Bayesian Audits are Average but Risk-Limiting Audits are Above Average

Amanda K. Glazer, Jacob V. Spertus, and Philip B. Stark⋆

University of California, Berkeley, Department of Statistics
`amandaglazer@berkeley.edu`; `jakespertus@berkeley.edu`; `pbstark@berkeley.edu`

Abstract. Post-election audits can provide convincing evidence that election outcomes are correct—that the reported winner(s) really won—by manually inspecting ballots selected at random from a trustworthy paper trail of votes. Risk-limiting audits (RLAs) control the probability that, if the reported outcome is wrong, it is not corrected before the outcome becomes official. RLAs keep this probability below the specified "risk limit." Bayesian audits (BAs) control the probability that the reported outcome is wrong, the "upset probability." The upset probability does not exist unless one posits a prior probability distribution for cast votes. RLAs ensure that if *this* election's reported outcome is wrong, the procedure has a large chance of correcting it. BAs control a *weighted average probability* of correcting wrong outcomes over a hypothetical collection of elections; the weights come from the prior. There are priors for which the upset probability is equal to the risk, but in general, BAs do not ensure a large chance of correcting the outcome of an election when the reported outcome is wrong. "Nonpartisan" priors, i.e., priors that are invariant under relabeling the candidates, lead to upset probabilities that can be far smaller than the chance of correcting wrong reported outcomes. We demonstrate the difference using simulations based on several real contests.

Keywords: election integrity · risk-limiting audits · Bayesian audits

1 Introduction

The 2016 U.S. Presidential election was attacked by Russian hackers, and U.S. Intelligence agencies warn that several nation-states are already mounting attacks on the 2020 election [29, 30, 31, 22]. Almost every U.S. jurisdiction uses computers to count votes; many use computers to record votes. All computerized systems

---

are vulnerable to bugs, misconfiguration, and hacking [26]. Voters, poll workers, and election officials are also bound to make mistakes [15]. Enough error from any source—innocent or malicious—could cause a losing candidate to appear to win.

The reported tallies will almost certainly be off by at least a little. Were the tallies accurate enough to ensure that the reported winner(s) really won—that the *reported outcome* is correct?

An election is *evidence-based* [26] if it provides convincing public evidence that the reported winners really won. The only federally certified technology that can provide such evidence is trustworthy paper ballots kept demonstrably secure throughout the election and canvass, then audited manually [2]. However:

- 14% of registered voters live in in jurisdictions using Direct Recording Electronic (DRE) Systems for all voters. DREs do not retain a paper ballot [27].
- Some paper ballots are not trustworthy. For instance, touchscreen voting machines and ballot-marking devices are vulnerable to bugs, hacking, and misconfiguration that can cause them to print the wrong votes [3, 4].
- Rules for securing cast ballots and for ensuring the paper trail remains trusworthy are uneven and generally inadequate.

Nonetheless, to focus on statistical issues, we assume here that elections produce a trustworthy collection of paper ballots containing voters' expressed preferences [3, 2, 11, 26]. A trustworthy paper trail allows audits to check whether errors, bugs, or malfeasance altered the reported outcome. ("Outcome" means who won, not the exact vote tallies.) For instance, we could tabulate the votes on all the cast ballots by hand, as some recount laws require. But full manual recounts are expensive, contentious, and rare: according to Richie and Smith [19], only 27 statewide U.S. elections between 2000 and 2015 were manually recounted; three of the recounts overturned the original outcomes (11%).

Some states conduct tabulation audits that involve manually reading votes from some ballots. For instance, California law requires manually tabulating the votes on ballots in 1% of precincts selected at random.[1] Such audits typically do not ensure that outcome-changing errors will (probably) be detected, much less corrected. In contrast, risk-limiting audits (RLAs) [23, 11] have a known minimum chance of correcting the reported outcome if the reported outcome is wrong (but never alter correct outcomes). RLAs stop without a full hand count only if there is sufficiently strong evidence that a full hand count would find the same winners, i.e., if the P-value of the hypothesis that the reported outcome is wrong is sufficiently small.

RLAs have been endorsed by the National Academies of Science, Engineering, and Medicine [15], the American Statistical Association [1], and many other organizations concerned with election integrity. There have been roughly

---

[1] The law is a bit more complicated, including provisions to ensure that every contest gets some scrutiny and options for sampling vote-by-mail ballots (including not sampling them if they arrive after election day).

