with the counties to update the documentation before each audit to ensure it's accurate, comprehensive, and understandable.

Trust and communication are critical to implementing large statewide projects—we can continue implementing big changes when there are simultaneous implementations that place a strain on the counties' resources. For example, in the 2018 general election, we implemented a significant change to how the statewide contest is audited. It worked and the audit ran smoothly because counties called with questions and the team spent a lot of time working one-on-one with them.

LAYERED TRAINING

Another factor of our success is approaching training as an iterative process. It's essential to establish and maintain a safe learning environment where everyone feels comfortable asking questions openly and honestly. If counties are honest about their challenges, the resulting discussion is more productive in terms of identifying real workable solutions.

We begin with general concepts and terms to help counties understand the legal and philosophical framework. Then we cover why the process or project is being implemented and any way in which the counties will benefit from the change. Helping counties understand 'why" is fundamental to gaining buy-in. It's also important to train on legislation or rules that are changing as well as the legislative and rule process, and to recognize that legislative and rule changes may need to be tackled iteratively to avoid unintended consequences.

We've found that it's important to cover the technical steps of the process, in this case the audit, at several points during the training cycle. But it's most critical during the process discussions and the hands-on training. As I discussed in the communications section, comprehensive guides for both software and processes are important. Guides should be step-by-step manuals that include screenshots and explanations of ancillary systems like the hashing tool. And they should be updated regularly to reflect technology updates and feedback from the counties.

Training around process improvement and change management in general is also a key factor of successful project implementation. In other words, how do we evaluate our processes, document them, and identify opportunities for efficiency. We also always try to bear in mind that one-size fits all processes generally aren't the most effective. What works for a large county with urban populations isn't going to translate well to a small rural county for several reasons, including resources, budget, and technology.

One approach we've found to be incredibly effective is to train to the goal and then crowdsource the solution. Get counties with similar populations, budgets, resources, and challenges together and work collaboratively to develop the business processes. This approach ensures that we develop good processes and it helps gain buy-in because it's not just a process mandated by the state.

When identifying process changes, we've also found that it's important to minimize significant disruptive changes, which can create confusion for pollworkers. Rather than implementing wholesale and overwhelming changes, which increase the risk of failure, we look for ways to streamline existing processes for efficiency and incorporate small adjustments. And with all changes, practice makes better. It's beneficial to practice and test the new processes to make sure things will work as expected. In addition to the counties conducting process walk-throughs, we also place a focus on hands-on training in the software.

Hands-on practice and mocks are one of our most effective training tools. We try to allow as much practice in the system as possible to build muscle memory. Colorado conducts a mock risk-limiting audit before every election. It gives the county staff and audit judges an opportunity to learn in a safe, but realistic, environment.

We believe it's important to treat a mock as a training exercise and respect the safe learning environment. We work to make sure it's safe to fail and learn from it. During the mock, the voting systems team spends a lot of time one-on-one with counties making sure they're comfortable and all of their questions are answered. They also work with counties to walk through any errors in the mock to explain how it would affect that county and the entire state in a real audit.

DEBRIEF AND IMPROVE

Following each audit, we solicit and listen to county feedback about the processes and the system. This has led to system enhancements for usability to help reduce errors as well as changes to the training and documentation. The team updates the instructions and documentation based on the feedback and resulting system changes, and we work with the county clerk's association to update the county process playbook.

It's also critical to continue providing refresher training. We survey following every training to ensure the training is meeting the counties' needs. And we've consistently found that the survey responses support a need for continual training. Finally, as we implement we try to keep the end goal in mind; what are we working to accomplish and why, and how can we work with our county partners and other stakeholders to ensure success.

LOOKING BEYOND COLORADO: CHANGING TECHNOLOGY AND STATE POLICY

LOOKING BEYOND COLORADO: RISK LIMITING AUDITS IN INDIANA

JAY BAGGA
BRYAN BYERS
Indiana Voting System Technical Oversight Program (VSTOP)
Ball State University

INTRODUCTION

The Indiana Voting System Technical Oversight Program (VSTOP) has recently been involved in risk-limiting audits and other election activities related to physical and cybersecurity of election systems in Indiana. This report presents a brief description of such activities.

VSTOP was established by Indiana statute in 2005 (P.L.221-2005, SEC.95). In 2008, the Indiana Secretary of State contracted with Ball State University to manage the operations of VSTOP. Since then, VSTOP has worked with the Indiana Secretary of State and the Indiana Election Division to manage many election-related activities, including developing and proposing procedures and standards for the certification, testing, acquisition, functioning, training, security for voting systems and electronic poll books used to conduct elections in Indiana, establishing and managing an inventory database of election equipment in the 92 counties in Indiana and offering a Certificate Program in Election Administration, Security and Technology (CEATS) to county election officials within the state.

LANDSCAPE OF ELECTION SYSTEMS IN INDIANA

The 92 counties in Indiana are served by five voting system vendors and five electronic poll book vendors. About half of the counties use DREs. The other half use OPSCANS or a combination with DREs.

To be certified for use in elections in Indiana, a voting system must comply with, among other requirements, the 2002 Voting System Standards (VSS), or the 2005 or 2015 Voluntary Voting System Guidelines (VVSG). VSTOP developed a protocol for certification of electronic poll books in In-

diana. According to the National Conference of State Legislatures, Indiana's 2013 e-poll book legislation is currently the most comprehensive in the country. Since 2013, the number of counties in Indiana using electronic poll books has grown rapidly (currently at about two-thirds).

INTEGRITY OF ELECTIONS AND SECURITY OF ELECTION SYSTEMS

Each year, the Bowen Center for Public Affairs at Ball State University conducts the annual Hoosier Survey, which aims to gather public opinion data on current issues and provides that data to policymakers. The 2017 survey included the following question:

> *What level of confidence do you have that your vote in the last election was properly recorded and accurately counted?*

The table below includes the responses of a random sample of 600 Indiana residents, showing that about 40% of the respondents are not very confident.

These findings indicate a lack of confidence among a sizable proportion of Indiana residents. Public perceptions are important indicators of areas where public officials may need to address concerns. A number of initiatives have been undertaken by the Indiana Secretary State and the State of Indiana to address issues with security and integrity around Indiana's elections. Elections are included in one important initiative within the state launched in 2017.

The Governor of Indiana established the Indiana Executive Council on Cybersecurity (IECC) in 2017 to *"... form an understanding of Indiana's cyber risk profile, identify priorities, establish a strategic framework of Indiana's cybersecurity initiatives, and leverage the body of talent to stay on the forefront of the cyber risk environment."* The IECC comprises several committees including the Elections committee, chaired by the Secretary of State.

Members of this committee include, as representatives, County Clerks, the Indiana Election Division, the Indiana Office of Technology, the Statewide Voter Reg-

CONFIDENCE LEVEL	PERCENTAGE
VERY CONFIDENT	60%
SOMEWHAT CONFIDENT	23%
NOT TOO CONFIDENT	8%
NOT CONFIDENT AT ALL	9%
DON'T KNOW/REFUSED TO ANSWER	<1%

istration Commission, and VSTOP. The committee has made several contributions including a review of the physical and cybersecurity aspects of elections, voting systems and electronic poll books, recommendation of best practices, and risk-limiting audits.