60 pilot RLAs in 15 U.S. states and Denmark. Currently 10 U.S. states require or specifically allow RLAs. There have been statewide RLAs or pilot RLAs in five U.S. states: Alaska[2], Colorado [8], Kansas[3], Rhode Island [7], and Wyoming[3], and a pilot RLA in Michigan in which 80 of 83 counties participated [13].

Bayesian audits (BAs, [20, 21]) have been proposed as an alternative to RLAs. BAs stop without a full hand count only if the "upset probability"—the posterior probability that the reported winner(s) actually lost, for a particular prior π, given the audit sample—is below a pre-specified threshold. They have been piloted in several states.

Bayesian and frequentist interpretations of probability are quite different. Frequentist probability is the long-run limiting relative frequency with which an event occurs in repeated trials. Bayesian probability quantifies the degree to which the subject believes an event will occur. A prior probability distribution quantifies beliefs before the data are collected; after the data are observed, Bayes' rule says how to update the prior using the data to obtain the posterior probability distribution.

Bayesian methods, including BAs, require stronger assumptions than frequentist methods, including RLAs. In particular, BAs require assuming that votes are random and follow a known "prior" probability distribution π.

Both RLAs and BAs rely on manually interpreting randomly selected ballots. In principle, both can use a wide range of sampling plans to accommodate differences in how jurisdictions handle and store ballots and variations in election laws and regulations. (To the best of our knowledge, BAs have been conducted only using "ballot polling" [9].) RLA methods have been developed to use individual ballots or groups of ballots as the sampling unit, to sample with or without replacement or to use Bernoulli sampling, to sample with and without stratification, and to sample uniformly or with unequal probabilities (see, e.g., Stark [23, 24, 25], Lindeman and Stark [11], Ottoboni et al. [18, 17]).

The manual interpretations can be used in two ways: *comparison audits* look at differences between the manual interpretation and the machine interpretation and tabulation, while *polling audits* just use the manual interpretation. (The two strategies can be combined in a single audit; see, e.g., Ottoboni et al. [18], Stark [25].) Comparison audits require more of the voting system and require more preparation than polling audits, but for a given size sampling unit, they generally require smaller samples. (The sample size scales like the reciprocal of the margin for comparison audits, and like the square of the reciprocal of the margin for polling audits.) Below, we focus on polling audits that use individual ballots as the sampling unit: *ballot-polling audits*. These are the simplest conceptually and require the least of the voting system: just the reported winner(s), but no other data export.

Both RLAs and BAs lead to a full hand count if sampling does not provide sufficiently strong evidence that the reported outcome is correct. If they lead to a full hand count, that hand count replaces the reported results. Thus, they

---

[2] Organized by J. Morrell; one of us (PBS) provided software and support.
[3] J. Morrell, personal communication, 2020

might confirm a wrong outcome, but they never overturn a correct outcome. They make different assumptions, use different standards of evidence, and offer different assurances, as we shall explain.

```
while (!(full handcount) && !(strong evidence outcome is correct)) {
    audit more
}
if (strong evidence outcome is correct) {
    reported result is final
}
if (full handcount) {
    handcount result is final
}
```

Fig. 1. Pseudo code for sequential auditing procedures

2   Risk

The *risk* of an auditing procedure, given a trustworthy set of cast ballots and a reported outcome, is zero if the reported outcome is correct and is the chance that the procedure will not correct the reported outcome if the reported outcome is wrong. Formally, let θ denote a set of cast votes. For example, in a contest between (only) Alice and Bob in which n ballots were cast, all containing valid votes, θ is an element of $\{\text{Alice}, \text{Bob}\}^n$. (For sampling with replacement, we could also parametrize the cast votes as the fraction of votes for Alice; see Figure 2.)

RLAs treat θ as fixed but unknown. The only probability in RLAs is the probability involved in sampling ballots at random—a probability that exists by fiat and is known to the auditor, because the auditor designs the sampling protocol.

In contrast, BAs treat θ—the cast votes—as random rather than simply unknown. The probability in BAs comes not only from the sampling but also from the assumption that votes are random and follow a probability distribution π known to (or believed by) the auditor.

Let $f(\cdot)$ be the social choice function that maps a set of cast votes to the contest winner(s). Then

$$\text{risk}(\theta) \equiv \begin{cases} \Pr(\text{audit confirms reported outcome}), & \text{reported winner} \neq f(\theta) \\ 0, & \text{reported winner} = f(\theta). \end{cases}$$

RLAs ensure that the risk does not exceed a pre-specified limit (denoted α), no matter what votes were actually cast. Because θ is fixed, probabilities in RLAs come only from the random sampling of ballots.

BAs control a weighted average of the risk rather than the maximum risk (whence the title of this paper). The weights come from the prior probability distribution on θ. In symbols:

$$\text{risk}_{\text{RLA}} = \max_{\theta} \ \text{risk}(\theta)$$

$$\text{risk}_{\text{BA}} = \frac{1}{c} \sum_{\theta} \ \text{risk}(\theta)\pi(\theta)$$

where $\pi(\theta)$ is the prior on θ and $c = \sum_{\theta:\text{reported winner}\neq f(\theta)} \pi(\theta)$ makes the weights sum to 1.

BAs can have a large chance of correcting some wrong outcomes and a small chance of correcting others, depending on the prior π. If π assigns much probability to wrong outcomes where it is easy to tell there was a problem (e.g., a reported loser really won by a wide margin) the average risk (the upset probability) can be much lower than the risk for the actual set of ballots cast in the election.

An RLA with risk limit α automatically limits the upset probability to α for any prior, but the converse is not true in general. (The average of a function cannot exceed the maximum of that function, but the maximum exceeds the average unless the function is constant.) Below, we demonstrate that the upset probability can be much smaller than the true risk using simulations based on close historical elections.

3   Choosing the Prior for a BA

In a BA, the prior quantifies beliefs about the cast votes and the correctness of the reported outcome before the audit commences. Beliefs differ across the electorate. To address this, Rivest and Shen [20] considered a "bring your own prior" BA: the audit continues until everyone's upset probability is sufficiently small (see Figure 2A). Of course, if anyone's prior implies that a reported loser is virtually certain to have won, the audit won't stop without a full hand count.

Ultimately, Rivest and Shen [20] and Rivest [21] recommend using a single "nonpartisan" prior. A nonpartisan prior is one for which every candidate is equally likely to win, i.e., a prior that is invariant under permutations of the candidates' names (see Figure 2B). We doubt this captures anyone's beliefs about any particular election. Beliefs about whether the reported winner really won may depend on many things, including pre-election polls and exit polls, the reported margin, reports of polling-place problems, news reports of election interference, etc.

For instance, it seems less plausible that the reported winner actually lost if the reported margin is 60% than if the reported margin is 0.6%: producing an erroneous 60% margin would require much more error or manipulation than producing an erroneous 0.6% margin if the reported winner really lost. On the other hand, when the *true* margin is small, it is easier for error or manipulation

to cause the wrong candidate to appear to win. Moreover, a tight contest might be a more attractive target for manipulation.

If every audit is to be conducted using the same prior, that prior arguably should put more weight on narrow margins. Taken to the extreme, the prior would concentrate the probability of wrong outcomes at the wrong outcome with the narrowest margin: a tie or one-vote win for a reported loser.

Indeed, Vora [28] and Morin et al. [14] show that in a two-candidate plurality contest with no invalid votes, a ballot-polling BA using a prior that assigns probability $1/2$ to a tie (or one-vote win for the reported loser) and probability $1/2$ to correct outcomes is in fact a RLA (see Figure 2C): the upset probability equals the risk.

Constructing priors that make BAs risk-limiting for more complicated elections (e.g., elections with more than two candidates, elections in which ballots may contain invalid votes, social choice functions other than plurality, and audit sampling designs other than simple random samples of individual ballots or random samples of individual ballots with replacement) is an open problem.[4]

4   Empirical Comparison

How are risk and upset probability related? The upset probability is never larger than the risk, but the risk is often much larger than the upset probability for BAs with non-partisan priors, as we show using data from three recent close U.S. elections: the 2017 House of Delegates contest in Virginia's 94th district, the 2018 Congressional contest in Maine's 2nd district, and the 2018 Georgia Governor contest. The simulations, summarized in Table 1, treat the reported vote shares as correct, but re-label the reported winner as the reported loser. "Simulated Risk" is the estimated probability that a BA with 5% upset probability corrects the reported outcome. The simulations use the nonpartisan prior recommended by [21], with initial "pseudo-counts" of 0.5. Each audit begins with a sample of 25 ballots. Each step of each audit simulates 1,000 draws from the posterior distribution to estimate the upset probability. If the upset probability is above 5%, then the sample is increased by 20%, and the upset probability is estimated again. Each audit stops when the upset probability falls below 5%, or all ballots have been audited. We simulate 10,000 ballot-polling BAs for each scenario. Code for the simulations is available at `https://github.com/akglazer/BRLA-Comparison`.