RECENT AND PROPOSED LEGISLATION

The 2018 Indiana Senate Enrolled Act (SEA 327) brought several physical and cybersecurity policies into law, including secure custody, sealing and storage, and inventory and disposal of election equipment. Under current law (IC 3-12-3.5-8), Indiana has some post-election audit requirements. Proposed legislation in the 2019 Senate Bill SB 570 includes voter verifiable paper audit trail (VVPAT) requirements and risk-limiting audits in the coming years.

SB 570 also includes national criminal history background checks of vendor employees, a requirement that polling places comply with the Election Infrastructure Outreach Security Checklist published by the United States Department of Homeland Security and a requirement that all problems or anomalies with the functioning of voting systems and electronic poll books be reported to the Secretary of State within 48 hours of its discovery.

RISK-LIMITING AUDITS

As part of its work with the IECC, VSTOP conducted the first ever RLA Pilot in Indiana in May of 2018. Dr. Ronald Rivest of MIT and Mr. Jerome Lovato of EAC assisted in this effort, among others. The RLA was conducted in Marion County, Indiana which includes the city of Indianapolis. Several weeks were spent in the preparation of this RLA. The RLA concept was totally new to Marion County and there was some initial reluctance. However, after discussion and several presentations, the county became quite interested in being part of the pilot. Substantial help was provided by the Marion County Clerk, Director of Elections, Deputy Director of Elections and their staff. The project was fully supported by the Indiana Secretary of State Connie Lawson.

In the Marion County RLA, three races were audited, the 2016 Presidential Election (5 precincts, Ballot Polling), the 2018 Primary Democratic Sheriff (10 Precincts, Ballot Polling) and the 2018 Republican U.S. Senator (10 Precincts, Comparison Polling). Both the Stark Method Risk Limit (10%) and the Bayesian Method (Bayesian Limit 5%) were employed. The first RLA confirmed Clinton as the winner in the precincts audited for the 2016 general election for President.

This was a fully completed RLA. The other two audits were ceased early due to time constraints. It is noteworthy that this was the first time that the Bayesian Audit Method was used in the field.

Jay Bagga and Bryan Byers presented the results of the Marion county RLA at the 8th annual national conference of the State Certification Testing of Voting Systems held in Raleigh, NC in June 2018. VSTOP also assisted with the organization of the RLA Pilot in Michigan that was led by Liz Howard of the Brennan Center.

The positive experience gained from the Marion county audit led the Secretary to ask VSTOP to conduct a second county wide audit of several races in Porter County, Indiana. This audit was conducted in January 2019 and included five countywide races (123 precincts): Public Question #1, Coun-

ty Prosecutor, County Auditor, County Recorder and County Coroner. The Porter County RLA was one of the most comprehensive local RLAs conducted in the United States and VSTOP was able to acquire valuable information about pre-election preparation, poll worker training, ballot chain-of-custody, post-election processes, and time and budget efforts for RLAs. The Three-Cut ballot sampling method facilitated efficient sampling and tabulation. Even with the ease of use and quick ballot polling, more time was still needed to sample additional ballots for the Recorder and Coroner races due to a substantial number of undervotes.

SUMMARY

The Voting System Technical Oversight Program (VSTOP) has been in existence since 2008. VSTOP's activities are wide and varied but are all concerned, in one way or another, with the integrity of elections and the security of election equipment. VSTOP has conducted two RLAs: Marion County (May 2018) and Porter County (January 2019. Both of these RLAs were successful, with the Porter County RLA being one of the most comprehensive ever performed.

The State of Indiana has taken many initiatives (including legislation) to secure elections and subsequently enhance voter confidence in election processes and outcomes. Should legislation regarding VVPATs and RLAs pass, VSTOP will be directly involved in the certification of VVPAT equipment and the implementation of Risk-Limiting Audits.

VOTING TECHNOLOGY & POST-ELECTION AUDITS

JEROME LOVATO
U.S. Election Assistance Commission

SUMMARY

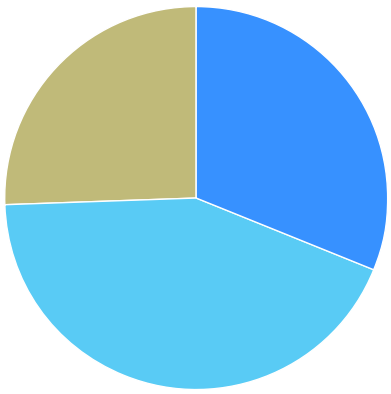
The decentralization of U.S. elections makes election administration very complex. One element of this complexity is voting technology and the ability to conduct risk-limiting audits (RLA). In this paper I will highlight three limitations of conducting RLAs, three ways to improve current voting system design,[1] and project what the future of RLAs look like with the advancement of voting technology and standards.

[1] *Voting system* refers to the total combination of mechanical, electromechanical or electronic equipment (including software, firmware, and documentation required to program, control, and support the equipment) that is used to define ballots; cast and count votes; report or display election results; and maintain and produce any audit trail information; and the practices and documentation used to identify system components and versions of such components; test the system during its development and maintenance; maintain records of system errors and defects; determine specific changes to be made to a system after the initial qualification of the system; and make available any materials to the voter.

WHERE WE ARE

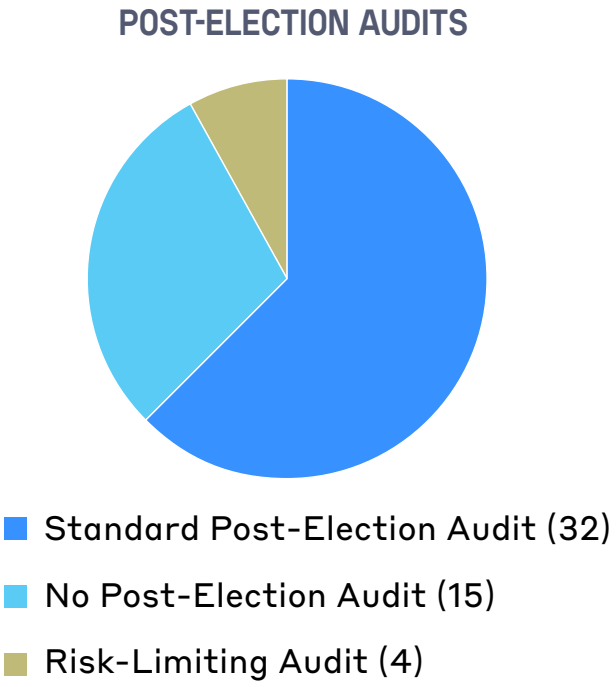
The U.S. Election Assistance Commission (EAC) is responsible for developing and maintaining the Voluntary Voting System

TESTING TO FEDERAL STANDARDS



Guidelines (VVSG). The VVSG are specifications and requirements by which voting systems are designed and tested. These specifications and requirements are voluntary, which means that states are not required to adopt these standards to test and certify their voting systems. At a minimum, most states (including Washing D.C.) require testing to federal standards (see graph on the previous page).

Along with the diverse voting system requirements are even more diverse post-election audit laws. Due to this diversity, I have placed post-election audits in two categories: Standard and RLA. The figure to the right here shows the number of states (including Washington D.C.) that conduct standard post-election audits, the number of states that do not require a post-election audit at all, and the number of states that require an RLA.