A recount of the 2017 Virginia 94th district contest gave a 1-vote win for Simonds over Yancey. (A three-judge panel later determined that a vote counted as an overvote should be attributed to Yancey; the winner was determined by

---

[4] This is related to the problem of constructing *least-favorable priors* in statistical decision problems. There is a deep duality between Bayesian and frequentist procedures: under mild regularity conditions the Bayes risk for a *least-favorable prior* is equal to the *minimax risk* [5]. (Here, risk is a term of art, a measure of the performance of the procedure.) That is to say, for a particular choice of prior, the Bayesian procedure is in fact the frequentist procedure that does best in the worst case. The least-favorable prior is generally not "flat" or "uninformative."
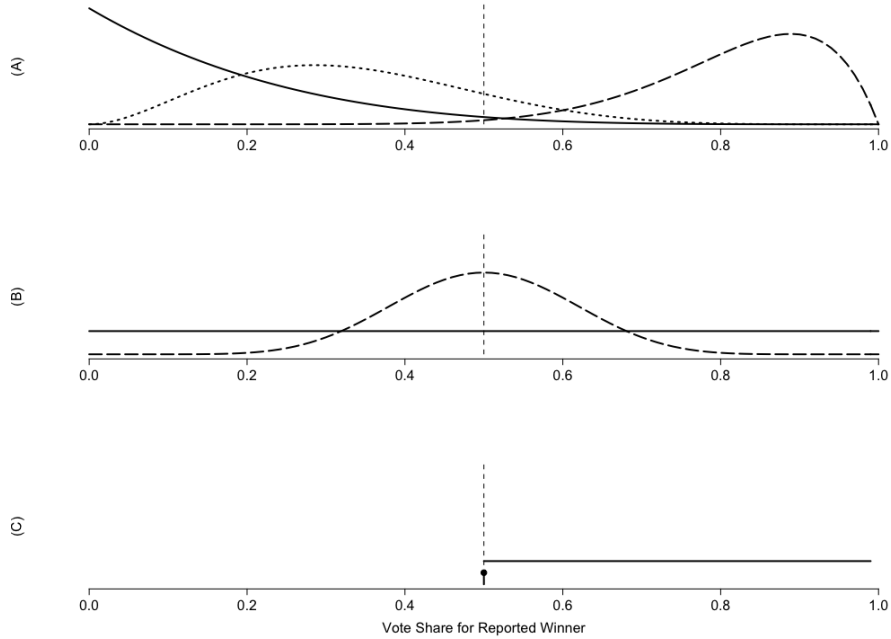
Fig. 2. *Exemplar priors for the true vote share for the reported winner in a two-candidate election. Values to the right of the vertical dotted line (at 1/2) correspond to correct reported outcomes: the winner got more than 50% of the valid votes. (A) plots three possible partisan priors. For BAs that allow observers to bring their own prior, a BA would stop only when all three posteriors give a sufficiently low probability to all outcomes where the reported winner actually lost: values less than or equal to 1/2. (B) plots two nonpartisan priors (the priors are symmetric around 1/2 and thus invariant under exchanging the candidates' names) including the flat prior recommended by Rivest and Shen [20]. The flat prior gives equal weight to all possible vote shares. (C) plots a least-favorable prior, a prior for which a BA is an RLA with risk limit equal to the upset probability. It assigns probability 1/2 to a tie, the wrong outcome that is most difficult to detect. The rest of the probability is spread (arbitrarily) across vote shares for which the reported outcome is correct. In this illustration, that probability is uniform. That choice affects the efficiency but not the risk.*

drawing a name from a bowl [12].) The 2018 Maine Congressional election used ranked-choice voting (RCV/IRV). While there are methods for conducting RLAs of IRV contests [6, 25], we treat the contest as if it were a plurality contest between the last two standing candidates, Golden and Poliquin, a "final-round margin" of 3,509 votes.[5]

	Number of Votes Cast	Margin	BA Risk (simulated)
Virginia 94th	23,215 votes	1 vote (0.004%)	43%
Maine 2nd	281,371 votes	3509 votes (1.25%)	23%
Georgia Governor	3,902,093 votes	54,723 votes (1.4%)	22%

Table 1. Simulated risk of a Bayesian Audit using 5% upset probability with a "non-partisan" prior for the 2017 Virginia House of Delegates District 94 contest, the 2018 Maine 2nd Congressional District contest, and the 2018 Georgia gubernatorial contest. Column 2: the margin for each election in number of votes and percentage. Column 3: risk of the BA, i.e., the estimated probability that the BA audit will fail to correct the outcome.

In these experiments, the actual risk of the BA is 4 to 9 times larger than the upset probability, 5%. For example, in the Virginia 94th District contest, the BA failed to correct the outcome 43% of the time, 8.6 times the upset probability. This results from the fact that the upset probability averages the risk over all possible losing margins (with equal weight), while the actual losing margin was small. Figure 3 shows the simulated risk of a BA with a nonpartisan prior and initial pseudo-counts of 0.5 for an election with 1,000,000 total votes cast. The risk is plotted as a function of the vote share for the winner. The empirical risk for a BA is very high for small margins, where auditing is especially important. As far as we know, there are situations where the risk can be an arbitrarily large multiple of the upset probability, depending on the actual cast votes, the social choice function, the prior, and details of the BA implementation (such as its rule for expanding the sample).

5   Conclusion

Elections are audited in part to rule out the possibility that voter errors, poll-worker errors, procedural errors, reporting errors, misconfiguration, miscalibration, malfunction, bugs, hacking, or other errors or malfeasance made losing candidates appear to win. We believe that controlling the probability that the reported outcome will not be corrected when it is wrong—the risk—should be the minimal goal of a post-election audit. RLAs control that risk; BAs control the upset probability, which can be much smaller than the risk.

Both RLAs and BAs require a trustworthy paper trail of voter intent. RLAs use the paper trail to protect against the worst case: they control the

---

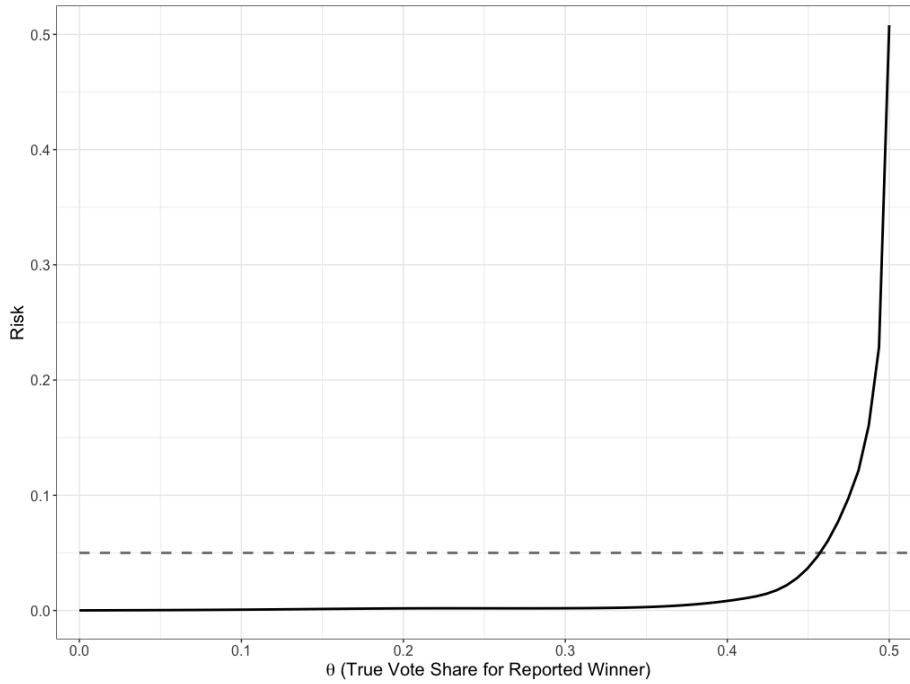[5] The final-round margin of an IRV contest is an upper bound on the true margin.