CURRENT LIMITATIONS

I have identified three limitations for conducting RLAs: a lack of paper ballot records, data exports from voting systems, and state-level certification requirements. This is not an all-encompassing list of limitations, but is a high-level "top 3" list.

The lack of a paper ballot record is the most obvious limitation for conducting RLAs. The simple solution is to just require that all voting systems produce a voter verifiable paper ballot record. However, this "simple" solution isn't so simple when election officials must consider other factors such as: legislation, budget, and training. How much will a new or modified voting system cost? How much will voting system certification cost and how long will it take to be certified? How much will it cost, and how long will it take, to implement a new voting system? What changes to election law will be needed to address RLAs? What resources are available to train local election officials on how to conduct RLAs, and how long will that take?

Voting systems produce a vast amount of data along with options to export that data. An essential export for conducting a ballot comparison RLA is a ***cast vote record***.[2] The export file formats vary for each voting system. For instance, some systems produce exports in JSON, others in XML, and others in CSV. Although these are commonly used data formats, the confusion arises in interpreting these files (i.e. what data is relevant for conducting the audit). Some formats are not human readable. For the files that are human readable, additional massaging of these files are required to make them intelligible.

[2] Permanent record of all votes produced by a single voter whether in electronic, paper or other form. Also referred to as ballot image when used to refer to electronic ballots.

A not-so-obvious limitation for conducting RLAs is varying state-level certification requirements. For example, consider the following scenario:

> *Voting System, Inc. develops Voting System Model Y2K and sells it to multiple states. Two states: State A and State Z, like the Model Y2K, but they have specific requirements that must be met. State A requires state-specific reports; State Z requires state-specific functionality. Voting System, Inc. decides to sell Model Y2K.1 to State A to produce the necessary reports, and Model Y2K.2 to State Z to address the necessary functionality.*
>
> *After the purchase, legislators in both states pass laws that require ballot comparison RLAs. Now, both states need additional exports that are only available in Model Y2K.7. Voting System, Inc. offers to upgrade State A to Model Y2K.7.1 and State B to Model Y2K.7.2, but it will require additional funds since a free upgrade was not included in the original contract.*

And that is a snippet of the complications that exist with varying state-level certification requirements.

CURRENT DESIGN

Three areas where voting system design can improve to assist with conducting RLAs are human-readable cast vote records, ballot imprinting, and independent verification.

Voting system manufacturers should work to produce human-readable cast vote records. Election officials and auditors should not have to use third-party utilities or devote additional resources to read cast vote records. It is recommended that the cast vote record be in a tabular format where each row of the table represents one paper ballot record.

Imprinting a unique ID on a ballot improves the efficiency of conducting a ballot comparison RLA. The unique ID should not be imprinted on sections on the ballot that will cause the ballot to be unreadable by the ballot scanner. The unique ID must not be able to tie a ballot back to the voter. Finally, the unique ID should be a field in the cast vote record. For example, if the unique ID on the ballot is "A-1111" then the cast vote record should reflect "A-1111" not "1111."

A basic principle of RLAs is that they provide independent verification of the results of an election. With that in mind, a voting system should not be designed to include an "RLA module" or any other self-auditing utility. Paper ballot records and proper ballot management and security are all that is needed to conduct an RLA.

FUTURE TECHNOLOGY

What lies ahead for the future of voting technology? Within the upcoming year, the EAC will publish ***Voluntary Voting System Gudelines 2.0***, which will include interoperability requirements. Part of the interoperability work includes creating common data format (CDF) standards for cast vote records, election results reporting, election event logging, and voter records interchange. CDF standards will make RLAs and other election-related audits much easier since it will eliminate the head scratching that exists today of wondering, "What am I even looking at?"

Other technology that is in the early stage of development is voting systems that use blockchain (i.e. UOCAVA ballot delivery) and end-to-end verifiable voting systems.

SOFTWARE SUPPORT FOR RISK-LIMITING AUDITS

MARK LINDEMAN
Verified Voting

Why do we need software to support risk-limiting audits? Many people I talk with assume it is because we need to do complex statistical calculations—because they've been told that risk-limiting audits are a statistical method. That's fundamentally mistaken, in much the same way it's mistaken to think of a meeting room as an engineering model. In audits, as in architecture, it's important to get the math right so nothing collapses, but the math itself is not the point.

Risk-limiting audits are a kind of tabulation audit, which means that at their heart, they're about having people manually examine a sample of voted ballots to check the voting system counts. Most of the work is about helping people manage paper, and to record what they see on the paper. The math is not the territory.

In this context, there's sometimes a disconnect between how statisticians talk about risk-limiting audits, and how election officials and others do. Are risk-limiting audits hard, or are they easy? From a statistical standpoint, many risk-limiting audits are easy: the underlying principles are explicable, the methods are straightforward, and sometimes the calculations can be done with pencil and paper.

In the real world, risk-limiting audits can get hard in at least two senses. First, in many jurisdictions, managing all the voted ballots in ways that support efficient auditing poses multifaceted logistical challenges. Second, election processes have a dizzying array of variations—voting method or methods, equipment, ballot design and differences, the number of sheets per ballot, and the time available to conduct audits – that efficient audit designs must or should

accommodate. Sometimes these accommodations complicate the statistics; more often, they require new features or subtle variations upon existing ones.

> *Risk-limiting audits are easy in some ways, but they aren't ramen-noodle easy: they're complicated because elections are complicated.*

Good audit design requires close collaboration between election officials and various kinds of domain experts to address specific goals in specific circumstances. Naturally, that affects software development.

A brief first-person case study: The city of Fairfax, Virginia, conducted a risk-limiting audit pilot in August 2018, in cooperation with the Virginia Department of Elections and Verified Voting. The pilot included a ballot-level comparison audit based on a retabulation of all the ballots cast in the June Republican primary (under 1,000 ballots), as well as a ballot-polling audit. I wrote the support software. It provided support for rescanning the voted ballots in batches, automatically interpreting the votes, manually reviewing apparent overvotes, selecting a random sample of ballots, retrieving those ballots from various batches, and entering the audit team's interpretation of the votes.

I spent maybe a few hours writing code to compute the statistics. Mostly I worked with election officials to design the audit procedures – specifying in detail what people would do with the paper ballots at every step – and then customized the software to be as helpful as possible.

In Fairfax, I wrote most of the audit code from scratch in Python, incorporating an open-source sampling algorithm written by Ron Rivest and the OpenCV computer vision library. Why did I do that? There is quite a bit of prior art on risk-limiting audit software, and much of it is open source. Let me briefly enumerate some of it.

- Philip Stark has two web pages that can be used to conduct audits from beginning to end. That's not an abstract possibility: several counties have used these tools.
- Ron Rivest and collaborators at MIT have developed several codebases including the GitHub ***bptool*** and ***bctool*** repositories, which provide support for Rivest's Bayesian audits.
- Open-source R and Python libraries referenced in the election audit literature support many of the basic concepts, although there has been no systematic effort to build out these offerings.
- Free & Fair developed the original open-source software implementation for Colorado's statewide risk-limiting audit in 2017—often called the "RLA tool."
- Democracy Works developed the 2018 version of the Colorado RLA tool.
- And a group of pro bono developers are working to integrate and extend some of these tools to support risk-limiting audits in Rhode Island, starting with a January 2019 pilot.