Fig. 3. Simulated risk (solid line) of a BA with nonpartisan prior for a two-candidate election with 1,000,000 total votes cast and no invalid votes. The x-axis is θ, the actual vote share for the reported winner. The reported winner really won if $\theta > 0.5$ and lost if $\theta < 0.5$. The y-axis is the actual risk, computed for $\theta < 0.5$ as the number of times the BA confirms the outcome over the total number of simulated audits. If $\theta > 0.5$ then the risk is 0. The dashed grey line at Risk $= 0.05$ is the upset probability threshold for the BA, and also the maximum risk for a RLA with risk limit 0.05.

chance of certifying the reported outcome if it is wrong, no matter why it is wrong.

BAs protect against an *average* over hypothetical sets of cast votes (rather than the worst case); the weights in the average come from the *prior*.

The priors that have been proposed for BAs do not seem to correspond to beliefs about voter preferences, nor do they take into account the chance of error or manipulation. Moreover, BAs do not condition on a number of things that bear on whether the reported outcome is likely to be wrong, such as the reported margin and the political consequences. As Vora [28] shows, some BAs are RLAs if the prior is chosen suitably. Bayesian upset probabilities can never be larger than the maximum risk, but it seems that they can be arbitrarily smaller. Conversely, Huang et al. [10] discuss finding a threshold for the upset probability in a BA (for any fixed prior) in a two-candidate, no invalid-vote contest so that using that threshold as a limit on the upset probability yields an RLA (with risk limit larger than the upset probability).

Sequential RLAs stop as soon as there is strong evidence that the reported result is correct. When the outcome is correct by a wide margin, they generally inspect relatively few ballots. Thus, even though RLAs protect against the worst case, they are relatively efficient when outcomes are correct. (When outcomes are incorrect, they are intended to lead to a full hand tabulation.)

Partisanship, foreign interference, vendor misrepresentations [29], and suspicious results [16] all threaten public trust in elections, potentially destabilizing our democracy. Conducting elections primarily on hand-marked paper ballots (with accessible options for voters with disabilities), routine compliance audits, and RLAs can help ensure that elections deserve public trust.

Bibliography

[1] American Statistical Association. American Statistical Association statement on risk-limiting post-election audits. `www.amstat.org/outreach/pdfs/Risk-Limiting_Endorsement.pdf`, 2010.

[2] A. Appel and P. Stark. Evidence-based elections: Create a meaningful paper trail, then audit. *Georgetown Law Technology Review*, 4.2: 523–541, 2020. `https://georgetownlawtechreview.org/wp-content/uploads/2020/07/4.2-p523-541-Appel-Stark.pdf`.

[3] A. Appel, R. DeMillo, and P. Stark. Ballot-marking devices cannot assure the will of the voters. *Election Law Journal, Rules, Politics, and Policy*, 2020. `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3375755`.

[4] M. Bernhard, A. McDonald, H. Meng, J. Hwa, N. Bajaj, K. Chang, and J. Halderman. Can voters detect malicious manipulation of ballot marking devices? *41st IEEE Symposium on Security and Privacy*, 2020. `https://jhalderm.com/pub/papers/bmd-verifiability-sp20.pdf`.

[5] P. Bickel and K. Doksum. *Mathematical Statistics: Basic Ideas and Selected Topics*. Pearson, 2006.

[6] M. Blom, P. Stuckey, and V. Teague. RAIRE: Risk-limiting audits for IRV elections. `https://arxiv.org/abs/1903.08804`, 2019.

[7] Brennan Center for Justice, Rhode Island RLA Working Group. Pilot implementation study of risk-limiting audit methods in the state of Rhode Island, 2019. URL `https://www.brennancenter.org/our-work/research-reports/pilot-implementation-study-risk-limiting-audit-methods-state-rhode-island`.

[8] Colorado Secretary of State. Audit Center, 2020. URL `https://www.sos.state.co.us/pubs/elections/auditCenter.html`.

[9] L. Howard, R. Rivest, and P. Stark. A review of robust post-election audits: Various methods of risk-limiting audits and Bayesian audits. Technical report, Brennan Center for Justice, 2019. `https://www.brennancenter.org/sites/default/files/2019-11/2019_011_RLA_Analysis_FINAL_0.pdf`.

[10] Z. Huang, R. Rivest, P. Stark, V. Teague, and D. Vukcevic. A unified evaluation of two-candidate ballot-polling election auditing methods. In *Proceedings of the 5th Annual Conference on Electronic Voting (E-Vote-ID '20)*, 2020.

[11] M. Lindeman and P. Stark. A gentle introduction to risk-limiting audits. *IEEE Security and Privacy*, 10:42–49, 2012.