So, with all this software available for reuse, what was I thinking? If you've developed software or used it in your research, you probably can imagine how things were for me. You want to solve a problem, Various people say, "Oh, no worries, there are some fabulous open-source tools that do what you want." So you start looking around, and you find a bunch of tools. You can't get some of them to run because of mysterious software dependencies. With others, the documenta-

tion is so crude that it's hard to tell exactly what they do, or how they could be adapted to your specific use case. Maybe one is great for a distributed application with dozens of clients (e.g., many counties conducting an audit simultaneously), but seems unwieldy for standalone use. Eventually you may find yourself writing software that has the functions you want, isn't cluttered with functions you don't need, and will be easier for you to customize because you understand its assumptions and limitations.

That's what I did. Unlike the Colorado RLA tool, the Fairfax software supported rescanning and automatically interpreting ballots, and the user interfaces were designed for readability when projected on a wall. Also unlike the RLA tool, it could only handle one ballot style and one plurality contest. It was exactly what we needed at the beginning of August, and what I could write in about six weeks while doing the rest of my day job at Verified Voting.

The June 2018 RLA pilot in Orange County, California took a different path. Neal McBurnett and Stephanie Singer, who collaborated on the technical support, reused and extended Free & Fair's version of the RLA tool. This required some interesting improvisations, because the RLA tool did not support ballot-polling audits – even though, conceptually, ballot-polling audits are simpler than the ballot-level comparison audits that the RLA tool does support! But this approach did prove workable.

So, on the software side, we have a growing number of codebases, many of which are written or customized for specific cases. That's partly because the development efforts tend to have small or nonexistent budgets, limiting the capacity to write code that can be readily extended beyond the problem at hand. The two iterations of the Colorado RLA tool are the most ambitious, but the state of Colorado could not, and did not, pay for an all-purpose customizable audit tool. We face a collective action problem: the governmental entities that could benefit from a large investment in open-source audit software for shared use have no means to pool their resources in order to obtain it. This problem seems eminently solvable, because the necessary seed investment is not very large: half a million dollars would go a long way. A collaborative project that engages software developers, election officials, other domain experts, and philanthropic support to support risk-limiting audits is well within our collective competence.

I have focused on how software development can address the problem of diverse needs—but we also have opportunities to simplify the problem itself. Currently, audit solutions must contend with a Babel of incompatible vendor data interfaces and election-office improvisations. NIST working group have been developing a series of Common Data Format (CDF) documents, including a forthcoming CDF specification for Cast Vote Records – the interpretations of individual ballots that are used in ballot-level comparison audits. Widespread adoption of CDFs and other interoperability standards will facilitate future audit implementations and other election innovations.

Working to implement statistically rigorous post-election audits sounds dreary; "a software developer, a statistician, and an election official walk into a bar" sounds like bad comedy. (It probably is.) But it turns out that we have a lot to say to each other, and we all enjoy solving problems together. Who knew that "limiting risk" could be so much fun?

NEW DIRECTIONS FOR COMPREHENSIVE AUDITING AND FORENSICS

A METHOD TO AUDIT THE ASSIGNMENT OF REGISTERED VOTERS TO DISTRICTS

BRIAN AMOS
University of North Florida

MICHAEL MCDONALD
University of Florida

Ensuring election officials give voters the correct ballot appears to be an easy task. However, three recent elections demonstrate consequential administrative errors happen:

In June 2018, dozens of voters in Habersham County, Georgia received a letter from their county's Office of Elections and Registration informing them that they had been assigned to the incorrect State House district. The 2018 Republican primary was decided by just 67 votes, the losing candidate challenged the results, and a judge ordered a re-vote.

In 2018, election officials discovered twenty-five homes along a stretch of road in Hamden, Connecticut were never assigned to their new district following the 2012 redistricting, leading to voters casting ballots in the wrong district across several elections.

In 2017, at least 384 registered voters in northeastern Virginia were assigned to incorrect State House of Delegates districts, of whom at least 147 cast a ballot. 125 of these were voters incorrectly assigned to House District 28, a number greater than the Republican candidate's 82-vote margin of victory.

From a naïve viewpoint, elections officials should easily determine which district voters' addresses are located in. In practice, election officials use data-driven representations of a jurisdiction's geography to manage the scale of assigning thousands of voters to the many overlapping districts and precincts in their jurisdiction. Intrigued by these situation, we developed a methodology to audit the assignment of registered voters to districts, and worked with Colorado and Florida election officials to identify three mechanisms that lead to district assignment errors:

Human error. This mundane error occurs when human operators make data entry errors into election management databases. These errors take different forms, depending on the management system. A frequent error involves databases of street address segments, which are street address ranges (e.g., 100 to 198 of the even side of Main Street) that are associated with districts. Election officials relate street address segments to voter registration database addresses, to assign districts to individual registered voters. A data entry error in a street segment database creates district assignment errors for an entire street segment, which is easily observed when affected residences are overlaid on satellite imagery maps. The district assignment errors in Hamden, Connecticut has the markers of such human error.

Geocoding error. Some election officials use geocoding processes to assign voters to districts. Geocoders have different levels of accuracy for the latitude and longitude coordinates they assign to an address. The most accurate level is what is known as "roof-top" accuracy, wherein a geocoding database provider has an accurate latitude and longitude point for a known address, often obtained from local government records. Geocoding programs use algorithms to guess at a latitude and longitude when they encounter an unknown address. For example, a geocoder may guess that 150 Main Street lies midway between the endpoints of the 100 to 198 even side of the Main Street segment. There are two necessary assumptions for such algorithms to work well: a street lies in a straight line, and the correct setback distance from the street to the building is used.

Geocoding processes are not panaceas. Geocoding databases and algorithms are often proprietary, created through different processes, such that they can produce different results. We identified and verified district assignment errors even when election officials use a geocoding process to assign voters to districts. Assignment errors that appear to be caused by non-linear streets and setback issues are more prominent in rural areas, but we have observed these issues in urban areas, too. In one case, we identified an assignment error for a large apartment complex with hundreds of registered voters.

Asynchronous data. Assignment errors may arise from a geocoding process that uses out of date data. Among the more esoteric errors we observed occurred in Colorado, where their geocoding process to assign registrants to districts used district boundaries based on 2013 Census Bureau geographic data, while their geocoder used updated 2017 data. Subtle changes in the 2013 to 2017 census geographical data resulted in district assignment errors.

Briefly, our audit methodology works in the following steps:

1. ***Obtain a voter registration file.*** These data contain two important pieces of information for our purposes: voters' addresses and the districts they are assigned to.
2. ***Geocode voters' addresses.*** We find using multiple geocoding databases provides greater successful geocoding of a voter registration database. A frequent issue we observe using a single geocoding database is street name changes, which may be updated in a geocoding database, but persist as legacy addresses in a voter registration database.
3. ***Obtain district boundary files.*** The Census Bureau disseminates boundary data for congressional and state legislative districts. Collecting data on other statewide districts and local district is deeply challenged.
4. ***Perform a spatial join.*** We overlay the point locations of geocoded voter registration addresses, and their associated district assignments per the voter registration file, on the district boundaries and note where the district identifiers are different.
5. ***Verify potential errors.*** We check each suspect address by overlaying the data we generate onto satellite imagery. This helps confirm that a building is indeed located at the latitude and longitude identified by the geocoding software.
6. ***Generate reports.*** We generate lists of suspect addresses, accompanied with maps of district boundaries and dots locating suspect addresses overlaid onto satellite imagery.

The good news is that we can audit district and precinct assignments. Technological innovations have progressed such that it is possible to develop and deploy auditing systems, and we recommend election officials to take advantage of them. Indeed, some vendors have deployed systems to report on the assignment of registered voters to districts, similar to the methodology we describe. However, even when election officials use such systems, we recommend an external audit since they depend on geocoding databases that may themselves have errors.

The result of these efforts will be better election data integrity, which will improve voters' experiences, reduce election costs, and improve voters' confidence in the electoral system.

PUBLICLY-VERIFIABLE ELECTIONS

JOSH BENALOH
Microsoft Research

Electoral systems in the United States are easy targets for attackers. As observed in the recent report from the National Academies of Science, Engineering, and Medicine, ***Securing the Vote: Protecting American Democracy***, our systems for casting and counting of votes are extremely vulnerable. The standards and practices in election systems do not compare with those of most industries, and they fall far short of the level that should be achieved by such a critical infrastructure.

However, it is important to recognize that while industry best practices should be applied, this isn't enough. Most of the over eight thousand election jurisdictions in the U.S. are small and lack a dedicated staff of information technology professionals. But many attacks come from nation-states with vastly superior resources and expertise. The battle is asymmetric, and it is simply not realistic to assume that it is possible to make our electoral system impervious to all possible attacks. We can, however, build a robust auditing infrastructure that allows us to know if any of our elections have been tampered with.

There are two basic varieties of audits:

1. Process audits allow administrators or third-parties to look at equipment and procedures to ensure that best practices are being applied.
2. Tally audits allow parties to verify the correct recording and counting of votes.

Within this second category, there are again two varieties:

1. Administrative audits allow election administrators to statistically sample ballots to confirm that they are consistent with the reported tallies.

2. Public audits allow independent observers and voters themselves to confirm that their ballots are correctly recorded and tallied.

While administrative audits, such as risk-limiting audits, are quite valuable and should be conducted for every contest in every election, they can be cumbersome and do little to provide confidence to a voter or observer who does not trust election administrators to properly maintain original ballots or to conduct their audits in ways that are fully independent of the original tallying.

In contrast, public audits allow skeptical parties to verify the accuracy of tallies themselves—without having to delegate trust to third-parties. The primary means for public auditing is a set of technologies that achieve what is known as ***end-to-end (E2E) verifiability***.

An election is said to be end-to-end verifiable if the following two properties hold:

1. *Voters can verify their own votes have been properly recorded.*

2. *Any observer can verify that all recorded votes have been correctly tallied.*

E2E-verification depends on the public, rather than election administrators, to perform auditing tasks. In high-profile elections, this may be commonplace. However, there is no guarantee that sufficient public attention will be paid to lower-profile elections. This is just one of several reasons for every election to also undergo administrative auditing.

It is easy to see how the requirements of E2E-verifiability can be achieved in open-ballot elections. Voters can convey their selections to election administrators who then post all votes—together with the names of the voters who cast them—in a public place such as a (digitally signed) web page. Voters can easily see that their votes are correctly recorded, and any observer can easily tally the votes to confirm that they correspond to the announced tallies. (A digital signature deters a malicious administrator from showing different posts to different viewers—since discovery of two distinct signed lists immediately implicates administrators as acting improperly.) The challenge is to achieve E2E-verifiablity in secret-ballot elections, and the typical mechanism is to post encrypted votes rather than open votes.

When posted votes are encrypted, achieving E2E-verifiability requires providing voters with means to confirm that the encrypted votes associated with them represent their actual selections (and this must be done in a way that does not allow voters to reveal their votes to others) and a mechanism must be provided to allow observers to verify that the encrypted votes accurately reflect the announced tallies.

There are multiple ways in which each of these tasks can be achieved. Numerous innovative mechanisms have been developed that allow voters to confirm the correct recording of their votes. Most don't require voters to take any extraordinary steps and do not impose additional burdens on voters who choose to avail themselves of this capability. The common element is that almost all of these systems provide voters with take-home receipts that can be used to track their votes. These receipts do not allow voters to see their actual selections nor

to show them to others. Instead, they allow voters to confirm that their votes have not been changed since the time that they were cast (when voters could confirm their actual selections).

The verification of tallying typically requires sophisticated cryptographic methods—often employing ***homomorphic encryption***. Homomorphic tallying allows encrypted votes to be directly amalgamated in encrypted form to construct an encrypted tally. This amalgamation can be repeated and checked by any observer and often is no more complicated than multiplying the encrypted ballots together. This aggregate encryption is then decrypted by election administrators who also provide a proof that allows observers to independently verify that the decryption is correct.

An alternative approach, known as a ***Mix-Net***, allows the encrypted ballots to be publicly shuffled while preserving their contents. Election administrators or others can serve as shufflers, and a proof must accompany each shuffle to demonstrate that the contents haven't been altered (the shuffling and proof typically use homomorphic encryption methods—although they do not employ homomorphic tallying). Once all of the encrypted ballots have been sufficiently shuffled, each ballot is individually decrypted by election administrators—who also provide independently verifiable proofs of each decryption. The open ballots (now dissociated from the voters who cast them) can be independently tallied by any observer to confirm that the announced tallies are correct.

With both homomorphic tallying and Mix-Nets, there is generally not a single decryption key. Instead, the key generation process is usually distributed so that multiple authorities must cooperate to form a decryption. Ideally, threshold encryption is used so that, for instance, three of five election authorities must cooperate to decrypt. This distributes the decryption capabilities so that a single rogue entity cannot compromise privacy while providing robustness so that a minority cannot prevent an election from completing.

The collection of technologies that enable public verifiability of election tallies provides a valuable complement to risk-limiting audits and similar administrative auditing methods. When used together, public and administrative audits can engender strong public confidence in the accuracy of election results.

DESIGNING BALLOTS FOR VOTERS AND ELECTION WORKERS—AND AUDITS

WHITNEY QUESENBERY
Center for Civic Design

At first, it wasn't entirely clear how ballot design fits into a conference on election audits. In all the discussions about the mathematics and 12-sided dice, it's easy to lose track of the goal: to ensure that the connection between voters and the results of an election is not broken. That is, an audit asks whether the ballots in an election were counted as they were cast, so perhaps we should also consider how to ensure that voters have the best possible opportunity to mark, verify, and cast a ballot that reflects their intent.

Phillip Stark famously said "an audit is no better than the paper trail it uses." I completely agree with that. Of course paper ballots are essential as a record of voter intent.

> *But, I disagree that a hand-marked, optical scan-style paper ballot is the only ballot design, or even a 'gold standard.'*

In fact, we have a long, rich history of ballot design that has fooled voters and has sometimes arguably affected the outcome of an election. Many of us are in this field today because of the butterfly ballot in Palm Beach County, Florida in 2000. But there have been many other examples before and since.