[12] S. McCammon. Virginia Republican David Yancey wins tie-breaking drawing. `https://www.npr.org/2018/01/04/573504079/virginia-republican-david-yancey-wins-tie-breaking-drawing`, 2018.

[13] Michigan Secretary of State. Pilot audit of march presidential primary results showcases security, accuracy of Michigan elections systems, 2020. URL `https://www.michigan.gov/sos/0,4670,7-127--531561--,00.html`.

[14] S. Morin, G. McClearn, N. McBurnett, P. Vora, and F. Zagorski. A note on risk-limiting Bayesian polling audits for two-candidate elections. *Voting '20*, in press, 2020.

[15] National Academies of Sciences, Engineering, and Medicine. *Securing the Vote: Protecting American Democracy*. The National Academies Press, Washington, DC, 2018. ISBN 978-0-309-47647-8. https://doi.org/10.17226/25120. URL `https://www.nap.edu/catalog/25120/securing-the-vote-protecting-american-democracy`.

[16] K. Ottoboni and P. Stark. Election integrity and electronic voting machines in 2018 Georgia, USA. In *E-Vote-ID 2019 Proceedings*, 2019. Preprint: `https://ssrn.com/abstract=3426250`.

[17] K. Ottoboni, M. Bernhard, A. Halderman, R. Rivest, and P. Stark. Bernoulli ballot polling: A manifest improvement for risk-limiting audits. In *Proceedings of the 4th Annual Workshop on Advances in Secure Electronic Voting (Voting'19)*, 2018. Preprint: `http://arxiv.org/abs/1812.06361`.

[18] K. Ottoboni, P. Stark, M. Lindeman, and N. McBurnett. Risk-limiting audits by stratified union-intersection tests of elections (SUITE). In *Electronic Voting. E-Vote-ID 2018. Lecture Notes in Computer Science*. Springer, 2018. `https://link.springer.com/chapter/10.1007/978-3-030-00419-4_12`.

[19] R. Richie and H. Smith. A survey and analysis of statewide election recounts 2000–2015. `https://fairvote.app.box.com/v/recounts`, 2015.

[20] R. Rivest and E. Shen. A Bayesian method for auditing elections. In *Proceedings of the 2012 Electronic Voting Technology Workshop / Workshop on Trustworthy Elections (EVT/WOTE '12)*. USENIX, August 2012.

[21] R. L. Rivest. Bayesian tabulation audits: Explained and extended. `https://arxiv.org/abs/1801.00528`, January 1, 2018.

[22] Select Committee on Intelligence. Russian active measures campaigns and interference in the 2016 U.S. election. `https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume1.pdf`, 2019.

[23] P. Stark. Conservative statistical post-election audits. *Ann. Appl. Stat.*, 2: 550–581, 2008. URL `http://arxiv.org/abs/0807.4005`.

[24] P. Stark. Election audits by sampling with probability proportional to an error bound: dealing with discrepancies. `https://www.stat.berkeley.edu/~stark/Preprints/ppebwrwd08.pdf`, 2008.

[25] P. Stark. Sets of half-average nulls generate risk-limiting audits: SHANGRLA. *Voting '20*, in press, 2020. Preprint: `http://arxiv.org/abs/1911.10035`.

[26] P. B. Stark and D. A. Wagner. Evidence-based elections. *IEEE Security and Privacy*, 10:33–41, 2012. `https://www.stat.berkeley.edu/~stark/Preprints/evidenceVote12.pdf`.

[27] Verified Voting. The Verifier, 2020. URL `https://verifiedvoting.org/verifier/#mode/navigate/map/ppEquip/mapType/normal/year/2020`.

[28] P. Vora. Risk-limiting Bayesian polling audits for two-candidate elections. `https://arxiv.org/abs/1902.00999`, 2019.

[29] K. Zetter. The crisis of election security. *The New York Times*, 2018. `https://www.nytimes.com/2018/09/26/magazine/election-security-crisis-midterms.html`.

[30] K. Zetter. Critical U.S. election systems have been left exposed online despite official denials. *Vice*, 2019. `https://www.vice.com/en_us/article/3kxzk9/exclusive-critical-us-election-systems-have-been-left-exposed-online-despite-official-denials`.

[31] K. Zetter. How close did Russia really come to hacking the 2016 election? *Politico*, 2019. `https://www.politico.com/news/magazine/2019/12/26/did-russia-really-hack-2016-election-088171`.