We know the problems and have strong research and empirical election evidence for best practices and the designs that cause problems: Contests that are split over two columns cause overvotes. Open primaries with two party elections on the same ballot invite people to vote in both, and throw away their vote. Confusing instructions, too small text, weak alignment between the marking target and the candidate. Sadly,

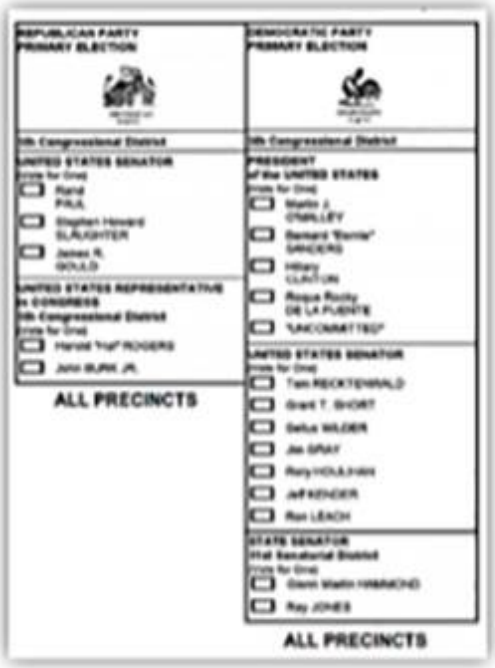
problems persist—even into the recent 2018 General Election, when a flawed ballot layout in Broward County triggered a spike in undervotes in a tight race for Senate (see the image on the next page).

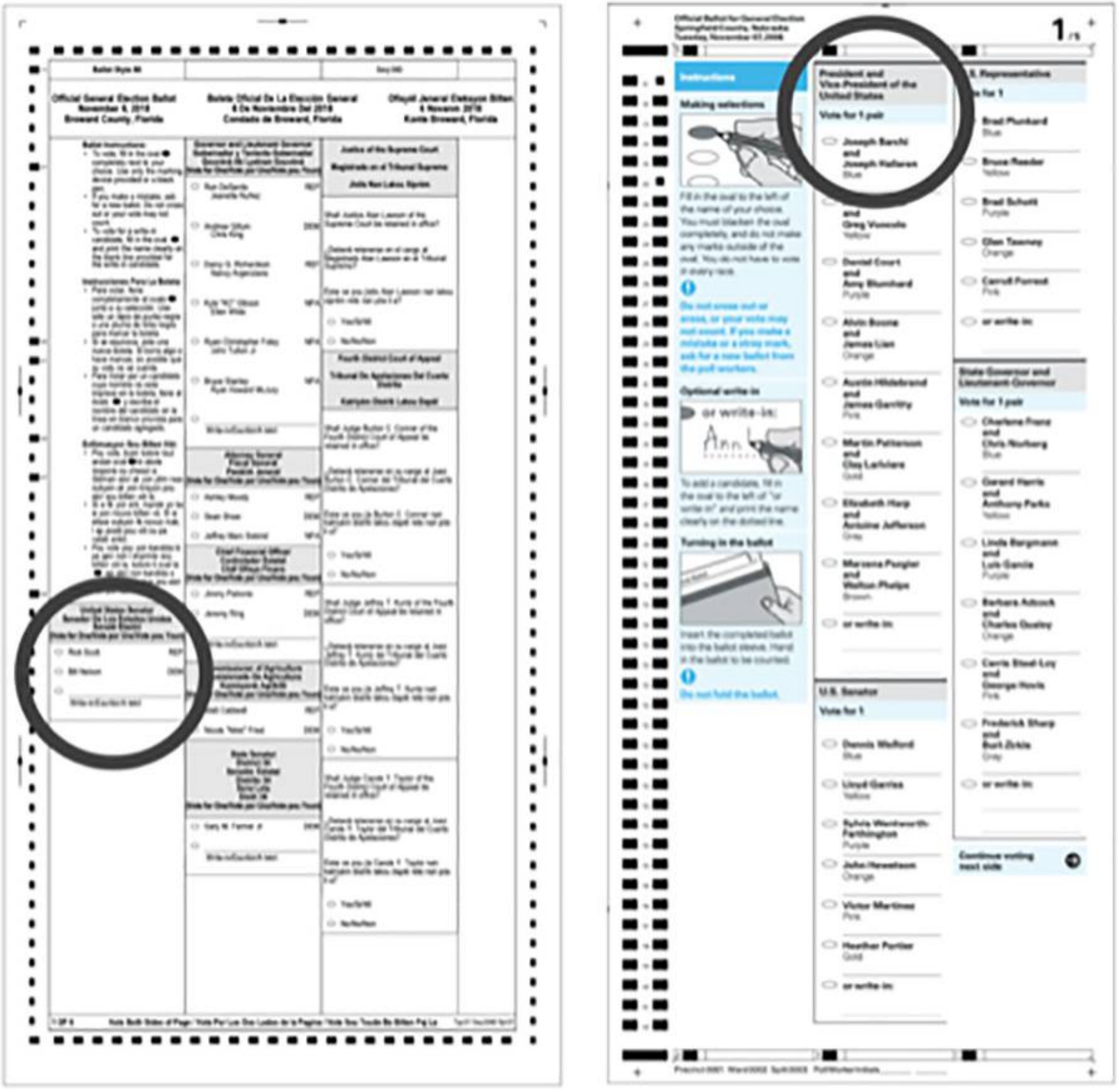
What makes this especially tragic is that we are not coming up with new problems…just new permutations of the same problems. A poorly designed ballot can result in voters making mistakes that result in a ballot that looks to be clearly marked (that is, it's been marked in an unambiguous way that the scanner easily reads) but in reality, they have not voted as they intended.

We need a better paper trail with ballots that are designed for capturing voter intent, that works for voters with the full range of civic literacy, elections savvy, physical and cognitive abilities. Ballots that don't rely on voters remembering or understanding the rules with no support from the voting system to verify their ballot. And ballots that are easy to read during an audit.

To get there, let's start with the process of voting. A ballot is the result of a conversation between the voter and the voting system to produce a paper ballot that reflects their intent with no ambiguous or inaccurate marks. We might think about ballots as a menu, showing all of the options. But when you go to your favorite restaurant, you don't order everything on the menu, so your bill at the end of the meal shows you what you selected. To make this metaphor work for a ballot, it also has to include any opportunity not taken in the list of selections.

Election history is a rogue's gallery of design defects. For more examples with the impact on real elections, see the Brennan Center's 2008 report, "Better Ballots." Left: Open primary with both parties on the same ballot. Center: Ballot with a contest split across two columns. Right: a contest in the left column, below the instructions.





The 2018 Broward County ballot, with contests below the instructions, compared to the EAC best practices and flawed Florida sample ballot. Left: The Broward County 2018 ballot. Right: Sample ballot from the EAC best practices.

Designing a voter selections ballot also requires understanding how marking a ballot moves through different information design needs for each stage of the process, in a form of progressive disclosure.

- As a voter first marks the ballot, the emphasis is on the rules and choices, helping the voter focus on each contest, one at a time.
- At the review screen of an electronic ballot marking interface, the focus shifts to a preview of all contests and selections, emphasizing missed opportunities to vote.
- Then, the printed ballot is a confirmation, with the ability to verify all of the contests and selections (and undervotes) before casting the ballot.

In addition to its accessibility features, the value of an electronic marking interface is that the voting system understands and can communicate the rules for the election: how many votes are allowed in each contest, voting variations like straight-party or ranked choice voting. It can also meet the goals of the Help America Vote Act by preventing overvotes entirely.

Seen in context, verification is not a proofreading task. It is the moment when a voter can say, "This is my ballot, and after all of the process for marking the ballot, this is how I am voting." For this moment to be meaningful, the ballot must be designed to be scanned quickly and accurately, with clear presentation of names, parties, and non-selections.

In addition to the design, the presentation of the physical ballot also matters. A piece of paper behind glass is not a useful verification artifact if voters can't read it because the text is too small, or the transport mechanism obscures part of the information, or glare from the glass makes it impossible to read. And no paper ballot supports verification for blind and low-vision voters unless is it can either be read back into a system or scanned with independent, trusted personal assistive OCR technologies.

Unambiguously marked ballots that are easy to read also have an effect on risk limiting audits. According to experts like Jennifer Morrell, much of the time in an audit is spent adjudicating voter intent on hand-marked ballots.

Because the Center for Civic Design is the voice of the humans in the process, we also have to mention the need to make risk limiting audits easy to run. Election workers need procedures, tools, and instructions that are clear, usable, and effective. There are best practices for writing good instructions for complex procedures and for designing forms. Let's bring them into election administration procedures, because elections work better when all of the materials are easy to use.

FURTHER READING

WEBSITES AND GUIDES ON ELECTION AUDITS

Lovato, Jerome. 2018. Risk-Limiting Audits – Practical Application. U.S. Election Assistance Commission.

Morrell, Jennifer. 2019. Knowing it's Right: A Two-Part Guide to Risk-Limiting Audits. Democracy Fund.

National Conference of State Legislatures. Post-Election Audits.

National Institute of Standards and Technology. Interoperability Public Working Group and Common Data Format (CDF) for Election Systems Project.

Stark, Philip. Tools for Comparison Risk-Limiting Election Audits.

Stark, Philip. Tools for Ballot-Polling Risk-Limiting Election Audits.

U.S. Election Assistance Commission. Post Election: Audits and Recounts.

U.S. Election Assistance Commission. Voluntary Voting System Guidelines.

Verified Voting. Post Election Audits.

STATISTICAL MODELS OF POST-ELECTION BALLOT AUDITING

Aslam, Javed A., Raluca A. Popa, and Ronald L. Rivest. 2008. On Auditing Elections When Precincts Have Different Sizes. Proceedings of the 2008 USENIX/ACCURATE Electronic Voting Technology Workshop.

Benaloh, Josh, Ronald Rivest, Peter Y.A. Ryan, Philip Stark, Vanessa Teague, and Poorvi Vora. 2015. End-to-End Verifiability. arXiv:1504.03778

Baxter, Patrick, Anne Edmundson, Keishla Ortiz, Ana Maria Quevedo, Samuel Rodriguez, Cynthia Sturton, and David Wagner. 2012. Automated Analysis of Election Audit Logs. Proceedings of the 2012 USENIX/WOTE Conference.

Bell, Susan, Josh Bealoh, Michael Byrne, Dana DeBeauvoir, Bryce Eakin, Gail Fisher, Philip Kortum, Neal McBurnett, Julian Montoya, Michelle Parker, Olivier Pereira, Philip Stark, Dan Wallach, Michael Winn. 2013. STAR-Vote: A Secure, Transparent, Auditable, and Reliable Voting System. JETS: USENIX Journal of Election Technology and Systems, 1.1.

Cordero, Arel and David Wagner. 2008. Replayable Voting Machine Audit Logs. Proceedings of the 2008 USENIX/ACCURATE Electronic Voting Technology Workshop.

Davtyan, Seda, Sotiris Kentros, Aggelos Kiayias, Laurent Michel, Nicolas Nicolaou, Alexander Russell, Andrew See, Narasimha Shashidhar, and Alexander A. Shvartsman. 2008. Pre-Election Testing and Post-Election Audit of Optical Scan Voting Terminal Memory Cards. Proceedings of the 2008 USENIX/ACCURATE Electronic Voting Technology Workshop.

Laurent, Michel, Alexander Shvartsman, and Nikolaj Volgushev. 2014. A Systematic Approach to Analyzing Voting Terminal Event Logs. JETS: USENIX Journal of Election Technology and Systems, 2.2.

Lindeman, Mark and Philip B. Stark. A Gentle introduction to Risk-limiting Audits. IEEE Security & Privacy 10, no. 5 (2012): 42-49.

Lindeman, Mark, Philip Stark, and Vincent Yates. 2012. BRAVO: Ballot-polling Risk-limiting Audits to Verify Outcomes. Proceedings of the 2012 USENIX/WOTE Conference.

Lindeman, Mark, Mark Halvorson, Pamela Smith, Lynn Garland, Vittorio Adona, and Dan McCrea, eds. Principles and Best Practices for Post-Election Audits.

Risk-Limiting Audits Working Group. 2012 (Version 1.1). Risk-Limiting Post-Election Audits: Why and How.

Rivest, Ronald L. 2018. Bayesian Tabulation Audits Explained and Extended.

Rivest, Ronald L. and Emily Shen. 2012. A Bayesian Method for Auditing Elections. Proceedings 2012 EVT/WOTE Conference.

Shenker, Jacob, and R. Michael Alvarez. 2014. Mitigating Coercion, Maximizing Confidence in Postal Elections. JETS: USENIX Journal of Election Technology and Systems, 2.3.

Stark, Philip B. 2009. Risk-Limiting Postelection Audits: Conservative P-Values From Common Probability Inequalities. IEEE Transactions on Information Forensics and Security, Vol. 4, No. 4.

Stark, Philip B. 2012. Ballot-polling Risk-limiting Audits in Two Pages (±1).

Stark, Philip B. Tools for Comparison Risk-Limiting Election Audits. Website.

Stark, Philip B. Tools for Ballot-Polling Risk-Limiting Election Audits. Website.

Stark, Philip B. and David Wagner. 2012. Evidence-Based Elections. IEEE Security and Privacy, Special Issue on Electronic Voting.

University of Connecticut. Center for Voting Technology Research. Computer Assisted Post Election Audits. Website.

Verified Voting. Post-Election Audits. Website.

QUANTITATIVE STUDIES OF POST-ELECTION BALLOT AUDITING

Alvarez, R. Michael, Lonna Rae Atkeson, and Thad E. Hall. 2012. Evaluating Elections: A Handbook of Methods and Standards. New York: Cambridge University Press, Chapter 5.

Alvarez, R. Michael, Lonna Rae Atkeson, and Thad E. Hall, editors. 2012. Confirming Elections: Creating Confidence and Integrity through Election Auditing. New York: Palgrave MacMillan.

Alvarez, R. Michael, Jonathan N. Katz, Sarah A. Hill, and Erin K. Hartman. 2012. Machines Versus Humans: The Counting and Recounting of Paper Ballots in LA County. In R. Michael Alvarez, Lonna Rae Atkeson, and Thad E. Hall, editors. 2012. Confirming Elections: Creating Confidence and Integrity through Election Auditing. New York: Palgrave MacMillan, 73-88.

Ansolabehere, Stephen, Barry C. Burden, Kenneth R. Mayer, and Charles Stewart III. 2018. Learning from Recounts. Election Law Journal: Rules, Politics, and Policy 17.2: 100-116.

Ansolabehere, Stephen and Reeves Andrew. 2012. Using Recounts to Measure the Accuracy of Vote Tabulations: Evidence from New Hampshire Elections 1946–2002. In Confirming Elections: Creating Confidence and Integrity through Election Auditing, eds. Michael Alvarez R., Atkeson Lonna Rae, and Hall Thad E.. New York: Palgrave.

Goggin, Stephen N., Michael D. Byrne, Juan E. Gilbert, Gregory Rogers, and Jerome McClendon. 2008. Comparing the Auditability of Optical Scan, Voter Verified Paper Audit

Trail (VVPAT) and Video (VVVAT) Ballot Systems. Proceedings of the 2008 USENIX/ACCURATE Electronic Voting Technology Workshop.

Kim, Eric, Nicholas Carlini, Andrew Chang, George Yiu, Kai Wang, and David Wagner. 2013. Improved Support for Machine-Assisted Ballot-Level Audits. JETS: USENIX Journal of Election Technology and Systems, 1.1.

McLaughlin, Katherine and Philip B. Stark. 2011. Workload Estimates for Risk-Limiting Audits of Large Contests.

National Academies of Sciences, Engineering, and Medicine. 2018. Securing the Vote: Protecting American Democracy. Washington, DC: The National Academies Press.

» *Further information about the report and related resources can be found at the committee's webpage.*

Stark, Philip B., and Vanessa Teague. 2014. Verifiable European Elections: Risk-limiting Audits for D'Hondt and Its Relatives, JETS: USENIX Journal of Election Technology and Systems, 3.1.

U.S. Vote Foundation. 2015. The Future of Voting: End-to-End Verifiable Internet Voting—Specification and Feasibility Study.

CASE STUDIES OF POST-ELECTION BALLOT AUDITING

Bader, Max. 2013. Do new voting technologies prevent fraud? Evidence from Russia. JETS: USENIX Journal of Election Technology and Systems, 2.1.

California Post-Election Audit Standards Working Group. 2007. Report: Evaluation of Audit Sampling Models and Options for Strengthening California's Manual Count.

Davtyan, Seda, Sotiris Kentros, Aggelos Kiayias, Laurent Michel, Nicolas Nicolaou, Alexander Russell, Andrew See, Narasimha Shashidhar, and Alexander A. Shvartsman. 2008. Pre-Election Testing and Post-Election Audit of Optical Scan Voting Terminal Memory Cards. Proceedings of the 2008 USENIX/ACCURATE Electronic Voting Technology Workshop.

Chilingirian, Berj, Zara Perumal, Ronald L. Rivest, Grahame Bowland, Andrew Conway, Philip B. Stark, Michelle Blom, Chris Culnane, and Vanessa Teague. 2016. Auditing Australian Senate Ballots. arXiv:1610.00127

The Colorado Risk-Limiting Project. Overview.

Douglas, Michael. 2014. Ballot Bungles: Lessons from the Australian Senate. Election Law Journal 13.4: 559-569.

Hall, Joseph Lorenzo. 2008. Improving the Security, Transparency, and Efficiency of California's 1% Manual Tally Procedures. Proceedings of the 2008 USENIX/ACCURATE Electronic Voting Technology Workshop.

ELECTION FORENSICS

A. Election Observation

Bland, Gary. 2015. Measuring the Quality of Kenya's March 2013 Election. Election Law Journal 14.2: 136-147.

Hall, Joseph, Emily Barabas, Gregory Shapiro, Coye Cheshire, and Deirdre Mulligan. 2012. Probing the Front Lines: Pollworker Perceptions of Security & Privacy. Proceedings of the 2012 USENIX/WOTE Conference.

Herron, Erik S., and Nazar Boyko. 2016. Conducting Credible Elections Under Threat: Results from a Survey of Election Administrators. Election Law Journal 15.4: 285-301.

Pomares, Julia, Ines Levin, and R. Michael Alvarez. 2014. Do Voters and Poll Workers Differ in their Attitudes Toward e-voting? Evidence from the first e-election in Salta, Argentina. JETS: USENIX Journal of Election Technology and Systems, 2.2.

B. Digit Tests

Beber, Bernd, and Alexandra Scacco. 2012. What the Numbers Say: A Digit-Based Test for Election Fraud. Political Analysis 20.2: 211-234.

Deckert, Joseph, Mikhail Myagkov, and Peter C. Ordeshook. 2011. Benford's Law and the Detection of Election Fraud. Political Analysis 19.3: 245-268.

Jiménez, Raúl, and Manuel Hidalgo. 2014. Forensic Analysis of Venezuelan Elections during the Chávez Presidency. Plos One 9.6: e100884.

Klimek, Peter, et al. 2012. Statistical detection of systematic election irregularities. Proceedings of the National Academy of Sciences 109.41: 16469-16473.

Mebane, Walter. 2008. Election Forensics: The Second-Digit Benford Law's Test and Recent American Presidential Elections. In Election Fraud: Detecting and Deterring Electoral Manipulation, eds. R. Michael Alvarez, Thad E. Hall, Susan D. Hyde, 162–81. Brookings.

Weidmann, Nils B., and Michael Callen. 2013. Violence and Election Fraud: Evidence from Afghanistan. British Journal of Political Science 43.1: 53-75.

C. Quantitative Forensics

Levin, Ines, Gabe A. Cohn, Peter C. Ordeshook, and R. Michael Alvarez. 2009. Detecting Voter Fraud in an Electronic Voting Context: An Analysis of the Unlimited Reelection Vote in Venezuela. EVT/WOTE '09 Proceedings, August 10-11, 2009, Montreal, Canada.

Levin, Ines, Julia Pomares, and R. Michael Alvarez. 2012. Using Machine Learning Algorithms to Detect Election Fraud. In R. Michael Alvarez, editor, Computational Social Science: Discovery and Prediction. New York: Cambridge University Press, 266-294.

Wand, Jonathan N., Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Michael C. Herron, and Henry E. Brady. 2001. The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida. American Political Science Review 95.4: 793-810.

D. Registration Database Auditing

Ansolabehere, Stephen and Eitan Hersh. 2010. The Quality of Voter Registration Records: A State-by-State Analysis. Report, Caltech/MIT Voting Technology Project.

Brater, Jonathan, Kevin Morris, Myrna Perez, Christopher Deluzio. 2018. Purges: A Growing Threat to the Right to Vote. Brennan Center for Justice..

Norden, Lawrence. Voting System Failures: A Database Solution. Brennan Center for Justice. 2010.

Ponoroff, Christopher. 2010. Voter Registration in a Digital Age. Brennan Center for Justice.

Weiser, Wendy. 2009. Voter Registration Modernization: Collected Reports and Papers. Brennan Center for Justice.


Chicago
Detroit
Cleveland
Columbus
Lansing
Milwaukee
IND.
OHIO
Baltimore
W.VA.
Kansas City
St. Louis
Jefferson City
STATES
MISSOURI
Frankfort
Charleston
Nashville
TENN.
Charlotte
N.C.
Raleigh
Greensboro
Memphis
Little Rock
ARK.
Huntsville
S.C.
Columbia
Atlanta
Dallas
MISS.
ALA.
GEORGIA
Charleston
Jackson
Montgomery
Savannah
LA.
Waco
Baton Rouge
Mobile
Tallahassee
Jacksonville
New Orleans
Houston
Lafayette
FLORIDA
Orlando
Cape Canaveral
Galveston
Corpus Christi
St. Petersburg
Tampa
Grand Bahama
Ft. Lauderdale
Miami

MIT ELECTION DATA + SCIENCE LAB

Caltech/MIT Voting Technology Project